IJACSA

WHERE WISDOM SHARES

International Journal of Advanced Computer Science and Applications

Volume 7 Issue 5

May 2016

ISSN 2156-5570(Online)
ISSN 2158-107X(Print)

SAI

www.ijacsa.thesai.org

Editorial Preface

From the Desk of Managing Editor…

It may be difficult to imagine that almost half a century ago we used computers far less sophisticated than current home desktop computers to put a man on the moon. In that 50 year span, the field of computer science has exploded.

Computer science has opened new avenues for thought and experimentation. What began as a way to simplify the calculation process has given birth to technology once only imagined by the human mind. The ability to communicate and share ideas even though collaborators are half a world away and exploration of not just the stars above but the internal workings of the human genome are some of the ways that this field has moved at an exponential pace.

At the International Journal of Advanced Computer Science and Applications it is our mission to provide an outlet for quality research. We want to promote universal access and opportunities for the international scientific community to share and disseminate scientific and technical information.

We believe in spreading knowledge of computer science and its applications to all classes of audiences. That is why we deliver up-to-date, authoritative coverage and offer open access of all our articles. Our archives have served as a place to provoke philosophical, theoretical, and empirical ideas from some of the finest minds in the field.

We utilize the talents and experience of editor and reviewers working at Universities and Institutions from around the world. We would like to express our gratitude to all authors, whose research results have been published in our journal, as well as our referees for their in-depth evaluations. Our high standards are maintained through a double blind review process.

We hope that this edition of IJACSA inspires and entices you to submit your own contributions in upcoming issues. Thank you for sharing wisdom.

Thank you for Sharing Wisdom!

Managing Editor
IJACSA
Volume 7 Issue 5 May 2016
ISSN 2156-5570 (Online)
ISSN 2158-107X (Print)

Editorial Board

Reviewer Board Members

CONTENTS

A Hybrid Method to Predict Success of Dental Implants

Reyhaneh Sadat Moayeri*
Department of Computer and Informatics Engineering,
Payame-Noor University, Tehran, Iran

Mehdi Khalili
Department of Computer and Informatics Engineering,
Payame-Noor University, Tehran, Iran

Mahsa Nazari
Department of Computer and Informatics Engineering,
Payame-Noor University, Tehran, Iran

***Abstract*—Background/Objectives: The market demand for dental implants is growing at a significant pace. Results obtained from real cases shows that some dental implants do not lead to success. Hence, the main problem is whether machine learning techniques can be successful in prediction of success of dental implants.**

Methods/Statistical Analysis: This paper presents a combined predictive model to evaluate the success of dental implants. The classifiers used in this model are W-J48, SVM, Neural Network, K-NN and Naïve Bayes. All internal parameters of each classifier are optimized. These classifiers are combined in a way that results in the highest possible accuracies.

Results: The performance of the proposed method is compared with single classifiers. Results of our study show that the combinative approach can achieve higher performance than the best of the single classifiers. Using the combinative approach improves the sensitivity indicator by up to 13.3%.

Conclusion/Application: Since diagnosis of patients whose implant does not lead to success is very important in implant surgery, the presented model can help surgeons to make a more reliable decision on level of success of implant operation prior to surgery.

***Keywords*—Data Mining; Dental Implant; W-J48; Neural Network; K-NN; Naïve Bayes; SVM**

I. Introduction

Dental plants are a sophisticated and unique technology with diverse applications which had a huge market in 2011 worth almost 7 billion USD[1]. Although it has been over a decade since the successful use of single tooth implants started, many uses and conditions for implants remain conditional and little understood. Many conditions that dentists consider important include success rates and long-term survival affected by several factors including the location of substitution (i.e., denture replacement), implant, denture anchoring, bone density, tissue health, age of recipient, prosthetic complications, abutment and implant types and also materials and post-operative medicines[2]. Therefore, this medical field of technology requires the combination of continuous clinical trials and technical innovation to improve implant reliability and survival rates and also reduce failure rates[3, 4].

Data mining is a computational process used to discover patterns in large data sets via methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The purpose of this process is to extract information from a data set and then transform it into a comprehensible structure for extra use. For data mining technologies, many data mining methods such as clustering, association, evolution, pattern matching, generalization, characterization, classification, data visualization and meta-rule guided mining have been developed[5].

Ensemble methods which combine the output of individual classifiers have been very successful for the production of accurate prediction for many complicated classification tasks. These methods become successful when they have the appropriate ability to consolidate accurate predictions and also correct errors in many diver base classifiers[6].

Two really good and well-known ensemble method s are as follows: a form of meta-learning which is called stacking and also ensemble selection. Stacking, makes a higher-level predictive model over the predictions of base classifiers, whereas, ensemble selection utilizes an incremental strategy to select base predictors for the ensemble while simultaneously balancing performance and diversity. These approaches have really superior performance in several areas because of their ability to utilize heterogeneous base classifiers[7, 8].

Stacking does not really manipulate the training dataset. Instead, based on two levels, an ensemble of classifiers is generated. In the base level, different learning algorithms are used to train multiple classifiers. The diversity is provided because different learning algorithms make different errors in the same dataset. A meta-classifier is utilized to general the final prediction. The meta-classifier is trained using a learning algorithm via a meta-dataset that combines the outputs of base-level classifiers and the real class label. A problem that exists in stacking is how to acquire an "appropriate" configuration of the meta-classifier and base-level classifiers for each domain-specific dataset. The type of meta-classifier matters to the function of the base-level classifiers. To determine the configuration of stacking, some researchers have proposed different methods[9].

Different approaches that combine many models, often called ensembles, have been explored. One of these approaches which is called "stacking", determines the optimally weighted average of many models through the minimization of predicted error. Wolpert introduced stacking in neural networks, whereas Breiman, extended the idea to uncensored regression models and then demonstrated that stacking can improve prediction error.

Breiman, discovered that combining completely different regression modes like ridge regression and subset regression, significantly reduced prediction error. LeBlanc and Tibshiranfi, discovered that stacking with a constraint of non-negative weights, is an efficient method to combine models. Van der Laan, *et al*. individually developed uncensored stacking as a "Super Learner" algorithm and offered results about the rate of convergence of the stacked estimator. Recently, Boonstra, *et al.* recently used stacking for the improvement of prediction when using different generation sequencing information in high-dimensional genome analysis[10, 11].

A. Related Works

In Text[12], Oliviera, *et al.* have studied the performance of the four techniques of RBF DDA, SVM, K-NN and NNSRM on dental implants using 10-fold cross validation. Input variables used in this study were age, gender, position, type, smoking and illness. Results showed that the accuracy of the classifiers were 75.49%, 75.96%, 75.93%, and 72.18%, respectively.

In Text[13], Braga, *et al*. proposed a decision model of prediction for dental implants by building a set of binary logistic models that could assess the probability of success or failure in oral rehabilitation process taking into account some genetic factors, individual habits, clinical and nonclinical factors. Results showed that they have obtained the AUC=0.789 (area under ROC curve).

II. Basic Theories

A. Decision tree technique

Definition of a decision tree: it is a decision support system that utilizes tree-like graph decisions and their probable after-effect, including resource costs, chance event results and utility. A classification tree or decision tree, is utilized to learn a classification function which concludes the value of a dependent attribute (variable) considering the values of the independent (input) attributes (variables). This verifies a problem which is known as supervised classification, because the dependent attribute and the counting of classes (values) are provided. Decision trees are the most powerful approaches in data mining and knowledge discovery. It includes the technology to research large and complex data to discover useful patterns. It is very important because it provides the possible to model and extract knowledge from the available data[14].

Specialists and theoreticians continually search for techniques to make the process more cost-effective, accurate and efficient. Decisions trees are very effective tools for many areas including information extraction, machine learning, data and text mining, and pattern recognition[15].

B. Neural Network

An Artificial Neural Network (ANN) is an information processing paradigm inspired by information processing methods used by biological nervous systems like the brain. The key component of it is the novel structure of the information processing system. It is comprised of many highly interconnected processing elements (neurons) working together to solve specific problems. Like people, ANNs learn by example. ANNs are typically comprised of hundreds of simple processing units wired together in a complicated communication network. Each unit or node is a simplified model of a real neuron which transmits a new signal or fires, in case it receives a sufficiently strong Input signal from the other nodes connected to it. A typical ANN is configured for a specific application like pattern recognition or data classification through a learning process. In biological systems, learning involves adjustments to the synaptic connections existing between the neurons[16].

Neural networks are one of the methods of making classifiers in which learning model are shown by a collection of joined nodes besides with their weighted connections. Neural networks are widely used to design black box classifiers. Black box means in neural network based classifiers, there is no way to express the hidden knowledge of neural networks clearly. Exactly, unlike decision tree based classifiers which are completely interpretable[17].

C. Support Vector Machine (SVM)

A Support Vector Machine (SVM) is a discriminative classifier defined by a separating hyper plane. In other words, considering labeled training data (supervised learning), the algorithm gives an optimal hyper plane which classifies new examples[18]. Besides performing linear classification, SVMs can perform a non-linear classification efficiently using what is called the kernel trick, mapping their inputs into high-dimensional feature spaces implicitly [19].

Given a dataset with n examples (x_i, y_i), where each x_i is an input data and y_i € {+1,1} corresponds to its bipolar label, i=1, 2,…, n. Using a nonlinear mapping $\phi(x)$, the input data is mapped into a high dimensional feature space F, in which the data are sparse and also possibly more separable. Then, the maximum margin which separates hyper-plane $w.\phi(x)+b=0$ is built in F, where w is a weight vector orthogonal to the hyper-plane, and b is an offset term. The margin is $1/||w||$. Maximizing the margin $1/||w||$is equivalent to minimizing $||w||2$, whose solution is found after solving the following quadratic optimization problem:

$$\min_{\omega, b} \frac{1}{2} ||\omega||^2 + C \sum_{i=1}^{n} \varepsilon_i \quad (1)$$

And

$$\text{Restrictions: } y_i(\omega\varphi(x_i) + b) \geq 1 - \varepsilon_i \text{ and } \varepsilon_i \geq 0 \text{ for } i = 1, \ldots, n \quad (2)$$

Here, C is the penalty parameter which causes a trade-off between training error and generalization and ξi is a slack variable[20].

D. K-Nearest Neighbors (K-NN)

The KNN is the simplest classification technique for the times when there is almost no prior knowledge about the distribution of the data. It simply preserves the entire training set during learning and assigns a class represented by the majority label of its k-nearest neighbors in the training set to each query. The performance of a KNN classifier is determined primarily by the choice of K as well as the distance metric. The estimate is influenced by the sensitivity of the selection of the neighborhood size K, and the reason for that is that the radius of the local region is determined by the distance of the K^{th} nearest neighbor to the query, and a different K yields different conditional class probabilities. If K is very small, the local estimate is usually going to be very poor due to the data sparseness and the noisy, ambiguous, or mislabeled points[21].

E. Naïve Bayes

The Naïve Bayes classifier is a probabilistic classifier based on the Bayes theorem, regarding Naive (Strong) independence assumption. Naïve Bayes classifiers assume that the effect of a variable value on a class is not related to the values of another variable. This assumption is referred to as class conditional independence. Naïve Bayes can usually perform more complex classification methods. It is especially suited when there is a high dimensionality of the inputs. When we want a more competent output, compared to other methods' output, we can utilize Naïve Bayes implementation. Naïve Bayesian is utilized to create models with predictive capabilities. An advantage of the naive Bayes classifier is that it merely requires a small amount of training data for estimating the parameters required for classification[22, 23].

F. Hybrid Method

A combinative classifier is a method which combines several classifiers in order to promote the robustness and achieving higher performance. In fact, this method increases the accuracy of the classification via using the results of predictions of classifiers. One of the popular methods of combinations is stacking which is usually used to combine several different classifiers such as decision tree, neural network, etc.[24]. In this method, a learning algorithm is trained to combine the predictions of many learning algorithms. First, all of the other algorithms are trained via using the available data, then a combiner algorithm is trained to make a final prediction with the use of all the predictions of the other algorithms as additional inputs. Usually, Stacking has a better performance than any trained models. It has been successfully used on both supervised learning tasks and unsupervised learning[25].

Stacking is related to combining multiple classifiers generated by using different learning algorithms L1, . . . , LN on a single dataset S, which is comprised of examples si = (xi , yi), i.e., pairs of feature vectors (xi) and their classifications (yi). In the first phase, a set of base-level classifiers C1,C2,…,CN is generated, where Ci = Li (S). In the second phase, a meta-level classifier is learned that combines the outputs of the base-level classifiers[26].

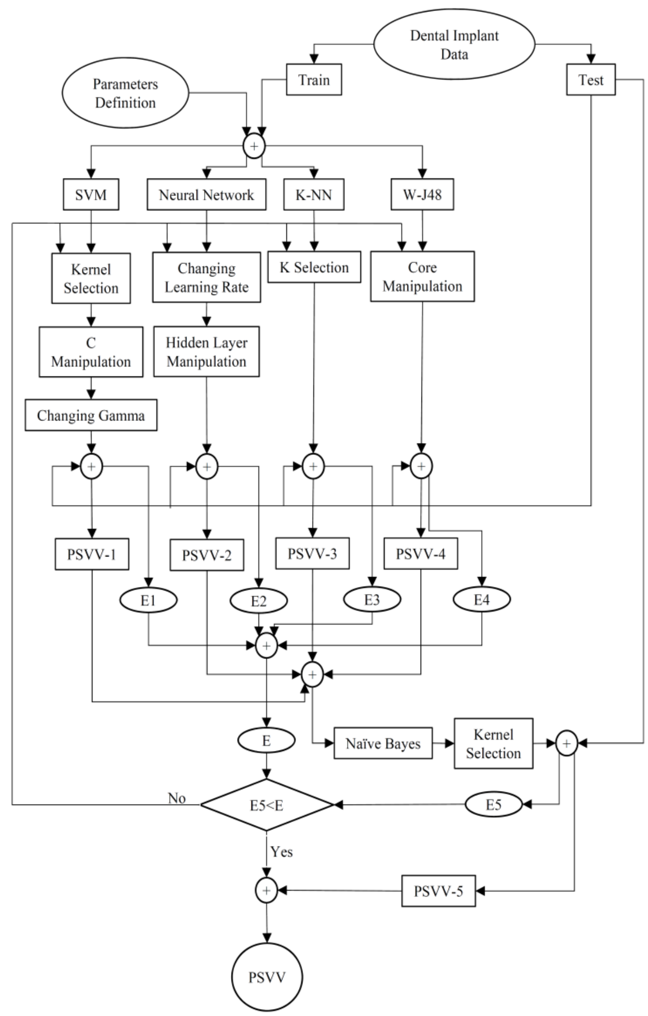
These methods can further be used to evaluate the necessity of a dental implant and reduce the risks of using one by providing prosthodontists with predictions of the dental implant results based on a patient's physical condition and dental implant characteristics prior to performing surgery.

G. Cross Validation

Cross-validation (CV) has been widely used to facilitate model estimation and variable selection. In a typical K-fold CV process, the data set is randomly and evenly split into K parts (when possible). A candidate model is made based on K−1 parts of the data set called a training set. Then the prediction accuracy of this candidate model is evaluated on a test set which contains the data in the hold-out part. By using each of the K parts as the test set and repeating the model building and evaluation process, we select the model with the smallest CV score as the 'optimal' model. In the K-fold CV procedure, each model is evaluated K times. The most common choice for evaluating a classification task is the accuracy. All other possible famous names of validation methods are seem to be as special cases of k-folds cross validation depending on the choosing value of k[27, 28].

III. Proposed Model

The block diagram of the proposed combined predictive model is shown in **Error! Reference source not found.**and the steps involved are described as follows:



1) Block diagram of the proposed model to predict dental implant success

2) Divide the dental implant data with the related defined parameters in two sections: training (to design the related classifiers) and test (to calculate the minimum error of the classifiers)

3) Apply the training section parameters to each of the classifiers: SVM, Neural Network, K-NN and W-J48

4) Change the core and Gamma in SVM classifier, after selection of the suitable kernel, to achieve PSVV-1

5) Manipulate the learning rate and hidden layer in neural network classifier, select the suitable k in K-NN classifier and change the core in W-J48 classifier, respectively to obtain PSVV-2 to 4

6) Compare the results of the above classifiers with test section parameters to achieve related errors: E1 to E4

7) Combine four predictive success variable vectors: PSVV1 to PSVV4 and enter the result to stacking learner 1 i.e. Naïve Bayes algorithm

8) Apply Naïve Bayes classifier on the combined input, select the suitable kernel and finally compare the related output results with test section parameters, to achieve predictive success variable vectors PSVV-5 and error E5

9) Determine the minimum error value E from E1 to E4 to choose the final predictive success variable vector PSVV as follows:

$$\begin{cases} E5 < E: & PSVV = PSVV5 \\ E5 > E: & Output\ of\ step\ 5 \end{cases}$$

IV. Experimental Results and Discussion

According to the experts' opinion, the most important factors which influence the success or failure of dental implants are different. To evaluate the effectiveness of the proposed algorithm, we used 224 patient cases which had bone graft. This data set belongs to School of Dentistry of Tehran University and consists of 16 different dental parameters which are: gender, age, systemic, smoking, location, placement, loading, diameter, length, system, type, platform, connection, parallel taper, over-denture, and sinus lift.

Well-known performance indicators in medical problems are accuracy, sensitivity, and specificity[29].

Accuracy is the degree of how accurate is the prediction of implant success or failure.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

Sensitivity is the degree of how accurate is the prediction of implant failure:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

Specificity is the degree of how accurate is the prediction of implant success:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (5)$$

Here, TN is True Negative, TP is True Positive, FP is False Positive and FN is False Negative.

Rapid miner software was used as the data-classification tool to analyse the 224 implants. Comparison between proposed method and other methods of W-J48, Neural Network, SVM, K-NN and Naïve Bayes with three different cross validation techniques of 5, 7, and 10-fold are shown in the following figures.

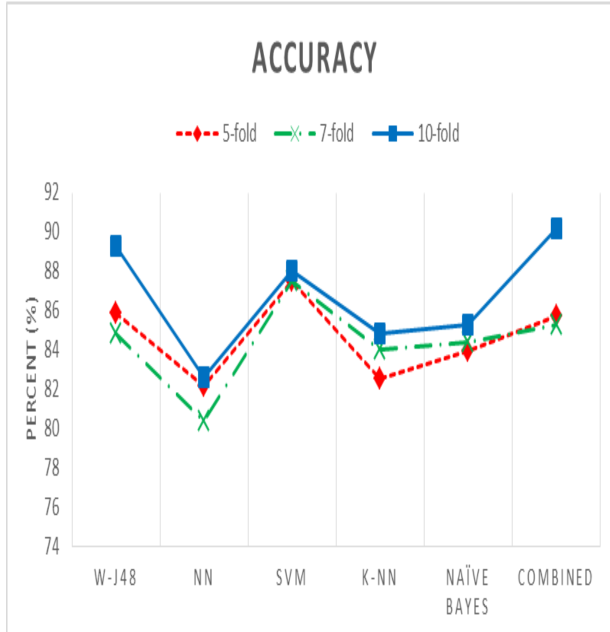


Fig. 1. Accuracy indicator for predictive models

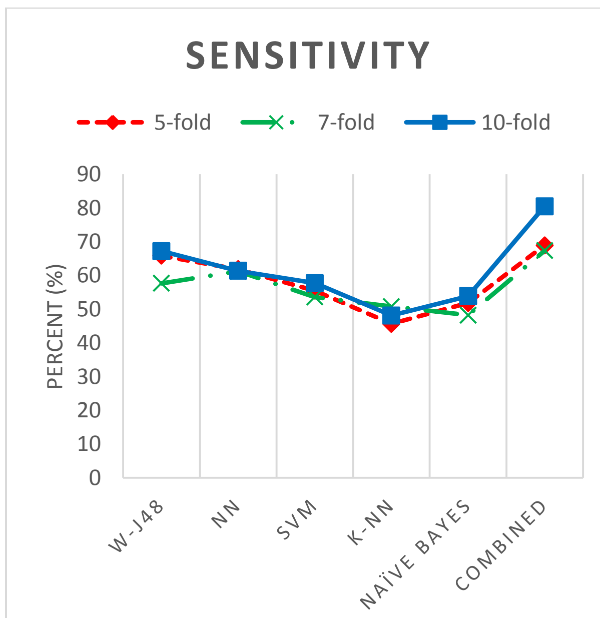


Fig. 2. Sensitivity indicator for predictive models

Comparison between different prediction methods with 5, 7, and 10-fold cross validation shows that the 10-fold cross validation gives higher performance than the others. Thus, the 10-fold cross validation has been selected for analysis. Results are summarized in **Error! Reference source not found.** as follows.

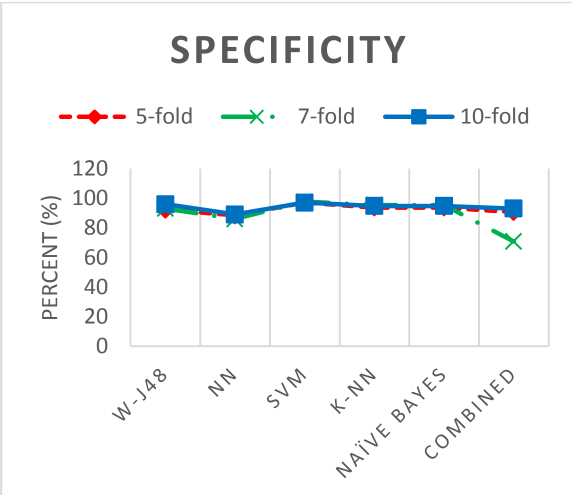


Fig. 3. Specificity indicator for predictive models

TABLE I. PERFORMANCE INDICATORS FOR DIFFERENT METHODS IN IMPLANTATION WITH 10-FOLD CROSS VALIDATION

Model	*Accuracy*	*Sensitivity*	*Specificity*
W-J48	89.31%	67.17%	95.92%
Neural Network	82.63%	61.33%	88.99%
SVM	88.00%	57.69%	97.09%
K-NN	84.05%	48.08%	94.77%
Naïve Bayes	85.30%	53.85%	94.77%
Proposed Method	90.22%	80.50%	93.01%

As shown above, the accuracy indicator of the proposed method with value of 90.22% is higher than prediction of W-J48 as most accurate single classifier with value of 89.31%.

From the viewpoint of sensitivity, the proposed method also gives better predictions than single component classifiers. The value of sensitivity indicator for the proposed model is 80.50% while it is 67.17% for W-J48 as the best single classifier.

However, from the viewpoint of specificity, the SVM model has highest value with prediction of 97.09% while the proposed model stands on the fifth rank before neural network model.

As another method for comparing models we can use ROC[1] diagram. ROC is graphical schematic which shows the performance of classifiers. It shows the False Positive Rate (FPR) versus True Positive Rate (TPR) [30].

The TPR defines the number of correct positive results that occur among all positive samples available during the test.

$$TPR = \frac{TP}{TP + FN} \tag{6}$$

On the other hand, FPR defines how many incorrect positive results occur among all negative samples which are available during the test.

$$FPR = \frac{FP}{FP + TN} \tag{7}$$

An ROC space is defined by FPR and TPR as x and y axes respectively. It shows relative trade-offs between false positive (costs) and true positive (benefits). TPR is equivalent to sensitivity and FPR is equal to 1-specificity[31, 32].

The more the surface under ROC curve, the more efficiency of the algorithm.

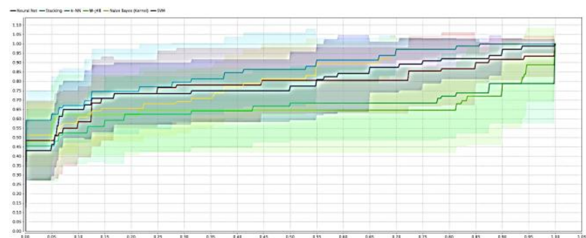


Fig. 4. ROC diagram for predictive models

In the above diagram, the six models of W-J48, Neural Network, SVM, K-NN, Naïve Bayes, and ensemble-based proposed model are shown with olive green, purple, red, green, yellow, and blue colors, respectively. From the above diagram, it is clear that the combined predictive model acts better than the other models.

V. CONCLUSION

In this study, we followed two important purposes. One is to show whether combination of algorithms has higher performance than the singular ones. The other purpose is to increase the prediction of implants which are not successful. This item is described by sensitivity indicator.

In order to substantiate the first purpose, we used five prediction classifiers using patients' data and then combined them in a way to achieve higher performance. The results of our study showed that the hybrid algorithm gains higher accuracy than using only one singular algorithm for classification of records. Also, it increased the sensitivity indicator significantly and since it is very important to identify the patients whose implant is unsuccessful, hence the second purpose of this study has been achieved.

ACKNOWLEDGEMENT

The authors are grateful to the School of Dentistry of Tehran University for Experimental Data Collection. Many people greatly contributed to this work. Special mention goes to Dr. Mohammad Javad Kharrazi Fard and Dr. Razieh Sadat Moayeri.

[1] Receiver Operating Characteristic

REFERENCES

[1] Ceramic Industry, Dental implants and prosthetics market continues growth, 2012, [Available at:
http://www.ceramicindustry.com/articles/92515-dental-implantsand-prosthetics-market-continues-growth].

[2] Mangano C,Piattelli A,Lezzi G,Mangano A,La Colla L. Prospective clinical evaluation of 307 single-tooth Morse taper-connection implants: a multicenter study. The International Journal of Oral and Maxiofacial Implants. 2010. 25 (2): 394-400

[3] Wang T, Trappey C, Hoang S, Trappey A. Constructing a dental implant ontology for domain specific clustering and life span analysis. Advanced Engineering Informatics. 2013. 27(3): 346-357

[4] Jung RE, Pjetursson BE, Glauser R, Zembic A, Zwahlen H, Lang NP. A systematic review of the 5-year survival and complication rates of implant supported single crowns. Clinical Oral Implants Research. 2008. 19(2): 119–130

[5] Liao S, Chu P, Hsiao P. Data mining techniques and applications – A decade review from 2000 to 2011. Expert Systems with Applications. 2012. 39(12): 11303-11311

[6] Yang P, Yang Y, Zhou B, Zomaya A. A Review of Ensemble Methods in Bioinformatics.Current Bioinformatics. 2010. 5(4): 296-308

[7] Whalen S, Pandey G. A Comparative Analysis of Ensemble Classifiers: Case Studies in Genomics, Icahn School of Medicine at Mount Sinai. New York, USA. 2013

[8] Altmann A et al. Comparison of Classifier Fusion Methods for Predicting Response to Anti HIV-1 Therapy.PLOS ONE. 2008. 3(10): 3470

[9] Chen Y, Wong M, Li H. Applying Ant Colony Optimization to configuring stacking ensembles for data mining. Expert Systems with Applications. 2014. 41(6): 2688-2702

[10] Boonstra P, Taylor J, Mukherjee B. Incorporating auxiliary information for improved prediction in high-dimensional datasets: an ensemble of shrinkage approaches. Biostatistics. 2013. 14(2): 259-272

[11] Wey A, Connett J, Rudser K. Combining parametric, semi-parametric, and non-parametric survival models with stacked survival models. Biostatistics. 2015. 16(3): 537-549

[12] Oliviera A, Baldisserotto C, Baldisserotto J. A Comparative Study on Machine Learning Techniques for Prediction of Success of Dental Implants. MICAI 2005: Advances in Artificial Intelligence. 2005. 3789: 939-948

[13] Braga et al. Decision Model to Predict the Implant Success. Computational Science and Its Applications. 2012. 7333: 665-674

[14] Korting T. C4.5 algorithm and Multivariate Decision Trees: IEEE Computer Society. 2013

[15] Rokach L, Maimon O. Data Mining with Decision Trees, Theory and Applications. 2nd Edition: World Scientific Publishing Company. 2014

[16] Sonali B, Wankar P. Research Paper on Basic of Artificial Neural Network. International Journal on Recent and Innovation Trends in Computing and Communication. 2014. 2(1): 96-100

[17] SanieiAbadeh M, Mahmoudi S, Taherparvar M. Applied Data Mining. 2nd Edition: NiazDanesh. 2012

[18] Huang X. Support Vector Machine Classifier with Pinball Loss.Pattern Analysis and Machine Intelligence. 2013. 36(5): 984-997

[19] Rahulamathavan Y, Phan R, Veluru S, Cumanan K. Privacy-Preserving Multi-Class Support Vector Machine for Outsourcing the Data Classification in Cloud. Dependable and Secure Computing. 2013. 11(5): 467-479

[20] Zhang X, Qiu D, Chen F. Support vector machine with parameter optimization by a novel hybrid method and its application to fault diagnosis. Neurocomputing. 2015. 149(B): 641-651

[21] Imandoust S, Bolandraftar M. Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background. International Journal of Engineering Research and Applications. 2013. 3(5): 605-610

[22] Manjusha K, Sankaranarayanan K, Seena P. Prediction of Different Dermatological Conditions Using Naïve Bayesian Classification. International Journal of Advanced Research in Computer Science and Software Engineering. 2014. 4(1): 864-868

[23] Kaur G, Oberai N. Naïve Bayes Classifier with Modified Smoothing Techniques for Better Spam Classification. International Journal of Computer Science and Mobile Computing. 2014. 3(10): 869-878

[24] Polikar R. Ensemble Based Systems in Decision Making. Circuits and Systems Magazine. 2006. 6(3): 21-45

[25] Gehi B, Siddavatam I. Network Bandwidth Predictive analysis using Stacking. International Journal on Recent and Innovation Trends in Computing and Communication. 2014. 2(8): 2166-2170

[26] Dzeroski S, Zenko B. Is Combining Classifiers with Stacking Better than Selecting the Best One? Machine Learning. 2004. 54(3): 255-273

[27] Jung Y, Hu J. A K-fold averaging cross-validation procedure.Journal of Nonparametric Statistics. 2015. 27(2): 167-179

[28] Abdalla M. Evaluation of Data Mining Classification Models. IUG Journal of Natural and Engineering Studies. 2014. 22(1): 151-165

[29] Hauser R et al. Sensitivity, Specificity, Positive and Negative Predictive Values and Diagnostic Accuracy of DaTscan™ (Ioflupane I123 Injection): Predicting Clinical Diagnosis in Early Clinically Uncertain Parkinsonian Syndrome. Journal of Neurology & Stroke. 2014. 1(1): 1-13

[30] Rehman A, Khanum A. Swarm Optimized Fuzzy Reasoning Model (SOFRM) for Diabetes Diagnosis. Life Science Journal. 2014. 11(3): 42-49

[31] Hernandez J. ROC Curves for Regression. Pattern Recognition. 2013. 46(12): 3395-3411

[32] Till R, Hand D. A Simple Generalization of the Area under the ROC Curve for Multiple Class Classification Problems.Machine Learning. 2001. 45(2): 171-186

ADBT Frame Work as a Testing Technique: An Improvement in Comparison with Traditional Model Based Testing

[1]Mohammed Akour, [2]Bouchaib Falah, [3]Karima Kaddouri
[1]Computer Information Systems Department, Yarmouk University
[2,3]School of Science and Engineering, Al Akhawayn University

***Abstract*—Software testing is an embedded activity in all software development life cycle phases. Due to the difficulties and high costs of software testing, many testing techniques have been developed with the common goal of testing software in the most optimal and cost-effective manner. Model-based testing (MBT) is used to direct testing activities such as test verification and selection. MBT is employed to encapsulate and understand the behavior of the system under test, which supports and helps software engineers to validate the system with various likely actions. The widespread usage of models has influenced the usage of MBT in the testing process, especially with UML. In this research, we proposed an improved model based testing strategy, which involves and uses four different diagrams in the testing process. This paper also discusses and explains the activities in the proposed model with the finite state model (FSM). The comparisons have been done with traditional model based testings in terms of test case generation and result.**

Keywords—Activity Diagram; Black Box Testing; Finite State Machine; Model-Based Testing; Software Testing; Test Suite; Test Case; Use Case Diagram

I. Introduction

Software testing is an important, if not the most important, activity in the software development cycle of any system without exception. It is an intellectually challenging activity aimed at evaluating the capability of a program or system to determine whether or not it meets requirements [1].

Software testing is defined as the validation and verification of the proposed system or product to ensure that it conforms to the agreed-upon requirements, that it is functioning as expected by both the developer team and the stakeholders, and that it satisfies the latter. As software testing can get very difficult or costly to perform, software testing engineers are always developing new or refining existing testing techniques tools, always having in mind the objective of developing the optimal approach that would ideally be cost-effective and efficient at the job simultaneously. In other words, the ideal testing methodology is one with maximum coverage and minimal cost. Since testing occurs right after the development phase, it is an on-going process which might take place earlier in the software development cycle. Consequently, there are two essential methods of software testing:

- Black box (also called functional testing) to be specific, which are used to test and validate the requirements and design of the system before moving on to the implementation.
- White-box testing (also called structural testing and glass box testing) is testing that takes into account the internal mechanism of a system or component.

Each method encompasses a lot of different testing techniques. One of the most important techniques of the black box testing method is called Model-Based Testing (MBT).

Our suggested approach aims to overcome some of the traditional MBT challenges. The idea behind this approach is to base the testing process on abstract representations of the system just as MBT does, but -unlike MBT- manually generating corresponding test cases. Our approach will test the software by testing its design. Based on the system requirements, this new testing technique which was inspired from the traditional MBT paradigm will use three UML diagrams instead of the FSM diagram. Detailed description is covered in the next sections.

The rest of the paper is organized as follows: Section II describes the traditional MBT. Section III explains our proposed approach. Section IV presents some related works. Section V presents a comparison between ADBT and MBT. Finally, we conclude the paper in section VI.

II. Traditional MBT

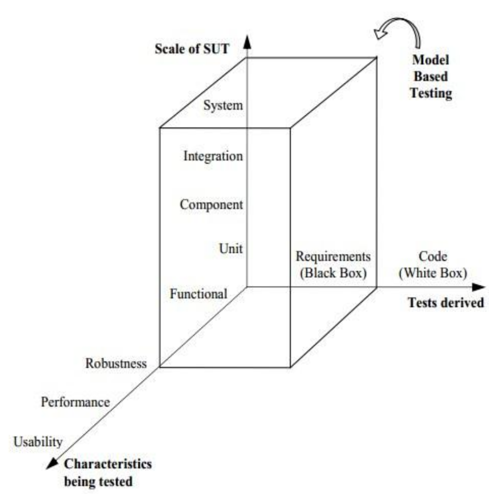


Fig. 1. MBT Context [3]

Model-Based testing refers to the black box testing technique where test cases are automatically generated based on a model, which represents the behavior of the system under test (SUT), and on the system's requirements and specification [2]. To clearly understand the scope of Model-Based Testing, Figure 1 illustrates the MBT context.

In the software development cycle, it is often required to model the system to be implemented and design an abstract view of its functionalities. There are many models available for testers to represent and model abstract depictions of systems. Some of which are UML diagrams, Markov chains, grammars, state charts, and finite state machines [4].

In traditional model-based testing, the model used to generate test cases is the finite state machine diagram (FSM) [4]. Finite state machines or finite state automata are mathematical models of computation. They are used at both hardware and software levels [5]. An FSM is the description of a finite set of states of a particular machine and the transitions between those states [5]. The events responsible for a transition from one state to another state are triggers [5]. In other terms, an FSM is a diagram which represents the set of system's states and the triggering events or conditions responsible for switching between states [5].

Figure 2 shows an example of a simple phone system. The SUT is a phone with a set of states [6]. Those states are represented as nodes, while the actions the user performs are represented as edges (triggering transitions) [6]. For instance, a possible system input to be tested could be: Pick Up. After getting the model of the SUT, the next step is to use a test case generator to generate test cases [6]. An example of a produced test case from this phone system model could be the following sequence of actions and states: <Ringing ->Hang Up -> On Hook ->Pick Up>.

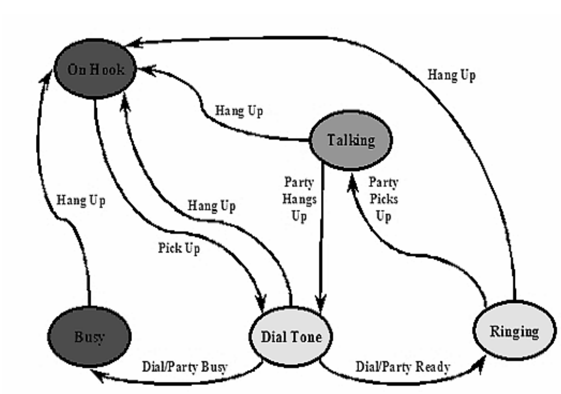


Fig. 2. Example of FSM (Phone System) [5]

The test case generator then proceeds to generate test cases for the given model, which could be based on some specific coverage criteria, such as a particular set of requirements [4]. The result of this process is an abstract test suite which needs to be concretized and converted into an executable set of test cases [4].

Test scripts or test drivers perform this operation and map each abstract model test case to an executable one using what is called an adaptor code, developed in C, Java, C#, or any other application language [4].The executable test suite is then run, and the final results are reported and analyzed. In the presence of faults, the failure is traced back; the model might be modified if deemed necessary and the testing process is repeated [4]. Figure 3 presents the process of MBT.

Model-based testing has many benefits [3]. Among its advantages is the fact that it fills the gap between the abstract and concrete levels of the system (enhances traceability between each executable test case and its corresponding part in the model and vice versa) thanks to the scripting tools [3].It also provides efficient fault detection and improves test quality, since it generates a set of non-repetitive test cases for a given SUT [3].

Some problems with traditional MBT are summarized as follows.

While having noticeable advantages, MBT also has its set of drawbacks and limitations [3].

Useless test cases

Not all the test cases generated automatically are useful for the testing process of the SUT. Traditional MBT produces a huge set of test cases, not all of them possible or beneficial for testing. This problem leads to additional cost and time to filter and get through all the test cases to choose the valid ones [3].

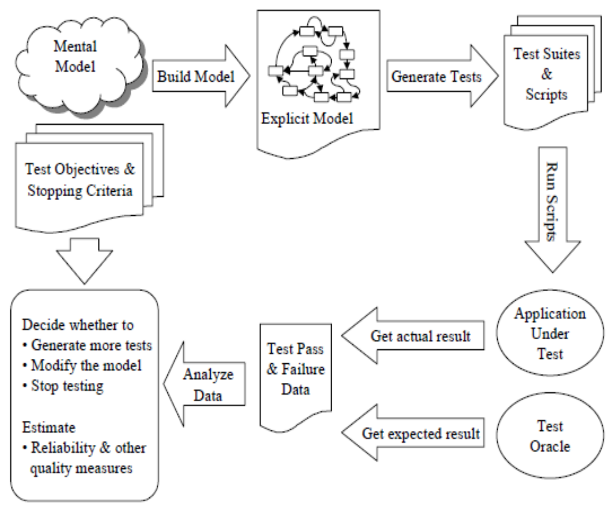


Fig. 3. MBT process [3]

Skills limitations

The test engineer has to demonstrate mastery of different skills such as high level of knowledge about computational system models or expertise in automation tools and scripts. This effort requires additional training costs.

FSM problem

The most issue with using finite state machines is state space explosion. For complex systems or large programs, the number of states in the FSM can grow uncontrollably and might exceed the given computational capabilities. This problem affects test coverage quality and efficiency and test case generation.

Failed tests issue

When a failure occurs, it can either be due to the system under test, the model, the test case generator, or the adaptor code used for conversion. Those many possible origins for failure increase the difficulty to trace back a failed test along with being extremely time-consuming [3].

III. Proposed Activity Diagram Based Testing Technique (ADBT)

The suggested testing approach aims to overcome these challenges by manually generating test cases instead of using a program for the task. The idea behind this approach is to base the testing process on abstract representations of the system just as MBT does, but – unlike MBT – manually generating corresponding test cases.

Our approach will test the software through testing its design. Based on the system requirements, this new inspired testing technique, from the traditional MBT paradigm, will use three UML diagrams instead of the FSM diagram: Use Case, Class and Activity diagrams. From the use case diagram, we derive the corresponding activity diagrams. The activity diagrams will present a set of numbered steps. Each path/scenario in the activity diagram of a single use case corresponds to a test case.

The test cases are then set up in test case tables divided into “Steps” and “Input/output”. “Steps” corresponds to the activity diagram numbered steps while Inputs are test points and Outputs are expected results from the system. The class diagram is needed to get the system’s variables (attributes of classes) used while setting up the test cases table.

The ADBT Steps are presented as follow:

1) Retrieve Use Case diagram from requirements: After getting the agreed upon set of specifications, the system designer will model the use case diagram for the system which represents the functional requirements.

2) Derive Class Diagram

3) Develop Activity Diagrams from Use Case Diagram: Each use case scenario in the use case diagram corresponds to an activity diagram.

4) Set test case for each Activity Diagram path: Each path from start to end in a given activity diagram is a test case.

5) Check test case results: In the case of error, the models used are checked for consistency problems or faults and use cases are generated again. In an optimal workflow setting, the design team will derive the necessary diagrams for the test engineers who will directly use them and only perform the last step of ADBT.

IV. Case Study

As a case study, Online Movie Tickets Purchase scenario is employed and discussed.

Let us consider the following system requirements:

- The customer should be able to search for a movie by title, with the output being its price, its time and its days of screening. In case it is unavailable, the user gets an error message and is asked to re-enter its title.
- The customer should be able to buy tickets for an available movie. He has to enter his/her credit card credentials. In case there is an error, the payment process is canceled, and the user has to re-enter the payment information.
- The customer should be able to get a ticket receipt.
- The customer should be able to display the list of all movies.

Step 1: Use Case Diagram

This system results in a properly simple use case diagram with four use case scenarios. Figure 4 shows the use case diagram for online movie tickets purchase system.

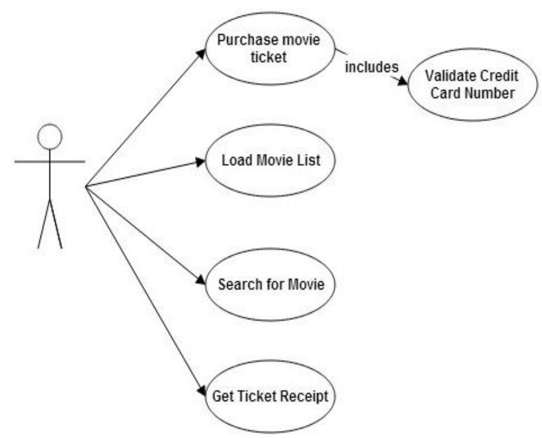


Fig. 4. Use Case Diagram

Step 2: Class Diagram

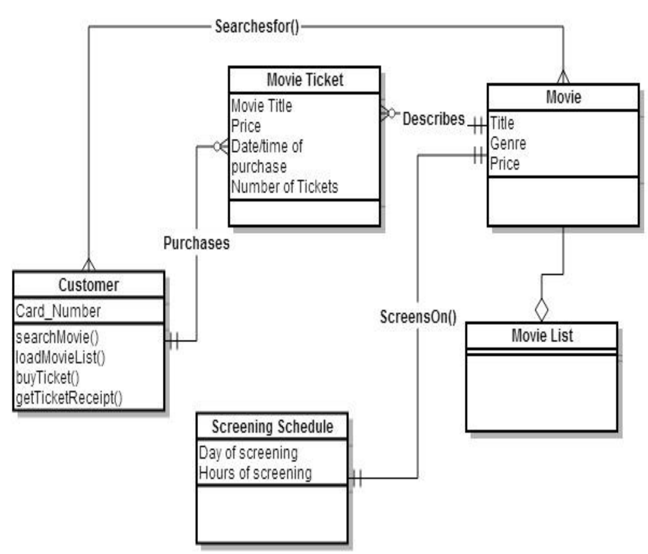


Fig. 5. Class Diagram

A tentative class diagram for this system could be as specified in Figure 5. The class diagram is vital for the I/O flow as it provides explicit information about the different system variables or attributes which we might need to test. Figure 5 shows the class diagram for online movie tickets purchase system.

Step 3: Activity Diagrams

According to the previous use case diagram, there are four use case scenarios. However, the last two requirements (get ticket receipt and load movie list) are impossible to test in the design phase and can be represented with an activity diagram. Thus, we will consider the two first requirements (search for the movie and pay movie ticket) for this particular example as shown in Figures 6 and 7.

Activity Diagram#1: Search for Movie

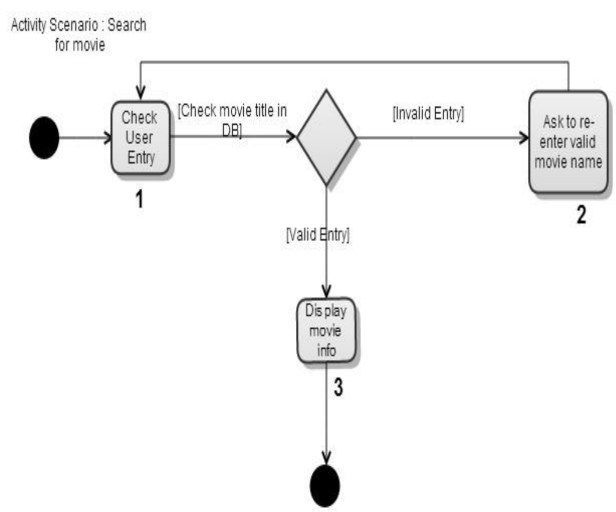


Fig. 6. Activity Diagram 1

Activity Diagram#2: Pay Movie Ticket

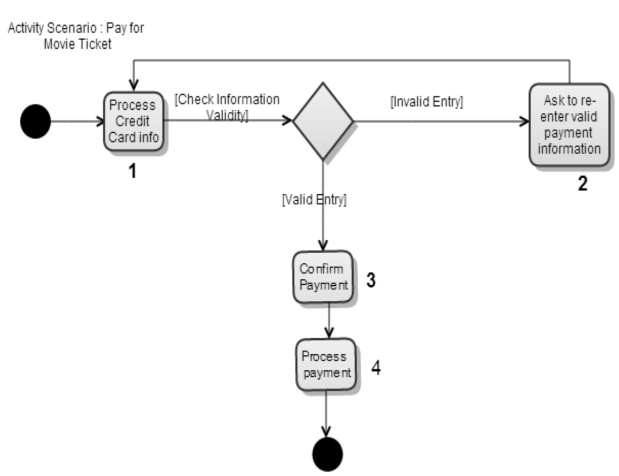


Fig. 7. Activity Diagram 2

Step 4: Test Cases

Since each path in each activity diagram results in test cases, we have four test cases in total.

Test case#1

TABLE I. TEST CASE 1

Step	I/O
1	Input: "HY7LO."
2	Output:"Not Available-Please Enter Valid Title"

Test case#2

TABLE II. TEST CASE 2

Step	I/O
1	Input:"Hunger Games"
3	Output:"Hunger Games, price: 30dhs, MWF 2h-4h."

Test case#3

TABLE III. TEST CASE 3

Step	I/O
1	Input: "zzz zzz zzz"
2	Output: "Invalid Payment Information-Try again."

Test case#4

TABLE IV. TEST CASE 4

Step	I/O
1	Input:"1234569921"
3	Output:"Your Payment has been confirmed."
4	(no I/O, only processing)

Step 5: Checking Test Cases Results

There are no errors in the test cases; the I/O flow is correct and expected, which indicates that our models are consistent.

V. COMPARISON OF ADBT WITH MBT

A. *Test case generation:*

The state machine diagram for the previous online purchase ticket was implemented, resulting in some number of 11 states. For state-based test case coverage, each state has to be visited at least once, meaning that we will generate at least 11 test cases in addition to other relevant or irrelevant paths/test cases (since we have automatic random path selection). ADBT for this simple example results in 4 test cases.

B. SUT Model:

The ADBT approach uses the Activity Diagram for testing; however, it needs to first design the Use Case and Class as well. MBT only uses FSM. However, and as suggested earlier, the test engineer does not necessarily design all three models as it could be and usually is part of another team's work.

C. Test case results:

The test case results for the four test cases are clear and significant and easily allow checking for possible model errors. Concerning MBT, generating and running test cases is beyond this project's scope due to the complexity of the process (needs automation tools, coding test scripts, etc…). However, due to the number of generated test cases, we can assume that it will be cumbersome to filter through all of them, choose the relevant ones for execution, and proceed to trace back errors found since it could have – as previously mentioned – many origins (code, test script, model, test case generation, etc…).

VI. Related Work

Hemmati, *et al.* [7] propose an alternative technique to MBT, which is supposed to overcome one of MBT's most important issues, that is, the enormous number of generated test cases which impact negatively on both time and cost.

Their approach implements a smart test case selection technique based on genetic algorithms which choose a test suite from the large pool of generated test cases to be executed based on resources and maximum fault coverage criteria, thus extending traditional MBT into a more time/cost saving version [7].

Arnold, *et al.* propose a scenario-based approach to traditional MBT [8]. Tests are executed automatically and exist within the scope of a large pool of states in MBT, which makes it harder to trace them back directly to the SUT.

This new approach makes test execution semi-automatic and introduces scenario-based test cases which are much more relevant and closer to the SUT because the set of these applicable states is manually selected.

The approaches presented in [9, 10] utilized a model based for software security testing and software test selection perceptively. In [10], authors implemented model-based approach to tracking vital items in test models and its corresponding item in structure model. When any modification occurs in the component model of the software under test, the component model identifies and conveys changes that should be performed to update the corresponding test model.

Mohacsi *et al.* [11] adopted a model-based test (MBT) approach for systematic test design and generation of their case study. They believed in that MBT assured modularity and abstraction, moreover, it leads to decrease the required effort for test maintenance. Their model based testing is build based on activity diagrams. One of the main lessons learned from their case study is the reduction of the test effort, especially the effort for test maintenance.

Yanjun, *et al.* [12] proposed new model-based testing process in order to improve structural coverage in functional testing. They concentrate on integrating three main parts, specification-based test generation tool, a model-checker and an environment for model test execution to enhance structural coverage rate. Their MBT process facilitates capturing suspicious code branches that require analysis to determine whether they are truly unreachable or a bug is occurring in a condition guarding this branch. Moreover, Model checking allows extending the functional test set by test cases derived from uncovered branches.

Amalfitano, *et al.* [13] proposed and implemented a new fully automatic technique to test GUI-based Android apps. Their technique is composed of 3 main steps namely, observation, extraction, and abstraction of the run-time state of GUI widgets. The abstraction is employed to develop a scalable state-machine model that, together with event-based test coverage criteria provide a way to automatically generate test cases. They performed their technique on 4 open-source software applications. The results showed that the test cases generated were useful at detecting serious and relevant bugs in the apps.

VII. Conclusion

Software Testing is and will remain the most important, but also the trickiest and most challenging activity in the software development cycle. There is an abundance of testing techniques in the literature, and one of them is a black box testing technique called Model-Based Testing.

This paper presents an improvement on testing technique to overcome some of the traditional MBT challenges. Our approach is based on Activity Diagrams.

Following are the advantages of using ADBT.

- No more useless or irrelevant test case generation problems.
- The FSM diagram state explosion is resolved since we changed the model.
- Errors can be easily traced back to the model in case of failure.
- The tester does not need extra training skills to conduct the testing process, since the three UML diagrams can be provided by the design team and made available for the tester to only generate and execute tests.

This new technique has been demonstrated as having some benefits.

It is more costly and less composite than traditional MBT, it allows for easily testing the consistency of the software design and checking if it conforms to what is expected from the customer, and finally it provides an easy and systematic way of generating test cases. As future works, we intend to conduct more rigorous validation to make our result well proven.

References

[1] B.Falah, K.Magel, O.ElAriss, "*A Complexity Based Regression Test Selection Strategy*," Computer Science & Engineering: An International Journal (CSEIJ), Vol.2, No.5, October 2012

[2] M. Utting, A. Pretschner, and B. Legeard.Taxonomy of Model-Based Testing.In *Working Paper Series*, vol. 04, April 2006.

[3] Y. Malik. *Model Based Testing: An Evaluation*. Master Thesis. Reading University: Blekinge Institute of Technology, May 2010.

[4] K. El-Far and A. James. Model-based Software Testing. In *Encyclopedia on Software Engineering*, J.J. Marciniak (ed). Wiley, 2001.

[5] T. Coquand. *Finite State Machines. In Automata*, pp 471-545, University of Gothenburg Press, September 2010.

[6] A. Chandran. *Model-Based Testing: Executable State Diagrams*. In proceedings of the STEP-AUTO 2011 Conference for International Testing (ISQT' 11), pp 10-17, 2011.

[7] H. Hemmati, L. Briand, A. Arcuri, and S. Ali.*An Enhanced Test Case Selection Approach for Model-Based Testing: An Industrial Case Study*. Simula Research Laboratory.Technical Report, 2010.

[8] D. Arnold, J.P. Corriveau, and W. Shi. *A Scenario-Driven Approach to Model-Based Testing*.2010. http://people.scs.carleton.ca/~jeanpier/VF_test_generation.pdf

[9] Bouchaib Falah, Mohammed Akour, Samia Oukemeni, An Alternative Threat Model-based Approach for Security Testing. International Journal of Secure Software Engineering (IJSSE), IGI, Vol. 6 issue 3 (2015): 50-64.

[10] Ahmad Saifan, Mohammed Akour, Iyad Alazzam, Feras Hanandeh, Regression Test-Selection Technique Using Component Model Based Modification: Code to Test Traceability, IJACSA, 2016

[11] Mohacsi, Stefan, Michael Felderer, and Armin Beer. "A Case Study on the Efficiency of Model-Based Testing at the European Space Agency." 2015 IEEE 8th International Conference on Software Testing, Verification and Validation (ICST). IEEE, 2015.

[12] Sun, Yanjun, Gérard Memmi, and Sylvie Vignes. "A Model-Based Testing Process for Enhancing Structural Coverage in Functional Testing." Complex Systems Design & Management Asia. Springer International Publishing, 2016. 171-180.

[13] Amalfitano, D., Fasolino, A. R., Tramontana, P., Ta, B. D., & Memon, A. M. (2015). MobiGUITAR: Automated Model-Based Testing of Mobile Apps.Software, IEEE, 32(5), 53-59

Classified Arabic Documents Using Semi-Supervised Technique

Dr. Khalaf Khatatneh
Al-Balqa Applied University

***Abstract*—In this work, we test the performance of the Naïve Bayes classifier in the categorization of Arabic text. Arabic is rich and unique in its own way and has its own distinct features. The issues and characteristics of Arabic language are addressed in our study and the classifier was modified and regulates to fit the needs of the language. a vector or word and their frequencies method is used to represent each document. We trained our classifier using both techniques supervised and semi-supervised in an attempt to compare between them and see if the classification accuracy will improve as a result of using the technique of semi-supervised. Many various experiments were performed, and the thoroughness of the classifier was measured using recall, precision, fallout and error. The outcomes illustrates that the semi-supervised learning can significantly enhance the classification accuracy of Arabic text.**

***Keywords*—Arabic Language; Naïve Bays; Classifier; Indexing; Stop word**

I. Overview and Introduction

The classification of text is the work of sorting a set of documents into different categories from a set that was predefined. Classification of text is considered as an old domain of research but it gained more concern because the number of online documents is becoming huge and getting larger each day. Manual manipulation of this massive amount of data is extremely expensive, consuming too much time, and requires human expertise that cannot be continuously obtainable 24 hours a day, thus the need for automatic classification. Automatic text classification helps reduce the time required for classifying hundreds, even thousands, of documents every day, and will also save on the expenses and efforts of human experts.

Various algorithms for machine learning have been used for the process of text categorization: support vector machines, *k* means the closest neighbor, naïve Bayes, and neural networks considered as some of the most common ones, most of which were found to work quite well through the area of text classification. In our work, the naïve Bayes was chosen as a classifier. An applying Bayes' theorem (from Bayesian statistics) was used as a straightforward probabilistic classifier for Naïve Bayes along with using strong naïve independence assumptions; it supposes that the existence/absence of a specific word in/from a class (category) is separated from the existence/absence of another word. Despite the fact that naïve Bayes is simple and makes over-simplified assumptions, it has assured to completely work in a good manner in many complex real-world positioning, and has been known to produce very good results, with high classification accuracy [9][14]. These reasons contributed to our decision to choose naïve Bayes to be our classifier.

Researches in the text categorization area were mainly restricted to English text. Many studies also contain various continental languages. For example, German, French, and Spanish as well as languages of Asian countries such as Japanese and chine's. Researches that address text that written using Arabic language is rare in literature. [2] Attempted to attain a better understanding of Arabic text classification by evaluating the rendering of two widespread algorithms of classification (SVM and C5.0) that used in the process of classifying text that written using Arabic. Another contribution is presented in the works of [10], in which highest entropy was the method that used for classifying Arabic documents. The *k*-nearest neighbor algorithm was used in [4] for Arabic text classification in an attempt to assess the performance of this algorithm in the process of classifying text written using Arabic. Finally, Rehab M. Duwairi contributed to the research in Arabic text classification in two of her papers; [8] and [9] where she compares the accuracy of three classifiers (distance-based, Naïve Bayes, and *knn*) when used for categorizing Arabic text.

In the current research we test the precision of the Naïve Bayes classifier in categorizing Arabic text using both supervised and semi-supervised techniques in an attempt to compare between them and see if the classification accuracy will increase as a result of using the semi-supervised technique. In the supervised approach the classifier first trains using a collection of labeled documents and is then given a collection documents that are unlabeled in order to automatically classify based on the information it has gained in the training process. The semi-supervised approach, however, takes it one step further; not only does the classifier train on labeled documents, but also uses unlabeled documents for training.

Training data set was collected from many online forums, magazines, and newspapers. We had a total of 1890 documents that varied in length and writing style. These documents fall into 9 different categories with a various documents' number for every category. The data set was divided into three groups: the labeled training documents, the unlabeled data, and the data set used for testing. The preprocessed documents were provided by removing all stopwords, symbols, and digits, and light stemming was used by removing some of the prefixes and suffixes from the keywords.

The dataset was divided into 70% labeled data, 20% unlabeled data, and 10% test data. The results of this experiment were measured in recall, precision, fallout and error rate. The recall value that represents the classification accuracy of the supervised learning method was 77% and the one for semi-supervised learning method was 87%. It was also proven that the semi-supervised learning method's accuracy could not be further improved if it was given an extra number of documents to classify and learn from.

The remaining of the work is arranged as: second part describes the unique features of the Arabic language and the main issues that were taken into account when the classifier was built. In third part we explain about the preprocessing we performed on our documents. Fourth part includes an explanation about the Naïve Bayes classifier, the supervised and semi-supervised approaches, the implementation of the classifier, and the final results. Finally, fifth part contains the entire conclusion for this work.

II. CHARACTERISTICS OF ARABIC

The central HYPERLINK "http://en.wikipedia.org/wiki/Central_Semitic_language"SemiticHYPERLINK "http://en.wikipedia.org/wiki/Central_Semitic_language" language is Arabic; therefore it has some relations with other Semitic languages. For example, languages of Hebrew and Neo-Aramaic. Arabic is the language that has additional speakers than any other Semitic language. It is used by huge number of users that exceeds 280 million people, and is considered as the official language of 22 countries [5].

The alphabet of Arabic consists of 28 characters:

ا ب ت ث ج ح خ د ذ ر ز س ش ص ض ط ظ ع غ ف ق ك ل م ن هـ و ي

Besides to the Arabic *hamza* (ء), which is often considered to be a letter. Three of the letters are vowels (ي و ا), while the rest of the letters are consonants. Arabic text is written from right to left. The letters of this language take different forms and shapes depending on two main things: first, their position within the word (first, middle, or end), and second, whether the letter is connectable to its next neighbors.

Arabic is a highly inflected language; In addition, a verb in its root pattern is augmented with prefixes, infixes, and suffixes to reflect the time during which the event occurred, whether the verb is plural or singular (plural is divided into *two* and *three or more*) as well as the gender of the participants in the verb.

Arabic used diacritics. Diacritics are short vowels that are written above or below a letter to indicate the pronunciation on the letter. There are four main diacritics: *fat-ha*, *damma*, *kasra*, and *shadda*, in addition to *double fat-ha* (called *tanween fateh*), *double damma* (called *tanween damm*), *double kasra* (called *tanween kaser*), and *sukun*.

Arabic may vary in meaning depending on the diacritics. Some words in Arabic vary in meaning if the diacritics change. In this case, if diacritics are not added to clarify the meaning of the word, it is considered to be ambiguous. For more illustration on ambiguous words, we site the following examples:

- The word (شِفاه) means lips. Note that there is a kasra under the first letter (ش or sheen). When this kasra is replaced with a fat-ha placed above the letter, the word becomes (شَفاه) which means cured.
- The word (نَهْرْ) means *river*. Note that there is *sukun* above both letters (هـ or *ha*) and (ر or *ra*). When both *sukuns* are replaced with a couple of *fat-has*, the word becomes (نَهَرَ) which means *scolded*.

On the other hand, many of the words do not change in meaning due to change in diacritics. The change in diacritics in their case produces words that have no meaning, and so the meaning of the word is clear even if diacritics are suppressed.

Diacritics have a big task in the meaning of the word. Unfortunately, the majority of Arabic text is written without diacritics. This is a big issue in the text classification problem, and leads to one of the complications of Arabic text categorization in contrast to the English language. The diacritics have a big task in the meaning of the word, however, most of the time they are ignored and omitted causing many words to lose their meaning or to be confused with other words (it is better to find a method to deal carefully with diacritics in "Preprocessing phase"). As a result, misclassifications are bound to happen, causing a decrease in the classification accuracy.

III. DOCUMENT PREPROCESSING

We pre-processed the documents in two ways: filtering and stemming. We filtered out any word that occurred less than 5 times in a document. Multiple experiments were held and we concluded that the removal of the words that appear less than 5 times improves the performance of the classifier.

Since Arabic is a highly inflected language, we, also, performed light stemming. As mentioned earlier, often an Arabic verb in its root form is augmented with prefixes, infixes, and suffixes. Fortunately, all Arabic words can be mapped to their root types. Arabic words can have three-, four-, five-, or six-letter roots. More than 80% of Arabic words have three-letter roots [8]. The process of root extracting from a word is called root stemming. Stemming in general includes removing any added prefixes and suffixes to the word, and it is much needed in the text classification problem for the purpose of reducing the dimensionality of the feature vector. According to [1], there are two kinds of stemming:

1) Root stemming: a technique that attempts to reduce the word to its original root.

2) Light stemming: a technique that attempts to remove only some of the prefixes and/or suffixes. It does not attempt to remove any infixes or reduce the word to its root form [3].

TABLE I. THE PREFIXES AND SUFFIXES THAT WERE ELIMINATED BY LIGHT STEMMING

Prefixes	وكال – كال – فال – وبال – بال – لل – وال – ال
Suffixes	ها

In this work, we chose to do some light stemming on the documents rather than root stemming. Our light stemming started by removing all the diacritics from our documents (it is more accurate when using diacritics), and then we removed the prefixes and suffixes, shown in Table 1 from all the words.

IV. Supervised and Semi-Supervised Naïve Bayesian Classification

One of the probabilistic classifier that considered simple is Naive Bayesian classifier [7] [6] that uses theorem of Bayes for conditional probabilities. It is called naïve; because it assumes that all values of the attribute are independent from each other given a class value (i.e. it supposes that the presence or absence of a certain feature of a class is unrelated to the presence or absence of any other feature). Despite this naïve assumption, Naïve Bayesian has been successfully used as a text classifier [13][12][11][8].

To classify a new document, the method calculates the probability of each class value, given the document's words. The maximum probability of the class is then taken as the predicted class of the document. The training set is used to estimate all needed probabilities.

Given a document that contains the words $w_1, w_2, \ldots, w_n$, a value of a class has the probability , C, is computed as

$$P(C / w_1, w_2, \ldots, w_n) = \frac{P(w_1, w_2, \ldots, w_n / C)P(C)}{P(w_1, w_2, \ldots, w_n)} \quad (1)$$

Where:

P(C) is considered as the probability of class *C*.

$P(w_1, w_2, \ldots, w_n)$ is the probability that words $w_1, w_2, \ldots, w_n$ occur in a document irrespective of their position in the document.

$P(w_1, w_2, \ldots, w_n / C)$ is the words $w_1, w_2, \ldots, w_n$ will appear in a document of class C.

Since, given a document, the probability $P(a_1, a_2, \ldots, a_n)$ is the same regardless of the class, therefore, formula 1 can be simplified as follows:

$$P(C / w_1, w_2, \ldots, w_n) = P(w_1, w_2, \ldots, w_n / C)P(C) \quad (2)$$

The approach got its name because it naively assumes that values of the attribute are conditionally distinct stated the value of the class. Therefore, it assumes that

$$P(w_1, w_2, \ldots, w_n / C) = \prod_i P(w_i / C) \quad (3)$$

The documents are modeled as groups of words where the i-th word of a certain document has the probability that happens from class C in a document is written as P (wi|C). It is assumed that the position of the word within the document is not relevant.

A. Semi-Supervised Text Classification

Supervised learning uses a training set that consists of manually classified documents. Naïve Bayesian uses this training set to estimate all required probabilities. Therefore, the larger this set, the more accurate the estimations are. However, preparing a large training set is a tedious task that requires effort and time.

Semi-supervised classification [14] attempts to make use of unlabelled (unclassified) documents to increase the classification accuracy of a classifier; Initially, just like a supervised approach, the classifier is trained using a set of classified documents. The classifier is, then, given a set of unlabeled documents to classify. The newly classified documents are then added to the pool of training documents and the new bigger set is then used to re-estimate all needed probabilities. So actually, the algorithm is learning partially from unlabeled data.

The two main steps in semi-supervised learning are called EM [14]:

(E-step): utilize the naïve Bayes classifier to approximate the classification for each unlabeled document.

(M-step): the classifier is re-estimated given the new labeled documents.

V. Experiments and Results

Our data set was gathered from online forums, magazines and newspapers. We used a total of 1893 documents that vary in length and writing style. The documents fall into 9 different classes: Economics, Computer Science, Education, Engineering, Politics, Law, Religion, and Sports, with a different number of documents for each class. The whole documents for each classification are shown in Table 2**Error! Reference source not found.**

TABLE II. The # of Documents in Each Class

Category	# of Documents
Computer	120
Economics	270
Education	118
Engineer	165
Law	147
Medicine	283
Politics	232
Religion	277
Sport	282
Total	**1893**

The number of fold cross validation that used in our experiments is ten. This means that the each experiment was repeated 10 times, using a different subset of 10% of a test set of as the training data, each time. In each fold, the training data, which comprise of 90% of the original data set, was partitioned into 70% labeled documents, used to train the classifier in a supervised way, and 20% unlabeled documents used to, further, train the classifier in a semi-supervised way.

At each fold, the classifier was trained in a supervised way using only 70% of the original training data, then the accuracy was measured using the test data. This accuracy is reported as the result of supervised learning. The classifier was then further trained, using the unlabeled documents, in a semi-supervised way. The same test data was used to measure the classification accuracy. This accuracy is reported as the accuracy of semi-supervised training. This process was repeated 10 times, using a different test set of 10% data each time. Table 4 shows the average 10-fold classification accuracy for each category of documents. Semi-supervised learning was performed as batch learning; in the sense that, all unlabeled documents were labeled (classified) first, and then the probabilities were re-calculated. Also, the vocabulary list was updated to include the new words that appeared in the unlabelled documents (as "features extraction" process).

We can determine the accuracy of the classifier by expressing terms of recall, precision, fallout, and error percentage . To enlarge elaboration on the formulae of the four terms consider a binary classification matter (i.e., there are only one category and n documents that require to be classified), so a given document either belongs to this category (i.e., positive example) or does not belong to that category (i.e., negative example). presume that the classification is carried out by two classifiers: the first is a human and the second is a computer program. Then *recall (Re), precision (Pr), both of fallout,* and *error rate* are calculated as

$$Re = \frac{a}{(a+c)}$$

$$Pr = \frac{a}{(a+b)}$$

$$Fallout = \frac{b}{(b+d)}$$

$$Error\ rate = \frac{(b+c)}{(a+b+c+d)}$$

Where a = number of documents that both the human and the computer classify as positive examples, b = number of documents that the human classifies as negative examples but the computer classifies as positive examples, c = number of documents that the human classifies as positive examples but the computer classifies as negative examples, d = number of documents that both the human and the computer classify as negative documents, and $a + b + c + d = n$ (all test documents) [8].

Table 3 shows the result of comparing supervised learning and semi-supervised learning in terms of four accuracy measures: recall, precision, fallout, and classification error. It is obvious from the table that semi-supervised learning improved the results in terms of the four accuracy measures. The average recall of supervised learning is 76.33%. It rose to 84.67% using semi-supervised learning. Similarly, precision rose from 78.87% using supervised learning to 85.37% using semi-supervised learning. The fallout and error, also, fall down from 2.58% and 4.73% to 1.81% and 3.09%, respectively.

TABLE III. THE CLASSIFICATION ACCURACY FOR EACH CATEGORY (CLASS) OF DOCUMENTS USING SUPERVISED AND SEMI-SUPERVISED METHODS

	Supervised				**Semi-Supervised**			
	Recall	Prec.	Fall	Error	Recall	Prec.	Fall	Error
Computer	88.00%	94.20%	0.40%	1.10%	88.00%	92.40%	0.50%	1.20%
Economics	79.00%	66.50%	6.10%	7.20%	80.00%	77.80%	3.50%	5.70%
Education	58.00%	83.60%	0.70%	3.20%	62.00%	83.50%	0.80%	3.00%
Engineer	83.00%	75.60%	0.30%	4.10%	89.00%	92.40%	0.80%	1.70%
Law	72.00%	50.80%	6.00%	7.70%	75.00%	67.10%	3.10%	4.80%
Medicine	81.00%	93.30%	1.10%	2.00%	93.00%	98.10%	0.30%	1.40%
Politics	76.00%	74.90%	3.70%	6.30%	85.00%	79.70%	3.10%	4.70%
Religion	87.00%	80.50%	3.80%	5.10%	92.00%	84.50%	3.00%	3.80%
Sport	63.00%	90.40%	1.10%	5.90%	98.00%	92.80%	1.20%	1.50%
Average	**76.33%**	**78.87%**	**2.58%**	**4.73%**	**84.67%**	**85.37%**	**1.81%**	**3.09%**

At this point, one issue merits further investigation. Will the classifier give better performance if it was fed more unlabeled documents to classify and then learn from (i.e. will semi-supervised learning continue to improve the results)?

To answer this question, we collected (downloaded), yet, another set of documents. This new set consisted of 90 documents, 10 documents of each category. We trained the classifier (that we got of the semi-supervised phase) in a semi-supervised way, using the new 90 documents. We compared the results of the two experiments to see if there is any improvement on the classification accuracy of the algorithm; the results showed no improvement and the classification accuracy of the classifier remained the same as the one in the original experiment. This experiment does not prove that no further improvement is possible using semi-supervised learning, but at least, it shows that further improvement becomes more difficult to achieve as the error rate becomes smaller.

VI. A ROUGH SET-BASED APPROACH

Rough set methods can be used and applied here to improve the classification accuracy by feature selection. These methods based on mathematical and statistical calculations drive the algorithm to eliminate some of attributes. [15][16]

VII. FEATURES EXTRACTION BY PFC

Systematic features extraction is a main process of documents classification. Hence, taking a care of this phase does not lost the time, it is a valuable investigation to choose a clever method and fast to extract feature from the given documents. Using PFC (principle-feature classification) to extract the feature using a sequential method and pruning the used data may give the algorithm more efficiency and accuracy. [17]

VIII. CONCLUSION

This work demonstrates that the learning of semi-supervised can develope the accuracy of classification for Arabic documents, but this improvement becomes more difficult as the error rate becomes smaller. We used the Naïve Bayesian algorithm as to train the classifier. As the Arabic language is a highly inflected language, we performed light stemming on the documents. Semi-supervised learning gave better results than supervised learning only when we used batch learning and allowed the list of vocabulary to be dynamic. It turned out that adding the new words that appeared in the new documents to the list of vocabulary during training was essential to improve the classification accuracy.

REFERENCES

[1] Al-Ameed H., Al-Ketbi S., Al-Kaabi A., Al-Shebli K., Al-Shamsi N., Al-Nuaimi N. and Al-Muhairi S., (2005), "*Arabic Light Stemmer: A new Enhanced Approach"*, The Second International Conference on Innovations in Information Technology (IIT'05).

[2] Al-Harbi S., Almuhareb A., Al-Thubaity A., Khorsheed M. S., Al-Rajeh A., (2008), "Automatic Arabic Text Classification", In *Proceedings of The 9th International Conference on the Statistical Analysis of Textual Data*, Lyon-France.

[3] Al-Kharashi I. And Al-Sughaiyer I. , (2004), "*Performance Evaluation of an Arabic Rule-Based Stemmer"*, The 7th National Conference on Infromation Technology and Computers, King Abdulaziz University, Saudi Arabia.

[4] Al-Shalabi R., Kanaan G., and Gharaibeh M. H., (2006), "*Arabic Text Categorization Using kNN Algorithm"*, Proceedings of The 4th International Multiconference on Computer Science and Information Technology, Vol. 4, Amman, Jordan.

[5] Britannica, 2011. Retrieved in April, 2011 from http://www.britannica.com/EBchecked/topic/31677/Arabic-language

[6] Domgos P, Pazzani M. (1996). "Beyond Independence: conditions for the optimality of the Simple Bayesin Classifier". The 13th International Conference on Machine Learning.

[7] Duda R. , Hart P. (1973). "*Pattern Classification and Scene Analysis"*. John Wiely and Sons.

[8] Duwairi R. M., (2006), "*Machine Learning for Arabic Text Categorization"*, Journal of the American Society for Information Science and Technology, Vol. 57, No. 8, p. 1005-1010.

[9] Duwairi R. M., (2007), "*Arabic Text Categorization"*, The international Arab Journal of Information Technology, Vol. 4, No. 2.

[10] El-Halees A. M., (2007), "*Arabic Text Classification Using Maximum Entropy"*, The Islamic University of Gaza, Journal of Series of Natural Studies and Engineering Vol. 15, No. 1, pp 157-167.

[11] Manning C., R. P. (2008). "*Introduction to Information Retrieval"*. Cambridge University Press.

[12] McCallum A. and Nigam K. (1998). "A comparison of event models for naive Bayes text classification". In AAAI workshop on learning for text categorization.

[13] Mitchell T. M., (1997), "*Machine Learning"*, McGraw-Hill.

[14] Nigam K., McCallum A., Mitchell T. M., 2006, "*Semi-supervised Text Classification Using EM"*, In Semi-supervised Learning, O. Chapelle, A. Zien, and B. Scholkopf (Eds.), MIT Press.

[15] Alexios Chouchoulas1, Qiang Shen1., 1999, "*A Rough Set-Based Approach to Text Classification*".

[16] Libiao Zhang∗, Yuefeng Li†, Chao Sun‡, Wanvimol Nadee§., 2013, "*A Rough Set-Based Approach to Text Classification*".

[17] Donald W. Tufts and Qi Li., 1997, "*PRINCIPAL-FEATURE CLASSIFICATION*".

Geographical Information System Based Approach to Monitor Epidemiological Disaster: 2011 Dengue Fever Outbreak in Punjab, Pakistan

Shahbaz Ahmad
Department of Computer Science
NTU
Faisalabad, Pakistan

Muhammad Yasir
Department of Computer Science
UET
Faisalabad Campus, Pakistan

Muhammad Majid
IT Department, GPGC Samnabad
Faisalabad, Pakistan

Muhammad Asif
Department of Computer Science
NTU
Faisalabad, Pakistan

Shahzad Nazir
Department of Computer Science
NTU
Faisalabad, Pakistan

Muhammad Umar Chaudhry
Deptt of elec & comp eng,
Sungkyunkwan University
Suwon, republic of Korea

***Abstract*—Epidemiological disaster management, using geo-informatics (GIS), is an innovative field of rapid information gathering. Dengue fever, a vector-borne disease, also known as break bone fever, is a lethal re-emerging arboviral disease. Its endemic flow is causing serious effects to the economy and health at the global level. Even now, many under-developed and developing countries like Pakistan lack the necessary GIS technologies to monitor such health issues. The aim of this study is to enhance the effectiveness of developing countries through disaster management capabilities by using state-of-the-art technologies, which provide the measures to relief the disaster burden on public sector agencies. In this paper, temporal changes and regional burden for distribution of this disease are mapped using GIS tools. For the prevention of disaster burden, these types of studies are widely used to provide an effective help and relief. This study concludes that a public sector institute can use such tools for surveillance purpose and to identify the risk areas for possible precautionary measures.**

Keywords—GIS; Dengue; Hemorrhagic fever; Aedes aegypti

I. INTRODUCTION

Dengue virus first appeared at the start of World War II [1]. Dengue fever is also known as break bone fever. It is a lethal re-emerging arboviral disease whose endemic flow [2] is having serious effects on the economy and health at the global level. It is a mosquito-borne fellow of Flaviviridae clan, which is the contributory fever [3]. Dengue is found in tropical and a sub-tropical area around the globe, mainly in urban and semi-urban region. The vital force behind the growth and development of dengue virus is temperature and climate. Climate and temperature are two known natural resources, which are not under human control.

Therefore, precautionary measures should be taken against dengue, and the breeding of mosquitoes should be avoided. It is observed that time period of growth from egg to adult is inversely related to temperature. It ranges from 7.2 ± 0.2 days at 35°C to 39.7 ± 2.3 days at 15°C[4]. Dengue virus is now almost blown out over the globe. In recent years, due to fast communication and transportation, it is estimated that there are 50 to 100 million cases of dengue fever and about 500,000 cases of dengue haemorrhagic fever per year, which requires hospitalization [5].Dengue fever is normally caused by the bite of an infected *Aedes aegypti* mosquito, while symptoms of the disease automatically appear in about 5 to 7 days.

Dengue hemorrhagic fever was described in Southeast Asia, Manila in 1953[6] and even in the start of 1950 about 9 countries were its victim. Dengue fever became more public in 1980 while at the end of 1990s, it became the most substantial mosquito-borne disease [7]. After malaria, dengue is the disease that is distressing humans in a time span of 40 years. In 2009, an epidemic broke out in Bolivia, where 31000 cases were registered [8].

The basic objective behind this research work is to portrait the affected area using GIS-based techniques into satellite images. The level of intensity and geographical identification of registered cases may help the government agencies to take preemptive measures before and also for future planning to recover from disaster situation.

II. MATERIALS AND METHODS

In order to find out the intensity of dengue in Punjab, we collected the registered cases from government and private hospitals inside Punjab. Figure 1 shows map of Pakistan highlighting the Punjab province. To carry out this research work, we practically adopted a qualitative research method as elaborated by John Creswell [9]. For the collection of data, a survey was designed specifically for the dengue patients. It includes all the symptoms related to this particular disease inside the patient. It also includes the demographical information of patients and their belongings. On the basis of collected data, we used the GIS-based application to embed the number of registered cases district wise.

III. STUDY AREA

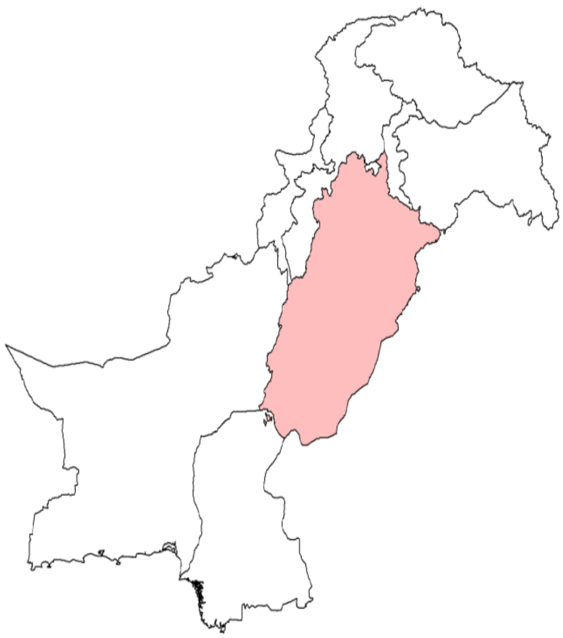


Fig. 1. Pakistan map reflecting punjab area

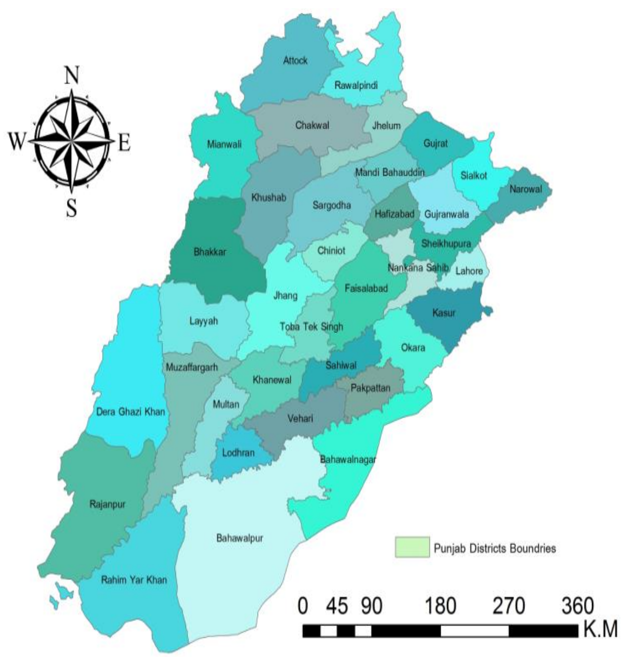


Fig. 2. Punjab Map reflecting Districts

Pakistan comprises four provinces while Punjab is major province of the Pakistan with respect to population, industry, and rich agricultural resource. Punjab is the most established, populated, and flourishing province of Pakistan. Figure 2 shows the map of the Punjab at district level. As it is hub of cultural, historical economic and administrative activities. Almost 62% population of the Pakistan residing in the province of Punjab [10].We can say that in the development and progress of Pakistan this region of Punjab play vital role while if any dead lock occur in this region then whole country is affected as it having massive impact upon country. In the last few years, the dengue virus has spread throughout the country particularly in Punjab. It became an epidemic in a couple of years, especially in 2011.

Pakistan is a developing country with lack of resources and poor management. The number of patients is increasing yearly due to lack of awareness [11].In year 2011, about 300 people died while over 14000 were infected by dengue. Majority of the population affected due to dengue was from Lahore which is also the provincial metropolis of Punjab. In Pakistan, normally patient data in hospitals is not in electronic form. It is noted in registers in manual fashion so it was very hard to move from district to district and hospital to hospital to collect data of dengue patients.

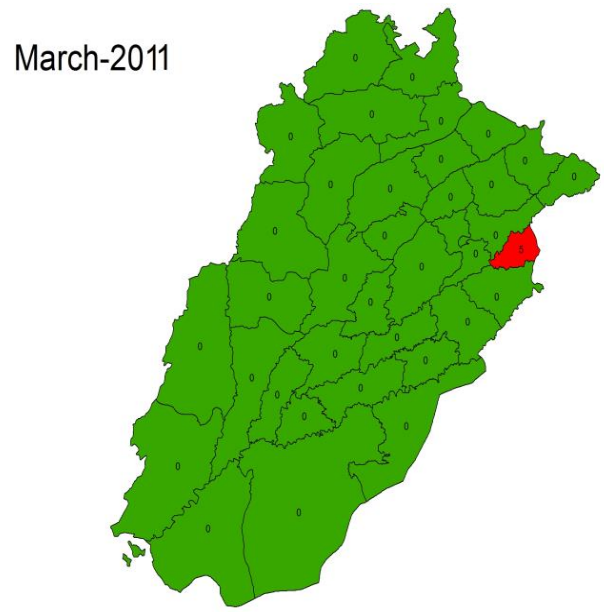


Fig. 3. Indicating origin of dengue in Punjab

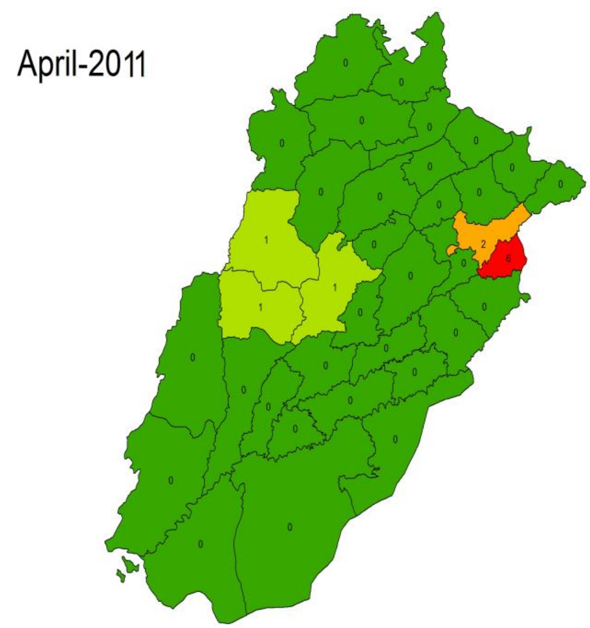


Fig. 4. Indicating flow and intensity of dengue in April

IV. RESULTS

In March 2011, dengue started in Lahore accidently and about five patients were admitted to different hospitals. Figure 3 indicates origin of dengue in Punjab in March. So it was dengue debut in Lahore in the province of Punjab as Lahore is located on the border of India and most populated city of the province in Pakistan. March is the end of winter and this period is not suitable for the growth of dengue virus as this virus was transferred by some foreign country through people and products [12]. Thus, Lahore proved to be a launching pad for dengue and in the very next month of April, the virus moved to the nearby city Sheikhupura which is just 37 km away from Lahore as it was natural for the virus to reach the nearest area as local bodies from Sheikhupura daily travel to Lahore and it is the commercial hub for the Punjab area. Figure 4 indicates the flow and intensity of dengue in April.

Meanwhile in the same month, viral attack moved directly to the Jhang, Leyya, and Bhakkar. This was due to daily transport of passengers. However, it was very minute in these four districts because only a single case was observed per district. So it was actually a start of the endemic flow of dengue. Actually this climate was not suitable for the growth of dengue virus. Treatment was also provided soon and dengue moved away from these four districts. However, it was still present in Lahore though on a small scale.

In the month of April, the weather starts getting slightly hot without any rain and it is the season of ripening of wheat and not favorable for dengue. In May, dengue moved to the district of Sargodha, Sahiwal, and Saialkot and this was due to transportation of passengers from Lahore to these areas. Figure 5 shows flow and intensity of dengue in May. The common point here is that Lahore is the dengue hub in Pakistan. In May, dengue cases decrease due to hot weather.

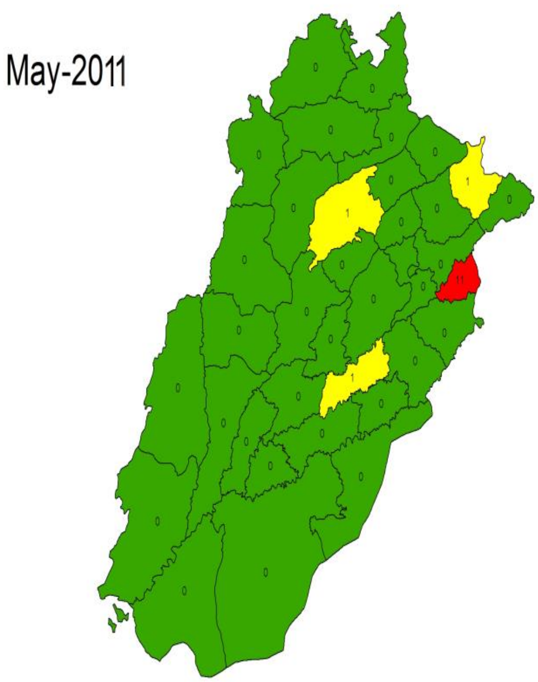


Fig. 5. Indicating flow and intensity of dengue in May

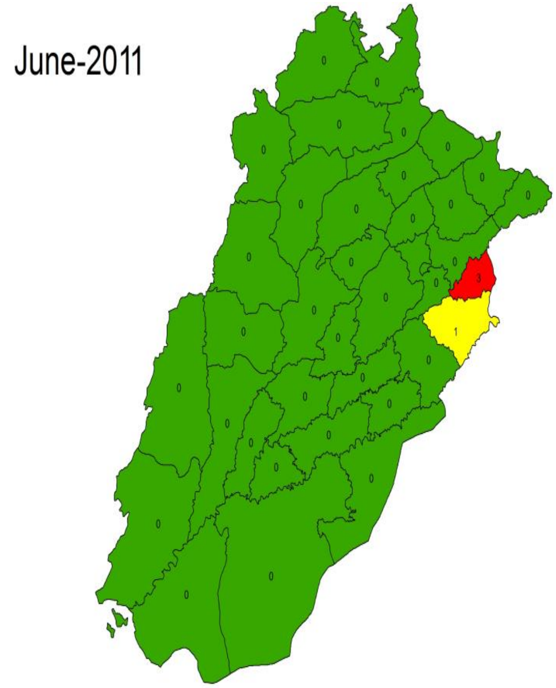


Fig. 6. Indicating flow and intensity of dengue in june

From March to May, the dengue virus moves through different districts of Punjab. However, its roots are strong in Lahore. In June, due to the weather, mosquitoes are completely wiped out. In June, the Kasur district was infected with just a single patient as this area is just 25 km away from Lahore. Figure 6 shows flow and intensity of dengue in June. July is also a hot month but with humidity and the growth of dengue still remained limited to Lahore with eight patients and a single patient was found in Attock. Figure 7 shows flow and intensity of dengue in July. Thus, in July, dengue created no problem for the government and public while the next month saw the spread of dengue in all of Punjab.

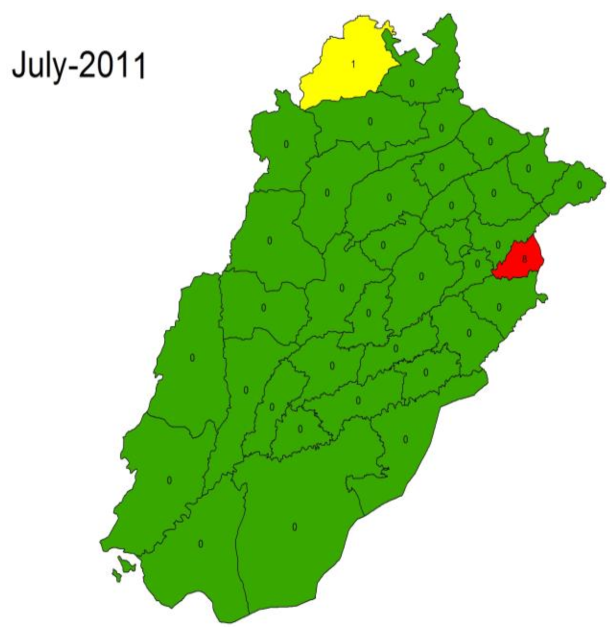


Fig. 7. Indicating flow and intensity of dengue in july

Pakistan is located at a place where rains gear up at the end of July or beginning of August and humidity is also more. This climate is favorable for the growth of dengue virus and it widely spreads out in the month of August covering 27 of the 36 districts (75% of the area). The affected area can be seen in the map below and Lahore, Faisalabad, Multan, and Sheikhupura were most badly affected. Figure 8 shows the flow and intensity of dengue in August.

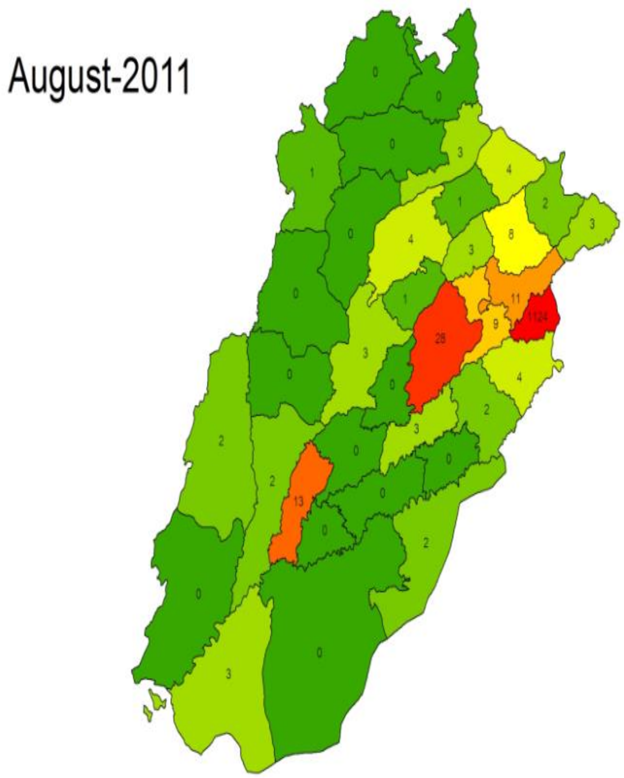


Fig. 8. Indicating high flow and intensity of dengue in August

In the month of August, almost whole of Punjab was affected and this caused loss of lives and fear among people. The government started giving out advertisements on prevention of dengue in an organized way for people to take precautionary measures. Some schools and colleges also announced holidays to control the intensity of the disease. In August, dengue covered almost 75% of the area of Punjab but its concentration was limited to four districts.

In the month of September, it covered 100% area of Punjab province and no district was without the viral infection. Its intensity also increased and spread over 14 districts as almost 39% area of Punjab was so badly infected. Figure 9 shows the flow and intensity of dengue in September. While 17% of the area was facing average intensity, about 44% area had a below average level. In September, about 11,000 patients were registered in different hospitals of Punjab. In the month of October, dengue devastation boomed much when compared to the month of September. Government agencies started spraying pesticides on the possible growth areas to control the spread of dengue.

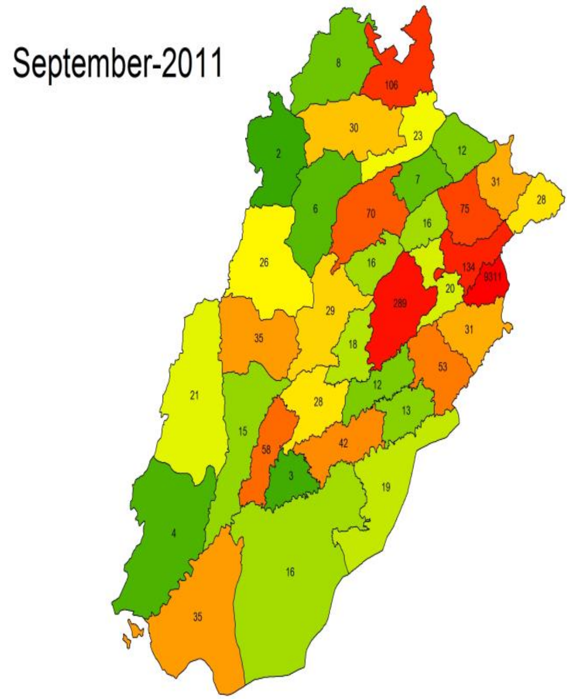


Fig. 9. Indicating high flow and intensity of dengue in September

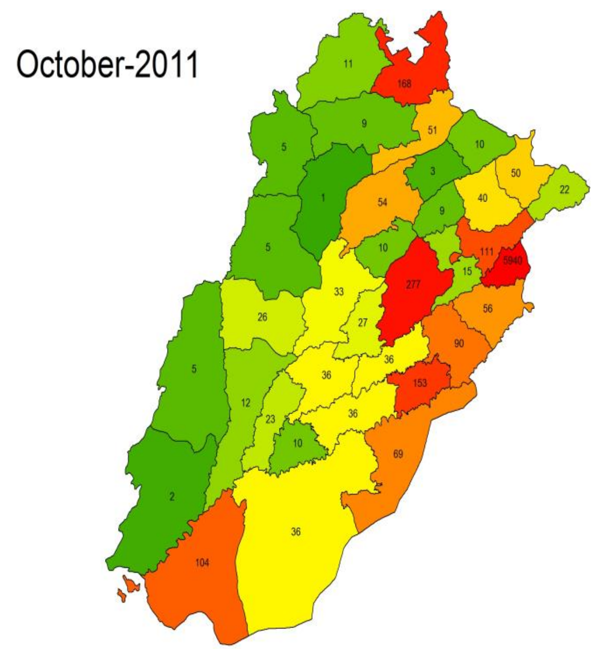


Fig. 10. Indicating high flow and intensity of dengue in October

In Lahore, Faisalabad, Sahiwal, Rawalpindi, and Rahim Yar khan, the intensity level of dengue remains the same in September and October. Figure 10 shows flow and intensity of dengue in October. While in November, the intensity lowers to about 19% compared to previous months. This also helped the government to fumigate the doubtful areas, and educate public through media and other modes of awareness. Figure

11 shows the intensity and flow of dengue in November. Dengue-hub Lahore, Faisalabad along with Rawalpindi had the same intensity of dengue viral cases while the area of Bahawalpur, Rajanpur, and Layya were completely off dengue. Kasur, Okara, Pakpatan, and Sahiwal have more than the average incidents of dengue cases. Figure 12 indicates scaling criteria with respect to distance.

In December, surprisingly, the intensity was way lower (20%) with respect to the number of cases arriving from different districts. Figure 13 shows the intensity and flow of dengue in December. Thus, 80% of the public got relief from dengue in December as in this month, the temperature falls by about 0-8 centigrade. and dengue death occur due to high chill climate and in this way in December the dengue story come to and which left permanent mark upon the families of Punjab who lost their own, restless sleep, government spent millions of rupees for advertisement and equipment, heavy economic loss, public face heavy strain, education sector was also effected, many doctor teams from Sri Lanka, Thailand, and other countries also came for rescue of public from dengue.

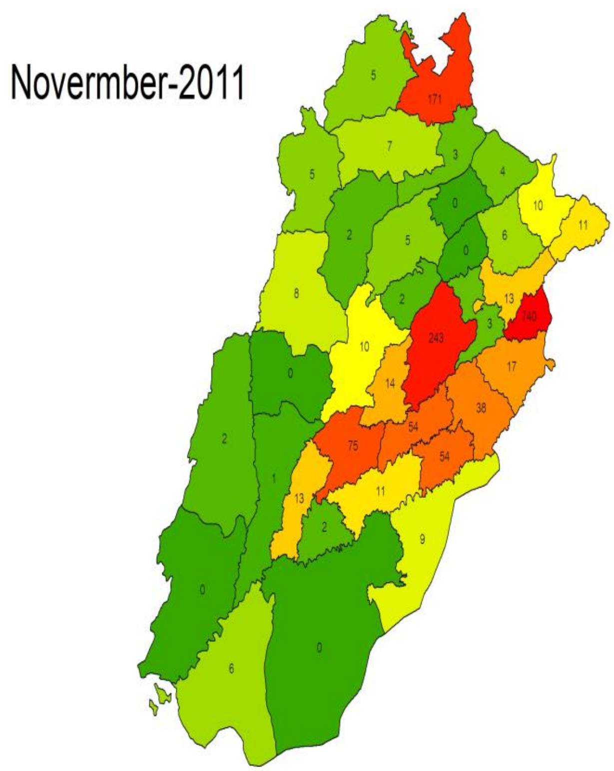


Fig. 11. Indicating high flow and intensity of dengue in November

When we observe the final map, we come to know that about 21,000 patients were influenced through this viral disease. Figure 14 shows the overall intensity and flow of dengue in the year 2011. Lahore proved to be a hub for dengue flow and top district with about 17,234 patients while Rajanpur was the least-affected district with just 6 patients.

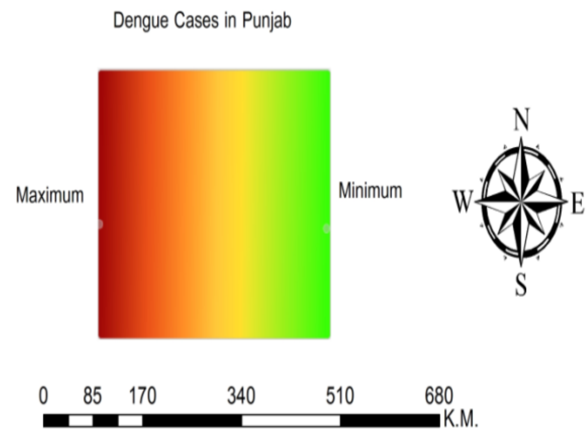


Fig. 12. Indicating scaling criteria with respect to distance

The dengue rampage was observed at Lahore, Faisalabad, Okara, Pakpatan, Khanewal, Rahim Yar Khan, Sheikhupura, Rawalpindi, and Sargodha, as these districts are adjacent to each other except, Rahim Yar Khan and Rawalpindi. Figure 15 indicates scaling criteria with respect to distance for final results.

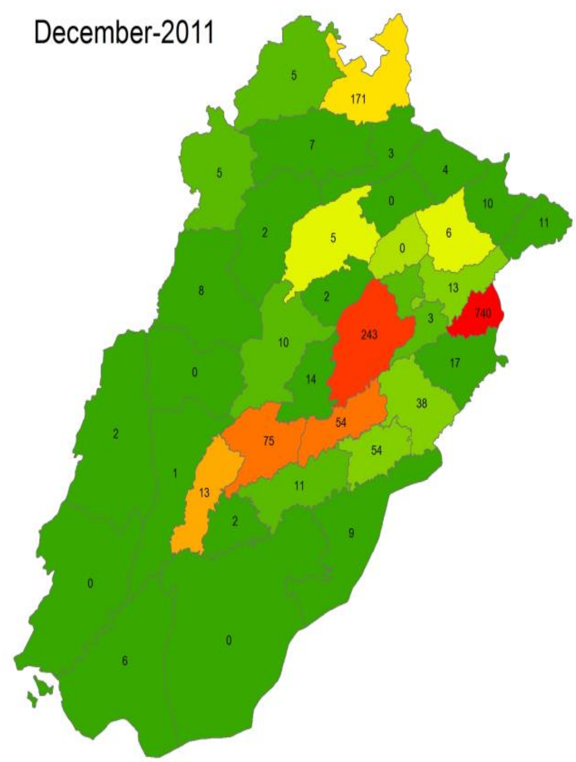


Fig. 13. Indicating high flow and intensity of dengue in December

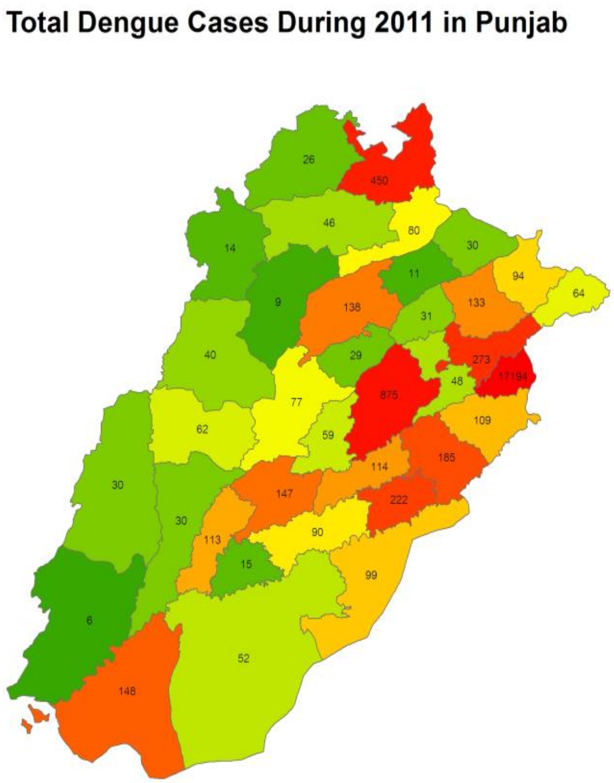


Fig. 14. Indicating over all flow and intensity of dengue in 2011

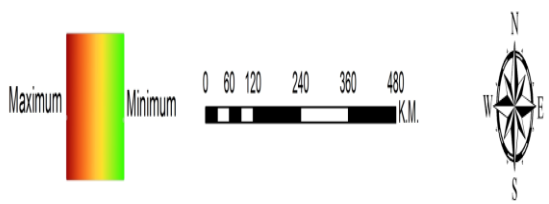


Fig. 15. Indicating scaling criteria with respect to distance for final results

V. Discussion

We know that dengue disaster geared up from Lahore accidently in the month of March 2011 while the weather of Lahore around that time was the end of winter season. This means that it was not a suitable climate for mosquitoes. The question that arises is from where did dengue arrive in Lahore. Investigations indicated that it arrived from Bangkok,Thailand while its real source was 1000 second hand tractor tyres. These tyers contain some water inside them where dengue eggs were present and this resulted in the spread of the disease. We know that Lahore is the capital city of Pakistan and the commerical hub so due to transportation this virus moved to Layya, Jhang, and Bhakkar However, soon the virus vanished from these districts and emerged in others. The incidence was however not regular due to unfavorable weather round the year.

The real outbreak occurred in August when the rainy season starts since the temperature and weather are in favor. It covered the 75% area of Punjab and in later months it cover 100% area and now the expension of dengue was due to public movement because a dengue can cover maximum 300 meters area.so real source from district to district is fast transportation which cause to deploy viral infection from one place to other place.In tropical parts, dengue diffusion occurs throughout the year,while the temperature and humidity favour the existence of adult mosquitoes beyond their extrinsic incubation period. We know that winter and summer are not suitable for spread of dengue. It spreads mostly during the rainy months. The sprays destroy the eggs and future growth can be avoided. One female mosquito produces about 300 eggs in the life span of 14 days. Thus, the real task is to wipe away and destroy its eggs so that further growth can be avoided and life, economy, and health can be saved. Figure 16 indicates overall eruption district wise. Table 1 shows number of patients district wise.

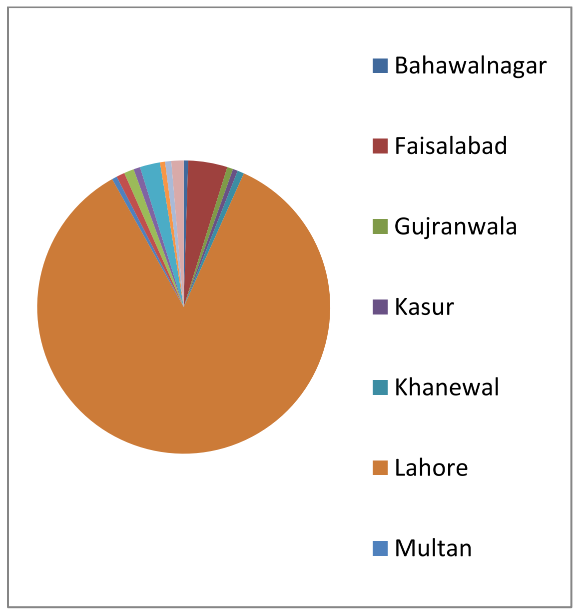


Fig. 16. Indicating over all eruption district wise

TABLE I. Reflecting Patients District Wise

Id	Dist code	District	Patient
1	601	Attock	28
2	602	Bahawalnagar	99
3	603	Bahawalpur	52
4	604	Bhakkar	40
5	605	Chakwal	46
6	606	Chiniot	29
7	607	D.G Khan	30
8	608	Faisalabad	875
9	609	Gujranwala	133
10	610	Gujrat	29
11	611	Hafizabad	31
12	612	Jhang	77
13	613	Jhelum	80
14	614	Kasur	109
15	615	Khanewal	147
16	616	Khushab	9
17	617	Lahore	17235
18	618	Layyah	62
19	619	Lodhran	15
20	620	M.Bahauddin	11
21	621	Mianwali	14

22	622	Multan	113
23	623	Muzaffargarh	30
24	624	Nankana	48
25	625	Narowal	64
26	626	Okara	185
27	627	Pakpattan	222
28	628	Rahim Yar	148
29	629	Rajanpur	6
30	630	Rawalpindi	450
31	631	Sahiwal	114
32	632	Sargodha	138
33	633	Sheikhupura	274
34	634	Sialkot	94
35	635	T.T Singh	59
36	636	Vehari	90

VI. Conclusion

GIS approach is very effective and suitable for observing and indicating the damage caused by viral diseases. It is not possible to monitor the whole area without using GIS approach. No other technique except GIS can be helpful to this extent. Data that we process throught GIS can generate information in different respects. In this research, we accumulate dengue patients data from different districts and embed it on sattellite images. Thus, we can observe and indicate the nature of the disaster and its intensity in different regions so that precautionary measures can be taken in the areas where the intensity of dengue is higher.

In future, if we develop an integrated network of all the governmental systems and private hospitals where data should be processed in realtime then it can help the governmnet to observe the current situation of disaster time to time and positive decision can be taken to see the real time map and necessary action can be take to observe this real-time GIS approach. Thus, we can save lives, indicate and observe disaster, and manage the resources according to the situation. Thus, in future, this dynamic environment real-time GIS approach should be implemented in the integrated network of all hospitals.

References

[1] Kurane I, Ennis FA. Cytokines in dengue virus infections: role of cytokines in the pathogenesis of dengue hemorrhagic fever. Sem Virol 1991;5:443–448.

[2] P. E. R. Dale, J. M. Knight, 2008.Wetlands and mosquitoes: a review, Wetlands Ecol Manage 16:255–276

[3] S.I.Hay, J.A.Omumbo, M.H.Craig (2000).Earth observation, geographic information system and plasmodium falciparum malaria in sub-Saharan Africa. Volume 47, 2000, Pages 173–174.

[4] W. Tun-Lin,T. R. Burkot,B. H. Kay effect of temperature and larval diet on development rates and survival of the dengue vector Aedes aegypti in north Queenland, Australia. Volume 14 issue 1, pages 31-37, march2000.

[5] Nisar Ahmad, Hina Afzal. Dengue fever treatment with Carica papaya leaves extracts. Asian Pacific Journal of Tropical Biomedicine (2011)330-333.

[6] Richard C Russell 1999.Mosquito-borne arboviruses in Australia:the current scene and implication of climate change for human health. Volume 28, Issue 6, 1 June 1998, Pages 955–969

[7] Duane J. Gubler, 2002.Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. TRENDS in Microbiology Vol.10 No.2.

[8] Oyewale Tomori (1999).Impact of Yellow Fever on The Developing World, Advance in virus research. Volume 53, Pages 5–34.

[9] John Creswell, 2008.Research Design, Qualitative, Quantitative, and mixed methods approach.

[10] http://dengue.punjab.gov.pk/

[11] C.G. Hayes, S. Baqar(2004). West Nile virus in Pakistan. 1. Sero-epidemiological studies in Punjab Province. Volume 76, Issue 4, 1982, Pages 431–436.

[12] Reich N. G. et al. Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. J R Soc Interface 10, 20130414 (2013).

Towards Face Recognition Using Eigenface

Md. Al-Amin Bhuiyan
Department of Computer Engineering
King Faisal University
Hofuf, Al-Ahsa 31982, Saudi Arabia

***Abstract*—This paper presents a face recognition system employing eigenface-based approach. The principal objective of this research is to extract feature vectors from images and to reduce the dimension of information. The method is implemented on frontal view facial images of persons to explore a two-dimensional representation of facial images. The system is organized with RMS (Root Mean Square) contrast scaling technique employed for pre-processing the images to adjust with poor lighting conditions. Experiments have been conducted using Carnegie Mellon University database of human faces and University of Essex Computer Vision Research Projects dataset. Experimental results indicate that the proposed eigenface-based approach can classify the faces with accuracy more than 80% in all cases.**

Keywords—Eigenvector; Eigenface; RMS Contrast Scaling; Face Recognition

I. INTRODUCTION

Face recognition has attained an overwhelming popularity in image processing, computer vision, pattern recognition, and so on. The major application areas for face recognition include human-machine interface, surveillance systems, credit card verification, security systems, financial transactions, criminal identification, and so on. Although human beings are excellent in recognizing faces, it is not evident how faces are being associated with human memory. Since faces represent complex and multi-dimensional visual information, developing a computational model for face recognition is, therefore, quiet a challenging job.

A number of approaches have been cited in literature on face recognition. Sirovich and Kirby [1] introduced the concept of eigenface for recognition and Turk and Pentland [2] have applied the approach in face classification. They derived the eigenvectors from the covariance matrix of the probability distribution over the high-dimensional vector space of facial images. Wiskott [3], *et al.* established a Gabor wavelet-based elastic bunch graph matching system to label and identify facial images where face is characterized as a graph. Each node of graphs contain a list of parameters known as jets. However, both dynamic link architecture and elastic bunch graph matching methods need huge volumes of computational cost due to point-to-point matching or graph matching and these are not appropriate for real-time requirements. Lades, *et al.* [4] represented facial images by Gabor filters and developed a face recognition system employing dynamic link topology. Shan, *et al.* [5] proposed an enhanced fisher approach employing AdaBoost architecture for face recognition. Zhang, *et al.* [5] developed a face recognition system employing histogram of Gabor feature arrangement. Liu and Wechsler [6] established a Gabor filter based identification method for dimensional reduction employing Fisher linear discriminate model. Kirby [8] and Terzopoulos, *et al.* [9] analysed the facial images on the bsis of facial characteristics. Wu [10] and Manjunath [11], *et al.* performed face recognition with feature vectors extracted from profile silhouettes. Kerin, *et al.* [12] have prposed a face recognition system employing neural network in tendem with self-organizing feature map. Nakamura, *et al.* [13] utilized isodensity maps to establish a face recognition system. Yullie, *et al.* [14] extracted feature vectors from the facial components like eyes, nose, mouth, and and employed deformable templates to the extraction of contours for facial images and outlined a face recognition procedure. Thakur et al. [15] proposed a face recognition system using principal component analysis (PCA) and radial basis function (RBF) neural network. The RBF network was designed by considering intra-class discriminating characteristics of the training images.

This paper explores Principle Component Analysis (PCA), which is applied on a characteristic dataset of facial images. The method is based on projecting images into a feature space that amounts the significant discrepancies among known facial images. These momentous features named as “Eigenfaces” are the principal components of the set of training facial images. Facial objects are then identified employing a nearest-neighbor classifier.

The rest of the paper is organized as follows. Section II highlights image pre-processing. Section III describes Principle Component Analysis. The process of face detection is addressed in section IV. Face image normalization is described in Section V. Section VI illustrates the algorithms for face recognition. The experimental result has been presented in Section VII. Section VIII draws the overall conclusions of the research.

II. IMAGE PRE-PROCESSING

The original images in the face databases contain both color and grey scale images with different illumination conditions. Therefore, to make the images contrast invariant in terms of bright or dark environments, these are processed with same RMS contrast equalization. The RMS contrast metric is given by [16, 17]:

$$C_{r,rms} = \left[\frac{1}{PQ} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} (r(p,q) - \bar{r})^2 \right]^{1/2} \quad (1)$$

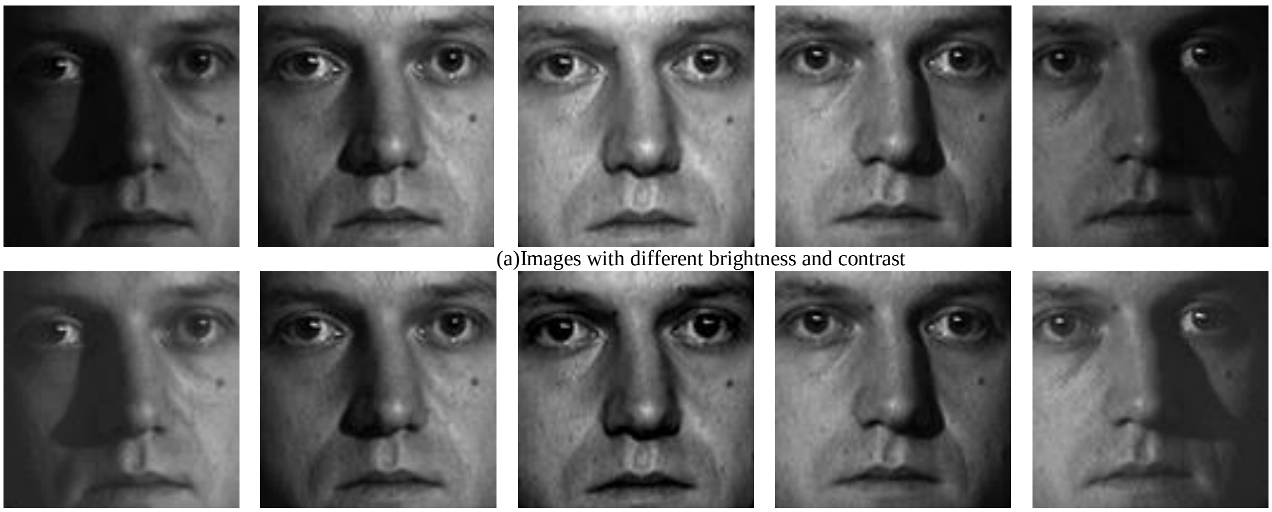
$$C_{g,rms} = \left[\frac{1}{PQ} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} (g(p,q) - \bar{g})^2 \right]^{1/2} \quad (2)$$

$$C_{b,rms} = \left[\frac{1}{PQ} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} (b(p,q) - \bar{b})^2 \right]^{1/2} \quad (3)$$

where $r(p,q)$, $g(p,q)$, $b(p,q)$ denote the illumination owing to red, green, and blue color constituents, respectively, and $\bar{r}$, $\bar{g}$, $\bar{b}$, are the mean illuminations owing to red, green, and blue color components. All facial images are retained the same lighting conditions applying the following equation:

$$\mathbf{f}_r=\alpha_r\mathbf{C}_r+\beta_r,\ \mathbf{f}_g=\alpha_g\mathbf{C}_g+\beta_g,\ \mathbf{f}_b = \alpha_b\mathbf{C}_b + \beta_b \quad (4)$$

where α_r, α_g, α_b, are the contrast due to red, green, and blue color constituents, respectively, and β_r, β_g, β_b denote the amount of brightness needed to be increased to or decreased from the respective red, green, and blue constituents $\mathbf{C}_r$, $\mathbf{C}_g$, $\mathbf{C}_b$ of the original color image **C** to the new color image **f**. The outcome of the RMS contrast equalization technique different images is shown in Fig. 1.



(a)Images with different brightness and contrast

(b) Images with same RMS contrast

Fig. 1. RMS contrast equalization. Images were captured at illumination angles of -38.4º, -21.6º, -0.2º, 20.6º, 37º, respectively, all images at a pose of 0º

III. Principal Component Analysis

Facial features are extracted employing principal components which are significant for face perception. This method is commenced with the computation of eigenvectors from the initial set of face images. New facial images are projected into the space stretched by eigenfaces and represented by weighted sum of the eigenfaces. These weights are applied to recognize the faces. The objective of the PCA is to yield the complete discrepancy on the training set of images and to describe this variation with a few variables. When the dimension of training data is increased, dimension of space complexity becomes a vital issue. The space complexity is extremely redundant when it represents faces because each pixel in a face is greatly correlated to other pixels.

PCA, a nonparametric statistical method concerned with explaining the covariance structure of a set of variables, is used to highlight the variation and convey strong patterns in a dataset. It allows us to explore, identify, and visualize the principal directions in which the data varies and representing the data to focus their resemblances and divergences. The principal reason behind using PCA is to decrease the dimension of space complexity. The maximum number of principal components is the number of variables in the original space [18]. The linear transformation maps the original n-dimensional space into an m-dimensional feature subspace. To reduce the dimension, some principal components can be discarded. The eigenfaces are the principal components of the original face images, achieved by the decomposition of PCA and eigenfaces are constructed employing eigenvectors [19].

A. Eigenvectors

An eigenvector, the representative vector of a square matrix, is a vector that does not change its direction under the associated linear transformation. Let **v** be a vector (**v**≠0), then it is an eigenvector of a square matrix **A**, if **Av** is a scalar multiple of **v**.

B. Eigenfaces

Eigenfaces are the set of eigenvectors which are used for human face recognition. These are represented as the eigenvectors which specify one of the dimensions of face image space. Eigenfaces provide significant characteristics that express the deviation in the group of face images. Each eigenvectors belongs to an eigenvalue associated with it and the eigenvectors having greater eigenvalues deliver more information on the face variation than the ones with lesser eigenvalues. Any new face image can be characterized as linear arrangement of these eigenfaces.

IV. Face Detection

A number of methods have been proposed for face detection, such as facial features extraction, knowledge-based approach, template matching, and color segmentation. This paper has combined template matching, skin color segmentation, and feature invariant approaches for face detection.

Skin color segmentation is established on the visual information of the human skin colors extracted from the image sequences. The human skin color differs from person to person and of different races, that is, chrominance and luminance components are different for different persons even in the same illumination environments. This research employs HSV color model for skin color segmentation.

In the HSV color space, a color is designated by three characteristics: hue, saturation, and value. Hue is the feature of visual impression that relates to color sensitivity linked with the prevailing colors, saturation infers the relative purity of the color component and value indicates the brightness of a color. The conversion from RGB space to HSV space is expressed by the equations [20-22]:

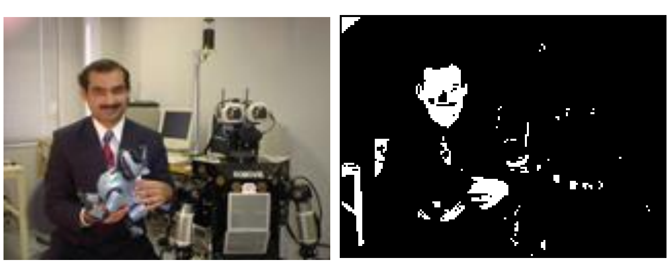
$$H = \begin{cases} \arccos \dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}, & B \le G \\ 2\pi - \arccos \dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}, & B > G \end{cases} \tag{5}$$

$$S = \frac{\max(R,G,B)-\min(R,G,B)}{\max(R,G,B)}, \qquad V = \frac{\max(R,G,B)}{255}, \tag{6}$$

where R,G,B are the red, green, and blue constituent values which exist in the range [0,255].

Facial images are thresholded employing the hue histogram of the respective image. In this research, the hue values are chosen h= [0, 40]. The detection of face area by such a hue segmentation process is illustrated in Figure 2.

To locate the face, an image pyramid is built from a set of facial images with different scales and resolutions. For this, a face template is moved from left to right and up to bottom over each image in the pyramid and calculate the matching probability at each position of the image segment under the template using minimum Manhattan distance. If the similarity value is greater than some threshold value, the existence of a face at that position and resolution is expected. From that position and resolution, the position and size of the face in the original image is being evaluated.



(a) Original image (b) Color thresholding

Fig. 2. Skin color segmentation

V. FACE IMAGE NORMALIZAION

The facial images are being normalized for face recognition. The original images are ostensibly the color images. These are transformed into gray scale images. The conversion from color image to grayscale image is given by the following equation:

$$G_n = \frac{R_n + G_n + B_n}{3} \tag{7}$$

where R_n, G_n, B_n denote the red, green, and blue color constituents of the n^{th} pixel of the color image and G_n is the gray level value of n^{th} pixel of the gray scale image, and the resolution of the image is $M \times N$. The gray scale image is then scaled to 120×120 pixel using Eq. 8.

$$N(x_m, y_m) = M(\frac{x_n}{120} x_m, \frac{y_n}{120} y_m) \tag{8}$$

where the coordinate of n^{th} pixel of original gray scale image, $M(x_n, y_n)$ is converted into the m^{th} pixel of the scaled image and $N(x_m, y_m)$ is the coordinate of that pixel.

VI. FACE RECOGNITION

This research considers the face recognition system dividing into two parts: initialization and recognition. In the initialization phase, the system is learned by creating eigenvectors of a training set of face images. The fundamental procedure employed for face recognition process is shown in Figure 3.

Algorithm 1 is applied for initialization of face recognition system.

Algorithm 1 (Initialization)

Input: A set of facial images known as Training set (Γ_k) .

Output: Form feature vectors for each image.

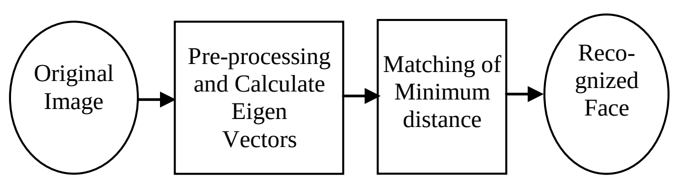


Fig. 3. Basic Steps used for Face Recognition

Method: The feature vector is constructed in the following steps.

1. Calculate mean matrix Ψ. Then subtract this mean from the original faces (Γ_k) to calculate the feature vector (φ_k), where

$$\psi = \frac{1}{N} \sum_{k=1}^{N} \Gamma_k \quad , \tag{9}$$

2. Find the covariance matrix **c** as follows:

$$\mathbf{c} = \frac{1}{N} \sum_{n=1}^{N} \left(\Gamma_k - \psi\right)\left(\Gamma_k - \psi\right)^T \quad (10)$$

3. Compute the eigenvectors and eigenvalues of **c**.

4. The N significant eigenvectors are selected on the basis of biggest equivalent eigenvalues.

5. Project all the face images into these eigenvectors and form the feature vectors of each face image.

Fig. 4. Training face images

Fig. 5. Eigenfaces with highest eigen values

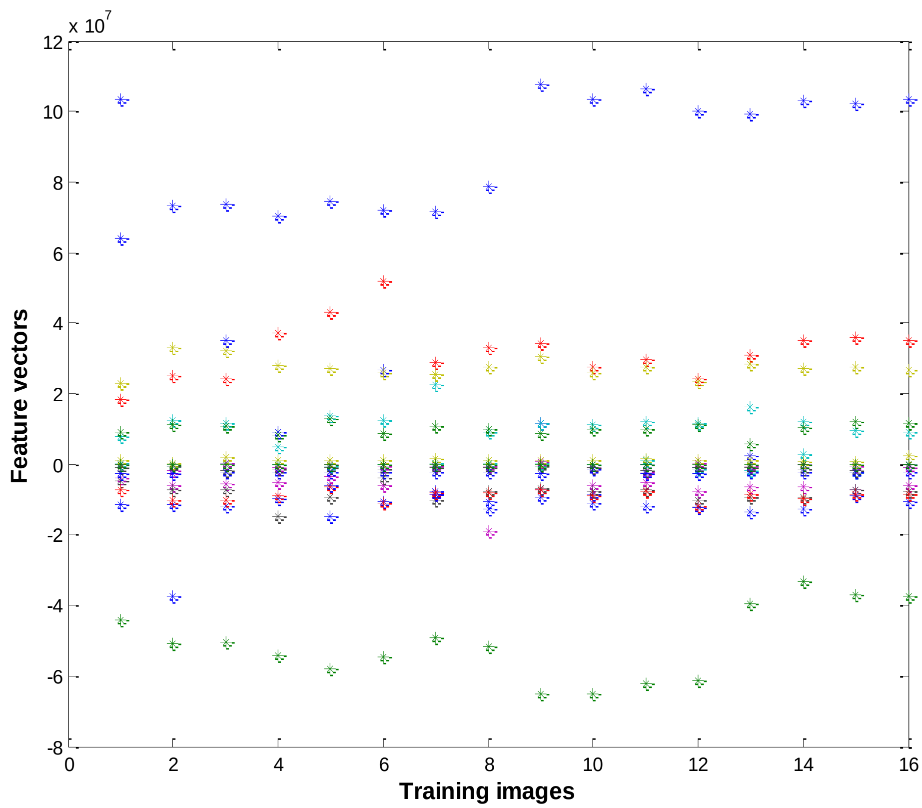


Fig. 6. Distribution of the feature vectors for a set of facial images

After getting the feature vectors, Algorithm 2 is employed to recognize an unknown face.

Algorithm 2 (Recognition)

Input: An unknown image ***I***.

Output: Recognize the image ***I***.

Method: The FP-tree is constructed in the following steps.

1. For a given input image ***I***, compute a set of vectors based on ***N*** eigenfaces by projecting the new image onto each of eigenfaces [2].

2. Compute the difference between the projected vector and each face image feature vector.

3. Classify the weight pattern as either known or unknown person.

4. The weight pattern can be compared with known weight patterns to match faces [2].

This analysis drastically reduces the order of the number of pixels in the images (N^2), to the order of the number of images in the training set (Γ_i).

A few sample face images and the corresponding eigenfaces are shown in Figures 4 and 5, respectively. Each eigenface departs from the original grayscale image where some facial features diverge among the set of training faces. Thus, eigenfaces can be regarded as a more or less mapping of the disparities between faces.

VII. Experimental Results

In order to justify the effectiveness of the algorithm several experiments were carried out with a training set of 500 images of 100 persons both male and female from CMU database [23] and University of Essex Computer Vision Science Research Projects dataset [24]. All images were in RGB color level which were normalized to gray level with dimension of 120×120. There were 70 subjects in the training set. Each subject had 5 images with frontal view with ±5° different poses (like left, right, up, and down). The training arrangement is summarized in Table 1.

TABLE I. Training Arrangement for Images

No. of images taken for the training procedure	$50\times 5 + 50\times 5 = 500$
Size	120×120
Format	BMP and PGM
Output	Normalized images of the face images and Eigenfaces

In this research, the eigenvectors of covariance matrix were computed by aggregating all deviations of training images from the average image. Since the training set contains 100 individuals, 100 eigenvectors have been used to represent the training set. The distribution of the feature vectors for a set of facial images is shown in Figure 6. Later on, images are being categorized in different lighting conditions. The overall results under different lighting conditions is shown in Table 2.

The performance of the proposed method has been compared to other similar methods which have been used the same experimental methodology. Table III shows the comparison of the performances between the proposed method (PCA+RMS contrast scaling) and the methods, as reported by [15] and [19].

TABLE II. Training Arrangement for Images

Lighting conditions	Total number of images taken for the test	Correctly recognized (Success rate)
Bright	350	97.5%
Dark	50	82%
Foggy	100	86%

TABLE III. Comparison of the performances between the proposed method and ohher similar methods

Methods	Total number of images taken for the test	Average Success rate
PCA+RBF	400	87.25%
PCA	350	80.86%
PCA+RMS contrast scaling	100	93.6%

VIII. Conclusion

This article presents an eigenface based approach for face recognition, where the eigenvectors of the covariance matrix of a set of representative images are explored. Recognition is accomplished by computing the distance between characteristic vectors from the eigenvectors space. This eigenface method is robust for face recognition which works fine under controlled environment. But the main limitation of this approach is contained in the management of face images with diversive facial expressions, lighting conditions and wearing glasses. Lighting problems are overcomed by RMS contrast stretching. The face database in this work contains 500 face images and the proposed method provides satisfactory result. It can recognize both the known and unknown images in the database in various conditions with accuracy more than 80%. Our next target is to extend the eigenface approach for live video stream so that any person can be identified at his or her own workplace.

References

[1] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces". Journal of Optical Society of America, 4(3):519-524 (1987).

[2] M. Turk and A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, 3(1):71-86 (1991).

[3] L. Wiskott, J.M. Fellous, N. Kruger, CV Malsburg, "Face recognition by elastic bunch graph matching", IEEE Transaction on Pattern Analysis and Machine Intelligence, 19(7):775–779 (1997).

[4] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C.V. Malsburg, C. Wurtz, W. Konen, "Distortion invariant object recognition in the dynamic link architecture, "IEEE Transaction on Computers", 42(3):300–311 (1993).

[5] S. Shan, P. Yang, X. Chen, W. Gao, "AdaBoost gabor fisher classifier for face recognition", Proceedings of IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 278–291 (2005).

[6] B. Zhang, S. Shan, X. Chen, W. Gao, "Histogram of gabor phase patterns: a novel object representation approach for face recognition", IEEE Transection on Image Processing, 16(1):57–68 (2007).

[7] C. Liu and K. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminate model for face recognition", IEEE Transaction on Image Processing, 11(4):467–476 (2002).

[8] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces", IEEE Transaction on Pattern Analysis and Machine Intelligence. 12(1):103-108 (1990).

[9] D. Terzopoulos, K. Waters, "Analysis of facial images using physical and anatomical models", Proceedings of 3rd International Conference on Computer Vision, 727-732 (1990).

[10] C.J. Wu and J.S. Huang, "Human face profile recognition by computer". Pattern Recognition, 23(1):255-259 (1990).

[11] B.S. Manjunath, R. Chellappa, and C. Malsburg, "A feature based approach to face recognition", Transection of IEEE, 373-378 (1992).

[12] M.A. Kerin, T.J. Stonham, "Face recognition using a digital neural network with self-organizing capabilities", Proceedings of 10th International Conference on Pattern Recognition, 738-741 (1990).

[13] O. Nakamura, S. Mathur, T. Minami, "Identification of human faces based on isodensity maps", Pattern Recognition, 24(3):263-272 (1991).

[14] A.L. Yuille, D.S. Cohen, P.W. Hallinan, "Feature extraction from faces using deformable templates", Proceedings of CVPR, 104-109 (1989).

[15] E. Peli E, "Contrast in complex images", Journal of Optical Society, 7(10): 2032-2040 (1990).

[16] S. Thakur, J.K. Sing, D.K. Basu, M. Nasipuri, " Face recognition using principal component analysis and RBF neural networks", IJSSST, 10(5): 7-15 (2012).

[17] M.A. Bhuiyan, F.A. Alsaade, "Genetic search for face detection", Proceedings of the World Congress on Engineering, I:157-162 (2015).

[18] M. Murali, "Principal component analysis based feature vector extraction", Indian Journal of Science and Technology. 8(35):1-4 (2015).

[19] D. Chakraborty, S.K. Saha, M.A. Bhuiyan, "Face recognition using eigenvector and principle component analysis", International Journal of Computer Applications, 50(10):42-49 (2015).

[20] M.A. Bhuiyan, "Content-based image retrieval for image indexing", International Journal of Advanced Computer Science and Applications, 6(6):71-79 (2015).

[21] D. Androutsos, K.N. Plataniotis, A.N. Venetsanopoulos, "A model vector-based approach to color image retrieval using a vector angular-based distance measure", Computer Vision and Image Understanding, 75(1):46-57 (1999).

[22] M.A. Bhuiyan, A. Vuthichai, S. Muto, H. Ueno, "On tracking of eye for human-robot interface", International Journal of Robotices and Automation, 19(1):42-54 (2004).

[23] T. Sim, S. Baker, M. Bsat, "The CMU Pose Illumination and Expression (PIE) Database", Proceedings of the 5th International Conference on Automatic Face and Gesture Recognition, 2002.

[24] M.A. Bhuiyan, C.H. Liu, "Intelliget vision system for human-robot interface", Proceedings of World Academy of Science, Engineering and Technology, 28:57-63 (2007).

Empirical Analysis of Metrics Using UML Class Diagram

Bhawana Mathur
Deptt. of CS&E
JECRC University Jaipur

Manju Kaushik
Deptt. of CS&E
JECRC University Jaipur

***Abstract*—Lots of Organizations before they are setup survey the maintainability of programming frameworks. To give quality program design there exists a critical strategy called Object-Oriented Framework. Object–Oriented estimations might be utilized to study the judgment skills suite of a class diagram's structure particularly programming valuations and how the models have been developed and portrayed. The UML Class Diagram metrics maintain the Object-Oriented software. It is maintained through the investigation of the association among object oriented metrics and maintainability. This paper shows the effects of a scientific evaluation of software maintainability forecast and metrics. The research aims at the software quality attribute of maintainability as opposite to the method of software maintenance. It also aims to find out the vital correlation between structural complexity metrics and maintenance time. Several investigators have done copious in it, got lots of theoretical outcomes, and subsequently established a chain of practical uses. Due to dynamic changes in object-oriented technology, in today's scenario the class diagram is an essential UML model, as, researcher must first get to know the use of software in a scientific manner. It is an affordable strategy which has had an exceptional result in recent times. This paper is related to UML class diagram metrics through which a way is provided to maintain UML class diagram complexity weights. UML Class diagram's qualities will efficiently and technically show the complexity of Object -Oriented Software. A more specific research study has shown that the technique is associated with individual's experience and also can be useful to improve software quality.**

Keywords—UML Class diagram; Maintainability; Object Oriented System; CK Metrics suite; Model; Software; UML

I. Introduction

The aim of the research is to find out the significant correlation between structure complexity during maintenance time and the role played by metrics in class diagrams for maintainability. In the arena of software engineering, the aim is to lessen the price and energy in developing a software which has improved features but low maintainability. Because of this, software maintainability of a code is a difficult task which each user is facing. The class diagram plays a significant role in software maintenance. An upgraded version of a software will help in solving various defects and problems and the users will become accustomed to the newer environment. With reference to the ISO/IEC-9126 maintainability, the capacity of a software should be changed as per the requirement [1]. The changes will be remedial, adaptive or perfective in nature and the software systems will purposefully fulfill the requirements of a user. The present requirement of a user is to acquire better features with low maintainability which poses a big challenge for researchers. Maintenance usually absorbs 40% to 80% of a software system's value. Subsequently, software maintenance is essential; it is significant for the designers to think about the maintainability side throughout the process of code changing. Failing to satisfy users in this regard can lead to an increase in their maintainability value. That is why software system metrics have become vital in the field of software engineering practice. Researchers are more inclined towards making quantitative software traits that would evaluate the labor cost, cost efforts, energy, etc.

Regardless of the problems that may come on the way, evaluating maintainability of a code continues to be an awfully troublesome and difficult task. In any case, some of the developers or users are of the view that practical valuation may appear to be difficult in software development process. As a result of the many necessary conclusions along with investigation associated with software metrics as well as maintainability, it is vital to ascertain various associations among object-oriented metrics and their maintainability. Chidamber and Kemerer (CK) object-oriented metrics is incontestable with its durable influence on programming feature and attempts a supportive outfit of item arranged software metrics [2].

II. Objectives

A. To use various metrics in UML class diagram to find out maintainability.

B. To demonstrate a controlled analysis so as to assess if there is a clear imprint in maintaining the integrity of the specifications.

C. To use the experiential data as structure models which portrays UML class diagram maintainability

D. To take into account the particular models to anticipate UML class diagram maintainability in the OOIS advancement life-cycle.

III. Literature Review

Varied kinds of maintainability models have been prescribed. Prototypes have expected maintainability with code as well as plan estimations; however certain are concentrating just on layout level estimations [3]. Nowadays, the principle points for programming building are to expand the properties of programming ancient rarities. Nearby is a

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

conjoint contract that the quality assertion are crucially assured since the beginning periods of the product development life cycle concentrating on high-level design articles like class diagrams. In the improvement of Object Oriented Information System, the class diagram is a significant initial object that sets the basis of completely imminent model and application task. Later, class diagram attribute is a vital problem that is required to be assessed (better-quality if essential) in form to get attributes OOIS, that is the prime thing of current era programming augmentation organizations. The underlying accentuation on class diagram quality is permitted with the help of IS innovators for enhanced OOIS, denied of inessential modification at future periods of the improvement when changes are further sumptuous and further hard to accomplish. In this field where software analysis is a significant part since the initial obtainability of metrics provides a new assessment to class diagram attributes and unbiased approach staying away from inclination in the quality assessment process. Besides, measurements convey an esteemed and fair-minded vision into accurate procedures of enhancing the greater part of the product qualities. Shown that maintenance was (then will remain) the key supply customer of the entire programming life cycle, common sense has made one of the item vestige qualities highlights that item change affiliations further focus on pretty much. Maintainability is not limited to code; it is nature of the divergent programming. However, researchers are cognizant that maintainability is an outside trademark that must be considered in the OOIS circuition. In this way, it is essential to have starting indications of such standards based on the physical attributes of class outlines. Most advanced of the obtainable works on OO methods is associated with methods which can first be real-world once a software artifact is finished or approximately finished. The above- mentioned data also makes the framework in OOIS. Later full examination of certain diagrams like reachable OO quantum, fitting to class diagrams at strong configuration stage researchers have future set up of UML class diagram structural complexity that helped over through viable methodology about UML relations, such as associations, generalizations, aggregations and dependencies[4]. However, use of metrics is of no worth if their applied procedure is not validated analytically, whichever in tricks of case studies occupied against tangible task beyond meticulous research. Observed justification is vitally aimed at achievement of some software evaluation task. During this research, class diagram is additionally perceived in experiential exploration, certainly attained initial signs of class diagram maintainability are obtained [4]. The results have shown good initial conclusions for the OOIS researchers to proceed in the OOIS improvement circuition subsidizing to the ascent of well component OOIS. A preceding skillful research has been achieved. The autonomous variables are present in the UML class diagram with structure complexity metrics. Prior examinations show that there are indigent variables of practicality particular understandability and modifiability using client appraisals gathered on seven semantic marks scale. In object oriented software, to investigate measurement like structure complexity metrics in UML ,Class diagrams are used to collect observational data in features like union and coupling [4] [5].

IV. Basic Concept

There dependably has been an interest to give proficient, compelling and amazing programming. There are numerous maintainability items to convey enhanced maintainability. The component of respectable programming arranges seriously strong element of programming. Finally, some finishing up comments and future patterns in measurements for OO models are displayed [3]. Quite recently specific estimations have been described, which can be associated with class diagram level when all is said and done (see Table 4). Additionally customary eminent like, the quantity of classes, quantity of properties and quantity of techniques are considered.

V. Empirical Data Collection

Information sources are exhibited and given a point by point portrayal of information accumulation technique through inspection of 22 distinctive UML Class diagrams.

Following steps have been followed:

A. The self-representing variables are the vital whim and the extent of UML class diagrams, measured concluded the 11 estimations showed up in Table I.

B. Individually, choose an inside of a subject class diagram test, i.e., each of the tests (trial undertakings) must be understood by all of the subjects. The trials are to be found as an alternate request used for every theme.

C. Themes are assumed a careful instructional meeting before the analysis occurred. Be that as it may, the subjects didn't know about what highlights are imagined to ponder. Nor are they mindful of the genuine speculation expressed.

D. So as an association between the measurements presented in Table 1 and the class diagram maintainability is clear.

VI. Case Study

Twenty two different UML class diagrams have been examined and maintainability through these Size metrics and structure complexity metrics have been found. Size metrics contains Response for a class, NOA, Number of Method, Total Number of methods, Weight methods per class. Structure Complexity metrics contains Number of Children, Number of classes, Number of Relation, NGen, MaxDIT, NAggH, NGenH [6][7].

VII. Significance of Study

Particularly measurements ratio the structure unpredictability of UML Class diagrams as a result of the use of associations, for instance, affiliations, speculations, conglomerations and conditions [3]. In like manner standard measurements, for instance, the number of classes and the amount of characteristics et cetera are considered (See Table 1).

VIII. These Class Diagram Structural Complexity Measures Permit OO Creators

A. *Numerical examination of outline choices, thus a target choice among various class diagram changes with proportional semantic substance.*

B. *The forecast of outer quality attributes, similar to maintainability in the starting periods of the IS period of existence is also an asset to assignment taking into account these expectations.*

IX. Observations

A. *Analyze:* UML class diagram structural complexity metrics

B. *Determination of Assessing*: Concerning their capacity of imperativeness castoff as class diagram maintainability markers.

C. *As of the perspective of Information system draftsmen:* Under the guidance of Research Scholar Computer Science and Associate Professor of the Computer Science in the JECRC University.

D. *Arranging. − Background choice.* The examination is not permanent and this one highlights on UML class diagram structural complexity metrics.The maintainability is improved since the particular setting is also progressed underneath one researcher [8] [9]. The Investigation reports are the genuine issue, particularly thae pointers can be utilized to measure the maintainability of class diagram. In the end, it inspects the connection between metrics and maintainability.

The components remain in UML class diagrams. The independent variable is controlled by the valuations. The destitute variable stays restricted all the times.

E. *Variables determination.* The free fickle in the UML class diagram is essentially unpredictable. The maintainability is called penniless fickle in UML class diagram.

F. *A researcher called time as "maintenance time."* Maintenance time incorporates an immaculate opportunity to get a hold on the class diagram so as to examine the compulsory variations and to execute them. The doubt here is that, for the similar adjustment assignment, the faster a class diagram can be changed, the less troublesome it will be to keep up.

G. *Hypothesis formulation.* Researcher desires to trial the subsequent hypotheses:

H. *The Null hypothesis, H0:* There is no significant correlation between structural complexity metrics like RFC, NOA, NOM, WMC, NOC, NC, NOR, NGen, MaxDIT, Nagg, NAggH, NGenH and maintenance time.

I. *An Alternative hypothesis, H1:* There is a significant correlation among structural complexity metrics (RFC, NOA, NOM, WMC, NOC, NC, NOR, NGen, MaxDIT, Nagg, NAggH, NGenH) and maintenance time.

TABLE I. Measurements for UML Class Diagram Basic Intricacy

Metric Name	**Metric definition**
Response for a Class (RFC)	Response for a Class(RFC)
Number of Attributes (NOA)	The Number of Attributes (NOA)
Total Number of Methods (NOM)	The Total Number of Methods (NOM)
Weight Method per Class (WMC)	The Weight Method per Class (WMC)
Number of Children (NOC)	The Number of Children (NOC)
Number of Classes (NC)	The Number of Classes (NC)
Number of Relation(NOR)	The Number of Relation(NOR)
Number of Generalizations (NGen)	The total number of Generalization relationships within a class diagram (each parent-child pair in a generalization relationship).
Maximum DIT(MaxDIT)	It is the maximum of the DIT (Depth of Inheritance Tree) values obtained for each class of the class diagram. The DIT value for a class within a Generalization Hierarchy is the longest path from the class to the root of the hierarchy.
A Number of Aggregations Hierarchies (NAggH)	The total number of Aggregation hierarchies (whole-part structures) within a class diagram.
A Number of Generalisations Hierarchies (NGenH)	The total number of Generalization hierarchies within a class diagram.

J. *Experiment design.* An inside of subject configuration test, i.e., completely the tests (trial undertakings) must be determined by all of the themes. The point indicates the trials in a various request.

X. Operation

Preparation. : While the examinations stayed finished, the subjects had involved two methods for Software Engineering. Twenty-two UML class outlines of exceptional use that were absolutely not hard to be appreciated by each of the topics. The charts have various basic multifaceted nature, all things considered of metric qualities. Each and every graph had an encased scan that integrated a quick depiction of what the outline meant and two new presents for the category chart. Every last zone anticipated that it would adjust the class diagrams permitting the novel necessities and demonstrate the begining and end time. The separation among the two is anything we call time (passed on in minutes and seconds). The alterations based on every classification graph have been simila in qualities, procedures, classes, et cetera.

Execution: The topics had been common and revealed figure out by what method to do the checks. The major part of the learning with the changed class diagrams is composed with the maintenance time acquired after the reactions of the checks and the estimations prices without a doubt figured by the process for a metric instrument orchestrated.

Information Validation: When the data was collected, it was assessed if the checks have been done and if the movements had been finished precisely.The expert used the information gathered as a touch of asking for to test the hypotheses definite. (see Table II) [4] [7].

Analyzing the Spearman´s correlation coefficients: There is a high association (rejecting hypothesis H0) among most of the UML class diagram's metrics and maintenance time is surmised. The way that every one of the metrics has an association more prominent than 0.5 is expected, which is a typical edge to survey correlation values. NOR is the emerge that has a lesser relationship, yet this could be cleared up by the route that in most of the class diagrams NOR appropriated the value 0 (see Table II).

TABLE II. MAINTENANCE TIME BETWEEN METRICS

	RFC	**NOA**	**NOM**	**WMC**	**NOC**	**NC**	**NOR**	**NGen**	**MaxDIT**	**NAggH**	**NGenH**
Maintenance	0.56	0.75	0.83	0.56	0.54	0.89	**0.41**	0.58	0.72	0.68	0.69

XI. VALIDITY EVALUATION

The few problems that undermine the validity of the experiential gain of knowledge are pondered [4] :

A. *Threats to inference validity*

The inference legitimacy describes the level to which inferences are quantifiably immense. The fundamental constraint that would affect the quantifiable legitimacy of this learn is the degree of the case know-how (242 qualities *and 22 subjects), that potentially are not suitable for each parametric and non-parametric size output [4] [10]*.

B. *Threats to develop Validity*

The dependent variable castoff as maintenance time, i.e., the time every subject spent implementation.

C. *Threats to inside Validity*

The inward legitimacy portrays the level of trust in an intention impact relationship amongst additional items of leisure activity and the experiential result.

The running with issues have been distributed with contrasts between subjects [9]:

In this examination, in building UML class diagrams. This is shown through the examination of the descriptive statistics in light of the aggregate maintenance time for every subject (see Table III). As the Kurtosis qualities are more prominent than zero and conclude that there are no great contrasts between in the investigation.

TABLE III. DESCRIPTIVE STATISTICS FOR THE TOTAL MAINTENANCE TIME

	Min.	**Max**	**Mean**	**Std. Deviation**	**Skewness**	**Kurtosis**
Total maintenance time (minutes)	35	45	40.9	2.9	1.3	2.1

1) *Mastery of the creation of exchange amongst type diagrams [3] .*
2) *Accuracy in the time standards*
3) *Knowledge impacts*
4) *Exhaustion impacts*
5) *Determination impacts*
6) *Subject inspiration*

B. *Threats to outer Validity*

Outer legitimacy is that UML diagrams used as an arrangement object for making OOIS and to various examination settings[1].

XII. RESULT

From the descriptive measurements, we noticed some observations and so actions were involved. Some of them have been said as takes after [13][14]:

A. *As NAggH Median and Mean value are minimum in 22 UML Class diagrams, so we conclude that the use of Aggregation in 22 UML Class diagrams is limited.*

B. *We have removed NOM, WMC, MaxDIT and NGenH from 22 UML Class diagrams because it has been observed that all data points for NOM, WMC, Maximum DIT and NGenH are zeros in the 22 UML Class diagrams.*

C. *We have observed that the classes, attributes between classes in 22 UML Class diagrams were high in RFC, NOM, and WMC 22 UML Class diagrams.*

D. *Values of Mean and Median of NC are almost same in 22 UML Class diagrams that mean UML Class diagrams have almost similar classes.*

E. *If WMC quality is high, it upsets the re-utilization of the class and to stay away from this we have to decrease the quantity of strategies or their complexity [13].*

F. *If the DIT is more than six, it expands the design complexity and subsequently decreasing the inheritance utilization while coding.*

TABLE IV. UML CLASS DIAGRAM WITH SIZE METRICS AND STRUCTURE COMPLEXITY METRICS

	Size Metrics				Structure Complexity Metrics						
Name of Software	**RFC**	**NOA**	**NOM**	**WMC**	**NOC**	**NC**	**NOR**	**NGen**	**MaxDIT**	**NAggH**	**NGenH**
U1	7	6	7	7	3	8	9	5	4	0	2
U2	22	7	22	22	2	5	4	4	2	1	2
U3	8	13	8	8	3	4	4	3	0	0	2
U4	9	6	9	9	1	3	2	3	0	0	1
U5	8	11	8	11	3	4	3	3	2	2	5
U6	5	13	5	5	2	3	8	8	2	6	2
U7	14	23	18	14	3	7	7	6	0	0	3
U8	30	23	30	30	5	6	8	5	0	1	3
U9	18	7	18	18	3	4	3	3	0	1	1
U10	6	4	6	2	3	4	5	4	2	0	2
U11	9	12	6	9	7	8	8	3	0	0	2
U12	16	2	16	16	6	7	7	7	5	0	4
U13	7	2	2	7	2	3	2	1	2	0	0
U14	3	3	3	5	3	4	4	2	4	0	4
U15	5	11	5	6	3	4	4	2	3	0	2
U16	10	6	10	10	4	5	5	3	4	3	0
U17	30	23	30	30	5	6	7	2	6	1	0
U18	5	13	5	5	6	7	7	2	2	1	3
U19	11	11	0	0	4	5	4	2	2	1	1
U20	8	8	8	8	2	3	2	2	2	1	0
U21	12	9	12	12	4	5	7	3	4	3	0
U22	16	27	17	17	8	7	9	2	2	0	4

TABLE V. DESCRIPTIVE STATISTICS

	RFC	**N0A**	**NOM**	**WMC**	**NOC**	**NC**	**NOR**	**NGen**	**MaxDIT**	**NAggH**	**NGenH**
Min	3	2	**0**	**0**	1	3	2	1	**0**	**0**	**0**
Max	**30**	27	**30**	**30**	8	8	9	8	6	6	5
Mean	**12.2**	11.2	11.5	11.7	3.8	**5.1**	5.4	3.5	2.3	**1.1**	2.0
Median	9	**10**	8	9	3	**5**	5	3	2	**0.5**	2
Std. Deviation	7.6	7.2	**8.4**	8.0	1.8	1.7	2.3	1.8	1.8	1.5	1.5

G. *If the NOC is high, weakening of abstraction of classes happens, and subsequently, it diminishes NOC as if an occurrence of DIT is applied here.*

H. *If the RFC of a class is high, it again achieves more measures of testing attempts that are an augmentation in the test gathering and over design complexity of the framework increments. Consequently, the number of operations that will be implemented because of a message received must be lessened.*

I. *An exceptionally compelling part is played by the quantities of subclasses in influencing the maintainability of software.*

J. *UML demonstrating is obvious to be high priced and not as a subject of course financially savvy.*

XIII. CONCLUSION

It is extensively acknowledged that the element of OO Software must be assessed from the underlying levels of its development life cycle. This statistic clue to describe a set of metrics for evaluating the structural complexity of UML class diagrams, through the knowledge that they are related through the maintainability of diagrams. The results obtained through research shows that most of the metrics like NAggH suggested. It is a good indicator of class diagram maintainability and their attributes because NAggH median and mean value are minimum in 22 UML Class diagram and the use of Aggregation is limited in these class diagram. NOM, WMC, MaxDIT, and NGenH have been removed from these 22 UML class diagrams because it has been observed that all data points are zero. Investigational outcomes design that the metric is to a great degree identified with human knowledge and can effectively quantify the complexity of object-oriented software measurement. On the other hand, the outcomes of this research can help the software quality mechanism and assessment demonstrating based on software measurement, increase the accuracy and can be useful to increase software quality, software maintenance work and it can enhance and improve software measurement research. Similarly, WMC quality is high so, the quantity of their complexity has been decreased. In the same way, if DIT is more than six, therefore decreasing the inheritance utilization while coding has been removed. The software maintainability is a noteworthy natural for an application model which is planned to lessen a framework's inclination, and to indicate when it creates much less high-priced and not more hazardous to change the code alternatively to alter it. The high-quality viable knowledge of maintainability allows for the aid crew to know, on what module to middle amid preservation.

XIV. FUTURE WORK

Even though the results acquired in this experiment are hopeful. It's major to place on these measures to data got from "actual tasks".

Various modifications that could be prepared to develop the research obtainable are:

A. *Expansion the size of the class diagrams*. By increasing the size of the class diagrams. Also, as the cases will be more genuine and if work with professionals, could make improved use of their possible ability and accomplish that the outcomes will more common.

B. *Increase the change between the values of the metrics.* This option could lead to more definite outcomes about the metrics and their relationship with the factor will try to control.

C. *Carry out the research in a further skillful environment.*

D. *Research with real data*. One more way to improve the legitimacy of the outcomes is by functioning with real data acquired from engineering environment .But, the lack of such data continuous to be difficult so must invention other ways to challenge validating metrics.

ACKNOWLEDGMENT

The JECRC University, Jaipur, Rajasthan, India supports this research.

REFERENCES

[1] Genero B., M., Moody D.L. and Piattini M. (2005), 'Assessing the capability of internal metrics as early indicators of maintenance effort through experimentation',*Journal of software maintenance and evolution: Research and Practice*, 17(3), pp.225-246.

[2] Dubey S.K. and Rana A. (2011), 'Assessment of maintainability metrics for object-oriented software system', *ACM SIGSOFT Software Engineering Notes,* 36(5), pp.1-7.

[3] Genero M., Piattini M. and Jiménez L. (2001), 'Empirical validation of class diagram complexity metrics', *In Computer Science Society, 2001. SCCC'01', XXI International Conference of the Chilean*, Chile, 9-9 Nov.2001. IEEE ,pp. 95-104.

[4] Genero M., Olivas J., Piattini M. and Romero F. (2002), 'Assessing object-oriented conceptual models maintainability', *In Advanced Conceptual Modeling Techniques*, Volume 2784 of the series Lecture Notes in Computer Science, Berlin, Germany,October 7,2002. Springer Berlin Heidelberg, pp. 288-299.

[5] Dubey S.K., Sharma A. and Rana A. (2011), 'Analysis of Maintainability Models for object-oriented System', *International Journal of Computer Science and Engineering,* 3(12), pp.3837-3844.

[6] Kuhl, Frederick S.(1990), 'Object-oriented programming applied to a prototype workstation', *Software: Practice and Experience,* 20(9), pp. 887-898.

[7] Genero M., Piattini M. and Calero C. (2002), ' Empirical validation of class diagram metrics', *International Symposium on Empirical Software Engineering*,Nara ,Japan, October 3-4,2002. IEEE ,pp. 195-203.

[8] Genero M., Olivas J., Piattini M. and Romero F. (2001), 'Using metrics to predict OO information systems maintainability', *In Advanced Information Systems Engineering* , Volume 2068 of the series Lecture Notes in Computer Science, Sweden,June 28,2001. Springer Berlin Heidelberg ,pp. 388-401.

[9] Genero M., Jiménez L. and Piattini M. (2002), 'A controlled experiment for validating class diagram structural complexity metrics', *In Object-Oriented Information Systems,* Montpellier, France,September 2,2002.Springer Berlin Heidelberg ,pp.372-383.

[10] Genero M., Poels G., Manso E. and Piattini M. (2005), 'Defining and validating metrics for UML class diagrams', *Metrics for Software Conceptual Models*, London, UK, January 4, 2005. Imperial College Press,pp.99-159.

[11] Genero M., Piattini M., Manso E. and Cantone G.(2003), 'Building UML class diagram maintainability prediction models based on early metrics', *Ninth International Symposium in Software Metrics,* Sydney ,NSW,Australia ,September 3-5,2003.IEEE ,pp. 263-275.

[12] Fernández-Sáez A.M., Genero M. and Chaudron M.R.(2011), 'Does the level of detail of UML models affect the maintainability of source code?', *Models in Software Engineering* , Wellington, New Zealand, October 16-21,2011. Springer Berlin Heidelberg, pp. 134-148.

[13] Al Dallal J. (2013), 'Object-oriented class maintainability prediction using internal quality attributes', *Information and Software Technology,* 55(11), pp.2028-2048.

[14] Genero M., Poels G. and Piattini M.,(2008), 'Defining and validating metrics for assessing the understandability of entity–relationship diagrams', *Data & Knowledge Engineering*, 64(3), pp.534-557.

A Reduced Switch Voltage Stress Class E Power Amplifier Using Harmonic Control Network

Ali Reza Zirak
Laser & Optics Research School,
NSTRI,
Tehran, Iran

Sobhan Roshani
Department of Electrical Engineering, Kermanshah Branch,
Islamic Azad University,
Kermanshah, Iran

***Abstract*—In this paper, a harmonic control network (HCN) is presented to reduce the voltage stress (maximum MOSFET voltage) of the class E power amplifier (PA). Effects of the HCN on the amplifier specifications are investigated. The results show that the proposed HCN affects several specifications of the amplifier, such as drain voltage, switch current, output power capability (C_p factor), and drain impedance. The output power capability of the presented amplifier is also improved, compared with the conventional class E structure. High-voltage stress limits the design specifications of the desired amplifier. Therefore, several limitations can be removed with the reduced switch voltage. According to the results, the maximum drain voltage for the presented amplifier is reduced and subsequently, the output power capability is increased about 25% using the presented structure. Zero-voltage switching condition (ZVS) and zero-voltage derivative switching condition (ZVDS) are assumed in the design procedure. These two conditions are essential for high efficiency achievement in various classes of switching amplifiers. A class E PA with operating frequency of 1 MHz is designed and simulated using advanced design system (ADS) and PSpice software. The theory and simulated results are in good agreement.**

Keywords—class E power amplifier; harmonic control network (HCN); MOSFET drain Impedance; ZVS and ZVDS conditions

I. Introduction

The main blocks in many communication systems are power amplifiers. Power amplifiers are high-consumption elements, and so, their efficiency is an important factor in the design procedure [1-3]. Class E power amplifiers are very interesting for communication systems due to their high efficiency. Class E amplifiers are also called DC-AC inverters, which are further divided into two main types: the ZVS power amplifiers and the zero current switching (ZCS) power amplifiers. Both ZVS/ZVDS conditions must be considered in the nominal operation of ZVS type amplifiers, while in sub-nominal operation only ZVS condition is considered. The ZVS type class E PA with nominal operation is studied in this paper. Several researches have been mentioned ZVS and ZVDS conditions as essential conditions to obtain high efficiency in high frequencies [4-6].

Recently, extra parameters have been assumed in some approaches to add one degree of freedom to class E PA design. For example, the shunt intrinsic capacitance (C_{ds}) of the MOSFET is considered nonlinear, and the grading coefficient is taken into account as a new parameter in class E amplifier design [7-8]. Besides, input voltage duty cycle is considered to analyse the class E PA in [9-11]. Varying input voltage duty cycle leads to change in the class E PA specifications. Also, recent approaches have introduced several structures with improved MOSFET voltage and current waveforms, such as inverse class E [12-14], class EF [15-17], and class DE power amplifiers [18-19]. Class DE amplifiers are introduced by adding two parallel capacitances to the switching MOSFETs. Class DE amplifiers have high efficiency, but their output power capability (C_p) is low. Inverse structure of class E is similar to typical structure, but has a dual circuit. ZCS and ZCDS (zero current derivative switching) conditions could be satisfied in Inverse type of class E PA. Combining the structures of class E and F PAs leads to class EF family amplifiers. Reduced switch (MOSFET) voltage or current could be achieved in this kind of amplifiers, according to the harmonics effects on the MOSFET voltage and current, while their efficiency is low, compared with the other switching classes. The ZVS and ZVDS conditions could also be satisfied in the class EF family amplifiers.

In comparison with the mentioned structures, the class E PA has high drain efficiency; however, its C_p factor is low. Therefore, increasing the C_p factor and decreasing the voltage stress in class E amplifier, are still subject to interest. In this paper, a harmonic control network (HCN) is designed and applied in the typical class E amplifier circuit to reduce the maximum value of the MOSFET voltage and also to improve the C_p factor of the amplifier.

The paper sections are summarized as follows. Sections II and III present the structure and specifications of the typical class E PA. Description of the presented HCN and C_p factor of the presented PA are described in Sections IV and V, respectively. Obtained results of the design examples are given in Section VI. Finally, the conclusions of the obtained results are presented in Section VII.

II. Typical Class E Pa Structure

The typical class E circuit is depicted Figure 1, which includes a transistor (MOSFET) as a switch, RF choke, series resonator, dc supply voltage (V_{dc}), and a parallel capacitance. The parallel capacitance (C_{sh}) in class E PA includes the MOSFET intrinsic capacitance (C_{ds}) and a shunt external capacitance. However, the external capacitance could be neglected to reach higher operating frequency.

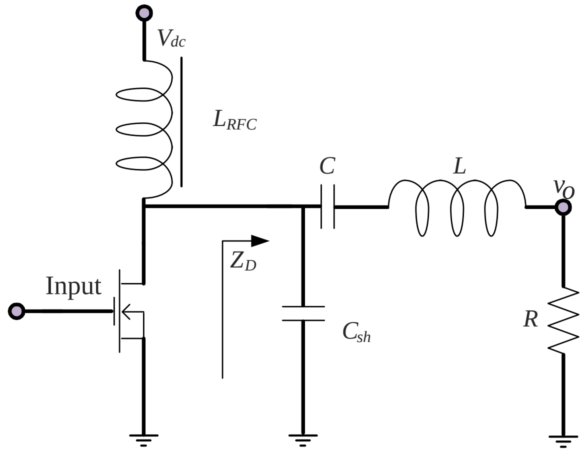


Fig. 1. Typical class E circuit

The normalized MOSFET voltage (v_s/V_{dc}) and normalized MOSFET current (i_s/I_{dc}) waveforms of a typical class E structure are illustrated in Fig. 2. According to this figure, the maximum values of v_s/V_{dc} and i_s/I_{dc} for typical class E amplifier are 3.56 and 2.86, respectively. According to breakdown voltage constraints of the MOSFET, V_{dc} and subsequently the output power could not exceed from a specific value and this problem limits the design procedure. With reduced maximum switch voltage, more output power could be achieved, and the design process of the presented structure would be more flexible.

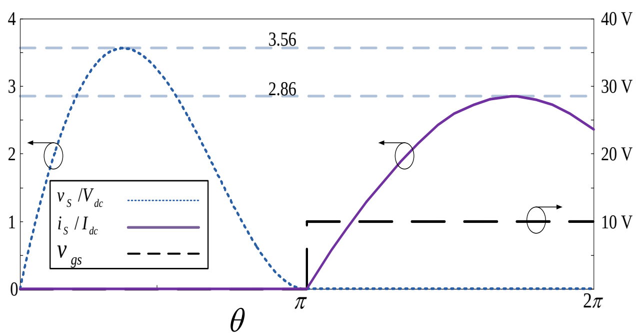


Fig. 2. v_s/V_{dc} and i_s/I_{dc} waveforms of the typical class E structure

III. Drain Impedance

In class E structure, the MOSFET plays the switch role in the circuit and threshold voltage (V_{th}) acts as a trigger for this switch. It means when the input voltage value is higher than V_{th}, the switch (MOSFET) is on, and the value of v_s is equal to zero. On the contrary, when the input voltage value is lower than V_{th}, the switch (MOSFET) is off, and the switch voltage can be calculated, according to the amplifier circuit. The obtained switch voltage waveform of class E amplifier depends on the drain impedance (Z_D). According to Fig. 1, the value of Z_D, can be calculated as follows

$$Z_D = \frac{LCS^2 + RCS + 1}{LCC_{sh}S^3 + RCC_{sh}S^2 + \left(C + C_{sh}\right)S}. \quad (1)$$

As can be seen in equation (1), the value of Z_D is a function of L, C, C_{sh}, R, loaded quality factor (Q) and frequency. From equation (1), the value of Z_D for the typical class E PA output structure could be achieved, which is depicted in Fig. 3. The drain impedance of the typical structure has a zero and a pole near the operating frequency.

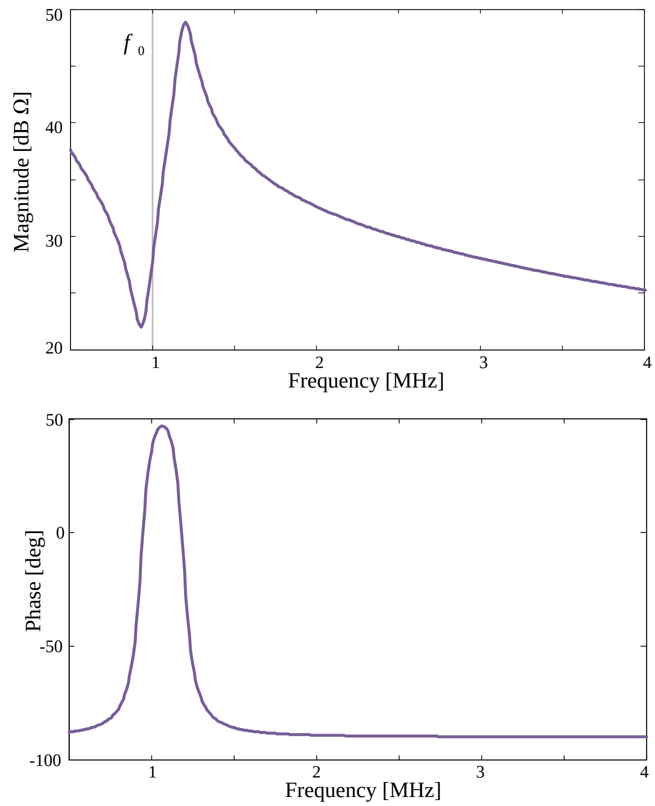


Fig. 3. The value of Z_D for the typical class E output structure

In the circuit structure, high value of Q for the resonant circuit provides a pure sinusoidal voltage at v_o. The value of loaded quality factor could be written as

$$Q = \frac{\omega L}{R}, \quad (2)$$

where ω is angular frequency. Effect of different values of Q on drain impedance is shown in Fig. 4. According to this figure, as the loaded quality factor increases, the zero and pole of the impedance become closer to operating frequency, and the ideal resonator will be achieved with a high value of Q.

IV. Harmonic Control Network

As mentioned, the class E PA drain voltage depends on the drain impedance. It means the switch voltage can be reduced using harmonic control elements. A harmonic control network is designed and inserted in typical class E circuit to decrease the MOSFET voltage. The proposed HCN structure includes two resonator branches, as illustrated in Fig. 5. The circuit values of the designed circuit are tabulated in Table 1.

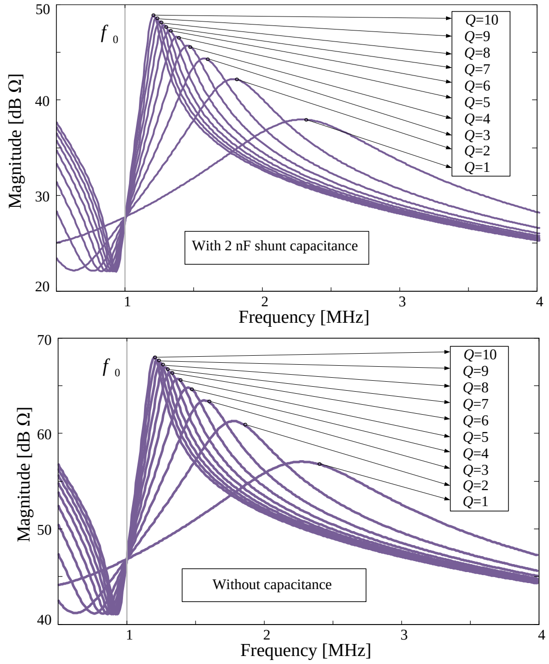


Fig. 4. Effect of Q on the typical class E PA drain impedance with and without shunt capacitance

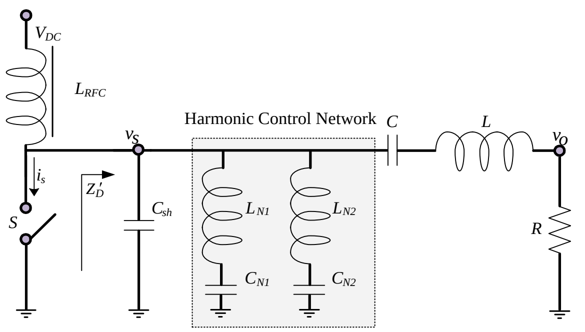


Fig. 5. Proposed harmonic control network in the amplifier circuit

The impedance of the HCN could be written as

$$Z_{HCN}=\frac{L_{N1}L_{N2}C_{N1}C_{N2}S^{4}+\left(L_{N1}C_{N1}+L_{N2}C_{N2}\right)S^{2}+1}{C_{N1}C_{N2}\left(L_{N1}+L_{N2}\right)S^{3}+\left(C_{N1}+C_{N2}\right)S}. \tag{3}$$

The impedance of the HCN is illustrated in Fig. 6. There is a zero in the impedance of the presented HCN in the third harmonic of the operating frequency. According to Fig. 5, the drain impedance of the proposed amplifier is

$$Z'_{D}=Z_{HCN}\frac{LCS^{2}+RCS+1}{F(S)} \tag{4}$$

where $F(S)$ is defined as

$$\begin{aligned}F(S)=Z_{HCN}LCC_{sh}S^{3}+\left(LC+Z_{HCN}RCC_{sh}\right)S^{2}\\+\left(Z_{HCN}\left(C+C_{sh}\right)+RC\right)S+1\end{aligned} \tag{5}$$

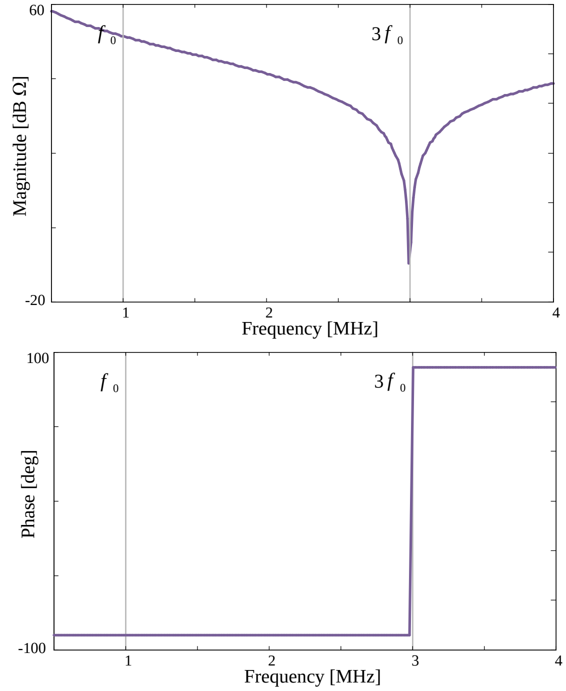


Fig. 6. Impedance of the presented HCN

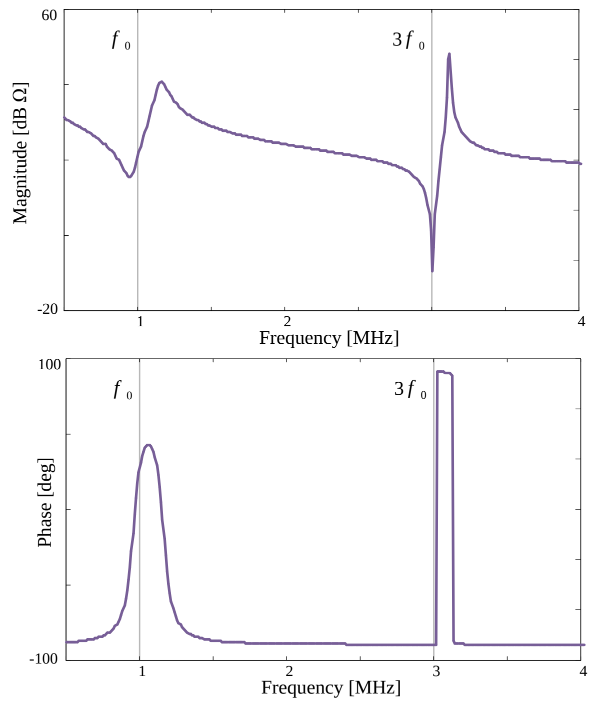


Fig. 7. Drain impedance of the presented amplifier using the proposed HCN

The drain impedance of the presented amplifier using the proposed HCN is shown in Fig. 7.

V. OUTPUT POWER CAPABILITY

To compare the output power of different amplifiers, the output power capability (c_P) factor is used. The c_P factor could be calculated as follows [20]

$$c_P = \frac{P_{o,\max}}{v_{s,\max} i_{s,\max}} = \frac{1}{\dfrac{v_{s,\max}}{V_{DC}} \dfrac{i_{s,\max}}{I_{DC}}}. \qquad (6)$$

As can be concluded from (6), the value of c_P is proportional to inverse values of $v_{s,max}/V_{dc}$ and $i_{s,max}/I_{dc}$. The value of c_P can be calculated for the typical class E PA, according to values of $v_{s,max}/V_{dc}$ and $i_{s,max}/I_{dc}$, which is 0.098.

VI. SIMULATION RESULTS AND DESIGN EXAMPLES

The presented amplifier is simulated at 1 MHz, using ADS and PSpice software. An IRF530 MOSFET transistor is applied as a switch in the designed class E circuit. The first design example is simulated using ADS software. A switch model with relevant parasitic elements is used in the ADS simulation to model the switching device. The MOSFET on resistance ($R_{DS(on)}$ = 0.16 Ω) is considered, according to the IRF530 MOSFET datasheet. To validate the ADS simulation results, the second design example is simulated using PSpice software. The circuit values and design parameters of two design examples are the same, except the load resistance. The small difference in the load resistance is due to the MOSFET parasitic elements, which were considered in the PSpice simulation. Parasitic resistances and capacitances of the MOSFET are considered in the PSpice model. The IRF530 level 3 PSpice model is used in this paper. Simulated results of the first and second design examples are shown in Fig. 8 and Table 2. According to the PSpice simulation results, the normalized maximum MOSFET voltage and MOSFET current of the presented PA have been achieved 2.76 and 2.95, respectively. Subsequently, from equation (6), the c_P factor for the presented PA can be obtained as 0.122. The c_P factors of different amplifiers are compared in Table 3. According to the results, the presented HCN can improve both the c_P factor and the MOSFET voltage. More parameters of class E amplifier could be improved using a modified HCN, which we will address in future work. As mentioned, the drain impedance shapes the MOSFET voltage and currents of the amplifier. Therefore, several limitations of class E PA could be removed using a modified HCN.

TABLE I. CIRCUIT VALUES OF DESIGNED AMPLIFIER

V_{DC}	12 V
f_0	1 MHz
L	11 µH
C	2.53 nF
C_{sh}	5 nF
L_{N1}	7.8 µH
C_{N1}	360 pF
L_{N2}	10 µH
C_{N2}	10 pF

TABLE II. DESIGN EXAMPLES RESULTS

	First design example	Second design example
V_{in}	10 V	10 V
R	5.8 Ω	6.2 Ω
$v_{s,max}$	33.12 V	33.4 V
V_o	11.8 V	12.3 V
P_{out}	12 W	12.2 W
η	92.6 %	93.3 %

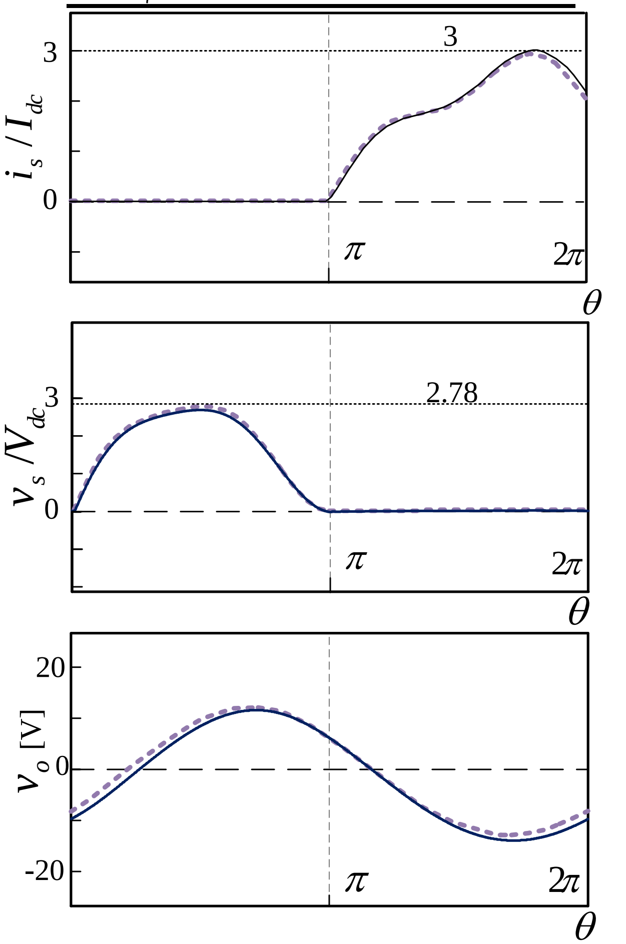


Fig. 8. Simulated waveforms of the presented amplifier using ADS (dashed line) and PSpice (solid line)

TABLE III. OUTPUT POWER CAPABILITIES (C_P) OF DIFFERENT AMPLIFIERS

Amplifier Type	c_P
Class B	0.125
Class D	0.159
Class E	0.098
Class DE	0.079
This work	**0.122**

VII. CONCLUSION

A new harmonic control network (HCN) is inserted in typical class E power amplifier structure. According to the results, several parameters of class E PA are improved, using the applied HCN. The voltage stress of the presented amplifier is reduced, which relaxes the design limitations of the presented PA. Besides, the output power capability is increased, when compared with the other amplifiers. According to the results, the proposed amplifier structure presents both high-efficiency advantage of class E PA and high output power capability advantage of class D PA.

REFERENCES

[1] A. Banerjee, L. Ding, and R. Hezar, "High efficiency multi-mode outphasing RF power amplifier in 45nm CMOS" In European Solid-

State Circuits Conference (ESSCIRC), ESSCIRC 2015-41, pp. 168-171, 2015.

[2] M. Hayati, S. Roshani, M. K. Kazimierczuk, and H. Sekiya, "A class E power amplifier design considering MOSFET nonlinear drain-to-source and nonlinear gate-to-drain capacitances at any grading coefficient," IEEE Transactions on Power Electronics, 2016, to be published, DOI: 10.1109/TPEL.2015.2512928.

[3] M. Hayati and S. Roshani, "A Class E Power Amplifier with Low Voltage Stress," Amirkabir International Journal of Electrical & Electronics Engineering, 47(1), pp. 31-37, 2015.

[4] M. Hayati, S. Roshani, M. K. Kazimierczuk, and H. Sekiya, "Analysis and Design of Class E Power Amplifier Considering MOSFET Parasitic Input and Output Capacitances," IET Circuits, Devices & Systems, 2016, to be published, DOI: 10.1049/iet-cds.2015.0271.

[5] K. Pengand and E. Santi, "Class E resonant inverter optimized design for high frequency (MHz) operation using eGaN HEMTs," In Applied Power Electronics Conference and Exposition (APEC), pp. 2469-2473, 2015

[6] S.C. Wong and C.K. Tse, "Design of symmetrical class E power amplifiers for very low harmonic-content applications," IEEE Transactions on Circuits and Systems I: Regular Papers, 52(8), pp.1684-1690, 2005.

[7] C. Chan and C. Toumazou, "Physically based design of a class e power amplifier with non-linear transistor output capacitance," In COLLOQUIUM DIGEST-IEE, pp. 5-5 1999.

[8] N. Ha-Van, N. Dang-Duy, H. Kim, and C. Seo, "Frequency limitation of an optimum performance class-E power amplifier," IEICE Electronics Express, 2016, to be published, DOI: 10.1587/elex.13.20160108.

[9] Y. Yusop, M.S.M. Saat, S.H. Husin, S.K. Nguang, and I. Hindustan, "Design and Analysis of 1MHz Class-E Power Amplifier for Load and Duty Cycle Variations," International Journal of Power Electronics and Drive Systems (IJPEDS), 7(2), 2016.

[10] A. Mediano, P. Molina-Gaudo, and C. Bernal, "Design of class E amplifier with nonlinear and linear shunt capacitances for any duty cycle," IEEE Transactions on Microwave Theory and Techniques, , 55(3), pp. 484-492, 2007.

[11] X. Du, J. Nan, W. Chen, and Z. Shao, "New solutions of Class-E power amplifier with finite dc feed inductor at any duty ratio," IET Circuits, Devices & Systems, 8(4), pp. 311-321, 2014.

[12] T. Mury and V. F. Fusco, "Sensitivity characteristics of inverse Class-E power amplifier," IEEE Transactions on Circuits and Systems I: Regular Papers, 54(4), pp. 768-778, 2007.

[13] S.H. Kam, O.S. Kwon, M.W. Lee, and Y.H. Jeong, "High-efficiency inverse class-E power amplifier with envelope tracking technique," Microwave and Optical Technology Letters, 55(4), pp.866-869, 2013.

[14] T. Cao, Y.J. Liu, R. Zeng, and L.M. LV, "S-band high efficiency GaN inverse-class E power amplifier," J Microw, 27(4), pp.49-52, 2011.

[15] G. Formicone and J. Custer, "Mixed-mode class EF− 1 high efficiency GaN power amplifier for P-band space applications" In Microwave Symposium (IMS), IEEE MTT-S International, pp. 1-4, 2015.

[16] S. Aldhaher, D. C. Yates, and P. D. Mitcheson, "Modeling and Analysis of Class EF and Class E/F Inverters With Series-Tuned Resonant Networks," IEEE Transactions on Power Electronics, 31(5), pp. 3415-3430, 2016.

[17] M. Thian, A. Barakat, and V. Fusco, "High-efficiency harmonic-peaking Class-EF power amplifiers with enhanced maximum operating frequency," IEEE Transactions on Microwave Theory and Techniques, 63(2), pp. 659-671, 2015.

[18] T. Kondo, and H. Koizumi, "Class DE voltage-source parallel resonant inverter," In Industrial Electronics Society, IECON 2015-41st Annual Conference of the IEEE pp. 002968-002973, 2015.

[19] H. Sekiya, X. Wei, T. Nagashima, and M. K. Kazimierczuk, "Steady-State Analysis and Design of Class-DE Inverter at Any Duty Ratio," IEEE Transactions on Power Electronics, 30(7), pp. 3685-3694, 2015.

[20] M. J. Chudobiak, "The use of parasitic nonlinear capacitors in class-E amplifiers," IEEE Transactions on Circuits and Systems I, 41(12), pp. 941–944, 1994.

Quizzes: Quiz Application Development Using Android-Based MIT APP Inventor Platform

Muhammad Zubair Asghar[1]
Institute of Computing and Information Technology
Gomal University, D.I.Khan, Pakistan

Khushboo Nasir[3]
Institute of Computing and Information Technology
Gomal University, D.I.Khan, Pakistan

Fazal Masud Kundi[5]
Institute of Computing and Information Technology
Gomal University, D.I.Khan, Pakistan

Iqra Sana[2]
Institute of Computing and Information Technology
Gomal University, D.I.Khan, Pakistan

Hina Iqbal[4]
Institute of Computing and Information Technology
Gomal University, D.I.Khan, Pakistan

Sadia Ismail[6]
Institute of Computing and Information Technology
Gomal University, D.I.Khan, Pakistan

***Abstract*—This work deals with the development of Android-based multiple-choice question examination system, namely: Quizzes. This application is developed for educational purposes, allowing the users to prepare the multiple choice questions for different examinations conducted on provincial and national level. The main goal of the application is to enable users to practice for subjective tests conducted for admissions and recruitment, with the focus on Computer Science field. This quiz application includes three main modules, namely (i) computer science, (ii) verbal, and (iii) analytical. The computer science and verbal modules contain various sub-categories. This quiz includes three functions: (i) Hint, (ii) Skip, and (iii) Pause/*life-lines*. These functions can be used only once by a user. It shows progress feedback during quiz play, and at the end, the app also shows the result.**

Keywords—Quiz; Android; MIT App Inventor; Interviews and test preparation

I. Introduction

Development of Android-based Quiz application is mainly required by students and learners to prepare themselves for different examinations directly through Smart-Phones and tablets in hands. The main aim of this project is to facilitate students in learning, gaining and improving their knowledge skills. At the meantime, our app provides them fun so that the users can prepare for interviews, entrance tests or any other corresponding purposes in a fresh mood and can't get bored or frustrated due to the dullness of application. We designed the application to facilitate the users to be able to take short quizzes using portable devices such as smart phones and tablets.

Byers and Alnarp [1] proposed an Interactive Learning Expert System for the Quizzes. They. In [2], authors proposed multiple-choice based quiz application using QuickBasic and JavaScript. Finally, the accuracy of correct answers is displayed after calculation. This quiz provides users the feature of making their own quiz. The operational version of the afore-mentioned system is available at [3].

The Web-based expert system proposed by [4], is the Student Edition for learning and preparation. It is a multiple choice quiz system. After each and every question, five choices are given. Users can select a single choice at a time. After giving an answer to all of the questions, users will submit the answers, and then a result or progress report is displayed containing total number and accuracy of correct, incorrect and un-answered questions. Email facility enables the users to send email to: their own id, an instructor, TA, and others.

The Quiz Hub [5] is an online Interactive Learning Quiz Games, focused on facts. This quiz has many sub-categories. It provides many fields to users, students and learners for the learning purpose. The categories are Math facts, U.S. History, Multiply fractions, Vocabulary Quiz, Spelling Quiz Game, Physics, and others. It is not a multiple choice quiz; one has to select the matching pairs in this quiz. Android is rapidly getting famous day by day, and the number of its users are increasing with each passing day, because it is easy to access the necessary Android-based applications on smartphones and tablets. Therefore, we found this idea easy and time efficient to facilitate the users in this way without any difficulty. There are many online quiz applications available on the internet, but most of them are only for entertainment and fun. Moreover, if one is going to appear in any test or interview, then it is time-consuming for them to read the full books or articles related to specific fields for the preparation or revising their knowledge.

However, the most attractive feature of our app is that we take learning and fun side by side. Our app provides them the facility to revise their knowledge or to learn something advantageous at one place without wasting their time.

The objective of this project is to develop an Android-based system with following features, namely: (i) Questions bank, (ii) Time frame, (iii) Life lines, (iv) Data Storage, and (v) Multimedia support (pictures, snapshots, tables). The objective of creating this Quiz app is to help the users to prepare for necessary educational purposes regarding Computer Science and IT field with an easy access to our app directly on their Android phones. Through our app, users can learn and prepare themselves for interviews, tests and exams on Android phones, and can also use this app for increasing their general knowledge about Computer Science, Verbal and Analytical, everywhere and anytime.

Material we used is Window 10 Haier laptop, MIT

App Inventor 2 software, Windroy, QMobile Noir LT700, and Nokia Smart Phone.

Although there are a number of web-based and Android-based applications which are, one way or other related to quiz, there are only few that help in learning and contribute to the academic enhancement of the students. Most of the available applications are aiming at having a fun or entertainment. Among the many applications, we review some Web-based and Android based applications that are quite famous and are successful regarding the amount of players and downloads.

Computer General Knowledge Quiz section is a repository of Multiple Choice Question that makes you aware about evolving nature of the competitive examination; this quiz is about subjects related to the computer field. It's a general computer quiz. This quiz is useful for the preparation of any computer field test. In this quiz app, questions are given along with four choices, and at the end, the correct choice is also given. After preparation, students can check their level of preparation through the quiz [6].

It is a Computer Science Quiz. It contains multiple choice questions and answers with explanations and examples. Operating System, Database Management System, Software Engineering, Computer Networks, Digital Electronics are the sub-fields present in this quiz. These Computer Science MCQs will help users for various Interviews, competitive exams, entrance exams, and others [7].

TreeKnox Computer Quiz is a quiz system for the help and preparation of computer science and IT students who are going to appear in any interview, tests or exams in computer science and IT field. Questions are given along with multiple choices and at the end of each question; a button named "Answer" is given. On clicking that button the correct answer is highlighted at the mean time [8]. This quiz application is very simple and interactive. In this there are two modes General and Aptitude, after selecting one of them it will be redirected to the Quiz interface which will contain the question with multiple answers (options) and contain three buttons "Submit", "Show Answer" and "Next" [9].

It is also a simple and interactive application [10]. It contains three modes "Easy", "Normal" and "Hard". After selecting one of them it will be redirected to the Quiz interface which contains questions with two options, True "T" and False "F". It also show hints when user wants but if the user will try to use this life line "Hint" more than one time then it will show the answer not hint. Thus, it is useless because user can't learn anything from it anymore.

Although there are many apps that focus on the quiz, there are limited applications with focus on learning or improving knowledge in the curriculum area. Most of the other apps are entertainment-based with little focus on the educational paradigm.

There are many limitations with the existing systems mentioned above. To overcome such limitations, we propose user-friendly application, namely "*Quizzes*," which mainly focuses on gaining the curriculum knowledge as well as entertainment. Therefore, where one is amazed at playing the quiz, he/she is gaining curriculum knowledge with emphasis on not only gaining good grades but also having a better understanding of the subject matter.

Another unique feature of Quizzes that is lacking in other apps is the life lines, which it provides to the user. Users can view the hints for the right answer, can skip a question and also pause the quiz app for thirty seconds. We provided the life lines for the particular questions or the category itself, but user can use these life lines only once.

Other features regarding Quizzes and other apps seemed to be quite similar, i.e. answering questions with multiple choices as fast as possible, scoring as high as possible among the group, and so on.

There are many systems on the quiz-related content analysis in the context of opinion mining and other disciplines of computer science [11, 12, 13, 14, 15, 16, 17], however, most of such studies are web-based and address the user generated contents.

The quiz above are either web-based recommendation systems or intelligent expert systems. Therefore, there is a need to develop an Android-based easy to use application.

II. Methodology

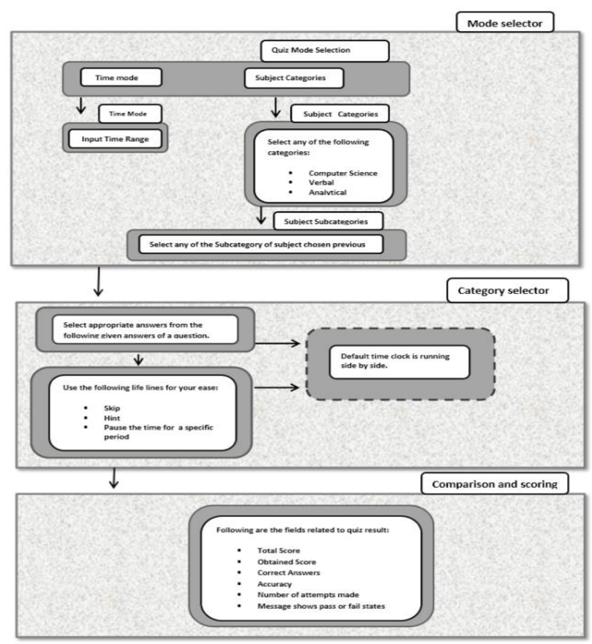


Fig. 1. The proposed system

The proposed framework Fig. 1. comprises the following modules: (i) Mode selector, (ii) Category selector, and (iii) Comparison and Scoring.

A. Mode Selector

The mode selection module allows the user to select a mode out of given modes, that is, Time mode and Categories mode. If the user selects any mode out of these modes then he/she has to give certain inputs, like set the time range, selecting categories and sub-categories. The detailed flowchart of mode selection is presented in Fig. 2.

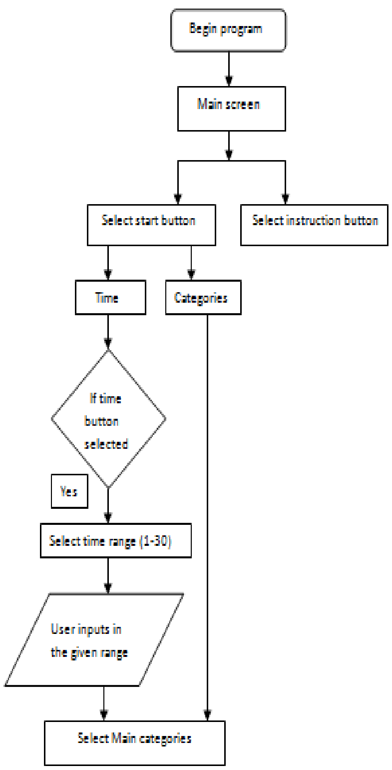


Fig. 2. Flow chart of mode selector

Algorithm.1 selecting the mode

Purpose: The purpose of this pseudo code is to select the mode. User can select any mode i.e. time or category.

Input: Time, Categories, Time Range

Output: time range, main categories

Begin

1. Lbtime← 0 // initialize time range variable
2. If Mode = Time Then
3. {Display time range}

// ……………………

4. If (time screen is initialized) then
5. {Set Listpicker.element = 1 to 30
6. Set continuebutton.enabled = false
7. Call DB.store value}
8. IF (listpicker.afterpicking =true) then
9. {Set continuebutton.enabled = true
10. Set lb time= listpicker.selection}
11. If listpicker.selection<10 then
12. {Append 0 ahead of Lbtime
13. Lbtime= lbtime – 1}
14. If (Mode = NonTime) Then
15. Display main categories
16. End

B. Category Selector

In this module, the questions are given along with the multiple choices Fig. 3. The user has to give input by selecting any of the given options, and then these inputs are used in the calculation of final result. Some life lines are given to the user which assists to play the quiz more efficiently.

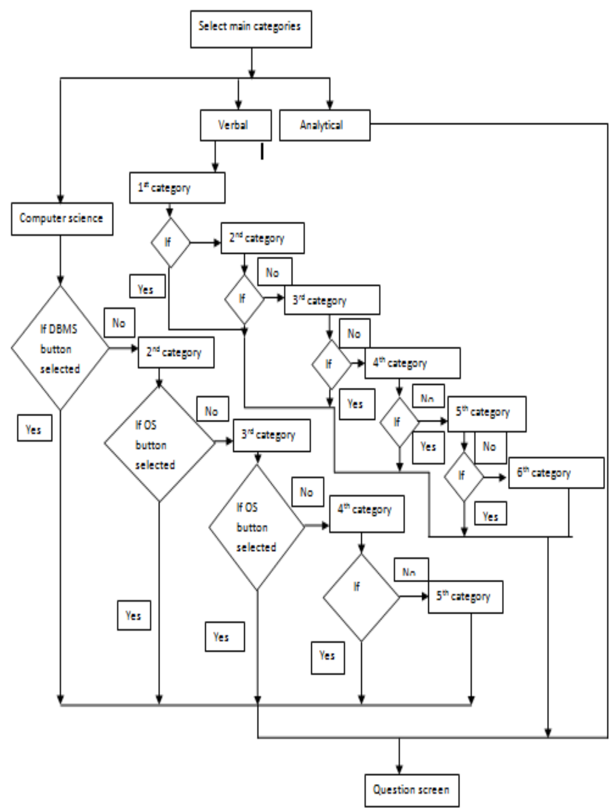


Fig. 3. Flow chart of category selector

Algorithm.2 selecting the main category

Purpose: The purpose of this pseudo code is to select the sub category of main category.

Input: Main_categories_list={Computer, Verbal, Analytical}, sub_categories_list1={Operating system, Database Management System, Computer Architecture, Data Structure, Computer Networking}, sub_categories_list2={Synonyms, Antonyms, Spellings, Ordering of words, Selecting words, Verbal Analogies}

Output: Sub categories_list1={Operating system, Database Management System, Computer Architecture, Data Structure, Computer Networking}, sub_categories_list2={Synonyms, Antonyms, Spellings, Ordering of words, Selecting words, Verbal Analogies}, Question screen

Begin

// Initialization

1. Initialize sub categories {Operating system, Database Management System,

Computer Architecture, Data Structure, Computer Networking, Synonyms, Antonyms, Spellings, Ordering of words, Selecting words, Verbal Analogies, Analytical } by assigning the list of questions and option answers.

2. If category = Computer Science Then

3. {Select any computer science's sub category from sub_categories_list1}

4. If (any of the subjects from sub_categories_list1 is selected) Then

5. {Set "category" tag to sub_categories_list1.selected

6. Display Question Screen }

7. Else If (category = Verbal)Then

8. {Select any Verbal's sub category fromsub_categories_list2}

9. If any of the sub_categories_list2 is clicked Then

10. {Set "category" tag to sub_categories_list2.selected

11. Display Question Screen}

12. Else category = Analytical Then

13. Display Question screen

14. End if

End

C. *Comparison and Scoring*

Here, outputs such as total score, obtained score, accuracy, and number of attempts made are displayed as also the messages showing whether the users passed or failed the quiz, based on the given inputs provided by the users. The entire process is shown in Fig. 4.

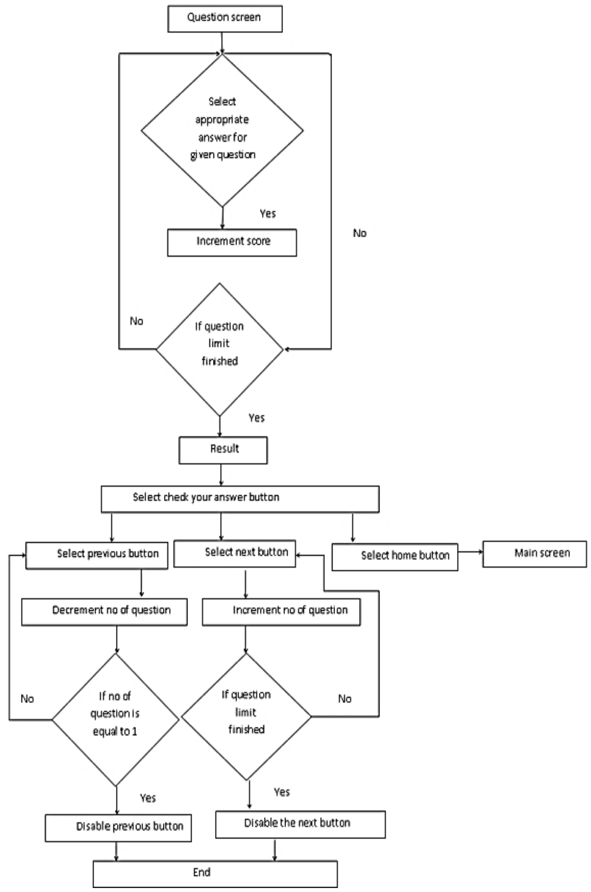


Fig. 4. Flow chart of comparison and scoring

Algorithm.3 Calculating the result and showing the right answer

Purpose: The purpose of this pseudo code is to show the appropriate answer of the question and to show the final results of the played quiz.
Input: appropriate answer for the given question
Output: result (total score, obtained score, correct answers, accuracy, number of attempts and message shows pass or fail states) , answer key

Begin:
1. Do while (question limit is not finished)
2. Select appropriate answer for the given question
3. score $\leftarrow$ score + 5
4. End while
5. Display the result screen
6. Select check your answer key
7. Select the previous button
8. (no_(of_ questions_ in_ answer key) $\leftarrow$ (no_(of_ questions_ in_ answer key) $-$ 1
9. If (No_of_ questions_ in_ answer key> 1) then
10. { Enable previous button
11. Else
12. Disable previous button}
13. **Or**
14. Select next button

15. (no_of_ questions_ in_ answerkey) ←
    (no_of_ questions_ in_ answerkey) + 1
16. If (no_of_questions _limit_finished_answerkey) then
17. { Disable next button
18. Else
19. Enable next button}
20. End if

End

III. Experimental Setup

A. Implementation

A partial list of coding of proposed modules, namely (i) mode selector, (ii) category selector and (iii) comparison and scoring.

Coding of mode selector module

In Fig. 5, there are two button events namely (i) btnTime and (ii) btncategories, and there is a list picker. In these button event, there is tinyDB in which time and categories mode selection are save user can select any mode. In list picker IF THEN statement is used, and the user can select any choice from the list.

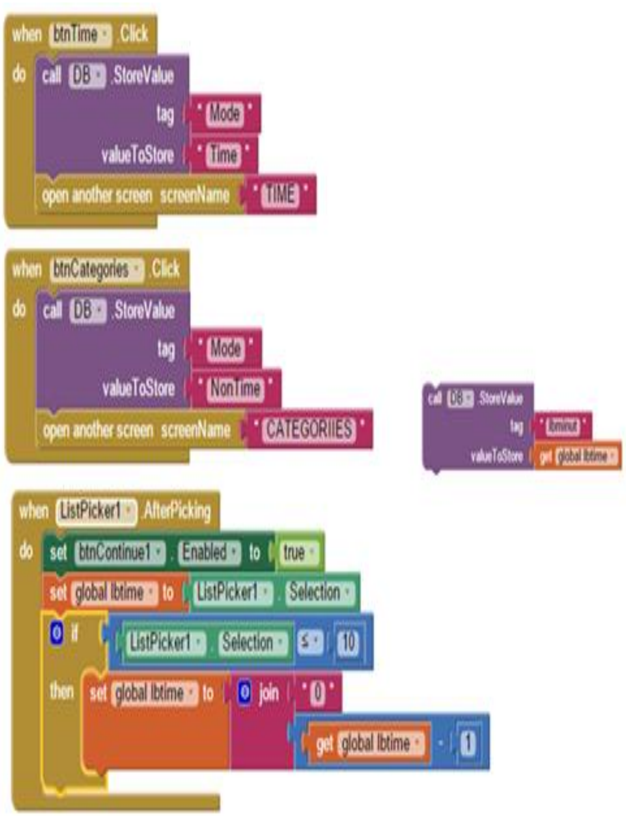


Fig. 5. Category selection and list picker

In Fig. 6, there is a list picker in which a list is made for selecting any time range. The continue button enabled property is made false.

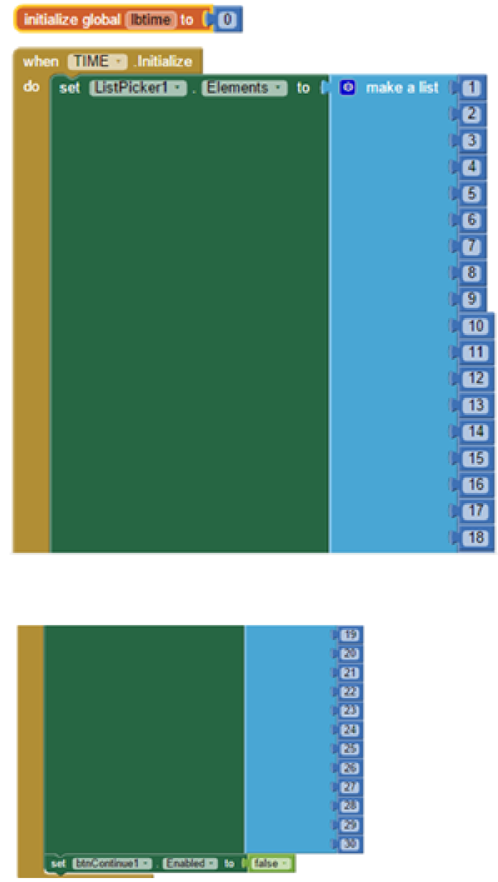


Fig. 6. Showing the list

In Fig. 7, there are five button events. In each button event, there is tinyDB; tinyDB is used to store the data and value. After selecting the desired category, the question screen will be opened.

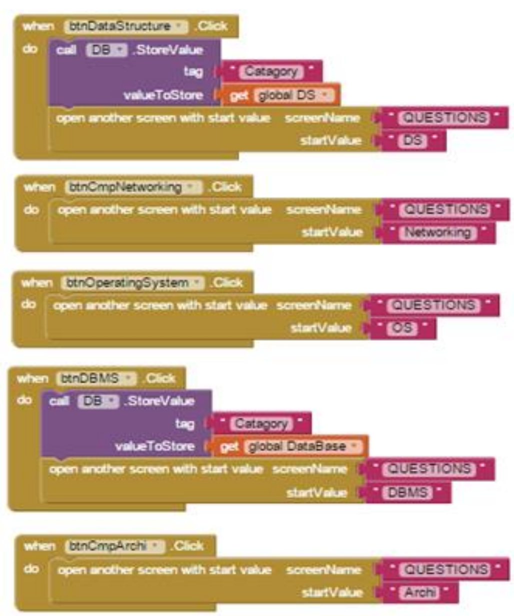


Fig. 7. Selecting sub-category and opening of question screen

In Fig. 8, there are different sub-categories buttons namely

(i) BtnAntonyms (ii) BtnSpottingErrors

(iii) BtnComprehension (iv) BtnSynonyms

(v) BtnSelectingWords (vi) BtnSpelling

(vii) BtnVerbalAnalogies (vii) BtnOrderingOfWords

These buttons are event handlers. In this event handler, the TinyDB is used to store the desired questions of the all sub categories.

Fig. 8. Sub categories of main verbal category

In Fig. 9, the btnAnalytical is event handler in which the TinyDB is used to store the questions of the analytical category.

Fig. 9. Showing the question of analytical main category

Coding of comparison and scoring module

In Fig. 10, these blocks of code life line time are enabled as false when question is initializing and there are three procedures. First procedure calls for the time and question category, the second procedure calls for the next category, and the last procedure calls for the next question.

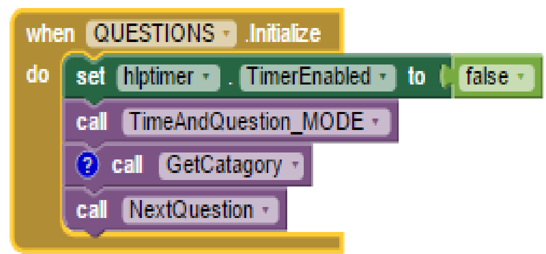


Fig. 10. Code blocks of life line

In Fig. 11, the procedure namely TimeAndQuestion_Mode is used in which IF THEN ELSE statement is used, TinyDB is also used and some textboxes and labels.

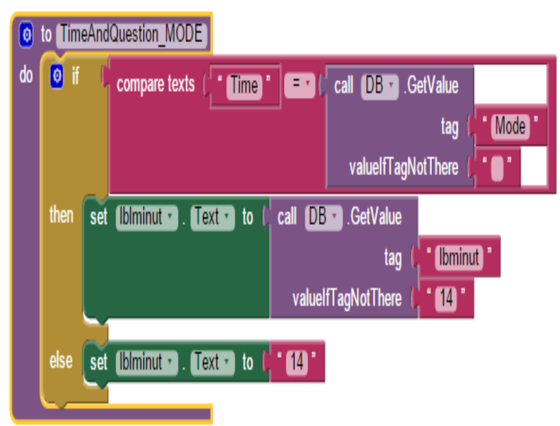


Fig. 11. Time and question mode

The procedure namely Final Result is used in which TinyDB is used nine times. IF THEN, ELSE IF THEN is used. All these instructions are executed in sequence. Set of labels and variables are also used and some of the blocks are dropped from the math Block editor. Fig. 12 shows the following code blocks. In final results, the results are calculated.

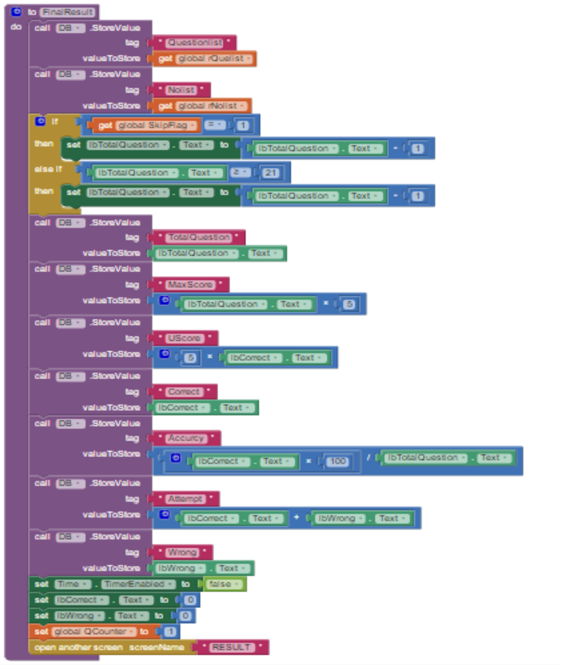
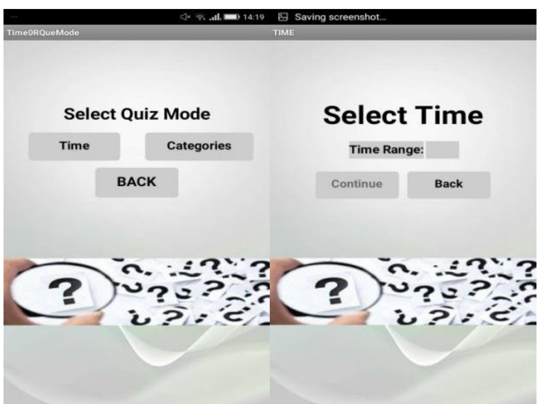


Fig. 12. Final result coding

B. Results

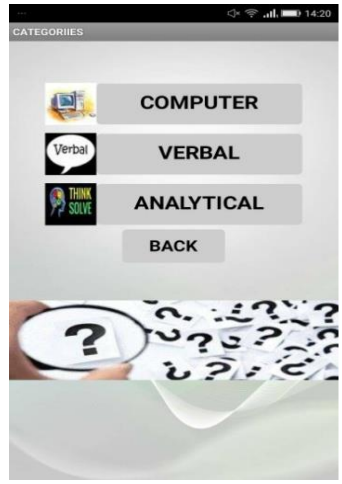
We executed our Quizzes application using Android-based platform. Fig. 13 to 7 shows the output screens of the main application.



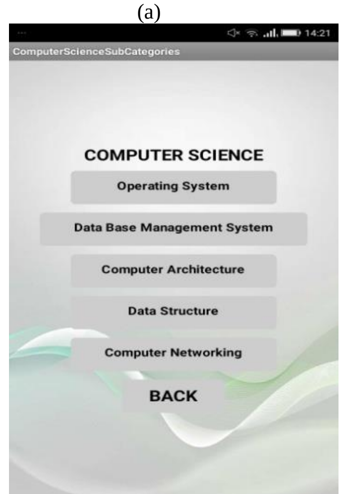
(a) (b)

(c)

Fig. 13. (a) (b) (c) Quiz mode selection module

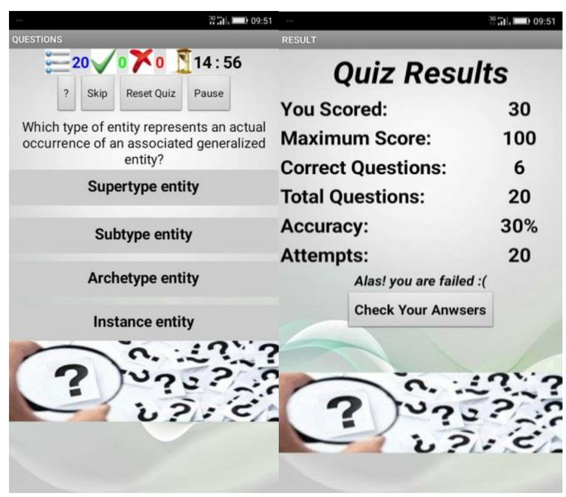


(a)



(b)

Fig. 14. (a) (b) Categories selection



(a) (b)

Fig. 15. Questions and Results (a) input (b) output

C. Descriptive Analysis of data

We asked some questions about our application to users and with the feedback we analysed our application.

1) Gender

TABLE I. SHOWING THE BASIC STATISTICS OF GENDER

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	2.00	2.00	1.56	0.50

The minimum and maximum show the smallest and largest number answer choice that obtained minimum one response. It is beneficial to find the range of answers. The minimum and maximum of 1 and 2 show that there were 16 responses in the uppermost answer (i.e. Female) and 20 responses in the lower most answer (i.e. Male). The median represents the answer choice in the center of all your responses, that is, 50% below and 50% above the middle answer choice. The median of 2.00 (higher than the 1.56 mean) shows that there were more respondents who were male than the respondents who were female. A mean is the average of whole responses by adds up all the numbers and then divide them by total amount of number. A mean of 1.56 represents the overall respondents came in somewhere between male and female. Ultimately, the standard deviation is 0.50 which shows the progress, dispersion, and variation of your responses.

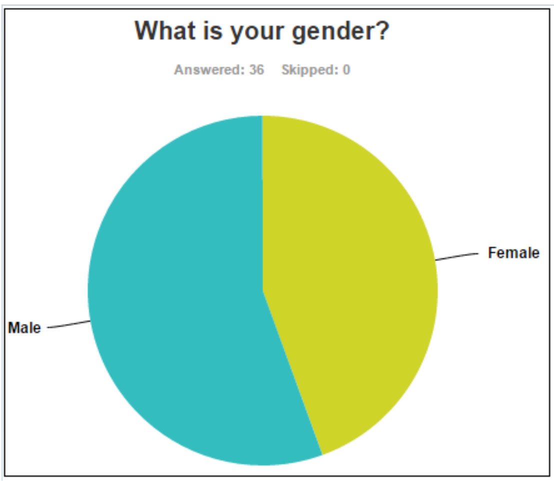


Fig. 16. Pie chart of gender

Fig. 16 shows the percentage distribution of the respondents according to gender: 44.44% of respondents were female, and 55.56% were male.

2) Age

TABLE II. SHOWING THE BASIC STATISTICS OF AGE

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	5.00	2.00	2.85	0.83

The minimum and maximum of 1 and 5 show that there was one response in the uppermost answer(i.e. age 18 to 24) and one response in the lower most answer(i.e. age 55 to 64). The median of 2.00 (lower than the 2.58 mean) shows that there were more respondents whose age was between 25 and 34 than the respondents whose age was in between 35 and 44. A mean of 2.58 represents that the overall respondents came in somewhere between 25 and 34 and 35 and 44. Ultimately, the standard deviation is 0.83 which shows the progress, dispersion, and variation of your responses.

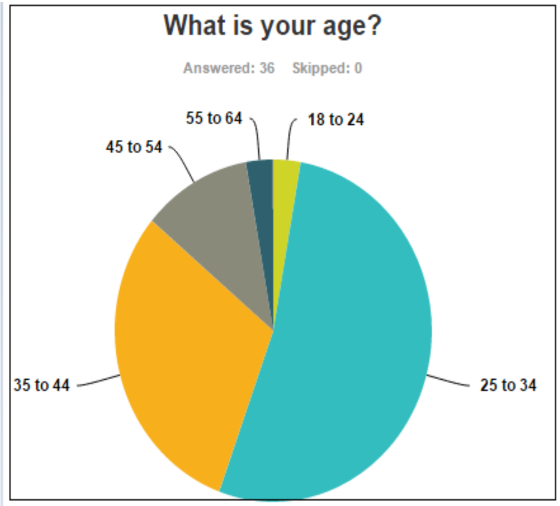


Fig. 17. Pie chart showing the age

Fig. 17 shows that there were 2.78% respondents whose age is between 18 and 24, 52.78% respondents whose age is between 25 and 34, 30.56% respondents whose age is between 35 and 44 and 2.78% respondents whose age is between 55 and 64.

3) Engaged in educational activities

TABLE III. SHOWING THE BASIC STATISTICS OF ENGAGEMENT IN EDUCATIONAL ACTIVITIES

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	2.00	1.00	1.36	0.48

The minimum and maximum of 1 and 4 shows that there were six responses in the uppermost answer (i.e. once in a while) and six responses in the lower most answer (i.e. almost never). The median of 2.00 (lower than the 2.29 mean) shows that there were more responses which were engaged with this app sometimes than the responses who were engaged almost all the time with the app. A mean of 2.29 represents the overall respondents came in somewhere between Sometimes or Once in a while. Ultimately, the standard deviation is 0.94 which shows the progress, dispersion, and variation of responses.

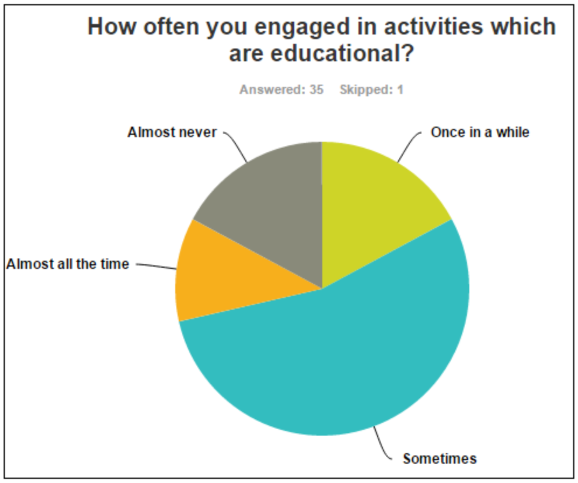


Fig. 18. Pie chart of activities in education

Fig. 18 shows that there were 17.14% respondents engaged in activities which are educational once in a while, 54.29% respondents engaged sometimes, 11.43% respondents engaged almost all the time, and 17.14% respondents who engaged almost never in educational activities.

4) Clear and understandable

TABLE IV. SHOWING THE BASIC STATISTICS OF THE IS CLEAR AND UNDERSTANDABLE

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	4.00	2.00	2.29	0.94

The minimum and maximum of 1 and 4 show that there was one response in the uppermost answer (i.e. strongly disagree) and three responses in the lower most answer (i.e. strongly agree). The median of 3.00 (higher than the 2.72 mean) shows that there were more respondents who were agreed than the respondents who were strongly agreed. A mean of 2.72 represents that the overall respondents came in somewhere between agreed or strongly agreed that the interaction with this application is clear and understandable. Ultimately, the standard deviation is 0.65 which shows the progress, dispersion, and variation of your responses.

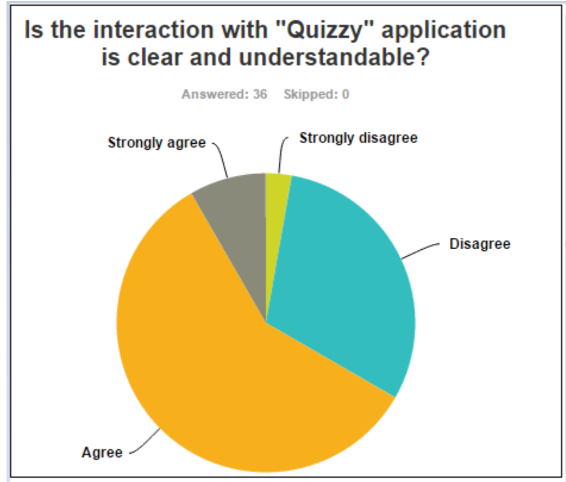


Fig. 19. Pie chart showing respondents calling the application as clearly understandable

Fig. 19 shows that 2.78% respondents strongly disagreed with the statement, 30.56% respondents disagreed, 58.33% respondents agreed, and 8.33% respondents strongly agreed with the statement that the interaction with “Quizzes” application is clear and understandable.

5) Effort and practice required

TABLE V. SHOWING THE BASIC STATISTICS OF THE EFFORT AND PRACTICE REQUIRED

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	4.00	3.00	2.72	0.65

The minimum and maximum of 1 and 4 show that there were four responses in the uppermost answer (i.e. strongly disagree) and one response in the lower most answer (i.e. strongly agree). The median of 2.00 (lower than the 2.26 mean) shows that there were more respondents who disagreed than the respondents who agreed. A mean of 2.26 represents the overall respondents came in somewhere between disagreeing and agree. Here the standard deviation is 0.69 which shows the progress, dispersion, and variation of responses.

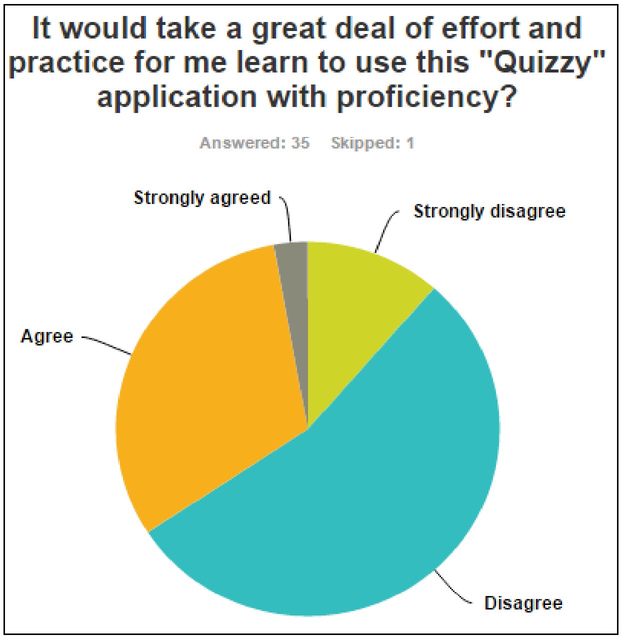


Fig. 20. Pie chart showing the effort and practice required

Fig. 20 shows that 11.43% respondents strongly disagreed with the statement, 54.29% respondents disagreed, 31.43% respondents agreed, and 2.86% respondents strongly agreed with the statement that it would take a great deal of effort and practice for them to learn to use this “Quizzes” application with proficiency.

6) Satisfactory user interface design application or not

TABLE VI. SHOWING THE BASIC STATISTICS OF SATISFACTORY USER INTERFACE DESIGN

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	4.00	2.00	2.26	0.69

The minimum and maximum of 1 and 2 show that there were 23 responses in the uppermost answer (i.e. yes) and 13 responses in the lower most answer (i.e. No). The median of 1.00 (lower than the 1.36 mean) shows that there were more respondents who said Yes than the respondents who said No. A mean of 1.36 represents the overall respondents came in somewhere between Yes and No. Ultimately, the standard deviation is 0.48 which shows the progress, dispersion, and variation of your responses

Fig. 21 shows that 63.89% respondents replied Yes that they are satisfied with the design of this application’s user interface and 36.11% respondents replied No that they are not satisfied with the design of this application’s user interface.

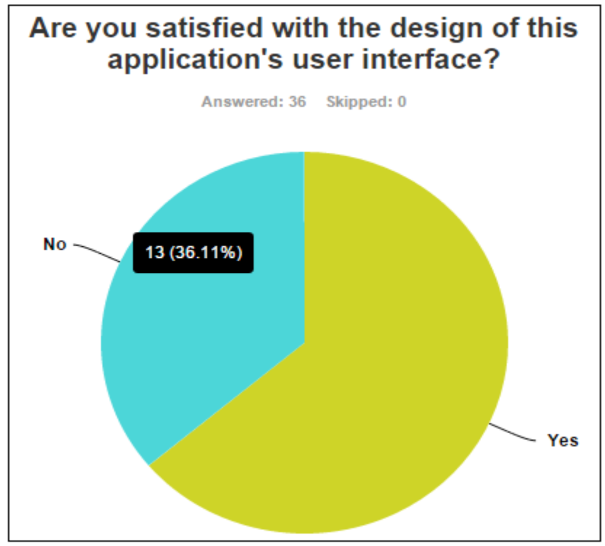


Fig. 21. Pie chart of user satisfaction for the design of the application

7) Why use Quizzes?

TABLE VII. SHOWING THE BASIC STATISTIC OF THE REASON FOR USING THE QUIZZY APPLICATION

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	4.00	4.00	3.11	1.07

The minimum and maximum of 1 and 4 show that there were three responses in the uppermost answer (i.e. For time pass) and 20 responses in the lower most answer (i.e. For education and entertainment purpose). The median of 4.00 (higher than the 3.11 mean) shows that there were more respondents who used this application for education and entertainment purpose than the respondents who used this application for education purpose only.

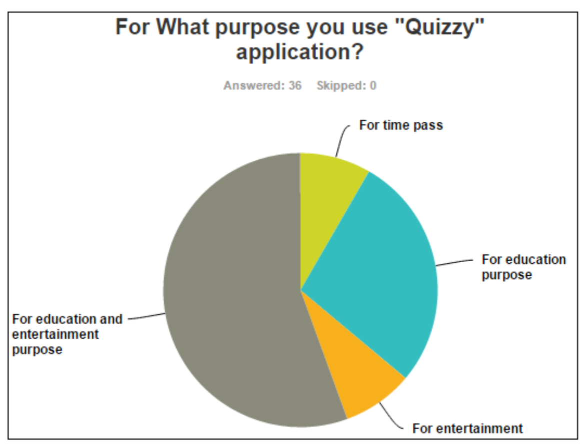


Fig. 22. Pie chart showing the purpose of Quizzes application

A mean of 3.11 represents the overall respondents came in somewhere between for education purpose and for education and entertainment purpose. Ultimately, the standard deviation is 1.07 which shows the progress, dispersion, and variation of your responses.

The above Fig. 22 shows that 8.33% respondents used this application for time pass, 27.78% respondents used it for education purpose, 8.33% respondents used for entertainment, and 55.56% respondents used this application for education and entertainment purpose.

8) Satisfied with Quizzes' ability or not

TABLE VIII. SHOWING THE BASIC STATISTIC OF APPLICATION ABILITY

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	4.00	3.00	2.74	0.65

The minimum and maximum of 1 and 4 show that there was one response in the uppermost answer (i.e.Very unsatisfied) and three responses in the lower most answer (i.e. Very satisfied). The median of 3.00 (higher than the 2.74 mean) shows that there were more respondents who were satisfied than the respondents who were unsatisfied. A mean of 2.74 represents the overall respondents came in somewhere between unsatisfied and satisfied. Ultimately, the standard deviation is 0.65 which shows the progress, dispersion, and variation of your responses

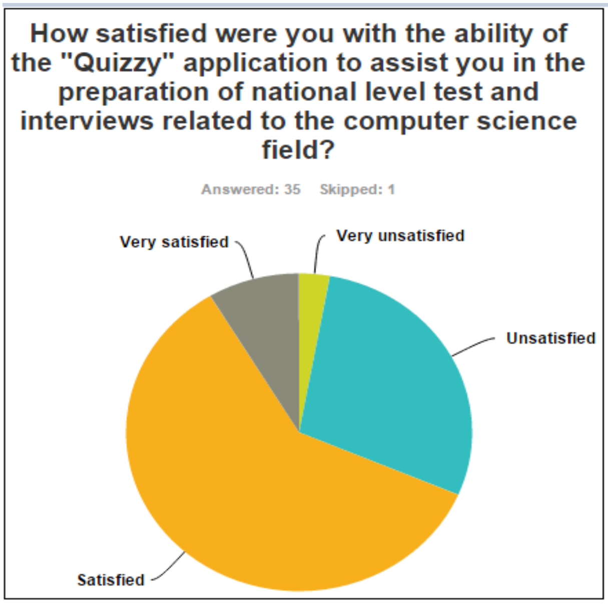


Fig. 23. Pie chart showing the ability of Quizzes application

Fig. 23 shows that 2.86% respondents were very unsatisfied, 28.57% respondents were unsatisfied, 60.00% respondents were satisfied and 8.57% respondents were very satisfied with the ability of the “Quizzes” application to assist them in the preparation of national level tests and interviews related to the computer science field.

10. Will you recommend this application or not?

TABLE IX. SHOWING THE BASIC STATISTICS OF RECOMMENDING THE APPLICATION OR NOT

Sr.no	basic statistics				
	Minimum	maximum	Median	mean	standard deviation
1.	1.00	2.00	1.00	1.31	0.46

The minimum and maximum of 1 and 2 show that there were 25 responses in the uppermost answer (i.e. I will recommend) and 11 responses in the lower most answer (i.e. I will never recommend).The median of 1.00 (lower than the 1.31 mean) shows that there were more respondents who will likely to recommend this application to others than the respondents who will never recommend this application. A mean of 1.31 represents the overall respondents came in somewhere between who will likely to recommend and who will never recommend. Ultimately, the standard deviation is 0.46 which shows the progress, dispersion and variation of your responses.

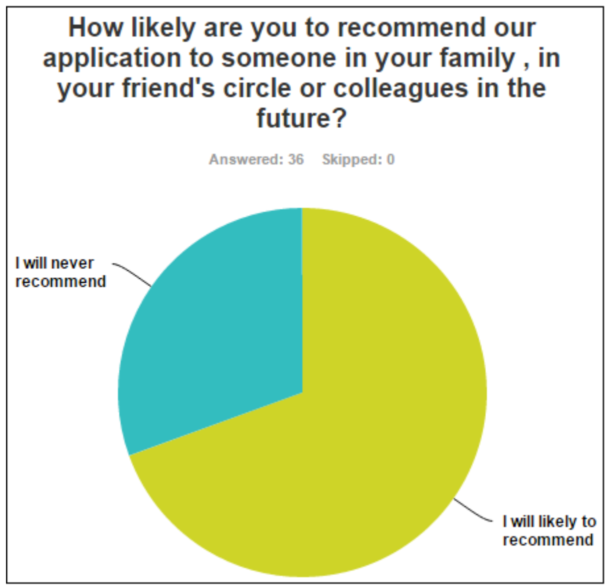


Fig. 24. Pie chart showing the recommendation of the application.

Fig. 24 shows that 69.44% respondents were likely to recommend this application and 30.56% respondents will never recommend this application to others.

REFERENCES

[1] Byers, J.A. 1999.” Interactive Learning Using Expert System Quizzes on the Internet. Educational Media International” 36:191-194. Available at: http://www.chemical-ecology.net/papers/expert.htm/ last accessed, 22 Nov, 2015

[2] Available at: http://www.wcrl.ars.usda.gov/download/itquiz.zip/ last accessed, 22 Nov 2015

[3] Available at: http://wcrl.ars.usda.gov/cec/udt/exam-f.htm/ last accessed, 22 Nov 2015

[4] Systems analysis and Design methods Available at: http://highered.mheducation.com/sites/0073052337/student_view0/chapter2/multiple_choice_quiz.html/ last accessed, 22 Nov 2015

[5] Quiz Hub Available at: http://quizhub.com/quiz/quizhub.cfm/ last accessed, 22 Nov 2015

[6] Jagran Josh (Simplifying Test Prep) Available at: http://www.jagranjosh.com/articles/computer-general-knowledge-quiz-1315979215-1/ last accessed, 24 Nov 2015

[7] EDU Zip TheKnowledge Hub Available at: http://www.eduzip.com/category/computer-science/ last accessed, 24 Nov 2015

[8] Tree Knox Available at:http://www.treeknox.com/gk/gk/computerquiz/24 Nov 2015

[9] Available at: https://play.google.com/store?hl=en/ last accessed, 24 Nov 2015

[10] Available at: https://play.google.com/store?hl=en/ last accessed, 24 Nov 2015

[11] Saqib SM, Asghar MZ, Ahmad S, Ahmad B, Jan MA. "Framework for Customized-SOA Projects". International Journal of Computer Science and Information Security. 2011 May 1;9(5):240.

[12] Asghar D, Zubair M, Asghar MJ. "Expert System For Online Diagnosis of Red-Eye Diseases". International Journal of Computer Science & Emerging Technologies (IJCSET). 2010;1(2):35-9.

[13] Saqib SM, Jan MA, Ahmad B, Ahmad S, Asghar MZ. "Custom Software under the Shade of Cloud Computing". International Journal of Computer Science and Information Security. 2011 May 1;9(5):219.

[14] Hussain S, Asghar MZ, Ahmad B, Ahmad S. "A Step towards Software Corrective Maintenance Using RCM model". arXiv preprint arXiv:0909.0732. 2009 Sep 3.

[15] Rashid A, Zubair MZ. An Intelligent Agent for a Vacuum Cleaner. International Journal of Digital Content Technology and its Applications. 2009;3(2):143-6.

[16] Asghar MZ, Ahmad S, Marwat A, Kundi FM. "Sentiment Analysis on YouTube: A Brief Survey". arXiv preprint arXiv:1511.09142. 2015 Nov 30.

[17] Asghar, Dr, Muhammad Zubair, and Dr Ahmad. "A Review of Location Technologies for Wireless Mobile Location-Based Services." Journal of American Science 10.7 (2014): 110-118.

Performance of Spectral Angle Mapper and Parallelepiped Classifiers in Agriculture Hyperspectral Image

Sahar A. El_Rahman[1,2]

[1]Electronics, Communication, and Computers Systems Engineering - Electrical Department
Faculty of Engineering-Shoubra, Benha University
Cairo, Egypt

[2] Computer Science Department, College of Computer and Information Sciences
Princess Nourah Bint Abdulrahman University
Riyadh, Saudi Arabia

***Abstract*—Hyperspectral Imaging (HSI) is used to provide a wealth of information which can be used to address a variety of problems in different applications. The main requirement in all applications is the classification of HSI data. In this paper, supervised HSI classification algorithms are used to extract agriculture areas that specialize in wheat growing and get a classified image. In particular, Parallelepiped and Spectral Angel Mapper (SAM) algorithms are used. They are implemented by a software tool used to analyse and process geospatial images that is an Environment of Visualizing Images (ENVI). They are applied on Al-Kharj, Saudi Arabia as the study area. The overall accuracy after applying the algorithms on the image of the study area for SAM classification was 66.67%, and 33.33% for Parallelepiped classification. Therefore, SAM algorithm has provided a better a study area image classification.**

***Keywords*—Accuracy Assessment; ENVI; Hyperspectral Imaging; Parallelepiped Classifier; Spectral Angel Mapper; Supervised Classification**

I. Introduction

Hyperspectral imaging (HSI) has been an active development and research area. Hyperspectral imaging is poised to enter the mainstream of remote sensing because the commercial hyperspectral imaging system's appearance. Remote sensing is the art and science of acquiring information about an area, object, or phenomenon by measuring the electromagnetic radiation emanating from the surface of the earth using indirect handle with the area, a phenomenon or object under implementation[1].

Hyperspectral images can be used in many various applications like as resource management, environmental monitoring, agriculture and mineral exploration as shown in Fig. 1 [2][3]. However, a recognition of the data and the nature restrictions and various processing strategies are considered an effective use of hyperspectral images.

HSI is described as a technique of spectral sensing that gathers hundreds of relatively small wave bands that supply spectral data to differentiate spectral unique objects or materials [4]. It is used in the analysis, measurement, and interpretation of the spectrum obtained from a particular object in a short, medium or long distance through remote sensing [5].

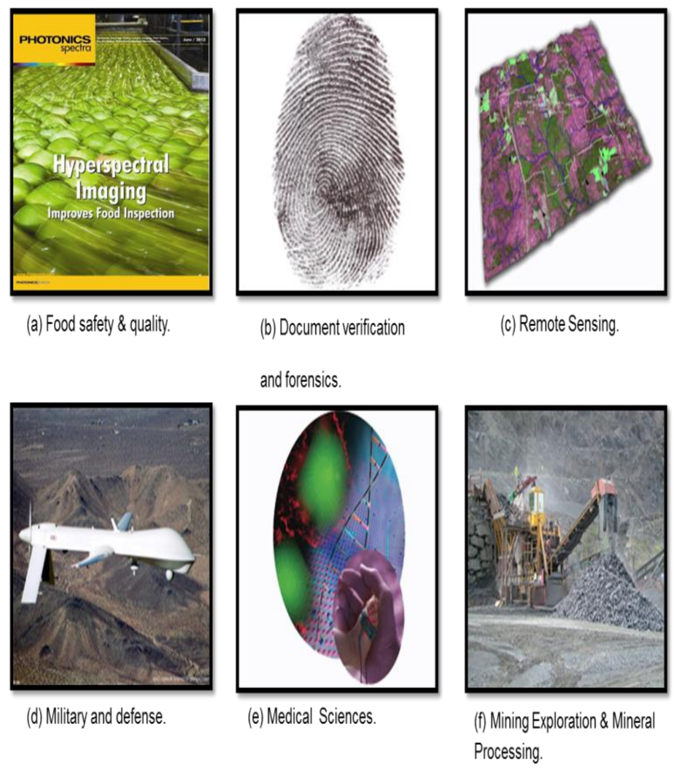


Fig. 1. Application of Hyperspectral Imaging [3]

II. Hyperspectral Imaging

Hyperspectral images provide large spectral information that is more accurate and detailed over another remotely sensed data type. It has an enhanced ability in that the probability of detecting interested materials is increased. It also provides additional information necessary for materials identification and classification [6]. The HSI pixels compose spectral vectors represent the materials spectral characteristic in the location image [4].

Hyperspectral imagers use hundreds of wavelength channels that have the capability to detect and select unique materials characteristics and features, much like DNA or a fingerprint have unique structures and features [7].

A. Advantages of Hyperspectral Over Multispectral Data

Hyperspectral sensors are expected to improve the ability to observe the earth's surface. For example, increasing the classification accuracy in contrast to multispectral imaging systems, where HSI gives an opportunity to extract additional exhaustive information over the data available using traditional multispectral imagery. The difference between hyperspectral and multispectral depends on the measurement type or on an arbitrary band number, where HSI systems provide exhaustive spectral information, which enables the analyst to classify and detect the pixels based on their spectral characteristics. However, in several cases, a multispectral imaging system has a spatial resolution higher than hyperspectral systems, however, it has less spectral channels [1].

The technologies of multispectral remote sensing have been usually utilized for remote sensing classification of vegetation since 1960s [8][9]. In a single observation, 3 to 6 spectral bands of data are generated by multi-spectral sensors which range from the visible end of the Electromagnetic Spectrum (EMS) to the NIR (Near Infrared) end [9]. This narrow window of spectral bands is the essential drawback of multi-spectral sensor. Through the previous decade, the improvement in spectrometer has helped to overcome the limitations of multispectral sensors. So, they are providing a good performance in identification, classification, and object detection of earth characteristics [10][11][12]. HSI sensor usually gathers further than 200 spectral bands that span between the visible part of EMS and SWIR (Short Wave Infrared) part. HSI sensor does not present only spectral data in detail consisting of hundreds of bands in one combination [9]. So, these features have resulted in new governmental and scholarly exploration of mapping and classification of vegetation and land cover with HSI application [10][13].

Furthermore, HSI handles narrow spectral bands through a spectral continuous range, and presents the all pixel spectra in the image, even as multi-spectral imagery deals with several images at somewhat narrow and discrete bands, as shown in Fig. 2.

The main advantage of HSI is the amount of spectral details it produces [14]. The main disadvantages are complexity and cost. Hyperspectral data analysis needs fast computers, large data storage capacities, and sensitive detectors. Also, the ways to programme hyperspectral satellite, which are found by one of the researchers, to arrange data on its own and transmits one and only the most significant images, as both storage and transmission of that large data could demonstrate the complexity and cost. The entire potential of HSI has not been explored yet [1][15].

III. Classification

Classification is a technique of information extraction, where it considered as one of the most often used techniques. Most of these techniques depend on the spectral reflectance property analysis of hyperspectral imagery and utilize specific techniques prepared to proceed several types of spectral analysis. The hyperspectral classification processes can be implemented using one of the two techniques: Unsupervised and Supervised [1][16].

A. Supervised and Unsupervised Classification

Supervised Classification, as shown in Fig. 3, needs to identify known prior as training sites out of a collection of personal experience, fieldwork, and map analysis. They are utilized for the classification algorithm training to the eventual land cover mapping of the rest of the images. Each pixel is then evaluated and assigned to the category (class) which has the maximum likelihood [17]. In the unsupervised, the pixels that have similar spectral features (covariance matrices, standard deviations, means, etc.) are grouped by the computer or algorithm into distinct classes as stated by several statistically defined specifications as shown in Fig. 4.

The Supervised hyperspectral algorithm uses the sample with recognized identity (i.e., information clusters with assigned pixels), and pixels with unknown identity are classified by the algorithm. The process begins with choosing and naming regions on the image by the user, which correspond to the clusters of concern. These clusters correspond to information clusters. Then, the algorithm of image classification will detect all analogous regions [18].

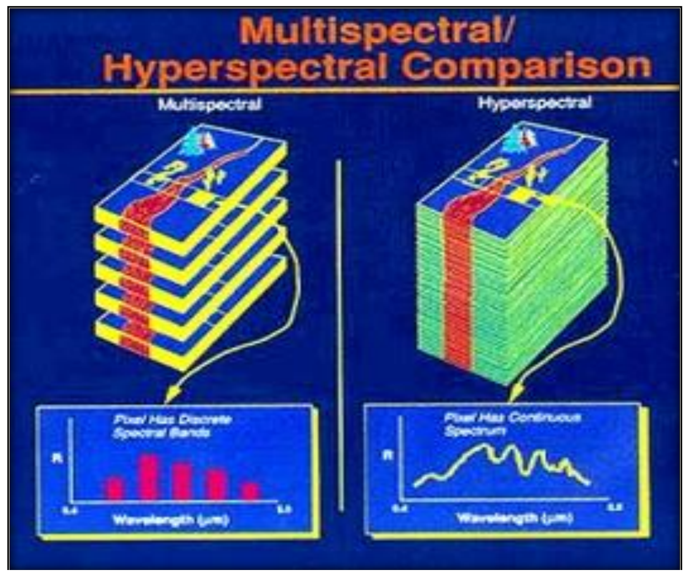


Fig. 2. Multispectral and Hyperspectral Comparison

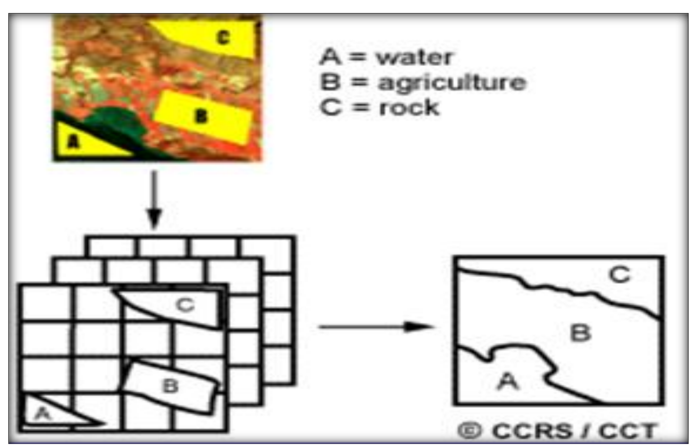


Fig. 3. Supervised Classification [19]

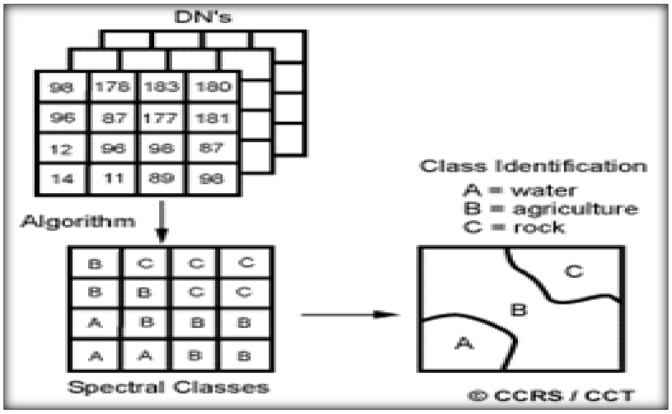


Fig. 4. Unsupervised Classification [19]

B. Supervised Algorithms

There are some supervised methods which have been designed to outline the problem of HSI classification. Various studies have applied algorithms and each algorithm provides different classification accuracy, but the most famous are Parallelepiped and SAM (Spectral Angle Mapper) algorithms.

1) Parallelepiped Classifiers: In this classifier, a common decision rule is used for classifying information. In the image data space, N dimensional parallelepiped classification is formed by the decision boundaries. From the mean of each specific cluster, a standard deviation threshold defines the dimensions of parallelepiped classification. A pixel is being classified and assigned to that class, if its value falls between the high threshold and the low threshold of every band in n bands. In ENVI, the pixel is allocated to the last class be matched, if the pixel value lies in more than one class. Areas are designated as unclassified that do not lie within any of the classes, as shown in Fig. 5 [20].

Parallelepiped classifier is considered as a very fast algorithm where every parallelepiped class demonstrates one information class. Each pixel with digital number into the bounds (range) determined by certain parallelepiped class will be chosen to the corresponding information class [21].

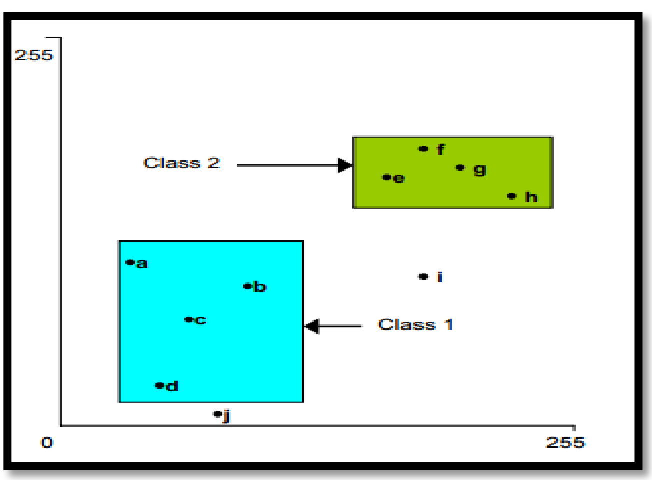


Fig. 5. Parallelepiped classifier

2) SAM (Spectral Angle Mapper) Classifier: SAM is an algorithm that allows quick mapping of the spectral symmetry of the spectra of the image to the spectra of the reference [18]. The spectra of the reference can either be a spectrum measured in a laboratory, a field spectrum or obtained straightway from the image. SAM classifier determines the spectral symmetry between both spectra, handling the two spectra as vectors in the space at the same dimension of the band number [22][23]. This can be easily explained, as shown in Fig. 6, where the spectra of the reference and the test are represented in a two-dimensional plot as two band data. The angle among every reference spectra and every test spectra is computed. The SAM program, in ENVI, assigns the angles to output channels, and then every pixel is allocated to the class defined by the reference spectrum. The class that is assigned to each pixel is saved in the output channel [18].

IV. Methodology

A. Study Area

As the Kingdom of Saudi Arabia is a desert with a dry climate, caring about its agriculture is important so it won't be worse. Studying its agriculture and which areas are infected and what areas are good for growing crops might improve its agriculture in the future. Al-kharj is selected as a study area because it is one of the important provinces in the Kingdom of Saudi Arabia for exporting wheat to markets around the world that it is existing in southeast area of Riyadh [24]. The detected area of study (as shown in Fig. 7), is downloaded from United States Geological Survey website [25].

AL-Kharj image when downloaded has the following information: Exact location is in Path 165 and Row 43. Acquisition date was in year 2007 and since Julian day is identified month and can be calculated on the basis that the Julian day was 80 so it was acquired on 23/3/2007 so we can say that Al-Kharj was scanned in the spring season. It was composed of 242 files in .tiff format where each band is represented by a certain file. Consequently, those bands have to be combined to create a single image with all bands [25].

B. Methods

An adequate preprocessing of HSI is requisite to extract helpful information from it. There are a number of preprocessing methods that should be done before starting the classifying [26]. The usage of these methods depends on the downloaded image.as shown in Fig. 8. First, we need to acquire the study area image. Then Preprocessing is needed on the image such as: collecting the bands into one image and applying wavelength for each band using Hyperion tool [27]. This tool is not found in ENVI, that it is utilized to support Hyperion data use in ENVI. Where the most basic functionality of this tool is converting Geo TIFF data sets into files of ENVI format which contain band information and wavelength [28]. Then, spatial and spectral subset is used to minimize the huge data and eliminate redundant data and atmospheric correction. After the preprocessing, the main processing takes place which is applied to two supervised classification algorithms: SAM, and Parallelepiped. Finally, these two algorithms will be compared using accuracy assessment as described in the following sections.

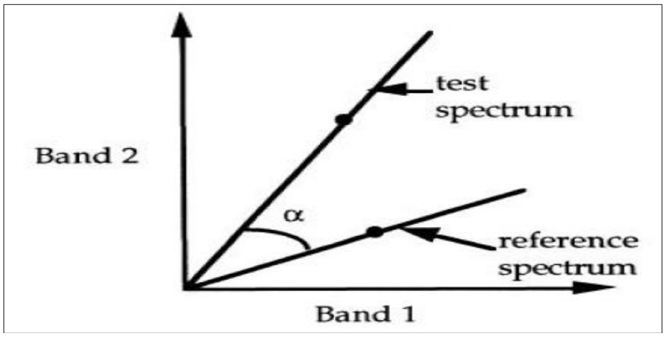
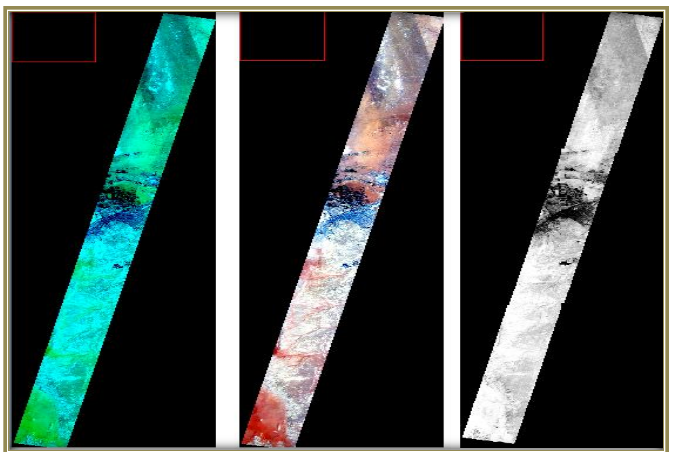


Fig. 6. The plot of a test and a reference spectrums of a two band image [18]



(a) true color (b) false color (c) gray scale

Fig. 7. Hyperspectral image for Al-Kharj Kharj [25]

1) Preprocessing: The data have been subjected to gathering all the bands into one image using the Hyperion Tool. Subsetting has been done to minimize the complexity and the big volume of the data produced by HSI sensors where it requires a lot of time to process, and to provide the base for comprehensive analysis, it is subjected to atmospheric corrections. Gathering all bands into one image using the Hyperion tool is accomplished first, then the sub setting before the atmospheric correction [27].

Transforming the image into an ENVI format file using Hyperion Tool. Hyperion tool is found in an external file in ENVI software, where it will get the study area image ready in an ENVI standard format instead of separated bands as images files with TIFF extension. The output of this process is six files: two with .DAT extension, two with .HDR extension, one with .txt extension, one with .STA extension.

Sometimes the hyperspectral image downloaded contains lot of unneeded information. This information will cause an overprocessing and minimize the performance. These unwanted data in the study will be cut off. This method is called Sub setting. This will enable the processor to focus only on data needed for a study area and get better performance of HSI classification. The process can account for the differing data quality and discernment capabilities among spectral bands, and use the spatial and spectral information with each other [29].

In spatial sub setting, the wanted area of study might be a part of the downloaded image from the satellite. Spatial subsetting is applied to the image to resize it, as well as utilize the function of resizing to produce new images with any aspect ratio or size [30]. It is focused on choosing the area of study, but it must be selected in a square shape. Fig. 9 shows a subset by image dialog showing the selected subset area.

The result after sub setting is viewed in three displays. Fig. 10-a is for false color display, Fig. 10-b is for true color display and Fig. 10-c is for gray scale display.

The Spectral Subsetting is established to identify bad bands. The ones which will don't support the analysis will only cause an overload on the processor. ENVI headers have related ancillary data (spectral library name, band name, bad bands list, wavelength, Full Width Half Maximum (FWHM)) depending on the image data type. These bad bands could be identified in ENVI, where Edit dialog of bad bands list contains 242 bands where each good band is highlighted and the rest is bad bands. In most hyperspectral images, last and first bands are considered bad. In the area of study has 155 good bands and the rest are bad.

a) Atmospheric Correction: Even a relatively clear atmosphere interacts with coming, and reflected solar energy. These interactions minimize the degree of coming energy arriving the ground for certain wavelengths. And also, they minimize the degree of reflected energy arriving an airborne or satellite sensor. So, in these regions, little useful information can be obtained from image bands. If atmospheric conditions are spatially variable, atmospheric effects may also differ between areas in a single scene, that's why we must work to rectify the image that is affected by the atmospheric gases. Also, ENVI can be used to rectify HSI image with atmospheric impact where it has multiple Atmospheric Correction methods [26][31].

Fig. 11 shows the influence of the atmosphere correction on the extent of absorption and reflectance of each pixel before and after atmosphere correction, In order to know the extent of the effect it we will view the spectral profile for a certain pixel in the image. The spectral profile for a certain pixel in the study area image before the atmospheric correction, as shown in Fig. 11-a. Output of spectral profile will be in the form of a curve, the top of each curve reflects the reversal, and the bottom of each curve reflects the absorption of each pixel, in zoom window there is the intersection of two lines point of this intersection represent the current pixel of the image and it will show the result of this pixel on spectral profile. If the pixel has a higher value reflection, it is better to detect materials [32].

Fig. 11 shows the image after applying atmospheric correction, where the spectral profile differs between the image before and after atmospheric correction. In Fig. 11-a, there is a lot of absorbed energy unlike Fig. 11-b which has a really high reflected energy.

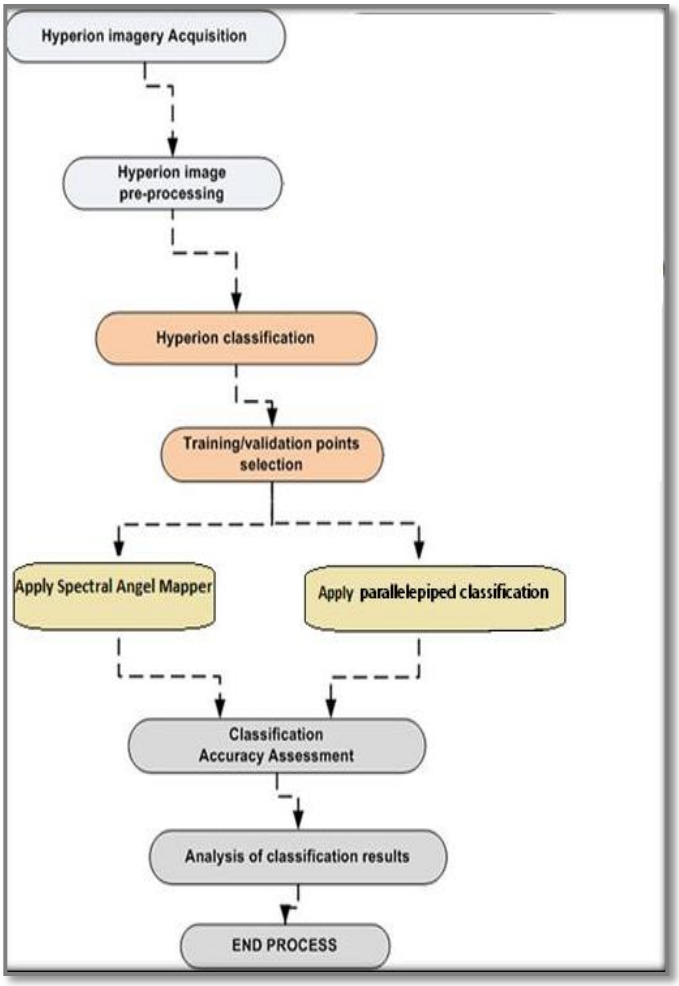


Fig. 8. Hyperspectral imaging analysis

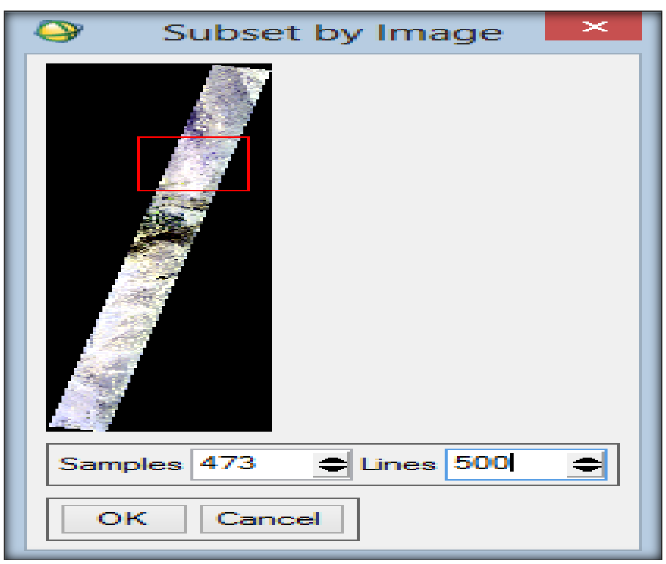
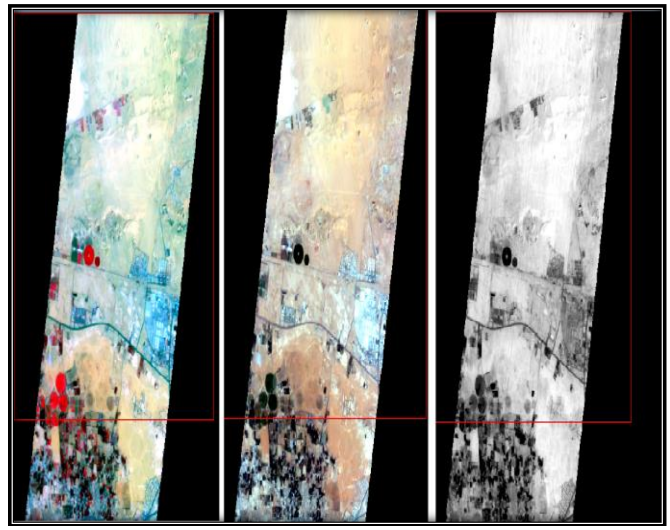
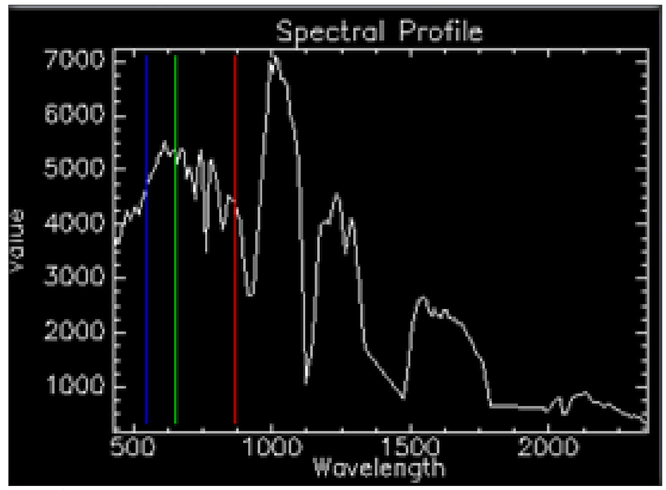


Fig. 9. Subset by Image Dialog where the red box contains the study area

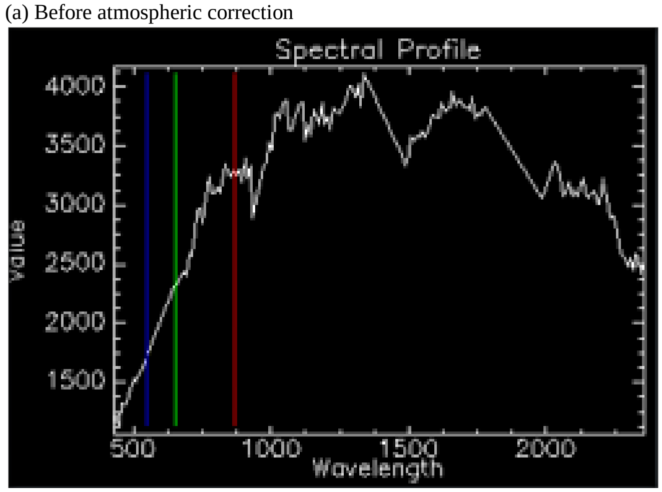


(a) False color (b) True color (c) Gray scale

Fig. 10. Spatial subset of Al-kharj study area



(a) Before atmospheric correction



(b) After atmospheric correction

Fig. 11. Spectral profile for a certain pixel

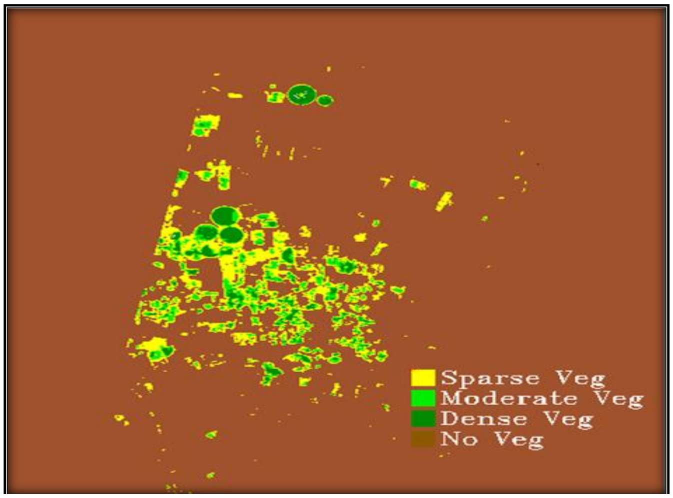


Fig. 12. Vegetation delineation and stress detection for processing complete

2) Processing: In the processing phase, NDVI (Normalized Difference Vegetation Index) values are utilized in corrected image, to show the green vegetation amount exist in the pixel. As more green vegetation is indicated, the NDVI values also increase. Such a step is just to view the amount of green area in a study area. Then, the classification methods SAM and Parallelepiped are applied.

a) NDVI: The study area image is analysed to check the presence of vegetation and the density of the vegetation. The Vegetation Delineation tool enables us to fast recognize the vegetation presence and to see its vigor level. Also, the Wizard supplies helpful tools to create graphics that are used in briefings and reports.

For most applications of spectral processing, dealing with data with atmospheric correction due to produce more accurate results. Since the study area image is already corrected, there is no need to perform any correction. The NDVI equation:

$$\text{NDVI} = \frac{(NIR-Red)}{(NIR+Red)} \quad (1)$$

The NDVI generates an image that ranges from -1 to 1. Pixels with no vegetation tend towards -1, while the pixels with vigorous vegetation tend towards 1. Examine results, as shown in Fig. 12.

b) SAM (Spectral Angle Mapper): SAM is a physically based spectral classifier which uses a n - D angle to identify the pixels to reference spectrum. It determines the spectral symmetry among two spectrums by computing the angle between them and handling them as vectors with dimensions equal to the band number. Such algorithm is comparatively insensitive to light and albedo effect when utilized in standardized reflectance data. The spectrum of endmember utilized by this algorithm can come from spectral libraries, or directly, they are extracted from the image (as an average spectrum of ROI) or ASCII files [2][20]. The steps followed in SAM are shown in Fig. 13-a. This work is based on spectral library for more accuracy. Therefore, it will match each pixel spectral signature in the image of the study area to the selected vegetation endmember spectral signature.

The result of study area classification is represented in Fig. 13-b after SAM classifier is applied, where the wheat is represented by the yellow color, and wheat (tan) is represented by the green color.

c) Parallelepiped Classification: It is a widely supervised algorithm. The bands of the image are utilized to define the pixels of the training area for every band based on least and most pixel value. Though, it is more accurate than other algorithms classification, it is not more widely utilized. Due to considerable unclassified pixels are left and likewise it have an overlapping among pixels of the training area. The candidate pixel values are compared to lower and upper boundaries [33]. This method needs the training areas that are selected by the user for utilization as the classifying base. When these data are collected, a different number of classification routines are ready to identify the concerning pixels and carry out the classification.

- *Defining ROIs (Regions of interest)*

ROIs are images parts, selected either graphically or by other ways like a threshold. They are with irregular shape and are normally utilized to make statistics for masking, classification, and different related processes [26].

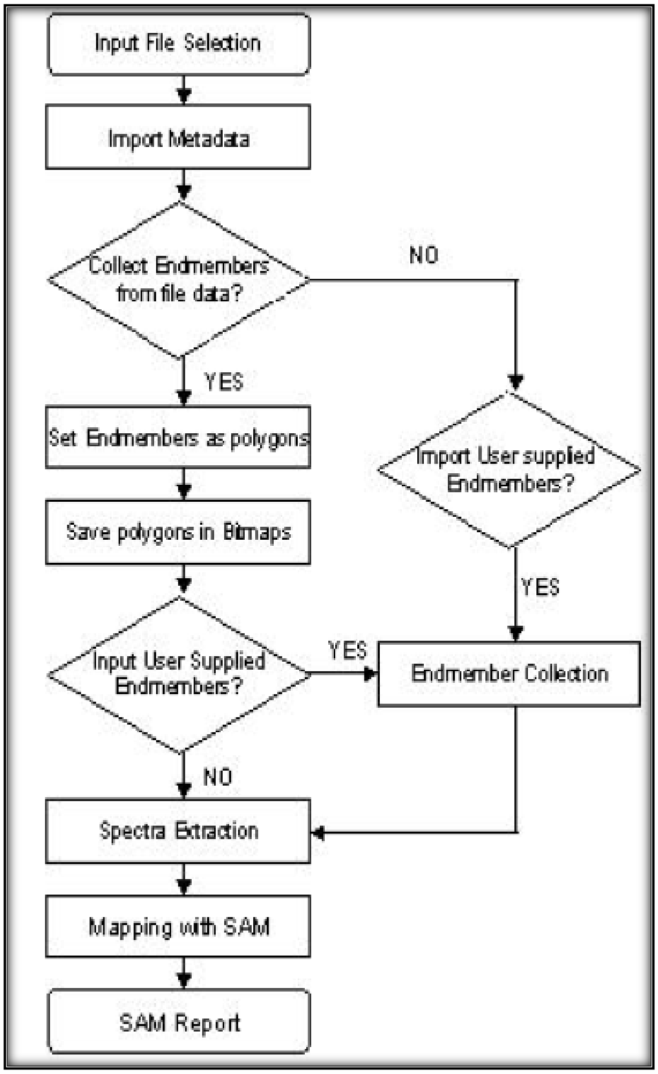
Number of ROI depends on the study area image work on. So, if the image has lots of features such as: mountains, sea, ocean grass, building. You will add region of interest based on the features you can see in the image. When the image represented by false color has red circles where they represent pivot which is an agricultural area, based on the knowledge on how pivot is represented by circles, we know it is agriculture. Since pivot is represented by red, we can say that everything with red color is agriculture, but the brightness of this red color is different from place to place and this is because the density of green land is different between them [34].

- *Parallelepiped Classification in ENVI*

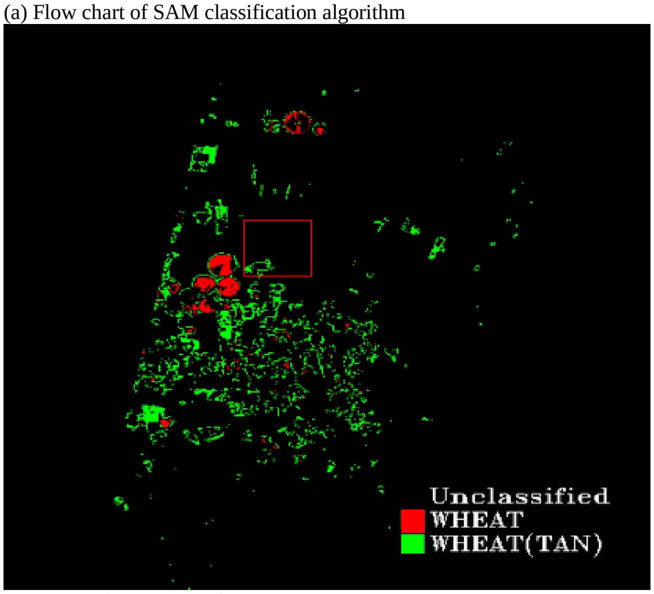
Using ROI represented above, the classes are created utilizing parallelepiped classifier. The default attributes and varied standard deviations from the mean of these regions of interest are used [17][21]. The steps followed in Parallelepiped, is shown in Fig. 14-a. Based on the ROI defined, it will search for the min, max digital values for each region in ROI in order to classify any pixel that is between the values as it belongs to this region.

The result of applying parallelepiped classification on the image has classified it only to the agricultural area where other areas were ignored, as shown in Fig. 14-b.

3) Post processing: Classified images require post-processing to perform generalizing for classes, which is called Majority Analysis. It is utilized in changing any spurious pixel within a considerable single class to that class. As shown in Fig. 15, the results after applying the majority of the two classified images. Whereas, the image is smothered but details such as certain pivot with a high density of green land has been lost. So, results before applying majority will be certified.

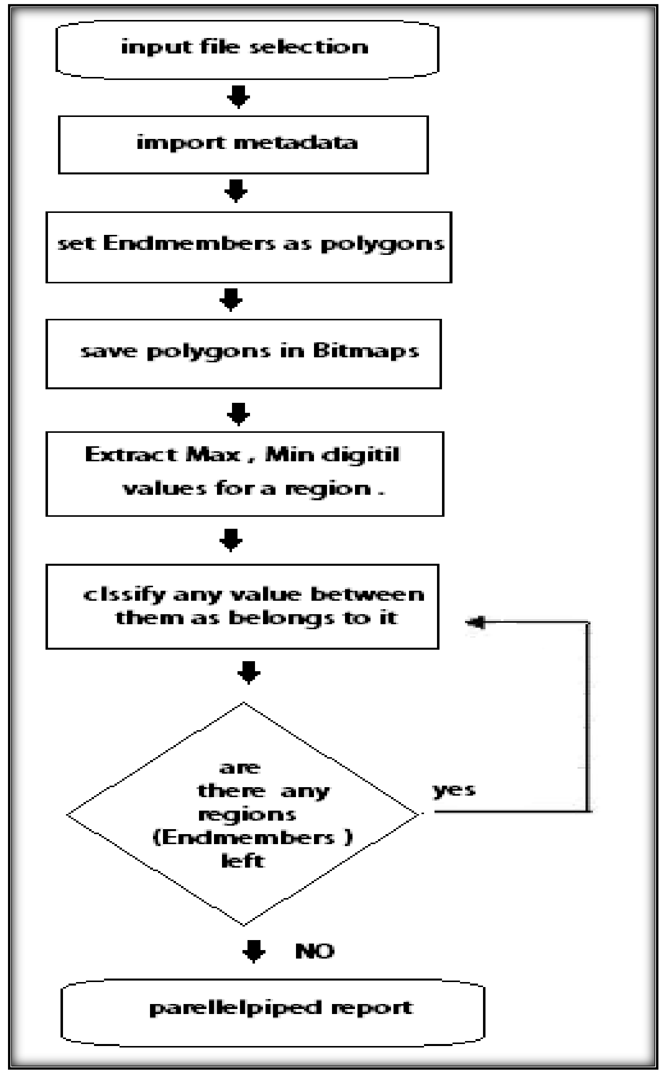


(a) Flow chart of SAM classification algorithm

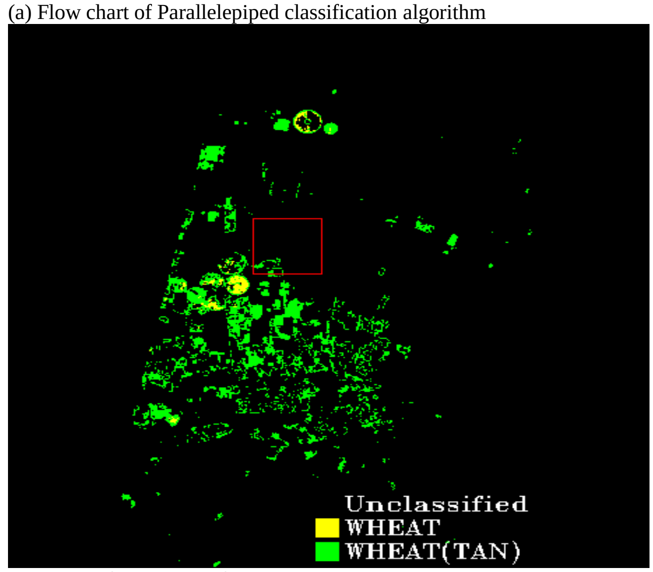


(b) Results for the area of study (Al-kharj)

Fig. 13. SAM classification

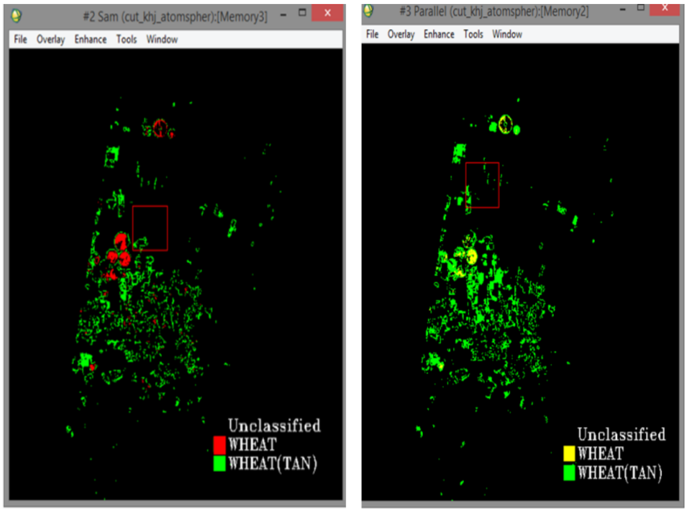


(a) Flow chart of Parallelepiped classification algorithm



(b) Results for the area of study (Al-kharj)

Fig. 14. Parallelepiped Classification



(a) (b)

Fig. 15. (a) SAM Classifier (b) Parallelepiped Classifier results for the study area (Al-Kharj)

V. Results

The image of the area of study (Al-Kharj in Kingdom of Saudi Arabia) has been classified using SAM classifier as shown in Fig. 13-b and by using Parallelepiped classifier as shown in Fig. 14-b, the validation or accuracy assessment is a significant stage in the remote sensing image process. Where it defines the data value of the resulting information to the user [35]. The aggregate accuracy is studied by taking the sum of No. of pixels classified right and divided by the total pixel number [36]. The correct class of the pixels is defined by the ground truth ROIs. The overall accuracy after applying it on the image of the study area for SAM classification was 66.67%, and 33.33% for Parallelepiped classification. Therefore, SAM provides better classification for the image of the area of study.

VI. Conclusion

This work has achieved the potential utilization of Parallelepiped and SAM classification algorithms integrated with EO-1 Hyperion imagery analysis to extract all areas of the wheat in the study region, that is Al-Kharj in Kingdom of Saudi Arabia, as it is considered as one of the regions greatly export wheat to the market. The Parallelepiped and SAM classifiers were implemented by using the identical training set and points of validation chosen on the gained EO-1 Hyperion image, that allow an explicit performance comparison of them.

SAM has better results usually because it is based on comparing pixels to reference objects using spectral signature. Parallelepiped differs from SAM, because it is depending on checking the digital value of the pixel if it ranges within the Min-Max values of the drawing ROIs.

Acknowledgment

The author would like to thank all the participants involved in this work at CCIS - PNU, especially to Abeer R. Alwalid, Ghadah A. Aljarallah, Manal I. Al Manna, Nada I. Al Rashed and Wateen A. Aliady.

References

[1] John A. Richards and Xiuping, “Remote Sensing Digital Image Analysis : An Introducion”, 4th Edition, © Springer-Verlag Berlin Heidelberg, 2006.

[2] Dongha Lee and Sunhui Sim, “The Application of Hyperspectral Sensing Data for Seabed Classification in the Coastal Area of Korea”, International Electronic Conference on Sensors and application, June 2014. www.mdpi.com/journal/sensors

[3] Headwall Photonics company, "Hyperspectral Imaging - Methods, Benefits and Applications," 2006. [Online]. Available: http://www.headwallphotonics.com/applications-old [Accessed 11 October 2013].

[4] G. Shaw and H. Burke, "Spectral imaging for remote sensing," Lincoln Laboratory Journal, VOL.14, NO. 1, pp. 3–28, 2003.

[5] R. Navalgund, V. Jayaraman and P. S. Roy, " Remote Sensing Application on overview, Current Science, VOL. 93, No.12, pp. 1747-1766, 2007.

[6] P. Shipper, “Introduction to Hyperspectral Image Analysis,” An International Electronic Journal, 2003. [Online]. Available: http://spacejournal.ohio.edu/pdf/shippert.pdf. [Accessed 10 october 2013].

[7] SpecTIR, “hyperspectral image,” SpecTIR, 2012. [Online].Available: http://www.spectir.com/technology/hyperspectral-imaging. [Accessed 10 November 2013].

[8] M. Govender, K. Chetty and H. Bulcock, " A Review of Hyperspectral RemoteSensing and its Application in Vegetation and Water Resource Studies”, vol. 3, no. 18, pp. 145-151, 2007.

[9] John R. Jensen,”Introductory digital image processing”, © Pearson Prentice Hall, 2007.

[10] U. Heiden, W. Heldens, S. Roessner, K. Segl, T. Esch and A. Mueller, "Landscape and Urban Planning.," in Urban structure type characterization using hyperspectral remote sensing and height information, New York, elsevier, pp. 361-375, 2012.

[11] S. Pignatti, M. R. Cavalli, V. Cuomo, L. Fusilli, M. Poscolieri and San, "Remote Sensing of Environment," Evaluating Hyperion capability for land cover mapping in a fragmented ecosystem: Pollino National Park, Italy,, vol. 3, no. 12, pp. 622-634, 2009.

[12] S. J. Purkis and V. V. Klemas, "Remote Sensing and Global Environmental Change", Remote Sensing and Global Environmental Change, vol. 3, no. 10, 2011.

[13] A. Jung, P. Kardevan and L. Tokei, "Physics and Chemistry of the Earth," in Detection of urban effect on vegetation in a less built-up Hungarian city by hyperspectral remote sensing, Italy, elsevier, pp. 255-259, 2005.

[14] A. Picon, O. Ghita, P. F. Whelan and P. Iriondo, “Fuzzy Spectral and Spatial Feature Integration for Classification of Nonferrous Materials in Hyperspectral Data,” IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 5, NO. 4, pp. 483-494,NOV., 2009.

[15] Schurmer, “Air Force Research Laboratories Technology,” J.H, U.k., 2003.

[16] Anthony Sgouros, “Hyperspectral Imaging and Spectral Classifcation Algorithms in Plant Pathology,”, PHD Thesies, Technical University of Crete Electronic and Computer Engineering Department,Chania, 2008.

[17] Randall Smith, “Tutorial: Introduction to Hyperspectral Imaging,” © MicroImages, Inc, USA,January 2012.

[18] F. A. Kruse, A. B. Lefkoff and J. W. Boardoman, "The Spectral Image Processing System (SIPS)- Interactive Visualization and Analysis of Imaging Spectrometer Data, Remote Sensing of Environment, © Elsevier puplishig co. Inc., pp. 145-163, 1993.

[19] Nrcan.gc.ca, “remote-sensing,” 2008. [Online]. Available: http://www.nrcan.gc.ca/earth-sciences/geography-boundary/remote-sensing/fundamentals/1920. [Accessed 27 septmber 2013].

[20] ENVI, “ENVI User’s Guide”, Copyright © ITT Visual Information Solutions All Rights Reserved, August, 2007.

[21] Randall B. Smith, "Tutorial: Image Classification”, © MicroImages, Inc, USA, April 2011.

[22] Helmi Z. Shafri, A. Suhaili and S. Mansor, "The Performance of Maximum Likelihood, Spectral Angle Mapper, Neural Network and Decision Tree Classifiers in Hyperspectral Image Analysis", Journal of Computer Science 3, ISSN 1549-3636, pp. 419-423, 2007.

[23] G. P. Petropoulos, K. P. Vadrevu, G. Xanthopoulos, G. Karantounias and M. Scholze, "A Comparison of Spectral Angle Mapper and Artificial Neural Network classifier Combined with Landsat TM Imagery Analysis for Obtaining Burnt Area Mapping", no. 1424-8220, p. 19, 2010.

[24] B. Salter, "kharj," Al Kharj: The Agri capital of saudi", vol. 15, no. 12, 2006.

[25] USGS, "Establishment of the U.S. Geological Survey," USGS Science for changing world, 2010. [Online]. Available: http://pubs.usgs.gov/circ/c1050/establish.htm. [Accessed 5 Febraury 2014].

[26] Benedicte Odden, "Comparison of a Hyperspectral Classification Method Implemented in Different Remote Sensing Software Packages", Diploma Thesis, Department of Geography . University of Zurich, Zurich, 2008.

[27] M. K. Griffin, S. M. Hsu, H.-h. K. Burke, S. M. Orloff and a. C. A. Upham, "Examples of EO-1 Hyperion Data Analysis", lincoln laboratory journal,VOL.15, no. 2, pp. 271-298, 2005.

[28] D. White, "Hyperion Tools 2.0 Installation and User Guide, 2013.

[29] Samuel Rosario Torres, "Implementation of the SVDSS in the ENVI/IDL Environment", vol. 15, no. 12, 2002.

[30] Sahar A. El_Rahman, Wateen Aliady, Nada Alrashed, "Supervised Classification Approaches to Analyze Hyperspectral Dataset", International Journal of Image, Graphics and Signal processing, IJIGSP Vol. 7, No. 5, April 2015, pp. 42-48 .

[31] Exelis Visual Information Solutions, "Products Services" , Exelis Visual Information Solutions, 2011. [Online]. available: http://www.exelisvis.com/ProductsServices/ENVI/ENVI.aspx [Accessed 10 November 2013].

[32] Exelis, "ENVI software", 2009. [Online]. Available: http://www.exelisvis.com/ProductsServices.aspx. [Accessed 10 October 2013].

[33] K Perumal and R Bhaskaran, "Supervised classification performance of multispectral images", journal of computing, volume 2, issue 2, february 2010, ISSN 2151-9617.

[34] ENVI, "ENVI Reference Guide", ENVI Version 4.7, Copyright © ITT Visual Information Solutions All Rights Reserved, August, 2009.

[35] S. P. Thenkabail, E. A. Enclona, M. Ashton and B. Van Der Meer, "Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications," in Remote Sensing of Environment, Italy, elsevier, pp. 354-379, 2004.

[36] G. Foody, "Status of land cover classification accuracy assessment," Remote Sensing of Environment, vol. 80, no. 1, pp. 185-201, 2002.

Identify and Manage the Software Requirements Volatility

Proposed Framework and CaseStudy

Khloud Abd Elwahab
MSc student /Information system Department
Faculty of computers and information, Helwan University line 3: Cairo, Egypt

Mahmoud Abd EL Latif
Associate Professor of information system department
Faculty of computers and information, Helwan University
Cairo, Egypt

Sherif Kholeif
Assistant Professor of information system
Faculty of computers and information, Helwan University
Cairo, Egypt

***Abstract*—Management of software requirements volatility through development of life cycle is a very important stage. It helps the team to control significant impact all over the project (cost, time and effort), and also it keeps the project on track, to finally satisfy the user which is the main success criteria for the software project.**

In this research paper, we have analysed the root causes of requirements volatility through a proposed framework presenting the requirements volatility causes and how to manage requirements volatility during the software development life cycle.

Our proposed framework identifies requirement error types, causes of requirements volatility and how to manage these volatilities to know the necessary changes and take the right decision according to volatility measurements (priorities, status and working hours). This framework contains four major phases (Elicitation and Analysis phase, Specification Validation phase, Requirements Volatility Causes phase and Changes Management phase). We will explain each phase in detail.

Keywords—software requirements; requirement errors; requirements volatility; reason for requirement changes and control changes

I. Introduction

The software engineering industry faces several issues; requirement changes are one of the most significant and critical issues during software development process. The project requirements are almost never stable and fixed as has been explained by [Jones (1996)]. Requirements are defined in [1] as "The information from the user about what will do and what is the main objective of this project and what is the deadline to deliver this project and so on".

Accuracy and focus through gathering requirements does not prevent requirement changes to take place during the software development life cycle. Requirements volatility is defined in [18] as “the emergence of new requirements or modification or removal of existing requirements”. Numbers of requirements change during software development process depending on the quality measures of requirements (Correct, Unambiguous, Complete, Consistent, Importance and Stability, Verifiable, Modifiable, Traceable, and Understandable) [5].

Requirements changes have a significant impact on project performance, project schedule and budget. This paper proposes a framework focus on how to manage requirements volatility during the software development life cycle and to limit the implications thereof. The remainder paper is structured as follows: **Section 2** presents the motivation to search for the requirements volatility topic. **Section 3** explains the requirements volatility definition, factors, causes, and the measures, and finally explains the impact of that on software development process. **Section 4** contains proposed framework and presents a case study in explanation of the benefits of this framework**. Section 5** concludes our results and provides some open research directions

II. Motivation

Understanding the requirements volatility and the impact thereof during the software development life cycle is a very interesting area that needs more studies to focus on how to manage these changes. Abeer. Al and Azeddine, in [18] according to previous studies suggest that 86% of the change requests are related to requirements volatility and it is often more than 50% of the requirements are changed before the delivery of a software project, furthermore implementing the requirements volatility in later phases causes 200 times costlier than implementing the requirements volatility in the analysis phase. All these facts are the motivation to start searching and try to find solutions for such critical point.

III. Research Questions

In this paper, we are going to discuss the following questions:

1) *What is Requirements volatility?*
2) *What are the causes of Requirements volatility?*
3) *What is the impact of Requirements volatility?*
4) *How to manage requirements volatility in different phases using proposed framework.*

IV. Requirements Volatility

Requirements volatility refers to additions, deletions and modifications of requirements during the system development life cycle, as defined in [4] "Stable requirements are the holy grail of software development." RV creates additional work in design and coding, which increases the system development cost and time and compromises the system quality. Ignoring requests for requirement changes can cause project failure due to user rejection, and failure to manage RV can increase the development time and cost.

A. REQUIREMENTS volatility factors

Environment changes are the main factor for requirements volatility; according to previous studies, there are Development Environment changes, and we also cannot avoid the Business Environment changes [11].

– **Business Environment changes** are, for instance, government regulations, market competition, financing sources, restrictions, management changes, organization policy, legal factors, technological factors and business laws.

– **Development Environment changes** are, for instance, requirement errors, evolving user and technological needs, new technology, missing members of team project, incorporation of cost upgrades; resolve requirement conflicts, missing requirements).

Fig.1 shows the various factors causing requirements volatility [11].

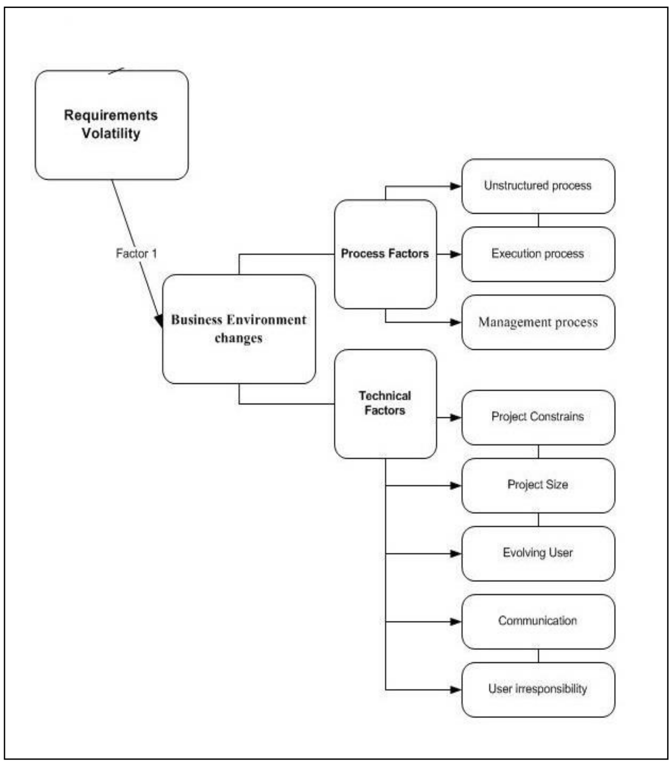


Fig. 1. (a)Requirements volatility factors [11]

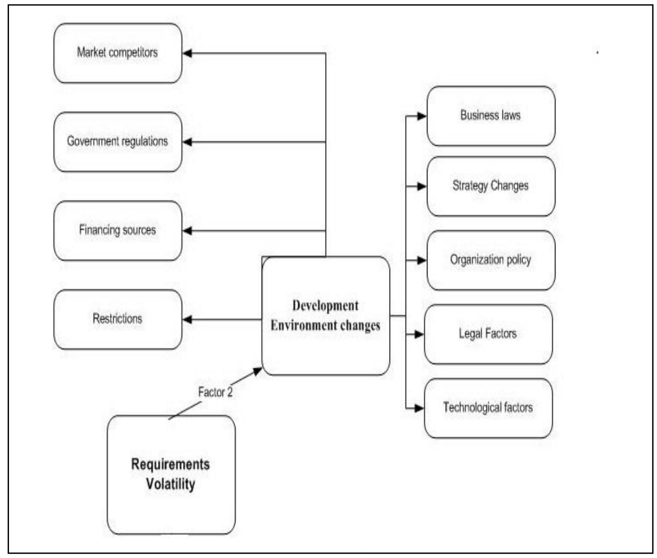


Fig. 1. (b)Requirements volatility factors [11]

B. Cause of REQUIREMENTS volatility

After considering the environment changes, we concluded that there are several causes for requirements volatility, Requirements errors are one of the main causes. For more accuracy, we classify requirement errors into three main groups: people errors, process errors and documentation errors [7], Fig. 2.

- **People Requirement Errors:**

– The communication gaps between stakeholders.

– Poor participation between the development team and management team.

– Less understanding domain knowledge, unstructured process execution.

– Users have unreasonable timelines and do not really know what they want

- **Process Requirement Errors:**

– Bad management process.

– Rush and bad analysis requirements, using old process and methodology.

– Inadequate method of achieving objectives and requirements change during the project

- **Documentation Requirement Errors:**

Team may do not understand the policy of the user's organization or no use of standards.

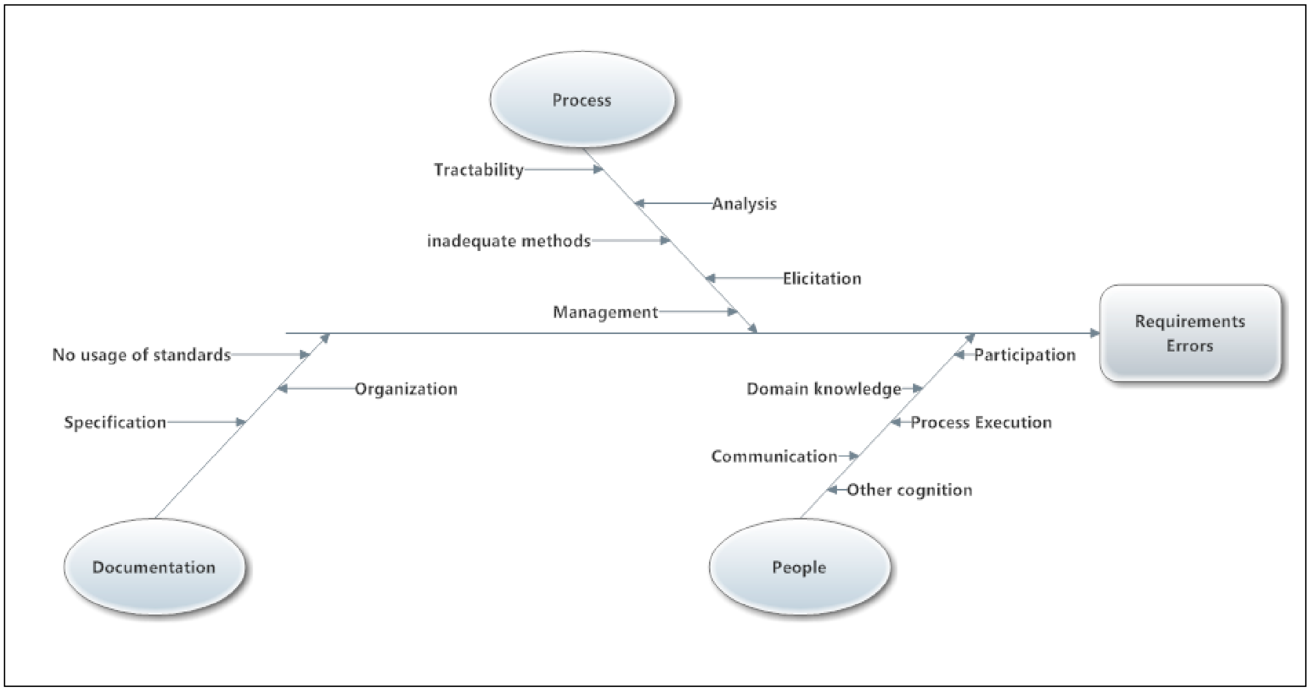


Fig. 2. Types of Requirements Errors [7]

C. Mesurements of REQUIREMENTS volatility

To avoid implementing unnecessary volatility changes, we should analyse the changes request depending on some measures; for instance, priority of changes (low, high, and medium) and the severity of changes (Critical, major, minor….), working hours needed, change types (added, deleted or modified), phase in SDLC and impact of each change. Figure-3

D. Impact of REQUIREMENTS volatility

Several research studies have found that requirements volatility is positively correlated with the increase in the size of the project, effort and cost (often >20%) and schedule duration [19].

Requirements volatility inevitably result in additional work and increased defect density, furthermore, it increased development working efforts that need a rework in code, design, and also increase team working hours and cost. Sometimes we need to reschedule the whole project, as the volatile requirements quality decreases. We need to re-design the test cases according to new requirements.

Fig.-4 presents the requirements volatilely over all (measurements, cause, and impact).

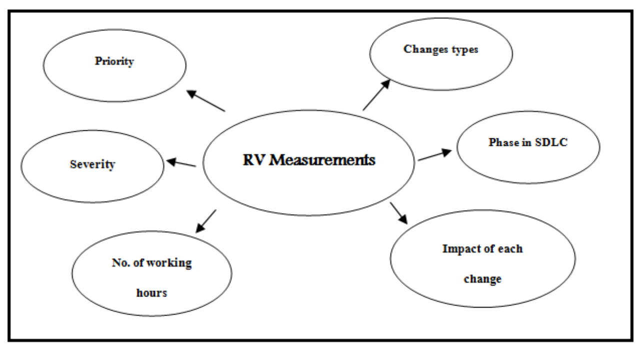


Fig. 3. Measurements of Requirements volatility

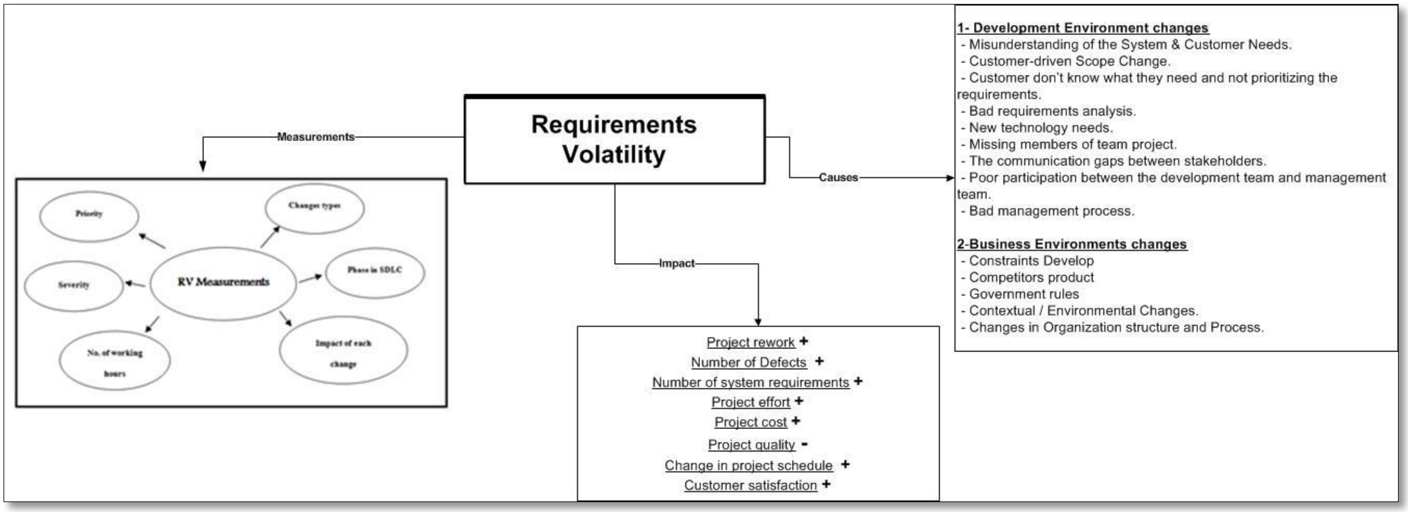


Fig. 4. Requirements volatility Causes, Impact and Measurements

V. Software Requirements Volatility Porposed Framework

After studying the causes and impact of volatility in requirements, we have proposed a framework using UML concepts for model-driven software development, which is given in Fig. 5. This framework consists of four major stages.

1) *Elicitation and analysis of business*
2) *Specification validation*
3) *Requirements volatility causes*
4) *Manage Changes phase*

The proposed framework identifies requirement error types, causes of requirements volatility and how to manage these volatilities to know the necessary changes and take the right decision according to volatility measurements (priorities, status and working hours).

The proposed framework is used to help the team to know the root cause of requirements volatility that will reduce number of changes happened in next release and other project.

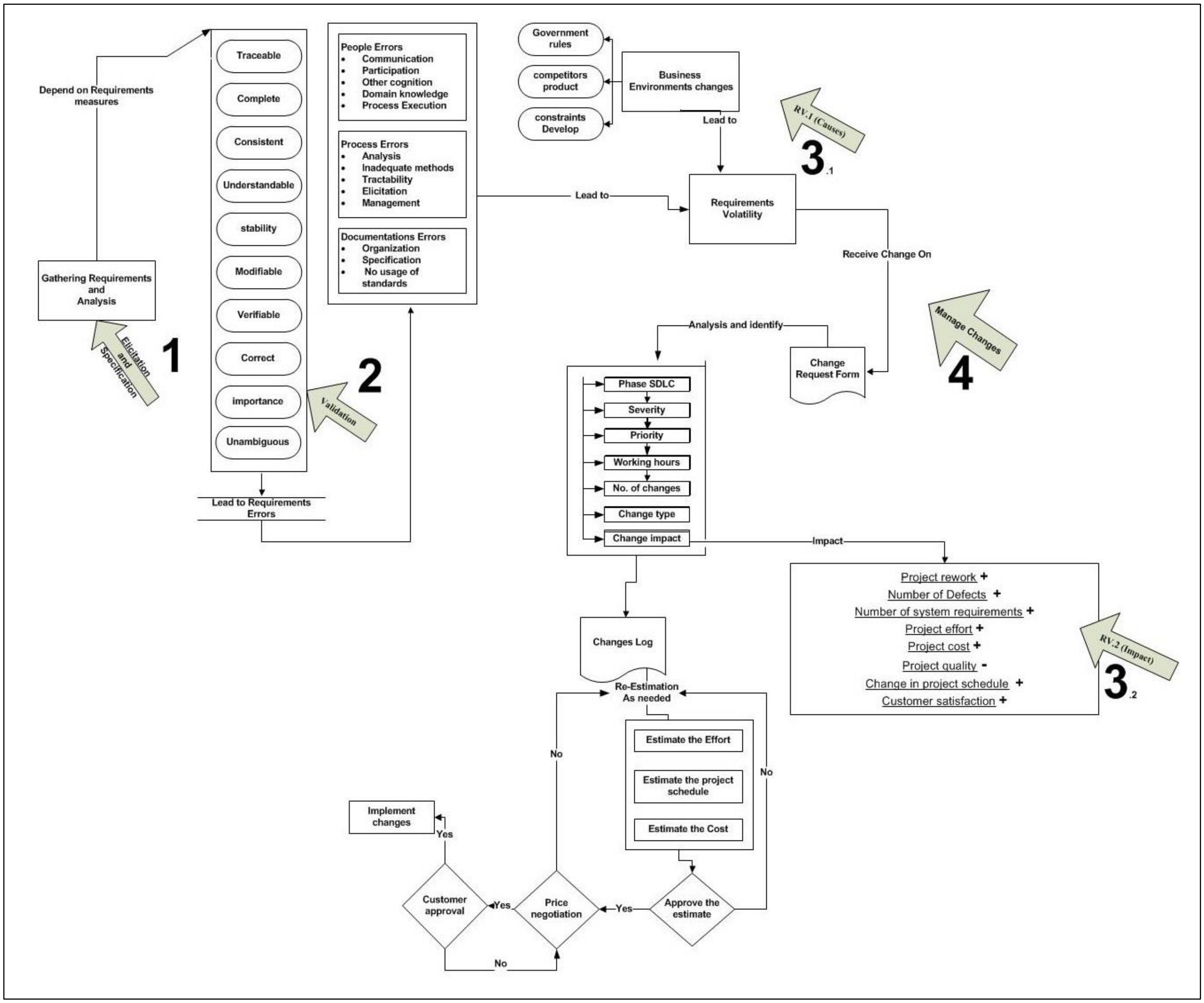


Fig. 5. Proposed framework to identify and manage the software requirements volatility

B. Elicitation and Analysis of business REQUIREMENTS:

Elicitation and analysis requirements is a critical phase in the software development life cycle, in this phase, requirements should discussed with user to make clear and then enter agreement therewith for all requirement needs. This process is to ensure that requirements are visible to and understood by all stakeholders.

There are some techniques that help in getting more accurate requirements such as: Brainstorming, Document Analysis, Focus Groups, Requirement Workshops and Interface. The results will be documented in Software Requirements Specification SRS, The SRS is full analysis and formalizing the requirements definition, what the software will do and how it will be expected to implement [16] [17]

C. SPECIFICATION VALIDATION:

Specification validation works with the final requirements document where a group of work team read and validate the requirements according to nine requirement measures "Correct, Unambiguous, Complete, Consistent, Importance and Stability, Verifiable, Modifiable, Traceable, and Understandable"[5] and user needs, and then look for errors to discuss and agree to actions to address such errors before beginning implementation to avoid volatility during software development life cycle (SDLC).

Fig. 6 shows the nine requirement measures and the classifications of requirements errors that will appear due to poor requirements gathered

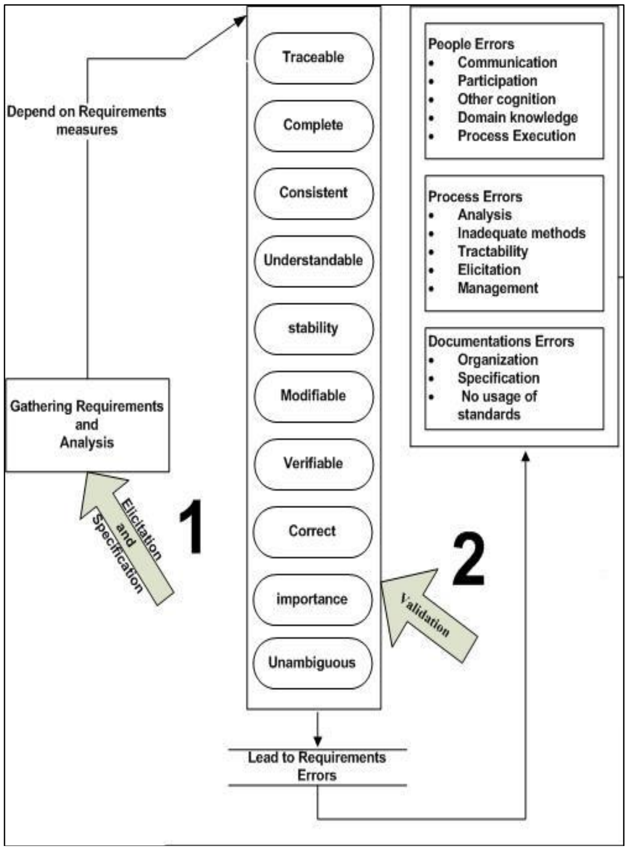


Fig. 6. Elicitation and analysis & Validation phase

D. Requirements VOLATILITY:

It is impossible to find software project without requirements volatility. We can define requirements volatility as missing requirements or misunderstanding and also gathering requirements without consideration of the nine requirement measurements to be added, deleted or modified, to keep project on track.

The development team needs to deal with changes and handle them. If the changes are not handled effectively, problems can occur needing extra efforts to manage the impact thereof on cost, quality, and schedule, as can be seen in the impact of requirements volatilely framework, both of **Project rework** + **Number of defects** + **Number of system requirements** + **Project effort** + **Project cost** + **and need to Change in project schedule.** At the end, all these changes may conclude into Customer satisfaction.

The causes of volatility can be classified into two main categories: development environment changes and business environment changes. These have already been discussed before.

Fig. 7 present causes of volatility and we are going to describe how to control the changes in next phase.

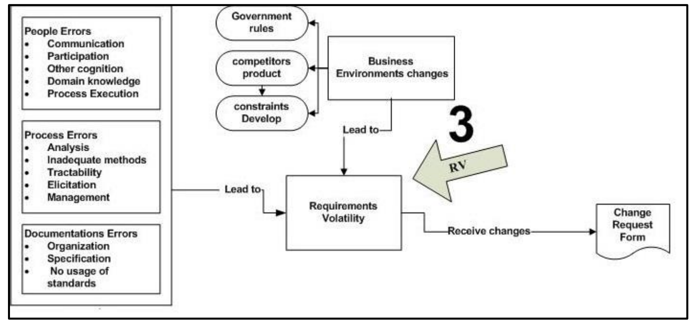


Fig. 7. RV Causes

E. Changes Management phase

It is almost certain that projects face changes during project life cycle development as explained above. These changes may help projects to cope with business needs. Change management is an important part of the project management process, thus each change should be considered carefully before approval, in order to deliver a project successfully.

The Proposed framework illustrates change control process, that each change requires a form properly defined by the user, that includes details of the change, and the business case, then the development team analysis may be considered and the changes approved prior to implementation to avoid unnecessary changes, and not to disrupt resources and delay the project delivery depending on seven measures (change impact, severity, priority, working hours, phase SDLC, change type and the number of changes), using all collected data to record all changes requested and decisions made in change log, the task group would re-estimate the project plan (effort, project schedule and project cost), developers are responsible for estimating the effort required to implement the new requirement changes which they will work on , project managers notify user and negotiate the cost of changes going to be implemented, after the change is done correctly, the case is closed in the change log.

Fig. 8 present the controlling of requirements volatilely process

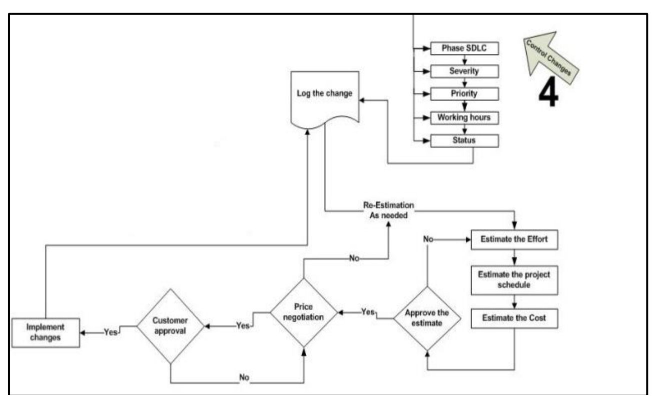


Fig. 8. Ccontrolling process of requirements volatilely

VI. Case Study

In this section, we will apply the proposed framework on ADJD system as a case study and discuss the impact of software requirement changes during the life cycle development in different project aspects (cost, schedule, and quality).

ADJD system refers to the Abu Dhabi Judiciary Department-Treasury System; the treasury system developed for the Abu Dhabi Judicial Department integrates with the existing case management system to facilitate the management of financial deposits and withdrawals associated with legal cases. By recording details of the beneficiaries and building an accounting structure, the Judicial Department is able to organize the financial processes involving receivables, check issuing and bank account management.

Basically clients are not technical people. They need software for their business but the requirements are unclear. Nobody has done an in-depth analysis of all the implications so during implantation phase client request some new features. The new features will probably break some assumptions development team made in their code and they start thinking immediately of all the things you might have to refractor, no matter how conscious you are of all these new features, you give shorter times than you originally suppose it might take. Special when you feel the pressure of deadlines and management expectations.

Measuring the Requirements volatility percent: Stark [20], derived a formula based on the statistics on different projects that:

Requirements volatility = (added + deleted + changed (modified))/ (# requirements in VCN)*100 [20].

Where VCN (version content notice) = set of requirements agreed by both the developer and customer.

In VCN 1.2 release has 55 requirements initially, later 2 new requirement are added, and 5 requirements are deleted from initial requirements, modified 11 requirements and at the end add 2 new features as business needs.

RV = (2+5+11+2)*100/55= 36.4 % of project has changes

TABLE I. Log of Changes

RV No.	Priority	Severity	Change Type	Effort *6 H (man days)
2	2	Critical	Additions	3
11	3	Major	Modifications	2
5	5	Minor	Delete	2
2	1	Critical	Add new feature	5
Total= 20				Total=12 man day * 6

- First when the proposed framework is used, it helped the team to know the root cause of requirements volatility that will reduce number of changes happened in next release and other projects.
- As is explained in case study, 36% of project requirements changed due to requirement errors that have not been discovered before development team start implementation. Proposed framework manage changes appeared during life cycle development by analysis of changes requested depending on some criteria (priority, severity, type of change, change type, impact of change and no of working hour's needs) to avoid implement unnecessary volatility change.

VII. Conclusion

In this paper we have described several aspects of requirements volatility, such as factors, causes, measurements and impact of requirement volatility on software life cycle development, and also proposed framework to manage requirements volatility.

The causes of requirements volatility cannot be overcome fully but we described some causes like (poor communication between stake holders and developers, technical aspects and bad management process, etc.)

Requirements volatility has an impact on the whole software life cycle development. It has impact on the project schedule, cost, and quality. It cannot be avoided, but we can manage it to reduce the effect of requirements volatility by following some methods in analysis, design, and coding. Due to the impact of requirements volatility many projects have failed.

The proposed framework will manage the requirement changes that help to reduce the effect of changes made all over the project. This framework contains four major phases (Elicitation and Analysis phase, specifications Validation phase, Requirements Volatility phase and Change management phase) that help project development team to know the necessary changes and take the right decision according to volatility measurements (priorities, status and working hours).

Future work: more research is needed to develop the flexible architecture which is suitable for requirements volatility. We need to define new methods to manage requirements volatility, and also more work on management of the requirements volatility, modified framework is needed by adding different software process models

References

[1] Amira A. Alshazly a, *. A. (2014). Detecting defects in software requirements specification. Alexandria Engineering Journal, 15

[2] Aybüke Aurum, C. W. (2005). Engineering and Managing Software Requirement. Germany: © Springer-Verlag Berlin Heidelberg.

[3] CS2 Software Engineering note 2. (2004, autumn 1). Software Requirements1. Software Requirements1, p. 9.

[4] Daniel D. Galorath, Galorath Incorporated. (2006). the 10 Step Software Estimation Process for Successful Software Planning, Measurement and Control. Galorath Incorporated, 13.

[5] Davis, A., Colorado Univ., C. S., Overmyer, S., Jordan, K., & Caruso, J. (1993). Identifying and measuring quality in a software requirements specification. Software Metrics Symposium, IEEE. Proceedings. First International, 12.

[6] Dhirendra Pandey1, U. S. (2011, May). A Framework for Modelling Software Requirements. IJCSI International Journal of Computer Science Issues, 8.

[7] Gursimran Singh Walia a, J. C. (2009). A systematic literature review to identify and classify software requirement errors. Information and Software Technology, 23.

[8] Kelly, A. (2008, Feb). Changing Software Development: Learning to Become Agile. Wiley.

[9] King, A. F. (2005, December). How to Detect Requirements Errors-A Guide to Slashing Software Costs and Shortening Development Time. Retrieved from ravenflow: www.ravenflow.com

[10] Loconsole, A. (2007). Definition and validation of requirements management measures. SE-90187 Umeå, Sweden, Umeå University, Thesis, 106.

[11] M.P.Singh, R. V. (2012, 9). Requirements Volatility in Software Development Process. International Journal of Soft Computing and Engineering (IJSCE), 6.

[12] N Nurmuliani, D. Z. (2007). Analysis of Requirements Volatility during Software Development Life Cycle. Australian Software Engineering Conference (ASWEC'04), 13.

[13] Sakthivel, S. (2010). Manage Requirements Volatility to Manage Risks in IS Development Projects. ISACA JOURNAL, 4.

[14] Sudhakar, M. (2005). Managing the Impact of Requirements Volatility, Master Thesis, Department of Computing Science,Umeå University,SE-90187 Umeå, Sweden, 42

[15] Awasthi, R. (2012). Development of a Structured Framework to Minimize Impact of Requirement Volatility. International Journal of Computer Applications (0975 – 8887), 7

[16] International Institute of Business Analysis (IIBA), "Business Analysis Body of Knowledge 2.0", 2009

[17] Wiegers, Karl E., "Software Requirements, 3nd Edition", Microsoft Press, 2009

[18] Abeer AlSanad and Azeddine Chikh., (2014). " The Impact of Software Requirement Change – A Review "

[19] G. Stark, P. Oman, A. Skillicorn, and R. Ameele, "An Examination of the Effects of Requirements Changes on Software Maintenance Releases" in Journal of Software Maintenance Research and Practice, Vol. 11, 1999, pp:293-309

[20] Mohd.Haleem, Mohd.Rizwan Beg, Sheikh Fahad AhmadInternational,Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, No 5, May 2013, pp: 1811-1815.

Carbon Break Even Analysis: Environmental Impact of Tablets in Higher Education

Fadi Safieddine
College of Business
The American University of the Middle East,
Kuwait

Imad Nakhoul
College of Business
The American University of the Middle East,
Kuwait

***Abstract*—With the growing pace of tablets use and the large focus it is attracting especially in higher education, this paper looks at an important aspect of tablets; their carbon footprint. Studies have suggested that tablets have positive impact on the environment; especially since tablets use less energy than laptops or desktops. Recent manufacturers' reports on the carbon footprint of tablets have revealed that a significant portion, as much as 80%, of the carbon footprint of tablets comes from production and delivery as opposed to the operational life-cycle of these devices. Thus rending some of previous assumptions about the environmental impact of tablets questionable. This study sets to answer a key question: What is the break-even analysis point when saving on printed paper offsets the carbon footprint of producing and running the tablet in higher education. A review of the literature indicated several examples of tablet models and their carbon emission impact; this is compared to the environmental savings on paper that green courses could produce. The analysis of the carbon break-even point shows that even when considering some of the most efficient and least carbon impact tablets available on the market with a carbon-footprint production of 153Kg CO_2e, the break-even point is 81.5 months; referring to 6 years, 9 months and 15 days of use. This exceeds the life-cycle of an average tablet of five years and average degree duration of four years. While tablets still have the least carbon-footprint impact compared to laptops and desktops, to achieve the break-even point of carbon neutral operations this study concludes that manufacturers need to find more environmentally efficient ways of production that would reduce the carbon-footprint product to a level that does not exceed 112.8kg CO_2e.**

Keywords*—*Environmental, Tablet; Higher Education; Carbon-footprint; Break-even Analysis

I. Introduction

The tablet industry is comparable to the mobile device industry in its growing pace and in its increased usage. Tablets are becoming a part of higher education delivery approach and is considered to turn into habitually usage in higher education replacing the longstanding method of course delivery: papers.

While the demand is growing and many advantages highlighted in previous publications about the usage of tablets in higher education [1], assumptions are made regarding the environmental impact of tablets and specifically their impact in reducing energy use and reducing paper usage.

In this paper, the team have set out to design a framework to better evaluate the environmental impact of tablets or computing devices in higher education thus setting the bar for future computing devices environmental impact studies. When justifying a technology as reducing environmental impact, the team suggest to look at the carbon foot print of the production and overtime use of such devices, compare it to the reduction in emission it help reduce, while considering the average life-cycle of these devices to appreciate the net impact. In doing so, the team conducted a literature review collecting valuable information that helped identify the breakeven point for tablets.

II. Literature Review

A. Tablets and higher education:

There are significant evidence in the literature regarding the positive financial impact for organizations to go paperless in their operations [1], [2], [3], & [4]. There are also several publication on how academics have been successful in going paperless in higher education [5-9]. One of the key aspects that have helped fan the move to paperless classrooms or totally green courses is the use of Virtual Learning Environment (VLE). In fact, three studies have reviewed the impact of switching courses from paper based to complete paperless on students' performance and in both cases the results have been favorable [10], [11], & [12].

The use of tablets in higher education has also been subject of several studies. A survey of higher education libraries in the U.S. has shown an increase use and dependency on educational technologies such as tablets [13]. In two separate studies, researchers have been able to prove that learning from electronic tablet screens does not affect the efficiency of learning compared to printed publications [14], [15]. In one of the largest study conducted on the use of tablets in higher education involving a sample of 280 students from several universities, researchers suggested the use of tablet promoted an active-learning environment with positive learning experiences from across the board [16]. In fact, increasingly more research into tablets in higher education is showing high acceptability rate [17 - 19], learners increase dependency on M-learning [20] & [21], and successful implementations of such devices [16] & [22]. Portability, mobility, and the longer battery are suggested as key success factor of tablets [18].

The underlining perception for many of the studies is that encouraging the use of tablets in higher education has positive environmental impact [1], [10], [13], [23], & [24]. And while this may well be the case when compared to other computing devices, the real data on this is still rather missing.

B. Environmental Impact of tablets

An independent research think-tank [25] suggested in 2012 that tablets produce less carbon emissions during their production and operation lifecycle compared to desktops and laptops. The study focused on a review of Apple Product Environmental Report [26]. In comparing the production footprint of two iPad tablets, two Mac laptops, and two Mac desktops, the research think-tank concluded that Apple iPad tablets carbon footprints are usually less than 10% compared to a Mac Pro Desktop while iPad tablets carbon footprint stood less than 30% of that of the average Mac laptop. The results looked further in favor of tablets when looking at the operational impact. A 10,000 hours average use of each device comparison showed that the average iPad tablets used as little as 1% electricity when compared to a Mac desktops and between 3 to 30% of that of comparable Mac laptops At least in the case of the Apple devices, the results seem to be conclusively in favor for a move to tablets as oppose to any other device.

Apple Inc. provides an interesting analysis that sheds light to the main source of emissions in the lifetime of a tablet. According to the Environmental Report for iPad Air 2 [27] the total carbon emission of the tablet in its lifetime is estimated to be 170Kg CO_2e in greenhouse gas emissions. However, the customer usage only amounts to 10% of that figure. 86% of the carbon emissions of an iPad Air 2 comes from the production of the tablet itself see Fig. 1.

The breakdown of carbon emission for the tablet is shown in Table 1.

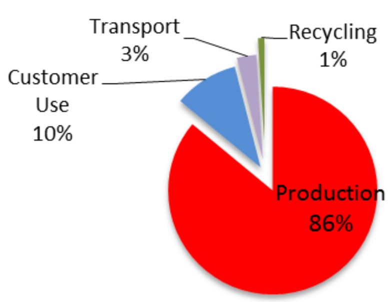


Fig. 1. Greenhouse Gas Emissions for iPad 2 (Wi-Fi + Cellular) [28]

TABLE I. BREAKDOWN OF THE CARBON-EMISSION

iPad Air2	Percentage	170 kg CO_2e
Production	86%	146.2 kg CO_2e
Customer Use	10%	14.62 kg CO_2e
Transport	3%	0.4386 kg CO_2e
Recycling	1%	0.004386 kg CO_2e

While the European Union Energy Efficiency Star rating of electronic products has focused on the energy usage during operation of a device [28], the literature review regarding tablets suggests that the rating of tablets' environmental impact purely on operational usage is misleading consumers.

Of course not all tablets rate the same. An internal comparative study of different means of studying by Wageningen UR University in the Netherlands reviewed several tablets available on the market and their environmental impact as opposed to buying books [29]. The study looked specifically at students' use and higher education. The study concludes that tablets rate better at environmental impact when compared to laptops and significantly better when compared to desktops.

The only other comparative study the team was able to find is that of an Australian independent consumer services [30]. In their consumer report, the authors suggests that Apple iPad do rate the highest for energy efficiency and lowest in carbon footprint while Toshiba and Google tablets ranked the least efficient. However, no data was provided that would justify the rating or ranking of these tablets. The ranking is not linked to specific tablets. What is more, the ranking doesn't focus merely on the environment impact, it mixes environmental and social impacts.

C. Paper Use in Higher Education

There is limited literature on the use of paper in higher education. The team will be relying in this instance on a study conducted at University of East London between 2006 and 2013 as the authors recorded the switch from paper based to completely Green courses [10]. In a follow up study, the team looked at specific use of paper, paper printing, and books part of cost-benefit analysis for running Green courses and their impact on the university, students, and faculty [1]. The following is an extract of that data quantifying the average use of papers per students for every course in each semester that the team will abbreviate as APUSCS (Average Paper Used per Student per Course per Semester).

University:

- Course Syllabus/Handbook: 37 APUSCS
- Handouts: 24 APUSCS
- Exam & Quizzes: 28 APUSCS

Student:

- Course work printing: 137.5 APUSCS
- Course book: One textbook per course and semester.

The total paper saving is 226.5 pages plus one textbook per student, per course, per semester.

D. Environmental Impact of Printed Papers

In the case of course books, the team have found one study by the US government Environmental Protection Agency department that suggested the carbon emission of printing an average textbook to be 2.4kg CO_2e [31]. And since almost all courses require one core textbook, the team have assumed the carbon-emission to be 2.4kg CO_2e per course. The same report looked at the environmental impact of using and printing papers with the study suggesting that 100 pages of papers generate 1.26kg CO_2e. This number has been cross checked with another two studies in which the team reached very similar figures of 1.251kg CO_2e for 100 printed papers [32] and consistent with another study in which carbon emission of 100 plain papers are calculated to be around 0.938Kg CO_2e [33].

III. Research Questions and Methodology

A. Research Questions:

The team set out to determine a key fact when it comes to evaluating the real environmental impact of tablets and to do so the team aimed to answer these questions:

1) What is the breakeven point when the use of a tablet becomes carbon neutral in Higher Education?

2) What is the carbon-emission target that companies need to aim at when designing greener tablets for higher education institutions?

B. Methodology:

From the literature review, the team has been able to determine the importance of tablets in the future of academia. The team has also been able to determine key details about the carbon-emission based on some key facts published by renounced manufacturers. Two important and distinctive variables have been identified: fixed carbon-emission of tablets representing the manufacturing, delivery, and recycling carbon footprint; and variable carbon-emission of tablets representing the impact of running the tablet. Unfortunately, there is one company who has provided those two details. Thus the team has decided to use the Apple iPad 2 carbon-emission data after careful consideration of the characteristics of that tablet. The tablet represents a well-known brand with some evidence of it being an energy efficient device.

On the other hand, the team has also been able to determine reasonable value to consider as average use of papers by students per course per semester. Finally, the team has been able to determine the carbon emission that would be fair means to judge the carbon impact of paper printing in academia. The printing of paper is a running variable that would in theory continue until a carbon neutrality impact is achieved. However, there will one important factor that the team will consider and is represented in the fact that an average tablet has a technological life-cycle that is unlikely to exceed five years. The aggregation of all these facts, fixed and variable carbon emission impact from tablets against the carbon emission saving from not printing are plotted on the break-even analysis diagram to determine the breakeven point in months. This would answer research question 1 and 2. Where the diagram exceeds the five-year life cycle of tablets, the team will perform breakeven sensitivity point analysis to determine what target should be for manufacturers of tablets in developing greener tablets.

IV. Results and Analysis

Case 1: Based on the literature review findings, the team has calculated the breakeven analysis where four factors have been considered:

Fixed Carbon-emission of a tablet for the manufacturing, delivery, and recycling to be 153kg CO_2e.Variable Carbon-emission of the tablet to be 17kg over 5 years or 0.2833kg CO_2e per month.

Fixed Carbon-saving from papers per student: 0kg since all the savings are variable.

Variable Carbon-saving from papers per student per course per semester to be: 226.5 papers or 2.85kg CO_2e. This is added to 2.4kg for each textbook bringing the total to: 5.2539 kg CO_2e per course per semester. Student take on average three course for two semester, so the total annual running carbon saving is 5.2439 x 3 course x 2 semester to be 31.5234kg CO_2e. Divided over 12-month period, the variable Carbon-saving from per student per month is 2.62695kg CO_2e.

When plotting these results onto a breakeven analysis diagram, the results show a breakeven point in 81.5 months, or in other words 6 years, 9 months and half. This is shown in Fig. 2.

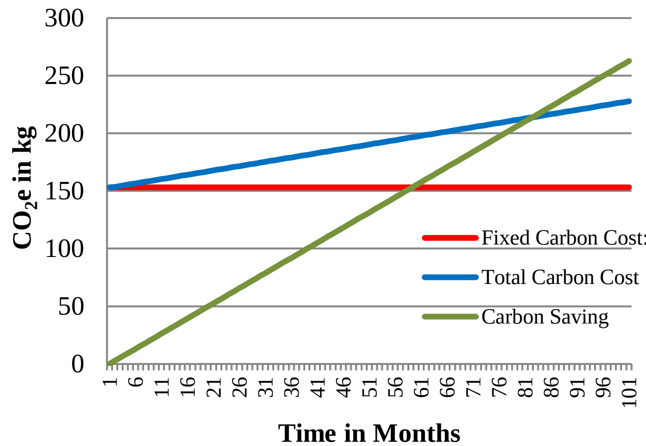


Fig. 2. Breakeven analysis for Case 1

This breakeven point exceeds the life cycle of the tablet and exceeds the length of most degree programmes, which is four years in the UK.

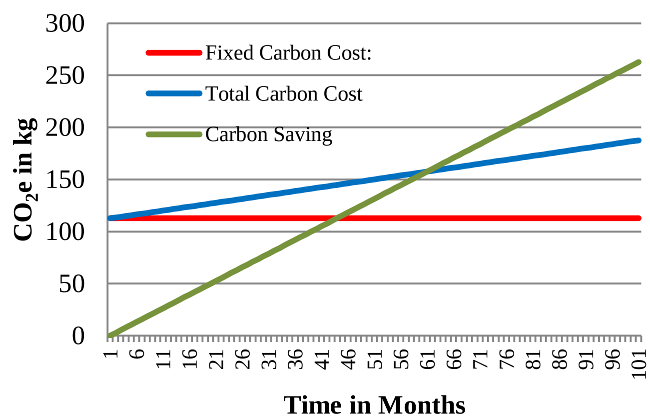


Fig. 3. Breakeven sensitivity analysis for Case 2

Case 2: The team then looked at the breakeven sensitivity analysis to determine what manufacturers and academics need to consider when looking for greener and carbon-neutral tablets and to achieve the life-cycle of five years for a tablet, considering in this case the production lines could be made to run more efficiently. The results showed that manufacturers need to produce, deliver, and recycle tablets at a rate that does not exceed 112.8kg CO_2e in which the new breakeven point would be in year 5 at 158 kg CO_2e see Fig. 3.

The evidence shows conclusively that the largest contribution to greenhouse gases is the production of the tablet. The same could be said regarding other devices. In academic environment where paper usage is considered high, the tablet failed to break-even in a considerable time. For tablets to be considered a product of neutral or positive impact on the environment, manufacturers need to consider several factors that could help achieve a greener status for higher education institution. For production, manufacturers would need to consider ways to reduce energy use, relying more on renewable energy sources. The supply chain network involved in the production of the material for the production of these tablets could be another area that could be optimized. Finally transport of these devices to the consumers

V. Research Limitations

Ascertaining what the carbon-footprint for an average tablet proved to be illusive, mainly because manufacturers are not required to provide this information. Instead the focus is on the energy usage during the operation of tablets. The team determination to select one model as an example is a research limitation and where possible the carbon-emission breakeven analysis could be applied for different tablet models.

Another concern comes from the fact that there are no academic publications on the actual carbon-footprint for the production of papers. Moreover, on the carbon-footprint of printed papers the team resorted to reviewing variety of figures from governmental and environmental agencies to reach a number that the team has confidence in being representative of the actual carbon-footprint of printed papers.

Finally, the study looks at only one aspect of the carbon saving a tablet can produce and that is for papers. There may be other aspects of saving on transportation where tablets can reduce transport of paper, books, and travel. Tablets could also contribute to reduction in use and need for labs in higher education; where students replace use of desktop computers for more energy efficient tablets. This may contribute to reducing investment by higher education institutions in ICT labs.

Despite these limitations, the team is confident that the data presented regarding the tablet carbon-footprint and paper saving footprint are reasonable. And while there may well be other factors that could be considered in the carbon saving associated with use of tablets, paper saving will remain the most significant and until further research reveals otherwise.

VI. Conclusion

Evidently, manufacturers have focused far and long on getting more energy efficient electronics to meet the demands of consumers seeking to reduce their energy bills. The running carbon-footprint of tablets could be seen as evidence of this trend. The focus could now be shifted towards the production and supply chain process as source of providing truly greener products.

More is needed to learn about the carbon-footprint of other computing devices. The industry should be seeking to not only produce products that reduce their running carbon-footprint but that surpasses in their life-cycle the carbon-footprint generated from their production. This study could be replicated as new devices are produced and new data is provided regarding these products carbon-footprint. This study could be replicated for other service and manufacturing organizations looking to reduce their carbon footprint. Presenting the carbon-footprint of a products as merely the running carbon footprint should be deemed as misleading consumers. Governments and energy rating agencies around the world need to enforce and enhance such criteria to allow better comparison of products by consumers.

References

[1] I. Nakhoul, and F. Safieddine, "Quantitative cost-benefit analysis of green courses: case study," International Technology, Education and Development Conference Proceedings, Valencia, Spain, 2014, pp. 3476-3482

[2] J.T. Ryan, "Document-management systems offer efficiency, save paper," Central Penn Business Journal, 2008, Vol 1(2).

[3] T. Thompson, "Less paper trumps paperless," Health Management Technology, 2008, pp. 42-43.

[4] F. M. Jurgens, "The paperless classroom goes to sea," Sea Power, 2000, Vol 43(2). Retrieved November 2, 2008 from ProQuest: ISSN 01991337.

[5] J. Slowinski, "Flaunt IT: Construction of a Paperless Classroom," in L. Mealy and B. Loller (eds) e-learning: Expanding the Training Classroom through Technology, 2000, pp. 117-127.

[6] A. Rea, D. White, R. McHaney, and C. Sanchez, "Pedagogical Methodology in Virtual Courses," in A. Aggarwal (ed.) Web-based learning and teaching technologies: opportunities and challenges, 2000, pp. 138-139.

[7] K. D. Lutes, and A. Harriger, "Essignments – A step toward the paperless classroom," Hawaii International Conference on Education, 2003.

[8] B. Meyer, "The Process of Implementing a Paperless Classroom in Teacher Education Using an Electronic Portfolio System. MountainRise," the International Journal of the Scholarship of Teaching and Learning, 2008.

[9] J. Arney, I. Jones, and A. Wolf, "Going green: paperless technology and feedback from the classroom," Journal of Sustainability and Green Business, 2010.

[10] F. Safieddine, and S. Wee Lee, "Green modules for sustainability in higher education: A longitude study on impact on students," International Technology, Education and Development Conference Proceedings. Valencia, Spain, 2013.

[11] J. Fei Wang, "Creating a Paperless Classroom with the Best of Two Worlds," Journal of Instructional Pedagogies. 2, 2010.

[12] S. De Bonis, and N. De Bonis, "Going Green: Managing a Paperless Classroom," US-China Education Review A 1 , 2011, 83-87. ISSN 1548-6613

[13] E. D. Cassidy, A. Colmenares, G. Jones, T. Manolovitz, L. Shen, and S. Vieira, "Higher Education and Emerging Technologies: Shifting Trends in Student Usage," Journal of Academic Librarianship, Mar 2014, Vol. 40 Issue 2, p124-133.

[14] L. Bayliss, C. Connel, and W. Farmer, "Effects of ebook readers and tablet computers on reading comprehension," Journal of Instructional Media, 2012, Vol 39(2), pp. 131-140.

[15] H. Dundar, and M. Akcayir, "Tablet vs. paper: The effect on learners' reading performance," International Electronic Journal of Elementary Education, 2012, Vol 4(3), pp. 441-450.

[16] M. Stickel, S. Edward, and Sr. Rogers, "Impact of lecturing with the tablet PC on students of different learning styles," Frontiers in Education Conference, 2009, FIE '09. 39th IEEE. pp. 1 – 6.

[17] T.J. Perez Decano, "Willingness of students to use tablets as a learning tool," Educational Media (ICEM), IEEE 63rd Annual Conference International Council for Education and Media, 2013, pp.1-9.

[18] E.M. Maina, R.W. Njoroge, P.W. Waiganjo, and R. Gitonga, "Use of tablets in blended learning: A case study of an Institution of Higher Learning in Kenya," IST-Africa Conference 2015 Proceedings, 2015, pp. 1-8.

[19] M. Moran, M. Hawkes, and O. Gayar, "Tablet Personal Computer Integration in Higher Education: Applying the Unified Theory of Acceptance and Use Technology Model to Understand Supporting Factors," Journal of Educational Computing Research, 2010, Vol. 42(1) p.p.79-101.

[20] Y. Park, "A pedagogical framework for mobile learning: Categorizing educational applications of mobile technologies into four types," The International Review of Research in Open and Distributed Learning. 2011, Vol 12.2. pp. 78-102.

[21] M. Sharples, J. Taylor, and G. Vavolva, "A Theory of Learning for the Mobile Age," Medienbildung in neuen Kulturräumen, 2007, pp 87-99

[22] J.C. Santamartaa, L.E. Hernández-Gutiérrezb, R. Tomásc, M. Canoc, J. Rodríguez-Martínd, and M.P. Arraizae, "Use of Tablet Pcs in Higher Education: A new Strategy for Training Engineers in European Bachelors and Masters Programs." Procedia - Social and Behavioral Sciences, 2 June 2015, Vol(191), p.p 2753-2757

[23] J. Arney, I. Jones, and A. Wolf, "Going green: paperless technology and feedback from the classroom," Journal of Sustainability and Green Business, 2010.

[24] S. De Bonis, and N. De Bonis, "Going Green: Managing a Paperless Classroom," US-China Education Review, 2011, Vol 1, 83-87. ISSN 1548-6613

[25] L. Wilson, "iPad or book? Is it time to go tablet for the planet?" Shrink Think-Tank, 2012, Last accessed (March 26th 2016) at http://shrinkthatfootprint.com/ipad-mini

[26] Apple Inc. "Product Report: Measuring Performance One Product at a Time," n.d., Last accessed (March, 26th 2016) at http://www.apple.com/environment/reports/

[27] Apple Inc. "iPad Air 2. Environmental Report," 2016, Last accessed (March 26, 2016) at http://images.apple.com/environment/pdf/products/ipad/iPadAir2_PER_oct2014.pdf

[28] EU Energy Star. "Energy Calculator for PC Equipment.", 2016, Last accessed (26th March 2016). https://www.eu-energystar.org/

[29] Green Office Wageningen. "CO2 footprints of Kindle vs iPad vs Books," Wageningen UR University Internal Press, 2014, Last accessed (March 26th 2016) at https://gowageningen.files.wordpress.com/2014/04/co2-footprints-of-kindle-vs-ipad-vs-books.pdf

[30] Choice. "Why every tablet has a footprint: Tablet manufacturers' environmental and social policies compared," n.d., Last access (March 26, 2016) at https://www.choice.com.au/electronics-and-technology/tablets-and-personal-media-devices/tablets/articles/tablet-environmental-footprint-report

[31] US Environmental Protection Agency. "Solid Waste Management and Greenhouse Gases A Life-Cycle Assessment of Emissions and Sinks-3rd edition," 2006, Chapter 2. Column g. p. 24.

[32] Environmental Paper Network (EPN). "Lifecycle Environmental Impact: Online Calculator," 2014, Last accessed 25th of April 2016 at http://c.environmentalpaper.org/baseline

[33] S.S. Muthu, "The Carbon Footprint Handbook," 2015, CRC Press. London. p.497.

Performance Analysis of Enhanced Interior Gateway Routing Protocol (EIGRP) Over Open Shortest Path First (OSPF) Protocol with Opnet

Anibrika Bright Selorm Kodzo, Koforidua[1]
Polytechnic (Computer Science Dept.),
Christ Apostolic University College, Kumasi
Box 981 KF Koforidua, E/R, Ghana, W.A

Mustapha Adamu Mohammed[2]
Departemnt of Computer Science, Koforidua, Ghana, West Africa

Ashigbi Franklin Degadzor[3]
Koforidua Polytechnic (Computer Sci. Dept.)
Box 981 KF Koforidua, E/R, Ghana, W.A

Dr. Michael Asante[4],
Computer Science Dept., KNUST, Kumasi, Ghana, W/A
Contact: +233 208168613

***Abstract*—Due to the increase in the easy accessibility of computers and mobile phones alike, routing has become indispensable in deciding how computes communicate especially modern computer communication networks. This paper presents performance analysis between EIGRP and OSPFP for real time applications using Optimized Network Engineering Tool (OPNET). In order to evaluate OSPF and EIGRP's performance, three network models were designed where 1st, 2nd and 3rd network models are configured respectively with OSPF, EIGRP and a combination of EIGRP and OSPF. Evaluation of the proposed routing protocols was performed based on quantitative metrics such as Convergence Time, Jitter, End-to-End delay, Throughput and Packet Loss through the simulated network models. The evaluation results showed that EIGRP protocol provides a better performance than OSPF routing protocol for real time applications. By examining the results (convergence times in particular), the results of simulating the various scenarios identified the routing protocol with the best performance for a large, realistic and scalable network.**

***Keywords*—Routing; Protocol; Algorithm; Throughput**

I. Introduction

The advancement in data communication technology facilitates users have easy access to services that enable users to use computers and mobile phones. Some of these services include file sharing through Bluetooth, print sharing, video streaming and voice conferencing services. The internet has created interconnected computer networks called the virtual underpinned by routing protocols. Currently the internet is playing a vital role in the life of communication networks. Data communication networks are solely based on technologies that provide the technical infrastructure base, where routing protocols transmit packets across the Internet. These routing protocols specify how routers communicate with each other by broadcasting messages. Also these routers update their routing tables based on prior knowledge of the adjacent networks that normally helps them in selecting the best routes possible between nodes that available on the network. These routing protocol differ in various like convergence, throughput, jitter delay and rout establishment

II. Related Work

Many researchers in the past have compared the performance of these two dynamic routing protocols that is Interior Gateway Routing Protocol (EIGRP) and Open Shortest path First (OSPF), based on dissimilar parameters used the analysis. (Ittiphon *et al*, 2005), showed the link recovery comparison that existed between OSPF & EIGRP and concluded based on the transmission time EIGRP is better choice than OSPF protocol whereas rerouting time after failure of a link also remains the same. (Shafiul *et al*, 2008), in his work explained his work on performance analysis of both EIGRP and OSPF routing protocols for real time applications including video streaming on wired and wireless networks and devices. The evaluation of these protocols based certain quantitative metrics such as Convergence Duration ,Packet Delay Variation, End to End Delay and above Throughput (Success rate of data transmitted), resulted in EIGRP performing far better than OSPF for real time video streaming and applications. Again (Sheela and Thorenoor, 2001), presented some implementation decisions on protocols that involved either distance vector protocols or link state protocols or even both and compared these protocols using different parameters. Finally it has proven from the results shown that EIGRP utilizes a far better network convergence time, with less bandwidth requirements and as well as efficient CPU and memory utilization when it is compared with other routing protocols like Open Shortest Path First Protocol (OSPFP) which is a link state routing protocol. This paper tested the two protocols on the basis of E-mail upload response time and Hypertext Transfer Protocol (HTTP) page response time, for different number of workstations (Holmes et al, 2002).

A. *Objectives*

- To simulate OSPF protocol and EIGRP protocol using OPNET based on two quantitative metrics (Throughput and Packet Delay Variation).
- Analyze the results of simulation.

- To determine a suitable and appropriate protocol for a scalable network.

B. Background Theory

In IP networks, a routing protocol usually carries packets by transferring them between different nodes. When considering a network, routing takes place hop by hop. Routing protocols have the following objectives:

- To establish communication among routers
- To construct routing tables based on routing loops
- To make routing decisions
- To learn existing and alternate routes
- To distributed information amongst autonomous neighbouring routers.

Routers perform routing by interconnecting several autonomous networks and routing packets through alternate routes and forwarding packets to different several networks based routing algorithms. The cardinal function behind routing protocols is designed to establish the best and alternate path from the source router to the destination router. A routing algorithm operate by employs several metrics, which are employed to resolve the best route that can be used to get to a network in which case this can be achieved through the use of a single or several properties of the path. For conventional routing protocols, networks are classified as Link State Routing Protocols and Distance Vector Routing Protocols. The conventional routing protocol is usually used for other types of communication networks such as Wireless Ad-Hoc Networks, Wireless Mesh Networks etc (Billings et al, 2002). Neighbor Discovery occurs by sending HELLO packets at intervals with a comparatively low overhead. After receiving a HELLO packet from its neighbors, the router ensures that its neighboring routers are active and that exchange of routing information will be possible. In the determination of the best path for transmission some specific metrics such as speed, node delay, congestion, and interference were used. OSPF is a type of routing algorithm that uses bandwidth as a routing metric while RIP (Routing Information Protocol) employs hop count whereas EIGRP uses a combination of bandwidth and delay as routing metrics.

C. Metric Parameters

In the case routing metric is measured in a manner to select the best and alternate routes as a means of ranking the routing protocol from most preferred to least preferred. In the case of routing, different metrics were employed for the purpose of different routing protocols. In the Internet Protocol routing, (Internet Protocol) routing protocols, below are some of the following routing metrics are used mostly:

- **Hop count:** It is used to determine the number of routers that are allowed to traverse the best route in a network in order to reach the desired destination.
- **Bandwidth:** Also a bandwidth metric is used to determine its routing path based on the best bandwidth speed possible.
- **Delay:** Delay is a measurement metric that specifies time for a packet to pass through a path. Delay depends on some factors, such as link bandwidth, utilization, physical distance travelled and port queues.
- **Cost:** It is the duty of the network administrator or Internet Operating System (IOS) engineer to determine the cost by specifying the best and alternate route to a destination. The cost of the routing metric can be used to represented either as a metric or a combination of metrics.
- **Load:** It is described as the traffic utilization of a defined link. The routing protocol use load in the calculation of a best route.
- **Reliability:** Reliability is used to determine the efficiency of the network as well as, it calculates the link failure probability and it can be calculated from earlier failures or interface error count (Douglas et al, 2006).

D. Routing Methodologies

A router is responsible for accomplishing the following procedure:

- Router are able to learn about directly connected networks and its own links.
- A router must have a connection with its directly connected and adjacent networks and this performed by the help of **HELLO** packet exchanges.
- Routers must send what are called a link state packet which contains the state of the available links.
- A router is able to stores a link state packet copy which is received by its neighbouring routers.
- A router must also independently establish the least cost path for the topology as proposed by (Lammle et al., 2005).

III. METHODOLOGY

During the implementation of a real world model of the simulation system that is designed by OPNET,a suitable algorithm was also adhered following the design using the packet simulator. Figure 3.1 shows a flow chart of the steps.

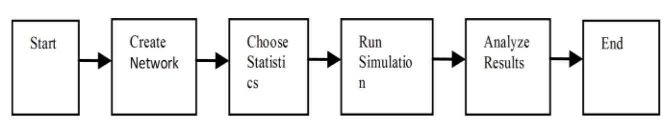


Fig. 1. Simulation in OPNET flow chart and Design Steps

OPNET Simulator

OPNET is a type of network simulator that stands for Optimized Network Engineering Tools modeler 14.0 was would be used as the network simulation environment. OPNET is a simulator built on top of Discrete Event System (DES) and also it is used to simulates the system functional characteristics and behavior often by modeling each event and process in the system by the help of user defined functionalities. OPNET is also suitable for the simulation of heterogeneous network coupled varying protocols.

The Anatomy of OPNET (Simulation Software)

OPNET is a high level discrete event simulator with striking user interface that is was developed by the C and C + + programming languages with their source codes.

A. Hierarchical Structure of OPNET Model

The OPNET simulator has three major functional models. These are:

B. Network Domain Model

The Network Domain Model has three sections. These are: Physical connection, interconnection and configuration. It is meant represent all system attributes like as network, sub-network on the geographical map to be simulated.

C. Node Domain Model

The Node domain is used to constitute all internal infrastructure of the network domain. Nodes can be routers, workstations, satellite as well

D. Process Domain Model

The Process domain are used to normally specify the intrinsic attributes of the processor and queue models by the use of use of source code C and C ++ libraries which is inside the node models as indicated above.

E. Measurements Characteristics

This section actually talks about measurement specifications, measurements that relate to the performance metrics specifically Throughput and Packet Delay Variation are done from the acquired results of Discrete Event Simulation in figure 3.6.Detailed information about the simulation and measurements are explained further below based on the various models created.

Network Topology under simulation

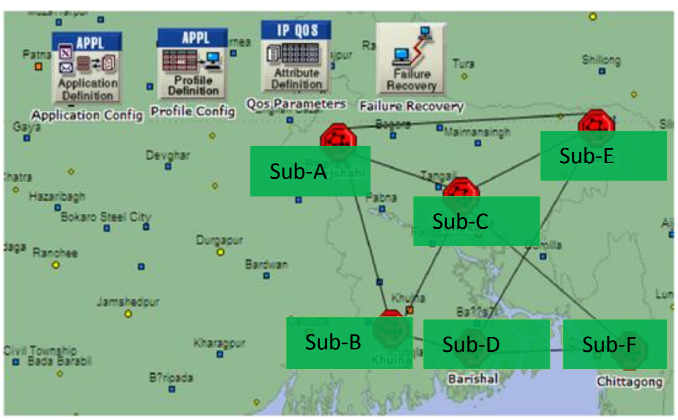


Fig. 2. The proposed network under simulation

In this thesis, three scenarios EIGRP, OSPF and EIGRP_OSPF were created that consists of six interconnected subnets where routers within each subnet are configured by using EIGRP and OSPF routing protocols. The network topology composed of the following network devices and configuration utilities:

- Switches
- CS_7200 Cisco Routers
- Ethernet Server
- PPP_DS3 Duplex Link
- PPP_DS1 Duplex Link
- Ethernet 10 BaseT Duplex Link
- Ethernet Workstation
- Six Subnets
- Application Configuration
- Profile Configuration
- Failure Recovery Configuration
- QoS (Quality of Service) Attribute Configuration

The network topology designed using OPNET as shown in figure3.6. Six subnets that are interconnected to each other were considered. All of the subnets contain routers, switches and workstations. An Application Definition Object and a Profile Definition Object have all been named correspondingly the figure 3.6. Application Config and Profile Config in the figure 3.6 are added from the object palette into the workspace. The Application Config allows generating different types of application traffic. As far as real time applications are concerned in this thesis, the Application Definition Object is set to support Video Streaming (Light) and Voice Conferencing. A Profile Definition Object defines the profiles within the defined application traffic of the Application Definition Objects. Weighted Fair Queuing (WFQ) is a scheduling technique that allows different scheduling priorities on the basis of Type of Service (ToS) and Differentiated Service Code Point (DSCP).The routers are connected using PPP_DS3 duplex link with each other. The switches are connected to routers using same duplex link. Ethernet workstations are connected to switch using 10 Base T duplex links and also links speeds of 44.76 Mbps for the first set of subnet connection with link type of PPP_DS3 and 1.544 Mbps for the second set of subnet connection with a link type of PPP_DS1 deployment to ensure standard data transmission across the links. The same numbers of bits were sent simulated for the various scenarios (EIGRP, OSPF, and EIGRP_OSPF). In this simulation three network models were created, simulated and measurements were carried out based on two performance metrics that is ***Throughput and Packet Delay Variation.***

"Three network models were simulated, which are configured and run as 1st scenario with OSPF alone, 2nd one with EIGRP alone and 3rd one with both EIGRP and OSPF concurrently".

Three network models were simulated, which are configured and run as 1st scenario with OSPF alone, 2nd one with EIGRP alone and 3rd one with both EIGRP and OSPF concurrently. One failure link between Sub-E and Sub-D has been configured to occur at 300 seconds and to recover at 500 seconds. The links that have been used in these scenarios are given in ***Table 1.0 below.***

TABLE I. LINK CONNECTION

Link Type	Connection between subnets	Link Speed
PPPDS3	Sub-C<->Sub-F, Sub-A<->Sub-C Sub-E<->Sub-C, Sub-B<->Sub-C Sub-E<->Sub-D, Sub-B<->Sub-D	44.736 Mbps
PPPDS1	Sub-A<->Sub-E ,Sub-B<->Sub-A Sub-C<->Sub-F	1.544 Mbps

F. *Results of simulation for the three models*

TABLE II. PACKET DELAY VARIATION RESULTS FOR EIGRP

No. of bits sent	Scenario Name	Routing Protocol	Packet Delay (sec)
5	EIGRP	EIGRP	0.026
10	EIGRP	EIGRP	0.028
15	EIGRP	EIGRP	0.030
20	EIGRP	EIGRP	0.032

TABLE III. PACKET DELAY VARIATION RESULTS FOR OSPF SCENARIO

No. of bits sent	Scenario Name	Routing Protocol	Packet Delay (sec)
5	OSPF	OSPF	0.043
10	OSPF	OSPF	0.045
15	OSPF	OSPF	0.047
20	OSPF	OSPF	0.049

TABLE IV. PACKET DELAY VARIATION RESULTS FOR EIGRP_OSPF SCENARIO

No. of bits sent	Scenario Name	Routing Protocol	Throughput (msec)
5	EIGRP_OSPF	EIGRP and OSPF	8,50,000
10	EIGRP_OSPF	EIGRP and OSPF	8,82,000
15	EIGRP_OSPF	EIGRP and OSPF	8,55,000
20	EIGRP_OSPF	EIGRP and OSPF	8,57,000

TABLE V. THROUGHPUT SIMULATION RESULTS FOR EIGRP

No. of bits sent	Scenario Name	Routing Protocol	Throughput (bits/sec)
5	EIRGP	EIGRP	8,80,000
10	EIRGP	EIGRP	8,20,000
15	EIRGP	EIRGP	8,85,000
20	EIGRP	EIRGP	8,87,000

TABLE VI. THROUGHPUT SIMULATION RESULTS FOR EIGRP_OSPF

No. of bits sent	Scenario Name	Routing Protocol	Packet Delay (msec)
5	EIGRP_OSPF	EIGRP and OSPF	0.026
10	EIGRP_OSPF	EIGRP and OSPF	0.027
15	EIGRP_OSPF	EIGRP and OSPF	0.028
20	EIGRP_OSPF	EIGRP and OSPF	0.029

IV. RESULTS AND FINDINGS

A. *Introduction*

In this section, the results obtained in chapter three are presented with their Performance Analysis of Enhance Interior Gateway Routing Protocol over Open Shortest Path First protocol. In all a model of three networks were designed and simulated, with configuration parameters and simulated based on 1st scenario with OSPF alone, 2nd scenario with EIGRP alone and 3rd scenario was a combination of both EIGRP and OSPF concurrently. A failure link established between Sub-E and Sub-D has been configured to occur at 300 seconds and to recover at 500 seconds tentatively.

B. *Packet delay variation graph*

Packet Delay Variation is measured based on the difference in the delay of the packets arriving at the destination. This performance and measurement metric has huge influence on the video and voice applications especially during streaming. The below figure 3.0, is the linear time variation increases starting from bits 5 to bits 20 showing that with the increase in traffic for voice and video applications the delay in packet transmission increase and EIGRP is slow to resolve packet delays when there is network congestion, and this also results in broken packet sizes before arriving at the destination, poor error detection and correction mechanisms and also poor bit synchronization may be the possible causes.

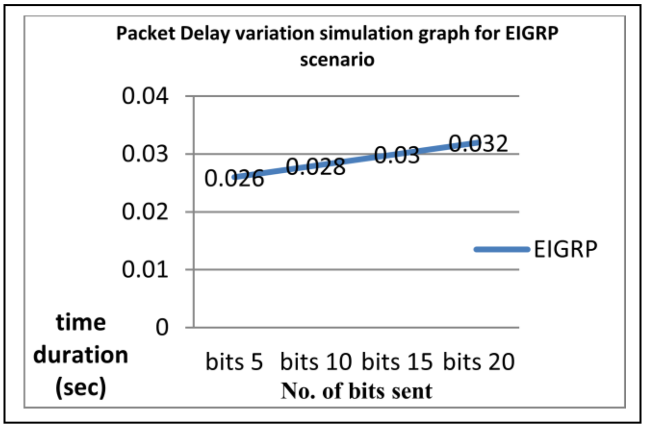


Fig. 3. Packet delay variation graph

Throughput simulation graph

The throughput is a key parameter to determine the rate at which total data packets are successfully delivered through the channel in the network. Figure 4.0 indicates that, bits 20 have high throughput and less packet loss than bits 5, bits 10 and bits 15 respectively. This means that EIGRP is efficient in handling throughput and packet loss during network congestion periods and therefore leads better error detection and correction, bit synchronization and faster routing table update interval time by EIGRP. Find below figure 4.0

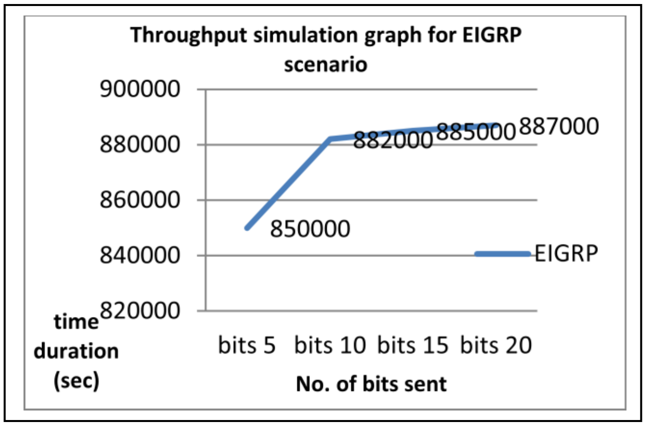


Fig. 4. Throughput simulation EIRGP

Packet delay simulation graph

This metric has huge influence on the manners of video applications. It is observed from the figure 5.0 that, the packet delay variation for OSPF networks are having higher values especially for that of bits 20 when it was sent through the OSPF network. Due to this, OSPF used triggered updates that allow efficient use of bandwidth and faster convergence time and not as susceptible to routing loops as EIGRP but requires more memory and processing power and harder to configure than EIGRP.

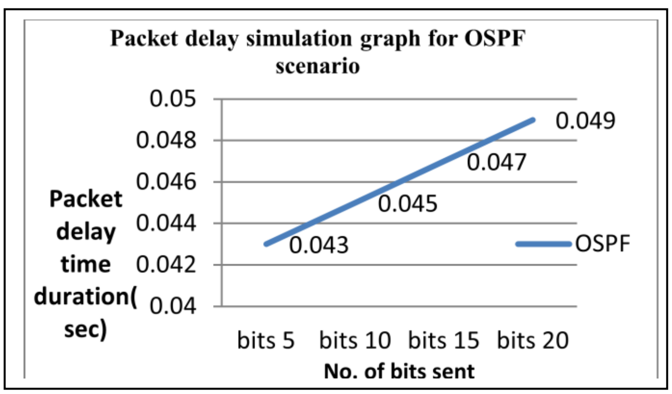


Fig. 5. Packet Delay simulation graph for OSPF protocol

Packet delay simulation graph for EIGRP_OSPF scenario

Packet Delay variation is measured by the difference in the delay of the packets. This metric has huge influence on the manners of video applications. It is observed from the figure 6.0 that EIGRP_OSPF has less packet delay variation than EIGRP and OSPF networks. Apparently, Figure 6.0 shown that despite of high congestion in the network, EIGR_POSPF is much better than OSPF and EIGRP network protocols in terms of packet delay variation and ensures efficient packet delivery.

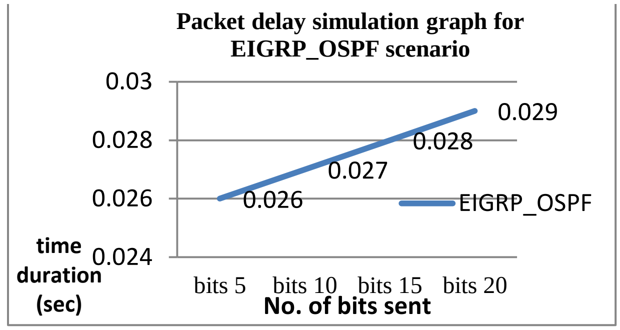


Fig. 6. Packet delay simulation EIGRP_OSPF

Throughput simulation graph for EIGRP_OSPF

The throughput is a key parameter to determine the rate at which total data packets are successfully delivered through the channel in the network. Figure 7.0 below indicates that El GRP_OSPF has higher throughput and less packet loss than OSPF and ElGRP networks especially for bits 15 and 20 respectively indicating an efficient network performance protocol suitable for voice and video applications .In effect EIGRP_OSPF has a better mechanism to ensure faster convergence and high throughput during data transmission especially where network congestion is rampant.

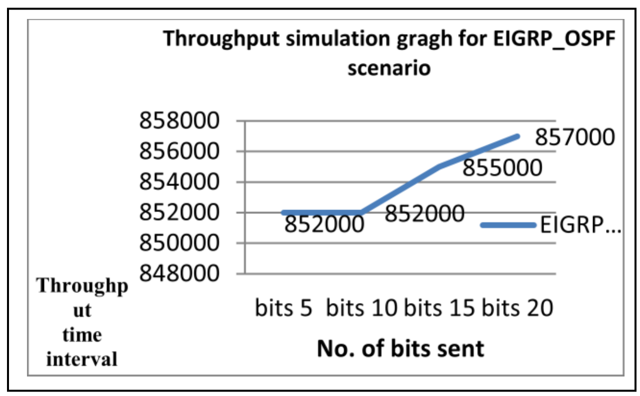


Fig. 7. Throughput simulation graph for EIGRP_OSPF

V. Conclusion/Recommendations

Network scalability is very important as it allows for future expansion of the network infrastructure can be enhanced by reducing network congestion and this demonstrates that the network convergence time is far better of as compared to EIGRP_OSPF and OSPF networks because EIGRP network is able learn the topological information and updates the routing table faster especially EIGRP and OSPF are widely being used in the computer networking. In this research work, I have presented a per analysis of selected routing protocols such as EIGRP, OSPF and the combination of EIGRP and OSPF

The result of the simulation has shown the difference between the end to end delay of EIGRP_OSPF network is relatively less than EIGRP and OSPF networks. As a result unstable network bandwidth, data packets in EIGRP_OSPF network reach faster to their destination as compared to OSPF. Another performance metrics for real time applications is packet delay variation that is a measurement of the difference between the delays of packets on transmission. The performance of packet delay variation for EIGRP_OSPF scenario is far better than OSPF and EIGRP relatively. Also, concerning the packet delay variations of EIGRP and OSPF networks is high while EIGRP_OSPF network is low. The case in the of context of packet loss, it was found that the packet loss in the EIGRP_OSPF network is less than OSPF and EIGRP networks. In final comparison, the overall simulation results have reflected that the maximum throughput in the combination of EIGRP and OSPF network is much higher than OSPF and EIGRP networks. In this research, the performance analysis among EIGRP, OSPF and combination of EIGRP and OSPF routing protocols for real time applications have been analyzed sequentially. By comparing comparing these protocols' performances, it can also be concluded that the combined implementation of EIGRP and OSPF routing protocols used in the network scenario performs far better than OSPF and EIGRP.

REFERENCES

[1] Christian Huitema, "Routing in the internet" 2. Ed. Prentice Hall PTR, cop. 2000.
[2] Cisco, "Internet Technology Handbook" ,2009
[3] Cisco, "IP Routing, Introduction to EIGRP" Document l D: 13669.
[4] Douglas E. Comer, 2006.
[5] Esuendale Lemm, Syed Hussain, WendwossenAnjelo 2010.
[6] FarazShamim, Zaheer Aziz, Johnson Liu, Abe Martey, 2002.
[7] J. Broch, D. Maltz, D. Johnson, Y. Hu and, 1998.
[8] K. Yamazaki and K. Sezaki, 2002.
[9] S. Basagni, I. Chlantac, V. Syrotiuk and B. Woodward, 1998.
[10] W. Su, S-J. Lee and M. Gerla, 2003.
[11] Y. Bae and N. Vaida, 1998.

Identify and Classify Critical Success Factor of Agile Software Development Methodology Using Mind Map

Tasneem Abd El Hameed
MSc student /Information system Department
Faculty of computers and information, Helwan University
Cairo, Egypt

Mahmoud Abd EL Latif
Assistant Professor of information system
Faculty of computers and information, Helwan University

Sherif Kholief
Assistant Professor of information system
Faculty of computers and information, Helwan University

***Abstract*—Selecting the right method, right personnel and right practices, and applying them adequately, determine the success of software development. In this paper, a qualitative study is carried out among the critical factors of success from previous studies. The factors of success match with their relative principles to illustrate the most valuable factor for agile approach success, this paper also prove that the twelve principles poorly identified for few factors resulting from qualitative and quantitative past studies. Dimensions and Factors are presented using Critical success Dimensions and Factors Mind Map Model.**

***Keywords*—Agile success factor; Agile principles**

I. Introduction

Most of the software is used in businesses and organizations all over the world. Nowadays, due to Volatile and unpredictable nature of system projects development, traditional approaches become inflexible, and are unable to adjust to the system projects.

Agile approach is developed to overcome the failures that result from traditional approaches, and to offer a lightweight framework for helping organizations and teams to respond faster and iteratively.

When it comes to adopting agile approach, the first step organizations start with is the agile manifesto stating that agile approach contains four values and twelve principles. Unfortunately, applying these principles, can create challenges, and these challenges result in failures or successes, which help in figuring out factors for success and failure.

Transition from being non-agile to agile in an organization is tricky and difficult. Agile adoption and transformation efforts are experiencing high failure rates in organizations. 84% of respondents in the Agile Development Survey reported that they had experienced a failed Agile project. Only 16% of respondents had not experienced failure[22]. In this paper, we clarify the success factors resulting from different previous studies in a graphical way called Mind Map for easy learning.

Mind map is a graphical way to represent ideas and concepts, and is like a visual thinking tool that helps structuring information, helping in better analysis. Mind Map represents semantic or other connections between portions of learned material hierarchically. Mind Maps are easy to learn and apply, and provide a concise hierarchic overview, The further advantages of mind map are easy to extend and add further content. They are Idiosyncratic, hard to read for others, can be inconsistent, and can become overly complex (loss of big picture) [8]. To overcome the complexity, we distinct each dimension with related factors in a particular map, to be easy in learning and reading.

The Remainder of the paper is organized into five sections. Section 2 discusses a background study, Section 3 discusses literature review, Section 4 discusses factors, and Section 5 is the conclusion.

II. Background

Software Process Models are based on one of the three models of software Development. The models are *waterfall approach, Iterative Development approach* [2].

A. Waterfall Approach

The waterfall approach emphasizes a structured progression between defined phases. Each phase consists of a definite set of activities and deliverables that must be accomplished before the following phase can begin [3].

The waterfall phases are requirements definition, system and software design, implementation and unit testing, integration and system testing, operation and maintenance [2]. Waterfall model has some disadvantages which are “1) some requirements may emerge after the requirements gathering phase, resulting in some problems. 2) Problems detected at a stage are not solved completely in the same stage. 3) There is no concept of changing (partitioning) the project into multiple stages. 4) New requirements by the client are very expensive and cannot be adjusted in the current edition of the software product. 5) Estimation of time and budget for each stage is very difficult. 6) No prototype before the finishing of the life cycle. 7) Testing in the last stage of the development. 8) If testing reveals some problems, then going to the design stage is very difficult. 9) Very high risk in the entire life cycle development. Not recommended for object oriented projects” [4].

B. Iterative development approach

Iterative Development starts with a simple implementation of a small set of the software requirements and iteratively enhances the evolving versions until the complete system is implemented and ready to be deployed. Process models have explicitly been designed to support (iteration development are incremental delivery and spiral development).

- Incremental delivery

Incremental Delivery customer identify, the services to be provided by the system and which are the most and the least important, however, there are some problems with incremental delivery. Increments should be small and each increment should deliver some of the system functionality. “It can be difficult to map the customer’s requirement onto the increments of the right size” [2].

- Spiral Development

Spiral Development Consists of Loops where each loop in the spiral represents a phase of software process. Each loop is split into four factors: Objective setting, risk assessment and reduction, development and validation, planning. The main difference between spiral and other software processes is the explicit recognition of risks [2]. Disadvantage of spiral model is a very complex solution suitable for big, complicated projects, and yet undoubtedly more flexible than the original waterfall method. The spiral model is an underlying inspiration for many modern methods, for example, Rational Unified Process [5].

C. Agile Software Approach

Agile is increasingly becoming the dominating developing method in software industry. For a successful software project, it is essential to identify what leads to success. Projects succeed when enough factors are well defined, and failure teaches us to overcome shortcomings in the future projects [12].

- Agile Manifesto

Agile approaches are introduced to overcome the failure factors for traditional SDLC, “In 2001 the agile manifesto states that four values and twelve principles, four Values are (1) Individuals and interactions over processes and tools (2) Working software over comprehensive documentation (3) Customer collaboration over contract negotiation (4) Responding to change over following a plan and twelve principles are (1) Our highest priority is to satisfy the customer through early and continuous delivery of valuable software. (2)Welcome changing requirements, even late in development. Agile processes harness change for the customer's competitive advantage. (3) Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale. (4) Business people and developers must work together daily throughout the project.

(5) Build projects around motivated individuals give them the environment and support they need and trust them to get the job done. (6)The most efficient and effective method of conveying information to and within a development team is face-to-face conversation. (7)Working software is the primary measure of progress. (8)Agile processes promote sustainable development. The sponsors, developers, and users should be able to maintain a constant pace indefinitely. (9)Continuous attention to technical excellence and good design enhances agility. (10)Simplicity--the art of maximizing the amount of work not done--is essential. (11)The best architectures, requirements, and designs emerge from self-organizing teams.(12) at regular intervals, the team reflects on how to become more effective , then tunes and adjusts its behavior accordingly”. [7]

Agile is a group of lightweight methodologies used to develop highly potential software. Agile methods universally rely on an iterative approach to software specification development and delivery, they are intended to deliver working software quickly to customers, who can propose new and changed requirement to be included in later iterations of the system [2].

- ***Agile Methodologies***

Agile software development methodologies share many features and practices which include the practice of whole team, measures, short release, test-driven development, Pair Programming, customer collaboration, Prototyping, refactoring, continuous integration and less documentation to produce valuable software [2]. Agile Methodologies are Extreme Programming (XP), Crystal Methods, Feature Driven development (FDD), system development Methods (DSDM), Scrum.

- Extreme Programming(XP)

Based on a set of practices like pair programming, customer collocation, customer satisfaction [3]. XP is perhaps the best known and most widely used of the agile methods [2].

- Crystal Methods

Crystal methods is a lightweight methodology, Based on premise that people impact software development projects more than tools or processes [2].

- Feature Driven development (FDD)

FDD is an iterative, incremental, and lightweight software development process. It is a combination of a number of industry-recognized best. These practices are all driven from a client- valued functionality. The main purpose it to deliver tangible, working software repeatedly in a timely manner [3].

- Dynamic system development Methods (DSDM)

It is mainly a framework more than a process. Dynamic system development method is about fixing quality, cost and time. DSDM is used for developing software and non-IT companies [3].

- Scrum

Is an incremental and iterative framework where practitioners can employ different processes and techniques to develop a complex product and project. It is specially designed to handle rapidly changing business requirements.

Scrum is a sprint; a time-boxed effort usually from two weeks to four weeks, in a sprint work is divided into parts and to be completed at the end of the sprint time. Scrum focuses more on management of the process than coding techniques, and it is used in small and large projects [3].

III. Literature Review

Search began in 1996 when Walid and Oya suggested a new framework to determine Critical Success and failure factors. They suggested a new design for critical factors and described the impact of these factors on performance of the project. They used Empirical study to test practicality of using the suggested design and grouped the factors into 4 areas: Project, Managers and team member, organization, and environment. The survey results "demonstrate that project managers, managerial skills, team members, commitment and their technical background, project attributes and environment factors are as viable and can be as a critical as an organization factor and the criticality of these factor varies between industries"[9].

A lot of researches are done and discussed from 1996 till 2006. Aniket Mahanti made a survey paper of major challenges in adopting agile practices by enterprises. Successful adoption of agile methodology includes obtain management buy-in, education and support, integrating to external processes, starting pilot projects, report and adapt, and sustain agility. The success of agile adoption is directly related to how the new methodology is introduced in the organization [10].

After that, in 2008 a survey study of critical success factors in agile software projects was done by Tsun chow, Dac-Buu Cao using quantitative approach. A study led by agile experts, gathered survey of 109 agile projects from 25 countries all over the world. Multiple regression techniques were used. The survey results obtained that only 10 out of 48 hypotheses were supported, and identified three critical success factors for agile software projects, "Delivery Strategy, agile software engineering techniques and team capability"[11]. They conclude that they should try to define different success factors or try to display the success of agile projects with different method [11].

Then Dragan Stankovica, Vesna Nikolicb, Miodrag Djordjevicc, Dac-Buu Cao continued the study (Chow and Cao, 2008). They tried to verify the classification of critical success factors previously described in study by Chow and Cao (2008). They made a regression analysis on the collected data which introduced three more factors that could potentially be considered as critical success factors. [12]

Subhas Chandra Misra, Vinod Kumar, and Uma Kumar developed a conjectured hypothetical success factors framework to address the research question. They used the data analysis techniques to validate the hypotheses. The study was made utilizing an extensive scale study-based methodology, comprising respondents who practice agile software development and who had experience in practicing arrangement driven programming advancement previously. The study demonstrates that 9 of the 14 hypothesized factors have significant relationship with ''Success". The important success factors that were found are: "customer collaboration, customer satisfaction, customer commitment, decision time, corporate culture, personal characteristics, societal culture, and training and learning" [13].

In 2010, Zulkefi, Saadiah and Noor carried out a literature review to gather information from previous study, they found that "Customer involvement, communication, minimum changes of requirement, corporate culture, time allocation, simplicity, active testing, code review and customer collaboration determine the successful in agile software development methodology" [14]. They developed a conceptual model in their study.

Jianping and Routing have designed P company success factors model "leading (recognition of top leaders, participation of top leaders), organization (creating clear vision, building the agile organizational culture, changing the way of management), tools and technology(configuring the necessary tools and infrastructure, using design patterns and other advanced design methods, using software reuse technology), appropriate import(selecting applicable import project, excellence implementation staff, selecting proper agile method practice), training and education(correct understanding and mastery of agile methodologies, enhancing the professional capabilities of the employee), measuring success(flexible and innovative development method, rapid response to demand, forward looking response to changing factors, successfully building learning organization) which they verified them by a questionnaire in P Company" [15]. They conclude that education and training play a positive role in agile improvement. Agile method must be established in agile culture with due attention to the design and application of technology.

In 2011 Claudia, Daniela, Fabio, and Reidar made two case studies in industry and analysed data from two projects. They identified three literature review and present the main factors based on most relevant factors, "the factors include product (reuse, software characteristics), project(resource constraint, schedule, team composition, communication), personnel (team experience and motivation) and process(customer participation, daily builds, documentation, early prototyping, incremental and iterative development, modern programming practice, programming language abstraction, software methods, tool usage)"[16]. They conclude that there are some factors impacting the productivity of agile teams. These factors are team decomposition and allocation, external dependencies, and staff turnover.

Ani Liza and Andrew M Gravell initiated a study that involves 13 participants including CEOs, project managers, founders, and developers. Their study resulted that social and human aspects are very important when they start using the agile methods. During the study, they used qualitative semi-structure interview and concluded that the issues and challenges during adoption were mindset, knowledge, project, people, knowledge transfer, management involvement, communication, technical aspects, and organizational structure[17].

Kumar and Goe (2012) illustrate the results of a survey conducted to demonstrate and explain the factors considered by software practitioners while adopting agile methodologies, and the effects of adopting agile methodologies on customers and business while practicing agile. They presented six hypotheses, which are “impact of team size, impact of requirement gathering for agile methodologies, effective requirement capturing method, time taken to resolve a problem and impact of small response time with customer on software development” [18]. The results of this research indicated that adopting agile increased the productivity of an organization and also increased customer satisfaction [18].

IV. Critical Success Factors

In this section, Critical Success Dimensions that grouped from previous quantitative and qualitative studies are displayed. The factors were arranged based upon the most ranked one to the least ranked ones. Similar factors were collected with each other and given the most common name between studies.

Finally, each principle was given a constant value (see Table 8) where we matched each factor with their related one as emphasized by Maarit [22], see Fig. 9. We found that most of studies are about certain dimension so each dimension has been given a weight to determine which dimension is the most important table [1-6]. The average displayed is the weight of each dimension as it determine how important is. All factors and dimensions are collected and displayed in one Mind Map to be more memorized, then Each dimension are illustrated with their related factors, at the end of this section, all Dimensions with their factors are presented in only one CSDF Model (Critical Success dimensions and Factors).

Fig. 1 includes Process, People, Project, Product, Organization, technical (4P OT), where each dimension is one of the most critical success dimensions grouped from previous researches.

Fig. 1. CSF Mind Map Model

A. People

Most of the researchers agreed that People Dimension plays an important role in any software development project. People dimensions are classified into eight factors, and these eight factors are demonstrated in Table 1.

People factors may include Education where team should learn agile techniques and how to apply and adopt them in non-agile companies, accomplished by learning how to support teams, as the project manager should have the ability to tradeoff, ability to coordinate, and participate in all aspects. Also the team member and the manager must be committed for their tasks and project.

Table 1 displays People Dimension and all factors according to the highest one. [9][10][11][12][13][14][15][16][18][19].

TABLE I. People Factors

Dimensions	Factors	{Belassi1996}	{Mahanti2006}	{Chow2008}	{Misra2009}	{MansorSandNoor2010}	{Wan2010}	{Melo2011}	{Kumar2012}	{Sheffield2013}	{Hummel2015}	Rate	Principles	AVG
People	Education and support	√	√	√	√		√	√			√	7	P5	4.5
	Customer Centric issues	√		√	√	√			√	√	√	7	P1,2,4	
	Management Style	√	√	√	√		√	√				6	All	
	Communication Skills	√		√	√	√		√			√	6	P6	
	Motivation			√	√			√	√	√		5	P5,11	
	Commitment	√		√							√	3	P3	
	Report and Adapt		√									1	P2	
	Project Champion	√										1	P2	

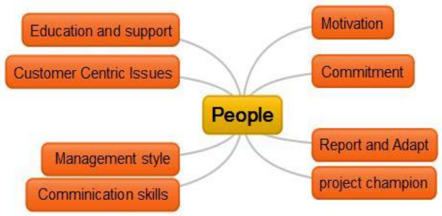


Fig. 2. People Mind Map

B. *Organization*

When an organization starts to adopt agile, the way of organizational culture and mindset have to change. An agile is way of working that deliver along new practices for teams and managers, and usually agile impacts the organizational culture and mindset. Although these factors are important but also changing everything at the same time might be too big challenge for an organization. Organization dimension with their factors are displayed in Table 2 [9][10][11][12][13][14][15][16][18][19]

Fig. 3. Organization Mind Map

TABLE II. ORGANIZATION FACTORS

Dimensions	Factors	{Belassi1996}	{Mahanti2006}	{Chow2008}	{Misra2009}	{MansorSandNoor2010}	{Wan2010}	{Melo2011}	{Kumar2012}	{Sheffield2013}	{Hummel2015}	Rate	Principles	AVG
Organization	Corporate Culture				√	√	√		√	√		4	P4	2.25
	Organizational Environment (Political economical, technological environment)	√									√	2	P5	
	Collocation of Whole team			√				√				2	P4,6	
	Sustain Agility		√									1	P8	

C. *Technical*

Technical Practices such as continuous integration, test-driven development, pair programming, refactoring, and collective ownership are what has made it possible for most of organization to deliver what the customer need efficiently and effectively. Teams will be become twice as productive, if they adopt some of these practices [20]. Table 3 displays the technical dimension and factors. [9][10][11][12][13][14][15][16][18][19]

Fig. 4. Technical Mind Map

TABLE III. TECHNICAL FACTORS

Dimensions	Factors	{Belassi1996}	{Mahanti2006}	{Chow2008}	{Misra2009}	{MansorSandNoor2010}	{Wan2010}	{Melo2011}	{Kumar2012}	{Sheffield2013}	{Hummel2015}	Rate	Principles	AVG
Technical	High expertise for team and organizational factor	√		√	√		√		√		√	4	P9	3.5
	Practices					√	√	√			√	2	P9	
	Trouble shooting for team	√		√					√			2	P9	
	Tool Usage		√				√	√				1	P9	

D. Process

There are many agile processes: SCRUM, Crystal, Behavior-Driven Development (BDD), Test-Driven Development (TDD), Feature-Driven Development (FDD), Adaptive Software Development (ADP), Extreme Programming (XP), and more. [9][10][11][12][13][14][15][16][18][19]

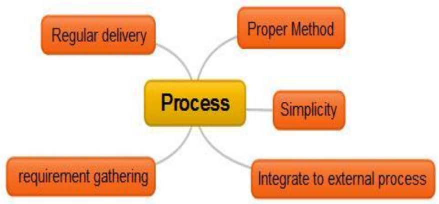


Fig. 5. Process Mind Maps

TABLE IV. PROCESS FACTORS

Dimensions	Factors	{Belassi1996}	{Mahanti2006}	{Chow2008}	{Misra2009}	{MansorSandNoor2010}	{Wan2010}	{Melo2011}	{Kumar2012}	{Sheffield2013}	{Hummel2015}	Rate	Principles	AVG
Process	Regular Delivery of software			√	√	√		√	√		√	5	P9	2.8
	Effective requirement gathering method			√				√	√			3	P9	
	Select proper Methodology			√							√	2	P9	
	Simplicity			√		√						2	P9	
	Integrate to external processes		√	√								2	X	

E. Project

Agile Methods are most applicable to projects where requirements are ill-defined and fluid since they seek to accommodate change easily. Projects that are unprecedented within an organization or use cutting-edge technology are examples of projects where change is likely to have a significant impact on the project [21]. Table 5 displays project factors [9][10][11][12][13][14]16].

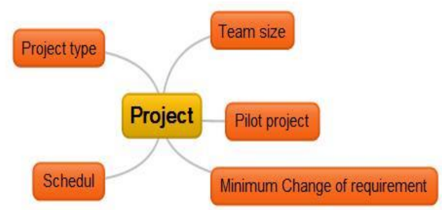


Fig. 6. Project Mind Maps

TABLE V. PROJECT FACTORS

Dimensions	Factors	{Belassi1996}	{Mahanti2006}	{Chow2008}	{Misra2009}	{MansorSandNoor2010}	{Wan2010}	{Melo2011}	{Kumar2012}	{Sheffield2013}	{Hummel2015}	Rate	Principles	AVG
Project	Project type	√		√	√						√	4	X	2.4
	Schedule		√	√				√				3	X	
	Tem Size			√		√		√			√	3	X	
	Pilot Project		√									1	X	
	Minimum changes of Requirements					√						1	X	

F. Product

Most of studies merge the product with project and process due to the big similarities between them but only few studies separate them, we illustrate it separate to be more specific and Precise. Table 6 displays product factor [15][16].

Fig. 7. Product Mind Map

TABLE VI. PRODUCT FACTORS

Dimensions	Factors	{Belassi1996}	{Mahanti2006}	{Chow2008}	{Misra2009}	{MansorSandNoor2010}	{Wan2010}	{Melo2011}	{Kumar2012}	{Sheffield2013}	{Hummel2015}	Rate	Principles	AVG
Product	Using Software reuse technology						√	√				2	X	1.5
	Software characteristics							√				1	X	

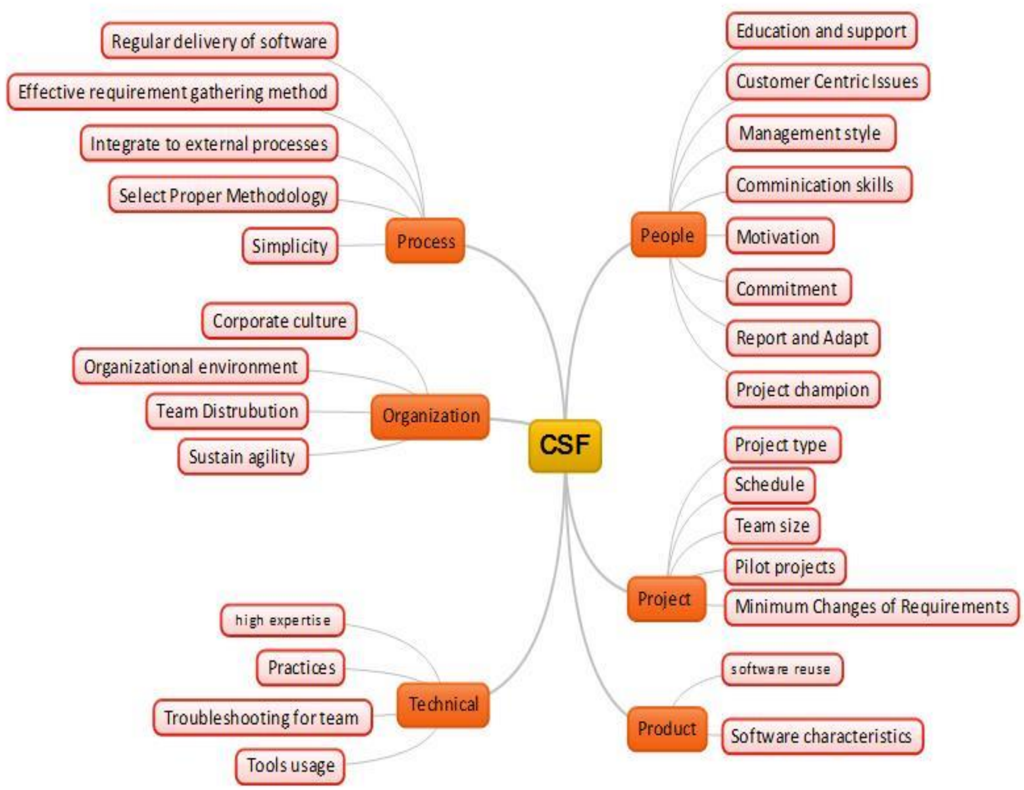


Fig. 8. CSDF Mind Map Model

TABLE VII. PRINICPLES AND WHAT THEY EMPHASIZE[22]

Agile Principles	**Emphasis**
Our highest priority is to satisfy the customer through early ad continuous delivery of valuable software	Customer satisfaction, continuous delivery, value, early deliveries
Welcome changing requirements, even late in development. Agile harness change for customer's competitive advantage	Adaptability, competitiveness, customer benefit
Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter time scale.	Working software frequently, shorter time scale frequent deliveries
Business people and developers must work together daily throughout the project.	Work together daily, collaboration
Build Projects around motivated individuals. Give them the environment and support they need, and trust them to get the job done	Motivated individuals, good environment, support, trust
The most efficient and effective method of conveying information to and within a development team is face to face conversation	Efficiency, Communication
Working software is the primary measure of progress	Measure progress via deliverable
Agile Processes promote sustainable development. the sponsors, developers, and users should be able to maintain a constant Pace identifility	Sustainability, People
Continuous attention to technical excellence and good design enhance agility	Focus on technical excellence, good design as enabler of agility
Simplicity – the art of maximizing the amount of work not don is essential	Simplicity – optimizing work
The best architectures, requirements, and design emerge from self-organizing teams	Self-organization
At regular interval, the team reflect on how to become more effective, then tunes and adjusts its behavior	Built-in improvement of efficiency

TABLE VIII. AGILE PRINCIPLES AND CONSTANT VALUES

Principles	**Factors**	**Constant**
Our highest priority is to satisfy the customer through early ad continuous delivery of valuable software	Customer satisfaction, continuous delivery, value, early deliveries	P1
Welcome changing requirements, even late in development. Agile harness change for customer's competitive advantage	Adaptability, competitiveness, customer benefit	P2
Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter time scale.	Working software frequently, shorter time scale frequent deliveries	P3
Business people and developers must work together daily throughout the project.	Work together daily, collaboration	P4
Build Projects around motivated individuals. Give them the environment and support they need, and trust them to get the job done	Motivated individuals, good environment, support, trust	P5
The most efficient and effective method of conveying information to and within a development team is face to face conversation	Efficiency, Communication	P6
Working software is the primary measure of progress	Measure progress via deliverable	P7
Agile Processes promote sustainable development. the sponsors, developers, and users should be able to maintain a constant Pace identifility	Sustainability, People	P8
Continuous attention to technical excellence and good design enhance agility	Focus on technical excellence, good design as enabler of agility	P9
Simplicity – the art of maximizing the amount of work not don is essential	Simplicity – optimizing work	P10
The best architectures, requirements, and design emerge from self-organizing teams	Self-organization	P11
At regular interval, the team reflect on how to become more effective, then tunes and adjusts its behavior	Built-in improvement of efficiency	P12

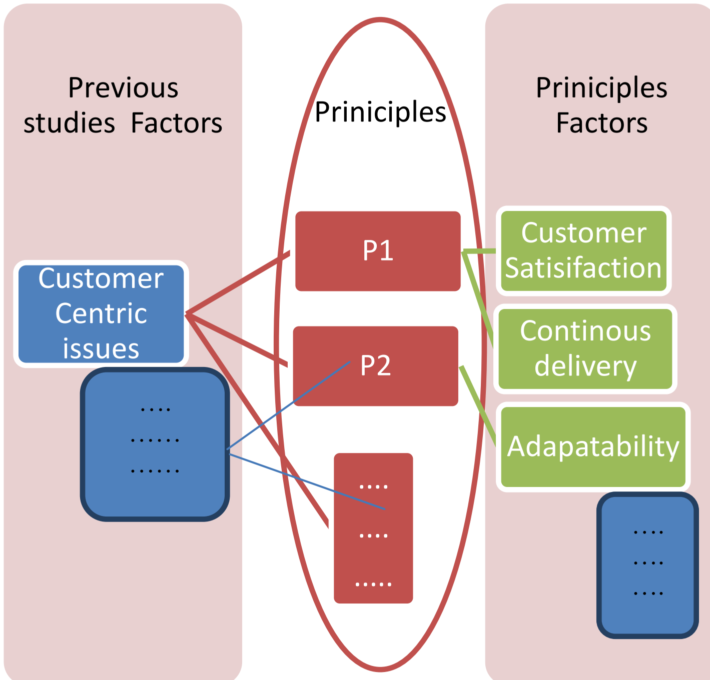


Fig. 9. Factors Matches Agile Principles

V. Conclusion and Future Work

This paper represents and identifies the success dimensions and factors for agile and how it matches the agile principles and concluded that few of factors resulted from people experts and companies' statistics and surveys. In the future, a quantitative study will be conducted by these factors to measure how these factors succeeded in Egypt companies.

In Process Dimension, the Factor (Effective requirement gathering method) is not explicitly appeared in P2 & P11, and also, continuous integration factor mentioned in the principles, but nothing mentioned about integration with external process and how adopting the agile in an old process, this means that gathering requirement and integrate with external process, is poorly identified in agile principles. In Process and Product Dimensions, there is no straightforward principle to clarify these factors, and all these factors result from qualitative and quantitative surveys resulting in turn from people and companies experiences.

Although there are some factors that are poorly identified, the remaining are clear, and all factors might have the probability of success and failure, but it depends on how, where, and when we use them in the organization.

We also suggest a framework to display the factors called (4P OT) which stand for People, Process, Product, Project, Organization, and Technical Model. In addition to another framework collect all dimensions with factors (CSDF) Model.

References

[1] Li, Tong. An Approach to Modelling and Describing Software Evolution Processes. UK: Software Technology Research Laboratory - De Montfort U, 2007.

[2] Sommerville, Ian. Software Engineering. 8th ed. Harlow, England: Addison-Wesley, 2007

[3] Ahmed Awad, Mohamed. A Comparison between Agile and Traditional Software Development Methodologies. U of Western Australia

[4] Ullah, Mehar. Comparison and Problems between Traditional and Agile Software Development Methods. Lappeenranta U of Technology, 2014

[5] Fianta, Roman. Iterative Web Systems Development. Brno, Spring, 2009.

[6] Boehm, B.W.(1988). A Spiral model of software development and enhancement. IEEE Computer, 21(5), 61-72. (Chs. 4, 5)

[7] "Manifesto for Agile Software Development." Manifesto for Agile Software Development. 2001. Web. <http://www.agilemanifesto.org/>

[8] J Eppler, Martin. "A Comparison between Concept Maps, Mind Maps, Conceptual Diagrams, and Visual Metaphors as Complementary Tools for Knowledge Construction and Sharing."Information Visualization 5 (2006): 202-10.

[9] W. Belassi and O. I. Tukel, "A new framework for determining critical success/failure factors in projects," Int. J. Proj. Manag., vol. 14, no. 3, pp. 141–151, 1996.

[10] A. Mahanti, "Challenges in enterprise adoption of agile methods-A survey," CIT. J. Comput. Inf. Technol., vol. 14, no. 3, pp. 197–206, 2006.

[11] T. Chow and D.-B. Cao, "A survey study of critical success factors in agile software projects," J. Syst. Softw., vol. 81, no. 6, pp. 961–971, 2008.

[12] Stankovic, D., et al., A survey study of critical success factors in agile software projects in former Yugoslavia IT companies. Journal of Systems and Software, 2013. 86(6): p. 1663-1678.

[13] S. C. Misra, V. Kumar, and U. Kumar, “Identifying some important success factors in adopting agile software development practices,” J. Syst. Softw., vol. 82, no. 11, pp. 1869–1890, 2009.

[14] Z. Mansor Saadiah Yahya, and Noor Habibah Arshad, “Success_Determinants_in_Agile_Software_Development_Methodology_ICMLC_Updated.” 2010.

[15] J. Wan and R. Wang, “Empirical Research on Critical Success Factors of Agile Software Process Improvement,” J. Softw. Eng. Appl., vol. 3, no. 12, p. 1131, 2010.

[16] C. Melo, D. S. Cruzes, F. Kon, and R. Conradi, “Agile team perceptions of productivity factors,” in Agile Conference (AGILE), 2011, 2011, pp. 57–66.

[17] A. L. Asnawi, A. M. Gravell, and G. B. Wills, “Empirical investigation on agile methods usage: issues identified from early adopters in Malaysia,” in Agile Processes in Software Engineering and Extreme Programming, Springer, 2011, pp. 192–207.

[18] A. Kumar and B. Goel, “Factors Influencing Agile Practices: A Survey,” 2012.

[19] J. Sheffield and J. Lemétayer, “Factors associated with the software development agility of successful projects,” Int. J. Proj. Manag., vol. 31, no. 3, pp. 459–472, 2013.

[20] Oluwole, Dele. "Agile X Factors." Https://www.scrumalliance.org. Web. <https://www.scrumalliance.org/community/articles/2013/december/agile-x-factors>.

[21] M. Coram and S. Bohner, “The Impact of Agile Methods on Software Project Management A Brief Look at Agile Methods,” Engineering, pp. 363–370, 2005.

[22] M. Sahota, An Agile Adoption and Transformation Survival Guide: Working with Organizational Culture. 2012.

An Enhanced Framework with Advanced Study to Incorporate the Searching of E-Commerce Products Using Modernization of Database Queries

Mohd Muntjir
College of Computers and Information Technology
Taif University
Taif, Saudi Arabia

Ahmad Tasnim Siddiqui
College of Computers and Information Technology
Taif University
Taif, Saudi Arabia

***Abstract*—This study aims to inspect and evaluate the integration of database queries and their use in e-commerce product searches. It has been observed that e-commerce is one of the most prominent trends, which have been emerged in the business world, for the past decade. E-commerce has gained tremendous popularity, as it offers higher flexibility, cost efficiency, effectiveness, and convenience, to both, consumers and businesses. Large number of retailing companies has adopted this technology, in order to expand their operations, across of the globe; hence they needs to have highly responsive and integrated databases. In this regard, the approach of database queries is found to be the most appropriate and adequate techniques, as it simplifies the searches of e-commerce products.**

Keywords—E-Commerce; Database; Database; Queries; Integration; Database Queries

I. Introduction

The purpose of this paper is to present the illustration of database queries as well as their use in searching e-commerce products. It has been assessed that the concept of e-commerce has gained tremendous popularity with the emergence of innovative and advanced technological tools. According to Li and Karahanna (2012), e-commerce or electronic commerce can be understood as the trading of services or products, by using computer networks, usually internet. It has been assessed that the functions or operations of e-commerce draws on different technologies, including automated data collection systems, inventory management systems, EDI (electronic data interchange), electronic funds transfer ,online transaction processing, supply chain management ,internet marketing, and mobile commerce. In this regard, Liu, et.al, (2010) has declared that contemporary e-commerce utilizes the World Wide Web, in order to conduct different transactions and retailing activities.

Recently researchers have developed various different techniques, in order to access and purchase e-commerce products and services. In this account, Endrullis, et.al, (2012) has asserted that integration and utilization of database queries may play a commendable role in searching e-commerce products, more conveniently. Database queries can be understood as one of the most advanced databases, which are based on the relational model, which was established by Codd. It is significant to notice that query form is found to be one of the most efficient and integrated user interfaces, which are being widely used for querying databases (VanderMeer, et.al, 2012). It has been inspected that database queries play an inevitable role in providing feasible and quick access to the required information or products.

It is due to the fact that in this paradigm, a database table usually represents a mathematical relationship amid set of different products or items, having similar attributes or characteristics. In other words, the entire framework of database queries is developed in such a manner, which assists its use to access ample amount of relevant and creditable information from complex and large databases. Thereby, it can be avowed that the integration and use of database queries may play an indispensable role in e-commerce product searches (Vander Meer, et.al, 2012). The proceeding paper will help in understanding the core concept of e-commerce as well as database queries. Mainly this study will focus on the integration of database queries in the process of searching e-commerce products.

II. Aim and Objectives

This study aims to investigate the integration and use of database queries in e-commerce product searches. Proceeding mentioned objectives would be fulfilled, in order to accomplish this research aim.

- To analyze the notion of e-commerce;
- To understand the basic concept of database queries;
- To examine different modules of database queries;
- To assess different techniques of integrating and utilizing database queries in order to search e-commerce products.

III. Significance of Research

According to Das-Sarma, et.al, (2014), technological developments have brought considerable changes in the lives of people, in terms of performing daily routine tasks. Most prominently, these technological developments have influenced business and retailing sector. In this regard, e-commerce can be considered as one of the most prominent advancements, which have been occurred in retailing industry, due to technological advancements. In this regard, Liu, et.al, (2010) has claimed that e-commerce has played a major role in transforming the actual faced of retailing industry, as this innovative approach allows the users to perform their transactions and purchasing activities by using different online and digital tools, regardless of their geographical locations (Telang, et.al, 2012).

It has been documented in the studies of Endrullis, et.al, (2012) in recent times; e-commerce has been developed at fast paced, across the world. E-commerce can be understood as the process of purchasing or selling services and goods, by using electronic tools. It is important to bring into the notice that e-commerce transactions can be conducted amid private and public organizations, governments, individuals, households, and businesses. It has been established that there are various different types of e-commerce transactions that takes place online ranging from sale of books, shoes, cloths to different services, like making hotel bookings or airline tickets.

During e-commerce activities, users have to face difficulties, in terms of searching their desired products and services. In this scenario, e-businesses have to implement and integrate high-tech systems, in order being ease and convenience for e-commerce users, in terms of searching their required products and services. It has been claimed by Li and Karahanna (2012) that database queries are found to be one of the integrated and valuable methods, which assists the users in accessing required products, in considerably massive and sophisticated databases.

Basically, the technique of database queries helps in optimizing the products through keywords; hence results in immediate and quick product searches. Thereby, it can be affirmed that the integration of database queries in e-commerce product searches is one of the greatest initiatives towards integrated and sustainable retailing activities. However, this technique also possesses various disadvantages, most prominently language gap. It has been established that the language gap is usually occurred in the keywords, which are used for search queries, and database's product specifications (Telang, et.al, 2012).

IV. Literatire Review

A. *E-Commerce-The Concept*

Li and Karahanna (2012) have claimed that e-commerce includes several activities, including exchanging, selling, and buying of services, products, and information, through computer networks, mainly internet. It is important to notice that the term "commerce", is usually referred to the transactions, which are conducted amid business partners. When the concept of e-commerce was evaluated on the pre-discussed definition of commerce, it was recognized that e-commerce is the same as electronic business or e-business, but this approach is not true. It is due to the fact that e-commerce does not only deal with selling or purchasing of goods, but also with collaborating with servicing customers,business partners, and performing electronic transactions within an association.

Accordingly, search marketing success for e-commerce sites is predicated less on innovation than possessing an expert understanding of factors that impact search marketing efforts.

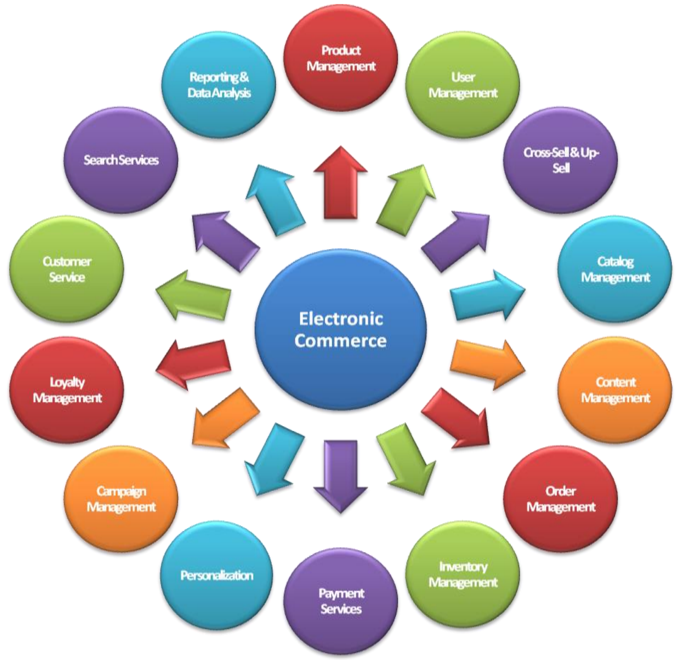


Fig. 1. Classifications of E-Commerce

It has been documented in the researches and studies, which were conducted by Liu, et.al, (2010) that e-business is all about reaching new customers, enhanced productivity, globalization, speed, time cycle, and sharing knowledge across institutions, in order to gain competitive advantages. According to Endrullis, et.al, (2012), e-commerce offers wide range of benefits to the organizations, customers, as well as to the entire society. When e-commerce benefits, for organizations, were analyzed it was revealed that e-commerce enables the businesses and companies to expand their market to international and national levels, while saving their cost and time.

Fig. 2. E-Commerce

In this account, Leszczynski and Stencel (2011) have affirmed that e-commerce plays an appreciable role in quickly locating more customers, suitable business partners, as well as valuable suppliers for business. In addition to this, e-commerce also allows the companies to procure services and material from other companies, in a cost effective and efficient manner. E-commerce also enables considerably specialized niche market. In customer's perspective, e-commerce provides less expensive services and products by enabling the customers to perform quick online comparisons.

More so, e-commerce also helps the consumers in getting customized products, ranging from computer systems to super luxury cars, at highly competitive prices. According to Leszczynski and Stencel (2011), e-commerce is also found to be one of the most beneficial approaches for the entire society. It is due to the fact that e-commerce allows people, living in remote areas, to get their desired products, while saving travelling cost. Furthermore, e-commerce also assists people to receive well-timed healthcare services, education facilities, etc., irrespective of their geographical locations.

Fig. 3. E-Commerce Framework

B. Database Queries

It has been established from the analysis of research and studies, which were carried out by Liu, et.al, (2010) that database queries can be considered as one of the most effective and integrated approaches, which allow its users to attain their desired information or data from substantially gigantic databases. It has been identified that the technique of database queries plays an inevitable and incredible role in enabling the users to point out their desired information, without accessing or making efforts to search for the entire table. Endrullis, et.al, (2012) has supported this idea by claiming that the approach of database queries allows its users to amalgamate wide range of tables. This phenomenon can be easily understood by considering an example, which is, if a user of the database is dealing with two different tables, including consumers and invoices, they can easily use this technique, i.e., database queries.

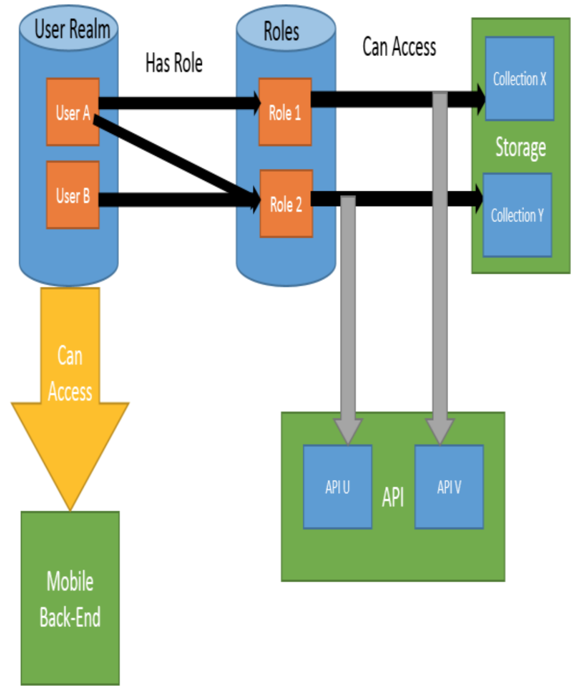


Fig. 4. A Framework of a System To Fetch the Data

It is due to the fact that this database queries inevitably helps in amalgamating the contents or information, which is being stored in two different tables. Afterwards, the names and complete data of the customers can be easily attained, by executing this query by the user. It is important to notice that the final results are according to the invoices of the consumers. Leszczynski and Stencel (2011) have presented an idea according to which the technique of database queries is not capable enough to store the data. Lu, et.al, (2013) has supported this approach by asserting that database queries can only identify the stored data, instead of storing it.

In accordance with the views of Telang, et.al, (2012) database query possesses wide range of benefits, which may play a major role in impacting the operations of e-commerce. In this regard, one of the most prominent advantages of database queries is that it merges or amalgamates wide range of data or information, which is being stored in different databases. In other words, database query combines required and valuable information from several different sources of data. In addition to this, database query also enable its users to choose their required fields from different sources, while identifying them, as per their needs. More so, this innovative and integrated framework also helps in specifying the records, which are similar to the criteria, which was predefined by the users (Li, et.al, 2013).

It has been documented in the researches of Endrullis, et.al, (2012), apart from various advantages; database queries also comprise different disadvantages. One of the most prominent disadvantages of database queries may include the language gap amid the requests (made by the users) and technical terms, which are used during the development of a database. It has been observed that researchers are intending to cope with this issue, as this feature considerably affects the reliability of database queries (VanderMeer, et.al, 2012).

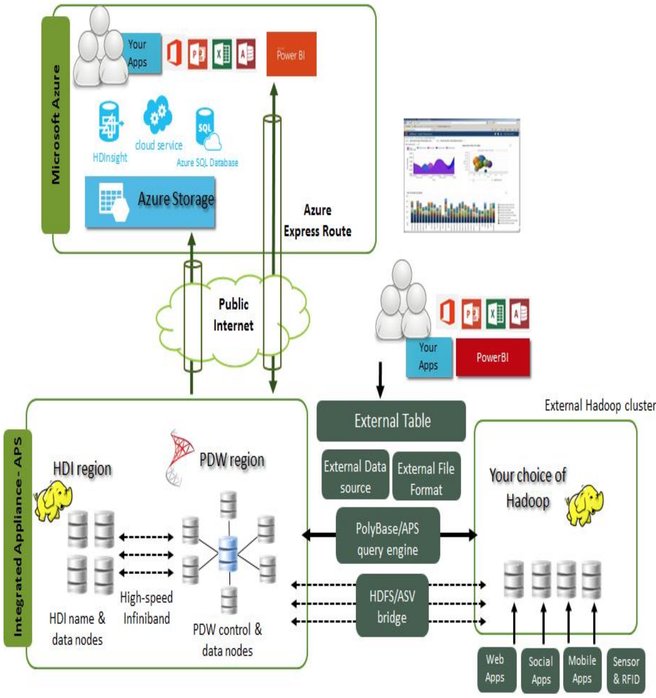


Fig. 5. An Architecture of Database To Fetch The Queries

C. Modules of Database Queries

It has been claimed by Leszczynski and Stencel (2011) that various modules have been developed by the researchers, in order to ensure the integrity and credibility of the database queries' operations. Some of the most prominent modules are briefly described in the proceeding paper.There are given two modules for Database Queries.

a) Query Breaking Module: According to Das-Sarma, et.al, (2014), query breaking module plays a crucial role in ensuring integrated operations of database query. It has been analyzed that massive queries to database often results in the malfunctioning of the database; hence results in stoppage of database or other unfavorable consequences. Therefore, it isnecessary to break the chosen queries into large number of small queries. It has been observed that the activity plays an inevitable and incredible role in reducing potential risks, which are usually occurred due to increased complexities. Li and Karahanna (2012) have asserted that the entire process can be easily conducted by the help of query breaking module. However, this module also possesses some issues, in terms of identifying adequate LECO (local e-catalog ontologies).

b) Query Reasoning and Expanding Module:

Query reasoning and expanding module is another most effective feature, which resides in database queries. It has been observed that this module is entirely different from traditional query system, as the primary features of the semantic query include the activity of expanding the query reason, during the process of querying. This function helps the users in identifying and attaining their desired information and statistics, in anappropriate and effectivebehavior (Wang, et.al, 2012).

V. Integrating and Utilizing Database Queries in E-Commerce Product Searches

In the current era, businesses are continually starving to expand their operations, across the globe, while controlling their operational cost.

Fig. 6. Demonstration of E-Commerce Product Search

In this regard, the emergence of e-commerce has commendably benefited and supported the businesses, in terms of expanding their operations, on local and international levels (Wang, et.al, 2012). It is a fact that e-commerce has played a vital role in facilitating the businesses as well as consumers in terms of performing their transactions and retailing activities. On the other hand, continually increasing use of e-commerce has also affected the activities of product searching. It is due to the fact that massive number of requests (queries), made by the consumers often affect the performance and functionality of the databases; hence influencing the integrity of e-commerce activities. In this account, Das-Sarma, et.al, (2014) has avowed that continuously increasing demands of e-commerce have pressurized the database developers to formulate such databases, which are capable enough to directly find out the query of the webpage content. In this scenario, the approach of database queries can be characterized amid one of the most appropriate solutions.

Recent studies, which are accomplished by Li, et.al, (2013) have revealed the fact that e-commerce has made it feasible for people to shop their desired products and services, in spite of living in remote areas. Emerging trends of e-commerce have also played an indispensable role in supporting mid as well as small sized companies to increase their revenues or profits through e-commerce. It is because; this tool (e-commerce) helps them in expanding their target market, without investing large capitals (Lu, et.al, 2013).

Therefore, it can be affirmed that e-commerce has opened new doors of feasibility, convenience, cost efficiency, and time efficiency, both for companies and consumers. While considering these trends, it can be stated that the integration of database queries in e-commerce activities may result in commendably profitable outcomes. It is due to the fact that database queries may considerably reduce certain ambiguities and indistinctness, while carrying out e-commerce product searches (Wang, et.al, 2012).

It has been asserted by Das-Sarma, et.al, (2014) that database queries may also facilitate various other areas of life, including ERP systems, CRM systems, personalized marketing applications, e-commerce apps, as well as custom-built apps. It is significant to notice that database queries encapsulate the traits of relational databases, which plays a noticeable role in improving the overall performance of the product search (Lu, et.al, 2013).

When database queries are incorporated with the e-commerce product searches, it usually results in higher atomicity, consistency, integrity, durability, accuracy, and higher efficacy, in terms of transactions. All of these features are considered as most important elements for the environment of e-commerce product searches. Therefore, it can be stated that integration of database queries with e-commerce product searches may commendablysupport the companies, which have adopted e-commerce approach. It has been observed that the paradigm of database queries inevitably improves the scalability and reliability of e-commerce product search operations and aidsthe consumers to find out their preferred products and services, in an accurate and efficient manner (Li, et.al, 2013).

Endrullis, et.al, (2012) has claimed that consumers have to submit their vouchers or payment information, while purchasing their desired products and services, from different websites (e-commerce websites). In this regard, the consumers have to choose their required items, while assessing the specialties (cost, quality, etc.) of the products. On the other hand, companies have to identify the requirements, demands, and current trends of the market. In this scenario, integrated and systematically developed database query systems may substantially help them in developing their products as well as updating their e-commerce websites (Lu, et.al, 2013).

VI. Conclusion

From above discussion, it can be concluded that integration of database queries in e-commerce product searches is one of the greatest initiatives towards integrated, smooth, and systematic e-commerce activities. It is due to the fact that the technique of database queries plays an indispensable role in eliminating the probability of different risks, which are often occurred in different databases. Highly advanced and innovative modules of database queries play a vital role in controlling excessive traffic of queries, by breaking large queries into small ones; hence results in more efficient and well-timed searches of e-commerce products. The proceeding paper has briefly discussed the concept of e-commerce and database queries. More so, the paper has also encapsulated the analysis of different aspects, which are associated with the integration of database queries and their use in e-commerce product searches.

Acknowledgment

We would like to gratefully and sincerely thank to The Dean of our College, and Chairman of our Departments for their guidance, understanding, patience, and most importantly, their friendly nature during this research paper writing at College of computers and Information Technology, Taif University, Saudi Arabia. We would also like to thank all of the members of the research group and friends who have always supported us for giving the opportunity to write this research paper.

References

[1] Das-Sarma, A., Parikh, N., & Sundaresan, N. (2014). E-commerce product search: personalization, diversification, and beyond. In Proceedings of the companion publication of the 23rd international conference on World wide web companion, International World Wide Web Conferences Steering Committee, available at, http://wwwconference.org/proceedings/www2014/companion/p189.pdf

[2] Endrullis, S., Thor, A., & Rahm, E. (2012). Entity search strategies for mashup applications. In Data Engineering (ICDE), IEEE, available at, http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6228073&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D6228073

[3] Leszczynski, P., & Stencel, K. (2011). Consistent caching of data objects in database driven websites. In Advances in Databases and Information Systems, Springer Berlin Heidelberg, available at, http://link.springer.com/chapter/10.1007/978-3-642-15576-5_28

[4] Li, S., & Karahanna, E. (2012). Peer-based recommendations in online B2C e-commerce: comparing collaborative personalization and social network-based personalization. IEEE, available at, http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6148983&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D6148983

[5] Li, Y., Wang, Y., Jiang, P., & Zhang, Z. (2013). Multi-objective optimization integration of query interfaces for the Deep Web based on attribute constraints.*Data & Knowledge Engineering*, available at, http://www.sciencedirect.com/science/article/pii/S0169023X13000049

[6] Liu, Z. H., Novoselsky, A., & Arora, V. (2010). XML Data Management in Object Relational Database Systems. Advanced Applications and Structures in XML Processing: Label Streams, Semantics Utilization, and Data Query Technologies, available at, http://www.igi-global.com/chapter/xml-data-management-object-relational/41498

[7] Lu, Y., He, H., Zhao, H., Meng, W., & Yu, C. (2013). Annotating search results from Web databases. Knowledge and Data Engineering, *IEEE*, available at, http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5989804&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D5989804

[8] Telang, A., Li, C., & Chakravarthy, S. (2012). One Size Does Not Fit All: Toward User-and Query-Dependent Ranking for Web Databases. Knowledge and Data Engineering, *IEEE*, available at, http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5710921&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D5710921

[9] VanderMeer, D., Dutta, K., & Datta, A. (2012). A cost-based database request distribution technique for online e-commerce applications. *MIS quarterly*, available at, http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3032&context=misq&sei-redir=1&referer=http%3A%2F%2Fscholar.google.com%2Fscholar%3Fq%3DAn%2BActive%2BWeb-based%2BDistributed%2BDatabase%2BSystem%2Bfor%2BE-Commerce%26btnG%3D%26hl%3Den%26as_sdt%3D0%252C5%26as_ylo%3D2010#search=%22An%20Active%20Web-based%20Distributed%20Database%20System%20E-Commerce%22

[10] Wang, Y., Lu, J., Liang, J., Chen, J., & Liu, J. (2012). Selecting queries from sample to crawl deep web data sources. *Web Intelligence and Agent Systems*,available at, http://iospress.metapress.com/content/6349501674330k72/

IAX-JINGLE Network Architectures Based-One/Two Translation Gateways

Hadeel Saleh Haj Aliwi
School of Computer Sciences
Universiti Sains Malaysia
Penang, Malaysia

Putra Sumari
School of Computer Sciences
Universiti Sains Malaysia
Penang, Malaysia

***Abstract*—Nowadays, Multimedia Communication has improved rapidly to allow people to communicate via the Internet. However, Internet users cannot communicate with each other unless they use the same chatting applications since each chatting application uses a certain signaling protocol to make the media call. The interworking module is a very critical issue since it solves the communication problems between any two protocols, and enables people around the world to make a voice/video call even if they use different chatting applications. Providing interoperability between different signaling protocols and multimedia applications takes the advantages of more than one protocol. Usually, each signaling protocol has its own messages which differ from other signaling protocol messages format. Thus, when two clients use different signaling protocols want to communicate phonetically, the sent/received messages between them will not be understood because the control and media packets in each protocol are different than the corresponding ones in the other protocol, The interworking module solves this kind of problem by matching the signals and media messages by providing translation gateways in the middle between the two protocols. Thus, many interworking modules have been proposed in order to enable many protocols' users to chat with each other without any difficulties. This paper compares two interworking modules between Inter-Asterisk eXchange Protocol and Jingle Protocol. An experimental implementation in terms of session time is provided.**

Keywords—media conferencing; VoIP; interworking; translation gateway; IAX; Jingle

I. Introduction

Over the last few years, the need to provide communication facilities for participants all over the world and at any time via computer network systems has increased. These network systems enable the use of multimedia applications with several kinds of media conferencing, such as audio, video, graphics, images, and text [6, 17].

Nowadays, many signaling protocols and techniques such as Multimedia Conferencing and Voice over Internet Protocols (VoIP) [9] have been created and developed according to their usages in providing services between at least two participants. Such protocols are Session Initiation Protocol (SIP) [3, 7], InterAsterisk eXchange protocol (IAX) [1], Real-time SWitching Control Protocol (RSW) [11], eXtensible Messaging and Presence extension Protocol (XMPP extension/ Jingle) [15], H.323 protocol [2, 8], etc. All the signaling protocols take place in the session layer (L5) in the Open Systems Interconnection (OSI) model. Layer 5 provides the mechanism for opening, closing and managing a session between the end-user application processes.

With the appearance of numerous signaling protocols, the decision to choose the appropriate protocol to be utilized in such a service has become very difficult since each protocol has its own privileges which differ from the corresponding privileges of the other protocols [12]. Choosing IAX and Jingle protocols to build an interworking module between them is due to many reasons; IAX is an interesting alternative compared to the conventional VoIP protocols. Nowadays, IAX is being deployed by service providers for their VoIP service offerings (e.g. H.323 and SIP). IAX protocol offers significant features that are not provided by other existent VoIP signaling protocols. Furthermore, many researchers have shown that IAX is slightly better than SIP, H.323, and RSW in terms of the quality of services.

Just as IAX protocol has many features, Jingle protocol is considered as the standard protocol for Gmail chatting application with regard to audio and video conferencing services. Most popular chatting applications use Jingle protocol to handle the call setup, audio/video chatting, and call teardown sessions. Such applications are Gtalk, Talkonaut, and Hangouts [10].

This paper compares two interworking modules between IAX and Jingle protocols with regard to call setup, call teardown, and media sessions.

II. Background

A. IAX Protocol

Mark Spencer has created the Inter-Asterisk eXchange (IAX) protocol for asterisk that performs VoIP signaling [14]. IAX is supported by a few other softswitches, (Asterisk Private Branch eXchange) PBX systems, and softphones [5]. Any type of streaming media can be managed, controlled and transmitted through the Internet Protocol (IP) networks based on IAX protocol. However, IP voice calls are basically being controlled by IAX protocol. Currently, IAX has been changed to IAX2 which is the second version of the IAX protocol [18]. IAX protocol is used for many purposes, firstly, it is to minimize bandwidth usage for both control and media transmissions with specific emphasis on individual voice calls, secondly, to provide Network Address Translation (NAT) transparency, thirdly, to support the ability to transmit dial plan information, and lastly, to support efficient implementation of intercom and paging features. IAX is considered as both signaling and media protocol since it has its own media transfer method (full and mini frames) to exchange the data, unlike the other signaling protocols which use Real time Transport Protocol (RTP) to carry the data during the media session.

B. Jingle Protocol

Jingle protocol is the extension of the eXtensible Messaging and Presence Protocol (XMPP) [4] which is a standard specified by the IETF for carrying instant message service. XMPP is an open eXtensible Markup Language (XML) protocol for a real-time messaging, presence, and request/response services, and it is an out-of-band signaling protocol. The XMPP architecture consists of three elements, XMPP client, XMPP server and gateways to foreign networks [16]. The developers have added media session capabilities (which have been defined as an XMPP-specific negotiation protocol called Jingle) to XMPP clients [13, 19]. Jingle has been designed to support many types of applications, such as voice and video conferencing, file transfer, application sharing, and others.

III. IAX-JINGLE INTERWORKING MODULES: A COMPARISON

Both IAX and Jingle protocols are widely used to provide two ways media transfer features. Each protocol differs from the other one in many ways, such as registration matters, transport methods, media transport, signals, header format, and media packet formatting. Each protocol has its own signals in order to manage the call setup/ teardown sessions which has the same task compared to the other protocols but different formats. Table 1 shows IAX/Jingle signals matching. Solving the aforementioned problems will enable people around the world to talk with each other without caring about the protocols used by their applications.

In order to enable the IAX users to communicate with people who use application base Jingle protocol without any difficulties, an interworking module between IAX and Jingle has been presented in order to help bridging the gap between them and to provide the capability of IAX-Jingle interoperability. The network architecture of the first interworking module consists of IAX domain (IAX Client, IAX Server), Jingle domain (Jingle Client, Jingle Server), and IAX-Jingle gateway in the middle of IAX and Jingle domains, whereas, the architecture of the second interworking module consists of IAX domain (IAX Client, IAX Server, IAX-to-Jingle gateway), and Jingle domain (Jingle Client, Jingle Server, Jingle-to-IAX gateway)

The presented translation gateways are considered as translation and database server. The translation gateways are considered as a translator when sending any type of messages from one protocol to the other. The tasks of the translation gateways are represented by translator of call setup and teardown signals, and real time media data.

TABLE I. IAX/JINGLE CALL SETUP AND TEARDOWN MESSAGES

IAX	**Jingle**
NEW	Session-initiate
ACCEPT	Session-info (ping)
0RINGING	Session-info (ringing)
ACK	IQ-Result (ack)
ANSWER	Session-accept
HANGUP	Session-terminate <success/>
BUSY	Session-terminate <busy/>
REJECT	Session-terminate <decline/>

A. Call Setup/Teardown Sessions

For each signaling protocol, the call setup has to go through four steps: call initiation, negotiation acceptance, ringing, and answering. The caller has to send a terminate signal to the callee to end the call. In case of two different protocols, the translation gateway is needed for the translation or matching matters between IAX and Jingle users.

By using only one translation gateway, the translation gateway will be responsible of checking first whether the packet received belongs to either IAX client or Jingle client before translates the packet and forwards it to the other party. The checking step has to be done for each received packet by the translation gateway as well as the gateway is responsible for handling sending and receiving directions of both IAX and Jingle, in addition to two methods of packet translation; IAX message format to Jingle message format and vice versa. These steps have to be done for all signaling messages. Concluding that, using one translation gateway will lead to larger delay time compared to using more than one translation gateway.

The IAX-Jingle architecture based two translation gateways distributes the function of translation gateway into two gateways (IAX-to-Jingle and Jingle-to-IAX), so each gateway receives only from one party and sends only to the other party, in this case no need from the gateway to check the sent/received packet belongs to which party, and since each translation gateway handles only one direction, the two translation methods (IAX-to-Jingle and Jingle-to-IAX) have to be distributed between the two translation gateways, so each translation gateway performs only one translation method. This makes the function of each gateway is simpler and lead to less delay time compared to using one translation gateway.

Figures 1, 2, 3, and 4 present call setup/teardown in case of one and two translation gateways.

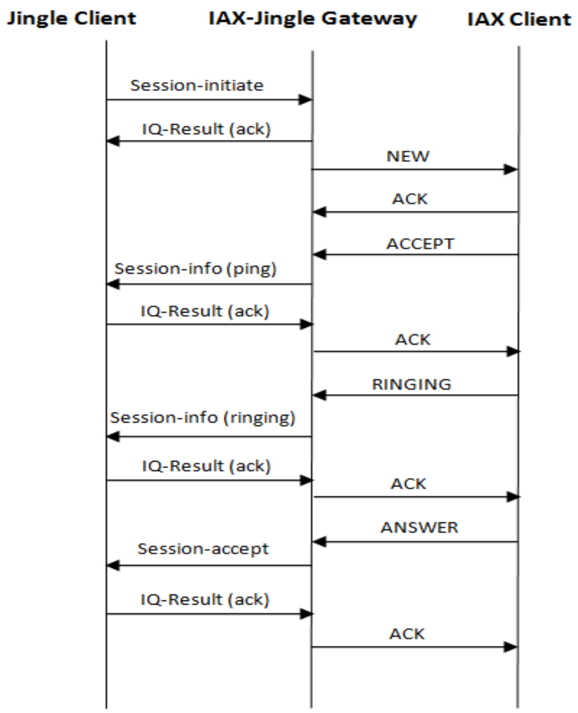


Fig. 1. IAX-Jingle Call Setup Session: One Translation Gateway

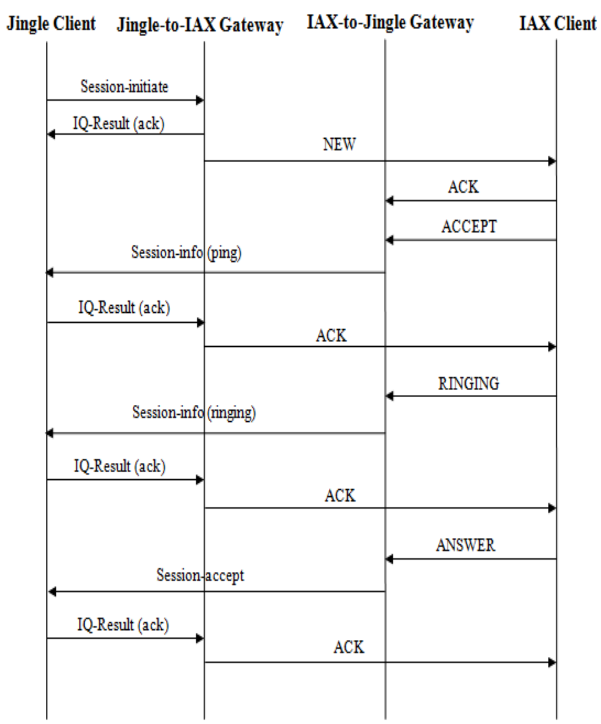


Fig. 2. IAX-Jingle Call Setup Session: Two Translation Gateways

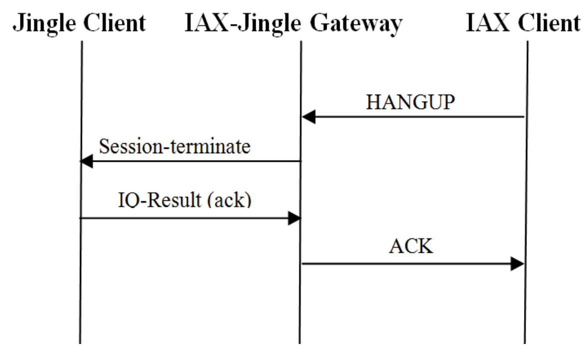


Fig. 3. IAX-Jingle Call Teardown Session: One Translation Gateway

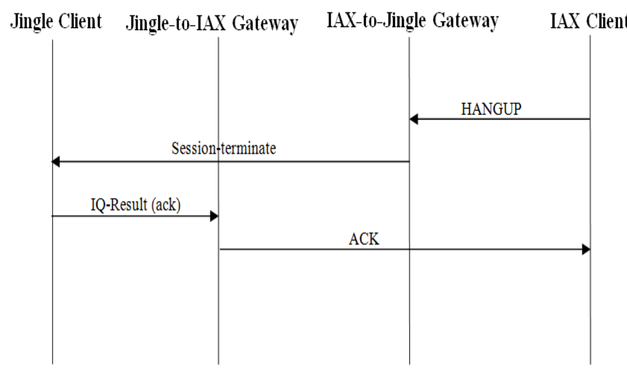


Fig. 4. IAX-Jingle Call Teardown Session: Two Translation Gateways

B. Media Session

The media session happens after the call setup session/ before the call teardown session. Just as the translation gateway translates the signals format in order to be exchanged between two different protocols, it is also responsible for translating the media packet format during the media session as each protocol has its own media packet format.

In case of using one translation gateway which is two ways gateway (IAX to Jingle/ Jingle to IAX), the gateway has to check whether the received packet from the client is carried by mini frame or RTP. If the packet is carried by mini frame, so it has been sent by IAX protocol, otherwise it is a Jingle packet. As a result, the translation gateway has to send the IAX packet to the Jingle client carried by RTP and vice versa. Thus, by using one translation gateway, its task is to both check the packet format and translate the packet format of the other protocol.

In case of using two translation gateways (IAX to Jingle and Jingle to IAX), the gateway is only one way gateway so no need check the packet format as each gateway receive from only one protocol and send to only the other one. So, the task of each of the two gateways is translating the packet format of the other protocol. Figures 5 and 6 show the IAX-Jingle media session in case of using one and two translation gateways.

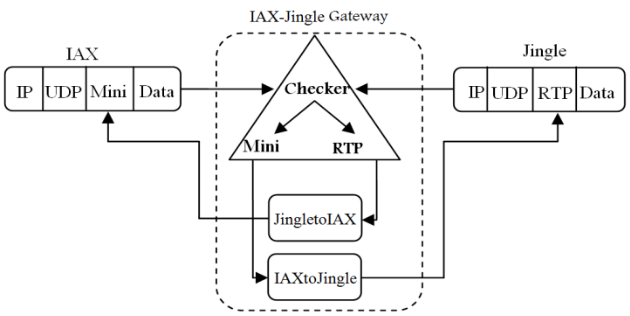


Fig. 5. IAX-Jingle Media Session: One Translation Gateway

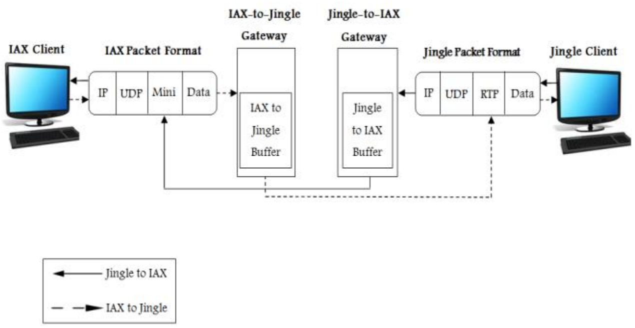


Fig. 6. IAX-Jingle Media Session: Two Translation Gateways

IV. Results

The IAX-Jingle interworking module has been implemented by using ns2.35 simulator. The results have been obtained in terms of session time. In IAX-Jingle Environment, two main sessions have to be considered namely signaling session and media session. Signaling session is divided into two sessions: setup session and teardown session. Table 1 describes the simulation parameters such as the packet size, the transport protocols used, and simulation time.

TABLE II. SIMULATION PARAMETERS

Parameter	Value Used in Scenarios
Nodes	1) IAX Client, IAX Server, IAX-Jingle Gateway, Jingle Client, Jingle Server 2) IAX Client, IAX Server, IAX-to-Jingle Gateway, Jingle-to-IAX Gateway, Jingle Client, Jingle Server
Number of Calls	Varies between 1 and 50 Calls
Codec	G.711
Network Protocol	IP
Transport Protocol	UDP, RTP
Signaling Protocol	IAX & Jingle
Transmission Range	375 m
Data Packet Size	512 Bytes
Simulation Time	50 Seconds

Eleven scenarios have been tested in order to find the session time of the IAX-Jingle media conferencing as shown in Tables 3, 4, and 5. Each scenario shows the IAX-Jingle network architecture with certain number of calls. This means that each scenario differs from the others in the number of clients/ calls. The different number of calls has been settled to 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 for the 11 scenarios respectively.

TABLE III. IAX-JINGLE SETUP SESSION TIME: A COMPARISON

Number of Calls	Setup Session Time by using 1 Gateway	Setup Session Time by using 2 Gateways
1	0.01655 Seconds	0.0152 Seconds
5	0.04739 Seconds	0.0467 Seconds
10	0.06557 Seconds	0.0649 Seconds
15	0.1032 Seconds	0.1025 Seconds
20	0.1225 Seconds	0.121 Seconds
25	0.1572 Seconds	0.1553 Seconds
30	0.2066 Seconds	0.2054 Seconds
35	0.2334 Seconds	0.2321 Seconds
40	0.2726 Seconds	0.2717 Seconds
45	0.336 Seconds	0.3354 Seconds
50	0.3626 Seconds	0.362 Seconds

TABLE IV. IAX-JINGLE TEARDOWN SESSION TIME: A COMPARISON

Number of Calls	Teardown Session Time by using 1 Gateway	Teardown Session Time by using 2 Gateways
1	0.01349 Seconds	0.01288 Seconds
5	0.04146 Seconds	0.04066 Seconds
10	0.0625 Seconds	0.0599 Seconds
15	0.0898 Seconds	0.0888 Seconds
20	0.1046 Seconds	0.1035 Seconds
25	0.1258 Seconds	0.1245 Seconds
30	0.169 Seconds	0.1678 Seconds
35	0.2163 Seconds	0.215 Seconds
40	0.2442 Seconds	0.2433 Seconds
45	0.2958 Seconds	0.2946 Seconds
50	0.3356 Seconds	0.3347 Seconds

In the experiments with more than one call, each session time value has been founded by calculating the average of the session time values for the whole number of calls.

For example, to find the setup/teardown session time within five calls, we have to find the summation of the setup/teardown session time of call 1, call 2, call 3, call 4, and call 5 divided by 5 which is the number of calls. This means that:

Setup/teardown session time (for 5 calls) = [setup/teardown session time (for call 1) + setup/teardown session time (for call 2) + setup/teardown session time (for call 3) + setup/teardown session time (for call 4) + setup/teardown session time (for call 5)] / 5.

TABLE V. IAX-JINGLE MEDIA SESSION TIME: A COMPARISON

Number of Calls	Media Session Time by using 1 Gateway	Media Session Time by using 2 Gateways
1	0.27061 Seconds	0.245445 Seconds
5	0.824824 Seconds	0.80373 Seconds
10	1.292835 Seconds	1.193071 Seconds
15	1.869356 Seconds	1.789603 Seconds
20	2.181975 Seconds	2.097207 Seconds
25	2.642627 Seconds	2.551829 Seconds
30	3.435142 Seconds	3.368528 Seconds
35	4.028682 Seconds	3.85267 Seconds
40	4.663689 Seconds	4.516233 Seconds

45	5.671236 Seconds	5.52836 Seconds
50	6.155735 Seconds	6.082079 Seconds

For media session experiments, the session time values have been founded for the first hundred packets only. To find the media session time during the first 100 packet, we have to calculate the summation of the end to end packet delay values (d) starting from the packet sequence number 1 until the packet sequence number 100 with considering the number of calls. Session Time for the first 100 packets with n number of call=$[\sum_1^n \sum_1^{100} d]/n$ (1)

V. Conclusion

This paper provides a comparison between two IAX-Jingle interworking modules in terms of call setup, call teardown, and media session time. It can be noticed from the experiments that the IAX-Jingle network architecture based two translation gateways has improvement of performance over the architecture based one translation gateway due to distributing the task of one translation gateway into two gateways in order to make the translation process consuming less time.

References

[1] T. Abbasi, S. Prasad, N. Seddigh, and I. Lambadaris, "A Comparative Study of the SIP and IAX," Canadian Conference on Electrical and Computer Engineering (CCECE), Saskatoon, Canada, pp. 179-183, 2005.

[2] I. Basicevic, M. Popovic, and D. Kukolj, "Comparison of sip and h.323 protocols," The Third International Conference on Digital Telecommunications ICDT '08, pp. 162–167, 2008.

[3] M. Boucadair, "Inter-Asterisk Exchange (IAX): Deployment Scenarios in SIP-Enabled Networks," Wiley InterScience, 2009.

[4] H.T. Chu, W. Chen, Y.H. Huang, and J.Y. Chen, "A novel design of instant messaging service extended from short message service with XMPP," Fifth IEE International Conference on 3G Mobile Communication Technologies, pp. 504- 508, 2004.

[5] Doxygen, "IAX2 Configuration," Asterisk - The Open Source Telephony Project, 2013.

[6] B. Forouzan, "Data Communications and Networking," 4th edition, McGrawHill, New York, USA, 2007.

[7] D. Geneiatakis, T. Dagiuklas, G. Kambourakis, C. Lambrinoudakis, and S. Gritzalis, "Survey of security vulnerabilities in session initial protocol," IEEE Communications Surveys & Tutorials, pp. 68-81, 2006.

[8] J. Glasmann, W. Kellerer, and H. Muller, "Service Architectures in H.323 and SIP: A Comparison," IEEE Communications Surveys & Tutorials, pp. 32-47, 2003.

[9] B. Goode, "Voice over internet protocol (VoIP)," Proceedings of the IEEE, pp. 1495 –1517, 2002.

[10] Google Developers, "Google Talk for developers," 2011.

[11] H.S. Haj Aliwi, S.A. Alomari, and P. Sumari, "An Effective Method for Audio Translation between IAX and RSW Protocols," World Academy of Science, Engineering and Technology 59 2011, France, pp. 253-256, 2011.

[12] H.S. Haj Aliwi and P. Sumari, "Audio/video Mapping Architecture between different Signaling Protocols: Problems and Suggestions," Journal of Applied Sciences, UAE, pp. 178-188, 2016.

[13] S. Kille, "JINGLE Implementation Growing: Open Alternative to Skype," 2008.

[14] M. Kolhar, M. Abu-Alhaj, O. Abouabdalla, T.C. Wan, and A. Manasrah, "Comparative Evaluation and Analysis of IAX and RSW," International Journal of Computer Science and Information Technology (IJCSIS), USA, pp. 250–252, 2009.

[15] S. Ludwig, J. Beda, P. Saint-Andre, R. McQueen, S. Egan, and J. Hildebrand, "Jingle," XSF XEP 0166, 2007.

[16] P. Nie, "An open standard for instant messaging: eXtensible Messaging and Presence Protocol (XMPP)," TKK T-110.5190 Seminar on Internetworking, Helsinki University of Technology, Finland, pp. 1-6, 2006.

[17] S. Ramadass,. "A distributed architecture to support multimedia applications over the internet and corporate intranets," In proceedings of SEACOMM'98, Penang, Malaysia, 1998.

[18] E. Reeves, " Devices That Handle Inter-Asterisk eXchange, Version 2 (IAX2) Protocol," 2011.

[19] P. Saint-Andre, "Interworking between the Session Initiation Protocol (SIP) and the Extensible Messaging and Presence Protocol (XMPP): Media Sessions," draft-saintandre-sip-xmpp-media-02, XMPP Standards Foundation, 2013.

Smoothness Measure for Image Fusion in Discrete Cosine Transform

Radhika Vadhi
Research Scholar, University College of Engineering
JNTUK, Kakinada A.P., India

Veera Swamy Kilari
Professor, QIS College of Engineering & Technology,
Ongole, A.P., India

Srinivas Kumar Samayamantula
Professor, University College of Engineering
JNTUK, Kakinada A.P., India

***Abstract*—The aim of image fusion is to generate high-quality images using information from source images. The fused image contains more information than any of the source images. Image fusion using transforms is more effective than spatial methods. Statistical measures such as mean, contrast, and variance, are used in Discrete Cosine Transform (DCT) for image fusion. In this paper, we use statistical measures, such as the smoothness of a block in the transform domain, to select appropriate blocks from multiple images to obtain a fused image. Smoothness captures important blocks in images and duly eliminates noisy blocks. Furthermore, we compare and analyze all statistical measures in the DCT domain. Experimental results establish the superiority of our proposed method over state-of-the-art techniques for image fusion.**

Keywords—smoothness; statistical measures; DCT; image fusion

I. Introduction

Today video surveillance is common in many public places such as hospitals, banks, offices, airports, military installations and other traffic control applications. The cameras are located in any corner of the place to capture the visuals. Proper visual information is capture in the center of the apex and at the edges and corners, where the camera located, have the poor quality of visuals due to the limited focal depth of optical lenses. Even though two or more cameras are located in various corners to capture depth, then also the problem of poor quality appears. To overcome this poor quality vision problem, Digital image fusion has been emerged.

Digital image fusion is a novel digital image processing technique that involves combining information from source images to form a single, final image. The fused image contains more relevant and accurate information than any of the source images, which are captured by visual sensors. The statistical measures involved are simple to compute, and play a vital role in identifying important information in images. The combination of transforms and statistical measures helps to identify vital information in the source images. Digital image fusion techniques can be classified into two domains: spatial domain and transform domain. Some researchers have proposed image fusion techniques that are implemented in the spatial domain [2, 3, 4, 5, 6, 9, 10, 11, 12, 24, 25, 26, 36, and 37]. Also the Fusion techniques based on multi scale decomposition are popular [7, 8, 13, 14, 15, 16, 18, 19, 20, 21, 22, and 23]. These involve grouping source images by observing a parameter called the activity level measure. The summation can be finalized by selecting coefficients with higher activity levels. The fused image is finally obtained by performing inverse multiscale operations. Discrete Wavelet Transform (DWT) is a multi-resolution transform used to fuse images. In the DWT domain, the maximum absolute value of the corresponding decomposed band coefficient at each position is selected as the activity level [20]. In image fusion using shearlet transforms [28], regional variance, and regional average gradients, regional spatial frequency is considered for high-frequency sub-band coefficients, and regional characteristics are used for low-frequency coefficients.

Most spatial domain image fusion techniques are complex and time consuming, and are hence not appropriate for real-time applications. In most communication, data is compressed prior to transmission, and images are coded in the JPEG format at present. Discrete Cosine Transform (DCT) is used in JPEG. Hence, we use the DCT domain in this study.

DCT-based methods, such as DCT + average [8], DCT + contrast [8], DCT + Variance1, DCT + Variance + Consistency Verification [1], are popular in the area. Prevalent DCT-based methods suffer from undesirable side-effects, such as blurring and blocking artifacts. In this work, we propose DCT + Smoothness to identify relatively less noisy blocks in source images for image fusion.

The rest of this paper is organized as follows: In Section 2, we describe smoothness calculation in the DCT domain, and the proposed fusion algorithm is detailed in Section 3. We discuss the results of experiments to test the performance of our proposed method in Section 4, and offer our conclusions in Section 5.

II. Smoothness Calculation in DCT Domain

Noise is ever presents in digital images during image acquisition, coding, transmission, and processing. Smoothness measures the relative smoothness of intensity in a region. It is high for a region of constant intensity, and low for regions with large excursions in the values of its intensity levels. Smoothness is a statistical method used to select relatively less noisy blocks in image fusion. Hence, smoothness algorithms tend to be superior in performance than others. Smoothness attempts to capture important patterns in an image while leaving out noisy blocks. Since our technique is implemented in the DCT domain, it saves time and computational complexity if the fused image needs to be stored or transmitted in the JPEG format. Smoothness is high when variations in AC coefficient values are low, and vice versa.

A two-dimensional (2D) DCT transform of an N × N block of an image $f(m,n)$ is defined as (1)

where k, l = 0, 1, ….., N – 1, and

$$\alpha(k) = \begin{cases} \frac{1}{\sqrt{2}} & if\ k = 0 \\ 1, & otherwise \end{cases} \tag{2}$$

In order to compute smoothness, the DC coefficient in Equation (1) needs to be eliminated to obtain results containing only AC coefficients because the DC coefficient is not useful in judging the smoothness of block

$$F(k,l) = \frac{2\alpha(k)\alpha(l)}{N} \times \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} f(m,n) \times \cos\left[\frac{(2m+1)\pi k}{2N}\right] \times \cos\left[\frac{(2n+1)\pi l}{2N}\right] \tag{3}$$

AC coefficients indicate variations in image blocks. The absolute value is considered to assign weight to all variations in AC coefficients. Hence, variations in the z block are computed as follows:

$$U(z) = \sum_{k,l} |F(k,l)| \qquad where \quad k \neq 0, l \neq 0 \tag{4}$$

In sum, the smoothness of the z block can be exactly calculated from its DCT coefficients by the absolute sum of the AC coefficients of the DCT block. Here, a high value of $U(z)$ indicates that smoothness is low, and vice versa.

III. Proposed Fusion Algorithm

The details of image fusion are shown in Fig. 1, which depicts the common framework of a JPEG encoder combined with our proposed image fusion method. This method can be extended to any number of source images.

Our proposed fusion algorithm is as follows:

1) *Consider two or more multi-focused source images.*
2) *Each source image is divided into sub-blocks*
3) *Apply 2D-DCT to each sub-block of each source image.*
4) *Variations in the DCT-transformed sub-bocks are calculated using Eq. (4).*
5) *The smoothness of each block is compared with that of a corresponding block from other source images.*
6) *The blocks with higher smoothness values are selected.*
7) *All sub-blocks are arranged into a single block.*
8) *2D- inverse DCT is applied to each sub-block of the fused image.*

The general fusion procedure is explained below. The source images are divided into sub-blocks, and 2D-DCT is applied to each sub-block. The statistical measure (smoothness) is calculated for all 2D-DCT sub-blocks, which are subsequently chosen based on smoothness values for fusion. Inverse DCT is applied to each sub-block to convert them into pixels.

IV. Experimental Results and Discussion

In this section, we describe the results of experiments on the proposed statistical measure for image fusion, and compare them with results from other important image fusion methods in the literature. The images used in our experiment are shown in Fig. 3

A. Measurement Criteria

Several objective evaluation methods are available to assess image fusion performance. Mutual Information (MI) [29, 32] is an important one used to test the quality of the fused image, and involves calculating information common to source images f_1, f_2 and fused image f_S. Edge Strength and Orientation Preservation (ESOP) ($Q^{f_1 f_2 / f_S}$) values can be calculated using Xydeas work [30, 33]. Lin et al. proposed a Feature Similarity Index method (FSIM), [31] and Normalized Cross-Correlation (NCC) [27] is another measurement parameter for image fusion. The NCC is computed between the mean of the ground truth image/ benchmark and that of the fused image.

B. Experimental Analysis

The algorithms were executed on six standard images shown in Fig. 3. In general, there is a problem in considering the images to be fused. We created out-of-focus or multi-focus images by blurring parts of the original image using low-pass filters. Blurring can be carried out by convolution with a Gaussian to reduce detail. The amount of blurring was considered in comparison with spatial frequency and visibility as in [26]. The original image and blurred images are shown in Fig. 4.

We used sub-blocks of size 8 × 8. Mutual Information (MI) is a measure used to test image quality using quantity of information. ESOP ($Q^{f1f2/fs}$) evaluates edge information, and FSIM is a metric for phase congruence and edge information between the source images and the fused image. The experimental results are listed in Table 1.

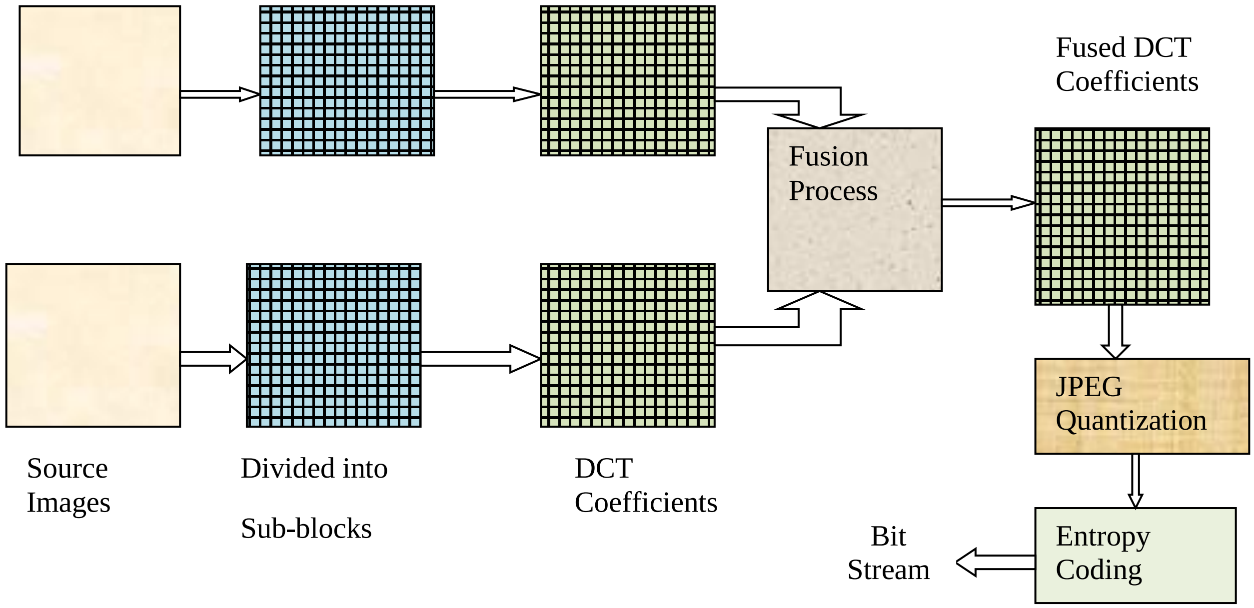


Fig. 1. General schematic representation of image fusion method

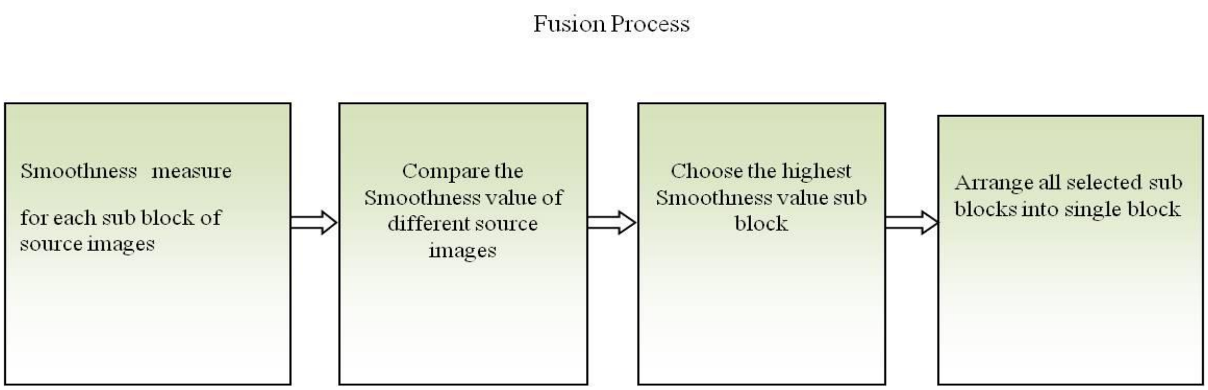


Fig. 2. Fusion process

Fig. 3. Images used in experiment

(a) (b) (c)

Fig. 4. Paper, (a) Original image (b) Left-side blurred image (c) right-side blurred image

TABLE I. OBJECTIVE ASSESSMENT OF IMAGE FUSION ALGORITHMS

	Fusion Rule	MI	$Q^{f_1 f_2 / f_S}$ (ESOP)	FSIM	NCC
Clock	DCT + Avg.	4.3933	0.9050	0.9997	0.9991
	DCT + Contrast	4.4910	0.9146	0.9997	0.9992
	DCT + Variance	4.5347	0.9149	0.9997	0.9992
	DCT+ Smoothness	**4.5940**	**0.9152**	**0.9997**	**0.9993**
Toy	DCT + Avg.	3.2999	0.0402	0.9998	0.9979
	DCT + Contrast	3.2803	0.8691	0.9996	0.9960
	DCT + Variance	3.5650	0.8748	0.9998	0.9987
	DCT +Smoothness	**3.8143**	**0.8785**	**0.9999**	**0.9991**
Disk	DCT + Avg.	3.3177	0.0354	0.9996	0.9974
	DCT + Contrast	3.9216	0.8958	0.9995	0.9978
	DCT + Variance	4.1189	0.9012	0.9997	0.9983
	DCT+ Smoothness	**4.1846**	**0.9039**	**0.9998**	**0.9987**
Pepsi	DCT + Avg.	3.8730	0.0501	0.9998	0.9988

	DCT + Contrast	3.9474	0.8995	0.9998	0.9986
	DCT + Variance	4.4756	0.9138	0.9999	0.9994
	DCT+ Smoothness	**4.5236**	**0.9148**	**0.9999**	**0.9995**
Paper	DCT + Avg.	2.9544	0.0197	0.9994	0.9843
	DCT + Contrast	3.1565	0.8497	0.9989	0.9688
	DCT + Variance	3.9129	0.8947	0.9997	0.9828
	DCT+ Smoothness	**3.9306**	**0.8950**	**0.9997**	**0.9928**
Lena	DCT + Avg.	3.7464	0.0477	0.9997	0.9973
	DCT + Contrast	3.9497	0.8925	0.9997	0.9979
	DCT + Variance	4.2831	0.8919	**0.9998**	0.9984
	DCT+ Smoothness	**4.3085**	**0.8928**	0.9987	**0.9999**
Cameraman	DCT + Avg.	3.6970	0.0264	0.9996	0.9973
	DCT + Contrast	3.7701	0.8440	0.9993	0.9973
	DCT + Variance	4.1361	0.8733	0.9997	0.9984
	DCT+ Smoothness	**4.1795**	**0.8748**	**0.9998**	**0.9986**
Woman	DCT + Avg.	3.6814	0.0336	0.9998	0.9963
	DCT + Contrast	3.7791	0.8720	0.9997	0.9941
	DCT + Variance	4.3559	0.8893	0.9999	0.9986
	DCT +Smoothness	**4.4202**	**0.8919**	**0.9999**	**0.9989**
F17	DCT + Avg.	3.5011	0.0205	0.9997	0.9978
	DCT + Contrast	3.5892	0.8983	0.9996	0.9974
	DCT + Variance	3.8001	0.8987	0.9998	0.9981
	DCT+ Smoothness	**3.8786**	**0.9017**	**0.9998**	**0.9986**
Fishingboat	DCT + Avg.	3.2273	0.0244	0.9998	0.9985
	DCT + Contrast	3.2305	0.8857	0.9996	0.9982
	DCT + Variance	3.4229	0.8889	0.9998	0.9990
	DCT+ Smoothness	**3.5059**	**0.8934**	**0.9999**	**0.9994**
Mandrill	DCT + Avg.	3.4635	0.0554	0.9995	0.9909
	DCT + Contrast	3.5544	0.8765	0.9993	0.9806
	DCT + Variance	4.4831	0.9206	0.9998	0.9963
	DCT+ Smoothness	**4.4875**	**0.9205**	**0.9998**	**0.9962**
Livingroom	DCT + Avg.	3.2742	0.0341	0.9996	0.9933
	DCT + Contrast	3.4754	0.8553	0.9994	0.9894
	DCT + Variance	4.2227	0.8830	0.9997	0.9966
	DCT+ Smoothness	**4.2728**	**0.8844**	**0.9998**	**0.9969**
Pirate	DCT + Avg.	3.5956	0.0335	0.9997	0.9963
	DCT + Contrast	3.9354	0.8598	0.9996	0.9949
	DCT + Variance	4.5513	0.8891	0.9998	0.9982
	DCT+ Smoothness	**4.6094**	**0.8910**	**0.9999**	**0.9985**
Peppers	DCT + Avg.	4.3024	0.0238	0.9999	0.9994
	DCT + Contrast	3.2305	0.8996	0.8857	0.9982
	DCT + Variance	4.5821	0.9193	0.9998	0.9994

	DCT + Smoothness	**4.6808**	**0.9222**	**0.9999**	**0.9996**

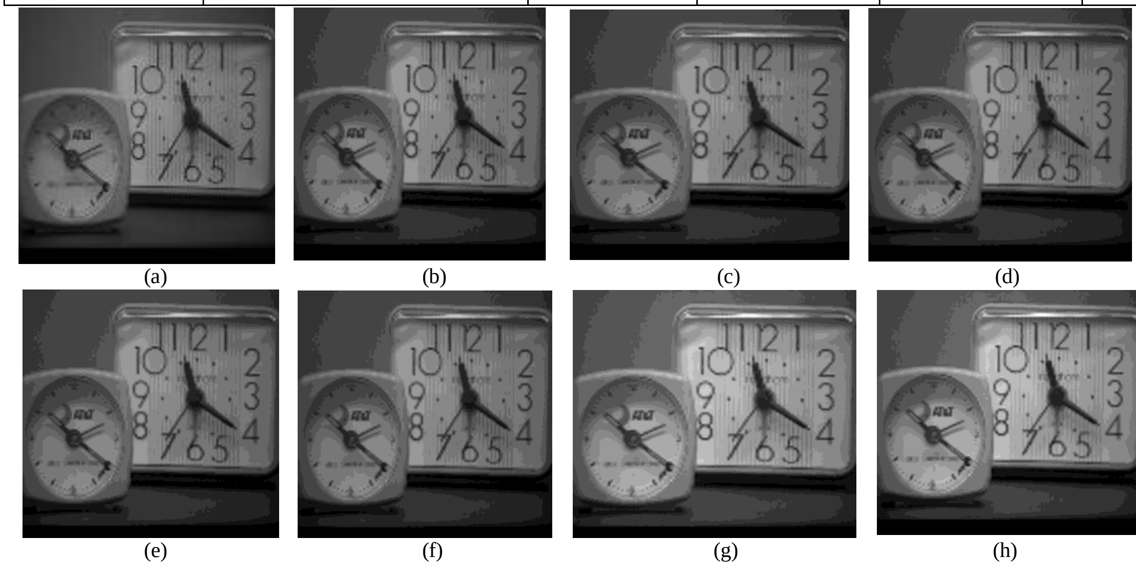


Fig. 5. "Clock." (a) Ground truth. (b) Left-blurred image. (c) Right-blurred image. (d) DCT + Avg. (e) DCT + Contrast. (f) DCT + Variance. (g) DCT + Smoothness. (h) Shearlet transform

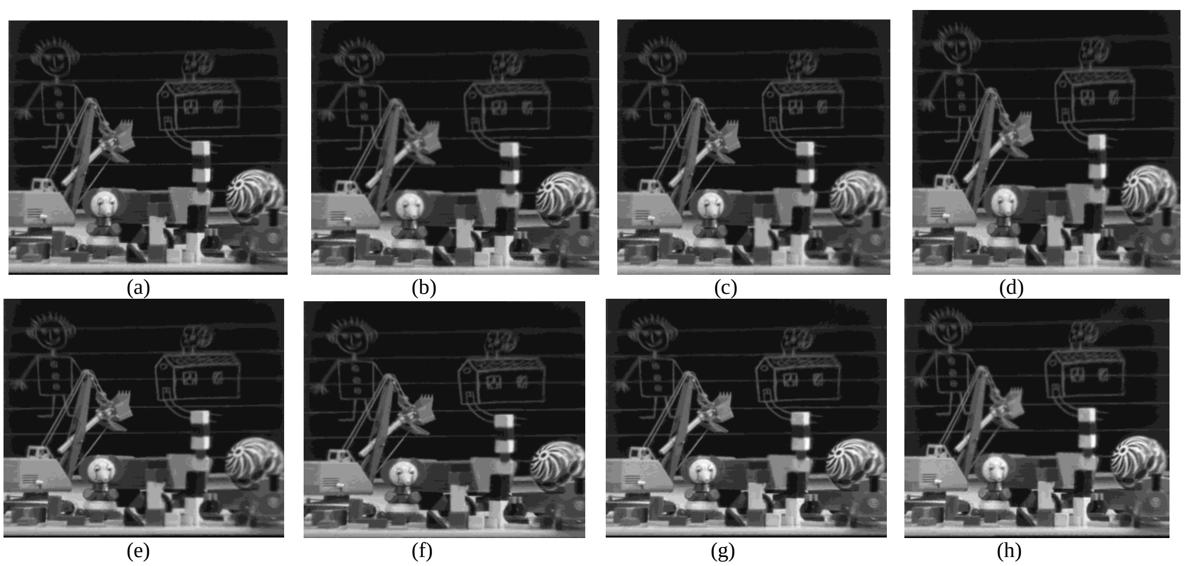


Fig. 6. "Toy." (a) Ground truth. (b) Left-blurred image. (c) Right-blurred image. (d) DCT + Avg. (e) DCT + Contrast. (f) DCT + Variance. (g) DCT + Smoothness. (h) shearlet transform

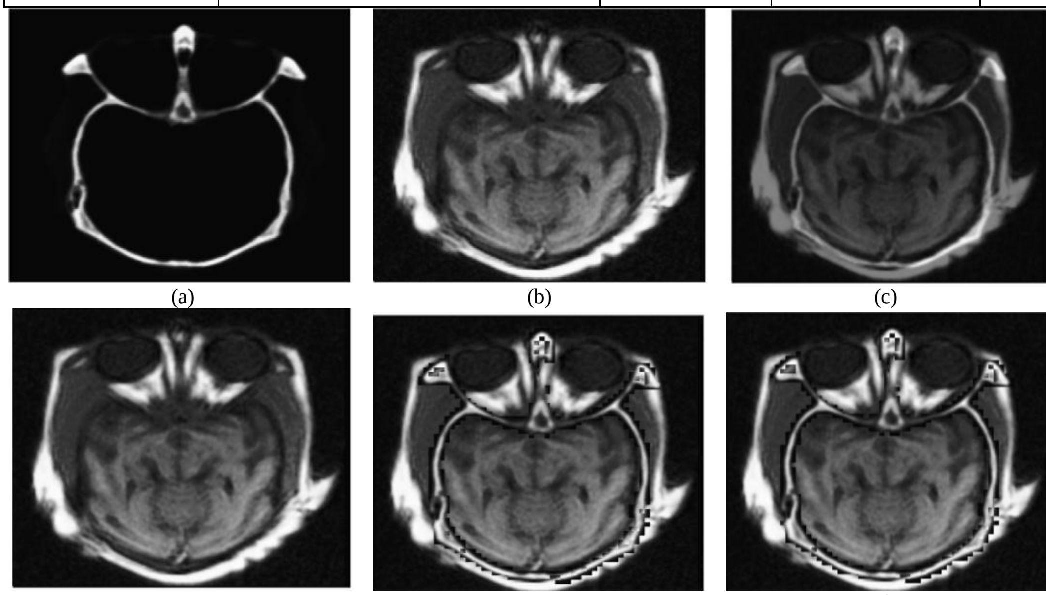
TABLE II. RUNTIME VALUES OF VARIOUS ALGORITHMS FOR "CLOCK."

DCT + Avg.	DCT + Con.	DCT + Var.	DCT + Smoothness
4.741602	6.236915	4.650012	4.530772

TABLE III. EXPERIMENTAL RESULTS (DCT AND SHEARLET TRANSFORM)

	Fusion Rule	MI	$Q^{f_1 f_2 / f_S}$ (ESOP)	FSIM	NCC
Clock	DCT+ Smoothness	**4.5940**	**0.9152**	**0.9997**	**0.9993**
	shearlet transform	4.49190	0.9010	0.9943	0.9992
Toy	DCT+ Smoothness	**3.8143**	**0.8785**	**0.9999**	**0.9991**

	shearlet transform	3.0648	0.8602	0.9862	0.9973
Disk	DCT+ Smoothness	**4.1846**	**0.9039**	**0.9998**	**0.9987**
	shearlet transform	3.2895	0.8877	0.9875	0.9973
Pepsi	DCT+ Smoothness	**4.5236**	**0.9148**	**0.9999**	**0.9995**
	shearlet transform	3.9992	0.8901	0.9919	0.9974
Paper	DCT+ Smoothness	**3.9306**	**0.8950**	**0.9997**	**0.9928**
	shearlet transform	2.5694	0.8539	0.9828	0.9797
Lena	DCT+ Smoothness	**4.3085**	**0.8928**	**0.9987**	**0.9999**
	shearlet transform	3.8757	0.8826	0.9915	0.9976



(a) (b) (c)

(d) (e) (f)

Fig. 7. “Medical images” (a) CT (b) MRI (c) DCT+ Avg (d) DCT + Contrast (e) DCT + Variance (f) DCT+ Smoothness

(a) (b) (c)

Fig. 8. Clock, (a) Original image (b) Noise image1 (c) Noise image 2

TABLE IV. EXPERIMENTAL RESULTS (FOR NOISE IMAGES)

	MI	ESOP	FSIM	NCC
DCT+ Avg	3.9190	0.9555	0.9992	0.9939
DCT+ Con	3.9188	0.9528	0.9992	0.9942
DCT+ Var	3.9180	0.9536	0.9990	0.9935
DCT+ Smoothness	**3.9194**	**0.9573**	**0.9993**	**0.9944**

The results in Table 1 show that MI, FSIM, and NCC improved with our proposed DCT + Smoothness approach. We can see that DCT + Smoothness is competent than the other DCT-based methods. The amount of blurriness was created by a [4 4] Gaussian filter, and its standard deviation is 6. The experimental results for the “clock” and “toy” images are shown in Fig. 5 and Fig. 6, respectively.

DCT + Avg., DCT + Contrast, and DCT + Variance are the existed algorithms. By carefully observing the fusion results, it is concluded that the method DCT + Average results blurring the fused image (Fig.5.d). The method DCT + Contrast and DCT+ Variance results some blocking artifacts (Fig.5.e and Fig.5.f). There are ringing artifacts for the shearlet transform based fusion method (Fig.5.h).The time taken by each method to perform the fusion operation is shown in Table 2. The runtime of the proposed DCT + Smoothness was shorter than that of existing methods for image fusion. All algorithms were executed on a Pentium IV processor with 3GHz ROM, and 504 MB of Random-Access Memory (RAM). The operating system used was Windows XP Professional 2002.

We also compared the image fusion performance of our method with that of the shearlet transform proposed by Liu and Wang[28]. Low-frequency sub-band coefficients were processed based on the energy of each sub-band, and high-frequency sub-band coefficients were processed based on the variance of each sub-band.

The results indicate that the smoothness measure in the DCT domain is suitable for selecting blocks for image fusion. This yields better results than the variance characteristic in the shearlet domain. The comparisons between DCT+ Smoothness and Shearlet transforms are given in Table3. The proposed algorithm is also tested on noise images and naturally acquired images. The experimental results of CT, MRI are given in Fig.7. Experiments are also performed on Noise images. We created noise images by considering impulse noise with different densities. The original image and noise images are given in Fig.8.

V. Conclusions

In this paper, we proposed a method for image fusion that used smoothness as a statistical measure in the DCT domain, and experimentally compared it with other methods that employ different statistical measures. The experiments established the superiority of our smoothness-based measure in the DCT domain in terms of complexity and execution time. Our method was also superior when compared with multi-resolution transform-based image fusion methods. It is thus more appropriate for real-time applications.

Aknowledgements

Thank you to the www.fusion org and http://utopia.duth.gr/~nmitiano/fusion.html; http://www.hindawi.com/journals/mpe/2009/128317/fig10/; http://www.ece.lehigh.edu/SPCRL/IF/disk.htm

http://utopia.duth.gr/~nmitiano/fusion.html]

https://www.pantechsolutions.net/basics-of-image-fusion] for providing the source images.

References

[1] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, “Multi-focus image fusion for visual sensor networks in DCT domain,” *Computers & Electrical Engineering.* 37(5), 789-797 (2011).

[2] A. A. Goshtasby and S. Nikolov, “Image fusion: Advances in the state of the art,” *Information Fusion*. 8(2), 114–118 (2007).

[3] V. S. Petrovic and C. S. Xydeas, “Gradient-based multiresolution image fusion,” *IEEE Transactions on Image Processing*. 13(2), 228-237 (2004).

[4] T. Stathaki, *Image fusion algorithms and applications*, Academic Press, (2008).

[5] N. Mitianoudis and T. Stathaki, “Pixel-based and region-based image fusion schemes using ICA bases,” Information Fusion. 8(2), 131-142 (2007).

[6] H. Li, B. Manjunath and S. Mitra, “Multisensor image fusion using the wavelet transform,” Graphical Models and Image Processing. 57(3), 235-245 (1995).

[7] J. Tang, “A contrast based image fusion technique in the DCT domain,” Digital Signal Processing. 14(3), 218-226 (2004).

[8] G. Piella, “A general framework for multiresolution image fusion: from pixels to regions,” Information Fusion. 4 (4), 259-280 (2003).

[9] Y. Xia and H. Leung, “A fast learning algorithm for blind data fusion using a novel L2-norm estimation,” IEEE Sensors Journal. 14(3), 666-672 (2014).

[10] M. B. A. Haghighat, Ali Aghagolzadeh, and Hadi Seyedarabi, “Real-time fusion of multi-focus images for visual sensor networks,” 6th Iranian Conference on Machine Vision and Image Processing (MVIP). 1-6, (2010).

[11] H. Liu, J. Yang, Z. Wu, and Q. Zhang, “Fast single image dehazing based on image,” Journal of Electronic Imaging. New. 24(1), 013020-013020 (2015).

[12] X.q. Luo, Z.C. Zhang, and X.j. Wu, “Adaptive multistrategy image fusion method,” Journal of Electronic Imaging.23 (5), 053011 (2014).

[13] O. Rockinger, “Image sequence fusions using a shift-invariant wavelet transform,” Proceedings of IEEE International Conference on Image Processing. 3, 288-291 (1997).

[14] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, Digital Image Processing Using MATLAB. Low price edition, 2002.

[15] A. Saleem, A. Beghdadi, and B. Boashash, “Image fusion-based contrast enhancement,” EURASIP Journal on Image and Video Processing. 1-17 (2012).

[16] B. Yang and S. Li, “Pixel-level image fusion with simultaneous orthogonal matching pursuit,” Information Fusion. 13(1), 10-19 (2012).

[17] H. Li, B. S. Manjunath, and S. K. Mitra. “Image Fusion Using the Wavelet Transform,” Proc. First International Conference on Image Processing ICIP 94, Austin, Texas. 1, 51-55 (1994).

[18] D. Drajic and N. Cvcjic, “Adaptive fusion of multi-model surveillance image sequences in visual sensor networks,” IEEE Trans Consum Electron. 8(2), 119-30 (2007).

[19] G. Bhatnagar and B. Raman, “A new image fusion technique based on directive contrast,” Electron Letter on Computer Vision and Image Analysis. 8(2), 18-38 (2009).

[20] Y. B.-Shoshan and Y. Yitzhaky, “Improvements of image fusion methods,” Journal of Electronic Imaging. 23(2), 023021 (2014).

[21] Q. Yuan, L. Zhang, and H. Shen, “Hyperspectral image denoising with a spatial–spectral view fusion strategy,” IEEE Transaction on Geoscience and Remote Sensing. 52(5), 2314 – 2325 (2014).

[22] T. Peli and E. Peli, “Contrast in complex images,” J. Op. Soc. Am. A 7. 2030-2040 (1990).

[23] R. Maruthi and Sankarasubramanian, “Multi-focus image fusion based on the information level in the regions of the images” Journal of Theoretical and Applied information Technology. 3(4), 80-85 (2007).

[24] V. Radhika, V. Swamy Kilari, and S. Kumar Samayamantula, “Uniform-based approach for image fusion,” ICECCS-2012, CCIS305, Springer-Verlag Berlin Heidelberg. 186-194 (2012).

[25] X. Liu and J. Wang, "Image fusion based on shearlet transform and regional features," International Journal of Electronics and Communications (AEÜ). 68(6), 1-7 (2013).

[26] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," Computers & Electrical Engineering. 37(5), 744 -756 (2011).

[27] G. Piella and Heijmans, "New quality measures for image fusion," Polytechnical University of Catalonia (UPC). Jordi Girona, 08034 Barcelona, Spain. 1-3 (2003).

[28] L. Zhang and X. Mou, "FSIM: A feature similarity index for image quality assessment," IEEE transactions on Image Processing. 20(8), 2378-2386 (2011).

[29] G. H. Qu and D. L. Zhang, "Information measure for performance of image," Electronic Letters. 38(7), 313-315 (2002).

[30] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," Electronic Letters. 36(4), 308-309 (2000).

The Factors of Subjective Voice Disorder Using Integrated Method of Decision Tree and Multi-Layer Perceptron Artificial Neural Network Algorithm

Haewon Byeon
Department of Speech Language Pathology
Nambu University, Gwangju, Republic of Korea

Sunghyoun Cho*
Department of Physical Therapy
Nambu University, Gwangju, Republic of Korea

***Abstract*—The aim of the present study was to develop a prediction model for subjective voice disorders based on an artificial neural network algorithm and a decision tree using national statistical data. Subjects of analysis were 8,713 adults over the age of 19 (3,801 males and 4,912 females) who completed the otolaryngological examination of the Korea National Health and Nutrition Examination Survey from 2010 to 2012. Explanatory variables included age, education level, income, occupation, problem drinking, coffee consumption, and pain and discomfort from disease over the last two weeks. A multi-layer perceptron artificial neural network and a decision tree model were used for the analysis. In this model, smoking, pain and discomfort from disease over the last two weeks, education level, occupation, and income were drawn out as major predictors of subjective voice disorders. In order to minimize the risk of dysphonia, it is necessary to establish a scientific management system for high-risk groups.**

***Keywords*—Neural Networks; Subjective Voice Disorder; decision tree; risk factor; data-mining**

I. Introduction

Voice disorders refer to problems of the voice due to abnormalities in the structure, function, or vagus nerve of vocal cords, and the term comprises both laryngeal disorders and subjective voice problems [1]. The prevalence rate of voice disorders is 5–7% [2, 3], and 30% of the community-dwelling population experience voice problems at least once in their lifetime [4]. Thus, it is presumed that among the total population of 50 million Koreans, more than 2.5 million Koreans suffer from voice disorders.

The measurement and assessment of voice disorders are classified into objective methods, such as acoustic and aerodynamic tests, and patients' subjective assessments of voice symptoms [5]. Preliminary tests for the diagnosis of voice disorders include the acoustic test and the laryngoscopy test, but it is difficult to detect functional dysphonia caused by psychological problems with these tests alone. Therefore, capturing the subjective perceptions of a voice problem that a subject reports as well as observing the voice problems objectively is important in the diagnosis of dysphonia [6]. In addition, although there are (individual) differences in the degree of self-recognition of voice problems depending on the characteristics of individuals, subjective voice problems play an important role in deciding the treatment of dysphonia [7].

Nevertheless, in many cases, subjects do not take voice problems seriously and, thus, do not report them to medical professionals. Moreover, even when subjects do report, diagnoses are terminated if no problems are discovered on the objective tests.

In order to effectively prevent dysphonia, investigations of its risk factors are vital. Over the last 20 years, smoking, drinking, misuse and abuse of vocal cords, and occupation have been reported to be the major risk factors of dysphonia [8–12]. Among them, subjective voice problems have been verified to be not only an independent risk factor of voice disorders [3, 4, 13, 14] but also the most predominant predictor among the various risk factors of dysphonia [15].

To date, numerous epidemiological studies have confirmed that subjective voice disorders are an independent risk factor of dysphonia [16]. However, it has not yet been verified as a risk factor in Korean adults.

Meanwhile, artificial neural network analysis, an analysis technique that increases problem-solving abilities through the learning of artificial neurons, is currently being widely used in classification and prediction. Inspired by the human brain, neural network analysis has several advantages [17]. For example, past experiences enable automatic learning, the analysis of qualitative and quantitative variables is possible, and it has excellent prediction power, as non-linear combinations among entered variables are possible.

Furthermore, the decision tree model, which displays the decision-making process in a tree-structure diagram, has the advantages of accommodating both continuous and categorical variables and enabling an understanding of the factors contributing to the dependent variable.

In order to determine the characteristics of high-risk groups for dysphonia, it is necessary to elucidate the complex factors that affect voice problems. Decision trees, widely used in the areas of pattern recognition and medical science, and data mining analysis, used in neural networks, can be effectively used in predicting the target group for dysphonia prevention programs [15].

This study presents basic materials to reduce dysphonia by developing a prediction model for subjective voice disorders based on an artificial neural network algorithm and a decision tree using national statistical data. The organization of our study is as follows: Section 2 describes the data resources, Section 3 explains the procedure for the development of the prediction model, and Section 4 presents the results of the developed prediction model.

II. METHODS

A. Data sources

This study analyzed 8,713 adults (3,801 males and 4,912 females) who completed the otolaryngological examination of the Korea National Health and Nutrition Examination Survey from 2010 to 2012. The Korea National Health and Nutrition Examination Survey is a nationwide health survey conducted by the Ministry of Health and Welfare on 11,520 households regarding education, economic activities, contraction of diseases, use of medical institutions, and health behaviors [18]. Education and economic activities were researched by individual face-to-face interviews, while health behaviors, such as smoking and drinking, were researched by self-administered questionnaires. The detailed methods of the research are specified in the preceding study [18].

B. Measurements

The dependent variable, subjective voice disorders, was classified (yes, no) based on the answers to the otolaryngological question: “Do you think that you have an abnormality in your voice?”

Explanatory variables included age, education level, income (quartile), occupation, problem drinking, coffee consumption, and pain and discomfort from disease over the last two weeks (yes, no). Age was classified as 19 to 39 years, 40 to 59 years, and more than 60 years old. Education levels were classified as below elementary school graduation, middle school graduation, high school graduation, and above college graduation. Occupations were classified as follows: economically inactive person, non-manual worker (e.g., managers & professionals, clerical support workers, service & sales workers), manual worker (skilled agricultural & forestry & fishery workers, craft & plant and machine operators and assemblers, and unskilled laborers). As for alcohol consumption, 8 points and over was classified as problem drinking by using Alcohol Use Disorders Identification Test in Korea (AUDIT-K)[19].

III. STATISTICAL ANALYSIS

A. Artificial neural network

Factors potentially related to subjective voice disorders were analyzed by using an artificial neural network. Artificial neural network analysis is a data mining modeling technique that finds hidden patterns from actual data through a repetitive learning process imitating the neural network of the human brain. It is a nonlinear model that is used to solve prediction problems in data with complex structures [20].

Artificial neural network analysis is a mathematical model composed of numerous processing factors with a hierarchical structure, and it learns the relationship between input and output by the repetitive adjustment of weights by comparing past input data values and corresponding output data values [21]. The structure of the neural network is composed of both an input layer made of nodes corresponding with input variables and a hidden layer made of multiple hidden nodes. The hidden nodes turn the linear combination of variable values delivered from the input layer into a nonlinear function and deliver it to the input layers or other hidden layers [22].

This study used a Radial Basis Function (RBF) neural network [23]. This study regarded variables with relative importance of inputs over 0.1 as major explanatory variables that affect the dependent variable and thus included them in the decision tree model.

B. Classification and regression tree algorithm

The decision tree model was established by using the Classification and Regression Tree (CART) algorithm. CART is an algorithm based on binary classification that measures impurities by using the Gini Index, and in it, only two children nodes are formed from a parent node [24]. The Gini Index refers to the probability that two elements randomly extracted from *n* elements belong to different groups from each other [25]. The alpha value for the criteria of splitting and merging was set at 0.05. The number of parent nodes was 200 and that of child nodes was 100, and the number of branches was limited to five. The validity of the developed model was assessed with the 10-fold cross-validation method.

IV. RESULTS

A. General characteristics of study subjects

The general characteristics of the study subjects are presented in Table 1. Out of the total 8,713 subjects, the prevalence of subjective voice disorders was 6.9% (*n*=602). According to the result of the chi-square test, healthy subjects and those with subjective dysphonia did not have significant differences in any of the variables.

TABLE I. GENERAL CHARACTERISTICS OF THE STUDY SUBJECTS, N (%)

Characteristics	**Subjective voice disorder**		***p***
	No (n=8,111)	***Yes (n=602)***	
Age			0.79
19-39	2,485 (93.0)	188 (7.0)	
40-59	2,050 (93.3)	218 (6.7)	
60 ⩽	2,576 (92.9)	196 (7.1)	
Sex			0.75
Male	3,542 (93.2)	259 (6.8)	
Female	4,569 (93.0)	343 (7.0)	
Education			0.31
Elementary school	1,885 (93.1)	139 (6.9)	
Middle school	835 (92.7)	66 (7.3)	

High school	2,444 (92.7)	192 (7.3)	
Collage	2,267 (93.9)	146 (6.1)	
Income			0.35
First quartile	1,579 (93.5)	109 (6.5)	
Second quartile	2114 (92.9)	162 (7.1)	
Third quartile	2,138 (92.4)	177 (7.6)	
Fourth quartile	2,150 (93.6)	148 (6.4)	
Occupation			0.46
economically inactive person	3,067 (93.3)	221 (6.7)	
non-manual worker	2,470 (92.8)	193 (7.2)	
manual worker	1,874 (93.7)	127 (6.3)	
Problem drinking			0.91
No	4,263 (93.1)	317 (6.9)	
Yes	2,095 (93.2)	154 (6.8)	
Smoking			0.53
Non-smoking	4,323 (93.4)	306 (6.6)	
Past smoking	1,547 (92.6)	124 (7.4)	
Current smoking	1,561 (93.2)	114 (6.8)	
Pain & discomfort from disease for the recent 2 weeks			0.59
No	5,734 (93.3)	412 (6.7)	
Yes	1,704 (92.8)	133 (7.2)	

B. *Factors potentially related to subjective voice disorders*

As a result of the artificial neural network analysis on 60.6% of the training sample, 29.6% of the test sample and 9.8% of the verification sample, five hidden layers were drawn out that produced the smallest data errors. The sum of the square error was 7.2%, and the classification accuracy of the training sample, test sample, and verification sample proved to be 92.8%, 93.8%, and 92.6%, respectively.

A synaptic weighted network diagram of the neural network model is presented in Figure 1. The synaptic weighted value in the network diagram demonstrates the relationship among layers, and the higher the combined weighted value, the thicker the line between layers. In this model, smoking, pain and discomfort from disease over the last two weeks, education level, occupation, and income were drawn out as major variables with high weighted values for subjective voice disorders.

The normalized importance sampling estimator drawn out from the neural network model is presented in Figure 2. According to the result of the normalized importance sampling estimator, smoking, pain and discomfort from disease over the last two weeks, education level, occupation, and income were deciding factors of subjective voice disorders.

The prediction model for subjective voice disorders using the CART algorithm is presented in Figure 3. According to the result of the classification model constructed using the CART algorithm, the most preferentially involved predictor was household income.

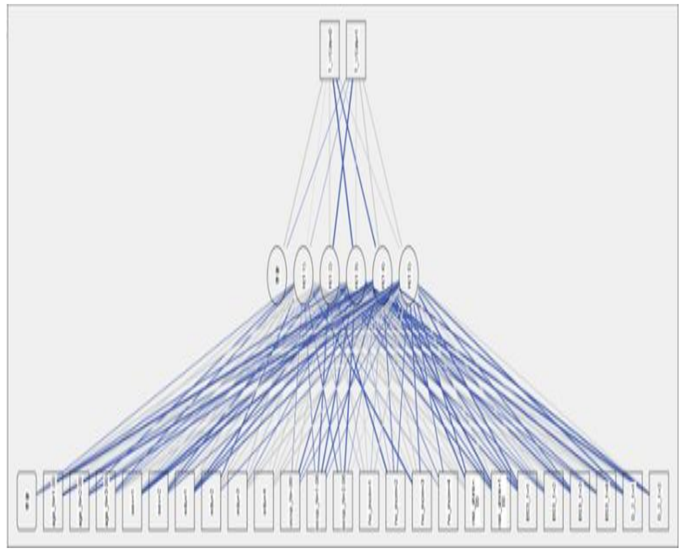


Fig. 1. Synaptic weighted network diagram

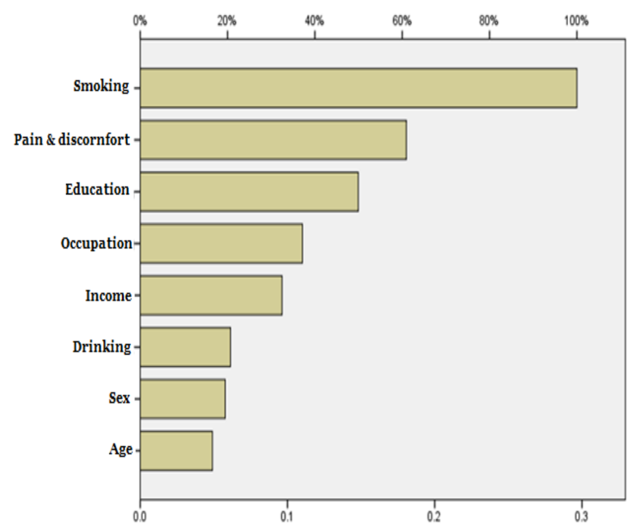


Fig. 2. The normalized importance sampling estimator drawn out from the neural network model

Table 2 is a gains chart of the final prediction model for subjective voice disorders created by the CART algorithm. The two nodes were confirmed as significant paths that effectively predicted subjective voice disorders.

The first path with the greatest profit index value (10th node) was "past smokers who belong to the lowest 25% in household income and are middle school graduates and lower or university graduates and over." Of them, 13.5% were classified as suffering from subjective dysphonia, and their profit index was 204.8%. The second path (16th node) was "manual workers who belong to the lowest 25% in household income and are primary school graduates and lower or high school graduates with past smoking experience." Of them, 11.1% were classified as suffering from subjective dysphonia, and their profit index was 169.1%.

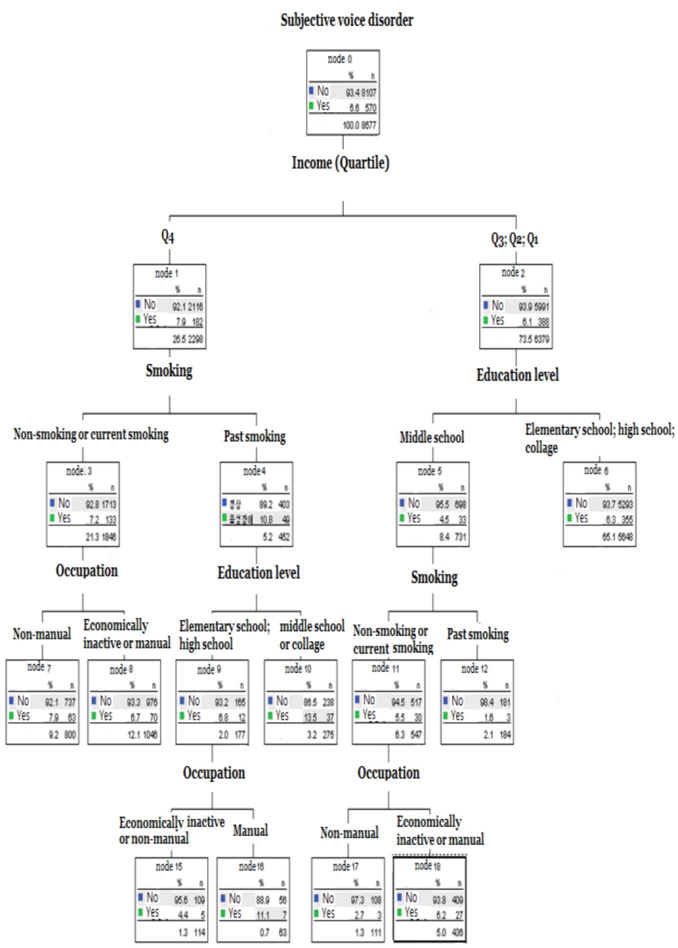


Fig. 3. The prediction model for subjective voice disorders using the CART algorithm

This study conducted a 10-fold cross-validation test to assess the final developed model. According to the result of the 10-fold cross-validation test to compare the stability of the drawn-out model, the drawn-out risk index was 0.32 and the misclassification rate was 32% for the cross-classification model, showing the same risk index (0.32) and misclassification rate (32%) of the prediction model.

TABLE II. GAINS CHART OF THE FINAL PREDICTION MODEL FOR SUBJECTIVE VOICE DISORDERS CREATED BY THE CART ALGORITHM

Node no	Node n (%)[1]	Gain n (%)[2]	Response %[3]	Gain Index %[4]	Group
10	275 (3.2)	37 (6.5)	13.5	204.8	Past smokers who belong to the lowest 25% in household income and are middle school graduates and lower or university graduates and over
16	63 (0.7)	7 (1.2)	11.1	169.1	Manual workers who belong to the lowest 25% in household income and are primary school graduates and lower or high school graduates with past smoking experience

[1] Node n(%); node number, % to 8,713
[2] Gain n(%); gain number, % to 602
[3] Response (%): The fraction of the subjective voice disorder
[4] Gain index (%):= 204.8 in total 7 node

V. CONCLUSION

This study analyzed predictors of subjective dysphonia by combining neural network analysis and a decision tree model.

In our study, "past smokers who belong to the lowest 25% in household income and are middle school graduates and lower or university graduates and over" and "manual workers who belong to the lowest 25% in household income and are primary school graduates and lower or high school graduates with past smoking experience" were high-risk groups for subjective dysphonia.

In order to prevent dysphonia effectively, it is necessary to develop prevention programs that consider the characteristics of high-risk groups.

ACKNOWLEDGMENT

The authors wish to thank the Korea Centers for Disease Control and Prevention that provided the raw data for analysis.

REFERENCES

[1] D. R. Boone, S. C. McFarlane, S. L. Von Berg, and R. I. Zraick, The voice and voice therapy. Boston, Allyn & Bacon Communication Sciences and Disorders, 2013.

[2] H. Byeon, Gender differences in risk factors of benign vocal fold disease in Korea: the fifth Korea National Health and Nutritional Examination Survey. Logoped Phoniatr Vocol, E-pub, doi:10.3109/14015439.2015.1004365, 2015.

[3] N. Roy, R. M. Merrill, S. D. Gray, and E. M. Smith, Voice disorders in the general population: prevalence, risk factors, and occupational impact. Laryngoscope, vol. 115, no. 11, pp. 1988-1995, 2005.

[4] S. M. Cohen, Self-reported impact of dysphonia in a primary care population: an epidemiological study. Laryngoscope, vol. 120, no. 10, pp. 2022-2032, 2010.

[5] A. E. Aronson, and D. Bless, Clinical voice disorders. New York, Thieme Medical Publishers, 2011.

[6] E. Vilkman, Voice problems at work: a challenge for occupational safety and health arrangement. Folia Phoniatrica et Logopaedica, vol. 52, no. 1-3, pp. 120-125, 2000.

[7] A. Behrman, L. Sulica, and T. He, Factors predicting patient perception of dysphonia caused by benign vocal fold lesions. Laryngoscope, vol. 114, no. 10, pp. 1693-1700, 2004.

[8] P. N. Carding, S. Roulstone, K. Northstone, and ALSPAC Study Team, The prevalence of childhood dysphonia: a cross-sectional study. Journal of Voice, vol. 20, no. 4, pp. 623-630, 2006.

[9] N. R. Williams, Occupational groups at risk of voice disorders: a review of the literature. Occupational Medicine, vol. 53, no. 7, pp. 456-460, 2003.

[10] S. L. Thibeault, R. M. Merrill, N. Roy, S. D. Gray, and E. M. Smith, Occupational risk factors associated with voice disorders among teachers. Annals of Epidemiology, vol. 14, no. 10, pp. 786-792, 2004.

[11] H. Byeon, Relationships among smoking, organic, and functional voice disorders in Korean general population. Journal of Voice, vol. 29, no. 3, pp. 312-316, 2015.

[12] H. Byeon, A population-based cross-sectional study of alcohol consumption and risk of benign laryngeal disease in Korean adults. Journal of Voice, E-Pub: doi:10.1016/j.jvoice.2014.10.014, 2016.

[13] H. Byeon, Prevalence of perceived dysphonia and its correlation with the prevalence of clinically diagnosed laryngeal disorders: the Korea national health and nutrition examination surveys 2010-2012. Annals of Otology, Rhinology & Laryngology, vol. 124, no. 10, pp. 770-776, 2015.

[14] N. Roy, J. Stemple, R. M. Merrill, and L. Thomas, Epidemiology of voice disorders in the elderly: preliminary findings, Laryngoscope, vol. 117, no. 4, pp. 628-633, 2007.

[15] H. Byeon, The risk factors of laryngeal pathology in Korean adults using a decision tree model. Journal of Voice, vol. 29, no. 1, pp. 59-64, 2015.

[16] J. H. Hah, S. Sim, S. Y. An, M. W. Sung, and H. G. Choi, Evaluation of the prevalence of and factors associated with laryngeal diseases among the general population. Laryngoscope, vol. 125, no. 11, pp. 2536-2542, 2015.

[17] R. C. Eberhart, Neural network PC tools: a practical guide. San Diego, Academic Press, 2014.

[18] Ministry of Health and Welfare, Korea National Health and Nutrition Examination Survey 2010-2012. Seoul, Ministry of Health and Welfare, 2014.

[19] J. S. Kim, M. K. Oh, B. K. Park, M. K. Lee, and G. J. Kim, Screening criteria of alcoholism by alcohol use disorders identification test(AUDIT) in Korea. Journal of the Korean Academy of Family Medicine, vol. 20, no. 9, pp. 1152-1159, 1999.

[20] V. Yashchenko, Multidimensional neural-like growing networks: a new type of neural network. International Journal of Advanced Computer Science and Applications, vol. 6, no. 4, pp. 1-10, 2015.

[21] M. K. Luka, I. A. Frank, and G. Onwodi, Neural network based Hausa language speech recognition. International Journal of Advanced Research in Artificial Intelligence, vol. 1, no. 2, pp. 39-44, 2012.

[22] A. G. Eldin, A data mining approach for the prediction of hepatitis C virus protease cleavage sites. International Journal of Advanced Computer Science and Applications, vol. 2, no. 12, pp. 179-182, 2011.

[23] M. A. Jayaram, G. K. Prashanth, and M. Anusha, Indexing of ears using radial basis function neural network for personal identification. International Journal of Advanced Computer Science and Applications, vol. 6, no. 7, pp. 28-33, 2015.

[24] H. Byeon, Development of prediction model for endocrine disorders in the Korean elderly using CART algorithm. International Journal of Advanced Computer Science and Applications, vol. 6, no. 9, pp. 125-129, 2015.

[25] H. Byeon, A prediction model for mild cognitive impairment using random forests. International Journal of Advanced Computer Science and Applications, vol. 6, no. 12, pp. 8-12, 2015.

SSH Honeypot: Building, Deploying and Analysis

Harry Doubleday, Leandros Maglaras, Helge Janicke
School of Computer Science and Informatics
De Montfort University, Leicester, UK

***Abstract*—This article is set to discuss the various techniques that can be used while developing a honeypot, of any form, while considering the advantages and disadvantages of these very different methods. The foremost aims are to cover the principles of the Secure Shell (SSH), how it can be useful and more importantly, how attackers can gain access to a system by using it. The article involved the development of multiple low interaction honeypots. The low interaction honeypots that have been developed make use of the highly documented libssh and even editing the source code of an already available SSH daemon. Finally the aim is to combine the results with the vastly distributed Kippo honeypot, in order to be able to compare and contrast the results along with usability and necessity of particular features. Providing a clean and simple description for less knowledgeable users to be able to create and deploy a honeypot of production quality, adding security advantages to their network instantaneously.**

Keywords—SSH Honeypot; Cyber Security

I. Introduction

There has been a variety of honeypots previously developed to work using the SSH protocol. The aim of this article is not to build software that can better these is every way, but more of a focus on a quick, simple, yet effective alternative to the pre-built packages available as well as providing a piece of software that can be available to professionals and unenlightened server users en masse. A honeypot is a wittingly vulnerable piece of software or system that is often used to emulate a service, system or network. The advantages of honeypots are that they are intentionally exposed in particular ways. The ruse and falsification used in honeypots is to hopefully entice attackers, which can be harder than it may seem as most attackers with some sort of knowledge, not a 'script kiddie', will soon realise that they are not in a real system when they try to run certain commands or processes that the honeypot doesn't understand. The results from different types of honeypots often vary significantly in depth, which will be further discussed in the results section of this document. Authors in [1] state that, a honeypot should be available to be attacked, as a security resource it has no value or purpose when it is not probed, attacked or compromised. The results that are produced from honeypots can cause vast improvements in computer security, including but not limited to; improved Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS) and Anti-Virus software [11],[14]. However, arguably the most important feature is that, when emulating a particular service or system, the honeypot is configured exactly the same as the regular services running on the system. The reason for this is that if an attacker succeeds at breaking into the honeypot with the same configuration it is very likely that the actual service could be compromised and is in need of some extra protection [2]. There are two main categories of honeypot that this article is concerned with and they are often used to gather very different information about the attacker. Low interaction honeypots, which can be referred to as facades, are much simpler to build and maintain, as they tend to be a simulation of a particular service, such as SSH [3]. Low interaction honeypots have been favoured by the industry due to the simplicity and ease to set up and collect meaningful results [4]. The limitations involved with these particular honeypots are vast as they only emulate a specific service and often will have no system beyond that particular service. Although they have their limitations, these types of honeypots have been the most prolific in recent years due to these limitations. The reason for this is that the user of this type of honeypot will be able to collect and analyse data that is only relevant to the service they are concerned with, which can give a much deeper understanding of the techniques and patterns that attackers tend to follow.

High interaction honeypots are what most people would consider as a typical honeypot. They provide a fully functioning system that will allow the attacker to interact with the system on all levels. Quite simply a high interaction honeypot can be any vulnerable system that is connected to a network and can be monitored for analysis. Authors in [5] describe these as truly vulnerable systems that can be probed, attacked and exploited, once the attacker gains access to the system the honeypot can be used in a botnet or to carry out other attacks. This gives light to some ethical issues with regard to continuing the research once a honeypot has been compromised, when should the system be taken back from the attacker and should it really be used in the type of attacks that it has been designed to prevent? It is for this reason that they take a lot of maintaining and will also need a system such as Honeywall [13], a gateway service monitoring all traffic, in order to complete a full forensic investigation. The example used throughout this article has been Kippo, which was deployed for this project. Other than interaction levels, honeypots can be classified in other ways such as; usage, virtual or physical.

Honeypots can take many forms and this means that they are regularly deployed in very different circumstances and positions within networks. They must also take into account the complexity of what they are researching, for example certain pieces of malware will not act in a malicious way when it finds itself in a virtual environment, this is obviously because the more we are allowed to research the methods that attackers use the more they must evolve in order to maintain the allusive nature and evade detection [6]. One of these methods is the minefield deployment system; this method will have a honeypot which is placed within the same subnet as a number of servers giving a better chance that the attacker will alert the honeypot if trying to breach a server on that system. It is well known that most attacks will scan an entire network or range of addresses and honeypots within this range will notice this scan, even if they use tools slowing down the scans to try and prevent the IDS from being alerted [2]. Other mechanisms of deployment include a Honeynet [12] which is a method of deploying an entire network of honeypots, that individually can collect information about particular services and as a whole can provide details on what is most likely to be attacked and whether the attacker will attempt to sit in the network attempting to perform attacks such as Man in the Middle.

II. Amazon Web Service

The hosting of this research was done on the Amazon Web Service (AWS). AWS provides a number of services but the Elastic Compute Cloud (EC2) is the web service which was used. EC2 provides resizable compute capacity in the cloud. It is designed to make web-scale cloud computing easier for developers and it is very useful for deploying honeypots. A main benefit of the AWS is that with its elastic computing it allows the volumes of instances to be attached, detached and reattached to instances. Being able to detach and reattach a volume may seem unnecessary but should the user become locked out of an instance, because of configuration modification, the whole server is not lost. One of the main issues surrounding honeypots is that if they are not attacked they are of no use [1]. The AWS, being part of one of the largest companies in the world, has a very high amount of traffic through its web servers and attackers know the range of IP addresses, making it much more likely that they honeypot will be able to collect an adequate amount of data. The AWS allows the user to select a particular region for where their cloud servers are deployed, putting it in a different bracket of IP addresses, which could give massively different results. The SmartHoney article has used AWS for running all manner of honeypots, focused on various services, one in particular is SSH where they found that placing their honeypots in certain regions meant a significant variation in the volume of these attacks (https://blog.smarthoneypot.com/tag/aws/). Considering the use of AWS has been very beneficial to much larger and full time honeypot projects; SmartHoney, Secure Honey it seems that it should more than suffice for a much smaller similar project.

III. SSH Protocol

The SSH protocol is designed to give the user a secure method of connecting to a system, to login or use the other services on a system, over an insecure network [7]. The SSH protocol uses a three step process in order to create the secure session; these steps are as follows, SSH transport layer, SSH user authentication and SSH connect. These steps are in fact sub-protocols that run on top of the previous sub-protocol respectively to create the SSH tunnel. The transport layer is the first sub-protocol when creating an SSH session, using TCP/IP to connect to port 22 of the server in order to provide authentication of the server and the key exchange. After the initial connect message there is a protocol-identification so that both parties are using the same protocol, SSH version 2 for example. The key exchange algorithm is then negotiated between the client and server and then the key exchange itself takes place using the agreed algorithm [8].

The user authentication process is the server confirming the identity of the user attempting to gain access. This can involve various methods, but must always include the public key authentication [7]. This is a check between the server and the client that the respective public and private keys are owned as this is used to encrypt the messages. Public key encryption uses two mathematically related keys, public and private, in order to encrypt and decrypt data. The private key is secret and only the owner should know it, whereas the public key is made readily available. Anything encrypted using the public key can only be decrypted using the corresponding private key and visa versa. Although this is the most secure method of authentication it is not always enabled and can sometimes be bypassed if the server will accept password authentication instead.

The final sub-protocol is the SSH connect, which runs on top of SSH transport layer and SSH user authentication. This sub-protocol is used to create channels used for data transfer, where each terminal session, forwarded connection, etc, are separate channels that are multiplexed into a single connection. It can provide channels for login sessions, TCP/IP connections and allows remote command execution along with file transfer using SFTP [7].

SFTP is not to be confused with FTPS, many things have changed since the introduction of protocols such as FTP and sending data over any public network without a form of encryption is considered very dangerous and in some cases prohibited. Regulations like PCI-DSS and HIPAA, for example, contain provisions that require data transmissions to be protected by encryption. When regulations such as these were initially discussed it was obvious for the need of a secure way to transfer files, which gave light to the Secure Socket Layer (SSL) being used on top of FTP to create FTPS. The issue with this is that it requires a minimum of two channels, one for the initial connection and subsequent commands and one for and data transfer, which causes a higher risk of a security breach as there must be a range of open ports on each system. SSL also does not offer any authentication per se as any certificates used can be self signed, therefore this is not an efficient method to determine the authenticity of any persons or servers that are being communicated with. Whereas SFTP uses only one channel as previously discussed to tunnel all information through. SSH is more specifically for remote login and has almost completely replaced Telnet for command-line access to remote computers.

IV. Brute Force Attacks

The most common form of initial attack involving SSH is brute force and in fact it is the most prolicfic form of attack against Internet facing servers [9]. The concept of a brute force attempt is simple; try every possible value until authentication has been achieved. The issue with using brute force is that given a 5 character password, where only letters that are of the same case are used, it could take 265 guesses (11,881,376). Given that the majority of passwords contain more letters and/or use numbers or special characters, the amount of time taken to gain entry could easily surpass the attacker's lifespan. In order to speed up this process and make it worthwhile for an attacker they will often use large lists of common passwords, called dictionaries. Dictionary attacks can be significantly much more efficient than brute force attacks because they are not sequentially trying password combinations but rather, known common passwords that are widely used. By default most SSH servers will have a limit to the number of authentication attempts that can be tried per connection, but as with many things involving connectivity it can be bypassed by the attacker, if the correct configuration is not it place, quite simply by adding an extra parameter to the initial connection command:

ssh -lusername -oKbdInteractiveDevices=`perl -e 'print "pam," x 10000'` targethost

The above command would allow the attacker up to 10000 password attempts before the connection is refused, which obviously is very useful while undergoing a brute force attack. (http://arstechnica.co.uk/security).

V. Building and Deploying

The aforementioned low interaction honeypots developed have been written in the C language, this is because there is a large amount of documentation involving available libraries, such as libssh, functions and source code that are readily available for inspiration and utilisation.There are many different ways to go about creating a low interaction honeypot of production standard, but with the aim of being simple to use and develop while maintaining the effectiveness of result gathering it can be a difficult trade off. The first method that was used was similar to many projects that already exist using the C SSH library, libssh, to employ the fuctions of the SSH protocol.

While conducting initial research about the SSH protocol and involved honeypot projects, there were quite a few production honeypots that are available and as most of these are open source projects the source code can be easily attained and edited to improve or configure on the users specific system. The most notable of these actually used the libssh for C was the SecureHoney project, which had modified a honeypot that has been previously written by another developer. This type of method to produce a honeypot is useful and most of all safe for the user to run, the reason for this is that the connection is never actually authenticated. The program uses the functions in the libssh library in order to listen for connections and begin the authentication process. The information gathered about the attacker is written into a file for later analysis. Issues with this is that an attacker with the know-how will realise that this is not an SSH daemon because information regarding the SSH can be collected while scanning and interrogating before attempting an attack. Given this information it was evident that, while this was exactly the type of honeypot that was to be produced during this project, an alternative to this could provide arguably better results with substantially less programming and development.

The alternative idea however does not emulate the SSH daemon, because it was created by editing the source code of by far the most prolific SSH daemon in use, OpenSSH. OpenSSH was originally part of the OpenBSD suite. Considering that in 2008 OpenSSH had 88% of the market share and in October 2015 announced that it will be natively supported on windows. The advantages of this are that the honeypot will be, to all intents and purposes, an actual version of the OpenSSH daemon. This means that an attacker is much less likely to be susceptible to suspicion when attempting to brute force the system.

Although this seems like a honeypot in the loosest of senses, it can be very beneficial as a production honeypot, as the software can be configured to provide an output, very similar to that in the SecureHoney project, including creating specific files for logging attempts and even collecting IP addresses of the attackers. There are many problems that can occur when attempting to use this method, as the source code for the daemon is being edited and recompiled, including making it difficult to actually use the SSH service for anything other than they honeypot, which can be devastating if this is being performed to a remote server.

VI. Methods

A. Honeypot in C

The first step to this process was becoming familiar with the libssh and the functions that were imperative to creating a valid SSH session that we would need as a basis for the honeypot. These functions are an example of how the libssh functions can be used to set up the standard configuration of a new SSH session, which include;

```
static ssh_session session;
static ssh_bind sshbind;
session=ssh_new();
ssh_options_set(session, ssh_options_timeout, &timeout)
sshbind=ssh_bind_new();
ssh_bind_options_set(sshbind, ssh_bind_options_banner, "ssh \r\n");
ssh_bind_options_set(sshbind, ssh_bind_options_bindaddr, listenaddress);
ssh_bind_options_set(sshbind, ssh_bind_options_bindport, &port);
ssh_bind_options_set(sshbind, ssh_bind_options_hostkey, "ssh-rsa"); ssh_bind_options_set(sshbind, ssh_bind_options_rsakey,rsa_keyfile);
```

The next step after making sure that the session has been set up and is listening on the desired port we must be able to accept incoming connections and drop them after the user authentication credentials that the attacker used have been logged and placed in a file called ssh_attemps. This forms the basis of the honeypot and used sections of an SSH honeypot that was found at as it fulfills the task of collecting the password attempts.

B. Modification of OpenSSH

This section is to describe exactly how the daemon can be modified to create a honeypot that is easy to maintain with little coding, although this can all be bypassed entirely by simply running the script that has been developed to automate the process. The automation of this via a script makes this method more efficient than developing a honeypot in C, having said this, the OpenSSH source files are written in C and manual editing of this would need some level of knowledge regarding programming in C.This method has been separated into two separate methods, this is because there is an instance where both methods could be doubled together in order to gather much more information from a selection of servers.

The first way of doing this is to simply modify the source code of the daemon. By doing this no SSH connection attempt will be authorised, the attackers IP address along with the username and password that was attempted, and these connection attempts will be written to a file, in the /var/log directory, called ssh_attempts. The most important part of this code is the return 0; segment, which is within the password authentication file in the source code. This line means that no matter what is entered by the attacker the authentication will always result in a failure. The problem with this method is that by doing this, the sshd is rendered useless for any sessions that the user may need in future without reverting the modifications.

The second method when editing source code requires a little more setting up and involves a second instance of the SSH daemon. The reason for this is that having the service running twice as two separate services allows different configuration for each daemon, therefore one should be configured as the façade daemon and one should be configure as a usable service. The usable service should be placed on a large port number preferably between 10000 and 65535 and designed with usual SSH security.

Finally, using a combination of both daemon modification methods a network of servers could each run multiple SSH daemons. Unlike the previous method though, this method has two fully functional daemons, one of which can be used by the user for their normal SSH activity and the other uses the ForceCommand in the sshd_config file. This will force all connections that are attempted on this daemon, to a central server that is running the aforementioned modified daemon that accepts no connections and logs all attempts, including IP address, username and password.

VII. Analysis

While running various honeypots, that have been partially developed or modified for the purposes of this project, the medium-interaction production honeypot Kippo was also deployed. The reason for initial deployment of this particular honeypot was to give a better understanding of the way that well known products, that are already available, record certain log attempts as well as the particular features that are available. This gives an insight into this type of technology available and provides an example of the reporting technique that's used. Another reason that this honeypot was deployed was to see if all the functions that are available in Kippo are of any use.

Interestingly the results of running this honeypot showed that large number of the attackers, once inside the honeypot, typed a single command and then exited. From this given information, it was deduced that the attackers knew they were within a honeypot. After experiencing this a little more research was conducted, via the SANS institute forums, and it would appear that this behavior could be a number of things, but most likely that they had in fact realised the honeypot for what it is. Accessing the server that is running Kippo can show this, and running the command that plays out a particular connection live.

```
$: ~/kippo/utils/playlog.py 20160316-100915-9940.log
```

Another idea is that this is part of an automated brute force attack. When the target system has finally been compromised, the machine that is conducting the attack saves the last password guess and logs out so that the owner can browse the compromised machine at their convenience. Another notable point was the large amount of IP addresses that had attacked this honeypot were predominantly Chinese and South Korean based internet service providers. This was also the case with downloads, using wget command. The downloads were directed to servers with Chinese IP addresses, many of which had been blacklisted online by various sites that provide lists of malicious hosts and reports it to relevant bodies. (https://cymon.io/222.186.15.61).

Kippo is a good tool but observation proves that fingerprinting may mean that by using a medium interaction honeypot such as this, we may not actually gain any better results than the low interaction SSH honeypots that never accept connections. Kippo can be difficult to use properly as a server admin with little experience of this type of technology, with more dependancies and longer set up time along with much more maintainence for sql databases, whereas a method that doesn't bother with what an attacker might possibly do once inside and a purely keeping them out strategy could provide just as valuble information with ease.

The idea was to use this as an inspiration in order to create something similar but more refined to the research needs. Although this honeypot has been successful with previous projects, it seemed to give a fair few problems when attempted to be run on one of the AWS instances. Naturally there were some dependencies to install and some configuration of the honeypot that was necessary before it could be used. The issues faced with running this on an AWS instance were initially compatibility errors. Errors including being unable to install a fully functional version of OpenSSL, which is a dependency with all SSH services as the libraries are used, this was resolved by using a different AWS instance because the package could not be located and installing from source on the server did not compile. More problems followed this, once the honeypot could be compiled and built it still wouldn't run due to the program being unable to find the private key file. On a final negative point, this technique should be used to create much more powerful projects and programs such as Kippo, when attempting to use such sophisticated techniques to emulate a daemon it makes no sense to limit the service by not implementing it into a medium to high interaction honeypot.

When emulating a service is required it seems to be far more efficient to modify a daemon that already has an enormous market share. Modification, as can be seen in this project, can be just as useful as developing a honeypot from nothing, if not more so because of the time saving. The reason that the method of two SSH daemons was used is because it allows the most amount of modification if necessary, as it is the source code being modified. This also makes the honeypot instance of the daemon incredibly secure, as the password authentication will always fail regardless of what is entered by the attacker. However, this procedure also offered some difficulties, such as modifying the incorrect files or missing out very necessary steps in the process. A solid understanding of the protocol, daemons, libraries and system files is necessary for developing any of these previously discussed honeypot designs.

VIII. Conclusions

Although this article has seemingly concluded with a tool that offers very similar services to those that are already available, this is by no means the limit to what is possible. Further work would involve the creation of a bash script. This script could then be used by 3rd parties who wish to conduct this sort of research or as an easier option when waiting to launch an SSH honeypot. Other possible development opportunities could include making this honeypot more available as a production honeypot. As the software that has been modified is open source, the redistribution of modified versions of it is permitted under its license [10]. Therefore it would not be difficult to produce a script that automates the whole process, using wget to obtain the modified code. The benefit of this easy method of install means that it could easily be placed on a large group of servers. Speculatively speaking, this would give light to even further development, using the sshpot.com as stimulus. The group of servers that are running the modified daemons would send the results to a main hub of results, being able to produce statistics and security enhancements alike. A thought on how this would be achieved, would be running a chronjob that ran another script. This script would check the hash of the sshd_attempts file and forward the results if any new ones had been recorded. Alternatively, editing the sshd_config file once again could also do this. These new additions would include a Match Group User section added that forced all connections made, to the modified daemon, straight to the main server utilising the ForceCommand option. Rather than beginning with a complete new build that is a honeypot, use existing well developed and highly distributed tools in order to develop a instrument that could be used on a commercial scale

References

[1] Spitzner L (2003) History and Definition of Honeypots, Pearson Education, Boston

[2] Jonsson, Erland, Alfonso Valdes, and Magnus Almgren. Recent Advances in Intrusion Detection: 7th International Symposium, RAID 2004, Sophia Antipolis, France, September 15-17, 2004, Proceedings. Vol. 3224. Springer Science & Business Media, 2004.

[3] Mokube, Iyatiti, and Michele Adams. "Honeypots: concepts, approaches, and challenges." Proceedings of the 45th annual southeast regional conference. ACM, 2007.

[4] Provos, Niels, and Thorsten Holz. Virtual honeypots: from botnet tracking to intrusion detection. Pearson Education, 2007.

[5] Alata, Eric, et al. "Lessons learned from the deployment of a high-interaction honeypot." arXiv preprint arXiv:0704.0858 (2007).

[6] Skoudis, Ed, and Lenny Zeltser. Malware: Fighting malicious code. Prentice Hall Professional, 2004.

[7] Ylonen, Tatu, and Chris Lonvick. "The secure shell (SSH) protocol architecture." (2006).

[8] Poll, Erik, and Aleksy Schubert. "Rigorous specifications of the SSH Transport Layer." Radboud University Nijmegen, Tech. Rep. ICIS–R11004 (2011).

[9] Owens Jr, James P. A study of passwords and methods used in brute-force SSH attacks. Diss. Clarkson University, 2008.

[10] Laurent, Andrew M. St. Understanding open source and free software licensing. " O'Reilly Media, Inc.", 2004.

[11] Zhang, Feng, et al. "Honeypot: a supplemented active defense system for network security." Parallel and Distributed Computing, Applications and Technologies, 2003. PDCAT'2003. Proceedings of the Fourth International Conference on. IEEE, 2003.

[12] Spitzner, Lance. "The honeynet project: Trapping the hackers." IEEE Security & Privacy 2 (2003): 15-23.

[13] Chamales, George. "The honeywall cd-rom." Security & Privacy, IEEE 2.2 (2004): 77-79.

[14] Simões, Paulo, et al. "On the use of Honeypots for detecting cyber attacks on industrial control networks." proc of 12th European Conf. on Information Warfare and Security (ECIW 2013). 2013.

MMO: Multiply-Minus-One Rule for Detecting & Ranking Positive and Negative Opinion

Sheikh Muhammad Saqib
Institute of Computing and information Technology
Gomal University
Dera Ismail Khan, PAKISTAN

Fazal Masud Kundi
Institute of Computing and information Technology
Gomal University
Dera Ismail Khan, PAKISTAN

Abstract—**Hit and hot issue about reviews of any product is sentiment classification. Not only manufacturing company of the reviewed product takes decision about its quality, but the customers' purchase of the product is also based on the reviews. Instead of reading all the reviews one by one, different works have been done to classify them as negative or positive with preprocessing. Suppose from 1000 reviews, there are 300 negative and 700 are positive. As a whole it is positive. Company and customer may not be satisfied with this sentiment orientation. For companies, negative reviews should be separated with respect to different aspects and features, so companies can enhance the features of the product. There is also a lot of work on aspect extraction, and then aspect based sentiment analysis. While on the other hand, users want the most positive reviews and the most negative reviews, then they can decide purchasing a certain product. To consider the issue from users' perspective, authors suggest a method Multiply-Minus-One (MMO) which can evaluate each review and find scores based on positive, negative, intensifiers and negation words using WordNet Dictionary. Experiments on 4 types of datasets of product reviews show that this method can achieve 86%, 83%, 83% and 85% precision performance.**

Keywords—Sentiment Classification; Preprocessing; Text Mining; Sentiment Orientation

I. Introduction

Positive or negative sentence is classified as opinion. With a single glance, anybody can understand either sentence is positive or negative. But automatic detection of sentence polarity requires some rules. Detection of polarity of sentence is also known as sentiment analysis. For sentiment analysis, subjectivity is very important [1][2]. Subjective sentences are user's opinion while objective sentence has no opinion. Different types of work has different accuracies i,e, final classification accuracies on reviews from various domains range from 84% for automobile reviews to 66% for movie reviews [2]. Adjective-noun pair's subjectivity found improved performance in sentiment classification [3]. This subjectivity can be done on document level [4], or sentence level [5]. Using any level, there is opinion related to some entity. Entity and aspect can be extracted using target relations, supervised learning or frequent noun [6][7][8][9]. After the extraction of aspect, sentiment analysis can be done on particular features [10][11]. Work of a supervised learning algorithm determines aspects and then sentiment classification shows the accuracies of 67.37% and 67.07% for the restaurants and laptops reviews, respectively [12]. In all above work used text is said as data. Incomplete, noisy, and inconsistent data requires preprocessing. Noisy data means incorrect attribute values, errors in data transmission, duplicate data etc. Preprocessing process includes [13] irrelevant opinion (non-English) are removed from the data set, duplicate of any opinion are deleted from the data set, stop words, numeric expressions and punctuations are removed, repeated spaces are replaced with single space character, characters repeated 2 or more times in any word are replaced with one or two occurrences for spelling correction if possible, words with all capital letters are identified used for expressing powerful emotions, all tokens starting with "http://", "https://", "http:", "http", or "www." are replaced with <URL>, negations "don't", "didn't" etc. are replaced with "do-not", "did-not". For tokenizing word_tokenize & tagging, word_tokenize & pos_tag will be used respectively. And large lexical database of English is implemented in WordNet. It consists of nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms known as synsets. Each synset express separate concept and sentiment scores. In proposed work scores will be taken from WordNet [15].

If there is single line review, then after preprocessing, directly take its polarity with some preprocessing. Now a days anybody want to purchase any product, then search can be made on online search engine. So there is no need to extract entity, because user direct take a jump on review page of required entity. User can purchase the product if as whole product reviews are positive. Aspect based classification is necessary for manufacturing, they can enhance less quality aspect. But end user need the reviews as whole ranked positive or ranked negative reviews.

A rule in which a positive word was given the sentiment score of +1 and a negative word was given the sentiment score of -1 [17], but there is no comparisons between the negative words and positive words, at the end decision (negative or positive) can be made on the basis of overall positive and negative score.

Sentiment shifter [18] also known as negation. Here in this rule a sentence with positive word followed by 'not' will get -1 score i.e. not good [-1]. In this polarity of a review can be found but ranking cannot be obtained. Suppose two sentences "This is bad" and "This is not good" here "not good [-1] = bad [-1]", both sentence has same scores and negative sentences but 1st one is less negative with respect to 2nd.

The algorithm [18] sums up the sentiment scores of the terms in the review considering negations and intensifiers, here positive score of a word is taken as 1 and negative score as -2. Whole algorithm is working well to handle negation and intensifier. But in case of two negative sentences, both can have same score i.e. this is very bad or this is very wicked. In proposed methodology author will determined 2nd sentence is more negative then 1st one.

Keeping this consideration, proposed work planned a method Multiply-Minus-One (MMO) which can determine the polarity of a review. In product reviews it is observed that user express their experiences with product using positive words, negative words, intensifiers and negations. Suppose there are two reviews, i.e. “this is good mobile”, “this is not bad mobile”, here both are positive but first one is more positive than second one. User says first sentence when they 100% satisfied with the product and says second sentence when they satisfied up to some extent. Here author proposed a strategy to determine eight rules based on opinionated words, negation and intensifiers. Experimental work on reviews of Hotel, Samsung J7, Lumia-520 and QX25, results have showed precision of 86%, 83%, 83% and 85% respectively.

II. MULTIPLY-MINUS-ONE (MMO) FOR REVIEW CLASSIFICATION AND RANKING

As main focus of proposed work is to handle intensifier and negation with opinionated words, so during preprocessing there is no need to remove intensifier and negation, if they exist in stop words list from get_stop_words of NLTK. Here author has manually created a intensifier list (i.e. I: very, more, lot, extra.. etc) and negation list (i.e. N: not, never, none, nobody, nowhere, neither,… etc).

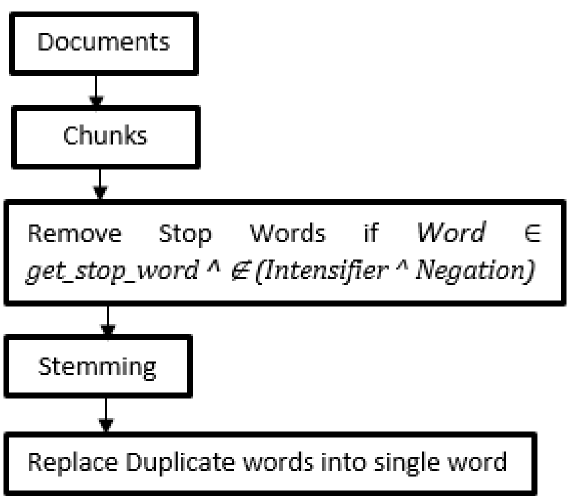


Fig. 1. Preprocessing work

After preprocessing each review consists of chunks and take each chunk with 4 tuples i.e. (W,T,PS,NS). Where w is any subjective word i.e. Adjective (JJ), negation (RB), Intensifier (RB) or even noun (NN). T is Tag and PS positive score & NS negative score. Now chunks of a review will be passed through five steps if word meets the objective of each steps.

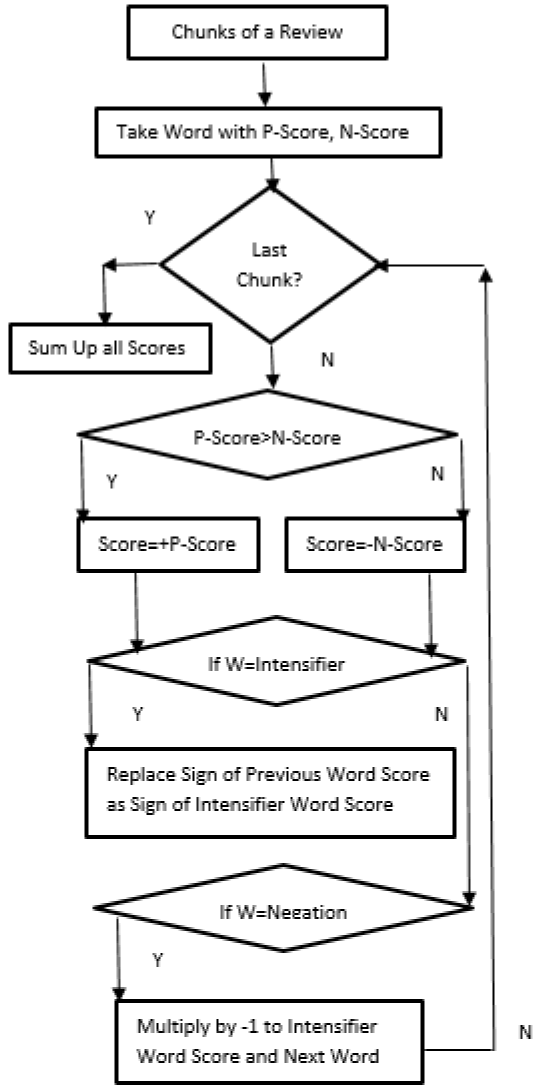


Fig. 2. Flow Chart of MMO

After exploring the whole work of proposed framework, it is concluded that whole work consists of five steps given in following table.

TABLE I. STEPS AFTER PRE-PROCESSING

Steps	Objective	Action
Step-1	Chunks	Words with Scores and Tag
Step-2	Capture Score	Take Greater Score from -ve and +ve Scores
Step-3	Intensifier Handling	Replace sign of Intensifier as next word (Adjective)
Step-4	Negation Handling	Replace sign of Intensifier and next word (Adjective) as negation
Step-5	Summation	Sum of All Scores from above Row

Positive words can express positive and negative opinion and vice versa. Keeping this thing in mind we have derived eight types of opinions, which will be handled in proposed model.

These steps will work on eight types of opinion. Here we will take simple examples to understand the concepts of proposed rules. Each examples consist of five steps as mentioned in Table-1.

A. Positive Opinion with Positive Words

It is clear that an opinion which contains positive word will depict positive opinion. We can extract polarity of a word using WordNet. If positive score is greater than negative score then then take positive score with +ve sign. Rule for this type of opinion is given below.

(PW,PS,JJ) => +PS ----------------------->Rule-1

Where PW is a positive word, PS is its positive score and JJ is adjective.

TABLE II. EXECUTED STEPS FOR RULE-1

Text-1	
Text	***This is good mobile***
Step-1	**['good', 'JJ', 0.5, 0.0]**
Step-2	**['good', 'JJ',+ 0.5]**
Step-3	['good', 'JJ', 0.5]
Step-4	['good', 'JJ', 0.5]
Step-5	**+0.5**

In above example there is no intensifier and negation, so only step-1, step-2 and step-5 will be performed. Score of this opinion is +0.5, means opinion is positive.

B. Positive Opinion with Positive Words and Intensifiers

As Intensifier with positive words increases the intensity of positive opinion, so we also consider the intensifier score with positive word.

(I,ISC,RB) ^ (PW,PS,JJ) => (+PS) + (+ISC) ----------------------->Rule-2

Where I is intensifier, ISC is its score, PW is positive word, PS is its positive score. Place +ve sign with ISC as sign of PS.

TABLE III. EXECUTED STEPS FOR RULE-2

Text-2	
Text	***This is very good mobile***
Step-1	**['very', 'RB', 0.25, 0.25, 'good', 'JJ', 0.5, 0.0]**
Step-2	**['very', 'RB', 0.25, 'good', 'JJ',+ 0.5]**
Step-3	**['very', 'RB', +0.25, 'good', 'JJ', +0.5]**
Step-4	['very', 'RB', 0.25, 'good', 'JJ', 0.5]
Step-5	**+0.75**

As there is intensifier, so step-3 will be performed to change the sign of RB (+0.25) as JJ (+0.5).

Hence Step-1, Step-2, Step-3 and Step-5 will be performed to get final score. Here final score is +0.75 means it is more positive opinion then without intensifier.

C. Negative Opinion with Negation and Positive Words

As if word is positive then its negation make its concept as negative, means multiply the positive score of JJ with -1 if previous word is negation.

(N,-NSC,RB) ^ (PW,PS,JJ) => (-NSC) + (PS * (-1.0)) ----------------------->Rule-3

Where N is negation, -NSC is negative score of negation, PW is positive word and PS is its positive score. In resultant side, there is PS * (-1) because previous word is negation.

TABLE IV. EXECUTED STEPS FOR RULE-3

Text-3	
Text	***This is not good mobile***
Step-1	**['not', 'RB', 0.0, 0.625, 'good', 'JJ', 0.5, 0.0]**
Step-2	**['not', 'RB', '-0.625', 'good', 'JJ',+ 0.5]**
Step-3	['not', 'RB', '-0.625', 'good', 'JJ', 0.5]
Step-4	**['not', 'RB', '-0.625', 'good', 'JJ', -0.5]**
Step-5	**-1.125**

As there is negation, so step-4will be performed to multiply positive score of JJ(0.5) with -1.0.

Hence Step-1, Step-2, Step-4 and Step-5 will be performed to get final score. Here final score is -1.125 means it is negative opinion.

D. Negative Opinion with Negation, Positive Words and Intensifiers

As positive word with negation shows that the opinion is negative, if there is intensifier then opinion will be more negative. So to handle intensifier, apply Rule-2 and for negation apply Rule-3. Resultant rule will be as:

(N,-NSC,RB) ^ (I,ISC,RB) ^ (PW,PS,JJ)

=> (N,-NSC,RB) ^ (I,+ISC,RB) ^ (PW,+PS,JJ)

=> (-NSC) + (ISC * (-1.0)) + (PS * (-1.0)) -->Rule-4

First of all replace the sign of Intensifier score (+ISC) as positive word score (+PS), then multiply both of them with -1, because there is negation before them.

TABLE V. EXECUTED STEPS FOR RULE-4

Text-4	
Text	***This is not very good mobile***
Step-1	**['not', 'RB', 0.0, 0.625, 'very', 'RB', 0.25, 0.25, 'good', 'JJ', 0.5, 0.0]**
Step-2	**['not', 'RB', '-0.625', 'very', 'RB',+ 0.25, 'good', 'JJ',+ 0.5]**
Step-3	**['not', 'RB', '-0.625', 'very', 'RB',+ 0.25, 'good', 'JJ', +0.5]**
Step-4	**['not', 'RB', '-0.625', 'very', 'RB', -0.25, 'good', 'JJ', -0.5]**
Step-5	**-1.375**

As there is intensifier and negation, so Step-3 and Step-4 both will be considered. For intensifier handling, replace the sign of intensifier (+0.25) as adjective (+0.5) and then to handle negation multiply them with -1.0 i.e intensifier (+0.25 * (-1.0)) as adjective (+0.5 * (-1.0)). So final score will be -1.375 means negative opinion.

E. Negative Opinion with Negative Words

If an opinion contains just negative word, then take its negative score with –ve sign. Its rule will be generated as:

(NW,NS,JJ) => -NS ----------------------->Rule-5

Where NW is a negative word, NS is negative score, resultant value with –ve sign shows that opinion is negative.

TABLE VI. EXECUTED STEPS FOR RULE-5

Text-5	
Text	***This is bad mobile***
Step-1	**['bad', 'JJ', 0.0, 0.875]**
Step-2	**['bad', 'JJ', '-0.875']**
Step-3	['bad', 'JJ', '-0.875']
Step-4	['bad', 'JJ', '-0.875']
Step-5	**-0.875**

In above example there is no intensifier and negation, so only step-1, step-2 and step-5 will be performed. Score of this opinion is -0.875, means opinion is negative.

F. Negative Opinion with Negative Words and Intensifiers

As intensifier with negative words also increases the intensity of negative opinion, so we also consider the intensifier score with negative word. As mostly intensifier are positive, so there is need to change its score as negative using following rule:

(I,ISC,RB) ^ (NW,-NS,JJ) => (ISC * (-1.0)) + (-NS) ------------------------>Rule-6

Where I is intensifier, ISC is its score, NW is positive word, -NS is its negative score. Place -ve sign with ISC as sign of NS.

TABLE VII. EXECUTED STEPS FOR RULE-6

Text-6	
Text	***This is very bad mobile***
Step-1	**['very', 'RB', 0.25, 0.25, 'bad', 'JJ', 0.0, 0.875]**
Step-2	**['very', 'RB', 0.25, 'bad', 'JJ', '-0.875']**
Step-3	**['very', 'RB', -0.25, 'bad', 'JJ', '-0.875']**
Step-4	['very', 'RB', -0.25, 'bad', 'JJ', '-0.875']
Step-5	**-1.125**

As there is intensifier, so, step-3 will be performed to change the sign of RB (-0.25) as JJ (-0.875). Hence Step-1, Step-2, Step-3 and Step-5 will be performed to get final score. Here final score is -1.125 means it is more negative opinion then without intensifier.

G. Positive Opinion with Negation and Negative Words

As if word is negative then its negation make its concept as positive, means multiply the multiply score of JJ with -1 if previous word is negation.

(N,-NSC,RB) ^ (NW,-NS,JJ) => (-NSC) + (-NS * (-1.0)) ----------------------->Rule-7

Where N is negation, -NSC is negative score of negation, NW is negative word and -NS is its negative score. In resultant side, there is (-NS) * (-1) because previous word is negation.

TABLE VIII. EXECUTED STEPS FOR RULE-7

Text-7	
Text	***This is not bad mobile***
Step-1	**['not', 'RB', 0.0, 0.625, 'bad', 'JJ', 0.0, 0.875]**
Step-2	**['not', 'RB', '-0.625', 'bad', 'JJ', '-0.875']**
Step-3	['not', 'RB', '-0.625', 'bad', 'JJ', '-0.875']
Step-4	**['not', 'RB', '-0.625', 'bad', 'JJ', 0.875]**
Step-5	**0.25**

As there is negation, so, step-4 will be performed to multiply negative score of JJ (-0.625) with -1.0.

Hence Step-1, Step-2, Step-4 and Step-5 will be performed to get final score. Here final score is 0.25 means it is positive opinion.

H. Positive Opinion with Negation, Negative Words and Intensifiers

As negative word with negation shows that the opinion is positive, if there is intensifier then opinion will be more positive. So to handle intensifier, apply Rule-6 and for negation apply Rule-7. Resultant rule will be as:

(N,NSC,RB) ^ (I,ISC,RB) ^ (NW,-NS,JJ)

=> (N,-NSC,RB) ^ (I,-ISC,RB) ^ (PW,-NS,JJ)

=> (-NSC) + (ISC * (-1.0)) + (-NS * (-1.0)) ----------------------->Rule-8

First of all replace the sign of Intensifier score (-ISC) as negative word score (-NS), then multiply both of them with -1, because there is negation before them.

TABLE IX. EXECUTED STEPS FOR RULE-8

Text-8	
Text	***This is not very bad mobile***
Step-1	**['not', 'RB', 0.0, 0.625, 'very', 'RB', 0.25, 0.25, 'bad', 'JJ', 0.0, 0.875]**
Step-2	**['not', 'RB', '-0.625', 'very', 'RB', 0.25, 'bad', 'JJ', '-0.875']**
Step-3	**['not', 'RB', '-0.625', 'very', 'RB', -0.25, 'bad', 'JJ', '-0.875']**
Step-4	**['not', 'RB', '-0.625', 'very', 'RB', 0.25, 'bad', 'JJ', 0.875]**
Step-5	**0.5**

As there is intensifier and negation, so Step-3 and Step-4 both will be considered. For intensifier handling, replace the sign of intensifier (-0.25) as adjective (-0.5) and then to handle negation multiply them with -1.0 i.e intensifier (-0.25 * (-1.0)) as adjective (-0.5 * (-1.0)). So, final score will be 0.5 means positive opinion.

III. ALGORITHM OF MMO

After merging the above eight rules in five steps (given in Table-1), algorithm of proposed work is plotted in Table-10.

TABLE X. ALGORITHM OF PROPOSED WORK

Input: All Reviews as a Text
Intensifier: List of Intensifier
Negation: List of Negation Words
Documents=Each Row Consist of a Words, Tags and Scores of a Review
for doc in Documents:
  for d in doc:
    if d→PosScore >= d→NegScore
     Score=PosScore
  else Score= NegScore * (-1.0)
   Add (W,T,Score) in NewdocScore
Intensifier Handling:
for doc in NewdocScore:
 for d in doc:
  if d ∈ Intensifiers:
    IH_docScore= Replace sign of Intensifier (RB) Score as sign of next adjective (JJ)
Negation Handling:
for doc in NewdocScore:
 for d in doc:
  if d ∈Negations:
    NH_docScore= Replace sign of Intensifier and adjective scores as sign of Negation Score
for doc in NH_docScore:
    ListOfScores = Take sum of all Scores in doc

IV. CONCLUSION AND RESULTS

User of a product only requires to know whether the product on the whole is negative or positive, while the manufacturing company would require a sorted list of reviews with respect to negative to positive impact i.e. most negative review should be placed at the top of the list, then next negative review and so on. So, on the basis that sorted list, they can enhance the quality of the features discussed in the negative reviews, according to their respective intensity. The purpose of the proposed framework is to sort all the reviews with respect to its sensitivity. Following table is representing the reviews from negative to positive score, calculating from eight said rules.

TABLE XI. SORTED REVIEWS BASED ON SENTIMENT SCORES

Text Tag	Text	Scores
T4	This is not very good mobile	-1.375
T3	This is not good mobile	-1.125
T6	This is very bad mobile	-1.125
T5	This is bad mobile	-0.875
T7	This is not bad mobile	0.25
T1	This is good mobile	0.5
T8	This is not very bad mobile	0.5
T2	This is very good mobile	0.75

Fig-3 is showing that T4 is high negative and T2 is high positive comment.

Proposed work has been applied on reviews of Hotel, Samsung-J7, Lumia-520 and QX25 products. Each product review consists of different types of intensifiers and negations. Positive predicted value is known as precision and recall means

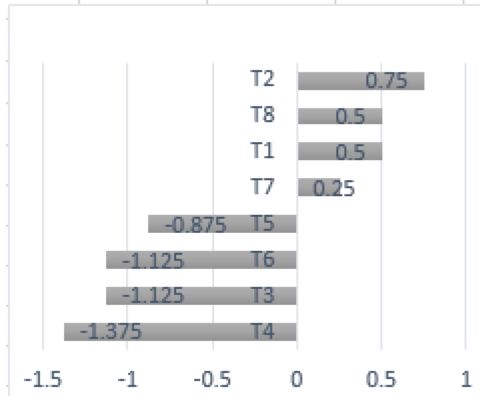


Fig. 3. Ranked Reviews

sensitivity i.e. large recall value means a few positive cases misclassified as a negative [16]. Both can be calculated through following formulas:

$$\text{Precision} = TP/(TP+FP) , \quad \text{Recall} = TP/(TP+FN)$$

Where TP is True Positive (Number of positive reviews classified correctly), FN is False Negative (Number of positive reviews classified incorrectly as a negative), TN is True Negative (Number of negative reviews classified correctly) and FP is False Positive (Number of negative reviews classified incorrectly as a positive).

TABLE XII. RESULTS BASED ON PROPOSED WORK

Category	TP	FN	TN	FP	Precision	Recall
Hotel	89%	11%	16%	14%	86%	89%
J7	71.43 %	28.57 %	42.86 %	14.29 %	83%	71%
Lumia-520	50%	10%	40%	10%	83%	83%
QX25	54.55 %	18.18 %	27.27 %	9.09%	85%	75%

Here we have mentioned some reviews of J7, with scores in sorted form, so user can read most sensitive review in advance.

TABLE XIII. SORTED REVIEWS BASED ON SCORE

Reviews	Scores
worst mobile, sensitive touch display+network problems bad in this mobile	-2.5
my J7 and my friends J7 mobile data network is not working properly sometimes it seems to be unavailable.	-1.875
worst performing smartphone till now.......low ram,low resolution only	-1.0
J7 is the Best Smartphone. Superbbbbbb...	0.25
I got this phone for about 4 months now..since then I have not face any logging issue, camera is good, very smooth, fast charging, for me it is the best phone at midrense, theres no issue even the high graphic games.	0.625
i never faced any issues during my use of the phone. also added is some new cloud function, performance and security updates.	1.0

REFERENCES

[1] Yu, Hong and Vasileios Hatzivassiloglou. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. in Proceedings of Conference on Empirical Methods in Natural Language Processing 2003.

[2] Turney, Peter D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. in Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2002).

[3] FelixHill, Anna Korhonen. Concreteness and Subjectivity as Dimensions of Lexical Meaning. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Short Papers), pages 725–731, Baltimore, Maryland, USA, June 23-25 2014. Association for Computational Linguistic

[4] Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? Sentiment classification using machine learning techniques. in Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2002).

[5] Wiebe, Janyce, Theresa Wilson, Rebecca F. Bruce, Matthew Bell, and Melanie Martin. Learning subjective language. Computational Linguistics, 2004. 30(3): p. 277-308.

[6] Jingbo Zhu Huizhen Wang, Benjamin K. Tsou, Muhua Zhu. Multi-Aspect Opinion Polling from Textual Reviews, Proceedings of ACM International Conference on Information and Knowledge Management (CIKM-2009). 2009.

[7] Chong Long, Jie Zhang, Xiaoyan Zhu.A Review Selection Approach for Accurate Feature Rating Estimation. Proceedings of Coling 2010, Poster Volume. 2010.

[8] Lei Zhang, Bing Liu, Aspect and Entity Extraction for Opinion Mining, Data Mining and Knowledge Discovery for Big Data Studies in Big Data Volume 1, 2014, pp 1-40

[9] Zhiyuan Chen, Arjun Mukherjee and Bing Liu, Aspect Extraction with Automated Prior Knowledge Learning, Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, pages 347–358, Baltimore, Maryland, USA, June 23-25 2014.

[10] D. K. Kirange1, Ratnadeep R. Deshmukh.ASPECT BASED SENTIMENT ANALYSIS SEMEVAL-2014 TASK 4.Asian Journal of Computer Science And Information Technology 4 : 8 (2014) 72 - 75.

[11] Tomas Brychcin, Michal Konkol, Josef Steinberger. Machine Learning Approach to Aspect-Based Sentiment Analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 817–822, Dublin, Ireland, August 23-24, 2014.

[12] Deepak Kumar Gupta,Asif Ekbal. Supervised Machine Learning for Aspect based Sentiment Analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 319–323, Dublin, Ireland, August 23-24, 2014.

[13] Fazal Masud Kundi, Aurangzeb Khan, Shakeel Ahmad, Muhammad Zubair Asghar. Web Lexicon-Based Sentiment Analysis in the Social 4(6)238-248, 2014. Journal of Basic and Applied Scientific Research

[14] Dive Into NLTK, Part II: Sentence Tokenize and Word Tokenize. http://textminingonline.com/dive-into-nltk-part-ii-sentence-tokenize-and-word-tokenize

[15] What is WordNet? https://wordnet.princeton.edu/

[16] David L. Olson, Dursun Delen, 2008. PerformanceEvaluation for Predictive Modeling. In: Advanced Data Mining Techniques pp: 137-139. Springer-Verlag Berlin Heidelberg.

[17] Kim, Soo-Min and Eduard Hovy. Determining the sentiment of opinions. In Proceedings of Interntional Conference on Computational Linguistics (COLING-2004). 2004.

[18] Polanyi, Livia and Annie Zaenen. Contextual valence shifters. In Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text. 2004.

[19] Wan, Xiaojun. Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. in Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2008).

Improving Accelerometer-Based Activity Recognition by Using Ensemble of Classifiers

Tahani Daghistani, Riyad Alshammari
College of Public Health and Health Informatics
King Saud Bin Abdulaziz University for Health Sciences – KSAU-HS
Riyadh, Saudi Arabia

***Abstract*—In line with the increasing use of sensors and health application, there are huge efforts on processing of collected data to extract valuable information such as accelerometer data. This study will propose activity recognition model aim to detect the activities by employing ensemble of classifiers techniques using the Wireless Sensor Data Mining (WISDM). The model will recognize six activities namely walking, jogging, upstairs, downstairs, sitting, and standing. Many experiments are conducted to determine the best classifier combination for activity recognition. An improvement is observed in the performance when the classifiers are combined than when used individually. An ensemble model is built using AdaBoost in combination with decision tree algorithm C4.5. The model effectively enhances the performance with an accuracy level of 94.04 %.**

Keywords—Activity Recognition; Sensors; Smart phones; accelerometer data; Data mining; Ensemble

I. Introduction

Health applications utilizing the built-in sensors in smartphones or those that are wearable are considered as system to simplify healthcare services such as monitoring. It is an efficient and innovative way to deliver healthcare to patients for improving healthcare outcomes and quality of life. There is a huge increase in the use of such technology. As a consequence, there is an increase in the generated data as well. In terms of health informatics, these data have received the greatest attention in various research areas such as diagnosis, decision making, and prediction. Sensed data need to be processed, analysed, and mined to derive valuable knowledge. In an attempt to address this need, classification techniques offer most capabilities need to identify physical activities by using accelerometer data [1, 5, 14]. Activity recognition is used for different purposes for a patient such as monitoring of chronic diseases, as well as fitness and wellness [8].

Despite the amount of research in activity recognition, enhancement for more accurate detection is a challenge in activity recognition problem. There is a recent advance in combining multiple classification techniques known as an ensemble of classifiers. In order to find the best combination, the best result is selected based on several experiments and using different evaluation criteria. Thus, the goal of this paper is to improve the overall performance and increase the ability to deal with more complex activities by applying ensemble of classifiers technique to improve the accuracy of recognizing various activities, as compared with other classification algorithms individually [1]. An investigation performed by Weiss and Lockhart showed that the performance of the personal model is higher than impersonal and hybrid model. Furthermore, the best algorithm that provided high performance of the personal model is MLP and Random Forests (RF) for impersonal model [4]. Lockhart and Weiss reviewed 34 AR papers; they observe many issues related to the datasets. Some issues could be found in datasets in terms of the number of subjects. They lack information about the type of developed model which is important in evaluating the performance [7].

The purpose of this study is to build activity recognition model to detect the activities by using an ensemble of classifiers technique. In this study, AdaBoost, meta classifier, is used in combination with C4.5, decision tree algorithm, for activity recognition.

The rest of the study is organized as follows: Section 2 presents the work of related activity recognition models. Section 3 describes the model development process. Section 4 presents result and Section 5 discusses results. Finally, Section 6 presents conclusion of the study.

II. Related Work

In line with the increasing usage of sensors and health applications, there is a tendency on collecting the sensor data to extract valuable knowledge. Till now, there are few applications for the activity recognition (AR), Lockhart, *et al*. recognized some AR applications such as health monitoring, self-managing systems, and fitness tracking [8].

Several studies applied data mining techniques to classify accelerometer sensor data to predict human physical activities. The summary of some articles reviewed is shown in Table 1. Kwapisz, *et al.* utilized the accelerometers in smartphones to design a system aimed at recognizing various activities. They applied three different algorithms, which are C4.5 decision tree, Logistic Regression, Multi-Layer Perceptron (MLP), on data collected from 29 users using 43 features. They reached an accuracy of 90% using MLP algorithm [6]. Catal, *et al*. conducted study based on Kwapisz, *et al*. study [6] and proposed model by using ensemble techniques of combing three classification algorithms, namely C4.5 decision tree, Multi-Layer Perceptrons (MLP) and Logistic Regression. They used the voting technique. They collected data from 36 users. The result showed that the performance of the proposed model is higher compared with applying the classification algorithms individually.

The model built by Bayat, *et al*., using six activities, achieved 91.15% accuracy. Moreover, a combination of three classification algorithms applied for the phone's potions, either in-hand or in-pocket. Based on several experiments that performed in this study, the best reported combinations that provided a high performance are MP, LogitBoost, SVM for in-hand position (91.15%) and MP, Random Forest, SimpleLogistic for in-pocket position (90.34%) [1]. While Wang, *et al.* achieved 94.8% accuracy for proposed algorithm which applied on Hidden Markov Model (HMM) [5]. Kwon et al. used suggested unsupervised learning algorithms. In this study, knowing the number of activities led to proper use of Gaussian method. Additionally, selecting K Calinski–Harabasz index achieved 90% accuracy [16]. Ayu et al. focused on the performance of the activity recognition model and the affection of the phone potion. To achieve this, they use machine learning algorithms and reach the highest performance of hand palm's position by IBk algorithm. For shirt pocket's position, Rotation Forest was the best algorithm [11]. Gao et al. investigated AR problem by using multiple sensors. The reported result was >=96.4% accuracy for ANN, decision tree and KNN which is better than the better performance by using Naïve Bayes, and SVM algorithms. Although the decision tree approach achieved the second accuracy rate, but it considered the best because training and test time consuming was less [9]. Hong, *et al.* suggested use three accelerometers in addition to RFID technology to build a model. The model with two accelerometers was able to classify the activities using decision tree with 95% accuracy. They have drawn an attention to utilize the smartphones to develop models similar to the suggested one without extra devices [17].

Recent studies motivated the use of meta algorithms such as AdaBoost, bagging and vote, which have the capability to combine one or more classifier. Dalton and O´ Laighin compared between basic and meta algorithms to find a better algorithm in terms performance, reliable and appropriate position of the sensors. The study aimed to recognize physical activities to develop monitoring system remotely. The accuracy for three highest basic algorithms was 89%, 86%, 83% for C4.5 graft, SVM and BayesNET, respectively. On the other hand, the accuracy of three meta algorithms is 95%, 92% and 91% for AdaBoostM1 with C4.5 Graft, Multiboost with AdaBoostM1 combined with C4.5 and AdaBoostM1 with SVM, respectively. The main remark from the study is the power of meta algorithms specifically AdaBoost which reached higher performance than basic algorithms [3]. Gupta and Kumar applied various algorithms to predict activities using data collected from a smartphone. The model built using AdaBoost, C4.5, Random Forest and Support vector machines (SVM). The activities classified with an accuracy level above 90% using four selected algorithms. The AdaBoost and C4.5 algorithms achieved an accuracy of 98.83% and 96.75%, respectively [13]. Wu and Song [15] used Random forest and AdaBoost to develop a model to classify activities on smart phones. They compared the result of both models and found that AdaBoost model is better performance than Random Forest model. The error rates of models were 1.10% for AdaBoost and 1.65% for Random Forest in addition to the lower time of AdaBoost model.

There are many researches focused on monitoring in healthcare by using data that generated from numerous monitoring devices. Advancements in activity recognition have demonstrated potential application in healthcare such as monitoring. Utilizing such systems and devices can improve quality of life for patients with different conditions. Massé et al. utilized stroke patients' information that generated from sensor system such as accelerometers and gyroscopes to develop activity monitoring system. As part of the system, classifier algorithms used to recognize the daily activities (standing, walking, sitting, lying) and barometric pressure to differentiate body elevation. For the purpose of improving the performance of the system, they experimented many classification algorithms and gain 82.5 %, 81.6 %, 87.1%, 85.6 %, for CCR , Naïve Bayes, Random Forest and K-Nearest-Neighbors, respectively [12]. Similarly, diabetes patients need to monitor their activities for a better lifestyle. Luštrek, *et al.* proposed using sensor data from smartphone to recognize activity for diabetes patients. Nine algorithms have been used in Weka, the classification accuracy was 88% [10].

TABLE I. THE SUMMARY OF SOME ARTICLES REVIEWED

Authors	***Classification algorithms used***	***Best Algorithm***	***Accuracy %***
Kwapisz et al. (2011) [6]	C4.5 decision tree, Logistic Regression, Multi-Layer Perceptron (MLP)	Multi-Layer Perceptron (MLP)	90%
Wang et al. (2011) [5]	Hidden Markov Model (HMM)		94.8%
Weiss and Lockhart (2012) [4]	C4.5 decision trees, Random Forest, RF, instance-based learning (IBk), neural networks, Multilayer Perceptron, NN) rule induction (J-Rip), Naive Bayes (NB), Voting Feature Intervals (VFI), Logistic Regression (LR).	MLP - personal model and Random	98.7 %
		Forests (RF) - impersonal model	75.9 %
Ayu et al. (2012) [11]	NaiveBayes NaiveBayesSimple NaiveBayesUpdateable SimpleLogistic IB1 Ibk RotationForest VFI DTNB LMT	IBk for hand palm's position.	>90%
		Rotation Forest for shirt pocket's position	97.19%
Dalton and O´ Laighin (2013) [3]	C4.5 Graft Naïve Bayes BayesNET IB1 IBK KStart JRip	Basic algorithm C4.5 Graft	89%
		Meta algorithm AdaBoost + C4.5 Graft	95%

Authors	***Classification algorithms used***	***Best Algorithm***	***Accuracy %***
	SVM Multi perceptron AdaBoost + C4.5 Graft AdaBoostM1 + SVM Bagging + C4.5 Graft MultiBoost + C4.5 Graft Vote + C4.5 Graft + SVM		
Gao et al. (2014) [9]	ANN Decision tree KNN Naïve Bayes SVM	Decision tree	96.4%
Bayat et al. (2014) [1]	Multilayer Perceptron SVM Random Forest LMT Simple Logistic Logit Boost	Combination of MP, LogitBoost, SVM	91.15%
		MP Random Forest SimpleLogistic MP LogitBoost SimpleLogistic Random Forest	90.34%
Massé et al. (2015) [12]	CCR Naïve Bayes Random Forest K-Nearest-Neighbors	K-Nearest-Neighbors	85.6 %
Luštrek et al. (2015) [10]	Naive Bayes C4.5 RIPPER SVM Random Forest Bagging AdaBoost Vote		88%
Gupta and Kumar (2015) [13]	AdaBoost C4.5 Support Vector Machines Random Forest	AdaBoost	98.83%
Catal et al. (2015) [2]	C4.5 MLP Logistic Regression Vote (C4.5+MLP+ Logistic Regression)	Vote (C4.5+MLP + Logistic Regression)	93.47%

III. METHODOLOGY

The study proposed activity recognition model by an ensemble of classifiers techniques, it aims to detect the human activities. The Wireless Sensor Data Mining (WISDM), which is publicly available on http://www.cis.fordham.edu/wisdm/dataset.php, is used in this study. This data is obtained from the transformation of time series accelerometer sensor data from smartphones during experiments of 36 people. It includes 46 features and label class. In the dataset, there are 5418 instances for six activities which are walking, jogging, upstairs, downstairs, sitting, and standing. WEKA software used to build the model using AdaBoost ensemble approach. According to previous studies, AdaBoost used effectively to enhance performance for activity recognition in combining with other classification algorithm. Several experiments were conducted by using AdaBoost in combination with C4.5 (decision tree) MLP (artificial neural network), Logistic algorithms. The three classifiers used in this study were decided due to the high performance achieved by those algorithms in previous studies. During experiments, 10-fold cross-validation (CV) approach was used. The confusion matrix presented the result of all experiments and performance compared among different parameters which are true positive (TP), false positive (FP), precision, recall, area under ROC Curve (AUC) and F-measure. Parameters employed as measure method to evaluate the model are as follows:

- True positive (TP): These are activities that correctly predicted.
- False positive (FP): These are activities that not predicted incorrectly.
- Precision: how often the prediction is correct.
- Recall: The number of correct activities predicted divided by the number of activities that should be predicted.
- Area under ROC Curve (AUC): The larger AUC indicates a high correct prediction and low incorrect prediction for activities.
- F-measure: it measures the accuracy of the test by a weighted harmonic average of precision and recall.

Furthermore, the experiments were repeated using different iteration numbers. NumIterations is one of the Adaboost algorithm parameters that determines the number of models that will be used in the decision step. Ensemble AdaBoost – C4.5 model re-build, repeatedly with altering iteration numbers from 10 to 100. The aim of this additional step is to enhance the performance of the selected combination of classifiers. The following section presents the results of the mentioned parts.

IV. RESULTS

The result of experiments confirms that AdaBoost used effectively to recognize activities in addition to power of C4.5 algorithm. Based on the height results of related work, AdaBoost selected and combined with each of the three algorithms which are C4.5, Logistic, Multi-Layer Perceptron (MLP). The performance achieved was over 90% most times but the best performance was achieved by combing AdaBoost with C4.5. It started from 94.034 % using default sitting (ten iteration numbers). Fig.1 shows the overall performance of proposed models that reached during experiments.

The performance for each classifier is individually calculated and presented to demonstrate the affectivity of ensemble classifiers. The overall performance is 89.46%, 84.94%, 92.65 for C4.5, Logistic, Multi-Layer Perceptron (MLP), respectively. The confusion matrix for each algorithm alone is shown in Tables 2 to 5. Table 5 presents the confusion matrix of proposed AdaBoost-C4.5 model with default sitting 10 iterations. The new model achieved 94.04% which is the highest compared with standalone classifiers or other classifiers combination.

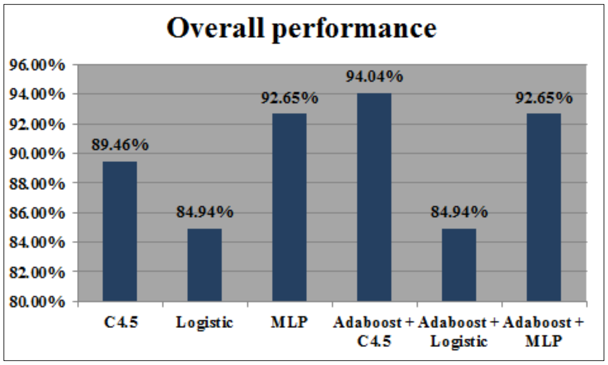


Fig. 1. Overall accuracy for different proposed models

TABLE II. CONFUSION MATRIX OF C4.5

Walking	Jogging	Upstairs	Downstairs	Sitting	Standing	TP Rate	FP Rate	Precision	Recall	F-measure	ROC Area
1988	19	37	34	2	1	95.53	4	93.7	95.5	94.6	97.2
17	1563	31	13	0	1	96.18	2	95.5	96.2	95.8	98
59	37	427	106	1	2	67.56	41	68.4	67.6	68	86
53	14	126	334	1	0	63.26	31	68.4	63.3	65.7	86.8
3	1	2	1	295	4	96.41	0.1	98.7	96.4	97.5	98.5
2	3	1	0	0	240	97.56	0.2	96.8	97.6	97.2	99
						89.5	2.9	89.2	89.5	89.3	95.3

TABLE III. CONFUSION MATRIX OF MULTI-LAYER PERCEPTRONS (MLP)

Walking	*Jogging*	*Upstairs*	*Downstairs*	*Sitting*	*Standing*	*TP Rate*	*FP Rate*	*Precision*	*Recall*	*F-measure*	*ROC Area*
2027	2	25	26	0	1	97.41	1.4	97.7	97.4	97.6	99.5
6	1609	6	3	1	0	99.02	0.1	99.7	99	99.4	99.9
14	1	520	93	3	1	82.28	4.2	72.3	82.3	77	95.7
21	2	161	340	1	3	64.39	2.5	73.3	64.4	68.5	93.3
3	0	2	0	292	9	95.42	0.2	97	95.4	96.2	99.8
3	0	5	2	4	232	94.31	0.3	94.3	94.3	94.3	99.4
						92.7	1.3	92.8	92.7	92.6	98.6

TABLE IV. CONFUSION MATRIX OF LOGISTIC RECOGNITION

Walking	*Jogging*	*Upstairs*	*Downstairs*	*Sitting*	*Standing*	*TP Rate*	*FP Rate*	*Precision*	*Recall*	*F-measure*	*ROC Area*
1980	9	57	34	0	1	95.15	9.8	85.8	95.1	90.2	96.9
18	1603	1	2	0	1	98.65	0.4	99	98.6	98.8	99.9
177	6	317	128	4	0	50.16	5.7	53.8	50.2	51.9	91.2
129	2	203	190	3	1	35.98	3.5	52.9	36	42.8	89.3
0	0	5	5	288	8	94.12	0.4	93.8	94.1	94	99.5
4	0	6	0	18	224	91.06	0.2	95.3	91.1	93.1	99.6
						0.849	4.9	83.7	84.9	84.1	96.7

In terms of Adaboost parameters, different values have been set to iteration number and reached our goal to improve the performance. The experiments repeated using different iteration numbers indicate a significant improvement in the performance as shown in Figure 2.

Table 6 also presents the confusion matrix of the proposed AdaBoost-C4.5 model that used 80 iterations to compare the results. Clearly, the improvement reflected on all parameters such as false positive rate, it decreased until 0.9%, which indicates reduced in a number of instances that were classified incorrectly.

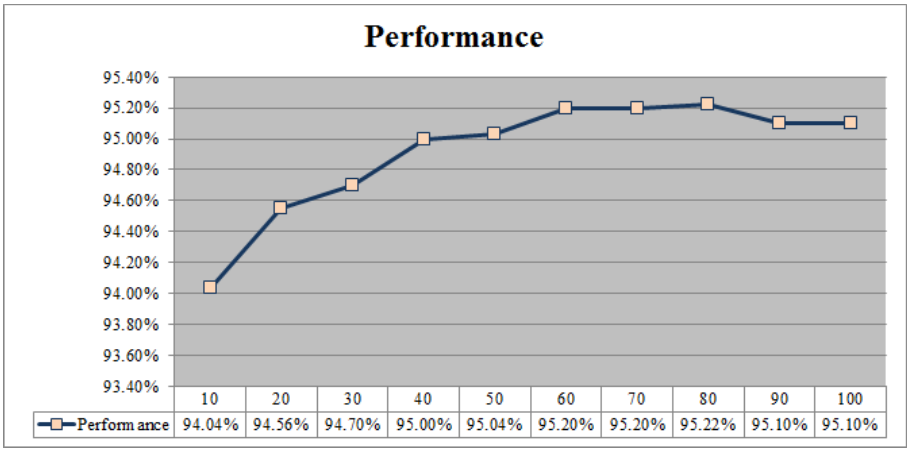


Fig. 2. the performance of the model using different iterations number

TABLE V. CONFUSION MATRIX FOR ADABOOST-C4.5 MODEL – 80 ITERATION NUMBER

Walking	*Jogging*	*Upstairs*	*Downstairs*	*Sitting*	*Standing*	*TP*	*FP*	*Precision*	*Recall*	*F-measure*	*ROC Area*
2051	4	13	12	0	1	98.6	0.7	98.8	98.6	98.7	99.8
6	1608	6	5	0	0	99	0.6	98.7	99	98.8	99.8
6	10	532	84	0	0	84.2	2.2	83.3	84.2	98.8	98.7
11	7	82	428	0	0	81.1	2.2	80.1	81.1	98.8	98.5
0	0	2	1	299	4	97.7	0	99.3	97.7	98.5	1
1	0	4	4	2	235	95.5	0.1	97.9	95.5	96.7	1
						95.1	0.9	95.1	95.1	95.1	99.6

V. DISCUSSION

In this study, an improvement is observed in the performance when combine classifiers than use them individually. C4.5 was the most effective classifiers although Multi-Layer Perceptron (MLP) achieved better accuracy alone, but it is not effective one to combine with AdaBoost. Also, Multi-Layer Perceptron (MLP) and C4.5 alone are slightly better than AdaBoost model for standing activity. Moreover, The C4.5 algorithm classified 97.56% of instances correctly compared to AdaBoost model 94.04%.

A comparison between the vote model proposed by Catal et al. study and the proposed model in this study is performed. As a result of the comparison, the proposed AdaBoost-C4.5 ensemble model achieved higher overall performance 94.04 % than vote model 93.47%. In addition to the shorter calculation time consumed by AdaBoost model. As mentioned above, re-building the model using different iteration number led to improve the performance. In fact, Adaboost build a model per iteration. As number of models increases the area under ROC Curve (AUC) also increases, although the prediction confidence slightly decreases. The possibility of recovering false negative will increase and classifying the new samples will be more accurate. The result showed improvement among various parameters as summarized as shows in Table 7. Increasing values of different parameters, except FP rate, indicates a better classification.

TABLE VI. COMPARISON OF MODELS AMONG VARIOUS PARAMETERS

	AdaBoost model 10 iterations number	*AdaBoost model 80 iterations number*
True positive	94%	95.2%
False positive	1.4%	0.9%
Precision	94%	95.3 %
Recall	94%	95.2 %
F measure	94%	95.2 %
ROC Area	99.5%	99.6%
Kappa statistic	91.87%	93.49%

According to the confusion matrix of Ababoost model, there is improvement in the performance of Downstairs activity reflected in true positive (81.1%) value and F measure measurements (98.8%). Furthermore, The results of walking and jogging activities were high due to the large number of instances for both activities compared to the others. In other hand, the lowest results were observed for upstairs and down-stairs activities due to the difficulty in differentiating between them. However, performance improvement observed in the downstairs activity using AdaBoost – C4.5 ensemble.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

Mining data collected from sensors provides valuable result in the activity recognition area. The improvement in performance is a requirement especially in the health field where such results are used to develop various health systems related to patient's lifestyle. The spread of smartphones made desirable data existing with huge volume. This increases opportunity in the data mining research area.

In this study, AdaBoost- C4.5 ensemble model is proposed using public data to recognize physical activities. The result shows a significant improvement in performance using meta classifiers instead of basic classifiers individually. Proposed model has an accuracy level starting from 94.034%.

B. Future work

The improved results motivate to conduct more studies in this field. Other combinations (meta and basic) and different machine learning methods can be used. The proposed models can be applied on different datasets to recognize more and complex activities.

REFERENCES

[1] Bayat, M. Pomplun, D.A. Tran, A study on human activity recognition using accelerometer data from smart phones, in: Proceedings of the MobiSPC-2014,Procedia Computer Science, vol. 34, 2014, pp. 450–457.

[2] Catal, C., Tufekci, S., Pirmit, E., & Kocabag, G. (2015). On the use of ensemble of classifiers for accelerometer-based activity recognition. Applied Soft Computing.

[3] Dalton, A., & OLaighin, G. (2013). Comparing supervised learning techniques on the task of physical activity recognition. Biomedical and Health Informatics, IEEE Journal of, 17(1), 46-52.

[4] G.M. Weiss, J.W. Lockhart, The impact of personalization on smart-phone based activity recognition, in: Proceedings of the AAAI Works hopon Activity Context Representation: Techniques and Languages, 2012,pp. 98–104.

[5] J. Wang, R. Chen, X. Sun, M.F.H. She, Y. Wu, Recognizing human daily activities from accelerometer signal, Procedia Eng. 15 (2011) 1780–1786.

[6] J.R. Kwapisz, G.M. Weiss, S.A. Moore, Activity recognition using cell phone accelerometers SIGKDD, Explor. Newsl. 12 (March (2)) (2011)74–82.

[7] J.W. Lockhart, G.M. Weiss, Limitations with activity recognition methodology& datasets, in: Proceedings of the UbiComp'14, Seattle, WA, 2014.

[8] J.W. Lockhart, T. Pulickal, G.M. Weiss, Applications of mobile activity recognition, in: Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp'12), ACM, New York, NY, 2012, pp. 1054–1058.

[9] L. Gao, A.K. Bourke, J. Nelson, Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems, Med. Eng. Phys. 36 (6) (2014)779–785.

[10] M.A. Ayu, S.A. Ismail, A.F.A. Matin, T. Mantoro, A comparison study of classifier algorithms for mobile-phone's accelerometer based activity recognition, Procedia Eng. 41 (2012) 224–229.

[11] Massé, F., Gonzenbach, R. R., Arami, A., Paraschiv-Ionescu, A., Luft, A. R., & Aminian, K. (2015). Improving activity recognition using a wearable barometric pressure sensor in mobility-impaired stroke patients. Journal of neuroengineering and rehabilitation, 12(1), 72.

[12] Sarthak Gupta and Ajeet Kumar. Article: Human Activity Recognition through Smartphone's Tri-Axial Accelerometer using Time Domain Wave Analysis and Machine Learning. International Journal of Computer Applications 127(18):22-26, October 2015. Published by Foundation of Computer Science (FCS), NY, USA.

[13] Suarez, I., Jahn, A., Anderson, C., & David, K. (2015, September). Improved activity recognition by using enriched acceleration data. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 1011-1015). ACM.

[14] Y. Kwon, K. Kang, C. Bae, Unsupervised learning for human activity recognition using smart phone sensors, Expert Syst. Appl. 41 (14) (2014)6067–6074.

[15] Y.-J. Hong, I.-J. Kim, S.C. Ahn, H.-G. Kim, Mobile health monitoring system based on activity recognition using accelerometer, Simul. Model. Pract. Theory 18 (4)(2010) 446–455.

[16] Wu, S., & Song, Y. (2014). Human Activity Recognition on Smartphone: A Classification Analysis. TELKOMNIKA Indonesian Journal of Electrical Engineering, 12(9), 7041-7045.

A Multimodal Firefly Optimization Algorithm Based on Coulomb's Law

Taymaz Rahkar-Farshi
Department of Computer engineering
Istanbul Aydin University
Istanbul, Turkey

Sara Behjat-Jamal
Department of Computer engineering
Gazi University (Faculty of Technology)
Ankara, Turkey

***Abstract*—In this paper, a multimodal firefly algorithm named the CFA (Coulomb Firefly Algorithm) has been presented based on the Coulomb's law. The algorithm is able to find more than one optimum solution in the problem search space without requiring any additional parameter. In this proposed method, less bright fireflies would be attracted to fireflies which are not only brighter, but according to the Coulomb's law pose the highest gravity. Approaching the end of iteration, fireflies' motion steps are reduced which finally results in a more accurate result. With limited number of iterations, groups of fireflies gather around global and local optimal points. After the final iteration, the firefly which has the highest fitness value, would be survived and the rest would be omitted. Experiments and comparisons on the CFA algorithm show that the proposed method has successfully reacted in solving multimodal optimization problems.**

Keywords—Swarm Intelligence; multimodal firefly algorithm; multimodal optimization; firefly algorithm

I. Introduction

Optimization is finding an optimum solution from a set of available options with the purpose of optimizing criteria for the problem in a limited time. The main challenge with single solution optimization algorithms, however, is that they are only able to find one optimum solution from a set of available options while most real-world problems have more than one optimum solution [1]. Hence, multimodal optimization algorithms which are among the novel inventions of evolutionary algorithms, have been designed to find a set of possible solutions from available options. Unlike unimodal optimization algorithms which try to avoid local optimal points, multimodal optimization algorithms recognize these points as a solution. Although normally the algorithms have not been basically designed to merely solve these problems, several algorithms have recently tried to solve these problems by modifying existing unimodal optimization algorithms. The majority of these algorithms are based on particle swarm optimization algorithms [1-6] and genetic algorithms [7-10]. The firefly optimization algorithm has been used successfully to optimize different kinds of problems, but all of them have been within the span of unimodal optimization problems. In this paper, the Coulomb's law has been applied to the firefly optimization algorithm in order to turn it into a multimodal algorithm.

EPSO algorithm [3] was introduced by J. Barbara and Carlos A. C. in 2009. In this method, the selection of global optimum mechanism, in PSO algorithm, was changed using Coulomb's law. Then, the particles that are to be selected as the global optimum can be separately calculated for each particle. In fact, particles may move towards different particles as the global optima. In other words, the global optima for every particle could vary from one particle to another. Hence, particles not only do not surround the global optima, but they also surround their local optima. It is evident that a particle with a more desirable cost function is surrounded by more particles. It is this mechanism's property that particles tend to move towards a point that has both an appropriate cost function value and an appropriate distance from the particle.

FERPSO [2] is a well-known algorithm that has been proposed for solving multimodal optimization problems which was introduced by Xiaodong Li in 2007. In terms of nature, this algorithm could be viewed as: more birds will gather where there is more food. In fact, if they find a good resource near themselves, they will not use farther resources. In FERPSO, the particles that are to be selected as global optimal point are selected for each particle regarding the Euclidian distance between particles. In essence, the overall structure of FERPSO and EPSO are highly similar, and they both have the same level of complexity.

B. Y. Qu et al have combined a local searching technique with some existing multimodal PSO optimization algorithms that have used niching [2, 11, 12] method trying to solve such problems. In this method, the personal best for particles are improved significantly by using a local searching method. In fact, the personal best is improved by generating a random point between the particle and the nearest point, that is, if the newer point is more desirable than the current personal best, the new point will replace the former, otherwise, the original point stays intact.

J.Zhang et al. [13] proposed a modified algorithm called the sequential niching particle swarm optimization (SNPSO). This algorithm divides the whole population into several sup-populations which can be located around optimal solutions in multimodal problems. They use space convergence rate (SCR), in which each sub-population detects global and local optimal solutions until the end of iteration.

Xiaodong li. [11] proposed an improved PSO algorithm called the (SPSO). In this method, the idea of species is used to specify each species' best value of neighborhood. The algorithm divides the whole population into several populations called species with regard to their similarity. Each species gather around a particle called species seed.

II. Firefly Algorithm

A. The Behavior of Fireflies in Nature

There are almost two thousand known species of firefly in nature, most of which emit flashes of light with a certain rhythm in order to attract a mating partner or bait. In addition to these reasons, fireflies can protect themselves against the attackers using the flashes which can also attract the opposite sexes. The distance between the fireflies and the environment, where the light is emitted, is somehow effective on the intensity of light received by fireflies. As the light intensity obeys the inverse square law at a particular distance r (I ∝ 1/r2), and because light is absorbed by air, most fireflies can just be visible to a limited distance.

B. Firefly Algorithm

The firefly algorithm is one of the novel optimization algorithms based on swarm intelligence which was first introduced by X.Yang in 2008 [14]. It was inspired by the natural behavior of fireflies. The firefly algorithm randomly distributes a number of artificial fireflies in the search space at the beginning. All of the fireflies are unisexual and thus regardless of gender, each firefly can be attracted by any other firefly. Each firefly produces a light whose intensity depends on the optimality of its position and is proportional to its fitness value. The next step is comparing constantly the intensity of the light of each firefly with that of other fireflies and less bright fireflies moves towards brighter ones. Evidently, depending on the distance, fireflies receive lights with varying intensities; however, the brightest firefly moves randomly in search of space to increase its chance of finding the global optimum solution. Movement of the less bright firefly towards the brighter one is expressed through equation (1).

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2}\left(x_j - x_i\right) + \alpha\left(rand - \frac{1}{2}\right) \qquad (1)$$

Where β_0 is the maximum coefficient of attraction between i_{th} and j_{th} fireflies, α is the coefficient of random displacement vector, γ is the light absorption coefficient for the environment, and r_{ij} is the Euclidean distance between two fireflies. Each firefly is compared to all others and if its fitness value is less than that of another one, it will be attracted according to equation (1). This trend continues to the last algorithm iteration when finally the optimum solution is obtained as the final solution. Main steps of the firefly algorithm can be expressed in the form of the pseudo-code briefed in Algorithm 1 (FA).

III. Multimodal Optimization

Constraints such as physical, temporal and economic limitations can prevent achievement of actual results; however, having knowledge of multimodal optimization solutions is very useful in engineering fields. In such cases, if multiple local and global solutions are available, the optimum system performance is obtained by switching between solutions. Since there are several solutions to many real-world problems, multimodal optimization algorithms are useful for solving these problems. Not only are these algorithms able to locate multiple optima in a single run, but they also preserve their population diversity. The reason why classic optimization techniques are not used to find multiple solutions shows their unreliability in finding more than one solution in multiple runs [15]. Evolutionary algorithms including Genetic Algorithms (GAs), Differential Evolution (DE), Particle Swarm Optimization (PSO), and Evolution Strategy (ES) are kinds of algorithms which has been tried to solve multimodal optimization problems. Referred algorithms [7, 10, 15-23] are among algorithms designed to the aforementioned criteria.

IV. Multimodal Firefly Algorithm

Studies on multimodal optimization have mostly focused on the PSO and genetic algorithms. In this paper, like other meta-heuristic algorithms which have used unimodal algorithms for solving multimodal optimization problems, some changes are made on FA algorithm without the need for any additional parameter, and it has been utilized to solve multimodal optimization problems. In the proposed algorithm, the Coulomb's law in equation (2) has been used to calculate the electrostatic interaction between two fireflies. This technique was successfully used by J. Barrera and A. Coello in [3] to obtain a multimodal PSO algorithm. They have used this method to calculate forces between two particles which has also been used in the present paper to calculate the attraction between two fireflies.

Algorithm 1 Pseudo-code for FA main steps

```
Objective function f(x), x = (x1, ..., xd)^T
Generate initial population of fireflies xi (i = 1, 2, ..., n)
Light intensity Ii at xi is determined by f(xi)
Define light absorption coefficient γ
while (t <MaxGeneration)
    for i = 1 : n all n fireflies
        for j = 1 : i all n fireflies
            if (Ij > Ii)
                Move firefly i towards j in d-dimension;
            end if
            Attractiveness varies with distance r via exp[−γr]
            Evaluate new solutions and update light intensity
        end for j
    end for i
    Rank the fireflies and find the current best
```

end while
Postprocess results and visualization

$$F_{(j,i)} = \frac{1}{4\pi\varepsilon_0} \cdot \frac{Q_i . Q_j}{r^2} \qquad (2)$$

In this equation, $\frac{1}{4\pi\varepsilon_0}$ is the proportionality constant (Coulomb constant); Q_i and Q_j denote the magnitude of two charged particles, and r is the distance between two charges. According to this formula, force magnitude is proportional to magnitude of charges but it obeys inverse square law for distance. Hence, the attraction between two fireflies is calculated through the following equation.

$$\vec{F}_{(i,\vec{j})} = \alpha . \frac{f\,(p_i) . f\,(p_{\vec{j}})}{\left\| p_i - p_j \right\|^2} \qquad (3)$$

In this equation, α is equal to 1 and $f(p_i)$ is the fitness value for the firefly which will be attracted to one of the present fireflies. $f(p_{\vec{j}})$ is the vector of the fitness value of other fireflies to which the ith firefly is being compared. Finally, the destination firefly is obtained by using equation (4).

$$Fmax_{(i)} = \operatorname{argmax}_{\vec{F}} \vec{F}_{(i,j)} \qquad (4)$$

Calculating the above equation yields to the maximum value of $F_{(i,j)}$. Finally, the index of the j^{th} firefly, with the highest value of F, is calculated and equation (5) is obtained by changing equation (1).

$$x_i = x_i + \beta_0 e^{-\gamma r^2_{ij\,\mathrm{Fmax}(i)}} \left(x_{j\,\mathrm{Fmax}(i)} - x_i \right) + \alpha \left(rand - \frac{1}{2} \right) \qquad (5)$$

As a result, the i^{th} firefly is attracted to the firefly which has the highest value of F. Therefore, destination fireflies are not selected just based on the value of fitness value but from calculating the electrostatic interaction between other fireflies. This method prevents the attraction of other fireflies by the best firefly. Instead, fireflies are attracted by fireflies which besides having sufficient fitness value, must be at a close distance since distance is an effective parameter in their attraction. In each iteration in this case, each firefly compares its electrostatic interaction with others and then moves toward the firefly which has the highest electrostatic interaction. As it was mentioned before, α is the coefficient of random displacement with a value considered to be [0.1-1] at the beginning. This makes the firefly's movements to be random to some extent and to search for new sources; however, this value of α results in a less precise solution at the end of iteration. To prevent it, the coefficient of random movement is reduced in each iteration so as to reduce the randomness of the movement of the firefly toward the destination. Moreover, the value of γ is increased in each iteration so that fireflies take smaller steps at the end of the iteration. These two actions take place using equations (6) and (7).

Algorithm 2 Pseudo-code for CFA main steps

Objective function f(**x**), **x** = $(x_1, ..., x_d)^T$
Generate initial population of fireflies $\mathbf{x}_i$ (i = 1, 2, ..., n)
Light intensity I_i at $\mathbf{x}_i$ is determined by $f(\mathbf{x}_i)$
Define light absorption coefficient γ

```
while (t <MaxGeneration)
    γ= γ +t / (MaxGeneration*5);
    α = α * (1-t / (MaxGeneration*2);
    for i = 1 : n all n fireflies
        for j = 1 : i all n fireflies
            if (Ij > Ii)
                push(F,( Ii * Ij)/norm(x(i)-x(j))^2); */ in d-dimension
            end if
        end for j
        Move firefly i towards j^max(f) in d-dimension;
        Attractiveness varies with distance r via exp[−γr]
        Evaluate new solutions and update light intensity
    end for i
end while
```

Postprocess results and visualization

TABLE I. TEST FUNCTION

	Test function	Range	Peaks Global/Local
1	$f(x_1,x_2)=20-(-20.\exp\left(-0.2\sqrt{\frac{1}{2}(x_1^2+x_2^2)}\right)-e^{\left(\frac{1}{2}\left(\cos(2\pi x_1)+\cos(2\pi x_2)\right)+20+e^1\right)}$	$-5<x_1<5$ $-5<x_2<5$	**1/1**
2	$f(x_1,x_2)=100-\left(20+\left(x_1^2-10\cos\left(2\pi x_1^2\right)\right)+\left(x_2^2-10\cos\left(2\pi x_2^2\right)\right)\right)$	$-5.12<x_1<5.12$ $-5.12<x_2<5.12$	**1/1**
3	$f(x_1,x_2)=220-\sum_{i=1}^{5} i\cos((i+1)x_1+i).\sum_{i=1}^{5} i\cos((i+1)x_2+i).$	$-5.12<x_1<5.12$ $-5.12<x_2<5.12$	**4/201**
4	$f(x_1,x_2)=20-\left(\left(4-2.1x_1^2+\frac{x_1^4}{3}\right)x_1^2+x_1x_2+\left(-4+4x_2^2\right)x_2^2\right)$	$-1.9<x_1<1.9$ $-1.1<x_2<1.1$	**2/6**
5	$f(x_1,x_2)=2500-((x_1^2+x_2-11)^2+(x_1+x_2^2-7)^2)$	$-6<x_1<6$ $-6<x_2<6$	**4/4**
6	$f(x_1)=5+sin^6(5\pi x)$	$0<x_1<1$	**5/5**
7	$f(x_1)=5+e^{-2log(2)\times\left(\frac{x_1-0.1}{0.8}\right)^2}\times sin^6(5\pi x)$	$0<x_1<1$	**1/5**

*f*1 = Ackley, *f*2 =Rastrigin, *f*3 = Shubert, *f*4 = Six-hump camel back, *f*5 = Himmelblau, *f*6 = Equal maxima, *f*7 = Decreasing maxima

$$\alpha = \alpha\times\left(1-\frac{\text{Iteration}}{\text{MaxGeneration}\times 2}\right) \quad (6)$$

$$\gamma = \gamma+\frac{\text{Iteration}}{\text{MaxGeneration}\times 5} \quad (7)$$

Main steps of CFA can be summarized into the pseudo-code shown in Algorithm 2.

V. EXPERIMENTAL RESULT

A. Test Functions

The experiments have been performed on benchmark functions common in multimodal optimization. Specifications of these algorithms are presented in Table (1).

B. Configurations

All algorithms were implemented in Matlab 2013 and were run in a computer equipped with an Intel Core(TM) i7-3632QM 2.2 GH processor and 8 gigabytes of RAM.

C. Performance measures

To assess the performance of aforementioned algorithms in section (6.4), the following 7 criteria are considered and measured 50 runs.

1) Success Rates (SR): The percentage of performances in which all the optimum points have been found successfully.

To calculate success rate, a user specified parameter, called the Level of Accuracy (LOA), is considered. This parameter is usually between (0,1] and is used to measure the difference between found solutions and the real optimum points in functions, so that if the difference between found solution and the real solution is less than the amount of LOA, then the found solution is counted as a successful solution [12].

2) Average Number Of Optima Found (ANOF): The average of optimum points found considering LOA for 50 runs.

3) Global Average Number of Optima Found: The average of global optima found considering LOA for 50 runs.

4) Average Function Evaluation is the average number of fitness function calling for 50 runs [1].

5) Success Performance (SP): This parameter is computable when the amount of SR is not zero [24]. SP is calculated from equation (8) :

$$SP=\frac{Avarage\ Number\ Of\ \ Function\ Evaluation\left(ANOF\right)}{SR} \quad (8)$$

.

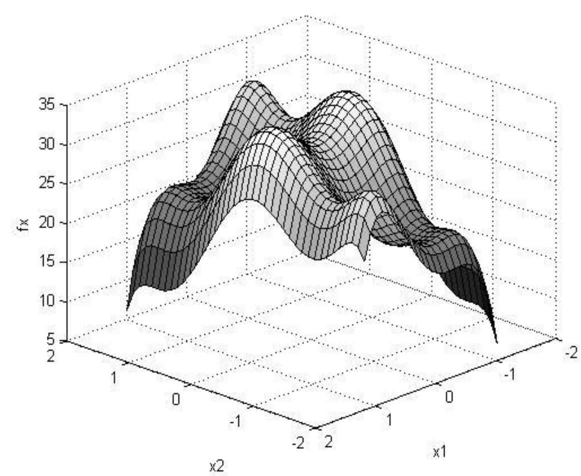


Fig. 1. Search landscape of f4

Based on the fact that algorithms with less ANOF and higher SR amount can be considered as a better one, it can be concluded that less SP amount is more acceptable.

6) Maximum Peak Ratio (MPR): The quality of optima is tested without considering the population distribution, and the performance metric, which is called the maximum peak ratio statistic (MPR), is adopted. The MPR is defined as follows:

$$MPR = \frac{\frac{1}{C}\sum_{i=1}^{C} f_i}{\frac{1}{q}\sum_{i=1}^{q} F_i} \qquad (9)$$

c: The number of found optimum point in the solution

q: The number of real optimum points in the solution

f_i: The quantity of fitness function obtained in the final population

F_j: The quantity of real fitness value in objective function

7) Precision : the ratio of found optimum points to the number of real optimum points

$$\text{Precision} = \frac{c}{q} \qquad (10)$$

D. Test and comparison results

The results of the experiments are shown in the tables 2-4 which have the accuracies of 10-1, 10-2 and 10-3 respectively. First columns of all three tables represent test functions; second columns are equivalent of implemented algorithms and the other columns are as written on top of each column. The best performance was reported in boldface. As it can be seen from the results, with the increase of the level of accuracy, the proposed algorithm has a better performance compared to FER-PSO [2], EPSO [3] and LS-FER-PSO [1] algorithms. It is shown in the tables that the presented algorithm and LS-FER-PSO algorithm have better performance compared to other algorithms. Comparing LS-FER-PSO algorithm and presented algorithm, it is shown that sometimes one performance is better than the other and vice versa. However it should be mentioned that the ratio of average function evaluation of LS-FER-PSO algorithm is 1.67 times higher than that of the presented algorithm. The performance of this algorithm proved its usefulness in solving optimization problems.

Error! Reference source not found.. shows the search space of f4. **Error! Reference source not found.** also indicates the position of fireflies during the running of the proposed algorithm with 60 fireflies and 60 iterations using the f4 function. **Error! Reference source not found.**. A-D shows the position of particles in the 1st, 10th, 20th and 30th iterations. In fact, 4 out of the 6 available points were successfully found in 1800 function evaluation. However all optimum points can be found by increasing the number of fireflies. As it can be clearly observed in **Error! Reference source not found.**, by getting close to the end of the iterations, fireflies around optimum solutions become gradually more and more concentrated. Since the firefly algorithm was designed for maximum optimization problems, the average value of cost functions of fireflies increased in each iteration **Error! Reference source not found.**. Moreover, as seen in **Error! Reference source not found.**, as the concentration of fireflies around optima (i.e. around each other) increased, the standard deviation of cost functions decreased. The reason was that the more the process of the algorithm got closer to the end of the iteration, the more the fireflies and their cost functions got closer to form neighborhoods.

VI. CONCLUSION

This paper proposed CFA multimodal firefly algorithm based on the Coulomb's law. This algorithm was successful in solving multimodal optimization problems. Results of experiments indicated that this unimodal optimization algorithm was successfully turned into a multimodal optimization algorithm through modifications. Two of the advantages of this algorithm are quickly yielding optimal results and not requiring additional parameter for being turned into a multimodal algorithm. According to the results, this algorithm can be considered as a reliable multimodal optimization algorithm.

TABLE II. THE RESULTS OF THE EXPERIMENTS - ACCURACY: 10-1

Function	Algorithm	Success Rate	Average Optima Found	Global Average Optima Found	Mean Peak Ration	Precision	Success Performance	Average Function Evaluation
Func-1	**CFA**	0	94.3	1	**1.013023**	**0.945670**	Inf	122000
	FER-PSO	0	37.38	1	0.94486	0.635588	Inf	102000
	EPSO	0	58.18	1	0.981776	0.664516	Inf	102000
	LS-FER-PSO	**0.08**	**108.68**	1	0.649761	0.908604	Inf	204000
Func-2	**CFA**	0	97.4	**1**	1.03153	0.944675	Inf	122000
	FER-PSO	0	58.9	0.98	1.235763	0.663146	Inf	102000
	EPSO	0	66.38	**1**	**0.996027**	0.766224	Inf	102000
	LS-FER-PSO	**0.52**	**110.3**	**1**	0.992886	**0.995518**	**927272.7**	204000
Func-3	**CFA**	0	147.34	**5.14**	**1.027689**	**0.962784**	Inf	170800
	FER-PSO	0	119.62	3.74	1.093445	0.805226	Inf	142800
	EPSO	0	62.18	3.02	1.040867	0.911363	Inf	142800
	LS-FER-PSO	0	**197.12**	4.2	0.190057	0.959919	Inf	285600
Func-4	**CFA**	0.14	4.56	2	**1.091253**	**0.984000**	87142.86	12200
	FER-PSO	0.02	4.2	2	1.135589	0.960333	510000	5100
	EPSO	0.12	4.7	2	1.022293	0.768690	**42500**	5100
	LS-FER-PSO	**0.98**	**5.98**	2	0.788084	0.976000	1020000	20400

Func-5	**CFA**	**1**	**4**	**4**	**0.888415**	0.456564	**12200**	12200
	FER-PSO	0.02	1.92	1.92	1.265238	0.03479	510000	5100
	EPSO	0.04	2.34	2.34	0.859798	0.037773	255000	5100
	LS-FER-PSO	**1**	**4**	**4**	2.205311	**0.466675**	20400	20400
Func-6	**CFA**	0.42	4.14	4.14	1.09713	0.855333	14523.80	6100
	FER-PSO	**1**	**5**	**5**	1.126117	0.648757	**5100**	5100
	EPSO	0	1	1	**1**	**1**	Inf	5100
	LS-FER-PSO	**1**	**5**	**5**	1.173182	0.868849	10200	10200
Func-7	**CFA**	0.12	3.14	1	**0.990534**	**1**	50833.33	6100
	FER-PSO	0.94	4.94	1	1.263032	0.812786	5425.532	5100
	EPSO	0	1	1	1.058738	**1**	Inf	5100
	LS-FER-PSO	**1**	**5**	1	1.144458	0.889643	**10200**	10200

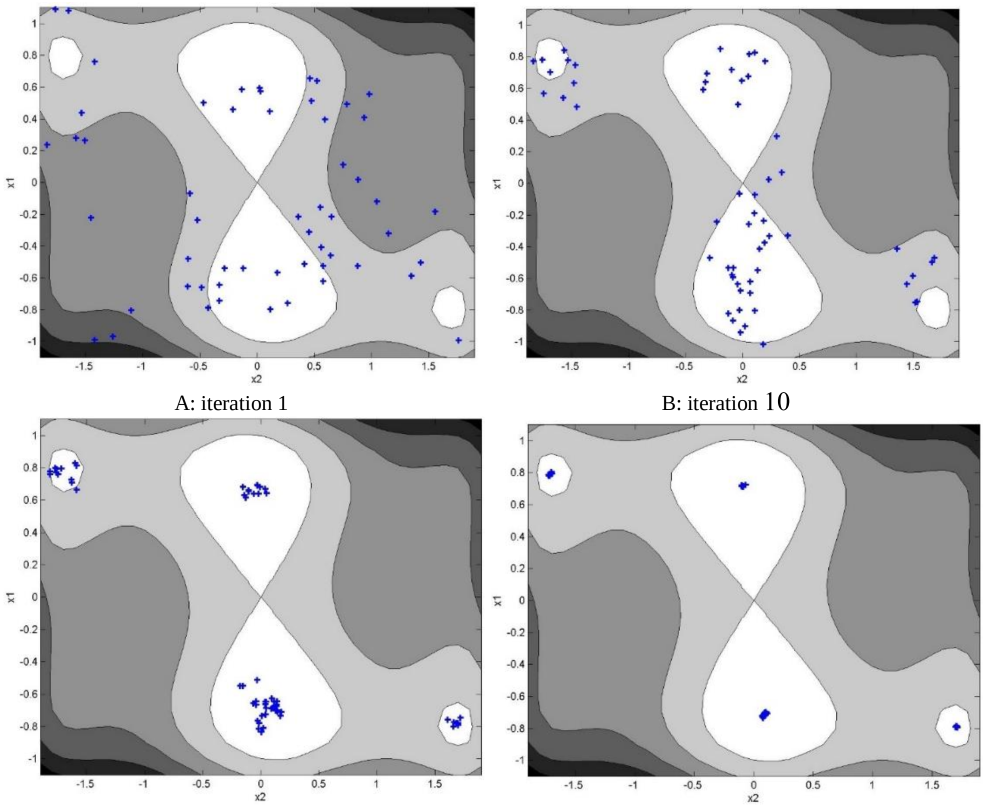
TABLE III. THE RESULTS OF THE EXPERIMENTS - ACCURACY: 10-2

Function	Algorithm	Success Rate	Average Optima Found	Global Average Optima Found	Mean Peak Ration	Precision	Success Performance	Average Function Evaluation
Func-1	**CFA**	0	**53.8**	**1**	**1.020978**	**0.621089**	Inf	122000
	FER-PSO	0	6.38	0.64	0.630055	0.117527	Inf	102000
	EPSO	0	20.88	0.94	1.026456	0.248781	Inf	102000
	LS-FER-PSO	0	43.14	**1**	1.109575	0.616820	Inf	204000
Func-2	**CFA**	0	57.88	**1**	1.054625	0.562889	Inf	122000
	FER-PSO	0	7.64	0.24	1.245746	0.091978	Inf	102000
	EPSO	0	21.52	0.78	1.055243	0.247988	Inf	102000
	LS-FER-PSO	0	**114.34**	**1**	**1.000549**	**0.953799**	Inf	204000
Func-3	**CFA**	0	**85.08**	**5**	**1.057686**	**0.627016**	Inf	170800
	FER-PSO	0	10.88	0.68	0.92454	0.072021	Inf	142800
	EPSO	0	38.32	2.7	1.072754	0.580346	Inf	142800
	LS-FER-PSO	0	84.04	4	1.060134	0.603456	Inf	285600
Func-4	**CFA**	**0.06**	**4.46**	**2**	0.951791	**0.931333**	**203333.3**	12200
	FER-PSO	0	3.88	**2**	0.916055	0.890000	Inf	10200
	EPSO	0	3.44	1.76	**1.038658**	0.552857	Inf	10200
	LS-FER-PSO	0	4.02	**2**	1.173043	0.898333	Inf	20400
Func-5	**CFA**	0	**4**	**4**	**0.926575**	**0.408128**	**12200**	12200
	FER-PSO	0	0.46	0.46	0.846641	0.008104	Inf	10200
	EPSO	0	0.58	0.58	0.861226	0.018275	Inf	10200
	LS-FER-PSO	**1**	**4**	**4**	1.513459	0.404559	20400	20400
Func-6	**CFA**	0.02	3.02	3.02	0.866844	0.575548	305000	6100
	FER-PSO	**1**	**5**	**5**	0.995434	0.636006	**5100**	5100
	EPSO	0	1	1	**1**	**1**	Inf	5100
	LS-FER-PSO	**1**	**5**	**5**	1.159774	0.87873	10200	10200
Func-7	**CFA**	**1**	2.96	1	0.980653	0.796000	Inf	6100
	FER-PSO	**1**	**5**	1	0.824626	0.777659	**5100**	5100
	EPSO	0	1	1	**1.058738**	**1**	Inf	5100
	LS-FER-PSO	**1**	**5**	1	1.090589	0.92869	10200	10200

TABLE IV. THE RESULTS OF THE EXPERIMENTS - ACCURACY: 10-3

Function	Algorithm	Success Rate	Average Optima Found	Global Average Optima Found	Mean Peak Ration	Precision	Success Performance	Average Function Evaluation
Func-1	**CFA**	0	**20.24**	0.86	**1.069871**	**0.228672**	Inf	122000
	FER-PSO	0	0.78	0.22	1.437689	0.014452	Inf	102000
	EPSO	0	8.06	0.5	1.103679	0.092075	Inf	102000
	LS-FER-PSO	0	13.22	**1**	1.123158	0.189530	Inf	204000
Func-2	**CFA**	0	18.1	0.62	1.030118	0.175755	Inf	122000
	FER-PSO	0	0.78	0	1.212728	0.009763	Inf	102000
	EPSO	0	8.44	0.42	0.986824	0.097171	Inf	102000
	LS-FER-PSO	0	**55.42**	**1**	**0.992886**	**0.461895**	Inf	204000
Func-3	**CFA**	0	**36.1**	**4.08**	**1.011276**	**0.266552**	Inf	170800
	FER-PSO	0	0.7	0.04	0.977142	0.004798	Inf	142800
	EPSO	0	16.04	1.08	1.03949	0.235603	Inf	142800
	LS-FER-PSO	0	24.14	4	1.039162	0.170459	Inf	285600
Func-4	**CFA**	**0**	**3.84**	1.88	**0.930002**	**0.845333**	Inf	12200
	FER-PSO	0	2.08	**2**	0.927387	0.470333	Inf	10200
	EPSO	0	1.28	0.78	0.785126	0.202500	Inf	10200
	LS-FER-PSO	0	3.14	**2**	1.182125	0.692333	Inf	20400
Func-5	**CFA**	0.9	3.88	3.88	0.798721	**0.412415**	13555.56	12200

	FER-PSO	0	0.24	0.24	1.702931	0.004307	Inf	10200
	EPSO	0	0.1	0.1	0.782353	0.001551	Inf	10200
	LS-FER-PSO	**1**	**4**	**4**	**1.054611**	0.170459	**20400**	20400
Func-6	**CFA**	0.06	2.76	2.76	1.29713	0.519389	101666.7	6100
	FER-PSO	**1**	**5**	**5**	1.012981	0.629383	**5100**	5100
	EPSO	0	1	1	**1**	**1**	Inf	5100
	LS-FER-PSO	**1**	**5**	**5**	1.208592	0.834246	10200	10200
Func-7	**CFA**	0	1.74	1	1.713465	0.528333	Inf	6100
	FER-PSO	0.96	4.96	1	1.368622	0.743413	**5312.5**	5100
	EPSO	0	1	1	**1.058738**	**1**	Inf	5100
	LS-FER-PSO	**1**	5	1	1.092660	0.922857	10200	10200



A: iteration 1 B: iteration 10

C: iteration 20 D: iteration 30

Fig. 2. the position of particles in different iterations (beginning to end)

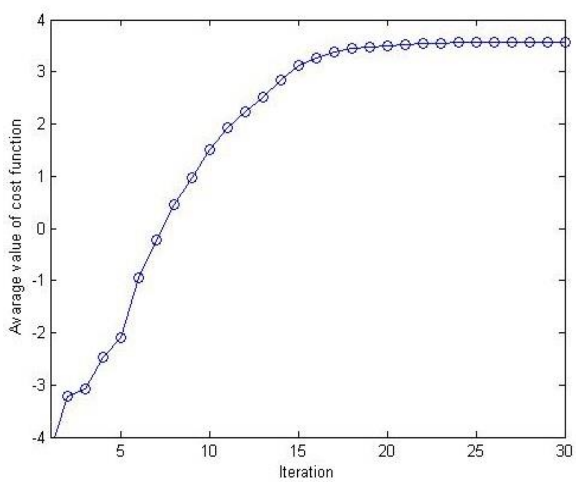


Fig. 3. The average value of cost function in each iteration

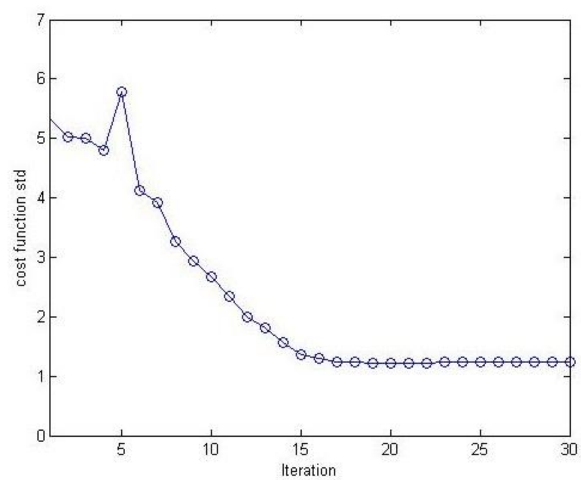


Fig. 4. The standard deviation of cost fucntions of fireflies in each iteration

References

[1] B.-Y. Qu, J. J. Liang, and P. N. Suganthan, "Niching particle swarm optimization with local search for multi-modal optimization," Information Sciences, vol. 197, pp. 131-143, 2012.

[2] X. Li, "A multimodal particle swarm optimizer based on fitness Euclidean-distance ratio," in Proceedings of the 9th annual conference on Genetic and evolutionary computation, 2007, pp. 78-85.

[3] J. Barrera and C. A. C. Coello, "A particle swarm optimization method for multimodal optimization based on electrostatic interaction," in MICAI 2009: Advances in Artificial Intelligence, ed: Springer, 2009, pp. 622-632.

[4] M. Li, D. Lin, and J. Kou, "A hybrid niching PSO enhanced with recombination-replacement crowding strategy for multimodal function optimization," Applied Soft Computing, vol. 12, pp. 975-987, 2012.

[5] E. Özcan and M. Yılmaz, "Particle swarms for multimodal optimization," in Adaptive and Natural Computing Algorithms, ed: Springer, 2007, pp. 366-375.

[6] T. Rahkar-Farshi, S. Behjat-Jamal, and M.-R. Feizi-Derakhshi, "An improved multimodal PSO method based on electrostatic interaction using n-nearest-neighbor local search," arXiv preprint arXiv:1410.2056, 2014.

[7] T. Grüninger and D. Wallace, "Multimodal optimization using genetic algorithms," Master's thesis, Stuttgart University, 1996.

[8] E. Dilettoso and N. Salerno, "A self-adaptive niching genetic algorithm for multimodal optimization of electromagnetic devices," Magnetics, IEEE Transactions on, vol. 42, pp. 1203-1206, 2006.

[9] R. K. Ursem, "Multinational GAs: Multimodal Optimization Techniques in Dynamic Environments," in GECCO, 2000, pp. 19-26.

[10] T. Rahkar-Farshi, O. Kesemen, and S. Behjat-Jamal, "Multi hyperbole detection on images using modified artificial bee colony (ABC) for multimodal function optimization," in Signal Processing and Communications Applications Conference (SIU), 2014 22nd, 2014, pp. 894-898.

[11] X. Li, "Adaptively choosing neighbourhood bests using species in a particle swarm optimizer for multimodal function optimization," in Genetic and Evolutionary Computation–GECCO 2004, 2004, pp. 105-116.

[12] X. Li, "Niching without niching parameters: particle swarm optimization using a ring topology," Evolutionary Computation, IEEE Transactions on, vol. 14, pp. 150-169, 2010.

[13] J. Zhang, J.-R. Zhang, and K. Li, "A sequential niching technique for particle swarm optimization," in Advances in Intelligent Computing, ed: Springer, 2005, pp. 390-399.

[14] X.-S. Yang, "Firefly algorithms for multimodal optimization," in Stochastic algorithms: foundations and applications, ed: Springer, 2009, pp. 169-178.

[15] S. W. Mahfoud, "Niching methods for genetic algorithms," Urbana, vol. 51, 1995.

[16] J.-H. Seo, C.-H. Im, C.-G. Heo, J.-K. Kim, H.-K. Jung, and C.-G. Lee, "Multimodal function optimization based on particle swarm optimization," Magnetics, IEEE Transactions on, vol. 42, pp. 1095-1098, 2006.

[17] M. Li and J. Kou, "Crowding with nearest neighbors replacement for multiple species niching and building blocks preservation in binary multimodal functions optimization," Journal of Heuristics, vol. 14, pp. 243-270, 2008.

[18] A. Passaro and A. Starita, "Particle swarm optimization for multimodal functions: a clustering approach," Journal of Artificial Evolution and Applications, vol. 2008, p. 8, 2008.

[19] K. E. Parsopoulos and M. N. Vrahatis, "On the computation of all global minimizers through particle swarm optimization," Evolutionary Computation, IEEE Transactions on, vol. 8, pp. 211-224, 2004.

[20] K. D. Koper, M. E. Wysession, and D. A. Wiens, "Multimodal function optimization with a niching genetic algorithm: A seismological example," Bulletin of the Seismological Society of America, vol. 89, pp. 978-988, 1999.

[21] A. Anderson, C. McNaught, J. MacFie, I. Tring, P. Barker, and C. Mitchell, "Randomized clinical trial of multimodal optimization and standard perioperative surgical care," British journal of surgery, vol. 90, pp. 1497-1504, 2003.

[22] K. Parsopoulos and M. Vrahatis, "Modification of the particle swarm optimizer for locating all the global minima," Artificial Neural Networks and Genetic Algorithms, pp. 324-327, 2001.

[23] R. Brits, A. P. Engelbrecht, and F. van den Bergh, "Locating multiple optima using particle swarm optimization," Applied Mathematics and Computation, vol. 189, pp. 1859-1883, 2007.

[24] J. Liang, T. P. Runarsson, E. Mezura-Montes, M. Clerc, P. Suganthan, C. A. C. Coello, et al., "Problem Definitions and Evaluation Criteria for the CEC 2006 Special Session on Constrained Real-Parameter Optimization," 2006.

The Impact of Privacy Concerns and Perceived Vulnerability to Risks on Users Privacy Protection Behaviors on SNS: A Structural Equation Model

Noora Sami Al-Saqer and Mohamed E. Seliaman
Department of Information Systems, King Faisal University,
P.O. Box: 400, Al-Hufof, Al-Hasa, 31982,
Saudi Arabia

***Abstract*—This research paper investigates Saudi users' awareness levels about privacy policies in Social Networking Sites (SNSs), their privacy concerns and their privacy protection measures. For this purpose, a research model that consists of five main constructs namely information privacy concern, awareness level of privacy policies of social networking sites, perceived vulnerability to privacy risks, perceived response efficacy, and privacy protecting behavior was developed. An online survey questionnaire was used to collect responses from a sample of (108) Saudi SNSs users. The study found that Saudi users of social networking sites are concerned about their information privacy, but they do not have enough awareness of the importance of privacy protecting behaviors to safeguard their privacy online. The research results also showed that there is a lack of awareness of privacy policies of Social networking sites among Saudi users. Testing hypothesis results using the Structural Equation Modeling (SEM) showed that information privacy concern positively affects privacy protection behaviors in SNSs and perceived vulnerability to privacy risks positively affects information privacy concern.**

Keywords—Social networking sites (SNSs); information privacy concern; perceived vulnerability; SEM; protection behavior

I. INTRODUCTION

TODAY, the number of internet users in Saudi Arabia has reached 19.6 million [1]. With the continuous development in internet technologies over the years, smartphones revolution and the web mobile internet, social networking sites became a need for every internet user. SNSs are now the most preferred communication choice of users in today's context. The benefits that the SNSs provide are indeed remarkable on many dimensions, including reducing the financial cost especially with the reasonable prices of internet services in Saudi Arabia. These Social networking services also help the users to go beyond the geographic locations, and make it easier to connect and communicate with people all over the world.

Every new invention in technology field including the SNSs, is intended to simplify users' lives by helping them to follow the modern era of speed and technology. However on another side, using SNSs is associated with some privacy risks such as the misuse of users personal information with serious personal and social implications. The users' lack of awareness of privacy policies and the consequences associated with it, might make it easier to breach the personal privacy and increase cyber-crimes. Some SNSs are requesting users to provide sensitive information including some personal or private details. Some users are providing these information easily without even thinking about the consequences of providing such information, and without knowing that these information and details are being sold and shared to third parties for marketing reasons. Most SNSs are clearly publishing their privacy policies (PP) regarding information sharing but many users do not pay enough attention to figure out the details components of each SNSs' PP.

Only few research studies investigated the relationship between SNSs' PP awareness level and its effects on raising privacy concern and privacy protecting behaviors.

The main objectives of this research paper are : (1)Measure the level of awareness among Saudi SNSs users of PP, (2) measure the information privacy concern of Saudi SNSs users, (3) measure the effect of information privacy concern and awareness of privacy policies on privacy protecting behaviors.

II. RELATED RESEARCH

Privacy has been interpreted as the ''boundary control process in which individuals regulate when, how, and to what extent information about them is communicated to others'' [2]. Maintaining Internet users privacy is a legal and human right of person regarding information disclosing, storing, miss maintaining, miss using, and transmitting through internet based applications including SNSs, web sites, and search engines [3].

Privacy policy is about principles of actions adopted by an individual or an organization in protecting their personal information and serves as guideline for users who would like to share their information [4]. In today's context where technology is interfering in everything we do, millions of users are vulnerable to privacy threats. SNSs provide easy to use privacy settings to users. These settings provide visibility and privacy options for a user profile to limit the access to certain people like: family and friends, or set the profile as public. However, most people don't change the privacy settings and set their SNSs account as default[3]. The use of SNSs varies from shopping, communication, expressing personal ideas and feelings to organizational matter of use.

Although there are large advantages offered by SNSs, there are also many challenges towards the information privacy from personal or organizational aspects. Moreover, most users don't know about the vulnerability of their information while using SNSs [4]. Some SNSs provide comprehensive, updated and detailed privacy policy, so the user can understand the potential risks when posting their information online. However, the more details and long comprehensive text make users ignore it, do not read it, and just clicking the agree button without understanding the implications of its content. People usually hate reading long texts especially when the PP includes technical jargons which are hard to understand for many users. Other reason to the lack of awareness of PP is the absence of engagements of users while developing the PP[4].

Privacy policies clearly mention about saving, sharing, and modifying the users' information and contents they post. Since many users do not read privacy policies, few people know about these details. The Saudi society in particular is conservative and religious society influencing its people attitudes and behaviors [5]. Hence, focusing on users' privacy and raising their awareness level of privacy threats is an important aspect. Information privacy is a critical issue in online environment as online companies depend on collecting large amounts of personal information about users [6].

Online users are concerned about information privacy when surfing the Web because the access to their personal information cannot be controlled meaning that information privacy is threatened [6, 7]. User concern about information privacy was found as an important factor which has a positive effect on his/her behavioral intention to practice privacy protection measures[8].

Previous studies of SNSs in Saudi Arabia focused on the reasons that motivate Saudi females to use Facebook. One study shows an existence of privacy concern among Saudi female users [5]. Studies [9, 10]found that Saudi women have higher levels of privacy concerns as compared to Saudi men. Regardless of gender differences for privacy behavior, gender does not predict worries about the ways third parties use personal information [11]. According to [11] users are aware of privacy protecting measures to protect themselves from threats and the users who are concerned about their information privacy are more likely to apply these privacy protecting measures.

While the research work in [11] used a simple binary variable (Yes, No) to measure the privacy protection measure use, this paper contributes to the SNS privacy research by developing and validating new measurement items operationalizing the construct of privacy protection behavior in SNS. In addition, this paper contributes to SNS research by developing new construct 'awareness level of privacy policies' and validating new self-developed measurement items operationalizing the construct. While research [11] was conducted in Malaysia, this paper validated the extended model with the new added construct of privacy protection behavior in the context of Saudi Arabia which is a different social context. Moreover, research [11] targeted a sample of only undergraduates at a public Malaysian university while this research targeted a more general sample of public respondents including different age groups and education levels.

III. Research Model and Hypothesis

This study followed the quantitative approach using a theoretical model as shown in Figure 1 to achieve the above-mentioned objectives. The research model developed for this study consists of the following five constructs:

1) Information privacy concern: It is the "extent to which an individual is concerned about organizational practices related to the collection and use of his or her personal information" [11].

2) Awareness level of privacy policies of SNSs: This construct is intended to understand the degree to which higher level of awareness of PP of SNSs affects the privacy protecting behavior in SNSs. It refers to SNSs users awareness and understanding of Privacy policies in SNSs.

3) Privacy protection behavior in SNSs: It is the adaption of protective behaviors to guard the privacy [12] .

4) Perceived vulnerability to privacy risks: is the degree to which a SNSs user believes a privacy threats will occur to him\her [11].

5) Perceived response efficacy: is the belief that a recommended protecting measure is effective in protecting the self and others from a threat [11].

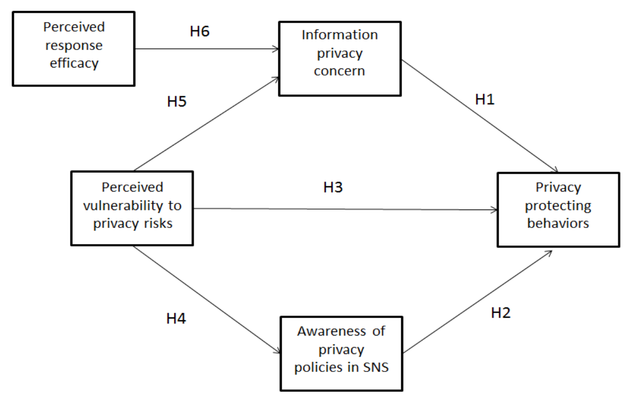


Fig. 1. Research model

The model is derived from previous research models grounded in several theories such as: (1) information privacy concern was introduced by two theories: agency theory and social contract theory, both suggest that privacy concern exist in online transactions due to incomplete information about online behavior of customer information [11][17[11]] . In our study we focus on the information privacy concern among SNSs' users regarding their information to understand its effect on privacy protection behavior. However, neither the agency theory nor the social contract theory provided applicable framework for empirical research [17]. (2) Awareness level of privacy policies of SNSs was self-developed to study the impact of reading and understanding the content of SNSs privacy policy on raising the adoption of privacy protection behaviors. (3) Privacy protection behavior in SNSs: the human behavior was introduced in many information systems theories such as: Social Cognitive Theory, Theory of Reasoned Action (TRA), and theory of planned behavior (TPB), and Protection motivation Theory (PMT) in each theory behavior was interpreted in different perspectives [12] [11][17]. In our model we intended to understand the factors that influence actions and practices SNSs users do to protect their information privacy by studying the relationships between perceived vulnerability to privacy risk, information privacy concern, awareness level of PP in SNSs, and privacy protection behavior. (4) Perceived vulnerability to privacy risks was derived from the protection motivation theory (PMT), our focus here is to understand the relationship between this construct and information privacy concern, awareness level of SNSs PP, and privacy protection behaviors. (5) Perceived Response – efficacy was derived from protection motivation theory (PMT); it was added to understand its effect on information privacy concern.

Based on the research model described above, the following hypotheses were developed:

H 1: Information privacy concern positively affects privacy protection behaviors in SNS.

H2: Higher awareness of SNSs' PP positively affects privacy protection behaviors in SNSs.

H3: Perceived vulnerability to privacy risks positively affects privacy protecting behaviors.

H4: Perceived vulnerability to privacy risks increase awareness level of privacy policies in SNS.

H5: Perceived vulnerability to privacy risks positively affects information privacy concern.

H6: Perceived response efficacy positively affects information privacy concern.

IV. Research Methodology

A. Sampling and Data collection

An online survey questionnaire was developed to collect data required for assessing the SNSs users' awareness of PP. The survey questionnaire consisted of 22 questions and was divided into two sections. The first section contained questions about user's demographic information. The second section contained questions related to the research model five constructs. A sample of (108) respondents filled out the survey. Table 1 shows the survey items. Table 2 shows the basic demographic data of respondents. Majority of participants are young females and adults belong to the age group of 2 to 30 years. Most of them are educated holding bachelor degree.

TABLE I. Survey Items

Construct	Item
Information Privacy Concern	**IPC1:** I am concerned about submitting my personal information in social networking sites because of what others might do with it [11]. **IPC2:** I am concerned about submitting my personal information in social networking sites because it could be used in a way I did not foresee [11].
perceived vulnerability to privacy risks	**PVPR1:** I could be subjected to a malicious computer/ information security problems (e.g. virus, privacy, identity theft, hacking and etc.) in social networking sites [11]. **PVPR2**: I feel my personal information in social networking sites could be misused [11]. **PVPR3** I feel my personal information in social networking sites could be made available to unknown individuals or companies without my knowledge [11]. **PVPR4:** I feel my personal information in social networking sites could be made available to government agencies [11]. **PVPR5:** I feel my personal information in social networking sites could be inappropriately used [11].
Awareness level of Privacy Policies in SNSs	**APPS1**: I read privacy policies of SNSs before using them. **APPS2:** I understand carefully what is mentioned in the Privacy policies of SNSs. **APPS3:** I read the updated versions of SNSs PP.
Perceived Response - efficacy	**PRE1**: If I used privacy protection measures in social networking sites, I could probably protect myself from losing my information privacy [11]. **PRE2:** I can protect my information privacy better if I use privacy protection measures in social networking sites [11]. **PRE3:** Utilizing privacy protection measures in social networking sites works to ensure my information privacy [11]. **PRE4:** If I utilize privacy protection measures in social networking sites, I am less likely to lose my information privacy [11].
Privacy Protection behavior in SNSs	**PPBS 1:** I do not share my personal information (like: mobile number, personal photos, personal events) on SNSs. **PPBS 2:** I tend to be careful about sharing my personal information (like: mobile number, personal photos, etc.) while using SNSs.

TABLE II. Demographic Data

Measure	Item	Frequency	Percent
Gender	Male	18	26.5%
	Female	50	73.5%
Age	Less than 20	13	12%
	From 21 to 30	63	58.3%
	From 31 to 40	24	22.2%
	From 41 to 50	7	6.5%
	More than 50	1	0.9%
Education	Pre high school	7	6.5%
	High school	14	13%
	College	52	48.1%
	Post graduate	35	32.4%

V. Results and Discussion

A. The measurement model

a) Reliability for the measurements: The sample consisted of 108 participants. WarpPls 5.0 was used to assess the reliability and validity of the the survey questionnaire as the main measurement instrument of this research study. The survey questionnaire measurement tool included 16 items forming 5 latent variables beside the moderators. Cronbach's Alpha (CA) was used to assess constructs reliability. CA estimates the inter-correlations of the indicators [18]. The acceptable score for CA is 0.7 and higher [18]. In addition, the constructs reliability was tested using Composite reliability (CR). CR or the internal consistency reliability readings unlike CA, takes into account the different loadings of the indicators. The acceptable score for CR should be 0.7 and higher [15]. As shown in table 3, all constructs reported Cronbach's Alpha values above the acceptable thresholds of 0.7. In addition, all CR values above the acceptable thresholds of 0.7.

TABLE III. MEASUREMENT RELIABILITY TESTING RESULTS

Construct	No of items	Cronbach's alpha	Composite reliability
Information privacy concern	2	0.791	0.905
Perceived vulnerability to privacy risks	5	0.832	0.882
Awareness level of privacy policies	3	0.851	0.910
Perceived response efficacy	4	0.868	0.911
Privacy protecting behaviors in SNS	2	0.761	0.893

b) Factor loadings: Factor loadings for the measured variables have to be at least 0.5 or the variable becomes a candidate for removal [15]. The factor loadings were calculated for the measured variables using confirmatory factor analysis as shown in table 4. All variables have statisified the loading value above 0.5 required for inclusion in the model.

TABLE IV. FACTOR LOADING FOR MEASURED VARIABLES

	IPC	PVPR	APPS	PRE	PPBS	SE	P value
IPC1	(0.909)					0.076	<0.001
IPC2	(0.909)					0.076	<0.001
PVPR1		(0.749)				0.079	<0.001
PVPR2		(0.846)				0.077	<0.001
PVPR3		(0.767)				0.079	<0.001
PVPR4		(0.757)				0.079	<0.001
PVPR5		(0.748)				0.079	<0.001
APPS1			(0.868)			0.077	<0.001
APPS2			(0.896)			0.076	<0.001
APPS3			(0.869)			0.077	<0.001
PRE1				(0.803)		0.078	<0.001
PRE2				(0.872)		0.077	<0.001
PRE3				(0.897)		0.076	<0.001
PRE4				(0.815)		0.078	<0.001
PPBS1					(0.898)	0.076	<0.001

c) The validity assessment:: Convergent validity was assessed by calculating composite reliability and the Average Variance Extracted (AVE) for each latent construct. Convergent validity is "the extent to which a measure is related to other measures which have been designed to assess the same construct" [14]. AVE is an indicator of convergence that is used to calculate the mean variance extracted for the construct items [15]. The composite reliability coefficients for all constructs are greater than the critical value of 0.7. In addition, all constructs reported an AVE score exceeding 0.5 as shown in Table 5.

TABLE V. CONVERGENT VALIDITY STATISTICS FOR THE CONSTRUCTS

Construct	Composite reliability	AVE
information privacy concern	0.905	0.827
perceived vulnerability to privacy risks	0.882	0.600
Awareness level of privacy policies	0.910	0.771
Perceived response efficacy	0.911	0.718
Privacy protecting behaviors in SNS	0.893	0.807

Discriminant validity refers to "the extent in which a construct is truly distinctive from other constructs" [15].The square root of the Average Variance Extracted (AVE) for each latent construct was calculated to assess the discriminant validity and then comparing the values with the other latent constructs correlations. All constructs reported an AVE score exceeding 0.5. In addition, the square root of AVE for each construct is greater than all correlations of other constructs supporting the measurement discriminant validity as shown in **Table 6**.

TABLE VI. DISCRIMINANT VALIDITY FOR THE CONSTRUCTS

	IPC	PVPR	APPS	PRE	PPBS
IPC	(0.909)				
PVPR	0.645	(0.774)			
APPS	0.147	0.173	(0.878)		
PRE	0.151	0.224	0.087	(0.848)	
PPBS	0.435	0.404	0.102	0.177	(0.898)

d) The structural model: Structural Equation Modeling (SEM) is a multivariate statistical technique used to test the model hypotheses that describes relationships between variables [13]. SEM was used to test the hypothesis of the research model. structural equation modeling (SEM) belongs to the second generation data analysis method. SEM is a multivariate statistical technique used to test the model hypotheses that describes relationships between variables [13]. SEM is mainly used by researchers because it considers the measurements errors when analysing data statistically [13, 17]. SEM is preferred by researchers because it is not only used to assess the structural model – the assumed relationship between multiple independent and dependent constructs – but at the same time, it also asseses the measurement model – the loadings of observed measurement items on their latent constructs. SEM allows for examining reliability and validity of the measurements with the testing of the hypotheses [13].

e) Goodness of fit measures (GOF): Goodness of Fit measures indicates "how well the specified model reproduces the observed covariance matrix among the indicator items" [15]. In structural equation modeling (SEM) it is important to assess whether a specified model fits the data or not. Goodness of fit measures provide the most necessary indication of how well the proposed theory fits the data. Different indices as shown in Table 7 reflect a different aspect of model fit and present the acceptable value for each fit measure [19].

The confirmatory factor analysis calculated ten Goodness of Fit measures as shown in Table 7.

TABLE VII. GOODNESS OF FIT MEASURES

GOF of research model	Acceptable value
Average path coefficient (APC)=0.250, P=0.002	Good if p< 0.05 [16]
Average R-squared (ARS)=0.237, P=0.003	Good if p< 0.05 [16]
Average block VIF (AVIF)=1.374	acceptable if <= 5, ideally <= 3.3 [16]
Average full collinearity VIF (AFVIF)=1.389	acceptable if <= 5, ideally <= 3.3 [16]
Tenenhaus GoF (GoF)=0.420	small >= 0.1, medium >= 0.25, large >= 0.36 [16]
Sympson's paradox ratio (SPR)=1.000	acceptable if >= 0.7, ideally = 1 [16]
R-squared contribution ratio (RSCR)=1.000	acceptable if >= 0.9, ideally = 1 [16]
Statistical suppression ratio (SSR)=1.000	acceptable if >= 0.7 [16]
Nonlinear bivariate causality direction ratio (NLBCDR)=1.000	acceptable if >= 0.7 [16]

f) Hypotheses testing results: significant level chosen for testing hypothesis is at <=0.05. The results of testing the research model are presented in Figure 2, and summarized in Table 8.

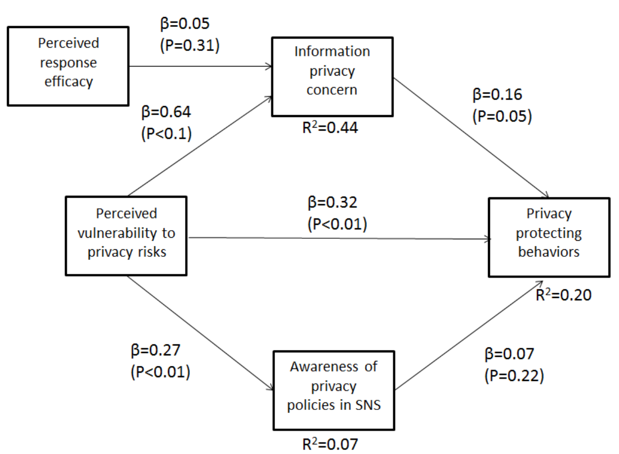


Fig. 2. Structural Equation Model Results

TABLE VIII. MODEL RESULTS

Hypothesis	Test result	Conclusion
H 1: Information privacy concern positively affects privacy protection behaviors in SNS.	(beta=0.16, p=0.05)	Supported
H2: Higher awareness of SNSs' PP positively affects privacy protection behaviors in SNSs.	(beta=0.07, p=0.22)	Not Supported
H3: perceived vulnerability to privacy risks positively effects privacy protecting behaviors.	(beta=0.32, p<0.01)	Supported
H4: perceived vulnerability to privacy risks increase awareness level of privacy policies in SNS.	(beta=0.27, p<0.01)	Supported
H5: perceived vulnerability to privacy risks positively effects information privacy concern.	(beta=0.64, p<0.01)	Supported
H6: perceived response efficacy positively effects information privacy concern.	(beta=0.05, p=0.31)	Not supported

Hypothesis 1 suggests a positive relationship between information privacy concern and privacy protection behavior. This Hypothesis was supported and it was found that information privacy concern positively influences privacy protecting behavior (beta=0.16, p=0.05). Hypothesis 2 proposes a positive relationship between awareness levels of privacy policies of social networking sites and privacy protection behaviors. It was found that awareness level of PP does not affect privacy protecting behaviors (beta=0.07, p=0.22) which does not support hypothesis 2. Hypothesis 3 suggests that perceived vulnerability to privacy risks positively affects privacy protecting behaviors. It was found that perceived vulnerability to privacy risks positively contributes to the adaption of privacy protecting behaviors (beta=0.32, p<0.01) supporting hypothesis 3. Hypothesis 4 proposes that perceived vulnerability to privacy risks increases awareness level of privacy policies in SNS. The perceived vulnerability to privacy risks positively contributes to awareness level of PP of SNSs (beta=0.27, p<0.01). Hence, hypothesis 4 is supported. Hypothesis 5 proposes that perceived vulnerability to privacy risks positively affects information privacy concern. Results showed that perceived vulnerability to privacy risks positively contributes to information privacy concern (beta=0.64, p<0.01) and that supports hypothesis 5. Hypothesis 6 suggests that perceived response efficacy positively affects information privacy concern. It was found that this hypothesis 6 is not supported.

VI. CONCLUSIONS

This study aimed to investigate Saudi users' awareness levels about privacy policies of Social Networking Services (SNSs) and the factors affect privacy protecting behaviors. A research model that consists of five constructs was developed for this purpose. The study sample consists of (108) participants who were surveyed to collect the data. The research model was assessed using WarpPLS 5.0 software.

Testing the hypotheses results supported all proposed hypotheses except for hypothesis two and hypothesis six. The study found that Saudi users' privacy concerns and perceived vulnerability to privacy risks in social networking sites are significant antecedents of their privacy protection behaviors. In addition, users' perceived vulnerability to privacy risks in social networking sites positively influences their awareness levels of the content of privacy policies of SNS. In contrast with our expectation, it was found that awareness level of privacy policies does not necessarily influence privacy protection behavior in social networking sites. Also, it was found that perceived response efficacy does not influence the information privacy concern which is consistent with previous studies in the literature.

Similar to any other study, this research has some limitations. Contrasting the expectations, the research results did not report any correlation between awareness level of privacy policies and user privacy protecting behaviors. This issue can be investigated further considering more sophisticated measurement items for some constructs in the current research model. The sample size can also be increased in future research and the quantitative methodology can be assisted by a qualitative inquiry for more in depth analysis.

Future studies may enhance the research model by adding more constructs, expand the sample size, and apply mixed methodology that includes qualitative approach to interpret the results.

ACKNOWLEDGMENT

The authors would like to acknowledge the support for this research provided by King Faisal University.

REFERENCES

[1] Anonymous "Communication and information technology commesion annual report 2014," communication and information technology commesion, KSA, 2014.

[2] E. Van De Garde-Perik, P. Markopoulos, B. De Ruyter, B. Eggen and W. Ijsselsteijn, "Investigating privacy attitudes and behavior in relation to personalization," *Soc. Sci. Comput. Rev.,* vol. 26, pp. 20-43, 2008.

[3] S. Talib, N. A. Ismail, A. Olowolayemo, S. Naser, S. Aina, S. Z. Haron, M. Yusof and A. Hanisah, "Social networks privacy policy awareness among undergraduate students: The case of twitter," in *Information and Communication Technology for the Muslim World (ICT4M), 2014 the 5th International Conference on,* 2014, pp. 1-5.

[4] S. Talib, A. Razak, S. Munirah, A. Olowolayemo, M. Salependi, N. F. Ahmad, S. Kunhamoo and S. K. Bani, "Perception analysis of social networks' privacy policy: Instagram as a case study," in *Information andCommunication Technology for the Muslim World (ICT4M), 2014 the 5th International Conference on,* 2014, pp. 1-5.

[5] Y. Al-Saggaf, "Saudi females on Facebook: An ethnographic study," *International Journal of Emerging Technologies and Society,* vol. 9, pp. 1-19, 2011.

[6] M. Zviran, "User's Perspectives on Privacy In Web-Based Applications." *Journal of Computer Information Systems,* vol. 48, 2008.

[7] E. Aimeur, S. Gambs and A. Ho, "UPP: User privacy policy for social networking sites," in *Internet and Web Applications and Services, 2009. ICIW'09. Fourth International Conference on,* 2009, pp. 267-272.

[8] M. L. Korzaan and K. T. Boswell, "The influence of personality traits and information privacy concerns on behavioral intentions," *Journal of Computer Information Systems,* vol. 48, pp. 15-24, 2008.

[9] J. Fogel and E. Nehmad, "Internet social network communities: Risk taking, trust, and privacy concerns," *Comput. Hum. Behav.,* vol. 25, pp. 153-160, 2009.

[10] M. G. Hoy and G. Milne, "Gender differences in privacy-related measures for young adult Facebook users," *Journal of Interactive Advertising,* vol. 10, pp. 28-45, 2010.

[11] N. Mohamed and I. H. Ahmad, "Information privacy concerns, antecedents and privacy measure use in social networking sites: Evidence from Malaysia," *Comput. Hum. Behav.,* vol. 28, pp. 2366-2375, 2012.

[12] Y. Feng and W. Xie, "Teens' concern for privacy when using social networking sites: An analysis of socialization agents and relationships with privacy-protecting behaviors," *Comput. Hum. Behav.,* vol. 33, pp. 153-162, 2014.

[13] J. Recker, *Scientific Research in Information Systems: A Beginner's Guide.* Springer Science & Business Media, 2012.

[14] D. Cramer and D. L. Howitt, *The Sage Dictionary of Statistics: A Practical Resource for Students in the Social Sciences.* Sage, 2004.

[15] J. Hair, W. Black, B. Babin and R. Anderson, "Multivariate Data Analysis Seventh Edition Prentice Hall," 2010.

[16] N. Kock, "WarpPLS 5.0 user manual," *Laredo, TX: ScriptWarp Systems,* 2015.

[17] Y. Li, "Theories in online information privacy research: A critical review and an integrated framework," *Decis. Support Syst.,* vol. 54, pp. 471-481, 2012.

[18] M. Moqbel, *The Effect of the use of Social Networking Sites in the Workplace on Job Performance,* 2012.

[19] D. Hooper, J. Coughlan and M. Mullen, "Structural equation modelling: Guidelines for determining model fit," *Articles,* pp. 2, 2008.

Assessment Model for Language Learners' Writing Practice (in Preparing for TOEFL iBT) Based on Comparing Structure, Vocabulary, and Identifying Discrepant Essays

Duc Huu Pham
Department of English
International University –VNU HCMC
Ho Chi Minh City, Vietnam

Tu Ngoc Nguyen
Department of English
International University –VNU HCMC
Ho Chi Minh City, Vietnam

***Abstract*—This study aims to investigate if learners of English can improve computer-assisted writing skills through the analysis of the data from the post test. In this study, the focus was given to intermediate-level students of English taking final writing tests (integrated and independent responses) in preparation for TOEFL iBT. We manually scored and categorized the students' writing responses into five-point levels for the data to make the software. The results of the study showed that the model could be suitable for computerized scoring for language instructors to grade in a fair and exact way and for students to improve their writing performance through practice on the computer.**

Keywords—Computer-assisted writing skills; computerized scoring; integrated and independent responses; model; posttest

I. INTRODUCTION

The computer has so far been used to assist the assessment of the writing ability of learners of English. This summative assessment helps language instructors to judge the success of their teaching and helps English language learners identify areas that need improvements.

In this paper, we suggest a model to help learners of English to improve their writing skills after an investigation of the Vietnamese students' English performance at a university in Vietnam. There has been significant research on how to assess foreign language students' performance [4]. However, more investigations are needed to develop computer-assisted writing skills for these learners. This study aims at exploiting language criteria with a reference to the scale of the Educational Testing Service [8] as the foundation to build a model that can help to language learners better their writing skills.

The study was carried out to present the development of a computerized assessment to enhance language learners' writing abilities. This study will lead to forming a scoring method, which is more objective and does not involve the participation of many scorers, especially when the individual human factor is always subjective. In this paper, we compared learners' responses with an answer text to find out how much they can match each other. According to [7], a text must consist of collections of clauses, and contextual coherence and cohesion (pronoun/noun reference, ellipsis, substitution). The following is the workflow of document processing.

Language learner's integrated or independent response

Fig 1a. Writing test

Introduction

The materials are concerned with the issue of whether dinosaurs were homoeothermic or poikilothermic creatures. ***The lecturer completely disagrees with the reading's position that*** they could have been homoeothermic. ***This belief is based on*** theories of hibernating patterns and body structure.

Body paragraph I

The reading suggests that because dinosaur fossils were found in the Arctic, they must have been warm-blooded homoiotherms. ***However, the professor contests this, claiming instead that*** dinosaurs were cold-blooded. ***The professor explains that*** the presence of dinosaur fossils in the arctic is a result of the dinosaurs migrating there to hibernate. ***He goes on to say that*** modern reptiles hibernate in cold weather.

Body paragraph II

Also, according to the reading, the adaptation of the dinosaurs' legs underneath their bodies is like that of a mammal or bird and does not resemble modern day poikilothermic, reptilian whose legs are on the sides of their bodies. ***In response, the lecturer says that*** dinosaurs could have adapted this way due to their size. In fact, the professor says that this adaptation was necessary in order for the dinosaur to carry its massive weight. ***This is understandable because dinosaurs*** were hundreds of times the size of modern day reptiles.

Body paragraph III

Finally, the passage claims that because dinosaurs bone structure is similar to that of a modern day homoeothermic mammal or bird, they must have been warm-blooded. ***However, the lecturer refutes this, suggesting that*** dinosaur bone histology is not a result of being warm-blooded. The professor explains that this is because the dinosaurs' rapid growth and evolution adapted their bone structure to carry their large body weight. ***Furthermore, it is noted that*** dinosaur bones would have had to be dense in order to carry their large bodies.

Fig 1b. The answer text [13] in the dataset document sample with highlighted language criteria for assessment.

Assessment: Learner's writing matched with the answer text

Fig 1c. The final stage of assessing a writing

Fig. 1. The process of assessing learner's writing skills

The comparison based on structure and vocabulary and identification of discrepant essays will contribute to transferring the manual scoring to automatic scoring with the higher accuracy. This high precision was enhanced based on the improvement in the comparison of the documents in not only structure and vocabulary but also the whole layout. Also, this model will help raise learners' test scores.

With the language features of a text, we can design the application of the model in which foreign language learners will have their responses assessed. The model will compare and match the responses with the features of a sample answer text based on the language criteria (addressing the topic, organization, coherence and language use) given in the model. This method has the following characteristics.

- Helping learners to raise their autonomy in acquiring a certain level of foreign language,
- The fastest way of practicing language writing skills for some formatted tests,
- Being objective in the assessment of language writing ability,
- Time-saving in marking learners' writing responses in writing tests,
- Being able to be used as a model for the comprehensive automatic scoring of the written tests.

The rest of the paper is organized as follows: In Related Works, we review some literature. In Development, we present the steps of writing the software. In Results, we show how the software work and compare our method with other methods available to validate our work the results. In Discussion, Conclusion and Future work, we propose using our method as the basis to enhance the use of the software in the semantic aspect.

II. Related Works

The literature review in this study analyses some investigations of computer software programs and the relationship between the issues of computer-assisted language learning and the second language acquisition. Accordingly, theory and practice in the second language learning can be matched together by using modern technology. Also, the development of technology has led to the dispensable incorporation of this medium into the instruction process. Therefore, the computer has become an integral part of the learning activity, through which learners can learn language skills [3].

Several studies on software programs reviewed by [5] and [10] have showed that the validity of the automated writing evaluation (AWE) or the automated essay scoring (AES) system, has not been thoroughly ascertained. Though they seem to be positive in some aspects, tools to review the second language through computer technology still do not meet the requirements of the standards of educational software for written communication such as assessing writing tests [4]. There was a correlation between AES scores and instructors' numerical grades and analytic ratings, which shows the usefulness of AES programs to classroom-based formative assessments and has provided support for us to write this paper.

[18] has developed a new version e-rater v.2.0 with 12 features: 4 in identifying errors in grammar, usage, mechanics, and style, 2 in organization and development, 3 in lexical complexity, 2 in pro-specific vocabulary usage, and one in essay length. However, e-rater v.2.0 still needs improving in three ways: (1) providing more different writing aspects through the theories of writing, (2) ameliorating the model process, and (3) expanding the identification of different essays [18].

Based on the characteristics already mentioned in e-rater v.2.0, we combined the treatment of grammar errors, the set sample vocabulary, and the identification of discrepant essays. We referred to the comparison of event models for naïve Bayes text classification [2], the support vector machines [16], and text categorization algorithms [15], construction of dictionary features covering word groups relevant to semantics or n-grams for text classification [12]. Then we used vector space model [9] to classify writings.

III. Development

This study used the collection data of typical 200 responses (100 integrated writing responses in which students had to combine different skills: reading a passage, listening to a lecture, and then write down the responses, and 100 independent responses in which we gave students a topic to write). Different test takers wrote 200 responses in different exams at a university in Vietnam, which were similar to those in the TOEFL iBT. After we had marked them manually, we found out the statistical difference ($p<0.05$) that we presented in the previous work [6]. We statistically list errors (spelling, grammar, vocabulary – content words and function words) that the students made in writing. Then we carried out the process of constructing a prototype for assisting language writing skills as follows.

Step 1: We define the language criteria about [8] and [14] regarding

1) Addressing the topic: Does the essay address the subject given?

2) Organization: Does the essay have an introduction, body paragraphs (including paragraph structure), and a conclusion?

3) Coherence: Does the essay have the connectives that join or make the sentences go smooth?

4) Language use: Does the essay have spelling errors or grammar mistakes?

Step 2: We collected the data to sample the dataset including main words and phrases according to the standard definition of an integrated and independent structure with the main words and phrases in response to separate parts. Based on that, we defined the structure of the integrated dataset.

1. Integrated

$S_1 = \{I, P_I, P_{II}, P_{III}\}$

In which:

- S_1: The dataset following the standard definition of an integrated structure with the main words and phrases in Introduction part, and in every Body Paragraph I, II, III part.

- I: The dataset of common words and phrases in Introduction part.

$I = \{w_1, w_2, \ldots, w_n, p_1, p_2, \ldots, p_n\}$

- P_1: The dataset of words and phrases in Body Paragraph I.

$P_I = \{w_1, w_2, \ldots, w_n, p_1, p_2, \ldots, p_n\}$

- P_2: The dataset of words and phrases in Body Paragraph II.

$P_{II} = \{w_1, w_2, \ldots, w_n, p_1, p_2, \ldots, p_n\}$

- P_3: The dataset of words and phrases in Body Paragraph III.

$P_{III} = \{w_1, w_2, \ldots, w_n, p_1, p_2, \ldots, p_n\}$

- W: Dataset common words in Introduction part.
- P: Dataset common phrases in Introduction part

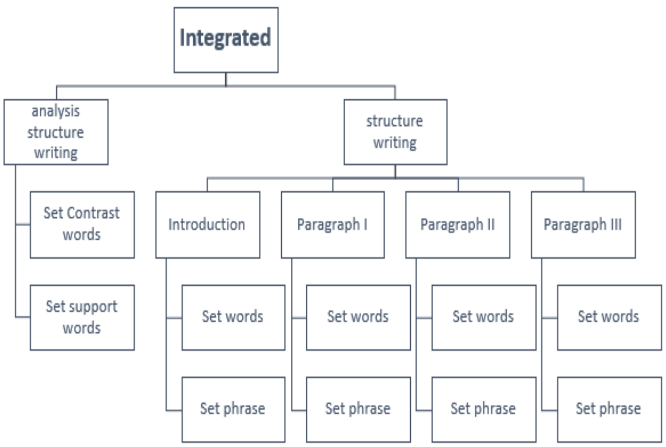


Fig. 2. The dataset structure of integrated part

2. Independent
$S_2 = \{w_1, w_2, \ldots, w_n, p_1, p_2, \ldots, p_n\}$

In which:

- W: Dataset common words in Independent.
- P: Dataset common phrases in Independent

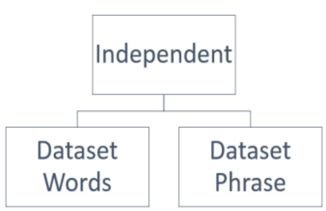


Fig. 3. The dataset structure of independent part

Step 3: Calculation of points for Integrated and Independent writings

Part 1: Comparing with words and phrases in the sample dataset.

Integrated:

Analysing the structure of integrated writing.

Based on the symbol Enter '\n' for recognizing the writing paragraph, we can construe the structure of Introduction, Body Paragraphs I, II, and III.

$$W_1 = \{I_1, P_1, P_2, P_3\}$$

In which:

- I_1: Introduction paragraph
- P_1: Body paragraph I
- P_2: Body paragraph II
- P_3: Body paragraph III

Checking the number that matches words or phrases in Introduction, Body Paragraphs I, II and III with the sample dataset. After that, based on number matching, we calculated the point of Part 1.

Algorithm:

Input:

- Dataset I, P_I, P_{II}, P_{III} and I_1, P_1, P_2, P_3

Output:

- R_1: Points of Part 1 user writing document **(A)**

- *T:* Array right answer

Initialization:

$R_1 \leftarrow 0$; $T \leftarrow 0$

// Introduction

For *i=0* **to length***(I) do*

If *I[i]* **stored in** *I_1* **then**

T = T + 1

// Body I

For *i=0* **to length***(P_I) do*

If *P_I[i]* **stored in** *P_1* **then**

T = T + 1

//Body II

For *i=0* **to length***(P_{II}) do*

If *P_{II}[i]* **stored in** *P_2* **then**

T = T + 1

//Body III

For *i=0* **to length***(P_{III}) do*

If *P_{III}[i]* **stored in** *P_3* **then**

T = T + 1

Return *T*;

Independent :

Checking the number matching of words or phrases in Independent writing.

Algorithm:

Input:

- Dataset w, p
- D: Document Independent

Output:

- R_1: Points of Part 1 user writing document **(A)**
- *T:* Array right answer

Initialization:

$R_1 \leftarrow 0; T \leftarrow 0$

//Word

For *i=0* **to length***(w) do*

If *w[i]* **stored in** *D* **then**

$T = T + 1$

Return *T*;

// Phrase

For *i=0* **to length***(p) do*

If *p[i]* **stored in** *D* **then**

$T = T + 1$

Return *T*;

Following the Rule of this table below for Point of R_1 based on T

T ➔ R_1

T	R_1	T	R_1
>14	5.0	9	2.5
13	4.5	8	2.0
12	4.0	7	1.5
11	3.5	6	1.0
10	3.0	5	0.5

Fig. 4. The table of point levels

Part 2: Comparing the Integrated or the Independent writing with the sample dataset standard document writing of this topic.

We used the comparison based on document classification method [17]. Then we checked some methods classifying documents such as Naive Bayes Text Classification [2], Support Vector Machines [16], and Vector Space Model [9]. After comparing some kinds of the algorithm [15], we saw that the vector calculation is done very quickly as well as very efficiently for the algorithm to optimize the selection of models, allowing for the revenue of the decreased dimensional vector and the visualization of vector space. Also, the vector space model and its variants are still appreciated as in the field of information retrieval. We chose Vector Space Model (VSM) to present the sample documents.

First, we carried out preprocessing which is one of the main components in a typical text classification model [1]. Then we set up the following model to describe the encoding of every document and the creation of a vector for every encoded document [17]:

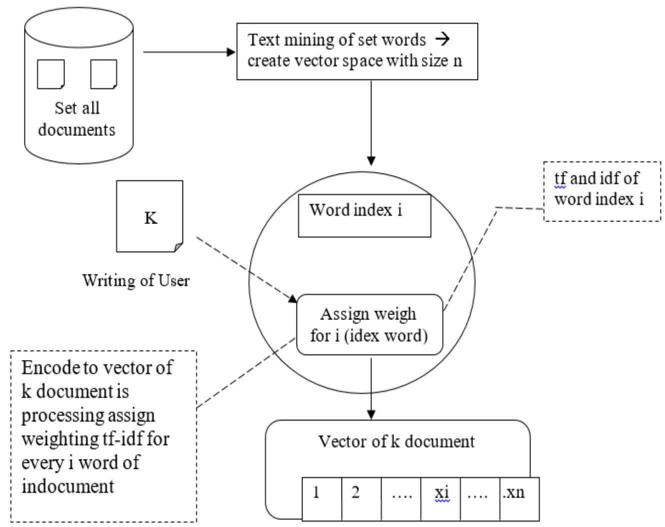


Fig. 5. The model creating vector space

- Creating vector space with size n.

$T = \{D_1, D_2, .., D_n, D_{n+1}\}$

$V = \{V_1, V_2, .., V_n, V_{n+1}\}$

In which:

- T: all documents
- $D_{i \to n}$: Every document in sample data set
- D_{n+1}: User writing document
- V: Vector set of all documents
- $V_{1 \to n}$: Vector of every document in sample dataset
- V_{n+1}: Vector of user writing document
- tf: *term frequency terms weighting*
- idf: *inverse document frequencies*

Algorithm:

Input:

- T: all documents

Output:

- R: Result of distance Vectors
- R_2: Points of Part 2 user writing document

Initialization

$V \leftarrow 0; N \leftarrow 0; R \leftarrow 0$

- N: Set of words for all documents
- S: Vector space

1. ***Creating vector space***

For i=0 **to length** (T)
    Separate words on T_i
    For j = 0 **to length** T_i
        N $\leftarrow T_i[j]$

For i = 0 to **length** (N)
    num = 0;
    For j = 0 to **length** (N)
        If N[i] is **equal** N[j]
            num=num+1
            if num = 3
                S $\leftarrow$ N[i]

2. **Creating vector for every document** [11]

For i=0 **to length** (T)
    Separate words on T_i
    For j = 0 **to length** T_i
        tf = (T_i **words stored in** S) / S
        idf = (T_i **words not stored** in S) / S
        $V_i[j] = tf_{ij} * \log_n(n / df_i)$
Return V;

3. **Comparing 2 vectors**

- Applying the calculation distance Euclide in group Minkowski

$$D_E(x,y) = \sqrt[2]{\sum_{i=1}^{n} | x_i - y_i |^2}$$

// calculating distance
For i = 0 **length to** (V-1)
    R $\leftarrow$ **Distance** (V_{n+1}, V_i)
// **The maximum of the percentage of the similarity of user writing document and the sample dataset document was presented by the minimum value of distance of both vectors.**
Double X = **Min**(R)
    $R_2 \leftarrow$ (1-X)*5(**Points**) **(B)**
    In which:

Double **Distance** (vector V_{n+1}, vector V_i):
    double dis = 0;
    int weigh(n+1), weigh(i);
        For i = 0 **length** to N
            String word = S[i];
                weigh(n+1) = V_{n+1}.searchHash(word);
                weigh(i) = V_i.searchHash(word);
                dis = dis + $|$ weigh(n+1) – weigh(i) $|^2$;
                dis = dis $^{1/2}$;

4. ***Total Points (A) & (B)***

TOTAL RESULT = $(R_1 + R_2) / 2$

IV. RESULTS

As this study was to compare the dataset and learners' responses in the posttest. The responses were the integrated and independent writings. We provided a two-box interface on the screen. The left box contained a reading text (for the integrated) or a topic (for the independent), and the right box was blank for learners to fill in their responses.

We scored both kinds of writings on the language-criterion basis. The language criteria are topic addressing, organization, coherence and language use. We provided the writing topics within a single theme or content area of language, which learners had acquired in the real world or the classroom.

The learners performed the integrated task first. They listened to a short lecture and read a passage from which they had to combine the information to give the responses. They spent five minutes reading the passage provided in the left box, and took notes. After that, they listened to a two-minute lecture, and took notes. Then they used the notes to write their answers in the right box in 30 minutes. After the learners had finished the integrated task, they went on to spend another 30 minutes on the independent writing about a given topic.

The model assessed both the writing tasks and gave the scores on the screen.

The parameters in the model:

- ➢ Input
- Dataset
- Integrated writing.
- Independent writing
- ➢ Output
- Errors
- Scores

The assessment appeared as follows:

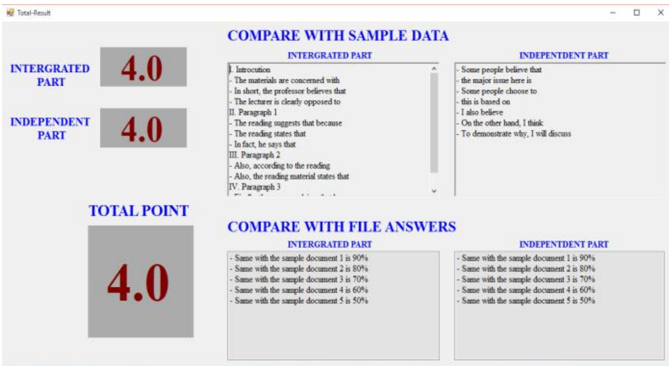


Fig. 6. The scoring model

This scoring method has some advantages over the other scoring methods which aim at the betterment of students' assignments through a continuous, iterative process of writing and revising [5] in that it can help learners to practice writing and get the results through matching words, phrases, and text documents of learners' work and the dataset. Also, this method relieves teachers of the burden of scoring essays which may involve subjective factors.

V. Discussion, Conclusion and Future Work

This performance assessment was for learners at the intermediate level of language proficiency. The design was in accordance with the Raw-to-Scale Score Conversion Tables (Converting Rubric Scores to Scaled Scores) [12] that rate writing performance based on whether it would meet the expectations, exceed the expectations, or not satisfy the expectations for the writing tasks. The performance assessment was valid and reliable according to the university requirements.

The flexible integration of both computer and humans (teacher and student) can increase students' autonomy and raise their awareness of language criteria through students' working with the software independently.

This method is a comprehensive performance assessment. The study contributes to identifying language errors and different kinds of essays to increase the language course outcomes and provide necessary feedback to work out the appropriate methods to improve English language learners' weaknesses. The proposed model can allow users with little knowledge of information technology to access the process of test performance. The software is user-friendly, which is a highly interactive between the software and the user. The analysis in this study ascertains learners' beliefs that they are competent to use computers in their choice of taking writing tests on the computer.

The model is supposed to be an open source so that language instructors can adjust their criteria to be suitable for specific requirements. Future work could use this research as the foundation to improve the implementation of this model in the direction of processing the contextual semantics of the writings for academic English proficiency.

References

[1] A. K. Uysal, and S. Gunal, "The impact of preprocessing on text classification," Information Processing & Management, vol. 50.1, 2014, pp. 104-112.

[2] A. McCallum, and N. Kamal, "A comparison of event models for naive bayes text classification," AAAI-98 workshop on learning for text categorization, vol. 752, 1998.

[3] B.B. Nomass, "The impact of using technology in teaching English as a second language," English Language and Literature Studies,vol. 3.1, 2013, p.111.

[4] C. A. Chapelle and D. Douglas, Assessing language through computer technology, Ernst Klett Sprachen, Cambridge: Cambridge University Press, 2006.

[5] C. E. Chen and W. E. Cheng. "Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes," Language Learning & Technology, vol. 12.2), 2008, pp. 94-112.

[6] D. H. Pham, "A Computer-Based Model for Assessing English Writing Skills for Vietnamese EFL Learners," unpublished.

[7] E. Suzanne, Introduction to systemic functional linguistics, A&C Black, 2004.

[8] Educational Testing Service, TOEFL iBT Scores, 2005. Retrieved from http://www.hhl.de/fileadmin/texte/_relaunch/Conversion_Table_TOEFL_(PBT,CBT,iBT).pdf

[9] G. Kanaan and A. Jafar, "A comprehensive comparative study using vector space model with k-nearest neighbor on text categorization data," Asian Journal of Information Management, vol. 2.1, 2008, pp.14-22.

[10] J. Choi, "The impact of automated essay scoring (AES) for improving English language learner's essay writing (Doctoral dissertation, University of Virginia)," 2010, retrieved from http://www.researchgate.net/profile/Jaeho_Choi6/

[11] L. P. Jing , H. K. Huang, and H. B. Shi, "Improved feature selection approach TFIDF in text mining," Machine Learning and Cybernetics, 2002, Proceedings, 2002 International Conference on vol. 2, IEEE, 2002, pp. 944-946.

[12] M. Brooks, S. Amershi,B. Lee, S.M. Drucker, A. Kapoor, and P. Simard, "FeatureInsight: Visual support for error-driven feature ideation in text classification," in Visual Analytics Science and Technology (VAST), 2015 IEEE Conference, pp. 105-112, IEEE.

[13] R. Hahn, A1 TOEFL writing (iBT) (Korean edition), 2008.

[14] Raw-to-Scale Score Conversion Tables by Educational Testing Service, TOEFL iBT Scores, 2005. Retrieved from http://www.etweb.fju.edu.tw/elite/ETS%20%20ibt%20TOEFL%20Converting_Rubric.pdf

[15] S.H.I Yong-feng and Z. Yan-ping, "Comparison of text categorization algorithms," Wuhan university Journal of natural sciences, vol. 9.5, 2004, pp.798-804.

[16] T. Joachims, Text categorization with support vector machines: Learning with many relevant features. Springer Berlin Heidelberg, 1998, pp. 137-142.

[17] V. Korde and C. N. Mahender, "Text classification and classifiers: A survey," International Journal of Artificial Intelligence & Applications, vol. 3.2, 2012, p.85.

[18] Y. Attali and J. Burstein, "Automated essay scoring with e-rater V. 2.0 (ETS RR-04-45),". Educational Testing Service, Princeton, NJ, 2005.

Parallel and Distributed Genetic Algorithm with Multiple-Objectives to Improve and Develop of Evolutionary Algorithm

Khalil Ibrahim Mohammad Abuzanouneh
Al Qassim University, College of Computer, IT Department
Al Qassim, Saudi Arabia

***Abstract*—In this paper, we argue that the timetabling problem reflects the problem of scheduling university courses, So you must specify the range of time periods and a group of instructors for a range of lectures to check a set of constraints and reduce the cost of other constraints ,this is the problem called NP-hard, it is a class of problems that are informally, it's mean that necessary operations to solve the problem will increases exponentially and directly proportional to the size of the problem, The construction of timetable is most complicated problem that was facing many universities, and increased by size of the university data and overlapping disciplines between colleges, and when a traditional algorithm (EA) is unable to provide satisfactory results, a distributed EA (dEA), which deploys the population on distributed systems ,it also offers an opportunity to solve extremely high dimensional problems through distributed coevolution using a divide-and-conquer mechanism, Further, the distributed environment allows a dEA to maintain population diversity, thereby avoiding local optima and also facilitating multi-objective search, by employing different distributed models to parallelize the processing of EAs, we designed a genetic algorithm suitable for Universities environment and the constraints facing it when building timetable for lectures.**

Keywords—Heterogeneous clusters; NP-hard; evolutionary multi-objective algorithm; parallel algorithms; Real-time scheduling

I. INTRODUCTION

Genetic Algorithm a heuristic used to find a vector x * of free parameters with associated values in an admissible region for which an arbitrary quality criterion is optimized as in given in figure 1.

$$f(\vec{x}) \to \max : \text{find } \vec{x}^* \text{ so that } \quad \forall \vec{x} \in M : f(\vec{x}) \leq f(\vec{x}^*) = f^*$$

Fig.1. A sequential Genetic Algorithm

The algorithm uses stochastic operator's selection, crossover and mutation on an initially random population in order to compute a whole generation of new strings.

Evolutionary Algorithm is part of the science of artificial intelligence so that it is similar to simulation solution including genetic evolution like the system through the representation of some genetic processes such as natural selection and the struggle for survival and mutations and live in groups [1].

Stages of the evolutionary algorithm begins by choosing a group of chromosomes so that each representing an individual solution to a problem, then you are causing a mutation on individuals process to produce new offspring , then a selection nearest the members of the ideal solution and the neglect of the solutions, we have to repeat this process for the production of successive generations, each with a larger number of qualities required prevailing until the arrival of the algorithm to termination as in given in figure 2.

```
BEGIN
 INITIALISE population with random candidate solutions;
 EVALUATE each candidate;
 REPEAT UNTIL (TERMINATION CONDITION is satisfied) DO
  1) SELECT parents;
  2) RECOMBINE pairs of parents;
  3) MUTATE the resulting offspring;
  4) EVALUATE new candidates;
  5) SELECT individuals based on their fitness for the next generation;
 DO
END
```

Fig.2.The General Scheme of an Evolutionary Algorithm

II. BACKGROUND

A. Parallel Genetic Algorithms (PGA): is important for improvements using parallel models of Genetic Algorithms, Parallel Genetic Algorithms (PGA) are faster to find sub-optimal solutions, and able of cooperating with other search techniques in parallel [10], PGA is independent of the problem and can yield alternative solutions to the problem, parallel search from multiple points, easy parallelization, better search, higher efficiency and efficacy than sequential GAs [1].

B. Parallel Multi-objective Evolutionary Algorithm

Steps of the Parallel Multi-objective EA:

1) Create a community, the creation of the initial community, according to the problem to be solved and is usually within the terms of the initial problem and the available data [5].

2) Individual's evaluation, we use to assess individual fitness function for each by calculating the optimal solution function, this value is used to determine the closest solutions to solve optimization problems because they need to modify less than others to reach the optimal solution.

3) Mutation, which occurs randomly on one of the characteristics of the individual. So that the value of this status change to a random value within the terms of individuals [9].

4) The overall structure of the proposed methodology is shown in the pseudocode Multi Clustered Parallel GA [14].

5) Divide the individuals into n clusters based on the fitness value.

6) For each cluster perform the following:

a) Using Selection Mechanism, Select the individuals from each node.

b) Convert the individual to Gray code.

c) Mutate the Parent.

d) Convert the offspring to Binary Value, calculate the fitness Value.

7) Group the clusters together.

8) Allow the migration of individuals based on fitness.

9) Until Termination condition is reached repeat from point 5.

10)Select the best Individual.

Multi-core cluster: is a cluster where all the nodes in the cluster have multi-core processors and multi-cluster architecture is a multiple cluster system that is connected via the cluster interconnection networks [4], [11].

C. Multi-objective algorithm

Most of the problems that we face in life are not easy to define it as one goal, but must all parameters within the framework of the problem to get a better definition of the problem, and this in turn facilitates access appropriate solution for all parties. From here began the pluralistic definition of goals and how to reach a solution that suits conflicting objectives to define the problem.

In our problem we had two goals must be achieved and taken into consideration when finding a suitable solution for hard constraints and soft constraints, during the search for a solution to implement these objectives to get a satisfactory solution for both instructor and student [12],[13].

D. Parallel algorithm

One of the most important issues that must be considered when dealing with an evolutionary algorithm redistribution of tasks, so that the tasks are evenly distributed to all devices to accomplish all the tasks at the same time, the importance of the redistribution of loads in process Heterogeneous environments specifications [4], each computer has its own specifications which appears in varying speeds completion operations [1], [3],to implement parallel software engineering apply the four processes are: partitioning, communications, aggregation, and planning [2].

The division of tasks in two ways, either partial operation is divided into tasks or data division. In both cases, the system needs to contact operations between devices for coordination and exchange of data and information between processors, and communication processes that have several classifications of them (local or global, or structured or unstructured, or a fixed or non-fixed, or synchronous or asynchronous),to reduce processing and communication operations, and then assemble a single tasks in groups, so be quick communication and less costly, and to increase the usefulness of this process (assembly process) ,we develop processes to execute parallel operations on different processors and all operations that frequently communicate with each other on the same processor, it must be a balance between these two points to improve the performance of the algorithm as possible. This process is called planning and is divided into three levels of complexity according to change the number of fragmented or stability tasks, and the division is organized or unorganized, or the communication process during implementation it will change or not [2]. There is a lot of evidence of the high efficacy and efficiency of PGAs over traditional sequential GAs (for example [15].

The best topology for a parallel dGA [16]. , the ring and hypercube are two of the best topologies. The ring topology is easy to implement on very different hardware platforms (cluster of workstations, multi-computers), and its use is very extended.

These results only mean that the dGA is an efficient and robust search procedure. The actual performance of this algorithm is often better than the sequential GA in many problems; see for example [17]. This research maintains samples of very different zones of the search space in every population, thus showing a higher efficacy, and probability of obtaining a solution, for complex applications, the parallel dGA is important in order to have lower computational times and using larger populations than with sequential GAs, and the high number of non-standard and machine-dependent PGAs has led to efficient algorithms in many domains, these call distributed and cellular GAs (dGA and cGA),we refer to their computation/communication ratio, while the actual differences can be also found in the way in which they both structure their populations (Figure 3), when a distributed GA has sub-populations a cGA has typically one single string in every sub-algorithm. In a dGA, the sub-algorithms are loosely connected, while for a cGA they are tightly coupled, and in a dGA there exist only a few sub-algorithms while in a cGA there is a large number of them.

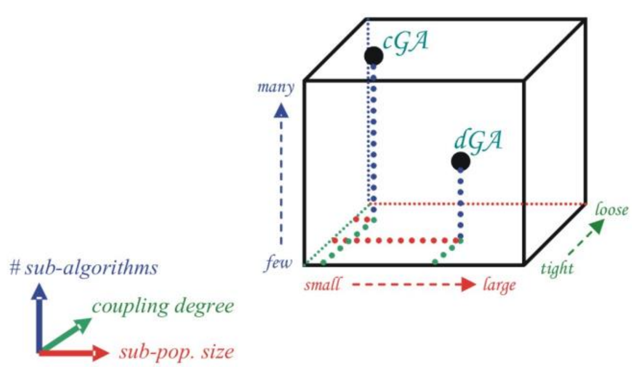


Fig.3. The Structured-Population Genetic Algorithm Cube

III. Related Work

The efficiency of schedule of courses is important and necessary for the satisfaction of instructors and students, as well as it is necessary for exploitation of human and material resources better. In the past, the solution to the scheduling problem was manually using trial and error and had access to resolve ideally devoid of errors is very difficult, even if any of this solution, it is not the best, for this is the use of scientific methods to solve the problem. The problem Study Began four decades ago, but not completed for a specific mechanism to resolve the problem until now. The study of the problem by (Gottlieb), the description of the problem that each lecture is one of a group of students and one teacher and a number of times that are randomly selected [6]. The problem of scheduling appointments in schools is easier than in the universities because the halls System (Class) is not changed ,but the students per group are variable and are chosen randomly for each course at the university. This increases in the complexity of the problem at the University greater than the complexity of halls system, so it is a traditional way to solve the problem is difficult, particularly if there are several objectives to be achieved at the same time [6]. Until the problem is defined as a Multi-objective, there must be more than one target fixes the problem and different goals from one university to another definition. It is noteworthy that most researchers had to solve the problem as a problem of a single goal , where the only goal is to reduce all constraints (Abramson and Abela 1992;Blum; 2002; Lima; 2001; Piola; 1994), but there are many options to define goals such as access to a timetable orderly and correctly, and a number of successive lectures to the instructor and others that make the problem of scheduling multiple objective problems, there are a few researchers who define the problem as a problem of multiple targets. Carrasc and Pat (2001) used the dual objective model to the problem of scheduling courses in school schedules to reduce soft constraints for the teacher and the class, Filho and Lorena (2001) they also define the problem as a dual goal, Dicev (2004) identified the problem as a dual goal, but the definition of different goals to reduce the gap between lecture time and the date of the completion of daily lectures, so that priority for early lectures given in one of the goals, and priority late for lectures given at the other goal [7].

In one of the other research, the researchers using the server and client architecture in the distributed evolutionary algorithm, so that there was a customer who apply evolutionary algorithm and named (O-clients) and others to resolve the problem by relying on the next transactions (O-clients) and those customers so-called (GEP-clients)[8]. In another research has the comparison between the two techniques for distribution, which is the first (Message Passing Programming) and the other (Shared Memory), so are easy to apply while the latter characterized the first efficient top except in special cases. This was after the application of the first three experiments on the sequential evolutionary algorithm, and the other on an evolutionary algorithm parallel technology (Message Passing Programming), and the third parallel evolutionary algorithm technology (Shared Memory) [9].

IV. Methodology

Description of the problem: we have implemented the algorithm on the data in the computer college at Qassim University. As an example of the experiments and development of the algorithm will be applied of all Qassim University Colleges. In the College of Computer, there are three disciplines: Computer Engineering, Computer Science and Information Technology. At the beginning of the semester, we have four student groups for each specialty, so it becomes a total 3 * 4 = 12 student groups. Where lectures system will apply on Sunday, Tuesday, and Thursday we have eight lectures of 60 minutes per lecture, but in the days Monday and Wednesday, there are five lectures each day by 90 minutes. To resolve the problem correctly, we define the problem to become composed of 6 different groups are as follows: groups of students, teachers, courses, halls, the time periods for lectures and restrictions.

The problem then being formulated to become as follows:

{St, Inst, Co, R, L, O}, where St = {st1, st2, st3 ..., sti} contain groups of students, each item which is stored the following values: (st_m, year , std_slot) where st_m: specialization, and year: the academic year, and st_slot: lectures time.

Ins= {inst1, inst2, inst3 ..., inst j} Instructors name, and every item in this collection contains the following information: (inst_no,inst_slot) where inst_no: instructors number, inst_slot: lectures times.

Co = {co1, co2, co3 ..., con} : reflect the, every element in this group stores the following values:

(sti, co_no, section, cap, inst_no, h_no, ph, ph_type, ph_inst), where the sti: student group number, and co_no: course number, section, cap : capacity for this course, inst_no: instructor number, h_th: the number of theoretical hours , h_lab: number of lab hours for this course, and h_type: the type of hours for this course, so the value of 0 if it is not for this course lab hours o and be 1 if the course has lab, 2 for multimedia lab, 3 if there electrical lab , 4 for electronics lab, either for h_inst lab they reflect Instructors hours number in the laboratory, shall be 0 if it is not necessary to the existence of the instructor in the laboratory, and is 1 if the instructor must stay a half lab time, and be 2 if the teacher must stay all time in the lab.

Ro = {ro1, ro2, ro3 ..., rok} expresses the halls and every item the stores by the following information: (cap, vcr, type_r) where that cap: reflect the capacity of the room, and VCR: expresses whether the room has a device for viewing or not shall be the value 0 if the room has a device for viewing and 1 if the room did not contain a device for viewing, but type_Ro: It reflects the room type is 0 if the lecture hall and 1 If a computer lab.

L = {lt1, lt2, lt3 ..., ltn} reflect the time periods in which every item contains the following data: (tn, dn) where tn: number The time period ranges from 1-8 days on Sunday, Tuesday, and Thursday , while its value ranging from 1 to 5 in the days Monday and Wednesday, and dn: thus ,expressed number of today ranges from 1-5.

Finally Cst = {cst1, cst2, cst3 ..., cst} a group which determinants that must be taken into account when searching for a solution, and Cst group: express the exact weight and value. Determinants are divided into two types difficult determinants (Hard), and determinants (Soft). The value should be easy or difficult determinants according to their importance.

G = {g1, g2, g3 ..., gj} represents a group of genes, which expresses the solution to the problem (chromosome), also shown in Figure data contained in the gene.

Each gene is a vector contains the following data (course, room_no, ph_tslot, lab_no, th_tslot),

Where course: The course number, but room_no: No. Hall, which will be given this course, and ph_tslot: expresses the number of time that will give the course, and lab_no: expresses the number laboratory or operator, and finally the th_tslot: expresses the time period in which will be given by the laboratory.

In this issue, we have a large and complex field of research, where it is to identify specific hall and laboratory hours if necessary process for each course suggested during the season, is also determined by the time period in which the course will be given and the laboratory, so that they are suitable for both the teacher and the student. To reduce the time and effort required to solve, our way to random selection for each of the halls and time periods to take into account the absence of a conflict with a teacher, and then later to improve the solution to meet the hard and soft student determinants (hard and soft constraints) easy for the teacher and determinants.

Constraints: the final solution takes into account the set of parameters that must be implemented so that the solution would be acceptable to both the student and the instructor, and these constraints are divided into hard constraints (infinity) soft constraints be valued according to their importance.

For hard Constraints are as follows:

1) You must not have an instructor lectures to the same time.

2) You must not be a student at the lectures in the same time.

3) You must not have an instructor lecture time at the time that does not exist in the university.

4) Lectures must not be in one hall at the same time and it accommodates the number of students.

For the soft constraints are as follows:

1) The students don't have lectures successive without intervals of rest.

2) Instructors or students do not have free time exceed three-hour.

3) Instructors don’t have a lecture late every day.

4) Students don’t have any consecutive lectures in places away.

V. Distributed Processing

In this section, we use distributed processing with the evolutionary algorithm for two main reasons:

1) To reduce the execution time by distributing operations on multiple processors

2) And the second by taking advantage of the division of society into several partial communities, which may help to give the best results [3], [4].

The researchers developed a set of definitions and methods relating to the process of distribution, and we'll show here what has been used in this project from these definitions and methods:

a) Acceleration ratio is the ratio of the time required to implement the algorithm on a single processor for the implementation of the algorithm on the processor.

b) The efficiency of the system: is the ratio of the acceleration for the number of processors that are divided by the algorithm.

c) Additional operations in the search process come from communication processes, and the waiting time for processing, for sharing data and processing.

There are several classifications of the process of distribution of any algorithm, including how to distribute the algorithm in terms of distribution operations or data which we used, called (Single Instruction, Multiple Data), which means that all processors will execute the same instructions, but on different data problem of this kind. The distribution of the process, he needs to redistribute loads after all the setups, especially if the application on a non-homogeneous system (heterogeneous environment) [4].

There are two methods that can be used in the distribution process : shared storage space and message passing computer, where the message passing computer used in this project, which means that each processor has its main memory, and can communicate with each other processors by sending.

The tasks divisions (partial division of society communities) at the start of the implementation of the algorithm in the central device, regardless of the needs of other devices, while not sent these partial societies only when a request for this data from their processors as I mentioned earlier, the redistribution of loads is important to us in this project to improve waiting time process, and this process is different from time to time because the evolutionary algorithm's dependence on random variables prevent the expected implementation the end of time, so the redistribution of loads is a dynamic process, and in the Central process device, we used algorithm centralized dynamic load balancing algorithm.

VI. Cluster Computing

This section provides the technique of linking set of computers in LAN in order to take advantage of the parallel processing power of those computers, clusters are designed to exploit the parallel processing power of multiple nodes and clusters operate by routing all work through one or more load-balancing front-end nodes, which then distribute the workload efficiently between the remaining active nodes and processing power can be optimized, the time required to solve the problem on N processors with the time required on a single processor. This is shown as:

$$S(n) = T(1) / T(n); \tag{1}$$

where S (n) is the speedup achieved with n processors, T (1) is the time required on a single processor, and T (n) is the time required for N processors.

The concept of efficiency is defined as

$$E(n) = S(n) / n. \quad (2)$$

It measures how much speedup is brought per additional processor and the task is to be computed on a single processor, the time needed can be represented as:

$$T(1) = s' + np', \quad (3)$$

The scaled speedup can be written as:

$$s'(n) = \frac{T(1)}{T(n)} = \frac{(s' + np')}{(s' + p')} = n - (n-1).\frac{s'}{s' + p'}$$
$$= n - (n-1).s'' \quad (4)$$

where s'' is defined as $s'/(s'+p')$. s'' is the ratio of serial code, execution, while s is with reference to all code in the whole program for the problem, and also be noted that s is a constant that is only relevant to the computation problem, under the precondition that problem scale is fixed; where s'' is a constant under the precondition of problem scale changes as Gustafson described. Under Gustafson's Law, the speedup can be linearly increased with the number of processors hired in the computation [18].

VII. Parallel and Distributed EA

Parallel and distributed EA computing are a key technology in the present days of networked and high-performance systems. And for increased performance e by adding processors, memory and an interconnection network and putting them to work together for sharing the workload, it is hoped that an N-processor system will give rise to a speedup in the computation time, the MIND class of parallel architectures multiple processors work together through some form of interconnection, the different programs and data can be loaded into different processors which mean that each processor can execute different instructions at any given point in time, The processors will require some form of synchronization and communication in order to cooperate on a given application. 4 Figure 4 is the most commercially and useful of the parallel and distributed architectures that belong to it.

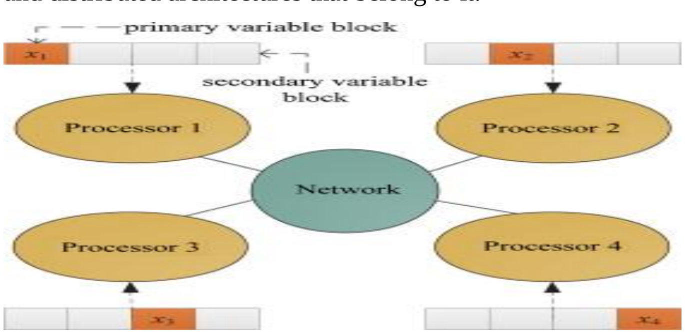


Fig. 4. Distributed GAs and homogeneous Implementations Used MIMD

The connectivity of the islands in a parallel distributed GA. We define the migration policy as a tuple of four values:

$$M = (m, \zeta, \omega_S, \omega_R) \quad (5)$$

Where m is the migration rate, ζ is the frequency of migration, $\zeta \in \{0, 1, ..., \infty\}$, being $0 \approx \infty$ (partitioned dGA),where ω_S is the *policy for selecting migrants*, we send a copy of the selected individual to the neighboring island; alternatively, the individual itself could be sent. It defines the connected nodes (topology) and the shared individuals, where ω_R is the *migration replacement policy*, used for integrating an incoming individual in the target island.

Some work is available at the convenience of using asynchronous communications [2], [12], [24]. This can be achieved by inserting an individual whenever it arrives, thus avoiding blocking every ζ steps, i.e., Emissions and receptions of migrants are managed in separate portions of the code.

The set of parameters Θ_M controlling the migration operator ω_M consists in the migration policy 0, plus a synchronization parameter (sync/async, labeled as **s** and **a** in the forthcoming graphs). The reproductive cycle of a parallel distributed GA is a composition of the island reproductive cycle and a migration operator:

$$\omega_d = \omega_M \circ \omega_{island} \quad (6)$$

To compute super-linear speedups we need to reduce both the number of necessary steps and the expected execution time Tnproc in relation with the sequential one T1, (Equation 7), [19].

$$S(n_{nproc}) = \frac{T(1)}{T(n_{nproc})} \quad (7)$$

There exists an exponential relationship between the speedup and the number of processors. In this relationship (Equation 8), s is the acceleration factor (super-linear speedup when s>1).

$$S(n_{nproc}) = n_{nproc.}\ s^{nproc-1} \quad (8)$$

VIII. Scheduling System

We have implemented in a cluster of 1 workstations and 5 processors, these machines have a technique employed to achieve parallelism has a number of processors that function asynchronously and independently, at any time, also different processors for executing different instructions on different data, implementations of homogeneous ,we used Java programming to solve the problem by using the principle of programming and the definition of the entities of the problem classes: teachers, students, and courses, and the halls, finally the genes which represents the solution.

A. The Results of Experiments

After many experiments and observations on the environment, Figure 5 shows the evolution of the average best fitness and the number of generations. Relationship between sequential execution and different variables Teachers=50, Sections=90, Classes=25.Generation=1000, Best Fitness=0.98.

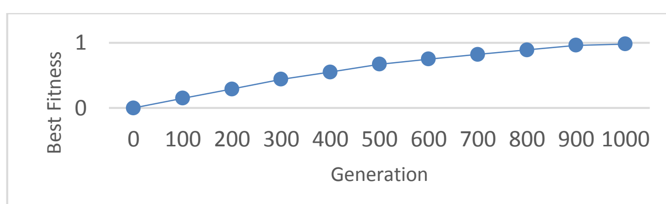


Fig.5.The Best Fitness Evolution – Sequential GAs

Figure 6 describes the results when we run code to test the Superliner speedup in a shared memory system we used the data to access a specific element in the memory. The sequential execution will have a cache miss time for next element is read. Average running time = 4.52274 in a second.

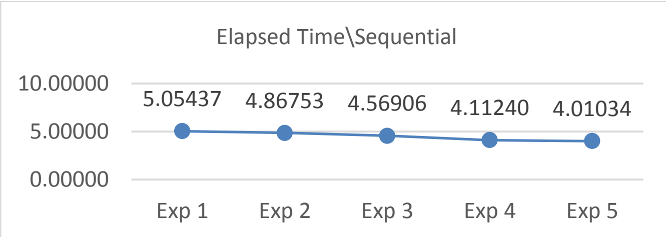


Fig.6.The Sequential System: S*ISD*

Figure 7 Shows the relationship between the number of different variables and the time required for the execution of the algorithm, note that the increased rate of time implementation and repetition stage with increasing variables of sections=100, Generation=1000 , Best Fitness=0.91.

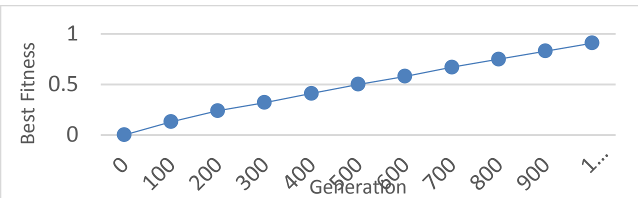


Fig.7.The best fitness evolution with increasing variables

Figure 8 shows the relationship between processor number, execution time, generation number and fitness in diagrams, and the analysis of the results. The execution time of the parallel multi-objective algorithm decreases efficiently as the number of processors increases. Processors=5 Fitness=0.91.

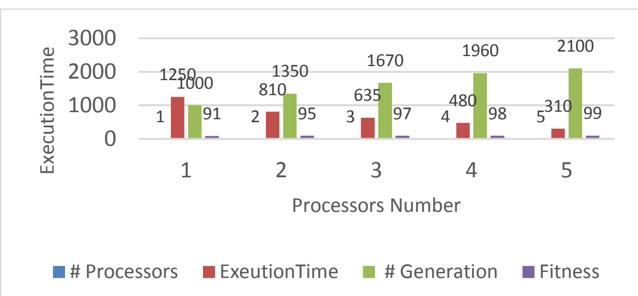


Fig.8.The Relationship between Execution Time and the Number of Processors- Parallel GAs

Figures (9, 10) Refer to the relationship between the number of generations, the hard cost, and soft cost respectively. We note through decreasing the cost of the hard and soft determinants with repeat stages of implementation, which means that it has been achieved two objectives at the same time.

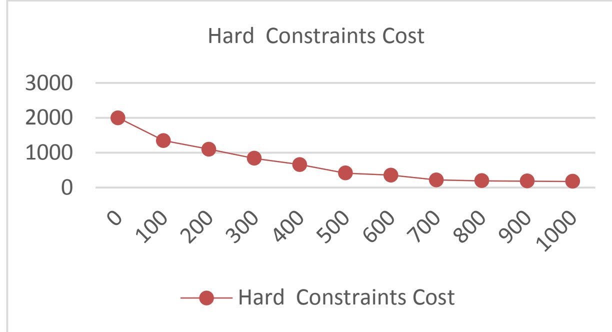


Fig.9. The Hard Cost

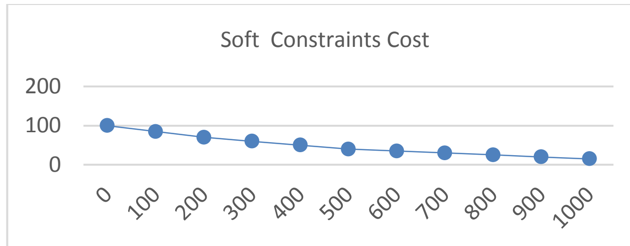


Fig.10.The Soft Cost

Figure 11 describes the result when we run code to test the Superliner speedup in a shared memory system in parallel execution, each element just fits in the level-1 cache memory to find the data it needs for its next operation, it will save time compared get it from random access memory, so there is no cache miss.

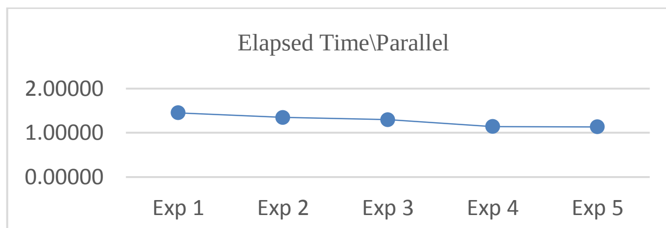


Fig.11. Parallel System: SIM*D*

TABLE I. SPEEDUP (AVG. SEQUENTIAL \ AVG. PARALLEL)

in second	Avg.Sequential	Avg.Parallel	Speedup
Elapsed Time	4.52274	1.27381	3.55055

Figure 12 describes the result when we run code with five slaves in one machine, we tested separately and the average running time of these five experiments. Elapsed time for each experiment is shown in Figure 12. Average running Time =2.80957s.

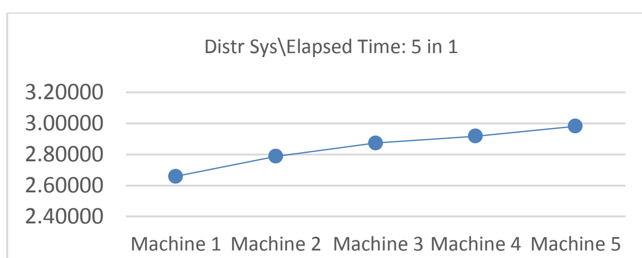


Fig.12. Distributed System: *MISD ((in second))*

Figure 13 describes the result with five slaves running on five different machines, we did five experiments with the same code and the same machines. Elapsed time for all experiments is shown in Figure 14. Average running time = 1.52401s

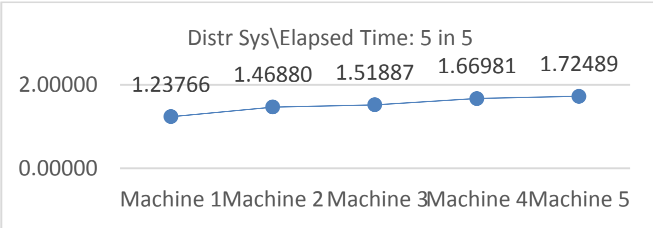


Fig.13. Distributed System: *MIMD (Five slaves in Fives machines)*

In table 2 describes the Ethernet latency slows down the application program. After comparing the results from table1 and table2, our conclusion we can gain Superliner speedup in distributed model to reduce access time, we need to use a fast Ethernet and store the data in a cache memory with each machine running with distributed system.

TABLE II. SPEEDUP IN DISTRIBUTED SYSTEM

Speedup in Distributed Model	
Five slaves in one machine	2.80957
Five slaves in Five machines	1.52401
Speedup	2.16679

IX. CONCLUSIONS AND RECOMMENDATIONS

A. Conclusions: In this paper, we have improved the performance of an evolutionary algorithm after research and tests, we concluded the following: 1. the application of multipurpose and techniques of the parallel and distributed algorithm gave the best performance in solving the problem. 2. After the application of the methodology presented with real data in all experiments of Computer College at Qassim University in all its complexity and constraints where the results were satisfactory. 3. The time required for the processes of communication between processors in parallel and distributed programming case, it takes less time for processing in the same conditions with one processor. 4. The process of redistribution of loads increased complexity as a result of the adoption of the original algorithm on the random variables, which makes it difficult to determine the time required for the implementation of the algorithm. 5. The main idea has been in parallel distributed GAs, so the impact of the research in this kind of algorithms is a larger than others of Parallel GAs.

B. *Recommendations*: After my research has been completed, I recommend the following: 1- Improve the performance Used a new architecture for multi-core multi-cluster [14].

2. Use algorithms to solve many problems in other ways and techniques 3. Applying algorithms in medical fields through the definition, diagnosis and analysis of complex diseases and finding the optimal treatment. 4. Use and application of these algorithms in industrial areas to get the best solutions and results after the definition of the project and identify targets. 5. The distance and locations between the halls we can use it as additional conditions and restrictions on the algorithm to enhance our solutions. 6. We can improve the application of the algorithm through the redistribution of tasks to get the final solutions. 8. We can improve the project used DBMS, which leads creating tables, relationships and reduce the execution time, and size of the data.

REFERENCES

[1] Entropic and Real-Time Analysis of the Search with Panmictic, Structured, and Parallel Distributed Genetic Algorithms. Enrique Alba, Carlos Cotta, Jose M Troya. LCC Technical Report ITI 99-7, 1999.

[2] A Parallel Implementation of Genetic Programming That Achieved Super-Linear Performance. David Andre, John R Koza, 1997.

[3] Designing Efficient and Accurate Parallel Genetic Algorithms. Erick Cantu-Paz. IlliGAL Report 99017, 1999.

[4] M. Aldasht, J. Ortega, and C. Puntonet, "Dynamic load Balancing in Heterogeneous Clusters Exploitation of the Processing Power", 2007.

[5] D. Datta, K. Deb, and CM Fonseca, "Solving class timetabling problem of IIT Kanpur use multi-objective evolutionary algorithm ", 2006.

[6] P. Pongcharoen, W. Promtetn, P. Yenradee, and C. Hicks, "Stochastic optimization timetabling tool for university course scheduling", 2007.

[7] H. Park, A. Grings, M. Santos, and A. Soares, "Parallel hybrid evolutionary computation: Automatic tuning of parameters for parallel gene expression programming", 2008.

[8] Hamid, Norhazlina, Walters, Robert John and Wills,Gary Brian (2014) Performance evaluation of multi-core multi- cluster architecture. In, Emerging Software as a Service and Analytics, Barcelona, ES,03-05 Apr2014. pp, 46-5.

[9] de Toro Negro, F., Ortega, J., Ros, E., Mota, S., Paechter, B., Martin, J.: PSFGA: Parallel processing ,and evolutionary computation for multi-objective optimization . Parallel Computing 30, 721–739 (2004).

[10] Miki, M., Hiroyasu, T., Watanabe, S.: The new model of the parallel genetic algorithm in multiobjective genetic algorithms. In: Congress on Evolutionary Computation CEC 2000, vol. 1, pp, 333–340 (2000).

[11] Alvaro Garcia-Piquer, Andreu Sancho-Asensio, Albert Fornells, Elisabet Golobardes, Guiomar Corral, Francesc Teixidó-Navarro. (2015) Toward high-performance solution retrieval in multiobjective clustering. Information Sciences 32012-25. 2015.

[12] Dr. P.M.G. Moreira and Dr. Paulo J. Tava, Kim C. Long, William S Duff, John W Labadie, Mitchell J Stansloski, Walajabad S Sampath, Edwin K.P. Chong. (2015) Multi-objective fatigue life optimization using Tabu Genetic Algorithms. International Journal of Structural Integrity 6677-688.7-Dec-2015.

[13] Lothar Thiele, Kaisa Miettinen, Pekka J. Korhonen Julian Molina, A Preference-Based Evolutionary Algorithm for Multi-Objective Optimization No Access Evolutionary Computation Fall 2009, Vol. 17, No. 3, Page 411-436.

[14] Hamid, Norhazlina, Walters, Robert John and Wills, Gary Brian, " An Architecture for Measuring Network Performance in Multi-Core Multi-Cluster Architecture (MCMCA)," International Journal of Computer Theory and Engineering vol. 7, no. 1, pp. 57-61, February 2015.

[15] J. L. Ribeiro Filho, C. Alippi, P. Treleaven. "Genetic Algorithm Programming Environments". Parallel Genetic Algorithms: Theory & Applications, J. Stender (ed.), IOS Press. 1993.

[16] S. Lin, W. F. Punch and E. D. Goodman. "Coarse-Grain Parallel Genetic Algorithms: Categorization and New Approach". Parallel & Distributed Processing. October 1994.

[17] S. Lin, W. F. Punch and E. D. Goodman. "Coarse-Grain Parallel Genetic Algorithms: Categorization and New Approach". Parallel & Distributed Processing.Oct1994.

[18] Yuan Shi. Reevaluating Amdahl's law and Gustafson's law.Available:http://joda.cis.temple.edu/~shi/docs/Amdahl/amdahl.html.

[19] T. C. Belding, "The Distributed Genetic Algorithm Revisited". In L. J. Eshelman. Proceedings of the Sixth ICGA. Morgan Kaufmann, CA, pp. 114-121, 1995.

Gender Prediction for Expert Finding Task

Daler Ali
Dept. of Computer Science and IT
The Islamia Univ. of Bahawalpur, Pakistan

Nadeem Akhtar
Dept. of Computer Science and IT
The Islamia Univ. of Bahawalpur, Pakistan

Hina Asmat
Dept. of Computer Science
Govt. S.E. College, Bahawalpur

Malik Muhammad Saad Missen
Dept. of Computer Science and IT
The Islamia Univ. of Bahawalpur, Pakistan

Nadeem Salamat
Dept. of Basic Science and Humanities
Khawaja Farid Univ. of Eng. And Tech. RYK

Amnah Firdous
Dept. of Computer Science
COMSATS Institute of Information Technology

***Abstract*—Predicting gender by names is one of the most interesting problems in the domain of Information Retrieval and expert finding task. In this research paper, we propose a machine learning approach for gender prediction task. We propose a new feature, that is, *combination of letters* in names which gives 86.54% accuracy. Our data collection consists of 3000 Urdu language names written using English Alphabets. This technique can be used to extract names from email addresses and hence is also valid for emails. To the best of our knowledge, it is the first-ever attempt for predicting gender from Pakistani (Urdu) names written using English alphabets.**

Keywords—Urdu; Semantic Web; Gender Prediction; Expert Profiling; Machine Learning

I. Introduction

As internet becomes an intrinsic part of our lives, organizations tend to focus on automated solutions that can exploit the information available on the web. With the volume of increasing information on the web, the motivation for generating increased mass of knowledge is also increasing. However, it is very obvious that if technology is meant to bring benefits, it has to be able to support not only access to documented knowledge but also, most importantly, knowledge held by individuals [1]. To find and process such knowledge, expert finding task has been proposed by Information Retrieval research community.

The objective of an expert finding system is to help find people with the appropriate expertise through some intelligent automated techniques [2]. This task is very challenging because of rich set of information needs related to it. For example, finding the experts with particular set of skills within a particular domain or finding an expert from a specific geographical location. One of the most interesting and challenging tasks associated with expert finding task is gender prediction through expert's names. When searching for experts with the data available on the web, finding experts with a particular gender could be a very pertinent information need.

Gender prediction through names (or emails) is not only important for expert finding task only but also for many tasks like Co-reference Resolution, Machine Translation, Textual Entailment, Question Answering, Contextual Advertising and Information Extraction [3]. In literature, most of the work regarding gender prediction can be associated with author profiling tasks [4, 5] or gender prediction using names [3, 6] for expert finding tasks.

In this paper, we propose a machine learning approach to predict gender when written using English alphabets as these are mostly found on the web. We propose a feature (named as *combination of letters*) which gives better results when combined with existing proposed features (proposed for other languages). To the best of our knowledge, there is no previous work on this problem.

II. Related Work

Generally four types of work can be found when talking about gender prediction i.e.

- gender prediction using text,
- gender prediction using names,
- gender prediction using images,
- gender prediction using voice.

Gender predictions using images [7, 8] and voice [9, 10] are beyond scope of our work so we only discuss first two categories of work in this section.

A. Gender Prediction Using Text

Gender prediction using text is a sub-task of author profiling task. Author profiling, in general, is used to determine an author's gender, age, native language, personality type, etc [11]. It is a problem of growing importance in a variety of areas, including forensics, security and marketing. This is why it was also introduced as part of PAN (CLEF)[1] in year 2013 and continues to year 2016 as one of its core tasks. Gender prediction from text has been performed in several forms like blogs [12], electronic discourse [13], online social networks [14], and email [15].

[1] http://pan.webis.de/clef13/pan13-web/author-profiling.html

Researchers have been using style-based features (N-grams of POS tags in documents, punctuation symbols and number of *href* links [16, 17] etc.) as well as topic-based features for gender prediction from text (for example, males usually use words like 'daily life' to describe their work and whereas females use 'daily life' to describe their love or spiritual life).

B. Gender Prediction from names

Gender prediction from names is a challenging task; hence, one cannot find lot of work already done for this particular task. One of the foremost works done in this regard was on North American names [6]. In this work, researchers used morphological features of English language and find out many handful features of sound and language. Similarly, Tripathi and Faruqui [3] used support vector machine (SVM) approach for gender classification using Indian names. They used n-gram suffix along with other morphological features to classify males and females names.

C. What makes our Work Different

As discussed above, we could only find works on English and Indian names for gender prediction. Therefore, to the best of our knowledge there exists no work for Urdu names. Morphological analysis of American, Indian and Urdu names reveals their differences [3]. This makes our work different from the existing work.

Second thing that makes our approach distinct from existing approaches is the use of a new feature, that is, combination of letters.

Another difference is the size of data collection. We use around 3047 names (1729 female names while 1308 males) for our work while work for Indian names [3] used a collection of 2000 names (890 female and 1110 male names) while work on North American names [6] included 489 names (222 females and 267 male).

Last but not the least is the use of only textual features for identification of gender. We did not use sound-based features like syllables and sonorant consonant ending.

The following table shows feature-based analysis of our training data. We consider names long if they contain six or more letters.

TABLE I. TRAINING DATA FEATURE ANALYSIS

Feature	Male	Female
Length	36.03%	63.97 %
Vowel Ending	4.00%	49.77%

III. MACHINE LEARNING FEATURES

Previous works [3, 6] have used the following features differentiating between male and female names. We have only used a subset of the following features in our work because our focus is on using only text-based features.

Vowel Ending: Names of females generally end in a vowel while that of males in consonants.

Number of syllables: A syllable is a unit of pronunciation uttered without interruption, loosely a single sound. Female names tend to have more number of syllables than males.

Sonorant Consonant Ending: A sonorant is a sound that is produced without turbulent airflow in the vocal tract. Hindi possesses eight sonorant consonants [19]. Compared to females, male names generally end with a sonorant consonant.

Length of the Word: Even though length of a name does not relate to its gender, our data shows that females have longer names than males in Pakistani names when compared to Indian names where opposite trend has been reported [3].

A. Issues with Previously Used Features

Previously used features for gender prediction through names are a mixture of textual and speech based characteristics. However, we focus on using only textual based features which is more practical when predicting gender through names in real time. Each language has its own conventions inherited from the region where it is spoken. Therefore, textual features like vowel ending and length of the word might not behave the same way for Urdu language as for other languages. Therefore, we propose a new feature called "combination of letters" for gender prediction through Urdu language names. This feature tries to capture the consecutive or nonconsecutive combination of letters in names. "Combination of letters" could prove very useful when Urdu names (written in Roman Urdu) have the same ending letters and length because in that context these can't accurately distinguish between "Male" and "Female" names. Table 2 describes some examples of Urdu names (in Roman Urdu) with all three textual features.

TABLE II. DIFFERENT FEATURES EXAMPLES

Examples of Some Names with Features				
Name	Length	1gram	Combination of Letters	Gender
DANYAL	6	L	AN	M
FARYAL	6	L	AR	F

In this table, we can see that names "Danyal" and "Faryal" have same lengths and 1 gram ending but it is "combination of letters" which can help recognizing the gender of the name. To compute this feature, that is, "Combination of letter", we develop an algorithm which extracts this information automatically.

IV. EXPERIMENTS

A. Data Set

We prepared a dataset of Urdu (Pakistani) names ourselves from online web sites (containing Urdu names and their meanings) and old PTCL (Pakistan Telecommunication Limited) telephone directories available. All names are written in Roman Urdu script i.e. using English language alphabets. It is to be noted that most of the Urdu linguistics resources have been developed by Centre of Language Engineering[2] but we could not find a collection for Urdu names on their web site even they claim that they have developed one already [18].

[2] http://www.cle.org.pk/software/ling_resources.htm

Our data collection consists of 3047 Pakistani names. It consists of 1729 female while 1308 male names.

B. Classifiers

We use Decision Tree (J48), Support Vector Machine (SVM), K-nearest neighbor (Lazy-IBK) and Random Forest classifiers for individual as well as for different combination of features and compare their performances on results of testing data. We use 1828 (almost 60 percent of total data) name instances for building training model while rest of the 1219 instances are used as testing data. We use Weka[3] toolkit for our experimentation.

C. Results and Discussions

In this section, we describe the results obtained through different classifiers using individual or combination of different features.

a) Decision Tree: Following tables shows results for decision tree classifier.

TABLE III. RESULTS FOR DECISION TREE CLASSIFER

Classifier	Feature	Accuracy
J48	Length	61.53%
	Unigram	83.92%
	Combination of letters	57.91%
	2-gram	83.84%
	3-gram	78.261%
	Length and unigram	85.24%
	Length and combination of letters	61.03%
	Length and 2-gram	84.00%
	Length and 3-gram	78.67%
	Unigram and combination of letters	**86.55%**
	Length, unigram and combination of letters	84.74%
	2-gram, length and combination of letters	84.66%
	3-gram, length and combination of letters	79.33%
	2-gram and combination of letters	84.49%
	3-gram and combination of letters	78.75%

b) Support Vector Machine: Following tables shows results for SVM classifier.

TABLE IV. RESULTS FOR SVM CLASSIFER

Classifier	Feature	Accuracy
SMO	Length	57.01%
	Unigram	83.92%
	Combination of letters	57.91%
	2-gram	84.74 %
	3-gram	81.46%
	Length and unigram	84.74%
	Length and combination of letters	56.77%
	Length and 2-gram	84.74%
	Length and 3-gram	81.46%
	Unigram and combination of letters	**84.74%**
	Length, unigram and combination of letters	83.92%
	2-gram, length and combination of letters	84.74%
	3-gram, length and combination of letters	81.46%
	2-gram and combination of letters	84.74%
	3-gram and combination of letters	81.46%

[3] www.cs.waikato.ac.nz/ml/**weka**/

c) K-Nearest Neighbour (KNN): Following tables shows results for KNN classifier.

TABLE V. RESULTS FOR KNN CLASSIFIER

Classifier	Feature	Accuracy
Lazy-IBK	Length	61.53%
	Unigram	83.92%
	Combination of letters	57.91%
	2-gram	83.98%
	3-gram	78.26%
	Length and unigram	84.90
	Length and combination of letters	61.61%
	Length and 2-gram	84.05 %
	Length and 3-gram	81.46%
	Unigram and combination of letters	**86.30%**
	Length, unigram and combination of letters	83.92%
	2-gram, length and combination of letters	85.15%
	3-gram, length and combination of letters	82.44%
	2-gram and combination of letters	84.90%
	3-gram and combination of letters	80.06%

d) Random Forest: Following tables shows results for random forest classifier.

TABLE VI. RESULTS FOR RANDOM FOREST CLASSIFIER

Classifier	Feature	Accuracy
Random Forest	Length	61.53%
	Unigram	83.92%
	Combination of letters	57.91%
	2-gram	83.92%
	3-gram	78.51%
	Length and unigram	84.66%
	Length and combination of letters	61.61%
	Length and 2-gram	83.75%
	Length and 3-gram	78.99%
	Unigram and combination of letters	**86.30%**
	Length, unigram and combination of letters	84.24%
	2-gram, length and combination of letters	85.23%
	3-gram, length and combination of letters	79.82%
	2-gram and combination of letters	85.23%
	3-gram and combination of letters	79.16%
	2-gram and combination of letters	85.06%
	3-gram and combination of letters	79.16%

Among individual features, unigram seems to be outperforming all other individual features for all classifiers. However, when unigram feature is combined with combination of letters, it further boosts up its performance to 86.55% from 83.92 % for decision tree, to 84.74 % from 83.92 % for SVM, to 86.30 % from 83.92% for KNN and Random Forest classifier. Accuracy of different classifiers for this combination is also shown in figure 1 (at end of the document). It is also very interesting to observe that 2-gram features seem to be playing more effective role than 3-gram features. We think it is because of the relatively shorter length of the names than other type of general words. Another positive aspect of these results is using 2-gram features with combination of letters which helps in further improving accuracy of 2-gram features.

V. Conclusions and Future Work

Results show that newly purposed feature. that is, finding combination of letters in name and unigram both together is the best feature for predicting gender by name written in URDU (Roman like English). We got highest accuracy of 86.54% with J48 classifier. It proves that using only textual based features can also improve gender prediction from names while existing works have achieved similar level of accuracy by using both i.e. speech and textual features.

While we have mentioned above that gender prediction from names is very important for expert finding task. We have focused on gender prediction in this work and we keep the task of predicting geographical background of the author from his /her text. For example, we might collect a data collection on same topic for authors from different locations and then find hidden patterns in their writings to determine their geographical background automatically. This task can be very helpful in determining political orientations or extremists attitudes for different topics.

References

[1] Yimam, Dawit. "Expert Finding Systems for Organizations: Domain Analysis and the DEMOIR approach. ECSCW 99 Beyond Knowledge Management: Management Expertise Workshop." (2000): 276-283.

[2] Karimzadehgan, Maryam, Ryen W. White, and Matthew Richardson. "Enhancing expert finding using organizational hierarchies." Advances in Information Retrieval. Springer Berlin Heidelberg, 2009. 177-188.

[3] Tripathi, Anshuman, and Manaal Faruqui. "Gender prediction of Indian names." Students' Technology Symposium (Tec hSym), 2011 IEEE. IEEE, 2011.

[4] Argamon, Shlomo, et al. "Automatically profiling the author of an anonymous text." Communications of the ACM 52.2 (2009): 119-123.

[5] Estival, Dominique, et al. "Author profiling for English emails." Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics (PACLING'07). 2007.

[6] A. S. Slater and S. Feinman, "Gender and the phonology of north american first names," Sex Roles, vol. 13, pp. 429–440, 1985, 10.1007/BF00287953

[7] Quanzeng You, Sumit Bhatia, Tong sum, Tiebo Luo," The eyes of the beholder: Gender prediction using images posted in Online Social Networks",2014 IEEE International Conference on Data Mining Workshop, Department of Computer Science, University of Rochester, NY

[8] Moghaddam, Baback, and Ming-Husan Yang. "Learning gender with support faces." Pattern Analysis and Machine Intelligence, IEEE Transactions on24.5 (2002): 707-711.

[9] Cornwall, Andrea. "Whose voices? Whose choices? Reflections on gender and participatory development." World development 31.8 (2003): 1325-1342.

[10] Wu, Ke, and Donald G. Childers. "Gender recognition from speech. Part I: Coarse analysis." The journal of the Acoustical society of America 90.4 (1991): 1828-1840.

[11] Rangel, Francisco, et al. "Overview of the author profiling task at pan 2013."CLEF Conference on Multilingual and Multimodal Information Access Evaluation. CELCT, 2013.

[12] Zhang, Cathy, and Pengyu Zhang. Predicting gender from blog posts. Technical Report. University of Massachusetts Amherst, USA, 2010.

[13] Thomson, Rob, and Tamar Murachver. "Predicting gender from electronic discourse." British Journal of Social Psychology 40.2 (2001): 193-208.

[14] Peersman, Claudia, Walter Daelemans, and Leona Van Vaerenbergh. "Predicting age and gender in online social networks." Proceedings of the 3rd international workshop on Search and mining user-generated contents. ACM, 2011.

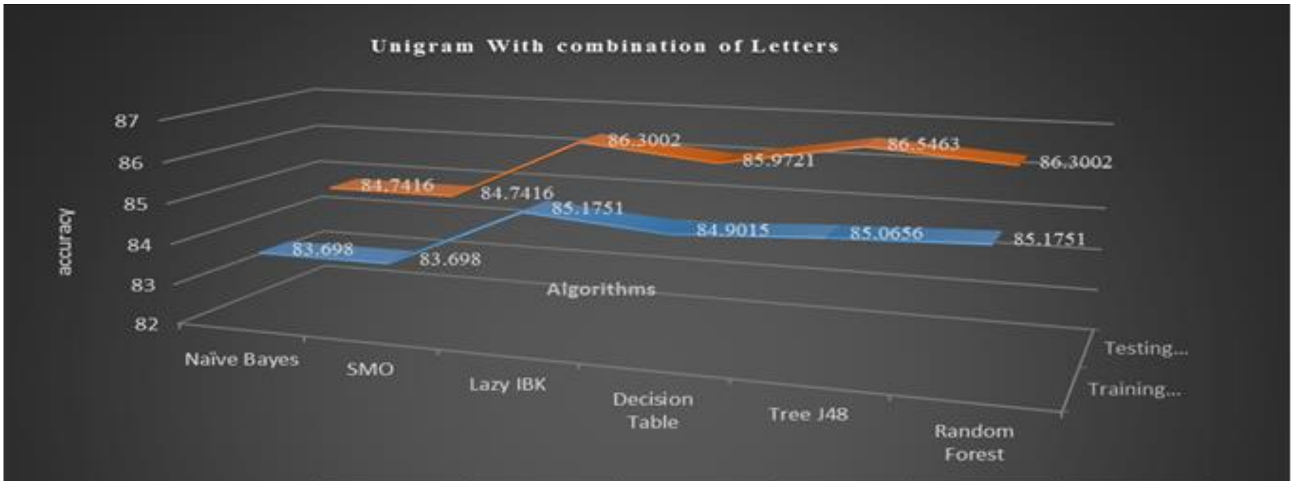
[15] Estival, Dominique, et al. "Author profiling for English emails." Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics (PACLING'07). 2007.

[16] Argamon, Shlomo, et al. "Automatically profiling the author of an anonymous text." Communications of the ACM 52.2 (2009): 119-123.

[17] Schler, Jonathan, et al. "Effects of Age and Gender on Blogging." AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs. Vol. 6. 2006..

[18] Hussain, Sarmad. "Resources for Urdu Language Processing." IJCNLP. 2008.

[19] Gordon, Matthew, et al. "Vowel and consonant sonority and coda weight: A cross-linguistic study." *WCCFL*. Vol. 26. 2008.



	Naïve Bayes	SMO	Lazy IBK	Decision Table	Tree J48	Random Forest
Training Accuracy	83.698	83.698	85.1751	84.9015	85.0656	85.1751
Testing Accuracy	84.7416	84.7416	86.3002	85.9721	86.5463	86.3002

Fig. 1. Result

A Robust Approach for Action Recognition Based on Spatio-Temporal Features in RGB-D Sequences

Ly Quoc Ngoc, Vo Hoai Viet, Tran Thai Son, Pham Minh Hoang
Department of Computer Vision and Robotics, University Of Science, VNU-HCM, Viet Nam

***Abstract*—Recognizing human action is attractive research topic in computer vision since it plays an important role on the applications such as human-computer interaction, intelligent surveillance, human actions retrieval system, health care, smart home, robotics and so on. The availability the low-cost Microsoft Kinect sensor, which can capture real-time high-resolution RGB and visual depth information, has opened an opportunity to significantly increase the capabilities of many automated vision based recognition tasks. In this paper, we propose new framework for action recognition in RGB-D video. We extract spatiotemporal features from RGB-D data that capture both visual, shape and motion information. Moreover, the segmentation technique is applied to present the temporal structure of action. Firstly, we use STIP to detect interest points both of RGB and depth channels. Secondly, we apply HOG3D descriptor for RGB channel and 3DS-HONV descriptor for depth channel. In addition, we also extract HOF2.5D from fusing RGB and Depth to capture human's motion. Thirdly, we divide the video into segments and apply GMM to create feature vectors for each segment. So, we have three feature vectors (HOG3D, 3DS-HONV, and HOF2.5D) that represent for each segment. Next, the max pooling technique is applied to create a final vector for each descriptor. Then, we concatenate the feature vectors from the previous step into the final vector for action representation. Lastly, we use SVM method for classification step. We evaluated our proposed method on three benchmark datasets to demonstrate generalizability. And, the experimental results shown to be more accurate for action recognition compared to the previous works. We obtain overall accuracies of 93.5%, 99.16% and 89.38% with our proposed method on the UTKinect-Action, 3D Action Pairs and MSR-Daily Activity 3D dataset, respectively. These results show that our method is feasible and superior performance over the-state-of-the-art methods on these datasets.**

Keywords*—Action Recognition; Depth Sequences; GMM; SVM; Multiple Features; Spatio-Temporal Features

I. Introduction

Automatic human action recognition is attractive research topic in the fields of computer vision and machine learning since it plays an important role in the applications such as human-computer interaction, intelligent surveillance, human action retrieval system, health care, smart home, and robotics. Due to its wide range of applications, automatic human action recognition has attracted much attention in recent years [11, 19, 20, 31, 37]. The goal of human action recognition is to automatically analyze ongoing action from an unknown video (i.e. a sequence of image frames).

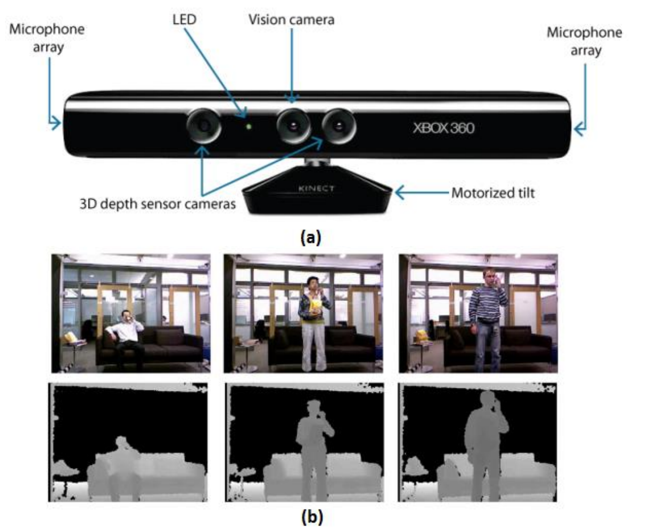


Fig. 1. Illustration of 3D camera and RGB-D data: a) Microsoft Kinect Device; b) Some examples of RGB-D data is captured by Kinect

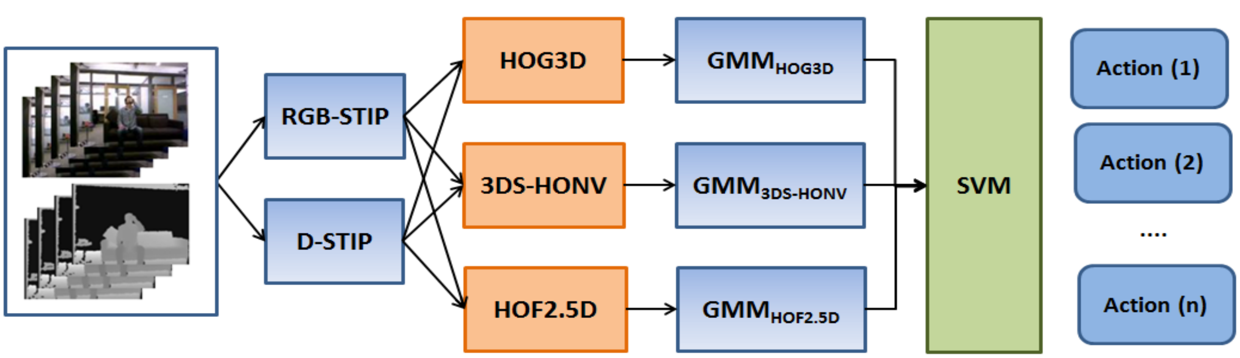


Fig. 2. Our proposal framework for action recognition in RGB-D video

Generally speaking, action recognition framework contains three main steps namely feature extraction, action representation, and pattern classification. Though much progress has been made [1, 6, 9, 11, 13, 19, 20, 31, 37], the problem of classifying action is currently of the most difficult challenges, especially in the presence of within-class variation, occlusion, background clutter, pose and lighting condition. These challenges address that the combination of different kinds of features action because action representation based on single feature is not enough to capture the imaging variations (view-point, illumination, etc…) and attributes of individuals (appearance, shape, motion, etc…).

Following the previous researches [1, 3, 6, 13, 15, 16, 18, 20], human action could be defined by structured patterns of the human's movements and poses. With the perspective, a robust feature extraction and description must capture shape and motion properties in action representation. As such, human action can be modeled by spatiotemporal features, where encode shapes and movements of the whole body or body parts, for instance temporal progression, e.g., one human action as the whole can decompose into local shapes and movements of parts. In the past two decades, a significant amount of research has been done in the area of human action recognition using a sequence of 2D images [1, 6, 13, 14, 15, 16, 18, 46]. A single spatiotemporal structure, however, is unlikely to be sufficient to represent a class of action in all but the simplest scenarios. Firstly, the execution of the action may differ from subject to subject, involving different body parts or different space-time progressions of body part movements. Secondly, the video capture process introduces intra-class variations due to occlusions or variations in camera viewpoint. Thus, the resulting space-time and appearance variations necessitate using a collection of spatiotemporal structures that can best represent the action at large. In addition, another property be also considered in action representation is evolution of action by time. It indicates that action also contains temporal structure for each action class. In this work, we apply video segmentation and max-poling technique which help to model temporal structure of action.

With the recent advent of the cost-effective Kinect, depth cameras have received a great deal of attention from researchers. It is excited to promote interest within the vision and robotics community for its broad applications [27]. The depth sensor has several advantages over the visible light camera. Firstly, the range sensor provides 3D structural information of the scene, which offers more discerning information to recover postures and recognize actions. The common low-level difficulties in RGB imagery are significantly alleviated. Secondly, the depth camera can work in total darkness. There is a benefit for applications such as patient/animal monitoring systems which run 24/7. With these benefits, the Kinect has been opened a new opportunity to improve the performance of human action recognition significantly. Recently, researchers have paid more attention to using 3D spatiotemporal features for describing and recognizing human actions [3, 4, 29, 35, 36, 44, 46] based on depth information from Kinect. Compared with conventional color data, depth maps provide several advantages, such as the ability to reflect pure geometry and shape cues, or insensitive to changes in lighting conditions. Moreover, the range sensor provides 3D structural information of the scene, which offers more discerning information to recover postures and recognize actions. These properties help depth data provide more natural and discriminative vision cues than color or texture. Furthermore, the depth images provide natural surfaces which can be extracted to capture the geometrical structure of the observed scene in a rich descriptor. However, depth sensors cannot differentiate between objects of the same depth but different color, which is trivial for color cameras. Clearly the color and depth information are correlated but also complementary to a large extent, so it would be expected to have considerable benefits by fusing them appropriately together aiming at more robust pervasive action recognition systems.

In all case, furthermore, it is commonly believed that in order to obtain high recognition rate, it is important to select an appropriate set of visual features that usually have to capture the particular properties of a specific domain and the distinctive characteristics of each action class. The most important aspect of any action recognition system is to seek an efficient action representation. The target of the feature extraction is to find an efficient and effective representation of the action which would provide robustness during recognition process. Besides, in case action representation from multiple feature vectors will need a robust method to combine feature vectors in the right way so that the system achieves good performance. In this work, we use the average-pooling technique to aggregating visual words in BOW model and the max pooling technique to aggregating the segment feature vectors into the final feature vector for action representation.

In this manuscript, we build a new framework for action recognition upon our previous works in [35] and [36]. The proposed action recognition system consisted of a flowchart is shown in Fig. 2. The main contributions of this paper are summarized as follows: Firstly, we propose a new framework for action recognition, which takes profits of multi-modal RGB-D data by fusing information from both RGB images and depth maps. The spatiotemporal features are applied to capture shape and motion. A new action presentation method is proposed by using segmentation and max pooling technique in order to capture temporal structure of human action. In addition, we use GMM instead of k-means in BOW model in order to be more distinctive for action representation. Secondly, we systematically evaluate our frameworks on three challenging datasets. Moreover, we also evaluate the impact of video segmentation technique and spatiotemporal descriptors on the performance of the system in overall accuracy.

The rest of this paper is organized as follows: Section II gives a concise review of existing works on feature extraction from a sequence of images and depth. Section III presents feature extraction and description. Section IV introduces a scheme of action representation. Section V presents action classification. Section VI shows the experiment results on relevant benchmarks. Finally, section VII draws conclusions of our work and indicates future studies.

II. Related Works

Comprehensive reviews of the previous studies can be found in [19, 20, 31, 37]. Our discussion in this section is restricted to a few influential and relevant parts of literature, with a focus on RGB, depth and RGB-D for feature extraction and representation.

There has been a lot of works on human action recognition from images in recent decades, that could be divided into two types of approaches: global-based and part-based method. Global features is temporal templates is introduced by Bobick and Davis [6]. They use the two components of motion template (MEI and MHI) and Hu Moments for representation and recognition of human movement. Xinghua Sun [42] use Zernike moments instead of Hu moments for action representation. Beside global feature approaches, the local features methods such as: histogram of 3D oriented gradients (HOG3D) [1], histogram of optical flow (HOF) [17], 3D speeded up robust features (SURF3D) [13] extends from SURF [5], 3D scale invariant feature transforms (3D-SIFT) [34] extends from SIFT [10], local trinary patterns [26] and dense trajectories with HOG/HOG/MHB[15, 16] are used to extract the most salient features (edges, corners, orientation, and motion), the choice of which would greatly influence the performance of high-level vision tasks such as recognition. Viet Vo and Ngoc Ly .al [40] also proposed hybrid features that combine local and global features for action representation. In addition, soft-weighting scheme was used to achieve more descriptive in BOW representation.

Recently, with the availability of low-cost RGB-D sensors, The similarly to recognizing human action from 2D video, the depth map-based methods rely mainly on features, either local or global, extracted from the space time volume. Lu Xia at [29] proposed DSTIP based on STIP's idea in RGB images Liet al [28] sample representative 3D points extracting the points on the contours of the projections of the 3D depth map onto the three orthogonal Cartesian planes. To reduce the size of the feature vector, the method selects a specified number of points at equal distance along the contours of the projections. Wang al. [22] fuses the skeleton information and a local occupancy pattern based on the 3D point cloud around each joint. In a different approach, J.Wang al.[23] treat an action sequence as a 4D shape and propose random occupancy pattern features, which are extracted from randomly sampled 4D sub-volumes with different sizes and at different locations. These features are robust to noise and less sensitive to occlusions. Furthermore, holistic approaches for action recognition from depth sequences are recently becoming popular. Vieira al. [3] proposed Space-Time Occupancy Patterns. The depth sequence is represented in a 4d space-time grid. Then, a scheme is used to enhance the roles of the sparse cells which typically consist of points on the silhouettes or moving parts of the body. Oreifej and Liu [33] describe the depth sequence using a histogram that captures the distribution of the surface normal orientation in the 4D space of time, depth, and spatial coordinates. The similar Oreifej's idea, Quang D. Tran and Ngoc Q. Ly [35] proposed 3DS-HONV descriptor that uses Euler angles-based quantization to create 3D histogram for action representation. This approach is simpler than the Oreifej's approach in angle quantization step. In addition, optical flows are extracted from depth channel to obtain more descriptive. Xiaodong Yang .al [41] also proposed SNV based on the surface normal orientation with adaptive spatiotemporal pyramid. Yang al. [43] project the depth maps onto three orthogonal planes and accumulate the whole sequence generating a depth motion map (DMM), the similar idea to the motion history images [6]. Histograms of oriented gradients [32] are obtained for each DMM. The concatenation of the three HOG represents an action. These features encode more information about shape, motion and context.

Nearly, some researches focus on combining both color and depth data for action recognition. Zhao Yang [47] used STIP to detect interest point and descriptor is described by combining HOG/HOF from RGB and LDP from depth data. These descriptors are used to build codebook for action representation. Quang D. Tran and Ngoc Q. Ly [35, 36] also used STIP to detect interest points but the descriptor is combined by 3DS-HONV and HOG-HOF2.5D. And, sparse coding is applied on these descriptors for representation. In [24], L. Liu proposed graph-based genetic programming by applying filters into RGB and depth data to automatically extract discriminative spatiotemporal features for action representation and SVM was used to classify actions. However, feature learning approaches have complexity in computing.

In this work, we propose a new framework for human action recognition that combines both RGB images and depth maps. This approach falls in the part-based method category. More details, we use spatiotemporal features based on the interest points that are detected by STIP in both RGB and depth channels. These interest points are represented by HOG3D, 3DS-HONV and HOF2.5D that capture shape, appearance and motion of action. Moreover, we also apply video segmentation and max pooling techniques to capture the temporal structure for action representation.

III. Feature Extraction and Description

The key to the success of part-based methods is that the interest points are distinctive and descriptive. Following the approach commonly used for local interest points in images and video, the detection and description of spatiotemporal interest points are separated in two different steps. This section describes local feature detector and descriptor used in our approach. For spatiotemporal interest points detector, we apply STIP detector [18] as a space-time extension of the Harris detector [8]. For spatiotemporal interest points descriptors, we use three descriptors such as HOG3D [1], 3DS-HONV [35], and HOF2.5D [36].

A. Prepocessing Stage

The 3D sensors such as Kinect based on structured light to estimate depth information, it is prone to be affected by noises due to reflection issues. These effects of noise could significantly decrease the overall performance of RGBD-based action recognition framework. Therefore, we firstly relieve the missing data and oultliers from the depth channel. As a result at [16], we adopted the bilateral filter for smoothing the depth channel. The bilateral filter [30] is a combination of a domain kernel, which gives priority to pixels that are close to the target pixel in the image plane, with a range kernel, which gives priority to the pixels which have similar labels as the target pixel. This filter is often useful to preserve edge information based on the range kernel advantages. The edge is important information to represent shape of action. The bilateral filter is defined as follows:

$$I^f(x) = \frac{1}{W_p}\sum_{x_i\in\Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|I(x_i) - I(x)\|)$$

$$W_p = \sum_{x_i\in\Omega} f_r(\|I(x_i) - I(x)\|) g_s(\|I(x_i) - I(x)\|)$$

Where I^f is the filtered image, I is the original input image , x are the coordinates of the current pixel to be filtered, Ω is the window centered in x, f_r is the range kernel for smoothing differences in intensities and g_s is the spatial kernel for smoothing differences in coordinates. In this research, f_r and g_s are supposed as Gaussian functions.

B. *Interest Point Detection*

The STIP or Harris3D detector was proposed by Laptev and Lindeberg in [18], which is an extension of the well-known Harris detector in the temporal dimension. The STIP detector first computes the second-moment 3 × 3 matrix μ of first order spatial and temporal derivatives. Then, the detector searches regions in the video with significant eigenvalues λ_1, λ_2, λ_3 of μ, combining the determinant and the trace of μ:\

$$H = |\mu| - k \cdot T_r(\mu)^3$$

where |·| corresponds to the determinant, Tr(·) computes the trace, and k stands for a relative importance constant factor. A commonly used value of k in the literature is k ≈ 0.005 As we have RGB-D data, we apply the STIP detector separately on the RGB and depth channels, so we get two sets of interest points for description step.

C. *HOG3D Descriptor*

The HOG3D descriptor was proposed by Kläser et al. [1]. It is based on histograms of 3D gradient orientations and can be seen as an extension of the well-known SIFT descriptor [10] to video sequences. Gradients are computed using an integral video representation. Regular polyhedrons are used to uniformly quantize the orientation of spatiotemporal gradients. The descriptor, therefore, combines shape and motion information at the same time. A given 3D patch is divided into $n_x \times n_y \times n_t$ cells. The corresponding descriptor concatenates gradient histograms of all cells and is then normalized. The process of computing the HOG3D descriptor for a patch in an action depth sequence is described in Fig. 3.

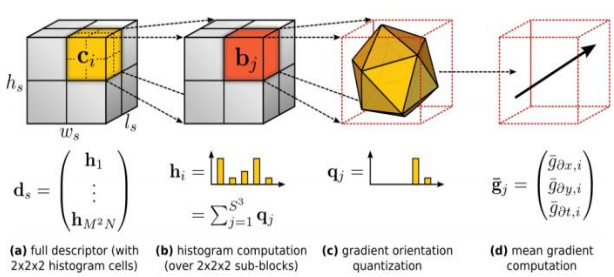


Fig. 3. Process of extracting HOG3D descriptor [1]

D. *3D Spherical Histogram of Oriented Normal Vectors (3DS-HONV) Descriptor*

The 3DS-HONV descriptor was proposed by Quang D. Tran, Ngoc Q. Ly in [35], which based on HONV in [38]. The process of computing the 3DS-HONV descriptor for a patch in an action depth sequence is described in Fig. 4. For each patch, the orientation of the normal vector at each depth point is first computed, quantized in spherical coordinate by using 3 angles θ, φ ψ, and voted into a 3D histogram $q_i \in R^{b_\theta \times b_\varphi \times b_\psi}$, where b_i is the relevant bin size. Those 3D histograms at all interest points are then accumulated to create a histogram of normal occurrences distribution. Implementation of this computing process is described as follows:

1) *Spatio-Temporal Surface Oriented Normal Vectors*

The depth sequence can be considered as a function $R^3 \rightarrow R^1$: $z = d(x, y, t)$ ($d(.)$ (is a function of depth sequence) which constitutes a surface in the 4D space represented as the set of points $\{p = (x, y, t, z)\}$ satisfying $S(p) = d(x, y, t) - z = 0$. The normal to the surface S is computed as:

$$n = \nabla S = (z_x, z_y, z_t, -1) = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial z}{\partial t}, -1\right)$$

where z_x, z_y, z_t are first derivatives of the depth map z over x, y, t, which can be computed by using the finite difference approximation respectively. Since only the orientation of the normal could describe the shape of the 4D surface, the computed normal vectors are then normalized to a unit length as follows:

$$\hat{n} = \left(\hat{z}_x, \hat{z}_y, \hat{z}_t, -1/\|(z_x, z_y, z_t, 1)\|_2\right)$$

2) *Spherical Quantization and 3D Histogram Representation:*

In our work, the orientation of spatiotemporal surface normal is characterized by three Euler angles { θ, ϕ, and ψ } ∈ [0; π] computed in spherical coordinate. The Euler angles are a classical way to specify the orientation of an object in space with respect to a fixed set of coordinate axes [21]. According to Euler's rotation theorem[21], any rotations may be described using three angles; therefore, we clarify that by just using 3 Euler angles θ, ϕ, ψ for quantization, the resulting histogram can encode any kinds of surface normal orientation in a rich representation. Euler angles-based quantization is simple, intuitive, but also more efficient than quaternions-based quantization. The approximate computation of Euler angles θ, ϕ, and ψ [21] are summarized as follows:

$$\theta = \tan^{-1}\left(\frac{\partial \hat{z}}{\partial \hat{y}} \Big/ \frac{\partial \hat{z}}{\partial \hat{x}}\right)$$

$$\phi = \tan^{-1}\left[\left(\left(\frac{\partial \hat{z}}{\partial \hat{x}}\right)^2 + \left(\frac{\partial \hat{z}}{\partial \hat{y}}\right)^2\right)^{1/2} \Big/ \frac{\partial \hat{z}}{\partial \hat{t}}\right]$$

$$\psi = \left(\left(\frac{\partial \hat{z}}{\partial \hat{x}}\right)^2 + \left(\frac{\partial \hat{z}}{\partial \hat{y}}\right)^2 + \left(\frac{\partial \hat{z}}{\partial \hat{t}}\right)^2\right)^{1/2}$$

In order to create 3D histogram representation for each depth point, the [0; π] interval is subdivided in b_θ, b_ϕ ,b_ψ bins, so that the histogram has a total of $b_\theta \times b_\phi \times b_\psi$ bins, and is then normalized to compute the proportion of normals falling into each bin. In this work, we use the tuple of bins' size which are $\{b_\theta = 5, b_\phi = 5, b_\psi = 6\}$, this means a 150-dimensions 3DS-HONV for each interest point.

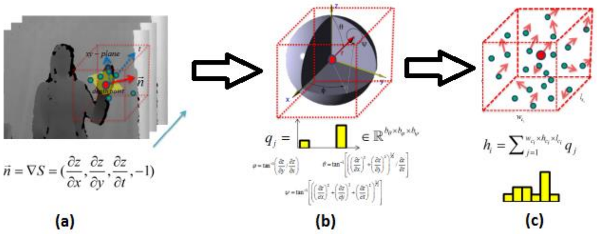


Fig. 4. Process of extracting 3DS-HONV descriptor from an interest point [36]: (a) Surface normal is computed at each point, (b) 3D histogram of normal distribution in spherical coordinate is constructed, (c) 3D histograms at all points are accumulated

E. HOF2.5D Descriptor

According to many previous researches [20, 31], the motion plays important role in human action analysis. In order to have a good representation for human action is feature descriptor must capture this property. The 3DS-HONV descriptor was proposed by Quang D. Tran and Ngoc Q. Ly in [36] which contains the human motion. This descriptor is not generated from a unified image sequence function $f(x, y, z, t)$, but instead of capturing separately the xy-motion from pairs of RGB images and the z movements from pairs of depth channel. With assuming that is the position of each pixel in RGB images can be mapped to the related cloud point in depth maps. In specific, each pixel ($p_t^{RGB} = \{x_t^{RGB}, x_t^{RGB}\}$) in RGB-D frame F_t can be easily projected to its corresponding position ($p_t^D = \{x_t^D, x_t^D\}$) in the depth map. The process of computing HOF2.5D descriptor is described as follows: each RGB frame F_t^{RGB}, the $\{V_x, V_y\}$ components of the optical flow fields (OF) at every pixels are computed using algorithm that was proposed by G. Farnebäck algorithm [12]. In order to create OF2.5D at each calibrated pixel ($p_t = \{p_t^{RGB}, p_t^D\}$), we utilize the information of available depth maps to compute the V_z component of the OF vector as this formulation:

$$V_z = F_{t+1}^D(p_{t+1}^D) - F_t^D(p_t^D)$$

As results, each RGB-D frame F_t, we obtain a feature descriptor $D = \{D_1, V_2, \ldots, V_n\}$, where each element $D_i = \{V_x, V_y, V_z\}$ is a 3D vector that captures satisfactorily 3D motion information of a particular pixel. As a final representation for each interest point, we perform a histogram quantization using three orthogonal planes xy, xz, yz as shown in Fig. 6. The orientations of each OF2.5D are computed on three projected planes as follows:

$$\alpha_1 = \tan^{-1}\left(\frac{V_y}{V_x}\right)$$

$$\alpha_2 = \tan^{-1}\left(\frac{V_z}{V_x}\right)$$

$$\alpha_3 = \tan^{-1}\left(\frac{V_z}{V_y}\right)$$

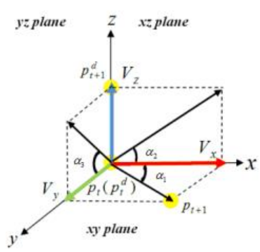


Fig. 5. Quantization scheme for computing HOF2.5D [36]

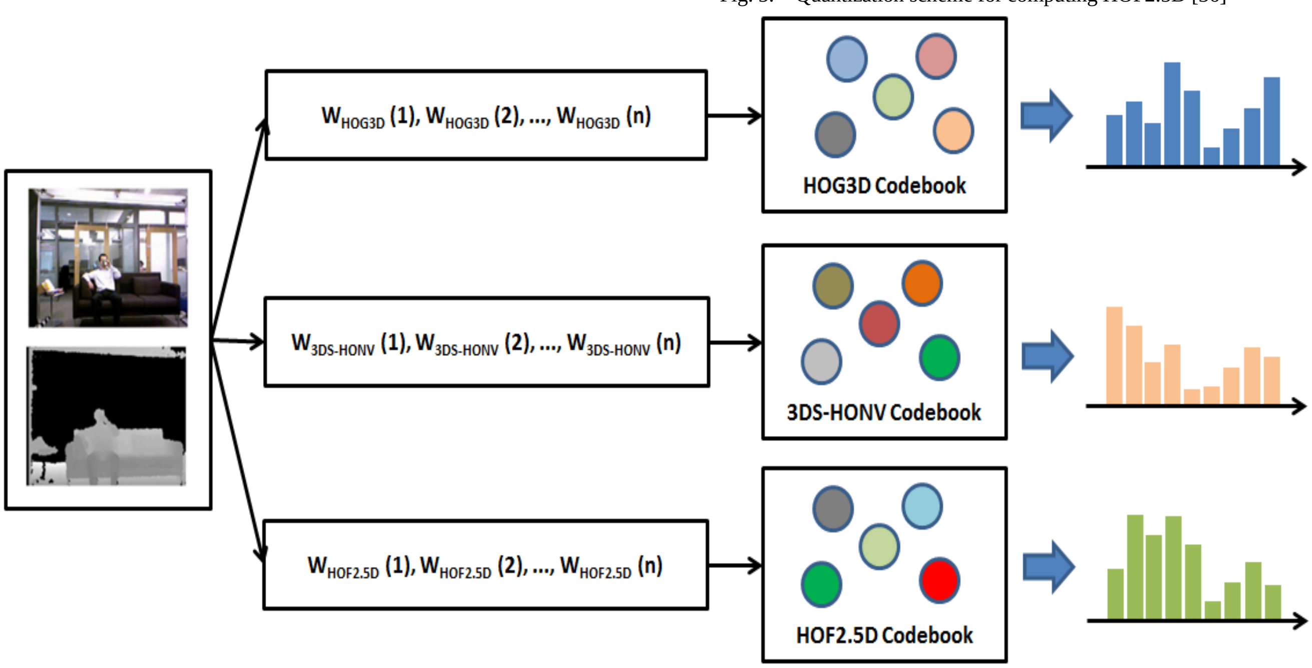


Fig. 6. Illustration of BOW for action representation in RGB-D data

We then evenly deploy $b_{\alpha 1}$, $b_{\alpha 2}$, $b_{\alpha 3}$ orientations binning on three orthogonal planes to finally generate a histogram representation of each semi-scene flow vector, namely as HOF2.5D. In all experiments, we set $b_{\alpha 1} = b_{\alpha 2} = b_{\alpha 3} = 8$. As a consequence, for each interest point descriptor, by accumulating all HOF2.5D descriptors at all pixels, we achieve a 24-bins histogram that captures the distribution of motion flows.

IV. Action Representation

A. Bag of Word

In part-based methods, a video is modeled by the bag of words (BOW) model which is the way of constructing a feature vector based on the number of occurrences of word. Each visual word is just a feature vector of patch. The major issue of BOW is vector quantization algorithms to create effective clusters. The original BOW used k-means algorithm to quantize feature vectors. Although k-means is used widely in clustering, its accuracy is not good in some cases. In addition, binary weighting for histogram of word occurrences which indicates the presence and absence of a visual word with values 1 and 0 respectively, was used. Generally speaking, all the weighting schemes perform the nearest neighbor search in the vocabulary in the sense that each interest point is mapped to the most similar visual word. Many researches argue that, for visual words, directly assigning an interest point to its nearest neighbor is not an optimal choice, given the fact that two similar points may be clustered into different clusters when increasing the size of visual vocabulary. On the other hand, simply counting the votes is not optimal as well. For instance, two interest points assigned to the same visual word are not necessarily equally similar to that visual word, meaning that their distances to the cluster centroid are different. Ignoring their similarity with the visual word during weight assignment causes the contribution of two interest points equal, and thus more difficult to assess the importance of a visual word in video.

In this work, we propose GMM instead of k-means in BOW model. We denote the parameters of the K-component GMM by $\lambda = \{w_k, \mu_k, \Sigma_k, k = 1, \dots, K\}$, where w_k, μ_k and Σ_k are respectively the mixture weight, mean vector and covariance matrix of Gaussian k and subject to $\sum_k w_k = 1$. In this work, we set K = 512 that are used in many researches. We estimate the GMM parameters on a large X training set of local spatiotemporal descriptors using the Expectation-Maximization (EM) [2] algorithm to optimize a Maximum Likelihood (ML) criterion. For GMM, soft quantization corresponds to assigning features partially to each of the GMM clusters, according to their posterior probabilities:

$$v_i = \left[P_{K|X}(1, x_i), P_{K|X}(2, x_i), \dots, P_{K|X}(K, x_i)\right]$$

$$P_{K|X}(k, x) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_k|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x - \mu_k)^T {\Sigma_k}^{-1}(x - \mu_k)\right\}$$

where v_i is the vector of soft-counts associated with feature x_i. The soft-weights of each visual word, contributed by all features in the video, are then pooled into a histogram:

$$H(V) = F(v_1, v_2, \dots, v_n)$$

which is the final video representation (n is the number of descriptors). The standard average pooling operator aggregates word counts into bins of $H(V)$ and normalizes as follows:

$$F_{av}(v_1, v_2, \dots, v_n) = \frac{1}{n}\sum_i v_i$$

$H(V)$ represents a histogram for the video V.

B. Video Segmentation

Video segmentation is the method that divides video into fixed length segments. These approaches can be divided into two types: non-overlapping and overlapping segments. For non-overlapping segments, a video is divided into continuous and equal length segments. The method does not take account information about the semantic boundary of a segment. However, this information is important because it keeps semantic meaning of each segment.

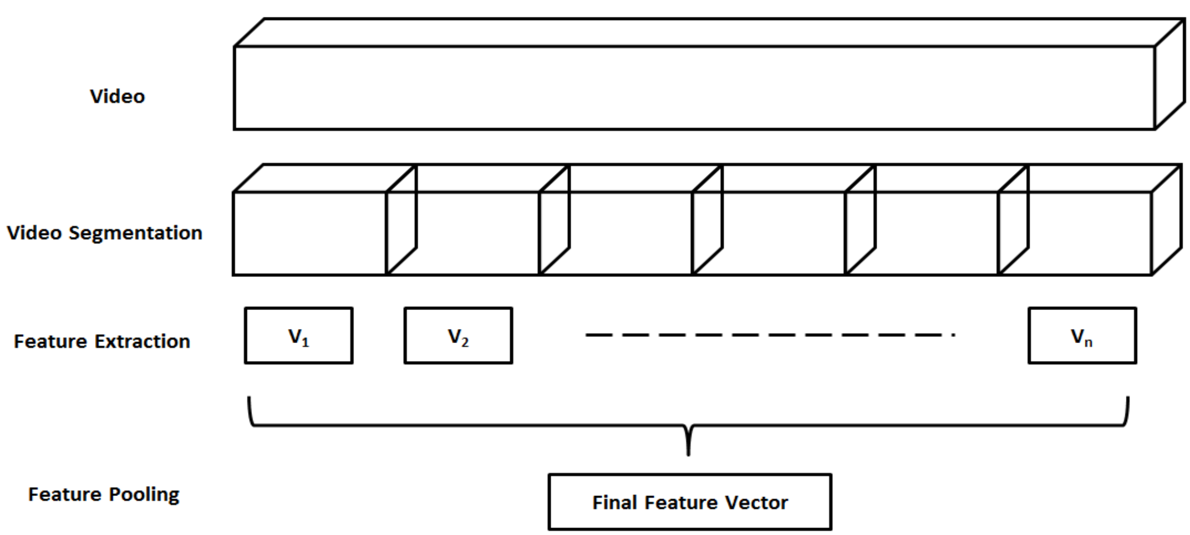


Fig. 7. Illustration of video segmentation method for action representation

This method also has the advantage that the subsequent ranking algorithm does not have to deal with problems arising from length differences. A variant of this fixed length method uses overlapping segments. In this method, a video is divided into overlapping and equal length segments. This approaches can be used that try to identify lexically and semantically coherent segments.

For all used methods we have to determine the length of the segments or the number of segments for a video. For the action recognition task as described above long segments clearly have two disadvantages: longer segments have a higher risk of covering several subtopics and thus give a lower score on each of the included subtopics. In the second place, long segments run the risk that they include the relevant fragment but that the beginning of the segment is nevertheless too far away from the jump-in point that should be found. Short segments, on the other hand, might get high rankings based on just a view words. Furthermore, short segments make the recognizing process more costly. In our approach, we choose different length segments to select the optimal one.

C. Action Represention with Feature Pooling

In video segmentation stage, we divide the video into the set of segments. Each segment is represented by three feature vectors (HOG3D, 3DS-HONV, and HOF2.5D) that are computed by BOW. We use the following temporal aggregation pools feature values for each feature dimension over time as Fig. 7. Pooling features over time means that the temporal structure of action will be modeled. With three descriptors, we have three feature vectors for action representation. Finally, we concatenate them into a final feature vector that presents for action. The vector feature will be provided to classifier to identify the label of action class which performed in video. In this research, the max pooling technique are proposed for aggregating feature vectors.

V. Action Classification

SVM is the most popular discriminative classifier and was proposed by Vladimir Vapnik [39]. It provide the state-of-art performance in many real applications such as text categorization, image classification etc… It is known as the maximum margin classifier. Consider the given training data set $\{(x_i, y_i)\}$ where i = 1, 2,…n and x_i, is N-dimensional feature vector with label y_i = +1 or −1 denoting the class it belongs to. The feature vectors are assumed to be normalized between [−1, 1] or [0, 1] to obviate the undesirable domination of any particular dimension(s) in deciding the decision boundary. SVM strives to find the hyper-plane $w \cdot x - b = 0$ that best separates the training data with regards to the distance from this hyper-plane. The optimal values for w and b can be found by solving a constrained minimization problem, using Lagrange multipliers α_i (i =1,…, n).

$$f(x) = \sum_{i=1}^{n} \alpha_i \, y_i \, K(x_i, x) + b$$

Where α_i and b are found by using an SVC learning algorithm. And $K(x_i, x)$ is a kernel function for the training sample x_i and the test sample x.

The multi-class classification problem is commonly solved by a decomposition to several binary problems for which the standard binary SVM can be used. The one-against-rest decomposition is often applied. In this case, the classification problem to k classes is countered by training k different classifiers, each one trained to distinguish the examples in a single class from the examples in all remaining classes. When it is desired to classify a new example, the k classifiers are run, and the classifier which outputs the largest (most positive) value is chosen. We use non-linear SVM with a RBF kernel which have shown good performance in many researches. In this work, we use LibSVM [7] for SVM classifier implementation. The penalty parameter is set as C = 100.

VI. Exprimental Results

We firstly evaluate the performance of the proposed approach on the three challenging 3D action datasets such as UTKinect-Action, 3D Action Pairs, and MSR-Daily Activity dataset. Then we compare our results to the state-of-the-art methods to demonstrate the superiority of the proposed approach.

Secondly, we evaluate the performance of separation of descriptors and combination of descriptors that are used BOW with k-means and GMM to yield the histogram represents for actions. Furthermore, in order study the effect of the size of the video segmentation on the final classification performance, we choose segment lengths of 10, 15, 20, and 25 frames on non-overlapping and overlapping segmentation. And we use uniform segment sampling with 50% of overlapping. Therefore, the number of segments will be doubled for each overlapping experiment.

A. UTKinect-Action dataset

UTKinect-Action dataset [25] contains 10 different action classes performed by 10 subjects, collected by a stationary Kinect sensor. The 10 action classes are: walk, sit down, stand up, pick up, carry, throw, push, pull, wave hands, clap hands. Each action was collected from 10 different persons for 2 times: 9 males and 1 female. Depth sequences are provided with resolution 320 × 240, and skeleton joint locations are also provided in this dataset. In our experiment, we used the same setting is the leave-one-out scheme in [25].

In this dataset, Table I shows the experimental results of our different methods. From the results one can see that 3D-HONV is the best descriptor in case only one descriptor is used and the fusion of HOG3D, 3DS-HONV and HOF2.5D outperforms the single descriptor in using BOW with k-means and GMM. Fig. 8 presents a comparison of the accuracy of overlapping and non-overlapping segmentation with the difference on the length of segment. The overlapping method is better than the non-overlapping method in all cases. And, the length of 15 frames for each segment achieve the best performance. Table II compares our approach results with state-of-the-art results on UTKinect-Action dataset. We can see that our result of 93.5% in accuracy is better than all previous results using the same settings. Our recognition rate is more than the current best rate by 1.6%.

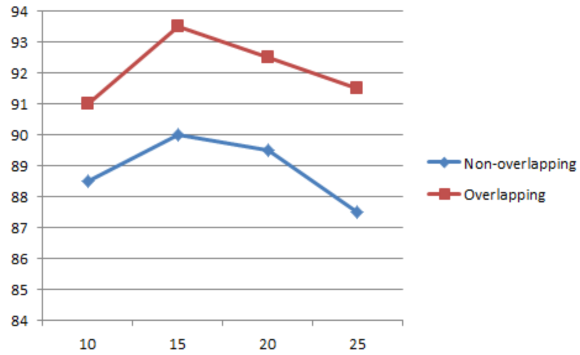


Fig. 8. Experimental results from non-overlapping and overlapping segmentation on UTKinect-Action

TABLE I. EXPERIMENTAL RESULTS OF OUR METHOD ON UTKINECT ACTION DATASET

Methods	**Accuracy (%)**	
	KM-BOW	**GMM-BOW**
HOG3D	83.5	85
3DS-HONV	88	91.5
HOF2.5D	86	87.5
Combined	**91**	**93.5**

TABLE II. COMPARISION OF THE PROPOSED METHOD WITH THE STATE OF THE ART METHODS ON UTKINECT ACTION DATASET

Methods	**Accuracy (%)**
HOJ3D [25]	90.92
STIPS + Joint [46]	91.9
Our approach	**93.5**

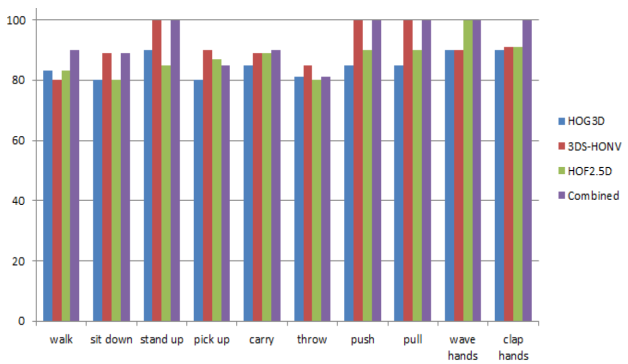


Fig. 9. Comparison of our proposed methods on UTKinect Action dataset

B. 3D Action-Pairs dataset

The 3D Action-Pairs dataset contains activities which are selected in pairs such that the two activities of each pair are similar in motion and shape. For example, "Pick up" and "Put down" actions have similar motion and shape. This dataset has six pairs of activities: "Pick up a box/Put down a box", "Lift a box/Place a box", "Push a chair/Pull a chair", "Wear a hat/Take off a hat", "Put on a backpack/Take off a backpack", and "Stick a poster/Remove a poster". The dataset includes 12 activities performed by 10 different subjects. Each action was performed three times by each subject. We used this dataset in order to emphasize two points: 1) to evaluate the performance of our proposed method in the case of actions that have similar trajectories and objects; 2) to show the advantage of using the feature fusion to enhance the recognition rate.

TABLE III. EXPERIMENTAL RESULTS OF OUR METHOD ON 3D ACTION PAIRS DATASET

Methods	**Accuracy (%)**	
	KM-BOW	**GMM-BOW**
HOG3D	92.22	94
3DS-HONV	92.27	91.38
HOF2.5D	93.61	95.28
Combined	**94.44**	**99.16**

TABLE IV. COMPARISION OF THE PROPOSED METHOD WITH THE STATE OF THE ART METHODS 3D ACTION PAIRS DATASET

Methods	**Accuracy (%)**
Skeleton+LOP [23]	63.33
Depth Motion Maps [43]	66.11
Skeleton + LOP + Pyramid [23]	82.22
HON4D [33]	96.67
SNV [41]	98.89
BHIM [45]	100
Our Approach	**99.16**

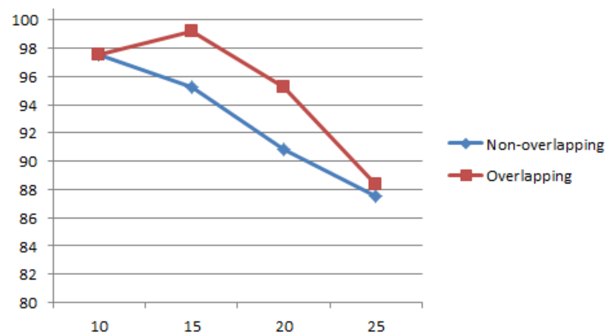


Fig. 10. Experimental results from non-overlapping and overlapping segmentation on 3D Action Pairs

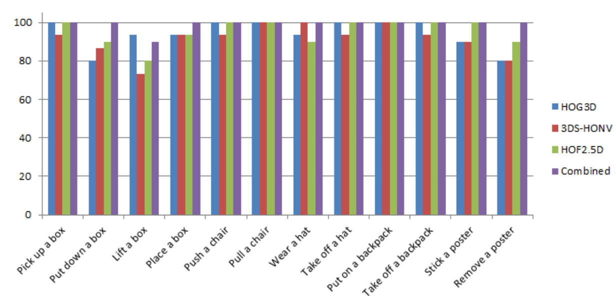


Fig. 11. Comparison of our proposed methods on 3D Action Pairs dataset

In this dataset, Table III shows the experimental results of our different methods. From the results one can see that HOF2.5D is the best descriptor in case only one descriptor is used and the fusion of HOG3D, 3DS-HONV and HOF2.5D outperforms the single descriptor in using BOW with k-means and GMM. Fig. 10 presents a comparison of the accuracy of overlapping and non-overlapping segmentation with the difference on the length of segment. The overlapping method is better than the non-overlapping method in most cases. And, the length of 15 frames for each segment obtain the best performance. Table IV compares our approach results with state-of-the-art results on 3D Action Pairs dataset. We can see that our result of 99.16% in accuracy is better than most previous results using the same settings. Our recognition rate is less than the current best rate is 100% by 0.84%.

C. MSR-Daily Activity 3D dataset

The MSR-Daily Activity 3D dataset contains 16 different human activities: drink, eat, read book, call cell phone, write on a paper, use laptop, vacuum cleaner use, cheer up, sit still, toss paper, play game, lay down on sofa, walk, play guitar, stand-up, sit-down, and each subject performs an activity in two different poses: a standing pose and a sitting on sofa pose. Each pose has 160 total samples, with each subject is one sample per activity in each pose. This dataset is created to cover daily activities and human–object interactions in the living room. These tests are more challenging than the other datasets because of frequent human–object interactions.

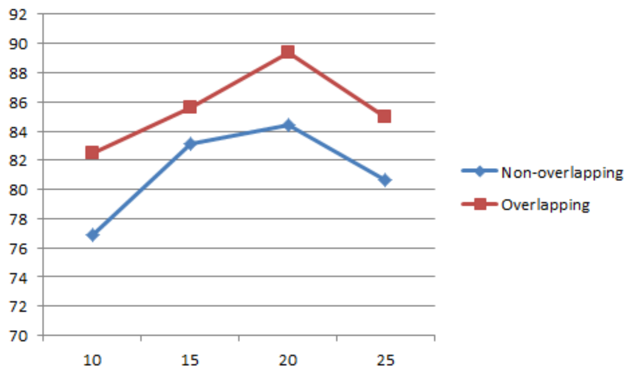


Fig. 12. Experimental results from non-overlapping and overlapping segmentation MSR-Daily Activity

TABLE V. EXPERIMENTAL RESULTS OF OUR METHOD ON MSR-DAILY ACTIVITY DATASET

Methods	Accuracy (%)	
	KM-BOW	**GMM-BOW**
HOG3D	78.75	80.63
3DS-HONV	84.36	87.5
HOF2.5D	81.25	82.5
Combined	**86.25**	**89.38**

TABLE VI. COMPARISION OF THE PROPOSED METHOD WITH THE STATE OF THE ART METHODS ON MSR-DAILY ACTIVITY DATASET

Methods	Accuracy (%)
LOP [23]	42.50
Depth Motion Maps [43]	43.13
Local HON4D [33]	80.00
Actionlet Ensemble [23]	85.75
SNV [41]	86.25
BHIM [45]	86.88
Our approach	**89.38**

In this dataset, Table V shows the experimental results of our different methods. From the results one can see that 3DS-HONV is the best descriptor in case only one descriptor is used and the fusion of HOG3D, 3DS-HONV and HOF2.5D outperforms the single descriptor in using BOW with k-means and GMM. Fig. 12 presents a comparison of the accuracy of overlapping and non-overlapping segmentation with the difference on the length of segment. The overlapping method is better than the non-overlapping method in all cases. And, the length of 20 frames for each segment achieve the best performance. Table VI compares our approach results with state-of-the-art results on MSR-Daily Activity dataset. We can see that our result of 89.38% in accuracy is better than all previous results using the same settings. Our recognition rate is higher than the current best rate is 86.88% by 2.5%.

VII. CONCLUSION

In this work, we present a new framework for action recognition in RGB-D video based on spatiotemporal features and segmentation technique. We use STIP detector to select interest points for both RGB and depth channels. Spatiotemporal descriptors consist of HOG3D, 3DS-HONV and HOF2.5D are extracted. These descriptors capture shape, appearance and motion information which are vital properties for action representation. We use GMM instead of k-means in BOW model to create more distinctive for action representation. Also, we apply segmentation and max pooling technique to capture the temporal structure of action. Our approach systematically is evaluated on several benchmark datasets such as UTKinect-Action, 3D Action Pairs, and MSR-Daily Activity 3D dataset with final recognition accuracies of 93.5%, 99.16% and 89.38% for fusion of descriptors, respectively. The experimental results have shown outcome performance compare to the-state-of-art methods in overall in most cases. For the spatiotemporal descriptors, 3DS-HONV has shown robust descriptor in most cases. However, HOG3D is better than 3DS-HONV in case that needs to distinguish these objects that have the similar shape as 3D Action Pairs dataset. And, HOF2.5D is better than HOG3D and 3DS-HONV in case that needs to differentiate these actions that have the similar motion. Thus, to improve the action recognition system, fusion of the descriptors is the best way. For the part-based model, from experimental results also show that GMM is more powerful than k-means when using to create visual words in BOW model. For segmentation method, in addition, we indicate that overlapping method performs the best in most cases. And, the length of segment also impacts to the performance of the system. However, the length is not fixed for all the dataset that it depends on the descriptors are used and the nature of dataset. In this work, the experimental results indicate that the length of segment is 15 and 20 frames are the best performances.

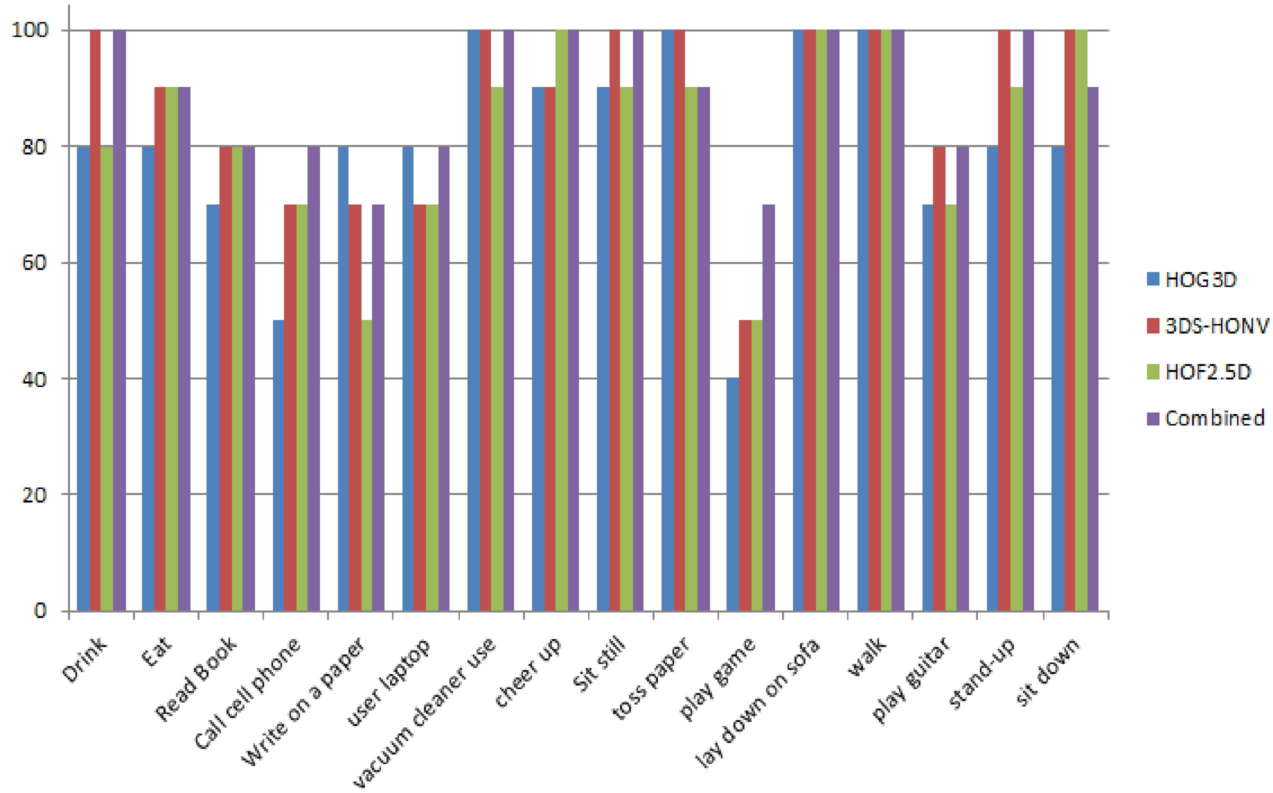


Fig. 13. Comparison of our proposed methods on MRS-Daily Activity dataset

In summary, the key problems of this research are summed up as follows: firstly, we have explored the utility of spatiotemporal features derived from RGB and depth information. These features are extracted to capture both shape and motion in action. Secondly, GMM used to instead of k-means in BOW model to have more distinctive and descriptive for action representation. Finally, we have modeled temporal structure of action based on video segmentation and max pooling technique.

In the future, we will investigate new method to improve appearance, motion properties as well as consider the impact of context and evolution of human when performing the action.

ACKNOWLEDGMENT

This research is funded by Vietnam National University, Ho Chi Minh City (VNU-HCM) under grant number *B2014-18-02*.

REFERENCES

[1] A. Kläser, M. Marszałek, and C. Schmid. A spatio-temporal descriptor based on 3D- gradients. In BMVC , 2008.

[2] A. P. Dempster, N. M. Larid, and D. B. Rubin, Maximum likelihood from incomplete data via the em algorithm, Journal of the Royal Statistical Society Series B(Methodological), vol. 39(1), pp. 1–38, 1977.

[3] A. W. Vieira, E. R. Nascimento, G. L. Oliveira, Z. Liu, and M. F. M. Campos, "Stop: Space-time occupancy patterns for 3d action recognition from depth map sequences," in CIARP, 2012, pp. 252–259.

[4] B. Ni, G. Wang, and P. Moulin, "Rgbd-hudaact: A color-depth video database for human daily activity recognition," in ICCV, 2011.

[5] Bay, H., Tuytelaars, T., and Van Gool, L, "SURF:Speeded Up Robust Features", In Proceedings of the Ninth European Conference on Computer Vision, May, 2006.

[6] Bobick, A. and Davis, J.: The Recognition of Human Movement Using Temporal Templates. IEEE Trans. On Pattern Analysis and Machine Intelligence, 2001.

[7] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. ACM TIST, 2(27):1–27, 2011.

[8] C. Harris and M.J. Stephens. A combined corner and edge detector. In Alvey Vision Conference, 1988.

[9] C. Schuldt, I. Laptev, and B. Caputo. Recognizing human actions: A local SVM approach. In ICPR , 2004.

[10] D. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91–110, 2004.

[11] Daniel Weinland, Remi Ronfard, Edmond Boyer, A Survey of Vision-Based Methods for Action Representation, Segmentation and Recognition, INRIA, 2010.

[12] G. Farnebäck, "Two-frame motion estimation based on polynomial expansion," in Proc. 13th Scandin. Conf. Image Anal. (SCIA), 2003, pp. 363–370.

[13] G. Willems, T. Tuytelaars, and L. Van Gool. An efficient dense and scale-invariant spatio-temporal interest point detector. In ECCV, 2008.

[14] Hao Zhang, Wenjun Zhou, Christopher Reardon, Lynne E. Parker Simplex-Based 3D Spatio-Temporal Feature Description for Action Recognition, CPVR2014.

[15] H. Wang, A. Kl ¨ aser, C. Schmid, and C.-L. Liu: Dense trajectories and motion boundary descriptors for action recognition. International Journal of Computer Vision, Mar. 2013.

[16] Heng Wang and Cordelia Schmid, Action Recognition with Improved Trajectories. ICCV 2013.

[17] I. Laptev, M. Marszałek, C. Schmid, and B. Rozenfeld. Learning realistic human actions from movies. In CVPR, 2008.

[18] I. Laptev and T. Lindeberg. Space-time interest points. In ICCV, 2003.

[19] J. Aggarwal and M. Ryoo, "Human activity analysis: A review," ACM Computing Surveys, Apr. 2011.

[20] J. Aggarwal and Q. Cai. Human motion analysis: a review. In Nonrigid and Articulated Motion Workshop, pages 90–102. IEEE, 1997.

[21] J. Diebel, "Representing attitude: Euler angles, unit quaternions, and rotation vectors," 2006.

[22] J. Wang, Z. Liu, J. Chorowski, Z. Chen, , and Y. Wu. Robust 3d action recognition with random occupancy patterns. In ECCV, 2012.

[23] J. Wang, Z. Liu, Y. Wu, J. Yuan, Mining actionlet ensemble for action recognition with depth cameras, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2012, pp. 1290–1297.

[24] L. Liu, L. Shao, Learning discriminative representations from RGB-D video data, in: Proceedings of International Joint Conference on Artificial Intelligence (IJCAI), 2013.

[25] L. Xia, C. Chen, and J. Aggarwal. View invariant human action recognition using histograms of 3d joints. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, pages 20– 27. IEEE, 2012.

[26] L. Yeffet and L. Wolf. Local trinary patterns for human action recognition. In ICCV, 2009.

[27] Leandro Cruz, Djalma Lucio, Luiz Velho: Kinect and RGBD Images: Challenges and Applications. SIBGRAPI Tutorials, pp 36-49, 2012.

[28] Li, W., Zhang, Z., and Liu, Z. Action Recognition based on A Bag of 3D Points. IEEE Workshop on CVPR for Human Communicative Behavior Analysis, 2010.

[29] Lu Xia and J.K. Aggarwal, Spatio-Temporal Depth Cuboid Similarity Feature for Activity Recognition Using Depth Camera, CVPR 2013.

[30] M. Camplani and L. Salgado, Efficient spatio-temporal hole filling strategy for kinect depth maps, A. M. Baskurt and R. Sitnik, Eds., vol. 8290, no. 1. SPIE, 2012, p. 82900E.

[31] Mao Ye, Qing Zhang, Liang Wang, Jiejie Zhu, Ruigang Yang, Juergen Gall, A Survey on Human Motion Analysis from Depth Data. Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications Lecture Notes in Computer Science Volume 8200, 2013, pp 149-187

[32] Navneet Dalal, Bill Triggs, Histograms of Oriented Gradients for Human Detection, CVPR, 2005.

[33] O. Oreifej and Z. Liu, "Hon4d: Histogram of oriented 4d normals for activity recognition from depth sequences," in CVPR, 2013.

[34] P. Scovanner, S. Ali, and M. Shah. A 3-dimensional SIFT descriptor and its application to action recognition. In MULTIMEDIA, 2007.

[35] Quang D. Tran, Ngoc Q. Ly, Sparse Spatio-Temporal Representation of Joint Shape-Motion Cues for Human Action Recognition in Depth Sequences, 2013 IEEE RIVF International Conference on Computing & Communication Technologies -Research, Innovation, and Vision for the Future (RIVF), 2013.

[36] Quang D. Tran, Ngoc Q. Ly, An Effective Fusion Scheme of Spatio-Temporal Features for Human Action Recognition in RGB-D Video, IEEE-International Conference on Control, Automation and Information Sciences (ICCAIS), 2013.

[37] Ronald Poppe, A survey on vision-based human action recognition, Image and Vision Computing 28, 976–990, 2010.

[38] S. Tang, X. Wang, X. Lv, T. X. Han, J. M. Keller, Z. He, M. Skubic, and S. Lao, "Histogram of oriented normal vectors for object recognition with a depth sensor," in ACCV, 2012.

[39] V.Vapnik, "Statistical learning theory", *John Wiley and Sons*, New York, 1998.

[40] Viet Vo, Ngoc Ly, Robust human action recognition using improved BOW and hybrid features, 2012 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2012

[41] Xiaodong Yang, YingLi Tian. Super Normal Vector for Activity Recognition Using Depth Sequences, CVPR, 2014.

[42] Xinghua Sun, Mingyu Chen, Alexander Hauptmann,Action Recognition via Local Descriptor and holistic features, Computer Vision and Pattern RecognitionWorkshop, IEEE, 2009.

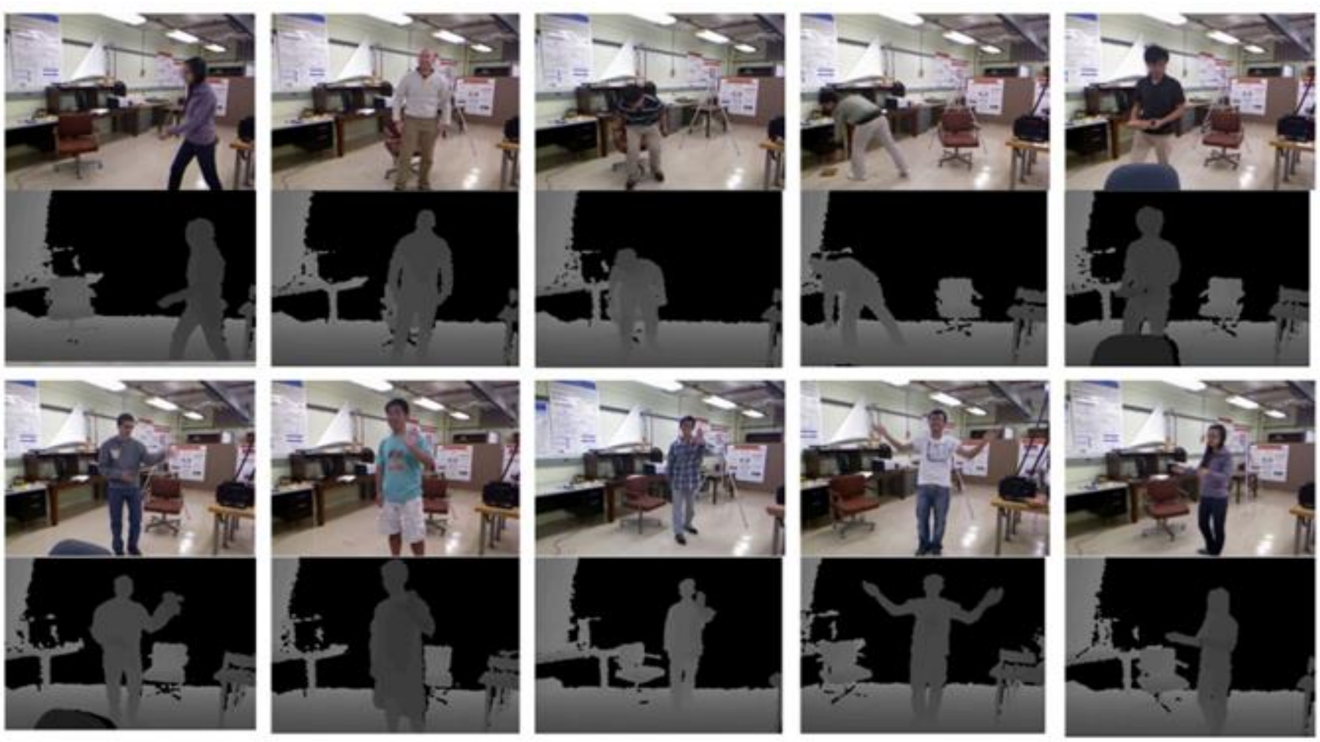
[43] Yang, X., Zhang, C., Tian, Y.: Recognizing actions using depth motion maps based histograms of oriented gradients. In: ACM International Conference on Multimedia. (2012) 1057-1060.

[44] Yen-Yu Lin, Ju-Hsuan Hua, Nick C. Tang, Min-Hung Chen, Hong-Yuan Mark Liao, Depth and Skeleton Associated Action Recognition without Online Accessible RGB-D Cameras, CVPR 2014.

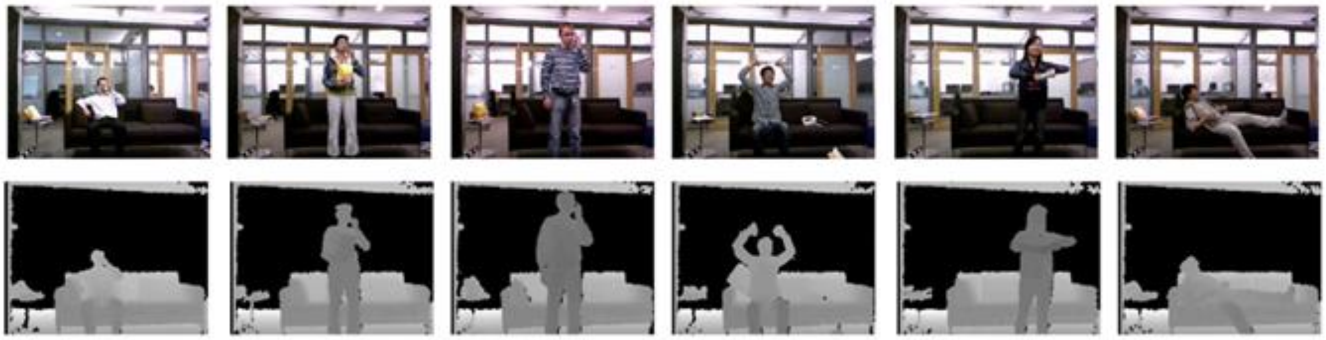
[45] Yu Kong, Yun Fu, Bilinear Heterogeneous Information Machine for RGB-D Action Recognition, CVPR 2015.

[46] Yu Zhu, Wenbin Chen, and Guodong Guo, Fusing Spatiotemporal Features and Joints for 3D Action Recognition, 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops.

[47] Zhao Yang, Liu Zicheng, Cheng Hong. RGB-Depth Feature for 3D Human Activity Recognition, Communications, China, 2013.



(a)



(b)

Fig. 14. Some sample frames from UTKinect-Action (a) and MSR-Daily Activity 3D (b) dataset

Using Business Intelligence Tools for Predictive Analytics in Healthcare System

Mihaela-Laura IVAN
Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

Manole VELICANU
Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

Mircea Răducu TRIFU
Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

Cristian CIUREA
Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

***Abstract*—The scope of this article is to highlight how healthcare analytics can be improved using Business Intelligence tools. Healthcare system has learned from the previous lessons the necessity of using healthcare analytics for improving patient care, hospital administration, population growth and many others aspects. Business Intelligence solutions applied for the current analysis demonstrate the benefits brought by the new tools, such as SAP HANA, SAP Lumira, and SAP Predictive Analytics. In detailed is analyzed the birth rate with the contribution of different factors to the world.**

***Keywords*—Healthcare Analytics; Business Intelligence tools; SAP HANA; SAP Lumira; SAP Predictive Analytics; Birth Rate; Big Data**

I. Introduction

According to a 2016 survey by the World Economic Forum [1], today's society is experiencing the fourth industrial revolution. A phenomenon that led to significant changes in the labor market by creating new jobs, while in a few years ago it was not possible or were unimaginable. The evolution of new technologies such as Business Intelligence, Big Data, Cloud Computing, Mobile Programming, Social Networks, Cyber-security and others have radically changed the way people communicate and collaborate. Educational programs offered by academic institutions must adapt to this technological revolution so that graduates are prepared for the new requirements imposed by the labor market, economic and social change. Areas such as healthcare, accounting, marketing, management, tourism, and many others have another perspective in this context because all were or are being influenced by IT&C technologies. Online platforms have emerged by which people perform medical appointments, electronic payments, reserves tickets, holidays and vacations reserves and many others. These operations in the past were taking place in bank branches or travel agencies.

Business Intelligence is very timely, being promoted both by large technology companies producing technology and professional associations fostering scientific research, such as ACM (Association for Computing Machinery) and IEEE (Institute of Electrical and Electronics Engineers). On 16 March 2016, ACM organized a Webinar on "From BI to Big Data - Architecture, Ethics and Economics", which shows that the domain is very relevant at this time.

The creative use of Big Data, especially for Internet of Things (IoT) transforms business models by supporting start-ups and strengthening or destruction of existing business values. Many supporters of technologies referred focus to use Big Data in the marketing area, but the greatest value will come from daily using and from new ways of transformation. To make appropriate technological and architectural choices is vital in designing and managing automated environments that Big Data should have and that the Internet of Things will ask. However, the real deal requires a close look at the ethical and economic aspects that Big Data poses. They concern personal privacy, employment and social disruption which must all be addressed urgently if the individual businesses and society support successful navigation in this transformation based on data for all aspects of business and technology [2] [3].

II. Business Intelligence Tools for Analytics

The technologies analyzed are implemented in SAP HANA and above this SAP HANA In-Memory database is deployed SAP Lumira or SAP Predictive Analytics. For this research, SAP Lumira [4] is used for illustration of factors affecting the birth rate in a country and SAP Predictive Analytics to see the contributions of the analyzed inputs into birth rate output with classification model, where the factors are classified depending on significance and contribution. The heart of HANA application development contains data and views. The main feature of SAP HANA is the ability to run different information views. The three possible information views that can be designed with the help of SAP HANA are presented in Table 1 [5].

TABLE I. Comparative Analysis Attribute Views/Analytic Views/Calculation Views

Attribute Views	Analytic Views	Calculation Views
- Used to model an entity that is based on joins between various	- Many have two types of columns: attributes and measures;	- Calculation Views provide the combination of tables

source tables and then to choose for output the relevant table's columns and rows; - Could model descriptive attribute data by using only attributes; - This type of view use the join Engine;	- Can contain attribute views inside of them to achieve more depth of attribute data; - Use the OLAP Engine in general, but if the view contains calculated attributes then it is used the Calculation Engine;	with other views like attribute views and analytic views; - The nodes that can be created with calculation views are union, join, projection, aggregation, and rank; - Two types of calculation view can be created: script calculation view using SQLScript and graphical calculation view;

Data sources for SAP Lumira can be Excel, SAP HANA Analytical Views, Software Development Kits (SDK) and many other data sources. SAP Lumira can connect data no matter where the data lives.

SAP Lumira's strength is the possibility to visualize data in different friendly ways like interactive maps, beautiful charts, and infographics. With the help of SAP Lumira, BI analysis with intuitive dashboards, and securely share insights and data stories can be shared with decision makers using SAP Lumira server and cloud platforms with browser and mobile-based experiences to further analyze data and collaborate with colleagues on datasets, stories, and other business intelligence artifacts.

One of the SAP HANA analytical views used to highlight the population analytics is presented below in Figure 1 [6].

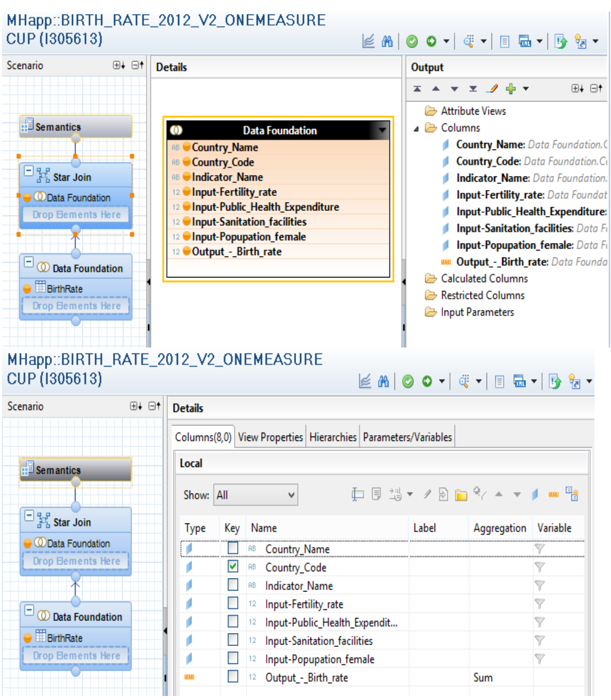


Fig. 1. SAP HANA Analytical view (used in SAP Lumira)

As can be seen, the table BirthRate was added into Data Foundation and all the columns were added for the output.

After that a new dataset is created in Lumira to connect to the presented analytical view. SAP Lumira is online connected to SAP HANA and the download of data from SAP HANA is made offline. The dataset have the columns from analytical view's output separated in measures and dimensions as can be shown in Figure 2.

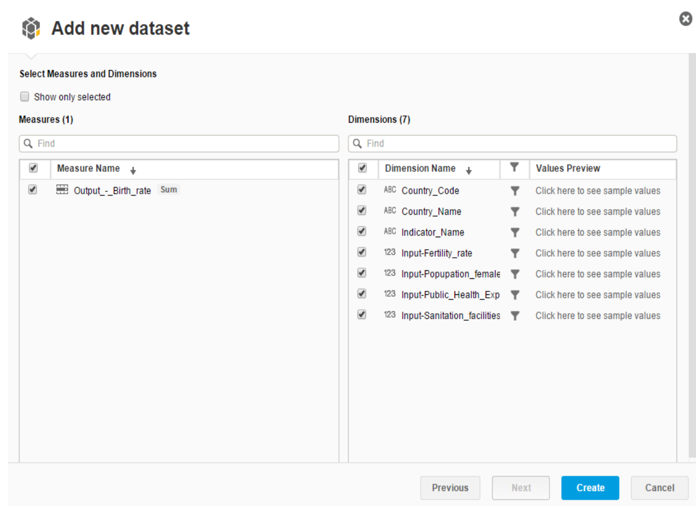


Fig. 2. SAP Lumira dataset of a SAP HANA analytical view

SAP Lumira datasets can be shared, exported or printed to collaborate with colleagues on the same datasets. The dataset sharing methods are the following:

- ➢ Save the dataset as .csv file or xls file;
- ➢ Publish the dataset to SAP HANA to create a new analytic view;
- ➢ Publish to SAP BusinessObjects Explorer. In this way the dataset can be used as an Information Space data source;
- ➢ Publish a dataset to SAP Lumira Cloud, to save documents and work together with colleagues on datasets;
- ➢ Publish a dataset to SAP BusinessObjects Business Intelligence platform;

III. Birth Rate Analytics with SAP HANA and SAP Lumira

The definition of Birth Rate indicator is the number average number of births for every 1000 people in a country. The birth rate indicator, BR, is calculated based on the formula below:

$$BR = \frac{NLB}{TP} x1000 \qquad (1)$$

where:

NLB – the number of live births

TP – the total of population

The analysis is of 207 countries with data related to birth rates for 2012 year.

The influences factors for birth rates presented in this paper are the following [7]:

- Fertility rate – is the number of children that could be born by a woman if she were to live to the end of her childbearing years and bear children in correspondence with her current age-specific fertility rates;
- Public health expenditure – is the recurrent and capital spending from government budgets, different borrowings and grants and social health insurance funds;
- Sanitation facilities – is the possibility to access the improved sanitation facilities. It is referred as the percentage of the population using improved sanitation facilities;
- Female population - is the proportion of the population that is woman;

In reality, there are more external factors affecting population depending on each country and region like the following factors, which will be presented on future researches [8]:

- Capital income - is income that comes from capital for a country;
- Age-sex structure - the distribution of the population by sex and age;
- Religious beliefs and social beliefs regarding contraception and abortion;
- Economic prosperity - in terms of the correlation between the economy growth and the willing of families to have more children;
- Female employment - Employment rate of women referred to percentage of female population;
- Poverty levels - babies are seen as a method for developing countries because they will become to earn money;
- Infant Mortality Rate - depending on the country's IMR, if it is high then some of the children will die;
- Typical age of marriage- is the age at first marriage [9] [10];

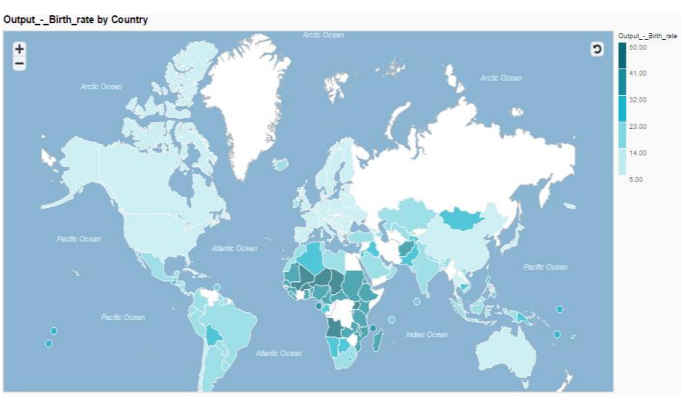


Fig. 3. Birth rate by country

The above image, Figure 3, shows the birth rate by country for 2012 year. This map was created with SAP Lumira where the global natality status is emphasized by color. As a first impression can be seen that Africa has the biggest birth rate, with the following three top countries: Niger (49.87), Angola (46.50), Mali (44.75). While the countries with the lowest level of births are Japan (8.20), Germany (8.40) and Portugal (8.50). Romania is part of the countries with low level of birth rate, the value is 8.80, which almost as small as the first range value.

In the following section are analyzed how the selected inputs affect the number of births. Figure 4 and Figure 5 present two tables for comparative analysis on how fertility rate indicator, female population indicator, public health expenditure indicator, sanitation facilities indicator make birth rate to increase or to decrease. One table contains the countries having the biggest birth rates and another one the countries having lowest birth rates. These two tables were obtained with SAP Lumira's ranking functions, top three and bottom three on a crosstab control as displayed in Figure 4 and 5. A crosstab allows complex multidimensional analysis which is useful to view the exact values or to examine data from multiple measures.

Top 3 Output_-_Birth_rate by Country_Name, Input - Fertility Rate, Input-Popupation_female, Input-Public_Health_Expenditure, Input-Sanitation_fac

					Measures
Country_Name	Input - Fertility Rate	Input-Popupation_female	Input-Public_Health_Expenditure	Input-Sanitation_facilities	Output_-_Birth_rate
Niger	7.64	49.63	33.06	10.10	49.87
Angola	6.25	50.42	62.16	49.10	46.50
Mali	6.40	49.56	38.83	23.30	44.75

Fig. 4. Top three Birth rate by fertility rate, population female, public health expenditure and sanitation facilities

Let us take a look on how the number of births occurring during one year, per 1000 population estimated at mid-year is correlated with the inputs for Niger country. We want to highlight the factors with the biggest influence and the reasons for this big birth rate.

Fertility rate that is calculated as $\frac{number\ of\ babies}{woman}$ has the greatest value of 7.64. For this country fertility rate is strongly correlated with birth rate.

Female population - proportion of population that is female, with 49.63% value. This proportion shows that the population is uniform divided and this argues the high value of birth rate.

Public health expenditure - percentage of total health costs is 33.06%, which means that the total expenditure on health by Niger is not so high and surprising this factor does not affect the number of births.

Sanitation facilities - proportion of population with access to sanitation, to safe water and those who practice good hygiene represent only 10.10% [11].

As a conclusion of the above analysis, the output value (49.87) of birth rate is strong influenced by fertility rate, female population. While the two other inputs, public health expenditure and sanitation facilities, does not affect the number of births. Although, the factors like religion and poverty are also important factors for birth rate, but these will be considered in future researches.

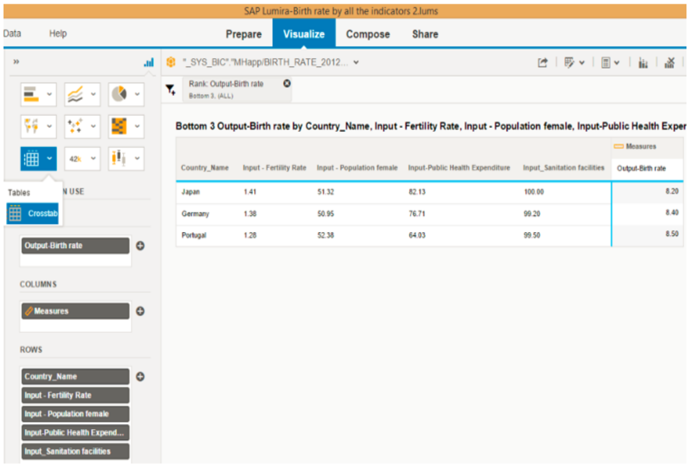


Fig. 5. Bottom three Birth rate by fertility rate, population female, public health expenditure and sanitation facilities

As a comparison to Niger, Japan has the littlest birth rate value, 8.20. The input factors for this output result are the following:

Fertility rate, the average number of children that a woman may give birth in her lifetime, with value 1.41 is also strongly correlated with the low value of birth rate.

Population female is 51.32 %, which is inversely correlated with birth rate and is not so relevant for the output.

Public health expenditure with the rate of 82.13%, which is a high value. Even if a big part of government budget is granted for the health sector, this endeavor is not enough to help the number of births to growth.

Sanitation facilities with the rate of 100.00% emphasize that all the population has gained access to improved sanitation facilities.

Further, in Figure 6, it is calculated with classification model of SAP Predictive Analytics the variables contribution into birth rate output.

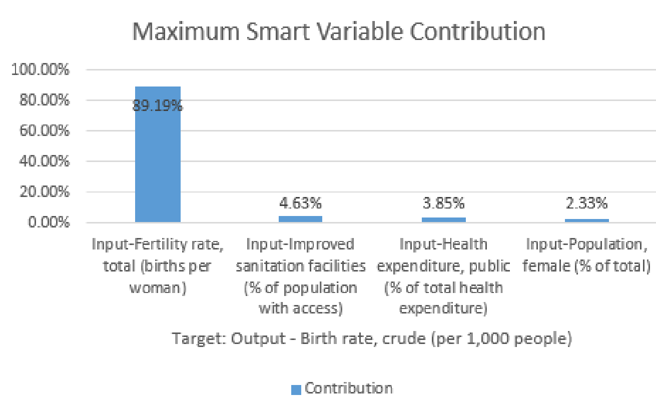


Fig. 6. Inputs contribution into birth rate output

This chart confirms the above assumptions. Also, the figures demonstrate that government efforts to help the low birth rate have not growth the desired effects. In reality, the government intention will not bring immediate fruit to straighten the nations demographic. That is why policymakers have to sustain the needed steps without being affected by down figures.

The trust of the model is calculated with Predictive Power indicator, abbreviated as KI, and Prediction Confidence indicator, abbreviated as KR. Prediction Confidence refers to the ability of the model to achieve the same performance on a new set of data. Models are power if the value is >= 0.95. While, Predictive Power, describes the percentage of information which exists in the target variable. Is good when the Predictive Power value is higher. That model with the highest sum of Predictive Power and Prediction Confidence will be chosen because of the robustness, accuracy and coherency it has. The range of possible values for these two predictions indicators is between 0 and 1. KI and KR values of classification model can be seen in Figure 7.

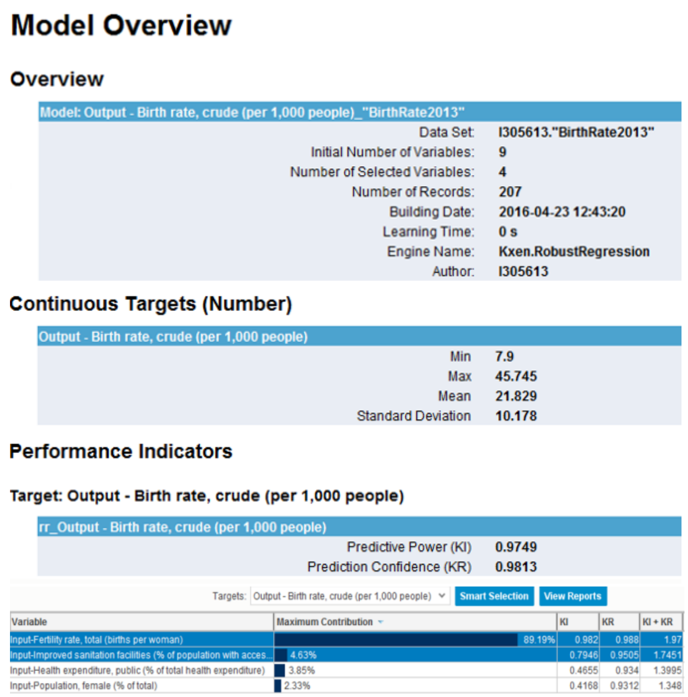


Fig. 7. Performance indicators

Also, valuable information from the model overview is regarding Min, Max, Mean, and Standard Deviation which extract relevant targets from the data set, very useful for analytics and decision makers.

IV. Conclusions

With the help of the new tools like SAP HANA, SAP Lumira and SAP Predictive Analytics data can be easily aggregated to make comparative or predictive analysis useful for important domains like it is healthcare field. In other words SAP Lumira 'predict the unpredictable' with the capability to load data from all sources, present valid insights and as a result to help on better decisions.

The results presented show the existence of the relationship between fertility rate in principal, female population, public health expenditure, sanitation facilities and birth rate. This relationship between four inputs and one output behaves differently among the countries due to geographical aspects, labor market characteristics and their welfare state model.

From our point of view, the humanity is in a change process. The lifestyle of people is tending to be robotized, even the process of making babies can in our time be helped mechanically with 'In Vitro Fertilization' process which brought five million children in 35 years [12].

REFERENCES

[1] World Economic Forum, The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution, January 2016, Available at: http://www3.weforum.org/docs/WEF_FOJ_Executive_Summary_Jobs.pdf

[2] B. Nedelcu, "Business Intelligence Systems", *Database Systems Journal*, vol. IV, no. 4/2013, pp 12-20.

[3] V. Farrokhi and L. Pokorádi, "The necessities for building a model to evaluate Business Intelligence projects- Literature Review", *International Journal of Computer Science & Engineering Survey (IJCSES)*, Vol.3, No.2, April 2012, pp. 1-10.

[4] D. Crockett, "Defining Predictive Analytics in Healthcare" [Online], 2013, Available: https://www.healthcatalyst.com/predictive-analytics

[5] ***, SAP SE or an SAP affiliate company, SAP HANA Modeling Guide, 2014.

[6] M. L. Ivan, "Improving Business Intelligence Applications by Using New Generation of Web and Mobile Technologies," *Informatica Economică Journal*, Vol. 19, No. 4, 2015, pp. 81-89.

[7] ***, World Bank [Online], Available: http://data.worldbank.org

[8] E. Jaba, I. A. Chirianu, C. B. Balan, I. B. Robu, M. D. Roman, "The Analysis of the Effect of Women's Participation in the Labor Market on Fertility in European Union Countries Using Welfare State Models", *Journal of Economic Computation and Economic Cybernetics Studies and Research*, No. 1 2016, Vol 50, pp. 69-84.

[9] ***, PAPP – Glossary, http://papp.iussp.org/sessions/glossary/glossary.html#population_pyramids

[10] ***, OECD library, http://www.oecd-ilibrary.org/employment/employment-rate-of-women_20752342-table5

[11] ***, International decade for action 'Water for Life' 2005-2015 http://www.un.org/waterforlifedecade/sanitation.shtml

[12] ***, The amazing story of IVF: 35 years and five million babies later [Online], Available at: http://www.theguardian.com/society/2013/jul/12/story-ivf-five-million-babies.

NISHA: Novel Interface for Smart Home Applications for Arabic Regionsubtitle as Needed

Muneer Bani Yassein*
Computer Science department
Jordan University of Science and Technology
Irbid - Jordan

Yaser Khamayseh, Maryan Yatim
Computer Science Department
Jordan University of Science and Technology
Irbid - Jordan

***Abstract*—Researchers have developed many devices and applications for smart homes to control home's appliances. The main goal of this research is to propose a touch-based interface (namely, NISHA) for smart homes to meet user needs and requirements and is able to control any appliance in the house. This study is designed for people and circumstances in the Middle East countries (Jordan and West Bank) and therefore, is set out to design a user interface for smart home applications taking into consideration the economic, social, and technological differences. Referring to those differences, NISHA was designed in a classical representational design instead of a modern advanced one, based on virtual images instead of text, full control instead of automatic control, and very restrictive privacy issues for people of these countries still look at smart homes as a technology that threaten their privacy. Moreover, NISHA was tested and evaluated using heuristic and cognitive walk-through evaluation techniques. Evaluation results showed that 80% of users and experts were satisfied with NISHA as a user friendly interface, 90% of users were satisfied that NISHA met their expectations, and finally, 93% of users strongly asked to have NISHA in their daily lives.**

Keywords—Human Computer Interaction (HCI); HCI Design and evaluation methods; User Interface Design; User Centered Design; Smart Homes

I. INTRODUCTION

People used to interact, control, and monitor smart homes using web based interfaces, however touch-enabled interfaces overshadowed the web based ones in the last few years, and the challenges remain on how to design better interfaces for controlling the appliances of smart homes. Many user interfaces nowadays are poorly usable. This research is concerned with the improvement of user interface design as a final product. There are plenty of methods on how to design a user interface, but only few of them concentrate on the overall design process. In this work, the design process is covered entirely from early user analysis, such as questionnaires and interviews to implementing or prototyping the actual product. Some limitations such as money, time, and people are taken into consideration while designing the interface. Although there are many HCI rules that should be included in any user interface design like safety, reliability, accessibility, consistency, etc., functionality and usability remain the most important terms that should be considered when implementing, designing, or evaluating any HCI application [18]. The functionality of any system defines the set of services this system can afford. Usability of the system defines how much the user can use the system efficiently and achieve the needed goals of the developed system. However, a system is said to be working effectively if there is a good balance between functionality and usability [5].

To achieve the goals of this study, a plan was first drawn to the design of the interface and methods of achieving the required goals were chosen. First of all, an initial closed ended questionnaire and open ended questions of interviews were performed for a test group of 41 users and 5 experts, and were performed to highlight the vital elements for users' and their goals. Also, expectations of users for extra services and features are identified. A list of interface requirements that should meet the users' needs efficiently was drawn, based not only on HCI rules, but also on the statistics and features analysis of questionnaires and feedback of the test group. Moreover, two evaluation techniques were used for the evaluation of NISHA; heuristic evaluation and cognitive walk-through. Finally, a prototype was built and connected to the interface so to have a full smart home system that can run in real life scenarios.

A. Interacting with Computing

Human Computer Interaction (HCI) is defined as a study field of interaction between people and devices; how they use, implement, design, affect, and be affected by computer systems [13]. In the past, developers' main concern was to program a code that works, and they never cared much about users and their needs. Nowadays, with the fast development of smart devices and with the variety of products in the market, the main concern remained on how to deliver a product that satisfies users' needs and requirements. Therefore "User – Centered Design UCD is an approach that is centered on determining the context of users and their requirements" [7].

Engineers and designers should have background knowledge of HCI rules so to design better interfaces that can achieve better user satisfaction. This knowledge can be achieved either by literature or by the designers' experience of interface design. Understanding what should be happening when a user is confronted with the interface is not an easy thing. Users want to achieve their goals easily, quickly, and their way. Therefore, designers should have experience and skills that would enhance their final design of the interface. However, it is not enough for designers to rely only on literature as it has limited knowledge. For a better User Interface design, personal experiences and explicit knowledge are essential. Design patterns can also be used as a way to capture interface design knowledge. These patterns contain a number of generalized solutions to certain problems and designers can use them in practice.

Smart home interfaces, industrial showcases and laboratories are spread all over the world and are available to all users; although some of them may be more optimized than others, there are still some similarities that are shared among these interfaces and smart home systems. User interfaces are used for monitoring and controlling these smart homes, while the decision making of tasks is managed using artificial intelligent middleware software. Regardless of the working team of the project, the aim of smart home project and smart home interfaces is to have a better system functionality, usability, and testing.

The rest of the paper is organized as follows: Section II presnts the evaluation method. Section III presnts the proposed system timeline. Results are presented and discussed in Section IV. Section V and VI concludeds the paper and presnts some future research directions.

II. NISHA EVALUATION METHODS

During the evaluation of NISHA, two types of evaluation were followed; heuristic evaluation and cognitive walk-through. A test group of 5 experts was involved in NISHA heuristic evaluation; each expert tested NISHA separately than 3 of the experts met together and discussed the usability problems. The reason that not all experts met is that 3 of them are in in Jordan and the others are in the West Bank. However, we didn't take the formal method of collecting reports from users, instead we highlighted their comments and recommendations.

Using a cognitive walk-through, we built a wooden prototype with two lights and one motor and connected these appliances with electricity to consider this prototype as the smart home. Then we connected NISHA interface with the prototype and prepared three real life scenarios for a test group of 5 experts to test. The scenarios included basic tasks and special tasks and are described in details in section 4.3. Moreover, the results of this method are shown in Tables 2 and 3.

Furthermore, five interviews were made by experts in engineering and development field, the interviewees were asked to highlight the desired design and characteristics of such an interface that will make it meet the HCI rules and user needs.

A. Iterative Strategy

In this work, three iterations of UCD phases are processed to have better and approximate results according to users. The reason upon why we chose to follow an iterative strategy of UCD phases is that many features and results are not clear from the first process, and only become obvious when first phase results are analyzed. For example, some features might not become clear and certain until users have the prototype to evaluate and revise their needs.

The first iteration of NISHA revealed that it has some weaknesses; experts complained that there is no administration page for administrators to add/remove important buttons, and since any house can have more or less rooms, administrators should be able to control these issues and should have a control panel for administration use only. Moreover, the experts recommended to add a "Weather button" in the front page of NISHA so to make it easier for users to check weather forecasts. Later on, these complaints and recommendations were taken into considerations and NISHA was enhanced and entered the second iteration of its design process. In the second iteration, "Weather" button and administration page were added, the interface was built with a touch enabled device, and evaluation of it has been made using the two evaluation techniques; heuristic evaluation and cognitive walk-through. The results of NISHA evaluation in the second iteration show that it does not include a turn-off-ALL button on all pages but only on the front page of the interface, a weakness that took users much more time to complete the basic tasks. Moreover, the "Back" button on each page of NISHA was placed on the right and not on the left as shown in Fig. 1. Users and experts complained that it is familiar for everybody that the "Back" button is placed on the left of any interface and people will get frustrated searching for this button if it was placed anywhere else. However, when specific tasks were given to experts (i.e. to turn on their morning mode), they recommended the interface to have a "Scenario" button in the front page of NISHA so people can finish the task with one click instead of having 3 clicks by doing each action separately. Furthermore, referring back to literature review findings, work in [8] proved that considering images instead of colors when designing buttons achieve more user satisfaction and higher learnability, at this point the solid color buttons of NISHA where replaced with Virtual image buttons as shown in Fig. 2.

After modifying NISHA based on the results of the second iteration, the interface was given again to a test group to evaluate. Third evaluation of NISHA showed better results but also revealed some missing features that should be included in it; when performing the basic tasks, users found it easier if there were a "Cameras" button that allow users to view their house's rooms in camera frames instead of entering each room

TABLE I. LITERATURE REVIEW FINDINGS (STRENGTHS AND WEAKNESSES)

Literature Review	Strengths	Weaknesses / Challenges
Virtual Place Framework for User-centered Smart Home Applications [9]	Guidelines on how to involve users in the design process to reduce the gap Virtual Reality increased accessibility of system	Generality of guidelines
Human Centred Design for Graphical User Interfaces [15]	Guidelines Images instead of colors Use color contrasts to separate	Usability Centered Only

	Contrasts are very effective Consider color blind people	
Design and evaluation of smart home user interface: effects of age, tasks and intelligence level [2]	Less time and Less errors using guidelines presented	Performance not clearly differentiated for senior people vs. young people
Smart Home, the Next Generation Closing the Gap between Users and Technology [1]	Full control over interaction by user achieved better results of usability	Problems with accessibility No predefined sets of guidelines
Design of web-based Smart Home with 3D virtual reality interface [16]	3D presentation of devices provides great convenience, flexibility, and immersive visualization to users	High complexity because of costs of much more software design and hardware's
Projection-Based User Interface for Smart Home Environments [4]	Very feasible approach More intuitive and natural way of user interaction because touch based Gesture recognition model for fingertips mapping	High response time
HouseGenie: Universal Monitor and Controller of Networked Devices on Touchscreen Phone in Smart Home [17]	Avoided confusion of too many devices in small screen display by adopting the filtered mode so users can focus on the important icons and hide unneeded ones	Not Pleasure to use - 2D panoramic two layered view
Control Your Smart Home with an Autonomously Mobile Smartphone [7]	Combination of voice and touch	Voice recognition Problems
3D virtual "smart home" user interface [3]	Virtual Reality technology	Users preferred cameras instead of the presented 3D interface
Design of a touch screen interface for a mobile position aware instant messaging client [6]	Low response time	Assumption of one hand not convenient Didn't involve users and hence user satisfaction was low

Fig. 1. "Back" button placed on the right of the page – weakness

Fig. 2. Virtual images buttons instead of solid colour buttons

TABLE II. NISHA VALUES FOR THE MEASUREMENT OF "TASK TIME" / TIME IN SECONDS

Iteration	Basic Tasks	Special Tasks	Avg. Total Task Time
1	41.4	37.2	39.3
2	32	37.2	34.6
3	26.7	28.5	27.6

TABLE III. NISHA VALUES FOR USABILITY PARAMETERS CONSIDERED

Iteration	Satisfaction	Avg. No. of Clicks	Avg. Time asking for	Avg. No. of user lost	Avg. No. a user got frustrated
1	1.92	4	10.2	2.7	2.7
2	1.78	3	6.6	1.2	1.2
3	1.54	3	2.5	0	0.4

Separately to view the camera frame. Also, experts complained that a "Recent Actions" page is missing and stated that people find it very important to check such a page so to make sure of the actions they have done recently. These two issues were compared to the HCI intersection model drawn at the research phase of the first iteration, and found that they are concerned with Learnability and Flexibility aspects. Finally, the third iteration of NISHA has been made based on the third evaluation and again given for users and experts to evaluate.

After modifying NISHA based on the results of the second iteration, the interface was given again to a test group to evaluate. Third evaluation of NISHA showed better results but also revealed some missing features that should be included in it; when performing the basic tasks, users found it easier if there were a "Cameras" button that allow users to view their house's rooms in camera frames instead of entering each room separately to view the camera frame. Also, experts complained that a "Recent Actions" page is missing and stated that people find it very important to check such a page so to make sure of the actions they have done recently. These two issues were compared to the HCI intersection model drawn at the research phase of the first iteration, and found that they are concerned with Learnability and Flexibility aspects. Finally, the third iteration of NISHA has been made based on the third evaluation and again given for users and experts to evaluate. The fourth evaluation of NISHA was promising and showed better results compared to the first and second iteration, NISHA also achieved positive experts' and users' comments.

B. Pre-Questionnaire Structure

The following topics were contained in the pre-questionnaire:

- General questions: Age, gender, and professional status.
- Requirement questions
- Interface features
- Special features questions

Furthermore, in the experts' interviews additional questions were asked to the experts about new or recommended ideas for the application. Most of the questions in the interviews were open. Some of those questions are:

- What are the buttons the users use most?
- How often do users ask for help?
- Is there any features users ask for, that are not found in the existing interfaces? Kindly provide us with examples.

III. METHODOLOGY TIMELINE

Fig. 1. NISHA designing process timeline presents the timeline of NISHA designing, development, and evaluation phases. First of all, huge research have been made on existing similar interfaces and on smart home implementation, then strengths and weaknesses of these interfaces and works have been taken into consideration and used to build a background knowledge and identify potential problems. Secondly, an online closed ended pre-questionnaire was conducted and given to a test group of 41 typical users to fill in order to have w background knowledge of these users and their mentality. The pre-questionnaire questions were chosen based on the experts' recommended features and on Nielsen's standards of usability [11]. At the point everything have become clear to draw an initial design of the interface, so mockups were made for NISHA based on the intersection and on the results of the pre-questionnaire.

At the testing design phase, the mock-ups were tested by the same experts and compared to the HCI rules chosen first and to the users' requirements collected from the pre-questionnaire. Moreover, at this phase, the mockups were tested by some of the usability aspects like learnability, flexibility, and robustness. Testing Design phase is iteratively repeated during the development cycle of NISHA until the initial design of the interface met users' requirements and HCI rules that were listed in the third phase. Later on, the real design of NISHA was made on a touch enabled device and linked to a wooden house-prototype that has two lights and one motor connected to the electricity.

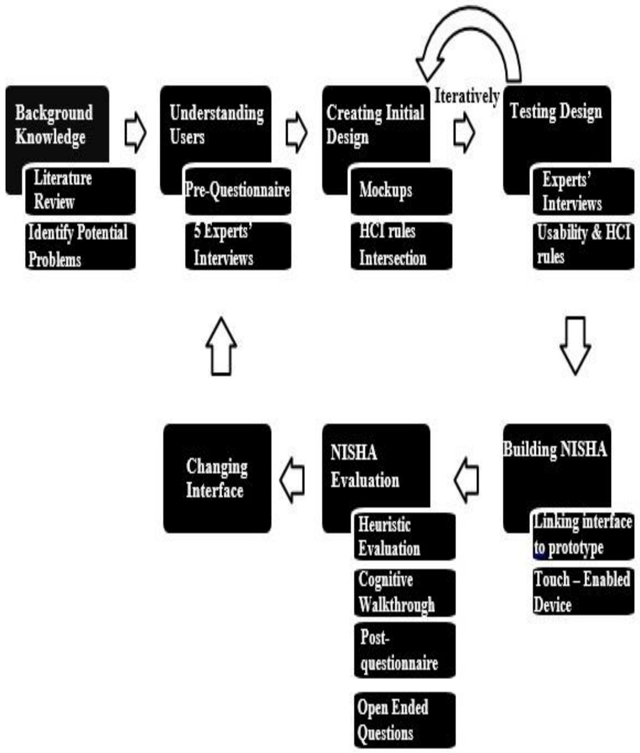


Fig. 3. NISHA designing process timeline

In the Evaluation level of NISHA two types of evaluation methods were considered; heuristic evaluation and cognitive walkthrough. In heuristic evaluation method another test group of 10 users were prepared to test NISHA as a user interface only without the use of a prototype only by providing these users with simple tasks to perform. The tasks given to the test group were mainly concerned to highlight interface features' issues such as consistency of pages, simplicity of use, and learnability of system, and did not consider other issues like response time and push up notifications because at this evaluation phase, the interface was not connected to the wooden prototype. On the other hand, we used cognitive walkthrough evaluation method by preparing 3 lifetime scenarios and given these scenarios to the same experts test group of the third phase to perform them. Each expert have been given NISHA on a touch-based device and was asked to perform certain tasks. Then, experts were asked to list the most negative and positive aspects of the interface and were asked an open ended questions like "how do you think NISHA can help you in your daily life?" Furthermore, number of clicks, time to finish a certain task, and users' reactions were recorded to evaluate NISHA. Finally, based on the interface evaluation, NISHA was enhanced and changed.

IV. Results

A. Pre-Questionnaire

The main goal of the online pre-questionnaire is to highlight the cultural, economic, social, and technological differences for people in the Arab Countries (Jordan and the West Bank). Since the circumstances and situations in the Arab world differ from other countries (ex. 3G technology is not available in the West Bank and hence users don't always have control over their homes), it is very important to build a background knowledge of these factors before designing such an interface.

B. Interface Post-Questionnaires

Another test group of 10 users were prepared to test NISHA as a user interface only, without the use of a prototype by providing these users with simple tasks to perform. The tasks given to the test group were mainly concerned to highlight interface features' issues such as consistency of pages, simplicity of use, and learnability of system, and did not consider other issues like response time and push up notifications because at this evaluation phase, the interface was not connected to the wooden prototype and hence there was no full system yet. The test group filled in an online questionnaire and results were satisfying; 100% of users agreed that the interface does not contain a lot of information on the screen, 90% agreed that the content fits well within the display and that information on the same screen is relevant, 20% found it misleading while 80% did not, 70% was able to use the interface with one hand, 60% agreed that there is no need to include narratives since interface is clear, 40% used a minimal number of click to reach their goal, 20% used the perfect number of clicks, and 30% stated that the number of clicks is normal. Also, 90% agreed that the interface ensure consistency while navigating from one screen to another and switching between screens, and 80% found it easy to diagnose and recover from errors.

When given a certain tasks to perform, 40% found the tasks neither difficult nor easy, 40% found them difficult, 10% only thought the tasks are easy. After completing the tasks, 40% stated that the tasks are very easy, 30% voted for "somewhat easy" option, and 10% only found the tasks difficult. There was a consensus of users that the characters on the screen are easy to read, 80% stated that they would recommend it to their friends and colleagues. Finally 50% strongly agreed to use the system frequently and 40% agreed to use it frequently.

C. Prototype Evaluation

To evaluate the developed interface based on real time actions, three scenarios have been prepared for a test group of the same experts of the interviews. Each expert have been given NISHA on a touch-based device and was asked to perform certain tasks that connect the interface with the built prototype. The prototype is a 1x1 meter wooden built house, with two lights and one motor. Users' reactions and comments were recorded while performing the tasks, and also they were asked to fill in the post-questionnaire after completing the tasks. The three scenarios given to experts are described as follows:

1) Scenario 1

Description: it is very desirable for users when leaving their house to switch off/on ALL the devices, for sure except those that always have to be turned on like the alarm system, refrigerator, etc. NISHA should be able to handle this and allow users to turn off/on these appliances before leaving or entering the house. This feature is one of the most important features that users asked for in the pre-questionnaire.

a) Scenario 1 – Task 1

After a long working day, it's time to have some rest and sleep. Unfortunately, while you are sleeping late at night, you hear footsteps and a strange movement around in your kitchen. You are very sleepy and still can't recognize whether it's an illusion or a real danger. The best thing to do is to turn on ALL of your kitchen's lights so everything becomes clear, but apparently it is a bad idea to risk and walk in the dark to turn on the lights. Moreover, if everything is ok, then it is time to turn off ALL unneeded appliances and lights in the house and sleep. How can you handle this problem using NISHA?

2) Scenario 2

Description: any appliance in the house can be switched on/off remotely from the application NISHA. A notification is shown to the users if any device is switched on/off or had a change in its current status.

a) Scenario 2 - Task2

Its morning, and you left your house heading to work, while driving you turn on the radio and the weather forecast predicts that a strong storm is coming at night. You remember that you left all your kitchen windows opened, and unfortunately you have a long working day and will not be back until very late. Use your NISHA to handle this problem.

3) Scenario 3

Description: Beside that smart home applications should be able to turn on/off ALL lights, any appliance or light in the house can be controlled separately by users.

a) Scenario 3 - Task3

On a Sunday morning you wake up with a fresh mind and a relaxed body, but absolutely you will stay in bed as long as your windows are closed and as long as your coffee is not ready, apparently this is you morning mode. How could you enjoy the morning mode without the need of leaving your bed? Also you need to turn on ONLY one of your kitchen lights so to be able to see the coffee machine on the camera.

User comments and reactions were recorded, and they were asked to answer some questions. The results of this evaluation is summarized in Table 4, Table 5, and Table 6.

V. CONCLUSION

This paper addressed the problem of inadequate usability of existing user interfaces for controlling home devices. Previous and existing user interfaces for controlling smart homes were promising, but were not usable and useful enough to meet users' expectations for a user friendly and "good" smart home interfaces. This led to the idea of designing a touch-based user interface for smart home systems that meets user expectations and follow as much as possible Human Computer Interaction rules. This study examined people, and circumstances of the developing countries of Middle East (Jordan and West Bank) and therefore is set out to design a user interface for smart home systems taking into consideration the economic, social, cultural, and technological differences that occur between these countries and other developed countries. This study included three stages; at first - in a preliminary study – users' expectations and needs were collected through an online pre-questionnaire and experts' interviews, resulting in an initial list of users' requirements; secondly, initial design of the interface was developed and given for typical users to evaluate through an online post-questionnaire, 3 iterations of interface design were then made;

Finally, the final design of NISHA has been developed and given to a number of experts to test. The Preliminary study of users in addition to experts' interviews examined the user interface elements that are highly influenced by economic, social, technological, and cultural values of Jordan and West Bank countries, and revealed that there are many differences that should be taken into consideration when designing a user interface for people in this region; Pre-questionnaire form showed that people preferred classical old fashion interfaces instead of modern advanced ones such as sliding of pages and fading of images. Moreover, experts' interviews revealed that people in these countries have high fear of smart technologies and consider these technologies as a threat to their privacy; hence NISHA was designed in a way that users can put a security key whenever they want to enter the application. Literature review studies also showed that people of the Arab countries in general don't like text-buttons and prefer to see virtual images instead of text [10], therefore NISHA buttons were designed based on virtual images. Moreover, as Blue color is considered to indicate protection in Arab cultures [14], it was considered for NISHA logo and design. Finally, as the 3G technology is not always available in West Bank countries, and since people don't always have access to the internet, NISHA was designed to be working automatically in special cases, and not to wait for users' confirmation (i.e. if there is a strong storm, close windows without waiting users' confirmation).

The final evaluation of NISHA by experts involved a 3 prepared real life scenarios that were given to those experts as a tasks to be done using the interface over a wooden house prototype that consists of two lights and a motor. Finally, results and feedback of post-questionnaire and experts' evaluation showed that users' and experts' satisfaction of using the interface was high and ranked an average of 8 out of 10 (80%) compared to the existing smart home interfaces. Moreover, 90% of users were satisfied that NISHA met their expectations, and finally 93% of users strongly asked to have NISHA in their daily lives.

VI. FUTURE WORK

First of all, as the feedback of users refers to the response-time of the interface is not very fast, we aim to enhance this feature. Secondly, a combination of touch-speech-recognition interface is planned to be developed since users find it easier for some tasks to be done by voice commands instead of touch commands. This is planned to be done by allowing users to use their voice for general tasks, like entering the kitchen, opening the windows, turning off the lights, etc. While specific tasks like oven settings, scenarios, etc. are controlled by touching commands.

TABLE IV. NISHA EVALUATION RESULTS BASED ON EXPERTS' REACTIONS / RECORDING

	Please list the most positive aspects you found using this interface	**Please list the most negative aspects you faced using this interface**
Expert 1	"Very simple to use" "User-friendly" "Clear Buttons"	"Not of a very high responding speed"
Expert 2	"Very simple and usable" "Contains every single detail the user may ask for"	"Presets Page is not clear"
Expert 3	"It saves my time" "It is Useful"	"Responding time is high"
Expert 4	"Navigation among screens is easy" "Very user friendly and simple"	"None"
Expert 5	"This interface is very easy to use and can save my time, I was pleased while testing and performing the scenarios"	"Presentational design is of an old fashion"

TABLE V. NISHA EVALUATION RESULTS BY EXPERTS / DIRECT CLOSED ENDED QUESTIONS

	Strongly Agree	**Agree**	**Disagree**	**Strongly Disagree**	**N/A**
It makes Things I want to accomplish easier to get done	20%	80%	-	-	-
It saves my time	60%	20%	-	-	20%
It is Simple to use	80%	20%	-	-	-
Using this interface is effortless	-	80%	-	-	20%
Shifting among buttons is easy	40%	60%	-	-	-
It has all the functions and capabilities I expect it to have	60%	20%	20%	-	-
I feel I need to have it	60%	20%	20%	-	-

TABLE VI. NISHA EVALUATION RESULTS BY EXPERTS / DIRECT CLOSED ENDED QUESTIONS

	Please list the most positive aspects you found using this interface	**Please list the most negative aspects you faced using this interface**
Expert 1	"Very simple to use" "User-friendly" "Clear Buttons"	"Not of a very high responding speed"
Expert 2	"Very simple and usable" "Contains every single detail the user may ask for"	"Presets Page is not clear"
Expert 3	"It saves my time" "It is Useful"	"Responding time is high"
Expert 4	"Navigation among screens is easy" "Very user friendly and simple"	"None"
Expert 5	"This interface is very easy to use and can save my time, I was pleased while testing and performing the scenarios"	"Presentational design is of an old fashion"

REFERENCES

[1] HwangA., HoeyJ., Smart Home, The Next Generation Closing the Gap between Users and Technology; [Online] [Accessed on 2014 March]. Available from URL https://cs.uwaterloo.ca/~jhoey/papers/AAAISS2012HwangHoey.pdf.

[2] Bin Z, Pei-Luen PR, Gavriel S, Design and evaluation of smart home user interface: effects of age, tasks and intelligence level; Behaviour & Information Technology,2009, vol. 28,pp.239-249.

[3] Borodulkin L, Ruser H, Trankler HR, 3D virtual "smart home" user interface. Virtual and Intelligent Measurement Systems; VIMS IEEE International Symposium, 2002,pp.111– 115.

[4] Chin-Yang Lin, Yi-Bin Lin, Projection-Based User Interface for Smart Home Environments; Computer Software and Applications Conference Workshops (COMPSACW) IEEE 37th Annual; Japan; pp. 546 – 549.

[5] Davis FD, Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology; MIS Quarterly, 1989, vo. 13,pp.319-340.

[6] StenmarkF., Design of a touch screen interface for a mobile position aware instant messaging client; [Online] [Accessed 2014 April]. Available from URL http://www8.cs.umu.se/education/examina/Rapporter/FiaStenmark.pdf

[7] WangH., Saboune J, El Saddik A,Control Your Smart Home with an Autonomously Mobile Smartphone; Multimedia and Expo Workshops (ICMEW) IEEE International Conference; San Jose, 2013CA, pp.1-6.

[8] Hwang A., Truong K., Mihailidis A;; Using participatory design to determine the needs of informal caregivers for smart home user interfaces; 6thInternational Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth); Diego, 2012CA; pp. 41-48.

[9] Jumphon L., Jinwon C.;; Virtual Place Framework for User-centered Smart Home Applications; Mahmoud A. Al-Qutayri; Smart Home Systems.Korea; 2010,pp.177- 193.

[10] Khanum, M. Akheela, F., Shameem, Chaurasia, Mousmi A;; Arabic Interface Analysis Based on Cultural Markers; International Journal of Computer Science,2012vol. 9, no. 1,pp 255-260.

[11] Nielsen, JUsability Engineering. Academic Press, Boston; ISBN 0-12-518405-0 (hardcover), 0-12-518406-9 (softcover1993; Japanese translation ISBN 4-8101-9009-9.

[12] Othmar Kyas, How to Smart Home; [Online] [Accessed 2014 February]; Available from URL http://www.openremote.com/wp-content/uploads/2013/12/How-To-Smart-Home-PDF-OR.pdf

[13] ZhangP.,BenbasatI., CareyJ., Fred Davis, Dennis F Galletta, Diane M Strong;; Human-Computer Interaction Research in MIS Discipline; AIS Transactions on Human-Computer Interaction; 2002, vol. 3, no. 1,pp. 55-107.

[14] Ruecian, Colors of Religion: Islam; [Online] [Accessed 2014 March]. Available from URL, http://www.colourlovers.com/blog/2007/09/08/colors-of-religion-islam/

[15] Van der Veer, Human centred Design for Graphical User Interfaces.[Online] [Accessed 2014 May]. Available from URL http://publik.tuwien.ac.at/files/pub-inf_2609.pdf

[16] Wenshan Hu, ZhouH., Chaoyang Lin, ChenX., ChenZ., Yiyan Lu;; Design of web-based Smart Home with 3D Virtural Reality Interface; Control (CONTROL) UKACC International Conference;; 2012,pp. 223 – 228.

[17] Yue S., Chenjun W., Yongqiang Q., Chun Yu, Yu Zhong, Yuanchun Shi;; HouseGenie: Universal Monitor and Controller of Networked Devices on Touchscreen Phone in Smart Home; Ubiquitous Intelligence & Computing and 7th International Conference on Autonomic & Trusted Computing (UIC/ATC); Xian, Shaanxi; 2010,pp.487-489.

[18] KhamaysehY., MardiniW.AljawarnehS., Bani_YasseinM.," Integration of Wireless Technologies in Smart University Campus Environment: Framework Architecture", International Journal of Information and Communication Technology Education, 2015, vol. 11, no. 2, pp.60-74.

SIP Signaling Implementations and Performance Enhancement over MANET: A Survey

Mazin Alshamrani
[1]Studies and Decision Support Center, Department for Planning and Development, Ministry of Haj, Saudi Arabia

Haitham Cruickshank
[2]Institute for Communication Systems (ICS), Faculty of Engineering and Physical Science, University of Surrey, Guildford, GU2 7XH, United Kingdom

Zhili Sun
[2]Institute for Communication Systems (ICS), Faculty of Engineering and Physical Science, University of Surrey, Guildford, GU2 7XH, United Kingdom

Godwin Ansa
[3]Department of Computer Science, Akwa Ibom State University,
Ikot Akpaden, Mkpat Enin, Akwa Ibom State, Nigeria

Feda Alshahwan
[4]Public Authority for Applied Education and Training, College of Technological Studies,
Electronics Engineering Department, Computer Section, Kuwait University, Kuwait

***Abstract*—The implementation of the Session Initiation Protocol (SIP)-based Voice over Internet Protocol (VoIP) and multimedia over MANET is still a challenging issue. Many routing factors affect the performance of SIP signaling and the voice Quality of Service (QoS). Node mobility in MANET causes dynamic changes to route calculations, topology, hop numbers, and the connectivity status between the correspondent nodes. SIP-based VoIP depends on the caller's registration, call initiation, and call termination processes. Therefore, the SIP signaling performance has an important role for the overall QoS of SIP-based VoIP applications for both IPv4 and IPv6 MANET. Different methods have been proposed to evaluate and benchmark the performance of the SIP signaling system. However, the efficiency of these methods vary and depend on the identified performance metrics and the implementation platforms. This survey examines the implementation of the SIP signaling system for VoIP applications over MANET and highlights the available performance enhancement methods.**

Keywords—SIP; VoIP; MANET; Peer-to-Peer; Back-to-Back User Agent (B2BUA); IMS

I. INTRODUCTION

SIP signaling is widely used to manage and control voice calls over IP-based network systems. The main functions of SIP signaling are: (1) inviting other parties to initiate a call, (2) adding media streams during a call, (3) changing the encoding system during a call, (4) transferring or holding voice calls. The capabilities of the SIP signaling system depends on the implementation systems of the SIP signaling that is used, and the level of support that the network system provides for the application layer services. On the other hand, a Mobile Ad Hoc Network (MANET) is a self-organizing, infrastructure-less, and multi-hop network that consists of unlike groups of nodes with limited capabilities and energy constraints.

The features of MANET include mobility of nodes, variable topology due to the dynamic nature of the network and multipath communication scheme. The communicating nodes in a MANET usually seek the help of other intermediate nodes to establish communication channels. Each node in a MANET works both as a host and a router. Unpredictable connectivity due to the dynamic nature of the network is another challenge faced by MANET. Developing efficient and dynamic routing protocols is a key challenge in MANET.

This review is focused on research in SIP signaling over MANET and the performance enhancement approaches for SIP-based VoIP applications. In this paper, the current state-of-the-art, results, gaps, the merits and demerits of the four types of SIP signaling systems over MANET mentioned here and the performance enhancement methods for SIP signaling over MANET are discussed in detail. Finally, two open issues have been identified and highlighted for future investigations.

A. SIP Signaling System

SIP is an Internet Engineering Task Force (IETF) standard for signaling protocol released as RFC 3261 [1]. SIP is commonly used for controlling multimedia communication sessions such as voice and video calls over Internet Protocol (IP). SIP is used in initiating, managing, and terminating multimedia sessions such as voice calls over IP based networks. This session can be either a two-way call, which is either unicast or collective multimedia calls, which is multicast. These features have made SIP a better choice for providing VoIP services in the last few years. SIP is an application layer protocol, which serves five main functions for multimedia calls [1]. These functions are: User Location, User Availability, User Capability, Session Setup, and Session Management.

User Location is used to determine the location of the end user, while User Availability examines the willingness of the end user to participate in the call session. User Capability supports the applications compatibility with different communication systems and users to determine the required methods and standards for the requested multimedia applications. Session Setup provides the resources to setup and establish the communication. Finally, the Session Management function supports the call management services in different ways such as adding, transferring and modifying the session parameters.

SIP is rather a component which works in a framework with other IETF protocols to build a complete multimedia architecture. The most common protocols which are used in this architecture are: Real-Time Transport Protocol (RTP) for real-time data transportation, Real-Time Streaming Protocol (RTSP) for controlled delivery of streaming media, and Session Description Protocol (SDP) for multimedia session description.

1) SIP Components

SIP works collectively and in conjunction with different protocols and technologies. SIP consists of two basic components known as User Agents and SIP servers. User Agents are the end points of the call, while SIP servers facilitate the sending of responses back to the requested client. User Agents are self-sufficient in initiating a session with other nodes in the network. Each node consists of two fundamental components known as User Agent Server (UAS) and User Agent Client (UAC). UAC is responsible for initiating a new session, while UAS handles all the connection requests of the clients.

A SIP server is responsible for handling the user name and the IP addresses of the User Agents which connect to it. There are four different SIP servers that are used to handle the calls' interconnection processes to different user agents in the network [2]. These SIP servers are: proxy server, location server, registration server, and redirect server. The proxy server is responsible for forwarding the requests on behalf of user agents. The location server is used to find the information about possible locations for the callee. The location server is most times incorporated within the proxy server features. The address registered to the register server is stored in the location server. The registration server is used for registering a user agent when it is logged into the network.

Hence, registration servers are responsible for registering the location of the user agents. The registration server is used to discover the IP address of the user agents and then map the IP address to the related user name. Finally, the redirect server is responsible for redirecting the clients to the user agents with whom they want to initiate the call session. The redirect server sends back the IP address of the user agent with whom other clients want to communicate. The main difference between the proxy server and the redirect server is that the proxy server forwards on behalf of the UAC, the redirect server on the other hand provides the IP address so that the UAC can contact other UACs directly.

2) SIP Messages

SIP is a text-based protocol similar to the Hyper-Text Transfer Protocol (HTTP), which is used for the forwarding of information between UAC and UAS, by using several requests and responses [3]. The request methods used in SIP are REGISTER, INVITE, OPTIONS, ACK, CANCEL and BYE. The REGISTER request is used for registration when a user agent initially logs on to the network. The INVITE request is used for inviting other UACs to establish communication and then to start a new SIP session between them. The OPTIONS request is used to query the server to find out the capabilities of other User Agents. The ACK request is used to acknowledge a session before exchanging the related messages. The CANCEL request is used to cancel a pending request, while the BYE request is used in terminating a session. The request methods are replied to with one of the response codes used by SIP.

The request methods used by SIP consist of six classes. The first class of response code belongs to an information or 1xx which is used to inform that the request is received and processed by having its provisional response, such as 180 ringing. The second class of response code belongs to success or 2xx which is used for acknowledgment, such as 200 OK. The redirection requests or 3xx is the third class of response code which tells that the request cannot be completed and needs redirection of the user agent, such as 302 moved. The fourth class of request code belongs to client error or 4xx which signifies that the server cannot process, such as 407. This means that SIP server authentication is required even for the Back-to-Back User Agent (B2BUA) where the SIP server is acting as a UAS. The fifth response class belongs to the server errors or 5xx which signifies that the server cannot process the request, such as 503, that means that the service is unavailable. The final class of response code is the server response code, known as the global error or 6xx. This code informs that the server cannot process globally, such as 603, which means decline. When a user agent wants to initiate a session with another user agent, the queries of the client are processed by specific servers.

The proxy-based SIP server on the other hand, relies on the SIP signaling system only for the registration stage of the SIP call processes. This is achieved by maintaining the transaction state of the SIP calls. The IP addresses and locations of the connected clients could be exposed by the callers because the proxy-based SIP server has a low level of security.

Fig. 1 shows the message flow for a simple scenario which depicts the invitation and termination transactions between two users through the B2BUA SIP server. There is a difference between the B2BUA-based SIP server and the proxy-based SIP server with regards to SIP signaling flow. The B2BUA maintains the whole call state and participates in all call requests. It is involved in the call initiation, management, and termination processes. Therefore, the B2BUA system of the SIP server provides a secure, reliable communication system for different User Agents (UAs) where all SIP signaling messages and voice data need to go through the SIP server.

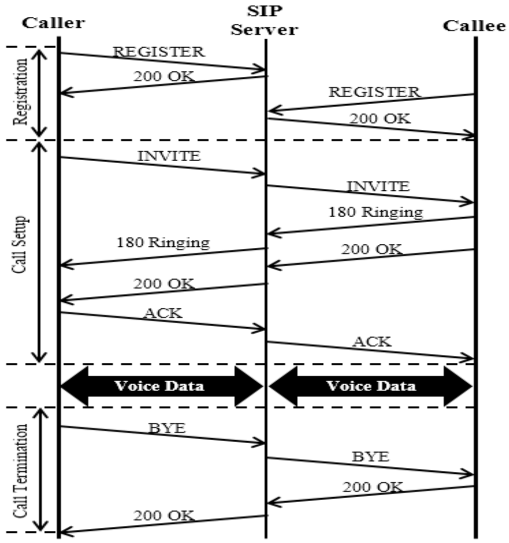


Fig. 1. The signaling flow for a SIP-based VoIP application using B2BUA-based SIP server

The IP addresses, port numbers, and locations of the users are only known to the B2BUA SIP server but hidden from each client thereby providing secure connectivity. The B2BUA SIP signaling system is commonly adopted for privacy approved VoIP implementations, such as military applications and secured call services. The single point of failure problem and congestion overhead are the main disadvantages of a B2BUA-based SIP server. The interactions in Fig. 1 show the use of the SIP methods INVITE, Ringing, and BYE through the SIP Sever. The SIP server depicted here records all the interactions. It is used as the coordinator of the Internet working system between the two ends, with exception of media transmissions. The Media Data mostly depends on the Real-Time Protocol (RTP) and Real-Time Control Protocol (RTCP). The call setup time consumes more time when compared with the termination time. The termination messages could be generated from both ends depending on the type of application and the connection system. In general, the proxy, redirect, register, and location servers are known as the B2BUA SIP Server as represented in Fig. 1. The interactions between the entities of the SIP server are integrated together to provide the SIP services depending on the connectivity methods.

B. SIP Implementations

Many VoIP phone companies allow clients to use their own SIP devices, as SIP-capable telephone sets, or soft phones. The market for consumer SIP devices continues to expand and there are many devices such as SIP Terminal Adapters, SIP Gateways, etc. The free software community has started to provide more and more of the SIP technology required to build both end points as well as proxy and registration servers. This will lead to a commoditization of the technology and accelerate global adoption. As an example, the open source community at SIP foundry actively develops a variety of SIP stacks, client applications, in addition to entire IP Private Branch Exchange (IP PBX) solutions that compete in the market against mostly proprietary IP PBX implementations from established vendors [4].

SIP-enabled video surveillance cameras can make calls to alert the owner or operator that an event has occurred. For example to notify that motion has been detected out-of-hours in a protected area. Other feasible application examples include video conferencing, streaming multimedia distribution, instant messaging, presence information, file transfer and online games. In general, there are four types of implementations for the SIP signaling system. These implementation types are: Peer-to-Peer SIP system, multiple server based SIP system, single server based SIP system, and IP Multimedia System (IMS) based SIP system.

1) Peer-to-Peer SIP System

Most of the SIP signaling or traditional SIP signaling is based on Client/Server architecture. In Peer-to-Peer (P2P) architecture, clients have capabilities of both client and server, and are capable of starting a new session with each other and requesting services [1]. Each node is capable of providing services and resources, and in case any node is unable to provide the services then the next node can be contacted. Nodes in a P2P architecture have the features of both UAC and UAS.

Therefore, P2P SIP provides instant messaging or VoIP services with the help of P2P architecture, where session initiation and communication between users is facilitated by the SIP protocol. The Client/Server architecture needs a SIP server for handling requests and responses. However, in the P2P based SIP architecture, there is no need of SIP servers. To-tag, From-tag and Call-ID are collectively used for handling the dialogue between UAC and UAS in P2P-based SIP [5]. Fig. 2 shows message exchange between two devices using P2P SIP.

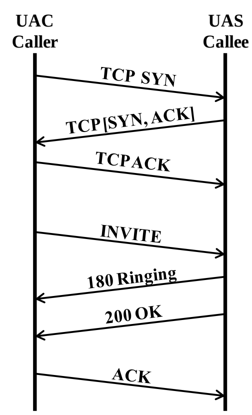


Fig. 2. Signaling flow of messages over Peer-to-Peer SIP

The SIP protocol stack is handled by the various protocols based on the media protocol stack, for example at the transport layer, the TCP/UDP protocol is used. In P2P SIP, two users are involved in the communication process and no SIP server is used as shown in Fig. 2. In this case, users is do not need to register at any SIP server. A TCP SYN packet is sent to open the connection since the TCP protocol is used at the transport layer. SYN consist of an initial sequence number to be used in subsequent communication between the two parties. The callee responds with the SYN message consisting of the initial sequence number and the ACK message, which confirms that the callee has received the SYN frame from the UAC. Then the UAC sends the TCP ACK message consisting of the UAC acknowledgement number and completes the 3-way handshake [5].

With the completion of the 3-way handshake, the connection is now open for communication. The UAC caller exchanges the message by sending a SIP INVITE message to the UAS callee. The INVITE message consists of various details, such as session type, which can be either a multimedia or a voice session. There are various other fields in the INVITE message. The first header field in the INVITE message is Via, which is usually a host name and further maps to the IP address using DNS query. In addition, the header field consists of the SIP version, transport layer protocol used, host name and port number. The next header fields are To and From, which dictate the sender and receiver details of the SIP request. Call-ID header field is the next header field which is used to keep track of a particular SIP session [6]. To-tag, From-tag, and Call-ID are known as tags which are collectively used as identifying parameters.

The initial INVITE message consists only of From-tag and the UAC caller generates an INVITE message which consists of both From-tag and Call-ID. In response to the INVITE message, the user agents who respond to this message will generate the To-tag. The SIP parameters From-tag, To-tag, and Call-ID are used to identify an initiated session. Furthermore, the Content-type and Content-length header fields are used to represent the message body as the SDP protocol. The SDP Content-type describes the media information using various SDP fields, such as media format port number, IP address, media transport protocol, media encoding, and sampling rate [6].

After receiving the INVITE message, the UAS callee responds back by sending 1xx or 180 ringing. The UAS callee creates a 180 ringing message by copying several header fields from the INVITE message [6], such as From, To, and Call-ID. The 180 ringing message consists of a header field known as the CONTACT header field, which specifies an address at which the UAC callee can be contacted. Once the UAC callee is ready to initiate the session, a 200 OK response is sent back to the UAC caller. The 200 OK message consists of the UAS callee SDP message using similar SDP fields. Finally, acknowledgement ACK is sent by the UAC caller to start the media session. Using another protocol for media data transfer, a media session is established between the UAC and UAS. The major advantage of P2P-based SIP is scalability [5]. As in P2P SIP, a user agent need not register with a central server. Instead, the user agent needs to register with an overlay network formed by UAC in the system [5]. Client/Server based SIP needs more maintenance and configuration. On the other hand, P2P-based SIP is more scalable and reliable as there is no single point of failure [7]. In addition, P2P SIP does not need maintenance and configuration including NAT and Firewall. All these benefits come at a cost of increased number of security threats and look-up delays [7]. As in Client/Server based SIP, look-up cost is very low, while in P2P SIP, look-up cost is comparatively very high. Security features such as authentication, and reputation is another major drawback of P2P SIP.

2) Multiple Servers Based SIP System

The multiple server SIP is based on the client/server architecture in which all the servers, such as proxy server, location server, and registration server, respond to the request sent by the UAC separately. Multiple servers use the Redirect server for initiating a session between a UAC caller and a UAS callee. The Redirect server does not forward the request on behalf of the UAC; it only returns the location shown in Fig. 3. The UAC caller registers itself with the Registration server by sending a REGISTER message. After receiving the REGISTER message from the registration server, it extracts the user name, IP address, and port number then stores them in the location server [6]. A contact header field of the REGISTER message holds information on the lifespan of the registration. Similarly, the UAS callee also registers itself at the registration server. The location details of both the UAC caller and the UAS callee are stored in the location server. An INVITE message is sent by the UAC callee to the redirect server.

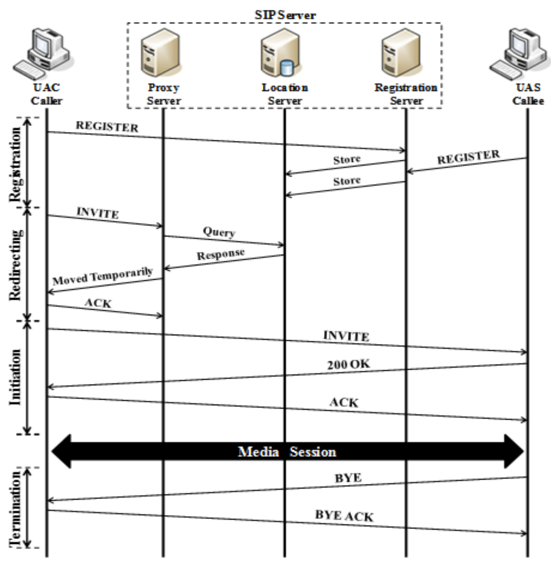


Fig. 3. Signaling flow of SIP messages over multiple SIP servers

The INVITE message consists of the header fields, such as INVITE, Via, Max-Forwards, To, From, Call-ID, CSeq, Subject, Contact-type, and Content-length [2]. The Redirect server performs a look-up within the database of the location server for the intended recipient. Then the location information of the user is sent back to the UAC in a redirection class response. The response Moved Temporarily (302) contains the message format having header fields SIP moved temporarily, Via, To, From, Call-ID, CSeq, Contact-type and Contact-length. After getting the response, the UAC callee acknowledges using an ACK response. At this stage, the redirection process and the exchange process are completed. A new INVITE message is sent directly to the UAS callee as the location is obtained from the control header field of Moved Temporarily in response to the redirect server. The new INVITE message contains a new Call-ID.

In response to the INVITE message, a direct 200 OK response is sent instead of the 180 ringing response. The UAC caller responds to the UAS callee by acknowledging it using an ACK response. Thus, a session is initiated between the UAC caller and the UAS callee using a redirect server. After initiating the session, the media session is started between the UAC caller and the UAS callee using the RTP protocol. Once the media session is completed, the session is terminated by sending a BYE request. Once it is acknowledged by the UAC caller, the complete session is terminated. In multiple server based SIP, the redirect server does not forward session initiation requests for the UAC caller as is done by the proxy server. Since the redirect server does not initiate the request, a lower state overhead is needed compared to a proxy server. Multiple server based SIP uses the redirect server which processes very few messages, therefore it has high processing capacity [6].

3) Single Server Based SIP System

The Single SIP server is based on Client/Server architecture in which the client sends requests to the server, and the server replies to the corresponding request of the client for establishing communication. A UAC requests the services and SIP servers, such as redirect server, or register server respond to those requests. The single server based SIP signaling system is a Back-to-Back User Agent (B2BUS) implementation, as shown in Fig. 1. Initially, the caller sends a REGISTER request to the SIP server. After receiving the REGISTER message, the information in the request message of the caller is updated in the database used by proxies. The REGISTER message sent by a caller consists of the address of the SIP server [6].

The REGISTER request contains To and From header fields. The To header field consists of the User Resource Identifier (URI) to be registered on the server. The next Contact header field containing the SIP URI is stored by the registrar [3]. Then the SIP server acknowledges the caller by sending a 200 OK response message. Similarly, the callee also registers himself on the SIP server. In this case, the SIP server is playing the role of both a registration and location service [6]. After completing the registration process, the caller is not aware of the callee's current location. The caller also needs to check whether the callee is available for the session initiation process or not. Hence, the SIP server is used for inviting the callee, as the SIP server forwards the request on behalf of the user agent. Initially the DNS look-up is performed by the caller SIP URI. It returns the IP address of the SIP server to handle the callee domain. Then the INVITE message is sent to that mapped IP address of the SIP server.

Furthermore, the SIP server looks up in its own database to locate the callee's current location. The process consists of two major steps: the DNS look-up step which is performed by the user agent to find the IP address of the SIP server, then the database look-up which is performed to locate the IP address of the SIP server. An INVITE message is then forwarded by the SIP server to the callee's IP address using a Via header field, having the address of the SIP server [3]. The callee becomes aware that an INVITE message has been routed through the SIP server because the INVITE message consists of two Via header fields. After receiving the INVITE message, the callee sends back a 180 ringing response code to the caller. The 180 response code is created by copying the header fields, such as To, From, Call-ID, and Cseq from the INVITE request. A response code is sent to the callee through the SIP server. The first Via header field contains the received parameters while the second Via header field contains the IP address in the URI. After receiving the 180 ringing response by the SIP server, the SIP server checks the contents of the first Via header field. Furthermore, when the SIP server finds the first Via header field consists of its own address, it removes the first Via header field and forwards the response to the address within the second Via header field.

Now, the callee is ready to start the session with the caller, it sends back a 200 OK message through the same set of proxies. The SIP server follows a similar process by removing the first Via header field and forwards a 200 OK message back to the caller. The contact header field of the callee in the 200 OK message allows the caller to send an ACK message directly to the callee by bypassing the SIP server. However, it needs to be noticed that the request is sent to the callee's contact URI not in the address of the contact header field. After getting the ACK message from the callee, the session is started between the caller and the callee. At this point, the transmission session is established between the caller and callee using the RTP protocol. In this scenario, the SIP server is used for contacting and locating both end points. The SIP server can drop the path if there is no exchange of media. In the SIP protocol, the path of the signaling message is different from the path of media packets. After the successful transfer of voice data, the connection is terminated using a BYE message. Once the BYE message is received by the callee, it responds by sending back a 200 OK message. On receipt of the 200 OK message, the media session and the transmission process is terminated.

In this case, the SIP signaling is performed using a single SIP server, which forwards the request on behalf of the user agent. A SIP server only forwards the message at the application layer level. It is allowed to modify both request and response, as defined in RFC 3261 [6]. Hence, the SIP server establishes end-to-end communication and preserves end-to-end transparency. As the SIP server can be either a stateful or stateless proxy. All the requests and responses that have been received in the past are tracked by a stateful proxy and can be beneficial for future processing of requests. One such example is the transactional stateful proxy [6]. Reliability is ensured when the TCP protocol is used in a stateful proxy. However, a stateless proxy does not keep track of the request and response messages. A stateless SIP server has higher processing capacity. Major benefits of the SIP server include reliability using replication, flexibility and the use of stateful or stateless proxies. If the number of proxies handling the message exceeds the limits calculated by the Max-forwards header field, then the SIP server discards the messages. If the SIP server is not properly scaled it can have a potential overload

4) IMS-based SIP System

The IP Multimedia System (IMS) is a concept for providing multimedia services regardless of the media type. The IMS provides a common architectural framework for most media. The IMS consists of multiple SIP proxies known as Call Session Control of Function (CSCF) for supporting multimedia services functionalities. The CSCF with other variants, such as P-CSCF (Proxy-CSCF) are used for SIP signaling. The P-CSCF is the first contact point for an IMS terminal and Internet with Gateway GPRS Support Node (GGSN) for resource allocation. The P-CSCF is assigned to an IMS terminal before registration. The I-CSCF (Interrogating-CSCF) performs similar functions to what the registration server does. The I-CSCF is responsible for routing to the S-CSCF. The S-CSCF on the other hand facilitates control and service triggers [8]. The IMS provides more efficient services and provisioning of capabilities than circuit and packet switched networks [14]. When any user initially registers to the IMS, a Subscriber Service Profile (SSP) is downloaded by S-CSCF from a Home Subscriber Server (HSS) [4]. The IMS-based SIP system is shown in Fig. 4.

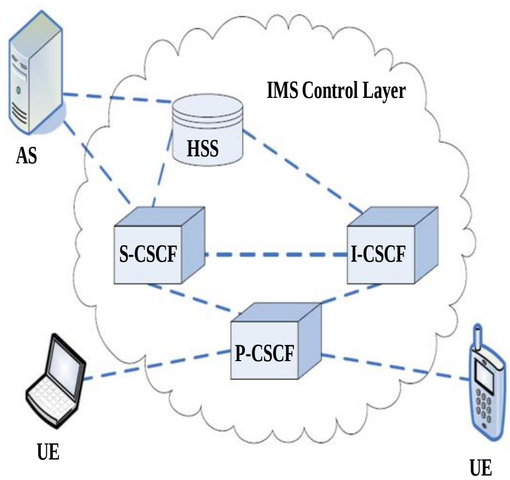


Fig. 4. Signaling flow of SIP messages over IMS-based SIP system

The first step is for the User Equipment (UE) devices to register themselves in the network. Session establishment between UE-1 and UE-2 can be such that either of UE-1 or UE-2 can originate and terminate a session. It is important that a UE has ready resources before sending INVITE and response messages [9]. The SIP-IMS message flow for the initiating session between the two UEs begins from the caller UE-1 to the callee UE-2. Initially UE-1 sends an INVITE message to the P-CSCF. The INVITE message contains various header fields, such as From, To, Call-ID, Cseq, Via, Max-forwards, Route, P-preferred identity, Privacy, Proxy-require, Security-verify, Contact, Allow, Content-type, and Content-length. After adding itself to record the route header, it forwards an INVITE message to S-CSCF then I-CSCF. The I-CSCF requests the DNS look-up for the location of user UE-2 and sends a Location Information Request (LIR) to the HSS. The HSS replies with a Location Information Answer (LIA) by providing the address of the S-CSCF of the terminating subscriber. Then an INVITE message is forwarded to the S-CSCF of the terminating visited network. The S-CSCF forwards the INVITE message to UE-2 via the P-CSCF. Then a message, 183 is sent back to UE-1 which indicates the session is in progress. After getting the 183 response code, UE-1 sends a Provisional Acknowledgement (PRACK) to UE-2. In responding to the PRACK, a 200 OK message is sent back to UE-1 for Policy Decision Point (PDP) activation, and resource reservation [10].

Next, an UPDATE message from UE-1 to UE-2 and a response code 200 OK is sent back to UE-1 for enabling QoS utilization. Since UE-2 has enough resources readily available, it sends a 180 ringing response to UE-1 via the S-CSCF, I-CSCF and the originating I-CSCF, S-CSCF and P-CSCF. It consists of the header fields, such as From, To, Call-ID, Cseq, Via, Record route, Contact, Privacy, P-Asserted identity, Privacy, and Content-type [4, 10]. UE-1 acknowledges 180 ringing message from UE-2 with a PRACK response. The PRACK consists of header fields, such as From, To, Call-ID, P-Access Network, Cseq, Via, Max-forward, Route, Ack, and Content-length [10]. A 200 OK response is generated and sent back to the UE-1 acknowledging the PRACK request. After acknowledging the PRACK request by an ACK, a session is initiated between UE-1 and UE-2 using the RTP protocol. The IMS SIP has made the provision of services such as multimedia services over IP, VoIP, and IMS possible. It has a very modular design with open interfaces. Hence, it provides flexibility for providing multimedia services over IP networks.

C. Classification of MANET Routing Protocols

Routing in MANET is a challenging task as it has a dearth of research efforts. This has led to the development of various routing protocol strategies for MANET. Each new proposed routing algorithm is supposed to be an improved version over some of the previous algorithms, considering the previous literature reviews by the authors. Since each protocol has its pros and cons when comparing it to other protocols, on the basis of certain attributes and different network scenarios. To analyze and compare Mobile ad hoc network protocols therefore, an appropriate categorization method is important. This will be helpful to understanding the nature and distinct properties of available routing protocols.

There are various ways to classify routing protocols in Mobile ad hoc networks. Most of these classifications are done on the basis of certain attributes such as routing strategy and network structure [11, 12]. Routing strategy is either table driven or source-initiated, so protocols can be categorized as either table-driven protocols or source-initiated protocols. On the structure of the network, protocols are classified as flat routing, hierarchical routing and geographical position as proposed by the authors in [13]. In general, there are three types of routing protocols in MANET [14, 15]:

1) Reactive Routing Protocols

Reactive routing protocols are on-demand protocols that discover the routes between the source and the destination when needed using the route discovery process. These routes are considered source-initiated. The most widely accepted and used reactive routing protocols are the Dynamic Source Routing (DSR) [16], Ad hoc On-Demand Distance Vector (AODV) [17], Temporally Ordered Routing Algorithm (TORA) [18], and Associativity Based Routing (ABR) [15].

2) Proactive Routing Protocols

Proactive routing protocols are traditional distributed protocols that use the shortest paths based on periodic updates. Proactive routing protocols are table driven where all possible routes to all destinations are determined at the start. Proactive routing protocols use periodic route updates and have a high routing overhead. The most widely accepted and used proactive routing protocols are the Optimized Link State Routing (OLSR) [19], Destination Sequenced Distance Vector (DSDV) [13], Fisheye State Routing (FSR) [20], and Topology Broadcast Reverse Path Forwarding Protocol Fisheye State (TBRPF) [21].

3) Hybrid Routing Protocols

Hybrid routing protocols have combined functionality from both reactive and proactive routing protocols but possess hybrid routing capabilities. The most widely accepted and used hybrid routing protocol is Zone Routing Protocol (ZRP) [22].

II. SIP SIGNALING SYSTEM OVER MANET

An overview of the existing literature on research focusing on SIP signaling performance over MANET and an extensive survey on the related work in this area is presented. This review is mainly focused on research on SIP signaling over MANET and the performance enhancement approaches for SIP-based VoIP applications. Generally, SIP is implemented over MANETs with four different types of SIP signaling systems as represented in Fig. 5.

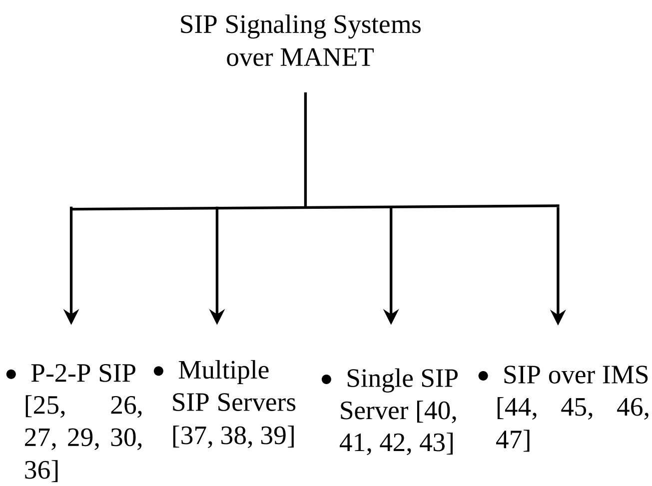


Fig. 5. A survey of types of SIP signaling system implementations over MANET

The first type of SIP signaling system is peer-to-peer SIP over MANET. The main purpose in this case is the elimination of the use of SIP servers. A detailed explanation and review of the existing research of this kind of system is given in section II (A). The second type of SIP signaling system is SIP with multiple servers over MANET. The SIP servers consist of the registration, redirect and proxy servers. The third type of SIP signaling system in the literature is SIP with a single SIP server that acts as a registration, redirect and proxy server over MANET. The fourth type of SIP signaling system over MANET is SIP with an IMS system.

In this section, we will review the current state-of-the-art, results, gaps, advantages and disadvantages with regards to the above-mentioned SIP signaling systems over MANET. Also, the available performance enhancement methods for SIP signaling over MANET will be discussed in section III. Four types for SIP signaling systems and their implementation have been introduced in section I (B). There are a number of research which primarily focuses on adapting SIP to MANETs. Such works can be categorized into two classes. This classification is based on which node(s) act as SIP server(s) in the network.

The first class is characterized with the implementation of the SIP servers in all nodes. Each node can register locally or broadcast location information in the entire network. The second class distinguishes some nodes which act as SIP servers. This survey presents the state-of-the-art in terms of the investigation, evaluation and various service enhancement techniques used in the implementation of SIP signaling system over MANETS. The simulation tools and test-beds for the implementation of SIP signaling systems for MANET will be discussed in this section.

A. Peer-to-Peer SIP Signaling Implementations over MANET

The authors of [23] propose two solutions for enabling SIP in MANETs: dSIP and sSIP. In dSIP, each node broadcast a REGISTER request to notify all nodes in the network with information about its location. Discovery of members in the network is accomplished by probing the cache locally. To enable Session Initiation Protocol in MANETs, the Service Location Protocol (SLP) [24] is used by sSIP [23]. An SLP request is broadcast from the node that wishes to connect to an ad hoc network in order to ask for bindings of users that are available. Every node that receives an SLP request responds using an SLP reply that includes its binding. As mentioned earlier, using this kind of solution can cause flooding. This can cause problems when used in larger ad hoc networks.

The authors of [25] employ peer-to-peer cover that is structured and related to Chord [26]. In order to map users with the relevant connection information, a Distributed Hash Table (DHT) protocol is used by the nodes. Hence, when some of the nodes connect to the Chord cover, they will be in charge for keeping the information related with the part of the cover that is mapping to its estimated Node-Identification (Node-ID). The maintenance of hash tables contribute to high control overhead. Registration in [27] is achieved by using the multicast mechanism with IPv6. A REGISTER request is multicast by a node to announce its presence to the whole network. The User-List-Cache is updated by each node when REGISTER updates are received. On receipt of the REGISTER updates, each node replies by providing the information to the correspondent using unicast. However, this solution gives poor results and is ineffective for large ad hoc networks, due to the preservation of a User-List-Cache

Research work on the subject of SIP over MANET was initiated in 2003 by the authors in [28]. Their research presents a framework for conference signaling using SIP which allows a MANET user to discover, initiate conferences, and join existing conferences with other users. Another research on SIP over MANET was carried out in 2004 by [29]. In this work, SIP is set up over OLSR using a cross layer, integrated application and routing layer to assist proxy-less and proxy-based systems. A proxy-less system is without a proxy server and a proxy-based SIP MANET contains at least one SIP proxy server.

Research in the field of proxy-less SIP MANET, i.e. SIP peer-to-peer over MANET without SIP servers is presented in [30]. The authors in this work propose a signaling system that is unique and is used for sessions in P2P ad hoc networks. Also, the framework proposed in [28] is enhanced in [30] by establishing a hierarchical clustering architecture. This concept is tested through computer emulations on a testbed running on eight computers. The benefit of the proposed system in [30], is that fewer overhead messages are generated when compared to [28].

Most P2P SIP over MANET approaches in the literature use resource discovery mechanisms in order to have the ability to provide SIP user location discovery. Hence, P2P SIP over MANET approaches could be also classified into P2P SIP without overlay network and P2P SIP on an overlay network.

Most approaches on SIP over MANET in the literature employ SIP register and user discovery operations in MANET. These approaches do not deal with the compatibility of their protocols in heterogeneous networks in order to support interoperability between MANET and Internet SIP users. Research solutions for Internet connected MANET environments are presented in [23, 31, 32, and 33]. The proposed solutions rely on a centralized SIP registrar/proxy that can be positioned at the Internet or at the MANET gateway. However, the centralized nature of the registrar/proxy in these solutions creates a traffic bottleneck when SIP requests are sent to the gateway. Another problem is that it creates a single point of failure in the system. With centralized architecture, users SIP binding information are stored by one or a few MANET gateways, called SIP gateways. Another function of a SIP gateway is to forward the received SIP register requests from MANET users to an external SIP registrar on the Internet.

In [27], the authors design and implement the pseudo Session Initiation Protocol (p-SIP) server. The p-SIP server is embedded in each mobile node in order to provide ad-hoc VoIP services. The contribution of this work is two-fold: first, the implemented p-SIP server is compatible with common VoIP user agents. Secondly, it integrates the standard SIP protocol with SIP presence in order to handle SIP signaling and discovery mechanism in the ad-hoc VoIP networks. One advantage is that the implementation of this work is based on real equipment. The implementation of p-SIP is done on IBM ThickPAD x32 laptops, equipped with IEEE 802.11g wireless communication. It uses the Ubuntu Linux 6.10 and Kphone 4.2 as UA which is applied on top of the embedded p-SIP server. With the implementation of the testbed and the performance measurements from the experimental setup, the authors in [27], have shown valuable analysis of the ad-hoc VoIP network.

The results of this work also demonstrate that it is possible to achieve ad-hoc VoIP services using the implemented p-SIP servers. However, the authors did not provide information on different UDP packet sizes, injection rates and contention scenarios. The work however provides information on the influence of TCP/UDP traffic that contend VoIP streams in ad-hoc networks. To improve on the work in [27], further research is needed to analyze the influence of ad hoc node density on performance and the limitation of forwarding hop counts to realize acceptable VoIP QoS in the ad-hoc network.

In [34], the authors suggest a framework for service provisioning in stand-alone MANETs. The contributions of this work provides a new model of business that is harmonized with the features of MANETs. This model allows the invocation and execution of services. It also supports the allocation system of the SIP servlets and overlay networks as a service execution environment. Any user can take part in possessing the required features since the proposed model does not have a central unit and its functionalities. The suggested functional distribution by the authors of this work deals with the number of independent units and the loose coupling.

Also in [34], the authors propose a covering network for execution of stand-alone services in MANETs based on the framework of the SIP servlets. Another contribution of this work is prototypes built to verify ideas for the model of business and the allocated system. This work attempts to prove that the model and the scheme are reasonable with a satisfactory response time. In the results presented, the covering network protocol is formally validated. Though more detailed validation would be needed.

The architecture of a MANET emulator suitable for SIP services is proposed in [35]. The proposed architecture supports real-time audio/video communication, node mobility, and peer-to-peer-type communication. The authors in [35] have developed a SIP_MANET emulator based on the proposed architecture, and it is confirmed that solid communication quality can be maintained with SIP applications. Communication quality evaluation is also conducted to confirm the effectiveness of the simulator. To make achievable usage of the MANET emulator for verifying a SIP application, it is suggested the capabilities to translate the IP address and port numbers be incorporated to give priority to AODV packets, and to process transmission/reception of packets in multiple threads.

When the nodes are stationary, the percentage of successful audio and video communication in a SIP application is approximately 95%. The communication quality in this case is satisfactory. However, when the nodes are mobile, this percentage drops to approximately 77%. It must be noted that multi-path protocols have not been taken into consideration and are not included in the test simulations presented in [35]. Therefore, to enhance the quality of communication when the nodes are mobile, further research on multi-path protocols is needed here.

An innovative Peer-to-Peer (P2P) framework for SIP on MANET is presented in [36]. The focus here is on distributed P2P resource lookup mechanisms for SIP that tolerate failures resulting from node mobility. The authors of [36] propose a novel P2P lookup architecture based on a Structured Mesh Overlay Network (SMON) that enables P2P applications to perform fast resource lookups in the MANET environment. Their approach extends the traditional SIP user location discovery. It utilizes DHT in SMON in order to distribute SIP object identifiers over SMON. In the simulation conducted, the results show that SIPMON provides the lowest call setup delay when compared with the existing broadcast-based approaches. In addition, a new OLSR Overlay Network (OON) is proposed in [36]. The OON is a single overlay network that contains MANET nodes and nodes on the Internet.

The testbed experiment results show that extended SIPMON (SIPMON+) gives better performance in terms of call setup delay and handoff delay when compared with MANET for Network Mobility. Another contribution made in [36] is a proof-of-concept and prototype of P2P multimedia communication based on SIPMON+ for post-disaster recovery missions. This concept is evaluated with experimentation in real disaster situations – Vehicle to Infrastructure scenarios - and it is concluded that the proposed prototype outperforms MANEMO-based approaches in terms of packet loss, call setup delay, and deployment time. The proposed framework in [36] can be easily implemented with the day-to-day growth in Internet connectivity. It will be interesting to see more research in this direction to address the issue of how TCP-based applications can be provided on SIPMON+. Session mobility is one issue that need to be investigated and addressed.

B. Implementation of Multiple SIP Servers over MANET

The authors in [37] propose a distributed protocol called AdSIP that allows SIP implementation in MANETs. This protocol is evaluated on the network simulator ns-2 where comparison is made with the Tightly Coupled Approach (TCA) using metrics such as average session establishment time, failure rate and consumed bandwidth. The evaluation shows that the proposed protocol in [37] has improved performance in terms of adaptability and scalability to node mobility. The proposed solution in [37] chooses a group of nodes that are mobile to act as SIP servers, and they establish a virtual infrastructure as overlay on top of the physical network. A new distributed algorithm is built to construct the topology and to assign dynamically previously explained functionality to a group of nodes in the network. The simulation results obtained using the ns-2 simulator clearly show that the proposed AdSIP protocol is well-adapted to mobile ad hoc network. AdSIP has a lower session establishment time, low control overhead and high service availability.

Apart from the results obtained using the ns-2 simulation tool, this work has not been verified using real results that could be obtained in a real life scenario. Proactive route optimization in SIP mobility is introduced in [38]. The authors' motivation for this work is to achieve latency reduction in session setup. In the proposed Session Initiation Protocol – Proactive Route Optimization (SIP-PRO), the mobility binding information is pre-fetched and used for session establishment during the location registration step. Using the proactive route optimization, reduced latency in session setup is achieved by eliminating the traversal over multiple SIP servers. When a session is initiated, direct establishment of the session with the callee is possible if the caller has valid mobility binding information.

A mobility-aware pre-fetching scheme is developed where only the lower mobility binding information is selected because it is most likely that such information could be used for session establishment. Also in [38], the authors propose a new session setup procedure where mobility information with a sufficient residual time is used. This work lacks extensive simulations using the developed analytical models in order to verify the proposed procedures and optimization level achieved.

C. Implementation of Single SIP Servers over MANET

In [39], an intelligent VoIP system with embedded pseudo SIP server in an ad-hoc network is proposed and implemented. The embedded pseudo SIP server presented in this work is compatible with common VoIP user agents using SIP. It acts like middleware between the application and the transport layer. The quality of the VoIP service is evaluated based on the transmission delay for signaling and voice packets. Based on conducted testbed experiments, the results show that an acceptable level of VoIP service quality is achieved. The pseudo SIP server utilizes SIP presence to discover the mobile device and exchange the signaling over an ad-hoc network. This work however lacks some performance metrics such as transmission delay in the experimental results to confirm the quality of the proposed SIP server.

A SIP-based mobile network architecture for Network Mobility (NEMO) in vehicular applications is developed in [40]. The focus of this work is on developing a MANET where the hosts are mobile. Hosts can be either in a vehicle or in a group of vehicles. The MANET is linked to a SIP-based Mobile Network Gateway (SIP-MNG) which connects to the outside. The SIP-MNG is equipped with external wireless interfaces and internal 802.11 interfaces. The SIP-MNG supports call admission control and resource management for the MHs. A boost mechanism with message service that is short has been proposed by the authors. The purpose is to wake up the wireless interfaces in an on-demand manner. The Signaling details of this mechanism is presented in [40]. Additionally, this system is completely well-matched with the SIP standards that are accessible. The prototyping practice and the outcomes of the performance measurements are also presented.

The proposed system saves internet access cost by allowing the sharing of one interface for multiple sessions which is beneficial for both operators and users of public transport. Furthermore, this kind of design supports group mobility where travelers in vehicles could easily access the Internet. A push mechanism which allows SIP-MNG to stay off-line when calling activity is dormant and activates SIP-MNG when there is a need is proposed. Maintaining global accessibility of users, the proposed push approach also helps in the reduction of call charges and energy. From the presented experimental results, it is demonstrated that for PHS, WCDMA, and 802.11 networks, it is possible for multiple stations to share one interface. Based on the proposed push mechanism, the call setup time is around 20s. The push server is also designed to select the session temporarily and to use the REFER scheme in order to transmit the session to the client within SIP-MNG. The downside of the proposed mechanism is that the lengthy delay in reconnection time of the wireless interface. Further research is needed in this direction to reduce the reconnection time.

Converting IP addresses, port number and rewriting SIP messages is required in order to enable a MANET emulator to provide SIP services. However, disruptions may arise between SIP clients, and real-time performance can decline. The authors of [41] propose an architecture for a MANET emulator and local multipath routing appropriate for SIP services. A SIP_MANET emulator is developed and the correct operation of the SIP-VoIP call has been verified. The proposed routing method provides high probability of retaining the required path. The developed system is well described and the evaluation results are presented in detail.

The proposed routing method is compared with AODV and the disjoint multipath routing, using the MANET emulator and the described evaluation model. A measurement of the call holding time is taken. Call holding time is defined as the time from the start of the call to the disconnection of the call is measured. Path retaining probability is also calculated and the effectiveness of the proposed local multipath routing is verified. The proposed routing method uses a spare path when some node in the used path fails. This is the reason why its path retaining probabilities are higher than that of AODV. It would be very useful if the proposed local multipath routing in [41] is compared with AODV on a variety of network models in order to have more detailed results in this domain.

The authors of [42] propose SIPHoc, a middleware infrastructure for session establishment and management in MANETs. SIPHoc is designed to be independent of the underlying network topology, and supports both mobile and static MANETs. Therefore, SIPHoc avoids the problem of having to elect nodes for specialized tasks and replacing them when conditions change. SIPHoc differs from the SIP standard in a fully decentralized implementation which does not require any centralized components, but they both provide the same interfaces.

SIPHoc is message efficient through routing message piggybacking and is independent of the routing protocols. It is also shown that SIPHoc does not impose any topology allowing seamless interaction with the Internet. The architecture, the implementation and performances of SIPHoc are evaluated in [42]. The results show that SIPHoc has a message efficient system and provides a low dial-to-ring delay. In addition, SIPHoc allows the usage of SIP-based applications in MANETs without modification. To support this claim, the authors in [42] show how SIPHoc supports VoIP conversations within MANET, between the end-points and the MANET on the Internet. A VoIP application is used in the evaluation of the performance of SIPHoc to prove that the resulting overhead is near the optimum and comparable with the results of the standard operations on MANETs.

Two approaches enabling SIP-based session setup in ad hoc networks are proposed in [43]. One of them is a loosely coupled method, where endpoint discovery of SIP is decoupled from the procedure of routing. The other approach is the tightly coupled method, which incorporates the endpoint discovery with a cluster supported routing protocol. This protocol is fully distributed and constructs a virtual topology for effective routing. Evaluation through simulation show that the tightly coupled method achieves improved results in terms of latency of the session setup of SIP over static multihop wireless networks compared with the loosely coupled method. On the other hand, results show that the loosely coupled method generally has improved performance in networks that are characterized with random node mobility.

In [43] the authors highlight the problem relating to basic deployment over ad hoc networks and propose solutions for the integration of ad hoc routing protocols with SIP. The use of SIP supported applications for ad hoc networks are not addressed in [43]. However, essential SIP supported session setup for the applications is provided in [43] with no consideration for special applications such as SIP supported conferencing application. Further research is needed to address issues such as load balancing methods and the design and deployment of SIP supported applications.

D. IMS-based SIP Signaling Implementations over MANET

IP Multimedia System (IMS) is a developing technology with enormous potential for its usage in MANETs. IMS offers a multimedia Internet experience for different kinds of users using various applications in a mobile environment. The deployment of IMS over MANETs and modern wireless and mobile networks has brought to the fore a plethora of needs and challenges. IMS uses a number of protocols, but its driving force is founded on the SIP. IMS [44] is a 3GPP/3GPP2 standard architecture for the Next Generation Networks (NGN). The goal of this system is to fill the gap that exists between the cellular and the Internet worlds. Hence, IMS offers operators the benefit of the interoperability and quality of telecoms and the modern progress of the Internet [45]. According to the work presented in [46], IMS proposes a SIP servlets-based application server. However, exploiting this technique in MANETs for service provisioning requires a signaling layer. SIP servlets as an option are the best alternative according to the proposed SIP-based architecture for signaling in MANETs in [47].

Three main entities are related to the service provisioning in IMS: HSS, CSCF and the SIP AS. The most important data stored in the HSS are user identities, registration information and security information. The main part represents the user profile. It resolves the services that are offered to each of the users and states the rules for triggering the services. The job of the S-CSCF is to download the user profile or its part from the HSS as soon as the user registers with that S-CSCF for the first time. The S-CSCF also evaluates the initial filter criteria and communicates with the proper application server. Connections between the HSS, the S-CSCF and the AS are achieved with standardized IMS interfaces.

III. PERFORMANCE ENHANCEMENT APPROACHES FOR SIP-BASED APPLICATIONS OVER MANET

The current performance enhancement methods for SIP-based applications over MANET vary in terms of system features, requirements, feasibility in implementation, integration with existing systems, and costs. In general, the main performance enhancement methods are related to the dynamic adjustments for SIP timers, dynamic adjustments for the routing protocol parameters, implementations for supportive signaling systems, infrastructural based solutions, or service distribution features for the system users.

The dynamic adjustments for SIP timers provide flexible implementation for SIP-based applications over different platforms. This assessment relates to theoretical studies, in reality however, the SIP adjustments need to consider the nature of the network systems that SIP signaling is working on. The wireless and mobility characteristics of MANET affect the SIP signaling performance [40, 46]. Therefore, applying the dynamic adjustments for SIP signaling systems is not a proper solution which can be applied over MANET systems unless the nature of MANET systems had considered this method.

On the other hand, the dynamic adjustments for the parameters of MANET routing protocols have shown an efficient enhancement for different implemented applications. This method depends on accommodating the routing parameters to provide the best level of service for the implemented applications [39]. SIP-based applications using this method show an enhanced level in performance for the SIP signaling and voice data transfer in general [39, 41]. This method is considered as one of the most effective performance enhancement methods. However, no efficient level of implementations has been shown for this method, especially for SIP-based VoIP over MANET for emergency and backup scenarios. The implementation for supportive signaling systems for SIP is considered as one of the effective solutions. Therefore, the SDP signaling system improves the SIP signaling performance over MANET as it supports the management features of the SIP signaling system. However, the lossy nature of MANET is also affects the performance of SDP which increases the performance problems of SIP signaling [41, 48]. Most research studies in the literature implement SIP without SDP.

Synchronization issues between SIP and SDP protocols has been a concern especially the performance of SIP signaling in network systems that are variable in nature or mobility related in their implementation [34, 36, 40, 43]. The infrastructural based solutions use methods to enhance the SIP implementations over MANET. One of the suggested methods implements multiple SIP servers with high performance in order to support larger numbers of MANET nodes [37]. However, this method is difficult to implement for emergency or communication backup scenarios because of the required synchronization functionality between multiple SIP servers for the mobile callers [39, 46].

This method could be supported by using the IMS infrastructure since the synchronization functions are secured by its infrastructure. The P2P SIP implementations are considered as the most direct and easiest infrastructural performance enhancement solutions, as described in section I (B.1) [36]. Regardless of the QoS issues, without a central SIP server, it will be difficult to communicate with a large number of MANET-based callers [23]. Other infrastructural methods suggested in the literature include controlling the speed of nodes, limiting the hop numbers, and reducing the background traffic of other simultaneous applications [27, 37, 39].

Other research efforts suggest the use of service distribution features over the system users by scheduling the calls' setup processes. These solutions control the ability of users to initiate voice calls in certain conditions relating to the number of users and amount of bandwidth. The main purpose of these methods is to reduce concurrent calls by applying the time distribution features over the service users to increase the QoS level for the provided services [27, 49]. The merits and demerits of the reviewed performance enhancement methods vary in terms of the enhancement level and implementation requirements. However, both dynamic adjustment for SIP parameters and MANET parameters methods show a good level of performance enhancement. Thus, the most efficient method for enhancing SIP signaling performance over MANET is to qualify the SIP signaling behaviour to conform to the mechanical nature of MANET systems.

Combining both dynamic methods for SIP and MANET has a promising level of performance enhancement with lower costs and simple implementation. However, this enhancement method needs to be based on the evaluation studies for the current state-of-the-art for SIP signaling over different MANET scenarios. In addition, the implementation of these enhancement methods has not been fully investigated over clearly identified mobility models for MANET nodes. The simulation or test-bed tools used do not reflect reliable results that can be considered as reference results for the investigated methods. In addition, none of the proposed solutions in this section have considered any performance metrics for both SIP signaling systems and MANET routing parameters for the SIP-based applications over MANET.

IV. Supportive Simulation Tools for SIP-Based Applications Over Manet

The implementation of SIP signaling over MANET protocols as defined in RFC 3261 [50] is available in few number of simulation and test-beds tools. Simulation tools have been used in SIP signaling and MANET research [51]. Although the consistency of the simulation results has been careful analyzed [52]. As a result of this, comparative researches have been published in order to confirm the achieved results [53, 54]. As mentioned in [51], there are a large number discrete-event network simulators that are accessible by the MANET community [55]. From the 80 papers analyzed in [51] and Fig. 1 in [51], [56] we can conclude that the most utilized simulator in MANET research is the Network Simulator-2 (ns-2) with 43.8% of the analyzed papers. According to this study, there are up to 27.3% of self-developed simulators

The Global Mobile Simulator (GloMoSim) is used in 10% of MANET simulations, 6.3% for QualNet, OPNET® with 6.3%, CSIM simulator with 2.5% and MATLAB with 3.8%. The OMNeT++ simulator is also used for simulations in MANET. Programs in the simulator are modular in structure. The OMNET++ simulator includes delay as a function of the distance of the nodes. In the ns-2 simulator, delay is defined as a constant in the configuration file. Because of this, the same kind of parameters will give diverse results although the simulation scenario for MANET could be exactly the same in both simulators. MANET can also be simulated with the ns-3 simulator, which is an improved version of the ns-2 simulator. MANET routing protocols such as Ad-Hoc on Demand Distance Vector (AODV), Destination-Sequenced Distance Vector (DSDV), Dynamic Source Routing (DSR), Optimized Link State Routing (OLSR) can be simulated in ns-3 [57].

The most popular and widely-used network simulators among researchers in MANET and SIP signaling are the ns-2 simulator [58] and OPNET® [59]. There are significant differences at various levels between the two simulators. Consequently, to repeat the results obtained using the ns-2 simulator with the OPNET® simulator, some form of modification is required. Most of the simulation parameters used in ns-2 and OPNET® simulators are the same. However, there are parameters such as the wireless buffer size and the transmission range which are different and influence the simulation results considerably.

For example, if the 802.11 technology with 54 Mbps data rate is used, in ns-2 the default transmission range is 250 meters, while in the OPNET® simulator the default transmission range is 371 meters. Another example worth examining is the buffer size parameter. In the ns-2 simulator, the default buffer size 204,800 bits which is equal to 50 packets (where the size each packet is 512 bytes). In the OPNET® simulator, the default buffer size is 256,000 bits. The differences noticed between these simulators has significant impact on key metrics, like throughput and load.

When these key metrics are processed with the ns-2 simulator, they are computed from the Application level perspective. To be more precise, the presented load is assessed by putting the transmitted data from the application layer on the source node. On the other hand, throughput is calculated by putting up the received data from the application level at a target node. The OPNET® simulator considers metrics such as load and throughput at the MAC level, and that is a reason for two straight outcomes. The first outcome is that overhead is included which is as a result of MAC frame headers, MAC control packets headers, and network protocol headers.

The other consequence is that all the nodes in the network, not just the source and destination nodes are taken into consideration when both statistics are assessed. This means that if any node retransmits packets, the entire load is also increased even when the transmitting node is an intermediate node. In the same vein, when an intermediate node is receiving some packets, the matching cumulative throughput is incremented. Based on these facts, there are differences in the end results. This problem can be solved if a statistic like end-to-end, which is on the Application level similar to the ns-2 simulator is assessed.

Another important issue when comparing ns-2 with OPNET® is that error bars are contained in the outcomes from the OPNET® simulator corresponding to the average of 90% confidence interval. In the ns-2 simulator, the graphs do not show error bars. The reason is simply because they are not observable, although a confidence interval of 99% can be reached [53]. This inconsistency is highlighted based on a few constraints that are discovered in the Random Number Generator (RNG) of the OPNET® simulator [60] or weak points of the RNG of the ns-2 simulator [61]. From the above comparison analysis between ns-2 and OPNET®, it can be concluded that the OPNET® simulator performs better and its results more realistic. The reason is that OPNET®, uses approved and supported simulation models and in turn produces more reliable results.

V. CONCLUSION

This literature survey has shown that all or some nodes in a MANET have functionalities of SIP, more precisely a proxy and a registrar. User location in SIP could be determined dynamically inside the MANET. With this kind of architecture, SIP implementations over MANET are secured from the single point of failure problem. This is not the case for a centralized SIP architecture. Research in SIP signaling over MANET presented in [23, 28, 29, 31, and 33] lack the inclusion of terminal mobility using SIP. To address this issue, the authors in [36] considered terminal mobility, low call setup delay and fast network operation by proposing the Easy Disaster Communication (EasyDC).

However, another mobility issue, session mobility is not addressed in [36] and should motivate further research. Furthermore, in [23, 28, 29, 31, 32, and 33] every node in a MANET has a role as a SIP register/proxy. Hence, a user of SIP within the MANET can flood the entire network with SIP REGISTER requests in order to register its presence. In contrast to this kind of broadcast-supported SIP register/proxy, is the group-supported SIP register/proxy [30] where registers/proxies take up roles as clusterheads only. These mechanisms utilize flooding SIP requests between the nodes. The result is high network overhead which is a big challenge with the adaptability level of these network systems. In a nutshell, the major problems of SIP signaling over MANET include SIP user lookup time, the mobility assistance of the terminal, and the interoperability between Internet users and SIP over MANET users.

This survey considered possible performance enhancement methods for SIP signaling over MANET. From the literature, it is suggested that enhancing SIP signaling performance is the most efficient method that can be considered when compared to other solutions. In addition, with efficient and simple implementation, combining the dynamic adjustment methods for SIP signaling and MANET routing parameters can improve the performance level.

In this article, we identify the following topics for future investigation to further motivate research interest in SIP signaling over MANET:

1) SIP user lookup time, the mobility assistance of the terminal, and the interoperability between Internet users and SIP over MANET users are some of the challenges SIP signalling over MANET. The use of reliable performance metrics to enhance SIP signalling performance for SIP-based applications over different platforms is still an open research issue which requires further investigation. Performance metrics need to consider the best and worst case scenarios during the dynamic implementation of the SIP signalling system over MANET

2) Security for SIP-based applications over MANET is another open research issue which needs attention by the research community. There is need to investigate how implementations will cope in the presence of security threats such as Denial of Service (DoS), Man-in-the-Middle and sniffing attacks.

References

[1] J. Rosenberg, et al. “RFC 3261: SIP: Session Initiation Protocol,” 2003, URL: http://www.ietf.org/rfc/rfc3261.txt.

[2] H. Sinnreich and A. B. Johnston, “Internet Communications Using SIP: Delivering VoIP and multimedia Services with Session Initiation Protocol,” vol. 27. Wiley, 2012.

[3] A. B. Johnston, “SIP: Understanding the Session Initiation Protocol,” Third Edition, Artech House Publishers, 2009.

[4] M. Ulvan, A. Ulvan, and R. Bestak, “IMS Signalling in LTE-based Femtocell Network,” UBICOMM 2010, The Fourth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies. 2010.

[5] G. Camarillo, “Peer-to-Peer (P2P) Architecture: Definition, Taxonomies, Examples, and Applicability,” (2009).

[6] A. B. Johnston, “SIP: Understanding the Session Initiation Protocol,” Artech House, 2009.

[7] D. Bryan, et al. “Concepts and Terminology for Peer-to-Peer SIP,” (2013).

[8] M. Handley, V. Jacobson, and C. Perkins, “SDP: Session Description Protocol,” IETF, RFC 4566, Jun 2006, URL: http://tools.ietf.org/html/rfc4566.html.

[9] M. Castro, and A. J. Kassler, “SIP Based Service Provisioning for Hybrid MANETs,” Proceedings of International Workshop on Telecommunications, Santa Rita do Sapucaí, Brazil. 2007.

[10] P. Agrawal, et al. “IP multimedia Subsystems in 3GPP and 3GPP2: Overview and Scalability Issues,” Communications Magazine, IEEE 46.1 (2008): 138-145.

[11] X. Hong, K. Xu, and M. Gerla, “Scalable Routing Protocols for Mobile Ad Hoc Networks,” 2002.

[12] S. Taneja, and A. Kush, “A Survey of routing protocols in mobile ad hoc networks,” International Journal of Innovation, Management and Technology 1.3 2010: 2010-0248, 2010.

[13] K. Khan, et al. “An Efficient DSDV Routing Protocol for Wireless Mobile Ad Hoc Networks and its Performance Comparison,” Computer Modeling and Simulation, 2008. EMS'08. Second UKSIM European Symposium on. IEEE, 2008.

[14] A. Gupta, H. Sadawarti, and A. Verma, “A Review of Routing Protocols for Mobile Ad Hoc Networks,” SEAS Transactions on Communications, 331-340, 2011.

[15] A. Boukerche, et al. “Routing Protocols in Ad Hoc Networks: A survey,” Computer Networks 55.13: 3032-3080, 2011.

[16] D. Johnson, Y. Hu, and D. Maltz, “The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4,” Vol. 260. RFC 4728, 2007.

[17] C. Perkins, E. Royer, and S. Das, “Ad hoc On-demand Distance Vector (AODV),” RFC 3561, July 2003.

[18] Z. Zhai, J. Du, and Y. Ren, “The Application and Improvement of Temporally Ordered Routing Algorithm in Swarm Network with Unmanned Aerial Vehicle Nodes,” In ICWMC 2013, The Ninth International Conference on Wireless and Mobile Communications, pp. 7-12, 2013.

[19] T. Clausen, C. Dearlove, and P. Jacquet, “The Optimized Link State Routing Protocol Version 2,” draft-ietf-manet-olsrv2-00, Work in progress, 2006.

[20] N. Sivakumar, and C. Chelliah, “Simulation and Evaluation of the Performance on Probabilistic Broadcasting in FSR (Fisheye State Routing) Routing Protocol Based on random Mobility Model in MANET,” Computational Intelligence, Communication Systems and Networks (CICSyN), 2012 Fourth International Conference on. IEEE, 2012.

[21] R. Ogier, F.Templin, and M. Lewis, “Topology Dissemination Based on Reverse-path Forwarding (TBRPF)”, IETF RFC 3684, February 2004.

[22] S. Pathak, R. Upadhyay, and U Bhatt, “An Efficient Query Packets Forward Algorithm in ZRP Protocol,” Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on. IEEE, 2014.

[23] S. Leggio, and et al. “Session Initiation Protocol Deployment in Ad-Hoc Networks: A Decentralized Approach,” 2nd International Workshop on Wireless Ad-hoc Networks (IWWAN). Vol. 5. 2005.

[24] E. Guttman, C. Perkins, J. Veizades, and M. Day, “Service Location Protocol, Version 2,” RFC 2165, 1999.

[25] A. O. Driscoll, S. Rea, and D. Pesch, “Hierarchical Clustering as an Approach for Supporting P2P SIP Sessions in Ubiquitous Environments,” IFIP/IEEE 9th International Conference on Mobile and Wireless Communications Networks (MWCN), 2007.

[26] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, “Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications,” Proceedings of the ACM SIGCOMM ‘01 Conference, San Diego, California, August 2001.

[27] L. Chang, C. Sung, S. Chiu, and Y. Lin, “Design and Realization of Ad Hoc VoIP with Embedded p-SIP Server,” Journal of Systems and Software, vol. 83, no. 12, pp. 2536–2555, 2010.

[28] H. Khlifi, A. Agarwal, J.-C. Gregoire, “A framework to Use SIP in Ad-Hoc Networks,” Canadian conference on electrical and computer engineering, 2003.

[29] L. Li, L. Lamont, “Service Discovery for Support of Real-Time Multimedia SIP Applications over OLSR MANETs,” OLSR interop & workshop, 2004

[30] C. Fu, R. H. Glitho, R. Dssouli, “A Novel Signaling System for Multiparty Sessions In Peer-To-Peer Ad Hoc Networks,” IEEE wireless communications and networking conference, 2005.

[31] M. C. Castro, A. J. Kassler, “Optimizing SIP Service Provisioning in Internet Connected MANETs,” International conference on software in telecommunications and computer networks, 2006, pp. 86-90.

[32] J. Manner, S. Leggio, K. Raatikainen, “An Internet SIP Gateway for Ad-Hoc Networks,” The 3rd annual IEEE communications society on sensor and ad hoc communications and networks, 2006, Vol. 3, pp. 740-745.

[33] P. Stuedi, M. Bihr, A. Remund, G. Alonso, “SIPhoc: Efficient SIP Middleware for Ad Hoc Networks,” Proceedings of the ACM/IFIP/USENIX 8th international middleware conference, 2007.

[34] S. Bah, R. Glitho, and R. Dssouli, “A SIP Servlets-based Framework for Service Provisioning in Stand-Alone MANETs,” Journal of Network and Computer Applications, 2012.

[35] H. Todoroki, T. Kagoshima, D. Kasamatsu, and K. Takami, “Implementation of a Peer-to-Peer-Type SIP Client Application on a MANET Emulator,” in TENCON 2012-2012 IEEE Region 10 Conference, 2012, pp. 1–6.

[36] T. Wongsaardsakul, “P2P SIP Over Mobile Ad Hoc Networks,” Institut National des Télécommunications, 2010, PhD Thesis.

[37] S. Yahiaoui, Y. Belhoul, N. Nouali-Taboudjemat, and H. Kheddouci, “AdSIP: Decentralized SIP for Mobile Ad hoc Networks,” in Advanced Information Networking and Applications Workshops (WAINA), 2012 26th International Conference on, 2012, pp. 490–495.

[38] S. Pack, P. Jeong, and Y. Kim, “Proactive Route Optimization in SIP Mobility Support Protocol,” in Consumer Communications and Networking Conference (CCNC), 2010 7th IEEE, 2010, pp. 1–2.

[39] L. Chang, C. Sung, S. Chiu, and J. Liaw, “Intelligent VoIP System in Ad-Hoc network with Pseudo SIP Server,” in Autonomic and Trusted Computing, Springer, 2008, pp. 641–654.

[40] Y. C. Tseng, J. J. Chen, Y. L. Cheng, “Design and Implementation of a SIP-based Mobile and Vehicular Wireless Network with Push Mechanism,” IEEE Transactions on Vehicular Technology, vol. 56, No. 6, November, 2007.

[41] T. Kagoshima, D. Kasamatsu, and K. Takami, “Architecture and Emulator in Ad Hoc Network for Providing P2P Type SIP_VoIP Services,” TENCON, 2011.

[42] P. Stuedi, M. Buhr, A. Remund, and G. Alonso, “SIPHoc: Efficient SIP Middleware for Ad Hoc Networks,” IFIP International Federation for Information Processing, 2007.

[43] N. Banerjee, A. Acharya, S. K. Das, “Enabling SIP-based Sessions in Ad Hoc Networks,” Wireless Networks, Springer, 2007, pp. 461-479.

[44] 3rd Generation Partnership Project, Technical Specifications Group; Services & Systems Aspects, IP Multimedia Subsystem (IMS), Stage 2, 3GPP TS 23.228 v8.7.0, December 2008.

[45] Ericsson white paper, Introduction to IMS, 284 23-8123 Uen Rrev A., March 2007.

[46] Slimane Bah, “SIP Servlets-Based Service Provisioning in MANETs,” A Thesis in the Department of Electrical and Computer Engineering, Montreal, Quebec, Canada, January 2010.

[47] C. Fu, F. Khendek, R. Glitho, “Signaling for Multimedia Conferencing in 4G: the case of integrated 3G/MANETs,” IEEE Communications Magazine, August 2006.

[48] S. Massner, C. Richter, and U. Hautzendorfer, “SIP Trunking—General Requirements for Interconnecting Enterprise Networks,” Journal of Networks 8, no. 10 (2013): 2195-2212.

[49] M. Voznak, and J. Rozhon, “Methodology for SIP infrastructure performance testing,” WSEAS Transactions on Computers 9, no. 9 (2010): 1012-1021.

[50] A. Gupta, H. Sadawarti, and A. Verma, “Review of Various Routing Protocols for MANETs,” International Journal of Information and Electronics Engineering 1.3 (2011).99, 2011.

[51] L. Hogie, P. Bouvry, and F. Guinand, “An Overview of MANETs Simulation,” Electronical Notes in Theoretical Computer Science, 150, 1, 2006, pp. 81-101.

[52] T. Andel, and A. Yasinsac, “On The Credibility Of MANET Simulations,” Computer 39, 7, July 2006, pp. 48-54.

[53] P. P. Garrido, M. P. Malumbres, C. T. Calafate, “ns-2 vs OPNET: A Comparative Study of The IEEE 802.11e Technology on MANET Environments,” Simutools ’08 Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops, 2008.

[54] G. Pandey, S. Kumar, and V. K. Patle, “Comparative Study of Ns2 and OPNET Simulator for AODV and DSR Routing Protocols in MANET,” International Journal of Engineering Research and Technology, 2, 7, July 2013.

[55] The SEACORN Project, SEA-CORN Simulation Tools, URL: http://seacorn.ptinovacao.pt/sim_tools.html.

[56] P. Sheeba and C.P. Vandama, “A New SIP-based Application Layer Protocol for VoIP in MANET,” International Journal of Engineering Research and Technology, vol. 4, issue 01, pp. 733-737, Jan., 2015.

[57] T. Henderson, ns-3 Tutorial, SIMUTools 2009, http://www.nsnam.org/tutorials.html.

[58] The Network Simulator – ns-2. http://www.isi.edu/nsnam/ns/.

[59] OPNET® Modeler 17.1 Documentation, Oct 2011 Release, OPNET® Technologies. http://www.opnet.com.

[60] M. Becker, T. L. Weerawardane, X. Li, and C. Gorg, “Extending OPNET Modeler with External Pseudo Random Number Generators and Statistical Evaluation By The Limited Relative Error Algorithm,” In Proceedings of Recent Advances in Modeling and Simulation Tools for Communication Networks and Services, 2007.

[61] M. Umlauft, and P. Reichl, “Experiences with the ns-2 Network Simulator – Explicitly Setting Seeds Considered Harmful,” Wireless Telecommunications Symposium, 2007.

An Efficient Audio Classification Approach Based on Support Vector Machines

Lhoucine Bahatti
Electrical Engineering Department
ENSET, Hassan II University
Mohammedia, Morocco

My Elhoussine Echhibat
Mechanical Engineering Department
ENSET, Hassan II University
Mohammedia, Morocco

Omar Bouattane
Electrical Engineering Department
ENSET, Hassan II University
Mohammedia, Morocco

Mohamed Hicham Zaggaf
Electrical Engineering Department
ENSET, Hassan II University
Mohammedia, Morocco

***Abstract*—In order to achieve an audio classification aimed to identify the composer, the use of adequate and relevant features is important to improve performance especially when the classification algorithm is based on support vector machines.**

As opposed to conventional approaches that often use timbral features based on a time-frequency representation of the musical signal using constant window, this paper deals with a new audio classification method which improves the features extraction according the Constant Q Transform (CQT) approach and includes original audio features related to the musical context in which the notes appear.

The enhancement done by this work is also lay on the proposal of an optimal features selection procedure which combines filter and wrapper strategies.

Experimental results show the accuracy and efficiency of the adopted approach in the binary classification as well as in the multi-class classification.

Keywords—Classification; features; selection; timbre; SVM; IRMFSP; RFE-SVM; CQT

I. Introduction

Music Information Retrieval (MIR) is a growing field that benefits signal processing development and communication media tools. It uses the pattern recognition techniques to solve problems of music digital transcription including classification process. This classification is aimed to identify the artist in a given musical track.

In order to improve the performance, many researches are focused on the choice of an efficient classification algorithm, and on the use of relevant features which are able to answer questions in query. To do so, the conventional solution is to choose a set of elements that reflect the musical timbre and reduce their dimensionality by using filters method, wrappers method or embedded strategy [1][2][3]. The purpose of this work is firstly to introduce and operate a new family of so-called transition features which characterize the musical context in which the notes appear and optimize the classification algorithm. Secondly, we improve the performances by combining features selections approaches.

The rest of this paper is organized as follows: Section 2 reviews the previous works as state of the art on the automatic audio classification algorithms and the features selection. The used classifier is presented in Section 3. Section 4 is devoted to the proposed approach which consists of extracting features and describes into details the reduction dimensionality technique. The parameters classifier and approach practical implementation results are presented in Section 5. Finally, we conclude the paper and suggest future related work in Section 6.

II. State of the Art

To achieve an audio recognition task, different classification algorithms have been designed and tested [4][5][6][7][8]. Among several approaches, it seems that a consensus has been developed around the use of Support Vector Machines (SVM) [9][10][11][7][12][13] because of their flexibility, computational efficiency, capacity to handle high dimensional data and their profits of the feature selection [11].

The features selection method has been a central topic in a variety of fields such as pattern recognition and machine learning .The objective is to find an optimal subset of relevant and not redundant features in order to guarantee classification accuracy, computational efficiency and learning convergence [13][14][15]. In filter strategy, the features selection is performed independently of the learning classifier and the Inertia Ratio Maximization using Feature Space Projection (IRMSFP) algorithm [6][13] is the most popular due to its simplicity, rapidity and efficiency. For the wrappers approach, the well-known feature selection method combined with Support Vector Machines is the Recursive Feature Elimination (RFE-SVM) algorithm [14]. It generates the ranking of features by using backward data elimination until the highest classification accuracy is obtained**.** However, the RFE-SVM is a greedy method that only hopes to find the best possible combination for classification [16] and may be biased when the features are highly correlated [17].

III. SVM Classifier

In this section, we first introduce the basic theory of the SVM for two-class classification problem, second the used strategy to solve multiclass problem and third the features selection method which is combined with Support Vector Machines.

A. Basic Theory of SVMs

Given a training set of instance-label pairs $(\mathrm{x_i}, \mathrm{y_i})_{\mathrm{i}=1,\ldots \mathrm{m}}$ where $\mathrm{x_i} \epsilon \mathbb{R}^{\mathrm{n}}$ and $\mathrm{y_i} \epsilon \{1, -1\}^{\mathrm{m}}$ .The support vector machines (SVMs)[18] are used to find a hyperplane $\mathrm{W.x} + \mathrm{b} = 0$ to separate the data with the maximum margin. This requires the solution of the following optimization problem:

$$\text{minimize } \ \mathrm{M} = \frac{1}{2}\mathrm{W^T W}$$

Subject to

$$y_i(\ W^T \emptyset(x_i) + b) \geq 1 \qquad (1)$$

Using a soft-margin instead of a hard-margin, the primal problem for SVMs is obtained:

$$\text{minimize } \ \frac{1}{2} W^T W \ + C \sum_{i=1}^{n} \xi_i \qquad (2)$$

Subject to

$$y_i(\ W^T \emptyset(x_i) + b) \geq 1 - \xi_i\ ;\ \ \xi_i \geq 0 \qquad (3)$$

Here:

- $\{\xi_i\}$ are slack variables which allow for penalized constraint violation through the penalty function $\mathrm{F}(\xi)$ which defined by (4):

$$F(\xi) = \sum_{i=1}^{n} \xi_i \qquad (4)$$

- C is the parameter controlling the trade-off between a large margin and less constrained violation
- $\emptyset(.)$ represents the mapping from the input space to the features space. However researchers prefer to use a kernel function K(.,.) given by the following expression: $\mathrm{K}(\mathrm{x_i}, \mathrm{x_j}) = \emptyset(\mathrm{x_i})^{\mathrm{T}} \emptyset(\mathrm{x_j})$. Practically, the most commonly used kernel functions are:
  - Linear: $K(x_i, x_j) = {x_i}^T x_j$ (5)
  - Polynomial: $K(x_i, x_j) = (\gamma {x_i}^T x_j + r)^d \ \ , \gamma > 0$ (6)
  - Radial basis function (RBF):

$$K(x_i, x_j) = exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \qquad (7)$$

  - Sigmoid: $K(x_i, x_j) = tanh(\gamma {x_i}^T x_j + r)$ (8)

Here, γ, r and d are kernel parameters, furthermore a practical use and implementation of the SVM classifier is presented in [19].

B. Multiclass Classification

As investigated in [20] the multiclass classification based on SVMs is commonly performed by one of the two methods “one-vs-one” or “one-vs-all”. Both consider the multi-class problem as a collection of “two-class classification” problems. For k-class classification, “one-vs-all” method constructs k classifiers where each classifier constructs a hyperplane between one class and the rest $(\mathrm{k} - 1)$classes. A majority vote or some other measures are applied over the all possible pairs for decision. For the “one-vs-one” approach, $\frac{\mathrm{K(K-1)}}{2}$ classifications are realized between each possible class pairs and similarly a voting scheme is applied for decision.

For more speed and reliability, Direct Acyclic Graph SVM (DAGSVM) [20][21] is often adopted. In this approach, the testing phase uses a rooted binary directed acyclic graph which has $\frac{\mathrm{k(k-1)}}{2}$ internal nodes and k leaves. Each node is a binary SVM of $\mathrm{i^{th}}$ and $\mathrm{j^{th}}$ classes.

Due to that hierarchical classification based on an acyclic tree structure, the DAG paradigm allows both of:

- Bringing a multiclass classification in a set of two-class classifications.
- Computing the classification process of k classes with $(\mathrm{k} - 1)$ comparisons instead of $\frac{\mathrm{k(k-1)}}{2}$ ones required for the basic “one-vs-one” approach.

C. Recursive Feature Elimination-SVM

The well-studied RFE-SVM is a feature selection algorithm for supervised classification which forms part of the wrapper method. It integrates filtering in the SVM learning process not only to evaluate each subset using SVM classifier but also to have information on each feature contribution in the separating hyperplane construction. The RFE-SVM is based on ranking all the features according to some score function and eliminating recursively one or more features with the lowest score.

According to [14], the RFE-SVM algorithm can be decomposed into four steps:

1) Train an SVM on the training set;

2) Order features using the weights of the resulting classifier;

3) Eliminate features with the smallest weight;

4) Repeat the process with the training set restricted to the remaining features.

IV. Proposed Approach

As already mentioned, the classification approach is based on three steps: using new features and timbral ones, features selection by combining filter and wrapper methods and optimizing the SVM classifier. The related block diagram is illustrated in Figure 1.

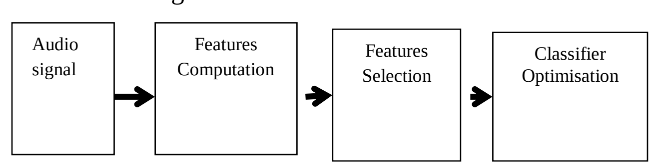


Fig. 1. Bloc diagram of our classification process

A. Features extraction

The basis of our approach focuses on the features of musical signal according to the block diagram scheme as in Figure 1. The obtained features are divided into the following two families.

- A classical family features that reflects the musical timbre and well known by the Music Information Retrieval "MIR" community [13][22][6][23][24].
- An additional and interesting family of features that reflects the transition segments between successive notes in the musical signal.

1) Features representing the musical signal timbre

The features part of this family are designed to represent the most important perceptual properties of a sound. They constitute a set of scalar parameters related to the spectral description of the musical signal. They were the subject of an extensive literature; further studies of their extraction are presented in [23]. The choice of these features, said low level, are generally depending on the desired application and on extraction duration. According to their modality of calculation as detailed in [24], they are organized as follow:

- Temporal features: convey information about the signal time evolution.
- Energetic features: features referring to various energy of the signal.
- Spectral features: those features are computed from signal time frequency representation without prior waveform model.
- Cepstral features: represent the shape of the spectrum with few coefficients using Mel-bands instead of the Fourier spectrum.
- Harmonic features: those features are computed from the detected pitch events associated with a fundamental frequency (F_0).
- Perceptual features: are computed from auditory filtered bandwidth versions of signals which aim at approximating the human perception of sounds.

The presentation of these features is illustrated by the following table No. 1

TABLE I. LIST OF FEATURES RELATED TO THE MUSICAL TIMBRE

Features list	number
Temporal Features	
Log Attack Time	1
Temporal Increase	1
Temporal Decrease	1
Temporal Centroid	1
Effective Duration	1
Signal Auto-correlation function	12
Zero-corssing rate	1
Energy Features	
Total energy	1
Total energy Modulation (frequency, amplitude)	2
Total harmonic energy	1
Total noise energy	1
Spectral Features	
Spectral centroid	6
Spectral spread	6
Spectral skewness	6
Spectral kurtosis	6
Spectral slope	6
Spectral decrease	1
Spectral rolloff	1
Spectral v ariation	3
Spectral flatness	4
Spectral crest	4
Cepstral Features	
MFCC	12
Delta MFCC	12
Delta Delta MFCC	12
Harmonic Features	
Fundamental frequency	1
Fundamental fr. Modulation (frequency, amplitude)	2
Noisiness	1
Inharmonicity	1
Harmonic Spectral Deviation	3
Odd to Even Harmonic Ratio	3
Harmonic Tristimulus	9
HarmonicSpectral centroid	6
HarmonicSpectral spread	6
HarmonicSpectral skewness	6
HarmonicSpectral kurtosis	6
HarmonicSpectral slope	6
HarmonicSpectral decrease	1
HarmonicSpectral rolloff	1
HarmonicSpectral v ariation	3
Perceptual Features	
Loudness	1
RelaitveSpecific Loudness	24
Sharpness	1
Spread	1
Perceptual Spectral centroid	6
Perceptual Spectral spread	6
Perceptual Spectral skewness	6
Perceptual Spectral kurtosis	6
Perceptual Spectral Slope	6
Perceptual Spectral Decrease	1
Perceptual Spectral Rolloff	1
Perceptual Spectral Variation	3
Odd to Even Band Ratio	3
Band Spectral Deviation	3
Band Tristimulus	9
Total Number of Features	166

In this work, we will extract the parameters of a signal related to the Oriental music known by its richness in melody [25] and generated by a lute. That signal is therefore relatively short, non-stationary and assumed to contain an almost percussive sound.

Spectral features, Cepstral features, Harmonic features, and Perceptual features are computed based on a Short Time Fourier Transform which is expressed according to (9):

$$S(t, f_K) = \int_{-\infty}^{+\infty} s(\tau - t) w(\tau) exp(-j2\pi . f_K . \tau) d\tau \qquad (9)$$

For best time-frequency localization, using a window w(t) with a variable length is more efficient [26]. Features extraction is then based on splitting the audio signal s(t) into successive frames where each frame $s_k(t)$ of index k and duration T_k is calculated by (10):

$$s_K(t) = s(t). w_K(t) \qquad (10)$$

$w_K(t)$ is the Hann window expressed by (11) :

$$W_K(t) = \begin{cases} 0.5 + 0.5\, cos\left(2\pi \frac{t}{T_K}\right); \; if \; \frac{T_K}{2} \leq t \leq \frac{T_K}{2} \\ 0 \quad elswhere \end{cases} \qquad (11)$$

According to the Constant Q Transform (CQT) approach applied in the oriental music context [26], the coefficient Q=37 and the duration T_K is given by (12)

$$T_K = \frac{Q}{f_K} \quad (12)$$

2) Transition features between notes

The majority of classical timbre features is restricted only to the spectral characteristics of short duration without modeling a music temporal aspect. Thus, in order to describe better a melody that is made for monophonic recordings, we propose, in this section, to use additional most representative features of musical timbre, taking into account the pattern musical note and the musical context in which the note appears.

So, when examining the envelopes of lute's notes, we see that they are closest to a well-known pattern modeled by the classical "ADSR" model [25] (Figure 2).

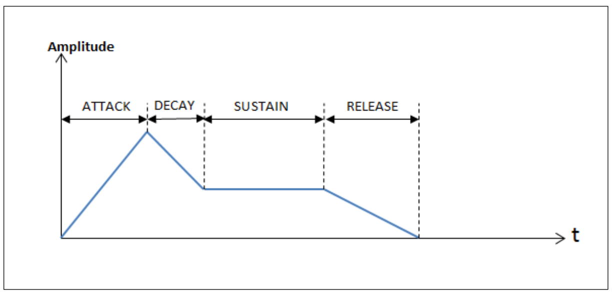


Fig. 2. Model ADSR of the note's envelope

The melody of a musical signal can only be achieved if the notes follow each other according to the sequence required by the musical score. In such a succession, energy E_{XX} of the resulting signal can be represented according to the ADSR model of Figure 3 as follows.

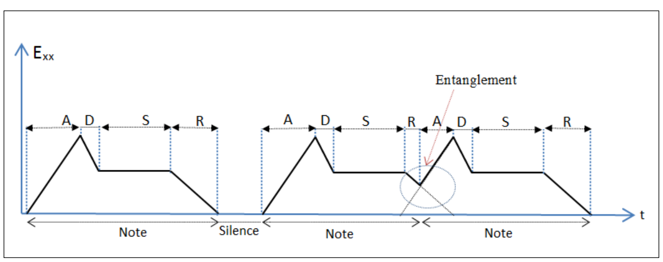


Fig. 3. Schematic view of the intra-note segmentation energy

Certainly, when two consecutive notes are separated by a "silence", the problem of intra-note correlation hardly arises, unlike the majority of cases where they have a temporal entanglement.

To characterize this phenomenon, we introduce a segment called "Transition" which starts at the beginning of the first note's Release phase and finishes at the end of Attack of the following one. We represent this articulation in Figure 4 while showing the evolution of the energy and fundamental frequency (F_0) during this new segment. We denote:

- T_c: Instant related to the energy envelope minimum.
- T_{init}: Instant of first note's Release phase beginning.
- E_{LT}: Energy of the first note at the instant T_{init}
- T_{end}: Instant of the following note Attack end
- E_{RT}: Energy of the following note at the instant T_{end}

Features associated with the transition between these types of notes are:

- The duration: T_D (13),
- The energy change during the interval T_D (14)

$$T_D = T_{end} - T_{init} \quad (13)$$

$$L_t = \frac{E_{RT} - E_{LT}}{T_D} \quad (14)$$

We also use two features introduced and investigated in [27] namely:

- The ratio between the instant T_c and the transition duration T_D (15). This parameter has proven useful when reconstructing the amplitude envelope's during transitions

$$E_{TPOS} = \frac{T_C}{T_D} \quad (15)$$

- A feature representing the legato which indicates how to link music notes together. This descriptor, whose relevance was assessed in [28], is described in (16).

$$LEG = \frac{A_1}{A_1 + A_2} = \frac{\int_{t_{init}}^{t_{end}} \big(L_t(t) - E_{XX}(t)\big)dt}{\int_{t_{init}}^{t_{end}} \big(L_t(t)\big)dt} \quad (16)$$

To compute LEG, we use the schematic view of figure 4. First, we join start and end points on the energy envelope contour using a line L_t which represents smoothest case of detachment. Then we calculate both the area A_2 below envelope energy and the area A_1 between the energy envelope and the joining line L_t. Our legato feature is finally defined as determined by (16).

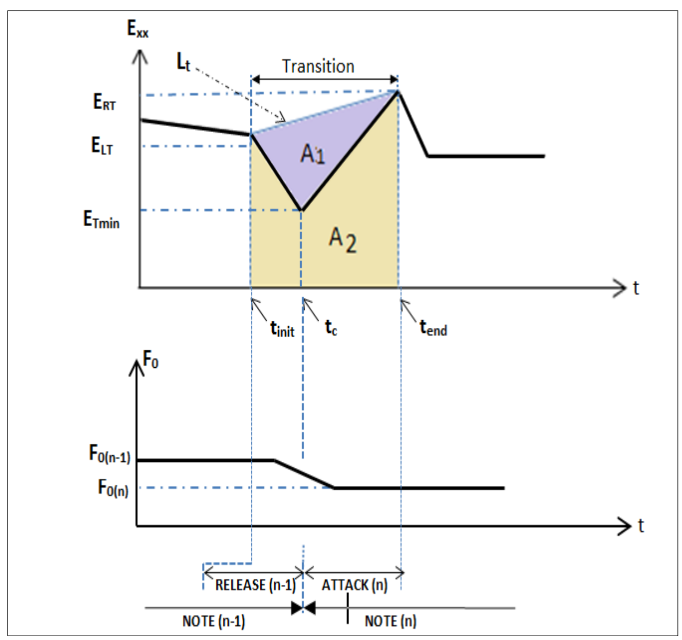


Fig. 4. Schematic view of the characterization of the "Transition" segment

B. Features selection

Since all calculated features are not relevant for the classification task, they must be previously processed by reducing dimensionality in order to keep only relevant candidates and therefore to facilitate the classification task. When the filter is adopted, the solution is very simple and fast, but features are selected based only on their intrinsic characteristics and regardless of the used classifier. In the wrapper method, the system is accurate but without any guarantee of rapid learning and features redundancy. Taking into account this fact, the adopted approach consists of combining the simplicity and the accuracy by associating both filter and wrapper methods as mentioned in the Figure 5.

In this topology, original features are firstly filtered in order to eliminate the redundant ones and the outliers. Due to this operation which selects variables as a pre-processing step, the obtained features are less correlated and a significant reduction is already performed. This will allow a better understanding of the contained information in this subset of selected features. The computing duration will be therefore short for the wrapper which selects only features that improve the prediction accuracy and optimize the classification performance.

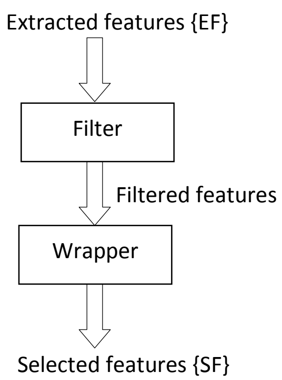


Fig. 5. Schematic view of features selection approach

To perform the filter, we choose a structure that simultaneously meets two fundamental criteria: Criterion A: Choosing a subset of informative features within each class. Criterion B: Selecting non-redundant and uncorrelated features.

For this, we adopt an algorithm based on linear discrimination analysis (LDA) strategy, called Inertia Ratio Maximization with Feature Space Projection (IRMFSP) [6][13]]. This simple and efficient filter whose relevance was assessed in [13], selects features to satisfy iteratively criterion A (Inertia Ratio Maximization) and criterion B (Feature Space Projection). The implementation of IRMFSP is composed of two steps.The first one selects at iteration l, the non-previously selected feature which maximizes the ratio between inter-class inertia and the total inertia expressed as follow (17):

$$\hat{d}^{(i)} = \underset{d}{arg\,max} \frac{\sum_{k=1}^{K} n_k(\mu_{d,k}-\mu_d)(\mu_{d,k}-\mu_d)^T}{\sum_{i=1}^{n}(f_{d,i}^{(l)}-\mu_d)(f_{d,i}^{(l)}-\mu_d)^T} \qquad (17)$$

Where n is the total number of features, n_k is the number of features belonging to class k, $f_{d,i}^{(l)}$ denotes the value of feature of index d affected to the vector i. $\mu_{d,k}$ and μ respectively denote the average value of feature d into the class k and for the total dataset. The second step of this algorithm aims at orthogonalizing the remaining feature for the next iteration as follows (18):

$$f_{d,i}^{(l+1)} = f_{d,i}^{(l)} - \left(f_d^{(l)}.g_{\hat{d}}\right)g_{\hat{d}} \qquad \forall d \neq \hat{d}^{(l)} \qquad (18)$$

Where $f_{\hat{d}}^{(l)}$ is the vector of the previously selected feature $\hat{d}^{(l)}$ and $g_{\hat{d}} = \frac{f_{\hat{d}}^{(l)}}{\left\| f_{\hat{d}}^{(l)} \right\|}$ is its normalized form.

Due to that filtering operation, the RFE-SVM wrapper will be able to measure rapidly and accurately the impact of each selected feature on the classification procedure.

V. Process Implementation

A. Materials

In this section, we are interested in the real tracks resulting from oriental music. Various artists and associated musical signals are shown in the following Table 2.

TABLE II. DATASET OF ARTISTS AND ASSOCIATED MUSICAL SIGNALS

Artist: Code	Learning database	Test database
Badr El Ouarzazi : BO	2 Tracks: BO1 : Duration : 4mn 42 s BO2 : Duration : 4 mn 52 s	8 Tracks : BO3, BO4,…BO10 : Duration: 5 mn
Farid El Attrach : FE	2 Tracks: FE1 : Duration: 5mn 30s FE2 : Duration: 4mn 48 s	8 Tracks : FE3, FE4,…… FE10 : Duration: 6 mn
Nasser Chamma : NC	2Tracks: NC1 : Duration:10 mn 30s NC2 : Duration: 7 mn 24 s	8 Tracks: NC3, NC4,…NC10 : Duration:10 mn
Nouamane Lahlou : NL	2Tracks: NL1 : Duration: 3mn 39 s NL2 : Duration :2mn 45s	2 Tracks: NL3: Duration : 4mn 31 s NL4: Duration : 3mn
Said Chraibi : SC	2 Tracks: SC1 : Duration: 5mn 10s SC2 : Duration: 5mn 24 s	8 Tracks : SC3, SC4,…… SC10 : Duration: 7 mn

As justified in [29], the most appropriate criteria function to evaluate the efficiency and accuracy of such classification process is to use the F-Measure indicator (FMSR), which is the harmonic mean of the recall (RCL) and precision (PRC). FMSR is given by (19):

$$FMSR = \frac{2.RCL.PRC}{RCL + PRC} \qquad (19)$$

B. Classifier Optimisation:

This section aims to set the classifier parameters based on binary classification. Beyond the fundamental principle of parsimony research, the SVM approach leaves, in practice, a number of options and settings to the user such as: the choice of the regularization parameter, and the choice of kernel type.

1) The kernel function:

The adopted kernel is the Exponential Radial Basis Function (ERBF) as it represents the best way to follow the non-linear decision surfaces.

For greatest robustness, instead of (7), we use a structure which takes into account the number of learning elements[18].The kernel function expression is given by (20).

$$k(x,y) = exp\left(-\frac{\|x-y\|^2}{m\sigma^2}\right) \tag{20}$$

- m is the dimension of the observation vectors:
- σ represents the width of the Gaussian function. It is a main parameter that affects the complexity of decision surfaces. Interesting choices are situated in the interval [0, 1]. The default value (before refining parameter) is σ = 1. Optimize the classifier involves determination of that parameter in order to maximize the FMSR performance on the training dataset. The obtained result is shown in Figure 6.

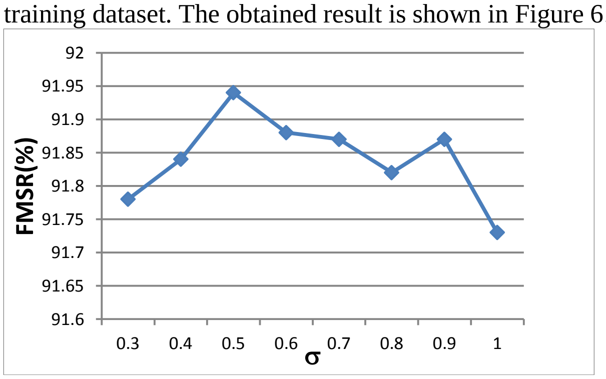


Fig. 6. Setting of SVM classifier

It can be seen that the overall performance shows a peak for σ=0.5 (F=91.94%). Nevertheless, the performance variations are very weak (about 0.24%) this justifies the robustness of the SVM classifier.

2) Controlling parameter

As mentioned in section III, this is a factor that controls the tradeoff between maximizing the margin of class's separation and minimizing of classification errors on the training set. It is a balancing parameter to set a priori, in order to make floppy the margin's SVM. Certainly, there is no strict relation to the exact calculation of C value. However the best practical results are usually obtained using an adaptive value ''C_{dat}''of that penalty parameter based on the number of « m » learning elements [18]. Thus, C_{dat} is obtained according to (21) wherein the kernel function $k(x_i, x_i)$ is defined by (20).

$$C_{dat} = \frac{1}{\frac{1}{m}\sum_{i=1}^{m} k(x_i, x_i)} \tag{21}$$

3) Performance of features selection approaches:

After the SVM optimization, we evaluate the efficiency of used features selection algorithms by comparing them according to the desired learning elements size.

The first comparison is about the complexity criterion. In this context, we set 50 optimal features and measure the processing duration for each features selection algorithm. The obtained result is presented in Table 3 below and shown clearly that the IRMFSP filter is faster than RFE-SVM wrapper. When two algorithms are used, the IRMFSP ensures well its role of preprocessing step which reduces the learning duration.

TABLE III. COMPLEXITY OF FEATURES SELECTION ALGORITHMS

Selection Features algorithm	Learning duration
IRMFSP	1,21 s
RFE-SVM	6,18 s
IRMFSP + RFE-SVM	2,34 s

The second comparison is made by F-measure criterion. Figure 7 compares that classification performance obtained as a function of the used features number.

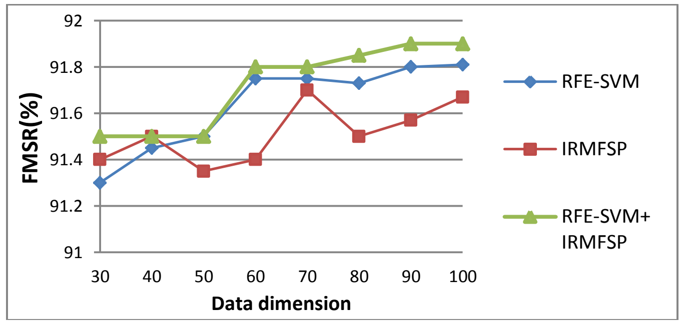


Fig. 7. Performance of the classification based on the features number

The performance of the IRMFSP filter is almost constant. The efficiency and robustness of the RFE-SVM wrapper alone or combination (RFE-SVM+ IRMFSP) are noticeable especially when the number of features is relatively high (>50).

Overall, both algorithms are less sensitive to the selected features number reduction and they perform a more reliable ranking of the most useful features by positioning those most effective at the forefront.

According to the results from Tables 6 and 7, a direct comparison between those two selection approaches (RFE-SVM vs IRMFSP) proves that obtained results are better when the two algorithms are combined regardless of filter simplicity and its efficiency.

4) Multiclass classification:

To evaluate this classification process in the multiclass context, we aim to highlight the effectiveness of the used features and their handling. The obtained results are presented as confusion matrix. (Table 4)

TABLE IV. CONFUSION MATRIX (EXPRESSED IN PERCENT) OF MULTI-CLASS CLASSIFICATION

SVM classifier : *« ERBF Kernel », ό =0.5 and C=C_{dat}*		***real classes***				
		Artist 1	*Artist 2*	*Artist 3*	*Artist 4*	*Artist 5*
estimated classes	***Artist1 (BO)***	76.25	4.6	16.9	0.87	1.39
	Artist2 (FE)	3.45	68.70	6.32	9,11	12.40
	Artist3 (NC)	17.58	5.12	74.94	2.02	0.43
	Artist4 (NL)	8.22	2.43	3.65	78.62	7.15
	Artist5 (SC)	1.22	3.28	4.81	5.20	85.50
	Overall average rate of recognition = 76.88%					

The resulting confusion matrix of dataset using 50 audio descriptors is presented in Table 4 and shows an average classification accuracy of 76,88 % where each artist is well classified with a minimal accuracy of 68,7% for the Artist N°2. These results are good and somewhat better than those described in literature [6][13] which uses only the timbral features without optimizing the SVM classifier.

VI. Conclusion and Future Works

In this paper, we have proposed a reinforcing audio recognition method by improving the extraction of well-known timbre features and taking into account features that reflect the transition between musical notes. We have also developed a practical method of the SVM classifier optimization as well as a features selection method that benefits from both of filters and wrappers advantages. In this approach, the filter has achieved a simple and straightforward features selection in order to get a subset of relevant factors most suitable to interpret and to deal with the Wrapper. Although the obtained results are interesting and encouraging, some aspects may be developed in future works. So, as perspective, we intend to investigate the use of other features kind and explore other features selection algorithms in such classification process.

References

[1] H. Liu and L. Yu, "Toward integrating feature selection algorithms for clas sification and clustering", IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 4, **(2005)**, pp. 491–502. doi>10.1109/TKDE.2005.66

[2] Verónica Bolón-Canedo · Noelia Sánchez-Maroño ·Amparo Alonso-Betanzos. A review of feature selection methods on synthetic data. Knowledge and Information Systems; Spring. March 2013, Volume 34, Issue 3, pp 483-519; ; DOI 10.1007/s10115-012-0487-8

[3] Tushar Ratanpara , Narendra Patel. Singer identification using perceptual features and cepstral coefficients of an audio signal from Indian video songs. *EURASIP Journal on Audio, Speech, and Music Processing* 2015, **2015**:16. doi:10.1186/s13636-015-0062-9

[4] Perfecto Herrera, Geoffroy Peeters, and Shlomo Dubnov. Automatic classification of musical sound. journal of New Music Research, 32:1 pages 2–21, 2003. http://dx.doi.org/10.1076/jnmr.32.1.3.16798

[5] Kaminsky, I., and A. Materka. "Automatic source identification of monophonic musical instrument sounds." Neural Networks, 1995. Proceedings, IEEE International Conference on. Vol. 1. IEEE, 1995. DOI: 10.1109/ICNN.1995.488091

[6] G. Peeters. Automatic classification of large musical instrument databases using hierarchical classifiers with inertia ratio maximization. In Proc. of AES 115th Convention, New York, USA,2003.http://recherche.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2003_AES_SoundClassification.pdf

[7] Dhanalakshmi, P. et al., Classification of audio signals using SVM and RBFNN, Expert Systems with Applications (2008); doi:10.1016/j.eswa.2008.06.126

[8] Guodong Guo and Stan Z. Li. Content-Based Audio Classification and Retrieval by Support Vector Machines. ieee transactions on neural networks, vol. 14, no. 1, january 2003. 209

[9] Guodong Guo and Stan z. Li. content-based audio classification and retrieval by support vector machines IEEE transactions on neural networks, vol. 14, no. 1, January 2003 209. http://www.ee.columbia.edu/~sfchang/course/spr/papers/guo-li-svm-audio00.pdf

[10] Qi, Zhiquan, Yingjie Tian, and Yong Shi. "Robust twin support vector machine for pattern classification." Pattern Recognition 46.1 (2013): 305-316. doi:10.1016/j.patcog.2012.06.019

[11] Moraes, R., et al. Document-level sentiment classification: An empirical comparison between SVM and ANN. Expert Systems with Applications (2012), http://dx.doi.org/10.1016/j.eswa.2012.07.059

[12] Rocamora, M., & Herrera, P. (2007, September). Comparing audio descriptors for singing voice detection in music audio files. In Brazilian Symposium on Computer Music, 11th. San Pablo, Brazil (Vol. 26, p. 27).

[13] Dominique Fourer, Jean-Luc Rouas, Pierre Hanna, Matthias Robine. automatic timbre classification of ethnomusicological audio recordings. International Society for Music Information Retrieval Conference (ISMIR 2014), Oct 2014, Taipei, Taiwan. **2014**.

[14] Pember A. Mundra and J. C. Rajapakse, "SVM-RFE with relevancy and redundancy criteria for gene selection, " in PRIB, J. C. Rajapakse, B. Schmidt, and L. G. Volkert, Eds., vol. 4774, Springer, **(2007)**, pp. 242–252.

[15] Y. Tang, Y.-Q. Zhang and Z. Huang, "Development of two-stage SVM-RFE gene selection strategy for microarray expression data analysis", IEEE/ACM Trans. Comput. Biology Bioinform, vol. 4, no. 3, **(2007)**, pp. 365–381.

[16] Mouhamadou Lamine Samb ; Fodé Camara ; Samba Ndiaye; Yahya Slimani and Mohamed Amir Esseghir; A Novel RFE-SVM-based Feature Selection Approach for Classification ; International Journal of Advanced Science and Technology; Vol. 43, June, 2012.

[17] Ke Yan; David Zhang. Feature selection and analysis on correlated gas sensor data with recursive feature elimination; Elsevier. Sensors and Actuators B 212(2015); 353–363; DOI: 10.1016/j.snb.2015.02.025

[18] T.Joachims. Svm light support vector machine. http ://svmlight.joachims.org

[19] Chang, Chih-Chung, and Chih-Jen Lin. "LIBSVM: A library for support vector machines." ACM Transactions on Intelligent Systems and Technology (TIST) 2.3 (2011): 27. doi>10.1145/1961189.1961199

[20] Chih-Wei Hsu and Chih-Jen Lin. A Comparison of Methods for Multiclass Support Vector Machines. IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 13, NO. 2, MARCH 2002

[21] Mostafa Sabzekar, Mohammad GhasemiGol, Mahmoud Naghibzadeh, Hadi Sadoghi Yazdi . "Improved DAG SVM: A New Method for Multi-Class SVM Classification". International Conférence on Artificial Intelligence ; ICAI'09. July 13-16, 2009, Las Vegas Nevada, USA.

[22] Guodong Guo and Stan Z. Li. Content-Based Audio Classification and Retrieval by Support Vector Machines. ieee transactions on neural networks, vol. 14, no. 1, January 2003. 209

[23] Richard, G.; Sundaram, S.; Narayanan, S., "An Overview on Perceptually Motivated Audio Indexing and Classification," *Proceedings of the IEEE*, vol.101, no.9, pp.1939, 1954, Sept. 2013. DOI: 10.1109/JPROC.2013.2251591

[24] G. Peeters A large set of audio features for sound description (similarity and classification) in the CUIDADO project. Technical Report 2004. http://recherche.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf

[25] Lhoucine Bahatti, Omar Bouattane, Mimoun Zazoui1 and Ahmed Rebbani: « Fast Algorithm for In situ transcription of musical signals: Case of lute music». IJCSI International Journal of Computer Science Issues, Vol.10, Issue 1, No1, January 2013. Page 444. http://ijcsi.org/papers/IJCSI-10-1-1-444-452.pdf

[26] L. Bahatti, M. Zazoui, O. Bouattane, and A. Rebbani: « Short-Term Sinusoidal Modeling of an Oriental Music Signal by Using CQT Transform ». *Journal of Signal and Information Processing*, 2013, 4, 51-56. doi:10.4236/jsip.2013.41006

[27] Ramirez Rafael, Esteban Maestre, and Xavier Serra. "Automatic performer identification in commercial monophonic jazz performances." *Pattern Recognition Letters* 31.12 (2010): 1514-1523. doi:10.1016/j.patrec.2009.12.032

[28] Maestre, E., Gomez, E., 2005. Automatic characterization of dynamics and articulation of monophonic expressive recordings. In: Proc. 118th AES Convention, Barcelona, Spain.

[29] Juan José Burred and Geoffroy Peeters. An Adaptive System for Music classification and Tagging. 3rd International Workshop on Learning Semantics of Audio Signals. Proceeding pages 3-16. December 2, 2009. SAMT Conference - Graz, Austria.

Educational Data Mining & Students' Performance Prediction

Amjad Abu Saa
Information Technology Department
Ajman University of Science and Technology
Ajman, United Arab Emirates

***Abstract*—It is important to study and analyse educational data especially students' performance. Educational Data Mining (EDM) is the field of study concerned with mining educational data to find out interesting patterns and knowledge in educational organizations. This study is equally concerned with this subject, specifically, the students' performance. This study explores multiple factors theoretically assumed to affect students' performance in higher education, and finds a qualitative model which best classifies and predicts the students' performance based on related personal and social factors.**

Keywords—Data Mining; Education; Students; Performance; Patterns

I. INTRODUCTION

Educational Data Mining (EDM) is a new trend in the data mining and Knowledge Discovery in Databases (KDD) field which focuses in mining useful patterns and discovering useful knowledge from the educational information systems, such as, admissions systems, registration systems, course management systems (moodle, blackboard, etc…), and any other systems dealing with students at different levels of education, from schools, to colleges and universities. Researchers in this field focus on discovering useful knowledge either to help the educational institutes manage their students better, or to help students to manage their education and deliverables better and enhance their performance.

Analysing students' data and information to classify students, or to create decision trees or association rules, to make better decisions or to enhance student's performance is an interesting field of research, which mainly focuses on analysing and understanding students' educational data that indicates their educational performance, and generates specific rules, classifications, and predictions to help students in their future educational performance.

Classification is the most familiar and most effective data mining technique used to classify and predict values. Educational Data Mining (EDM) is no exception of this fact, hence, it was used in this research paper to analyze collected students' information through a survey, and provide classifications based on the collected data to predict and classify students' performance in their upcoming semester. The objective of this study is to identify relations between students' personal and social factors, and their academic performance. This newly discovered knowledge can help students as well as instructors in carrying out better enhanced educational quality, by identifying possible underperformers at the beginning of the semester/year, and apply more attention to them in order to help them in their education process and get better marks. In fact, not only underperformers can benefit from this research, but also possible well performers can benefit from this study by employing more efforts to conduct better projects and research through having more help and attention from their instructors.

There are multiple different classification methods and techniques used in Knowledge Discovery and data mining. Every method or technique has its advantages and disadvantages. Thus, this paper uses multiple classification methods to confirm and verify the results with multiple classifiers. In the end, the best result could be selected in terms of accuracy and precision.

The rest of the paper is structured into 4 sections. In section 2, a review of the related work is presented. Section 3 contains the data mining process implemented in this study, which includes a representation of the collected dataset, an exploration and visualization of the data, and finally the implementation of the data mining tasks and the final results. In section 4, insights about future work are included. Finally, section 5 contains the outcomes of this study.

II. RELATED WORK

Baradwaj and Pal [1] conducted a research on a group of 50 students enrolled in a specific course program across a period of 4 years (2007-2010), with multiple performance indicators, including "Previous Semester Marks", "Class Test Grades", "Seminar Performance", "Assignments", "General Proficiency", "Attendance", "Lab Work", and "End Semester Marks". They used ID3 decision tree algorithm to finally construct a decision tree, and if-then rules which will eventually help the instructors as well as the students to better understand and predict students' performance at the end of the semester. Furthermore, they defined their objective of this study as: "This study will also work to identify those students which needed special attention to reduce fail ration and taking appropriate action for the next semester examination" [1]. Baradwaj and Pal [1] selected ID3 decision tree as their data mining technique to analyze the students' performance in the selected course program; because it is a "simple" decision tree learning algorithm.

Abeer and Elaraby [2] conducted a similar research that mainly focuses on generating classification rules and predicting students' performance in a selected course program based on previously recorded students' behavior and activities. Abeer and Elaraby [2] processed and analysed previously enrolled students' data in a specific course program across 6 years (2005–10), with multiple attributes collected from the university database. As a result, this study was able to predict, to a certain extent, the students' final grades in the selected course program, as well as, "help the student's to improve the student's performance, to identify those students which needed special attention to reduce failing ration and taking appropriate action at right time" [2].

Pandey and Pal [3] conducted a data mining research using Naïve Bayes classification to analyse, classify, and predict students as performers or underperformers. Naïve Bayes classification is a simple probability classification technique, which assumes that all given attributes in a dataset is independent from each other, hence the name "Naïve". Pandey and Pal [3] conducted this research on a sample data of students enrolled in a Post Graduate Diploma in Computer Applications (PGDCA) in Dr. R. M. L. Awadh University, Faizabad, India. The research was able to classify and predict to a certain extent the students' grades in their upcoming year, based on their grades in the previous year. Their findings can be employed to help students in their future education in many ways.

Bhardwaj and Pal [4] conducted a significant data mining research using the Naïve Bayes classification method, on a group of BCA students (Bachelor of Computer Applications) in Dr. R. M. L. Awadh University, Faizabad, India, who appeared for the final examination in 2010. A questionnaire was conducted and collected from each student before the final examination, which had multiple personal, social, and psychological questions that was used in the study to identify relations between these factors and the student's performance and grades. Bhardwaj and Pal [4] identified their main objectives of this study as: "(a) Generation of a data source of predictive variables; (b) Identification of different factors, which effects a student's learning behavior and performance during academic career; (c) Construction of a prediction model using classification data mining techniques on the basis of identified predictive variables; and (d) Validation of the developed model for higher education students studying in Indian Universities or Institutions" [4]. They found that the most influencing factor for student's performance is his grade in senior secondary school, which tells us, that those students who performed well in their secondary school, will definitely perform well in their Bachelors study. Furthermore, it was found that the living location, medium of teaching, mother's qualification, student other habits, family annual income, and student family status, all of which, highly contribute in the students' educational performance, thus, it can predict a student's grade or generally his/her performance if basic personal and social knowledge was collected about him/her.

Yadav, Bhardwaj, and Pal [5] conducted a comparative research to test multiple decision tree algorithms on an educational dataset to classify the educational performance of students. The study mainly focuses on selecting the best decision tree algorithm from among mostly used decision tree algorithms, and provide a benchmark to each one of them. Yadav, Bhardwaj, and Pal [5] found out that the CART (Classification and Regression Tree) decision tree classification method worked better on the tested dataset, which was selected based on the produced accuracy and precision using 10-fold cross validations. This study presented a good practice of identifying the best classification algorithm technique for a selected dataset; that is by testing multiple algorithms and techniques before deciding which one will eventually work better for the dataset in hand. Hence, it is highly advisable to test the dataset with multiple classifiers first, then choose the most accurate and precise one in order to decide the best classification method for any dataset.

III. Data Mining Process

The objective of this study is to discover relations between students' personal and social factors, and their educational performance in the previous semester using data mining tasks. Henceforth, their performance could be predicted in the upcoming semesters. Correspondingly, a survey was constructed with multiple personal, social, and academic questions which will later be preprocessed and transformed into nominal data which will be used in the data mining process to find out the relations between the mentioned factors and the students' performance. The student performance is measured and indicated by the Grade Point Average (GPA), which is a real number out of 4.0. This study was conducted on a group of students enrolled in different colleges in Ajman University of Science and Technology (AUST), Ajman, United Arab Emirates.

A. Dataset

The dataset used in this study was collected through a survey distributed to different students within their daily classes and as an online survey using Google Forms, the data was collected anonymously and without any bias. The initial size of the dataset is 270 records. Table 1 describes the attributes of the data and their possible values.

TABLE I. Attributes Description and Possible Values

Attribute	Description	Possible Values
GENDER	Student's gender	{Male, Female}
NATCAT	Nationality category	{Local, Gulf, Arab, Non-Arab}
FLANG	First Language	{Arabic, English, Hindu-Urdu, Other}
TEACHLANG	Teaching language in the university	{English, Arabic}
HSP	High School Percentage	{Excellent (90% to 100%), Very Good (High) (85% to 89.9%), Very Good (80% to 84.9%), Good (High) (75% to 79.9%), Good (70% to 74.9%), Pass (High) (65% to 69.9%), Pass (60% to 64.9%)}
STATUS	Student status depending on his/her earned credit hours	{Freshman (< 32), Sophomore (33 - 64), Junior (65 - 96), Senior (> 96)
LOC	Living Location	{Ajman, Sharjah, Dubai, Abu Dhabi, Al-Ain, UAQ, RAK, Fujairah, University Hostel}
SPON	Does the student have	{Yes, No}

	any sponsorship	
PWIU	Any parent works in the university	{Yes, No}
DISCOUNT	Student discounts	{Yes, No}
TRANSPORT	How the student comes to the university	{Private car, Public Transport, University Bus, Walking}
FAMSIZE	Family Size	{Single, With one parent, With both parents, medium family, big family}
INCOME	Total Family Monthly Income	{Low, Medium, Above Medium, High}
PARSTATUS	Parents Marital Status	{Married, Divorced, Separated, Widowed}
FQUAL	Father's Qualifications	{No Education, Elementary, Secondary, Graduate, Post Graduate, Doctorate, N/A}
MQUAL	Mother's Qualifications	{No Education, Elementary, Secondary, Graduate, Post Graduate, Doctorate, N/A}
FOCS	Father's Occupation Status	{Currently on Service, Retired, In between Jobs, N/A}
MOCS	Mother's Occupation Status	{Currently on Service, Retired, In between Jobs, Housewife, N/A}
FRIENDS	Number of Friends	{None, One, Average, Medium, Above Medium, High}
WEEKHOURS	Average number of hours spent with friends per week	{None, Very limited, Average, Medium, High, Very High}
GPA	Previous Semester GPA	{> 3.60 (Excellent), 3.00 – 3.59 (Very Good), 2.50 – 2.99 (Good), < 2.5 – (Pass)}

Following is a more detailed description about some attributes mentioned in Table 1:

- TEACHLANG: Some majors in the university are taught in English, and some others are taught in Arabic, and hence, it is useful to know the teaching language of the student, as it might be linked with his/her performance.
- STATUS: The University follows the American credit hours system, and hence, the status of the student can be acquired from his/her completed/earned credit hours.
- FAMSIZE: The possible values of this attribute are derived from the questionnaire as: 1 is "Single", 2 is "With one parent", 3 is "With both parents", 4 is "medium family", and 5 and above is "big family".
- INCOME: The possible values of this attribute are derived from the questionnaire as: < AED 15,000 is "Low", AED 15,000 to 25,000 is "Medium", AED 25,000 to 50,000 is "Above Medium", and above 50,000 is "High".
- FRIENDS: The possible values of this attribute are derived from the questionnaire as: None is "None", 1 is "One", 2 to 5 is "Average", 6 to 10 is "Medium", 11 to 15 is "Above Medium", and above 15 is "High".
- WEEKHOURS: The possible values of this attribute are derived from the questionnaire as: None is "None", 1 to 2 hours is "Very limited", 2 to 10 hours is "Average", 10 to 20 hours is "Medium", 20 to 30 hours is "High", and more than 30 hours is "Very High.

B. Data Exploration

In order to understand the dataset in hand, it must be explored in a statistical manner, as well as, visualize it using graphical plots and diagrams. This step in data mining is essential because it allows the researchers as well as the readers to understand the data before jumping into applying more complex data mining tasks and algorithms.

Table 2 shows the ranges of the data in the dataset according to their attributes, ordered from highest to lowest.

TABLE II. RANGES OF DATA IN THE DATASET

Attribute	Range
GPA	Very Good (81), Good (68), Pass (61), Excellent (60)
GENDER	Female (174), Male (96)
STATUS	Freshman (109), Sophomore (62), Junior (53), Senior (37)
NATCAT	Arab (180), Other (34), Gulf (29), Local (23), Non-Arab (4)
FLANG	Arabic (233), Other (18), Hindi-Urdu (16), English (3)
TEACHLANG	English (248), Arabic (20)
LOC	Ajman (123), Sharjah (90), Dubai (18), University Hostel (13), RAK (11), UAQ (10), Abu Dhabi (3), Fujairah (1), Al-Ain (1)
TRANSPORT	Car (175), University Bus (54), Walking (21), Public Transport (20)
HSP	Excellent (100), Very Good (High) (63), Very Good (50), Good (High) (33), Good (19), Pass (High) (4), Pass (1)
PWIU	No (262), Yes (8)
DISCOUNT	No (186), Yes (84)
SPON	No (210), Yes (60)
FRIENDS	Average (81), High (75), Medium (67), Above Medium (27), One (13), None (7)
WEEKHOURS	Average (122), Very limited (57), Medium (40), High (21), Very High (16), None (14)
FAMSIZE	Big (232), Medium (28), With both parents (6), With Two Parents (1), Single (1)
INCOME	Medium (83), Low (70), Above Medium (54), High (27)
PARSTATUS	Married (243), Widowed (17), Separated (6), Divorced (4)
FQUAL	Graduate (144), Post Graduate (41), Secondary (37), Doctorate (20), Elementary (11), N/A (10), No Education (7)
MQUAL	Graduate (140), Secondary (60), Post Graduate (25), No Education (16), Elementary (11), Doctorate (9), N/A (8)
FOCS	Service (166), N/A (42), Retired (32), In Between Jobs (30)
MOCS	Housewife (162), Service (65), N/A (22), In Between Jobs (11), Retired (10)

Furthermore, Table 3 includes summary statistics about the dataset, which includes the mode (the value with highest frequency), the least (the value with least frequency), and the number of missing values.

TABLE III. SUMMARY STATISTICS

Attribute	Mode	Least	Missing Values
GPA	Very Good (81)	Excellent (30)	0
GENDER	Female (174)	Male (96)	0
STATUS	Freshman (109)	Senior (37)	9
NATCAT	Arab (180)	Non-Arab (4)	0
FLANG	Arabic (233)	English (3)	0
TEACHLANG	English (248)	Arabic (20)	2
LOC	Ajman (123)	Fujairah (1)	0
TRANSPORT	Car (175)	Public Transport (20)	0
HSP	Excellent (100)	Pass (1)	0
PWIU	No (262)	Yes (8)	0
DISCOUNT	No (186)	Yes (84)	0
SPON	No (210)	Yes (60)	0
FRIENDS	Average (81)	None (7)	0
WEEKHOURS	Average (122)	None (14)	0
FAMSIZE	Big (232)	With Two Parents (1)	2
INCOME	Medium (83)	High (27)	36
PARSTATUS	Married (243)	Divorced (4)	0
FQUAL	Graduate (144)	No Education (7)	0
MQUAL	Graduate (140)	N/A (8)	1
FOCS	Service (166)	In Between Jobs (30)	0
MOCS	Housewife (162)	Retired (10)	0

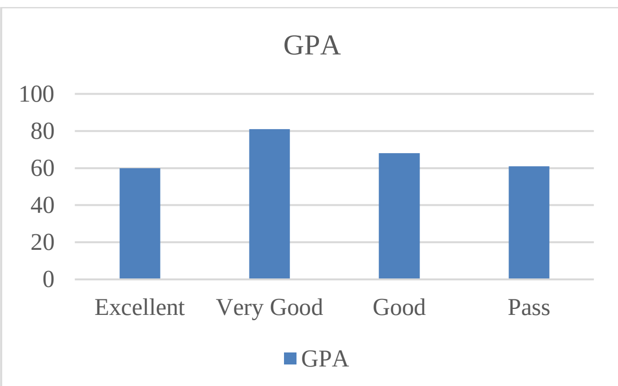


Fig. 1. Histogram of GPA attribute

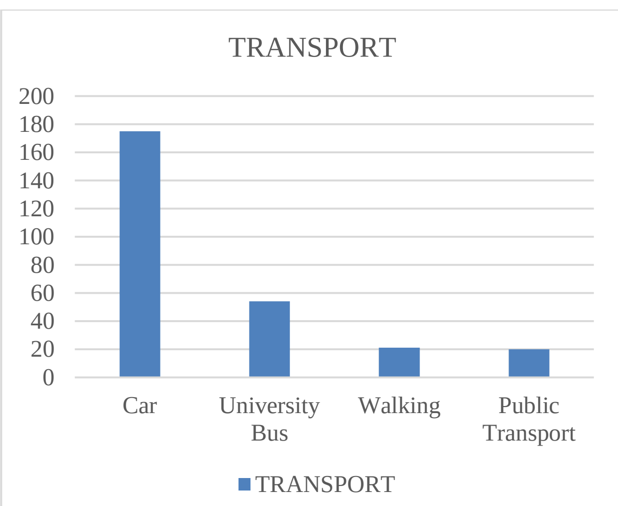


Fig. 2. Histogram of TRANSPORT attribute

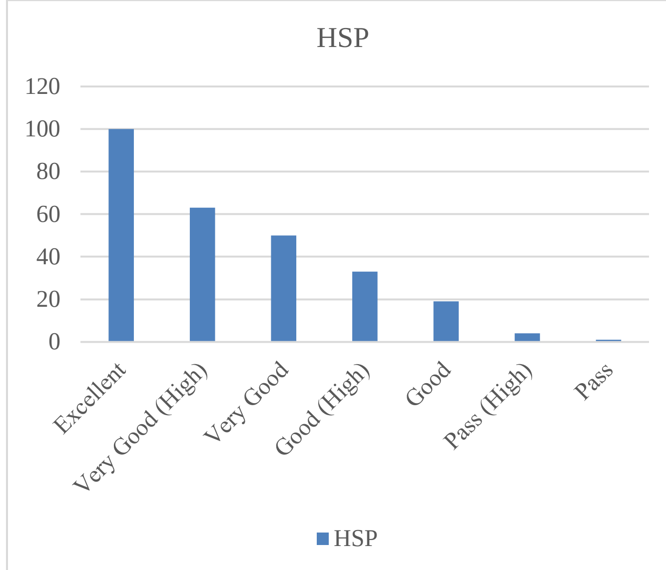


Fig. 3. Histogram of HSP attribute

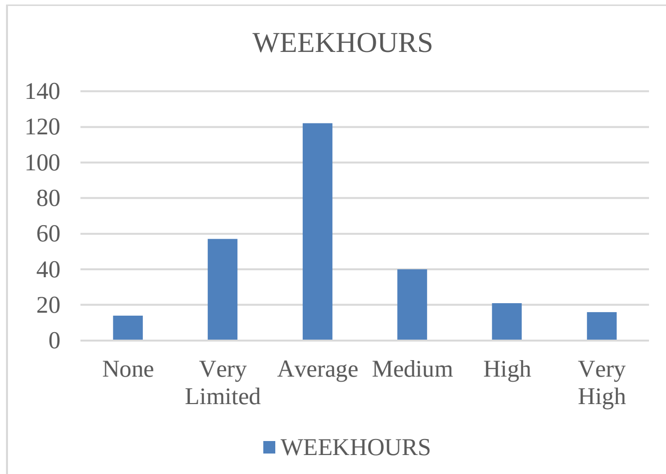


Fig. 4. Histogram of WEEKHOURS attribute

It is equally important to plot the data in graphical visualizations in order to understand the data, its characteristics, and its relationships. Henceforth, figures 1 to 4 are constructed as graphical plots of the data based on the summary statistics.

C. Data Mining Implementation & Results

There are multiple well known techniques available for data mining and knowledge discovery in databases (KDD), such as Classification, Clustering, Association Rule Learning, Artificial Intelligence, etc.

Classification is one of the mostly used and studied data mining technique. Researchers use and study classification because it is simple and easy to use. In detail, in data mining, Classification is a technique for predicting a data object's class or category based on previously learned classes from a training dataset, where the classes of the objects are known. There are multiple classification techniques available in data mining, such as, Decision Trees, K-Nearest Neighbor (K-NN), Neural Networks, Naïve Bayes, etc.

In this study, multiple classification techniques was used in the data mining process for predicting the students' grade at the end of the semester. This approach was used because it can provide a broader look and understanding of the final results and output, as well as, it will lead to a comparative conclusion over the outcomes of the study. Furthermore, a 10-fold cross validation was used to verify and validate the outcomes of the used algorithms and provide accuracy and precision measures.

All data mining implementation and processing in this study was done using RapidMiner and WEKA.

As can be seen from Table 3 in the previous section (3.2), the mode of the class attribute (GPA) is "Very Good", which occurs 81 times or 30% in the dataset. And hence, this percentage can be used as a reference to the accuracy measures produced by the algorithms in this section. Notably, in data mining, this is called the default model accuracy. The default model is a naïve model that predicts the classes of all examples in a dataset as the class of its mode (highest frequency). For example, let's consider a dataset of 100 records and 2 classes (Yes & No), the "Yes" occurs 75 times and "No" occurs 25 times, the default model for this dataset will classify all objects as "Yes", hence, its accuracy will be 75%. Even though it is useless, but equally important, it allows to evaluate the accuracies produced by other classification models. This concept can be generalized to all classes/labels in the data to produce an expectation of the class recall as well. Henceforth, Table 4 was constructed to summarize the expected recall for each class in the dataset.

TABLE IV. EXPECTED RECALL

Class (Label)	Excellent	Very Good	Good	Pass
Expected Recall	22.2%	30.0%	25.2%	22.6%

1) Decision Tree Induction

A decision tree is a supervised classification technique that builds a top-down tree-like model from a given dataset attributes. The decision tree is a predictive modeling technique used for predicting, classifying, or categorizing a given data object based on the previously generated model using a training dataset with the same features (attributes). The structure of the generated tree includes a root node, internal nodes, and leaf (terminal) nodes. The root node is the first node in the decision tree which have no incoming edges, and one or more outgoing edges; an internal node is a middle node in the decision tree which have one incoming edge, and one or more outgoing edges; the leaf node is the last node in the decision tree structure which represents the final suggested (predicted) class (label) of a data object.

In this study, four decision tree algorithms was used on the collected student's data, namely, C4.5 decision tree, ID3 decision tree, CART decision Tree, and CHAID.

C4.5 Decision Tree

The C4.5 decision tree algorithm is an algorithm developed by Ross Quinlan, which was the successor of the ID3 algorithm. The C4.5 algorithm uses pruning in the generation of a decision tree, where a node could be removed from the tree if it adds little to no value to the final predictive model.

Furthermore, the following settings was used with the C4.5 operator to produce the decision tree.

- Splitting criterion = information gain ratio
- Minimal size of split = 4
- Minimal leaf size = 1
- Minimal gain = 0.1
- Maximal depth = 20
- Confidence = 0.5

After running the C4.5 decision tree algorithm with the 10-fold cross validation on dataset, the following confusion matrix was generated.

		Actual				Class Precision (%)
		Excellent	Very Good	Good	Pass	
Prediction	Excellent	23	12	8	6	46.94
	Very Good	20	40	30	29	33.61
	Good	11	10	18	12	35.29
	Pass	6	19	12	14	27.45
Class Recall (%)		38.33	49.38	26.47	22.95	

The C4.5 algorithm was able to predict the class of 95 objects out of 270, which gives it an Accuracy value of 35.19%.

ID3 Decision Tree

The ID3 (Iterative Dichotomiser 3) decision tree algorithm is an algorithm developed by Ross Quinlan. The algorithm generates an unpruned full decision tree from a dataset.

Following are the settings used with the ID3 operator to produce the decision tree.

- Splitting criterion = information gain ratio
- Minimal size of split = 4
- Minimal leaf size = 1
- Minimal gain = 0.1

After running the ID3 decision tree algorithm with the 10-fold cross validation on the dataset, the following confusion matrix was generated.

		Actual				Class Precision (%)
		Excellent	Very Good	Good	Pass	
Prediction	Excellent	20	12	7	6	44.44
	Very Good	25	39	35	34	29.32
	Good	9	11	18	8	39.13
	Pass	6	19	8	13	28.26
Class Recall (%)		33.33	48.15	26.47	21.31	

The ID3 algorithm was able to predict the class of 90 objects out of 270, which gives it an Accuracy value of 33.33%.

CART Decision Tree

Classification and Regression Tree (CART) is another decision tree algorithm which uses minimal cost-complexity pruning.

Following are the settings used with the CART operator to produce the decision tree:

- Minimal leaf size = 1
- Number of folds used in minimal cost-complexity pruning = 5

After running the CART algorithm with the 10-fold cross validation on the dataset, the following confusion matrix was generated.

		Actual				Class Precision (%)
		Excellent	Very Good	Good	Pass	
Prediction	Excellent	43	16	10	6	57.33
	Very Good	12	40	38	26	34.48
	Good	4	10	2	6	9.09
	Pass	1	15	18	23	40.35
Class Recall (%)		71.67	49.38	2.94	37.70	

CART algorithm was able to predict the class of 108 objects out of 270, which gives it an Accuracy value of 40%.

CHAID Decision Tree

CHi-squared Automatic Interaction Detection (CHAID) is another decision tree algorithm which uses chi-squared based splitting criterion instead of the usual splitting criterions used in other decision tree algorithms.

Following are the settings used with the CART operator to produce the decision tree.

- Minimal size of split = 4
- Minimal leaf size = 2
- Minimal gain = 0.1
- Maximal depth = 20
- Confidence = 0.5

After running the CHAID algorithm with the 10-fold cross validation on the dataset, the following confusion matrix was generated:

		Actual				Class Precision (%)
		Excellent	Very Good	Good	Pass	
Prediction	Excellent	16	14	5	5	40.00
	Very Good	23	36	31	25	31.30
	Good	9	11	17	8	37.78
	Pass	12	20	15	23	32.86
Class Recall (%)		26.67	44.44	25.00	37.70	

The CHAID algorithm was able to predict the class of 92 objects out of 270, which gives it an Accuracy value of 34.07%.

Analysis and Summary

In this section, multiple decision tree techniques and algorithms were reviewed, and their performances and accuracies were tested and validated. As a final analysis, it was obviously noticed that some algorithms worked better with the dataset than others, in detail, CART had the best accuracy of 40%, which was significantly more than the expected (default model) accuracy, CHAID and C4.5 was next with 34.07% and 35.19% respectively, and the least accurate was ID3 with 33.33%. On the other hand, it was noticeable that the class recalls was always higher than the expectations assumed in Table 4, which some might argue with. Furthermore, it have been seen that most of the algorithms have struggled in distinguishing similar classes objects, and as a result, multiple objects was noticed being classified to their nearest similar class; for example, let's consider the class "Good" in the CART confusion matrix, it can be seen that 38 objects (out of 68) was classified as "Very Good", which is considered as the upper nearest class in terms of grades, similarly, 18 objects was classified as "Pass" which is also considered as the lower nearest class in terms of grades. This observation leads to conclude that the discretization of the class attribute was not suitable enough to capture the differences in other attributes, or, the attributes themselves was not clear enough to capture such differences, in other words, the classes used in this research was not totally independent, for instance, an "Excellent" student can have the same characteristics (attributes) as a "Very Good" student, and hence, this can confuse the classification algorithm and have big effects on its performance and accuracy.

2) Naïve Bayes Classification

Naïve Bayes classification is a simple probability classification technique, which assumes that all given attributes in a dataset is independent from each other, hence the name "Naïve".

"Bayes classification has been proposed that is based on Bayes rule of conditional probability. Bayes rule is a technique to estimate the likelihood of a property given the set of data as evidence or input Bayes rule or Bayes theorem is" [4]:

$$P(h_i|x_i) = \frac{P(x_i|h_i)P(h_i)}{P(x_i|h_i) + P(x_i|h_2)P(h_2)}$$

In order to summarize the probability distribution matrix generated by the Bayes model, the mode class attributes which have probabilities greater than 0.5 was selected. The selected rows are shown in Table 5.

TABLE V. PROBABILITY DISTRIBUTION MATRIX

Attribute	Value	Probability			
		Excellent	Very Good	Good	Pass
TEACHLANG	English	0.883	0.975	0.882	0.918
PWIU	No	0.950	0.963	0.971	1.000
PARSTATUS	Married	0.900	0.938	0.897	0.852
FAMSIZE	Big	0.800	0.877	0.897	0.852
FLANG	Arabic	0.833	0.802	0.912	0.918
DISCOUNT	No	0.250	0.778	0.838	0.836
SPON	No	0.867	0.753	0.721	0.787
GENDER	Female	0.733	0.691	0.676	0.459
NATCAT	Arab	0.683	0.630	0.647	0.721
TRANSPORT	Car	0.600	0.630	0.691	0.672
FOCS	Service	0.650	0.630	0.559	0.623
FQUAL	Graduate	0.550	0.580	0.456	0.541
MOCS	Housewife	0.550	0.580	0.574	0.705
MQUAL	Graduate	0.550	0.531	0.515	0.475
WEEKHOURS	Average	0.483	0.519	0.456	0.328

After the generation of the Bayes probability distribution matrix, in order to distinguish interesting probabilities from not interesting ones, a function was constructed to do that. The function calculates the absolute difference between the classes' probabilities for each row in the confusion matrix, and only if the absolute difference between two of them is more than 0.25 (25%), it will be considered as interesting, as well as, attributes with one or more class probability greater than or equal 0.35 (35%) was considered. Let's take an example to better clarify the idea; let's consider the following two rows from the generated confusion matrix.

Row	Attribute	Value	Probability				Interesting
			Excel-lent	Very Good	Good	Pass	
1	DISCOUNT	Yes	0.750	0.222	0.162	0.164	Interesting
2	TEACHLANG	English	0.883	0.975	0.882	0.918	Not Interesting

It can be seen that row 1 was considered as interesting because there are 2 probabilities greater than 0.35, and the absolute difference between some pairs of probability values are more than 0.25 (25%), hence, it is marked as interesting. Significantly, the interestingness behind the first row is that the probability of an "Honors" student to have a discount (value=Yes) is 86.7%, and it gets lower when it moves down to less GPA classes; Excellent 63.3%, Very Good 22.2%, etc... Furthermore, row 2 is considered not interesting because there are not much difference between the probabilities between the classes, even though they have high probabilities, henceforth, this attribute had almost the same probability across all types (classes) of students. Likewise, Table 6 shows all interesting probabilities found in the Bayes distribution matrix.

TABLE VI. INTERESTING BAYES PROBABILITIES

Row	Attribute	Value	Probability			
			Excellent	Very Good	Good	Pass
1	GENDER	Female	0.733	0.691	0.676	0.459
2		Male	0.267	0.309	0.324	0.541
3	HSP	Excellent	0.683	0.407	0.279	0.115
4	MOCS	Service	0.350	0.247	0.235	0.131
5	DISCOUNT	No	0.250	0.778	0.838	0.836
6		Yes	0.750	0.222	0.162	0.164

Following are the description for each one of the interesting Bayes Probabilities:

a) GENDER = Male: The probability of male students to get lower grades are significantly higher. Moving from higher to lower grades, the probability increases.

b) GENDER = Female: This scenario is opposite to the previous one, where the probability of female students to get higher grades are significantly higher. The probability decreases moving from high to low grades.

c) HSP = Excellent: Interesting enough, students who got excellent grades in High School had high grades in the university as well.

d) MOCS = Service: Interestingly, when the mother occupation status is on service, it appears that students get higher grades.

e) DISCOUNT: As illustrated earlier, students with higher grades tend to get discounts from the university more than low grades students.

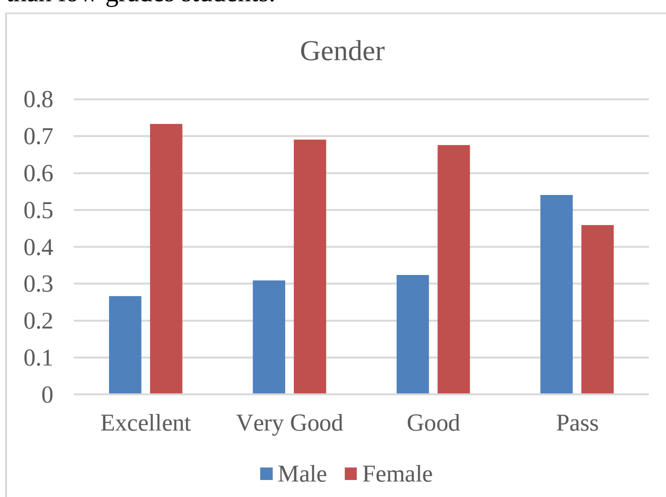


Fig. 5. Interesting probabilities of GENDER

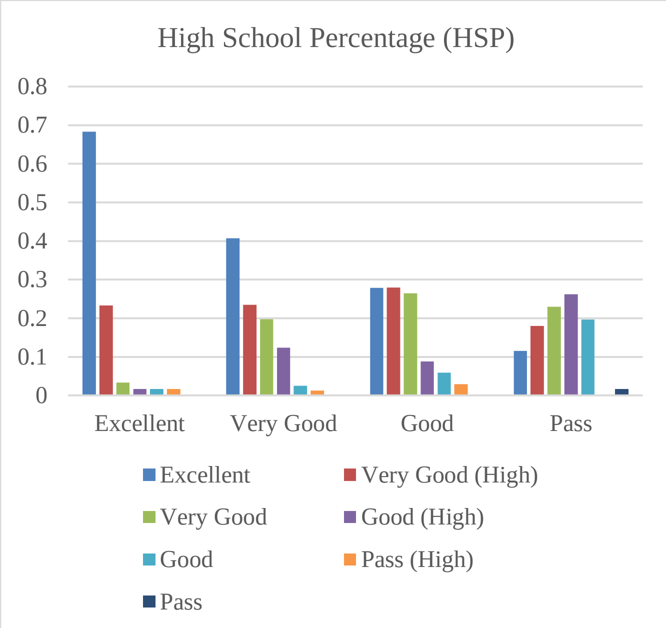


Fig. 6. Interesting probabilities of HSP

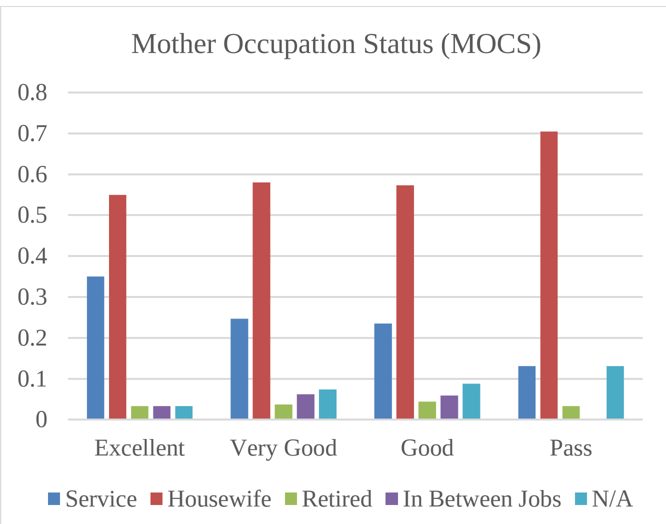


Fig. 7. Interesting probabilities of MOCS

Fig. 8. Interesting probabilities of DISCOUNT

Following is the confusion matrix of the Naïve Bayes classification model performance generated by the 10-fold cross validation:

		Actual				Class Precision (%)
		Excellent	Very Good	Good	Pass	
Prediction	Excellent	29	15	9	5	50.00
	Very Good	16	27	26	16	31.76
	Good	5	27	16	16	25.00
	Pass	5	11	14	23	43.40
Class Recall (%)		52.73	33.75	24.62	38.33	

The Naïve Bayes classifier was able to predict the class of 95 objects out of 270, which gives it an Accuracy value of 36.40%.

Analysis and Summary

In this section, a review of the implementation of the Naïve Bayes classification technique was presented on the dataset used in this research, as well as, its performance and accuracy have been tested and validated. Furthermore, this section has suggested some techniques to find interesting patterns in the Naïve Bayes model. As a final analysis, this section presented high potential results in the data mining analysis of the Naïve Bayes model, as well as, more interesting patterns could be drawn in the future from the Naïve Bayes model using other techniques.

IV. Conclusion

In this research paper, multiple data mining tasks were used to create qualitative predictive models which were efficiently and effectively able to predict the students' grades from a collected training dataset. First, a survey was constructed that has targeted university students and collected multiple personal, social, and academic data related to them. Second, the collected dataset was preprocessed and explored to become appropriate for the data mining tasks. Third, the implementation of data mining tasks was presented on the dataset in hand to generate classification models and testing them. Finally, interesting results were drawn from the classification models, as well as, interesting patterns in the Naïve Bayes model was found. Four decision tree algorithms have been implemented, as well as, with the Naïve Bayes algorithm. In the current study, it was slightly found that the student's performance is not totally dependent on their academic efforts, in spite, there are many other factors that have equal to greater influences as well. In conclusion, this study can motivate and help universities to perform data mining tasks on their students' data regularly to find out interesting results and patterns which can help both the university as well as the students in many ways.

V. Future Work

Using the same dataset, it would be possible to do more data mining tasks on it, as well as, apply more algorithms. For the time being, it would be interesting to apply association rules mining to find out interesting rules in the students data.

Similarly, clustering would be another data mining task that could be interesting to apply. Moreover, the students' data that was collected in this research included a classic sampling process which was a time consuming task, it could be better if the data was collected as part of the admission process of the university, that way, it would be easier to collect the data, as well as, the dataset would have been much bigger, and the university could run these data mining tasks regularly on their students to find out interesting patterns and maybe improve their performance.

ACKNOWLEDGMENT

I would like to express my very great appreciation to Dr Sherief Abdallah, Associate Professor, British University in Dubai, for his valuable efforts in teaching me data mining and exploration module in my Masters study, as well as, for his valuable and constructive suggestions and overall help in conducting this research.

I would also like to thank Dr. Ahmed Ankit, Assistant to the President, Ajman University of Science and Technology, and Mrs Inas Abou Sharkh, IT Manager, Ajman University of Science and Technology, for their great support during the planning and development of this research work, and for allowing me to conduct this research on the university campus.

As well as, I would like to thank and express my very great appreciation to Dr Mirna Nachouki, Assistant Professor and Head of Department, Information Systems Department, College of IT, Ajman University of Science and Technology, for her great support and inspiration during the whole time of conducting this research, as well as, her willingness to allocate her time so generously, along with:

- Mrs Amina Abou Sharkh, Teaching Assistant, Information Systems Department, College of IT, Ajman University of Science and Technology.
- Mr Hassan Sahyoun, Lecturer, Information Systems Department, College of IT, Ajman University of Science and Technology.

Last but not least, I would like to thank my wife, and my family for continuously supporting and inspiring me during my study and research.

REFERENCES

[1] Baradwaj, B.K. and Pal, S., 2011. Mining Educational Data to Analyze Students' Performance. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.

[2] Ahmed, A.B.E.D. and Elaraby, I.S., 2014. Data Mining: A prediction for Student's Performance Using Classification Method. World Journal of Computer Application and Technology, 2(2), pp.43-47.

[3] Pandey, U.K. and Pal, S., 2011. Data Mining: A prediction of performer or underperformer using classification. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (2), 2011, 686-690.

[4] Bhardwaj, B.K. and Pal, S., 2012. Data Mining: A prediction for performance improvement using classification. (IJCSIS) International Journal of Computer Science and Information Security, Vol. 9, No. 4, April 2011.

[5] Yadav, S.K., Bharadwaj, B. and Pal, S., 2012. Data Mining Applications: A Comparative Study for Predicting Student's Performance. International Journal of Innovative Technology & Creative Engineering (ISSN: 2045-711), Vol. 1, No.12, December.

[6] Yadav, S.K. and Pal, S., 2012. Data mining: A prediction for performance improvement of engineering students using classification. World of Computer Science and Information Technology Journal (WCSIT). (ISSN: 2221-0741), Vol. 2, No. 2, 51-56, 2012.

Performance Evaluation of 802.11p-Based Ad Hoc Vehicle-to-Vehicle Communications for Usual Applications Under Realistic Urban Mobility

Patrick Sondi[1], Martine Wahl[2], Lucas Rivoirard[3], Ouafae Cohin[4]
[1]ULCO, LISIC, EA 4491, F-62228 Calais, France
[2, 3]IFSTTAR, COSYS, LEOST, F-59650 Villeneuve d'Ascq, France
[1, 2]Univ Lille Nord de France, F-59000 Lille, France
[4]Hautes Etudes d'Ingénieur (HEI), F-59000 Lille, France

***Abstract*—In vehicular ad hoc networks, participating vehicles organize themselves in order to support lots of emerging applications. While network infrastructure can be dimensioned correctly in order to provide quality of service support to both vehicle-to-vehicle and vehicle-to-infrastructure communications, there are still many issues to achieve the same performance using only ad hoc vehicle-to-vehicle communications. This paper investigates the performance of such communications for complete applications including their specific packet size, packet acknowledgement mechanisms and quality of service requirements. The simulation experiments are performed using Riverbed (OPNET) Modeler on a network topology made of 50 nodes equipped with IEEE 802.11p technology and following realistic trajectories in the streets of Paris at authorized speeds. The results show that almost all application types are well supported, provided that the source and the destination have a direct link. Particularly, it is pointed out that introducing supplementary hops in a communication has more effects on end-to-end delay and loss rate rather than mobility of the nodes. The study also shows that ad hoc reactive routing protocols degrade performance by increasing the delays while proactive ones introduce the same counter performance by increasing the network load with routing traffic. Whatever the routing protocol adopted, the best performance is obtained only while small groups of nodes communicate using at most two-hop routes.**

***Keywords*—V2V; 802.11p; QoS; Urban mobility; Simulation**

I. Introduction

Vehicular ad hoc networking is an emerging paradigm where participating vehicles can exchange directly various information such as warnings, traffic conditions, and many other data. While network infrastructure can be dimensioned correctly in order to provide quality of service (QoS) support to both vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, there are still many issues to achieve the same performance using only ad hoc vehicle-to-vehicle communications. For several reasons such as the absence of infrastructure in some areas, its destruction after an accident or a disaster, or simply by opportunism, it may become necessary to rely only on ad hoc V2V communications in order to keep on providing the same services to vehicles. Performance of V2V communications is most evaluated at the level of wireless LAN based on differentiated traffic, thus regarding the characteristics of the data link and physical layers. However, due to the plethora of applications that are now available to the drivers through their smartphones connected to mobile communication technologies such as 3G/4G, V2V communications will not be developed further if the underlying technologies do not demonstrate their ability to support the same kind of applications. This work investigates the performance that can be expected from such communications, not only for differentiated traffic, but also for complete applications including their specific packet size, packet acknowledgement mechanisms, and QoS requirements. The main objective is to determine the performance that can be achieved with one of the technologies envisioned as a standard for V2V/V2I communications, namely IEEE 802.11p WAVE (Wireless Access for Vehicular Environments). For example: what performance a vehicle should expect when using a specific application ? Up to how many hops could be the vehicles sharing the application while keep a good QoS ? How many different traffic flows, of different types of service, can be involved in the same area simultaneously without degrading the performance for each application ? Particularly focused on real-world applications of vehicular ad hoc networks, this work targets evaluation of usual applications using realistic topology, mobility models, and network size while using standardized both ad hoc routing protocols and wireless LAN technologies.

The content of this paper is organized as follows. First, a related work about vehicular ad hoc networking and application is presented. Then, the system designed in order to perform simulation evaluation, including network topology, mobility models and simulation scenarios are described. Finally, the simulation results and performance analysis are reported and discussed, before the conclusion and prospective work are presented.

II. Related Work

This work has been firstly motivated by the key-role that vehicle-to-vehicle (V2V) communications will play in transportation and communication infrastructures with the growing penetration of electric vehicles [1] and the emergence of autonomous vehicle concept. On one hand, efficient resource management and service access could be achieved through vehicular cloud networks, and on the other hand, the passengers inside the vehicles could benefit of innovative applications while the vehicle will be in an almost autonomous driving mode most of the time. Several evaluations have been performed on V2V communications based on IEEE 802.11p [2][3][4][5][6]. The one presented in [2] is one of the most complete. Despite the quality of the investigations presented about the functioning of the data link and physical layers of this technology, this study does not allow catching 802.11p performance for concrete usual applications interacting with the users; that was not in the scope of the study. Thus, the second motivation of this work is to complete this part of the study of IEEE 802.11p-based vehicle-to-vehicle communication performance for usual applications. In absence of infrastructure, multi-hop V2V communications are dependent of ad hoc routing protocols: this is the only way to achieve the network management functions in a distributed manner. Despite the numerous proposals for vehicular ad hoc network routing protocols [7][8][9], the main routing protocols that are currently proceeding in the standardization process are those proposed for mobile ad hoc networks (MANETs) such as Dynamic Source Routing (DSR) [10], Ad hoc On-Demand Distance Vector (AODV) [11], Optimized Link State Routing Protocol (OLSR) [12], Topology Dissemination Based on Reverse-Path Forwarding (TBRPF) [13], and Open Shortest Path First-Overlapping Relay (OSPF-OR) [14]. It is interesting to evaluate how one protocol of each family of ad hoc routing protocols, namely reactive protocols such as AODV and proactive ones such as OLSR, affects the performance of V2V communications over IEEE 802.11p. Since they are all best effort routing protocols, it seems also interesting to evaluate some of their variants that provide an extension for quality of service (QoS) support. Quality of Service for Ad hoc Optimized Link State Routing Protocol (QOLSR) [15] is a variant of OLSR widely evaluated on applications with QoS requirements, but not yet with 802.11p. Its complete specifications are presented in [16], and its implementation is available in OPNET contributed models.

III. System Description

In a classical vehicular network infrastructure [17], Road Side Units (RSUs) are deployed along the roads and the streets in order to ensure a good coverage of the area where vehicles may request different services. Depending on the application, the communications between vehicles can transit through the RSUs or in ad hoc mode by multi-hop relaying from vehicle to vehicle. In this work, only this latter mode is evaluated due to the reasons previously mentioned in section 1.

A. *Network topology*

Intuitively, referring to the results of different researches presented in section 2, notably [2], [17] and [15], the area that may be covered efficiently in fully ad hoc mode has a limited size. Thus, the network topology considered in this study covers approximately an area of 5km x 3km dimensions.

Considering a sensing range of 1km, such an area allows reaching up to 5 hops on a straight line, which is sufficient for our study. In order to obtain accurate network density and diversity, the simulation is performed on an urban environment (Fig. 1) where the intersections of streets and roads allow avoiding particular cases such as isolated single roads or highways. The network is made of 50 nodes departed uniformly on the different routes in the streets of Paris. The same network sizes are used in similar evaluations [18], [19].

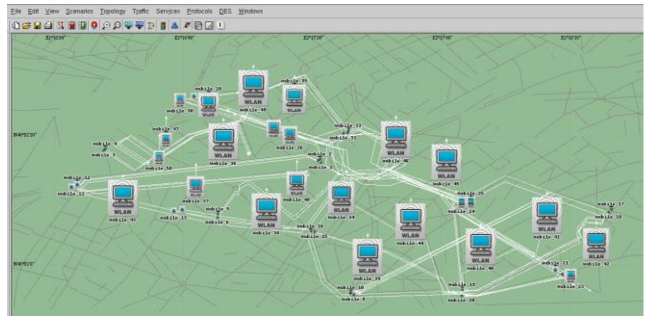


Fig. 1. Network topology for simulation

B. *Node mobility*

The nodes follow realistic trajectories defined along existing streets in Paris (Fig. 2). Up to 20 trajectories have been defined, each one followed by a pair of nodes initialized at different start point and time. Consequently, there are 10 fixed nodes departed over the area simulating stopped vehicles and 80% of mobile nodes in the network as in [18]. Each trajectory contains some steps where the node moves at 50km/h such as in long straight line, other where it moves at 30km/h such as when approaching intersections and 10km/h when turning around curves.

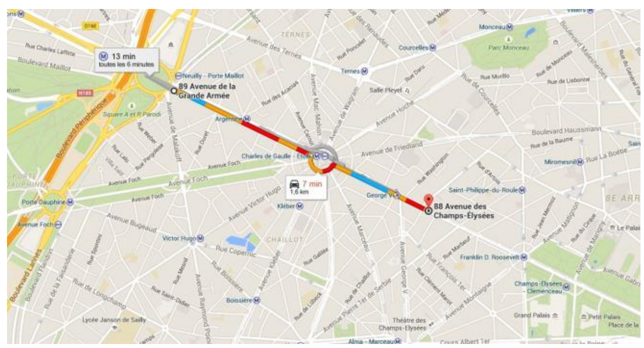


Fig. 2. Trajectory of node 28 on a real-world map (using Google maps)

In all the scenarios, each trajectory has a duration of about 6 minutes in one direction, then the node moves backward after a pausing time of 1 minute (Fig. 3). Some trajectories cover the same part of one street, in the same or in the opposite movement direction. It can be noticed that the trajectory of node 28 has the same duration in the simulation model (Fig. 3) as in a real-world itinerary (Fig. 2).

C. *Node configuration*

The nodes are equipped with IEEE 802.11p WAVE technology. As early proposed in [18], the transmit power is set at 0.02 W and receiver sensitivity at -95 dBm in order to obtain a communication range of 1km. These values were proposed before the model of IEEE 802.11p was available in Riverbed (OPNET) Modeler, and the evaluations carried out in this work confirm that they also work correctly in the official model now available. The wireless LAN configuration applied to the nodes is summarized in Fig. 4.

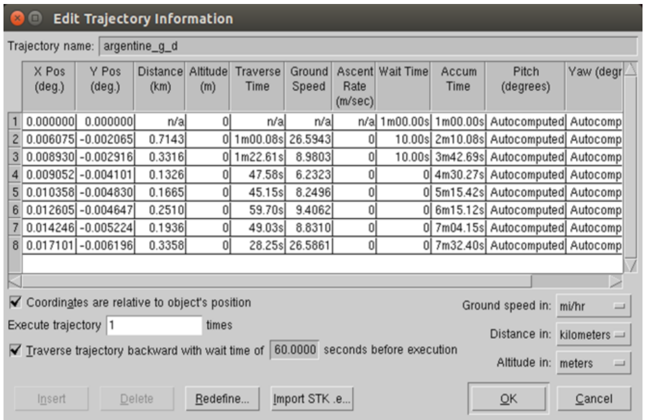


Fig. 3. OPNET model of trajectory of node 28 during simulation

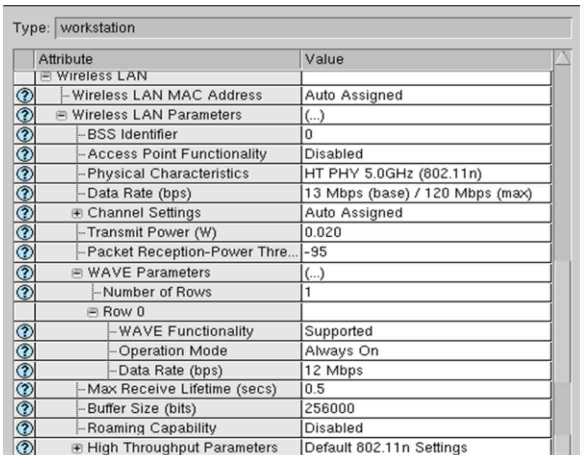


Fig. 4. Wireless LAN Parameters applied to the nodes

D. Ad hoc routing protocols

As mentioned in section 2, many routing protocols and their variants have been proposed especially for ad hoc vehicle-to-vehicle communications. However, many recent works still consider original standardized ad hoc routing protocols when evaluating the performance of V2V communications [20]. In this work, the main objective is to evaluate the performance that the users could expect for usual applications over IEEE 802.11p vehicle-to-vehicle communications independently from the routing protocol. Since an ad hoc routing protocol is mandatory to operate ad hoc vehicle-to-vehicle communications, AODV and OLSR, respectively one reactive and one proactive ad hoc routing protocols, are used in order to study the advantages and drawbacks of each family of ad hoc routing protocols on performance. Some of the applications evaluated have quality of service (QoS) requirements, but AODV and OLSR are best effort routing protocols that do not provide QoS support. In order to complete the study, two variants of OLSR with QoS support are also evaluate, namely:

- QOLSR : this routing protocol has been proposed in [15][16] as an extension of OLSR. The main idea is that each node uses the traffic received from the others in order to estimate locally the value of the metrics such as bandwidth, delay and loss on the route from each of them. This estimation is updated periodically and broadcasted through HELLO and TC messages so that all the nodes can take them into account when performing multipoint relay (MPR) selection and route computation to reach the originating node. It has been shown in [15] that QOLSR performs better than OLSR in large scale and congested ad hoc networks, particularly by maintaining a higher delivery ratio of the traffic from applications with QoS requirements such as voice and videoconference. However, the additional routing traffic introduced for QoS signaling increases the load on the WLAN thus consuming part of the resources that will lack to the applications. Previously mentioned work [15][16] realized the evaluation of QOLSR over 802.11g-based ad hoc networks and only with fixed nodes. In this work, the evaluations will be performed over 802.11p while considering realistic mobility of the vehicles.

- C231 : in order to avoid additional routing traffic introduced by QOLSR, another variant of OLSR that uses a single QoS metric instead of three is evaluated. In this variant, each node uses the location information of its neighbors in order to compute the expected received power from each neighbor and uses this information as a QoS metric. Since the network changes due to mobility, each node periodically updates the value of this metric and broadcasts it through HELLO and TC messages. The first assumption considered in this variant is that all the vehicles use the same transmit power value. The second one is that it is possible to compute accurately the value of the received power using an appropriate path loss propagation model. This latter subject has been an active area of research in recent years. Path loss arises when an electromagnetic wave propagates through space from transmitter to receiver. The power of the signal is reduced due to path distance, reflection, diffraction, scattering, free-space loss and absorption by the objects in the environment. It is also influenced by the different environment (i.e. urban, suburban and rural). The variations of the embedded transmitter and receiver antenna heights also produce losses. The losses present in a signal during propagation from transmitter to receiver may be classical and already existing. COST-231 Walfisch-Ikegami model is an extension of COST Hata model and has been proposed for urban areas [21], [22]. It can be used for frequencies above 2000 MHz when there is Line Of Site (LOS) between the transmitter node and the receiver node. It also takes into account various parameters such as the characteristics of buildings, roads and other obstacles which are important for a relevant prediction especially in urban areas. According to visibility conditions, only the case in LOS path loss calculation [23], [24] has been implemented. In this situation, there is no obstruction in the direct path between the transmitter and the receiver, and the estimated received power in dB is obtained using the equation (1):

$$L_{LOS}(dB) = 42{,}26 + 26log_{10}(d) + 20log_{10}(f) \qquad (1)$$

Where

f : Frequency of operation in MHz (5885 MHz in 802.11p standard which is greater than 2000 MHz)

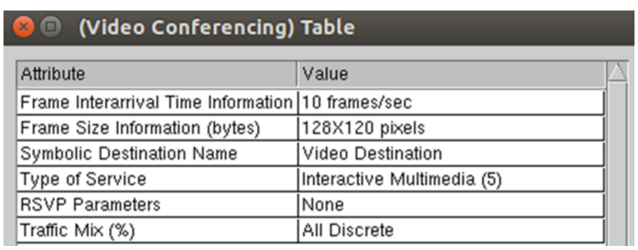
d: distance from the transmitter in kilometer

This equation is only accurate in the far-field where spherical spreading can be assumed. It is not applicable when the receiver is close to the transmitter. In this study, the nodes are spaced of 300 meters in average

E. Applications used by the vehicles

The variety of applications that can be deployed on a vehicular cloud with their specific constraints is very large. In this work, four types of applications have been modeled to which any other application could be attached, at least based on an approximation of its functioning. These applications are:

- App_1 an application generating broadcast traffic : the source sends one packet of 800 bytes every 25 milliseconds (40 packets/second) to the entire network with type of service (TOS) set to best effort;
- App_2 a safety application generating unicast traffic : the source sends one packet of 800 bytes every 25 milliseconds (40 packets/second) to a specific destination with TOS set to delay and reliability;
- App_3 a voice application : the source calls a specific destination and they start a voice session of GSM quality level;
- App_4 a videoconferencing application : the source calls a specific destination and they start a videoconferencing session (configuration in Fig. 5);


(Video Conferencing) Table

Attribute	Value
Frame Interarrival Time Information	10 frames/sec
Frame Size Information (bytes)	128X120 pixels
Symbolic Destination Name	Video Destination
Type of Service	Interactive Multimedia (5)
RSVP Parameters	None
Traffic Mix (%)	All Discrete

Fig. 5. Videoconferencing application configuration

IV. Simulation Results

The Riverbed (OPNET) Modeler is used as the modeling and simulation environment for all the evaluations realized in this work. A model of IEEE 802.11p is provided with the Modeler, and also process models of both AODV and OLSR routing protocols. A model of QOLSR [16] has been obtained in the contributed models of OPNET website from which was derived a model of C231. Each simulation session represents 12 minutes of the vehicular network functioning.

A. Simulation scenarios description

The simulation scenarios are designed in order to evaluate the performance of the applications in the context of the vehicular network described in section 3. Particularly, the behavior of both the application and the wireless LAN are studied when each routing protocol operates, and the key points that determine the performance are analyze. A description of each scenario follows:

- Scenario_1: in this scenario, node 26 (see Fig. 1) is the source of App_1 described in section III.E and it broadcasts packets through the entire network. The objective is to analyze how this traffic reaches one-hop, two-hop and, if any, farther neighbors;
- Scenario_2: in this scenario, each node in the vehicular network is a source of App_1 and broadcasts packets through the entire network. The objective is to analyze how both the wireless LAN and each routing protocol react to a great amount of generated traffic;
- Scenario_3: in this scenario, a node is the source of App_2 and sends packets to a destination located one-hop away. The objective is to evaluate the performance of a unicast traffic imposing a type of service similar to those of safety applications;
- Scenario_4: this scenario is the same as scenario_3, except that the destination is picked up three hops away and then it comes closer to the source;
- Scenario_5: in this scenario, a node is the source of App_3 and calls a destination located one-hop away. The objective is to evaluate the performance of voice conversation using vehicle-to-vehicle communication in urban mobility conditions;
- Scenario_6: this scenario is the same as scenario_5, except that the destination is picked up three hops away and then it comes closer to the source;
- Scenario_7: in this scenario, a node is the source of App_4 and calls a destination located one-hop away. The objective is to evaluate the performance of videoconferencing using vehicle-to-vehicle communication in urban mobility conditions;
- Scenario_8: this scenario is the same as scenario_7, except that the destination is picked up three hops away and then it comes closer to the source;
- Scenario_9: in this scenario, three pairs in the network are respectively the source and the destination of App_2, App_3 and App_4. The objectives are to evaluate the performance of each type of application in presence of concurrent traffic of other types, and to verify how the differentiated traffic management functionality of 802.11p is efficient according to the performance observed at the application level.

B. Result analysis

Each scenario is run several times with different seed values for the random number generator in order to avoid that the related sequence favor a particular routing protocol. The results presented and commented in this section are the average value of all the runs of the same scenario for each protocol. Consequently, the following results have been collected over a hundred simulations.

1) Results for scenario 1 : a single broadcast traffic

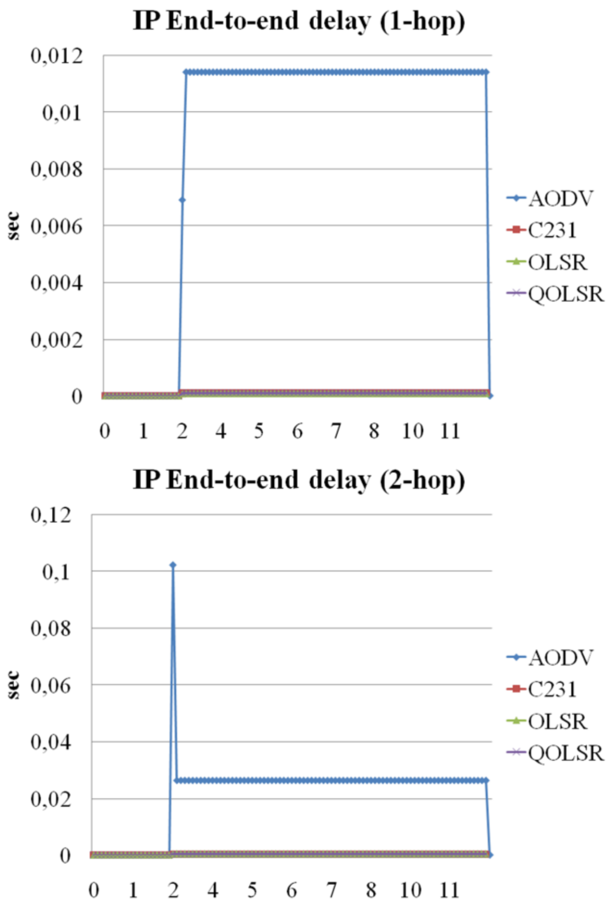


Fig. 6. End-to-end delay of broadcast traffic to 1-hop and 2-hop neighbors

In this scenario, two groups of nodes are obtained: the first is made of nodes that received the packets directly from the source (1-hop nodes), and the second that received them through a relaying node (2-hop nodes). Every nodes received correctly the packets (40 packets per second). AODV reaches the worst delay values in both cases, thus emphasizing its weaknesses in dealing with broadcast traffic. Taking into account that most applications involved in vehicular networks are broadcast-based, this is a critical issue about this protocol.

However, regarding the other protocols, the end-to-end delays achieved with 802.11p are lower than the milliseconds, thus ensuring very good performance for a safety application that requires a refresh time of 25 ms.

2) Results for scenarios 3 and 4 : a unicast application

The results presented in Fig. 7 show clearly that the application App_2 based on unicast traffic flow has good performance for a 1-hop destination. All 40 packets of 800 bytes sent per second are received with a low delay of 1.4 ms. On Fig. 8, results for progressive hop count show that the delivery ratio of application traffic is less than 50% when the nodes are 3-hop away, it approaches 70% at 2-hop and then all the packets are received when the source and destination have direct link.

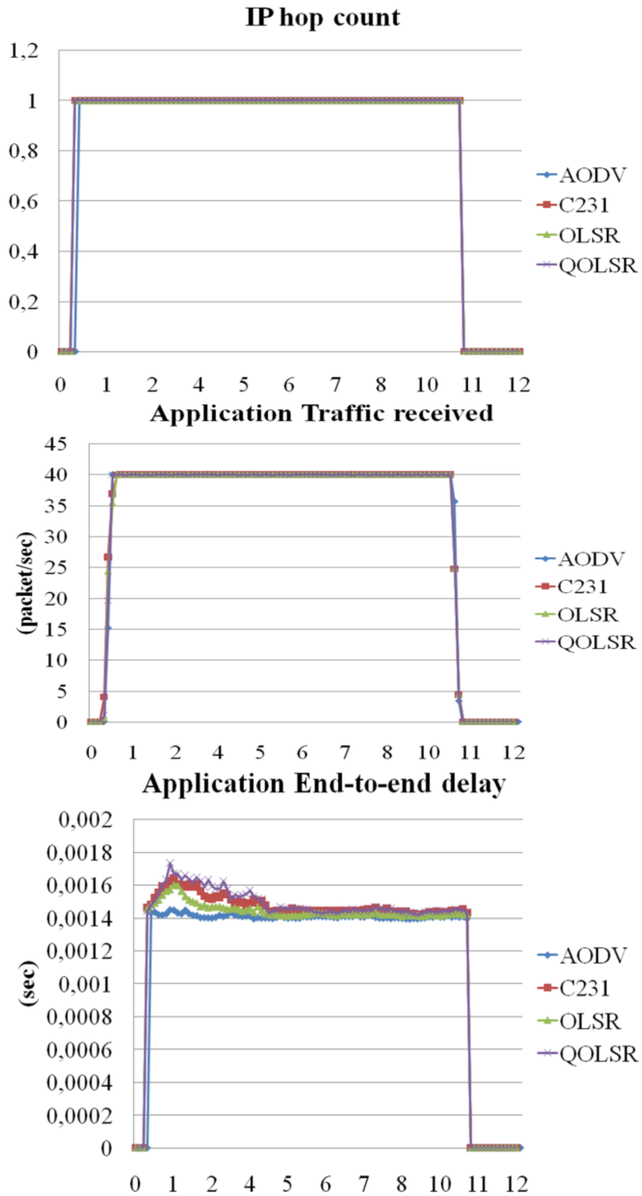


Fig. 7. Hop count, delivery, end-to-end delay for unicast traffic (scenario 3)

Many of the packets that were sent when the nodes were more than 1-hop away are not actually lost, but they are received with a delay. This explain why the destination receives up to 70 packets per second when only 40 were sent by the source. This evolution of the delivery ratio demonstrate that it is not suitable to target more than 2-hop destinations in V2V communication with 802.11p, unless the application was neither loss-sensitive nor delay-sensitive. The best-effort routing protocols, AODV and OLSR, provide better performance than QoS variants which have the worst delays and delivery ratios. Due to the overhead introduced by QoS mechanisms, such solutions should not be used when only best-effort unicast traffic are involved. For such unicast best-effort traffic flow, AODV is clearly the best routing protocol.

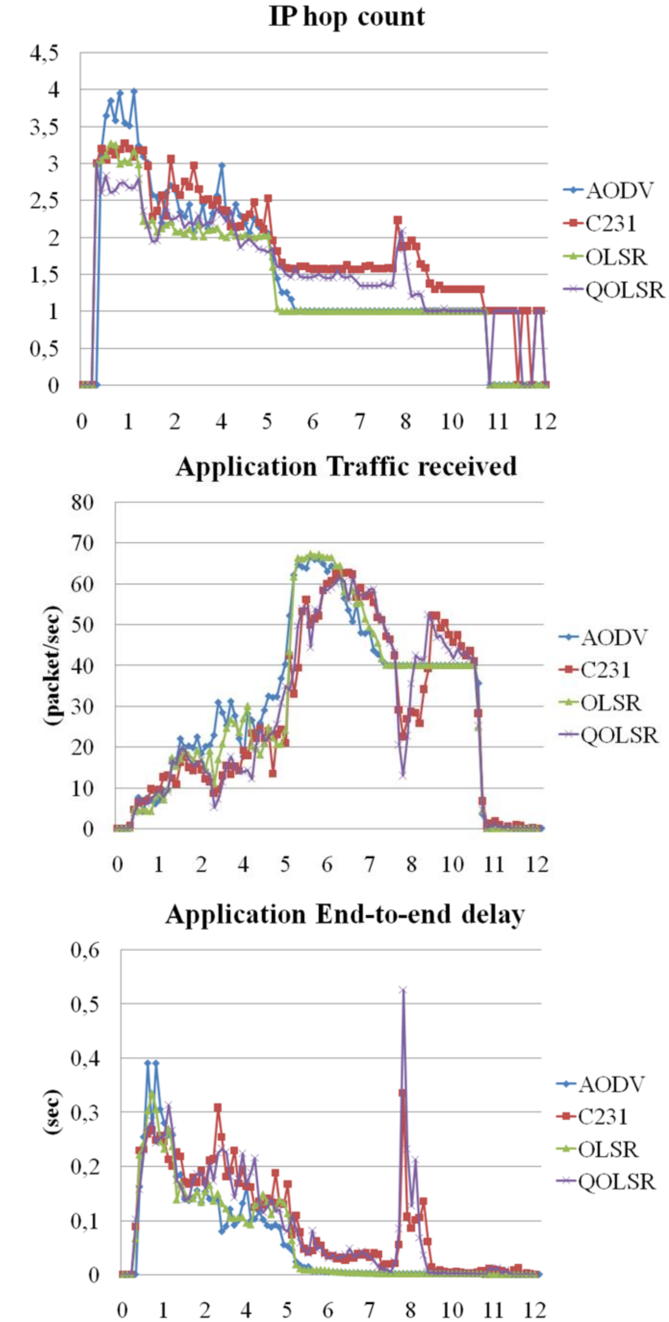


Fig. 8. Hop count, delivery, end-to-end delay for unicast traffic (scenario 4)

3) Results for scenarios 5 and 6 : a voice application

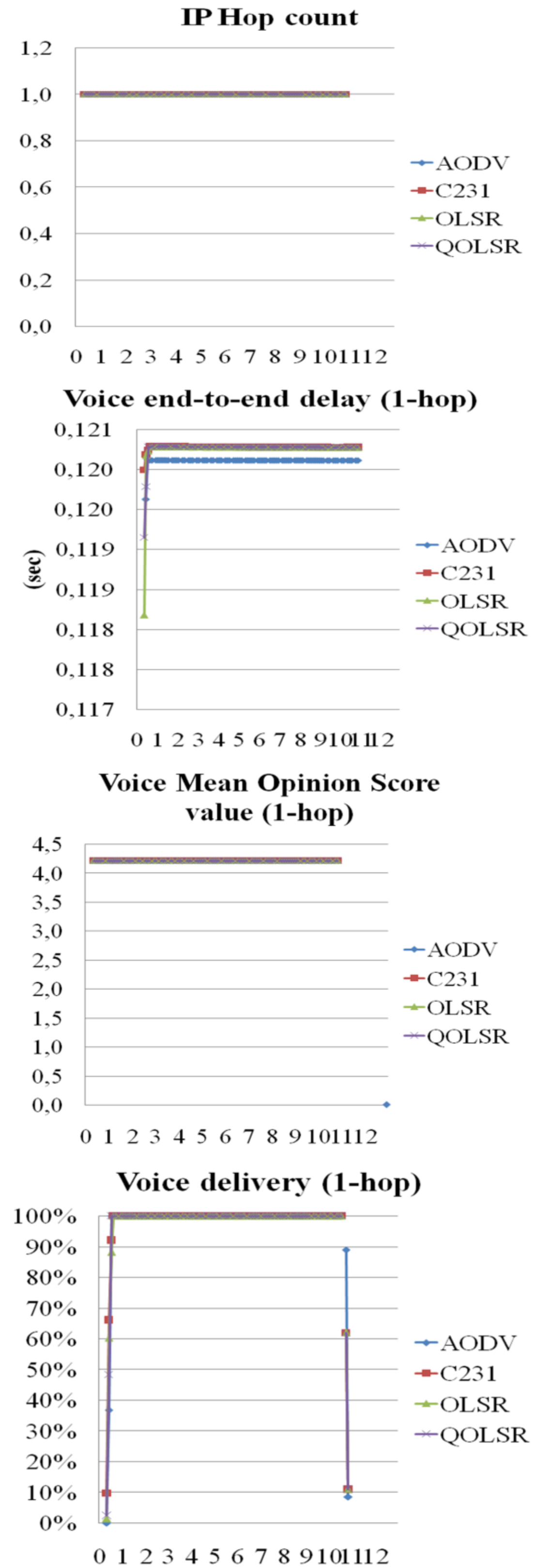


Fig. 9. Hop count, end-to-end delay, MOS, delivery for voice (scenario 5)

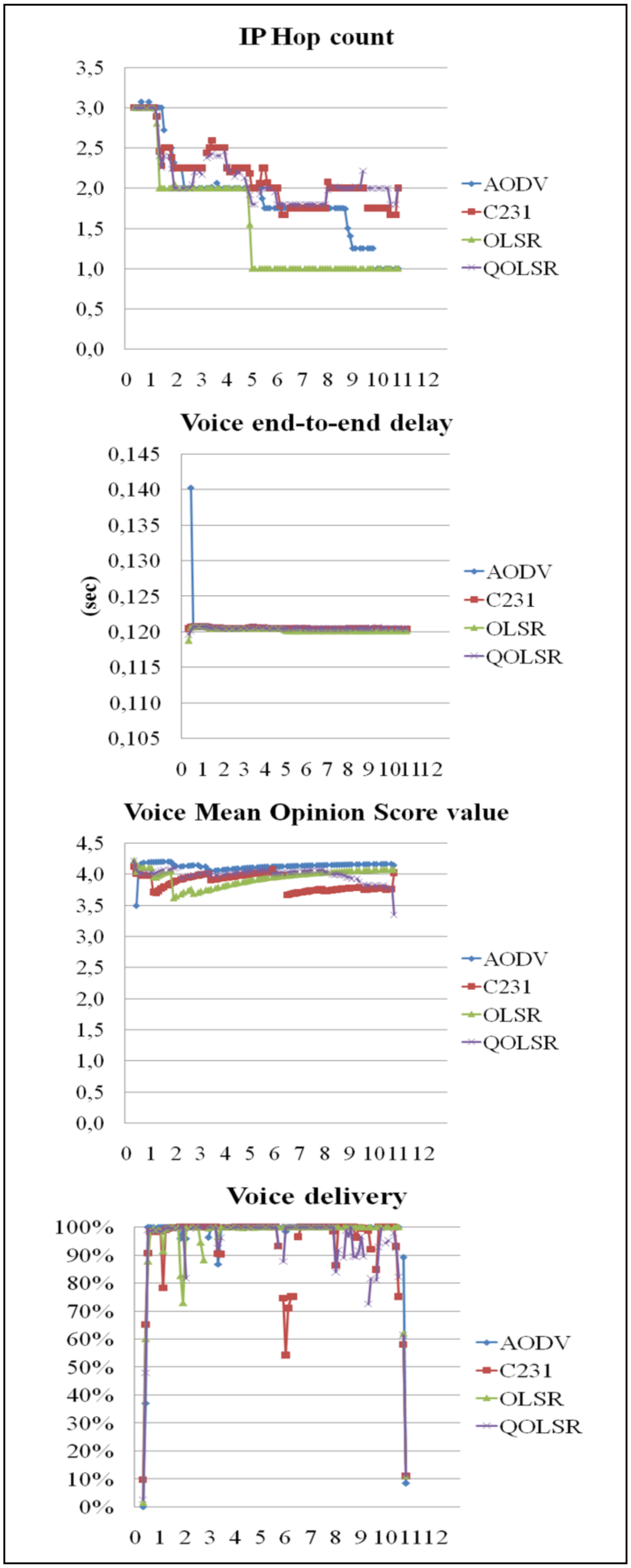


Fig. 10. Hop count, end-to-end delay, MOS, delivery for voice (scenario 6)

As shown on Fig. 9, voice communications operate perfectly 1-hop whatever the routing protocol, and despite the mobility of the vehicles. The delivery is 100%, the delay is inferior to QoS delay constraint for voice (150 ms), and the mean opinion score (MOS) value superior to 4 indicates good communication quality comparable to GSM. When voice session is operated by 2-hop communicating pairs or farther (Fig. 10), delivery may fall to 80% for every routing protocols, even 60% for C231.

The MOS value still indicates good communication quality, but also clearly degradations when losses occur. AODV provides the best performance every time, while OLSR has the worst until the communication is 1-hop again. QOLSR and C231 operate better with this application with QoS requirements, QOLSR being the more efficient. The good results obtained with AODV demonstrate that 802.11p can offer a good support for voice application up to 3-hop communicating pairs provided that the related flow is the sole.

4) Results for scenarios 7 and 8 : videoconferencing

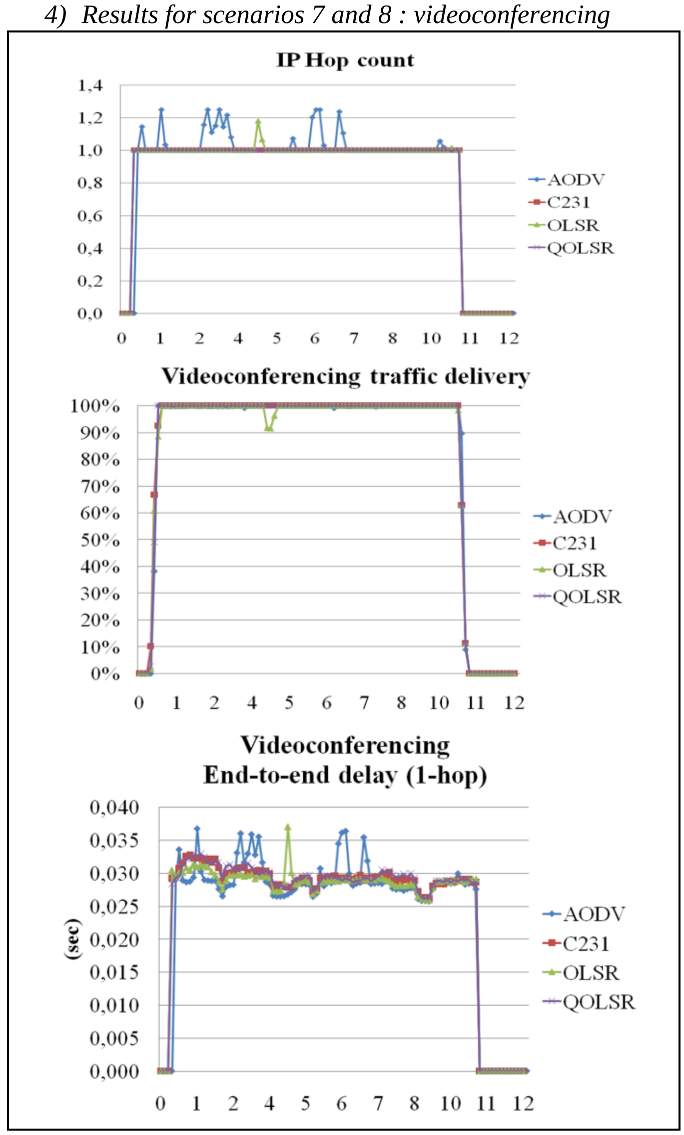


Fig. 11. Hop count, delivery, and delay for videoconferencing (scenario 7)

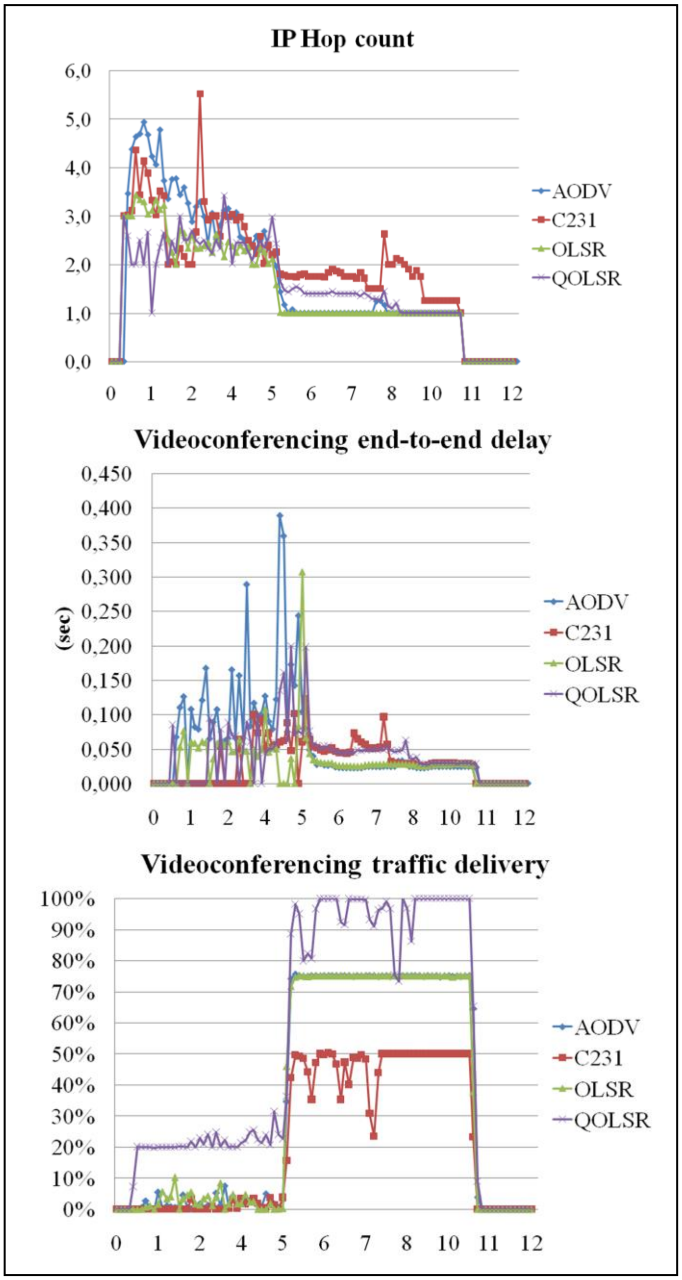


Fig. 12. Hop count, delivery, and delay for videoconferencing (scenario 8)

Videoconferencing application operates correctly when the communicating nodes are 1-hop away even under mobility conditions and for any of the routing protocols (Fig. 11). When multi-hop communications are necessary between the source and the destination, videoconferencing performance decreases dramatically (Fig. 12). These evaluations confirm the results in [4]: above 3-hop, the delivery of videoconferencing packets falls to 20% with OLSR and QOLSR, but also with AODV and C231. Despite additional routing traffic introduced by QOLSR, this variant still obtains higher packet delivery for videoconferencing than OLSR. AODV obtains the same packet delivery as OLSR, but due to longer routes AODV causes higher delays at the limit of videoconferencing threshold (300 ms) and very bad packet delay variation (above the limit of 50 ms for most of the packets). The strategy implemented in C231 based only on a single metric (i.e. received power) favors longer routes, thus degrading performance more than with other protocols. Combining three metrics which values are frequently updated as in QOLSR seem to be a better indicator for a good selection of relaying nodes for videoconferencing traffic. Indeed, QOLSR achieves the best performance despite the additional routing traffic necessary to its functioning.

5) Results for scenarios 9 : concurrent traffics

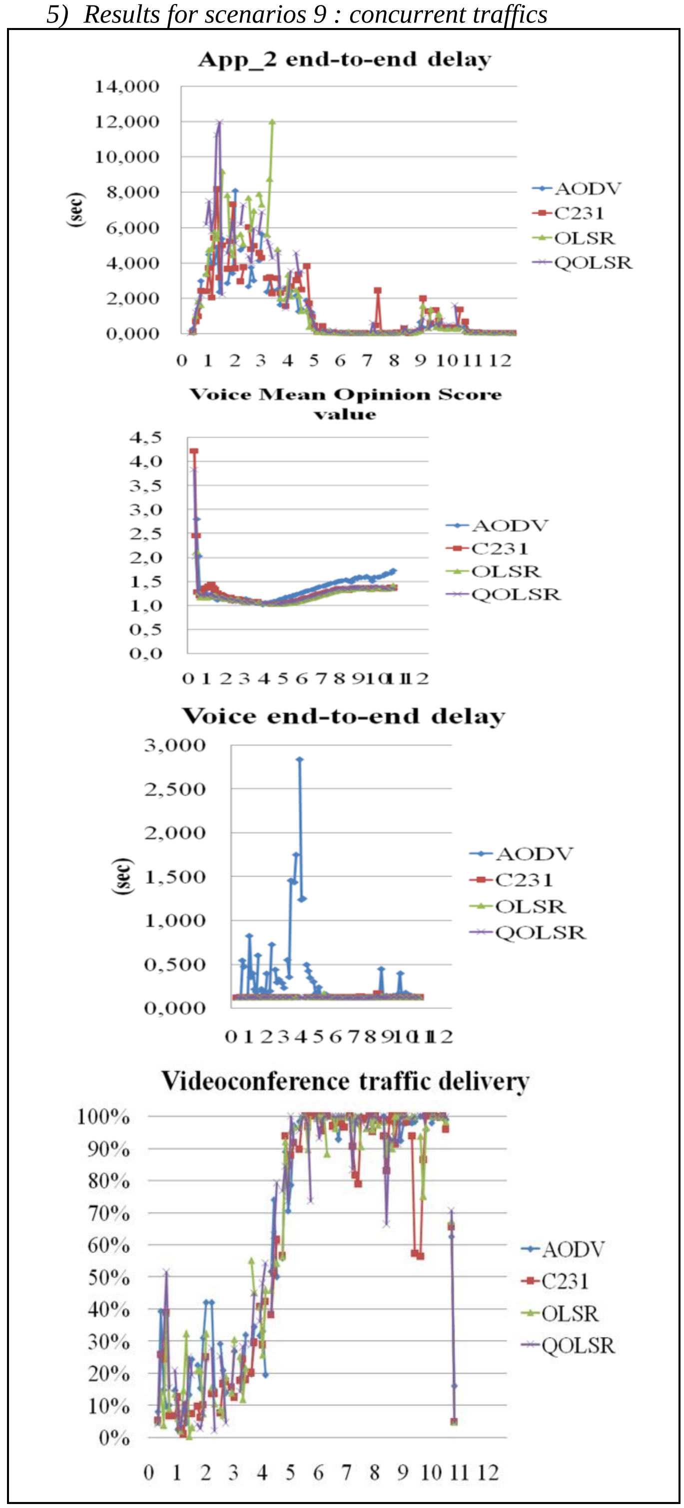


Fig. 13. Summary of results for the scenario 9 : concurrent traffics

As scenario 9 was described in section IV.A, three communicating pairs exchange simultaneously a different application each: one pair exchanges unicast traffic, another voice and the other videoconferencing. The end-to-end delay for the unicast traffic related to the safety application (App_2 described in section III.E) increases above the seconds, out of the bounds for real-time application constraints. Voice application is affected completely, falling to very bad quality (mean opinion score value) even in 1-hop case despite a relative correct packet delivery. The routing protocols have almost no real determining effect on the performance, and only AODV behaves slightly differently by reaching very high delays for voice application. Videoconferencing delivery is still good only when the communicating nodes are 1-hop away, but the related delays are slightly higher compared to scenarios 7 and 8 where there were no concurrent traffic. It seems that it should be avoided to operate several critical or QoS constrained applications in the same time and in the same group of vehicles. Thus, in absence of infrastructure, there is a need for distributed coordination between the nodes in order to operate admission control to ensure that a only a single critical traffic occur. These results confirm the conclusions of the evaluations performed in [2]. Indeed, in presence of different types of traffic with different priority levels, 802.11p does not yet achieve good delivery for all of them.

6) Scalability of the wireless LAN in presence of traffic

The results presented in Fig. 14 represent respectively the wireless LAN state when only one application is running (first row: the results are the average over scenarios 1, 3, 4, 5, 6, 7, 8), when three applications are running (second row: the results come from scenario 9) and when a full mesh traffic occurs between every pairs in the network (third row: the results come from scenario 2 where 50 broadcast traffic flows were sent simultaneously). When reasonable traffic is sent (one or three applications running), AODV realizes a throughput close to the load. As a reactive routing protocol, AODV tries to find enough resources to fulfill the demand. OLSR and its variants tend to ensure half more and even double throughput values as compared to the load. Proactive routing protocols try to gather most resources possible in order to be ready to fulfill any demand upon request. In the full mesh traffic case, AODV throughput and delays dramatically increase when OLSR and its variants scale better. Designed for dense networks, OLSR optimizes broadcast using multipoint relaying techniques.

V. Conclusion

In this work, the performance evaluation of several applications that could be provided as services to vehicles over ad hoc vehicle-to-vehicle communications has been presented. The simulation evaluation have been performed using Riverbed (OPNET) Modeler on a network topology made of 50 nodes equipped with 802.11p technology and following realistic trajectories in the streets of Paris at regular and authorized speeds. The results show that almost all application types are very well supported provided that the source and the destination have a direct link. Particularly, it has been shown that introducing supplementary hops in a communication has more effect on end-to-end delays and loss rates than having more packets, packets of higher size, or even higher mobility. It has been observed that when several types of traffic are sent simultaneously in the network, those having stringent QoS requirements undergo higher degradation, especially voice application. Another result of this study is the following. Ad hoc reactive routing protocols degrade performance by increasing the delays, whereas proactive ones introduce the same counter performance by increasing the network load with their routing traffic. Those latter, especially OLSR and its variants, are more efficient for broadcast traffic, while AODV allows better performance for unicast best effort traffic.

Whatever the routing protocol type adopted, the best performance seems to be obtained by maintaining small group of nodes reaching each other through at most two-hop routes. It will be particularly relevant, when the nodes can organize themselves, to avoid introducing other traffic while a session of an application with stringent QoS requirements is already running. Future study will investigate such self-organizing mechanisms, and try to lightening routing traffic induced by QoS signaling by improving the solutions such as the one evaluated in this work as C231 protocol.

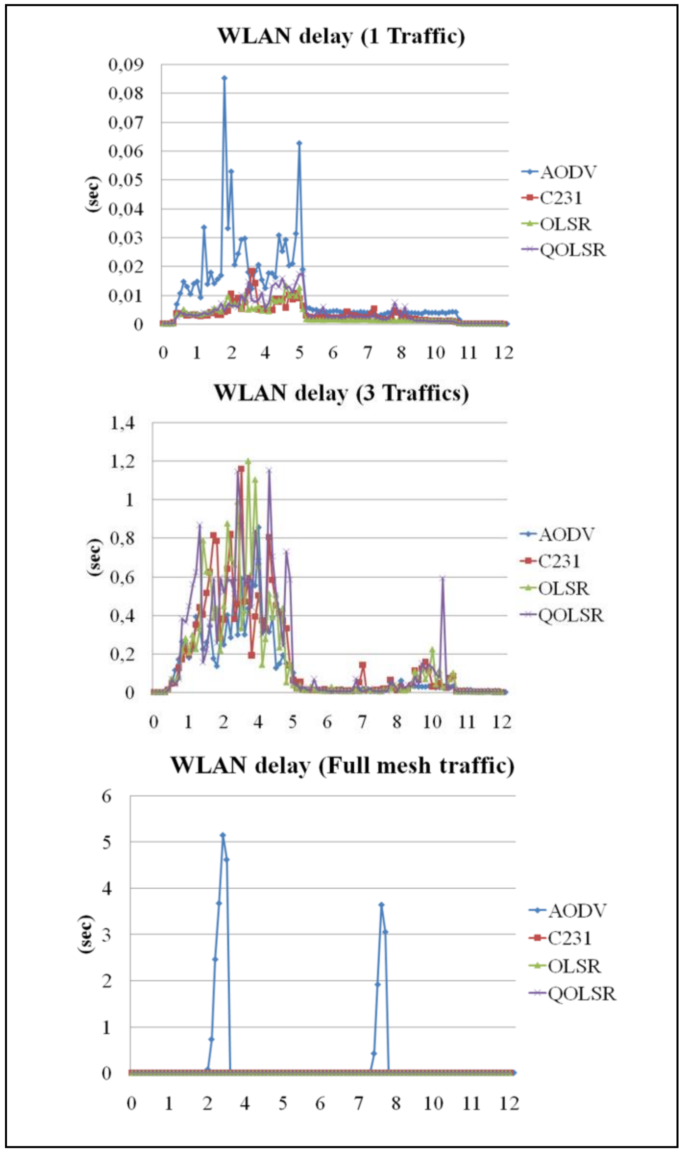


Fig. 14. Wireless LAN state evlolutions in presence of traffic

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the CPER ELSAT 2020, the Nord-Pas-de-Calais Region, the European Community and the French National Center for Scientific Research.

REFERENCES

[1] M. Gerla, "Electric and autonomous vehicles in smart cities: impact on energy, transport and communication infrastructures", IEEE MMTC COMSOC E-Letter, vol. 10, n° 3, may 2015.

[2] Ning sun, “Performance study of IEEE802.11P for vehicle to vehicle communications using OPNET”, a thesis presented in partial fulfillment of the requirements for the degree of Master of engineering in Telecommunications and network; Massy University, Auckland, New Zeland, November 2011.

[3] V.D. Khairnar,Dr. KetanKotecha , “Performance of Vehicle-to-Vehicle Communication using IEEE 802.11p in VehicularAd-hoc Network Environment”, International Journal of Network Security & Its Applications (IJNSA), Vol.5, No.2, March 2013.

[4] Bilstrup K.,Uhlemann E., Strom E.G., Bilstrup U., “Evaluation of the IEEE 802.11p MAC Method for Vehicle-to-Vehicle Communication”, Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th, p.1-5, September 21-24, 2008.

[5] Stephan Eichler, “Performance Evaluation of the IEEE 802.11p WAVE Communication Standard”, Proceedings of the 66th IEEE Vehicular Technology Conference, VTC Fall 2007, Baltimore, MD, USA., 30 September - 3 October 2007.

[6] B. E. Bilgin and V. C. Gungor, “Performance Comparison of IEEE 802.11p and IEEE 802.11b for Vehicle-to-Vehicle Communications in Highway, Rural, and Urban Areas”, International Journal of Vehicular Technology, Vol. 2013, Article ID 971684, 10 pages, http://dx.doi.org/10.1155/2013/971684, 2013.

[7] Duddalwar P., Deshmukh A., Dorle SS., “A Comparative Study of Routing Protocol in Vehicular Ad Hoc Network”, International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 3, pp. 71-76, March 2012

[8] Felipe Domingos da Cunha, AzzedineBoukerche, Leandro Villas, Aline CarneiroViana, Antonio A. F. Loureiro, “Data Communication in VANETs: A Survey, Challenges and Applications”, Research Report RR-8498, INRIA Saclay. 2014. HAL Id: hal-00981126, https://hal.inria.fr/hal-00981126v3

[9] Dua, A., Kumar, N., and , S. " A systematic review on routing protocols for Vehicular Ad Hoc Networks", Vehicular Communications (Elsevier) Journal, Volume 1 (1), 2014, Pages 33–52

[10] Johnson, D., Hu, Y. and Maltz, D. (2007), “The dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4”, IETF, RFC 4728 (2007)

[11] Perkins, C., Belding-Royer, E. and Das, S. (2003), “Ad hoc On-Demand Distance Vector(AODV) Routing”, IETF, RFC 3561 (2003)

[12] Clausen, T., and Jacquet, P. (2003), “OptimizedLink State Routing Protocol (OLSR)”, IETF, RFC 3626 (2003)

[13] Ogier, R., Templin, F. and Lewis, M. (2004), “Topology Dissemination Based on Reverse-Path Forwarding (TBRPF)”, IETF, RFC 3684 (2004)

[14] Roy, A. and Chandra, M. (2010), “Extensions to OSPF to Support Mobile Ad HocNetworking”, IETF, RFC 5820 (2010)

[15] Sondi P., Ganstou D. and Lecomte S. (2013), « Design Guidelines for Quality of Service Support in Optimized Link State Routing-Based Mobile Ad Hoc Networks », Ad Hoc Networks (Elsevier) Journal, Volume 11 issue 1, pp. 298-323, 2013

[16] Sondi P., Ganstou D. and Lecomte S. «A Multiple-Metric QoS-Aware Implementation of the Optimized Link State Routing Protocol»,Int. J. Communication Networks and Distributed Systems, Vol. 12, N° 4, 2014

[17] Euisin Lee; Eun-Kyu Lee; Gerla, M.; Oh, S.Y., "Vehicular cloud networking: architecture and design principles," Communications Magazine, IEEE , vol.52, no.2, pp.148,155, February 2014

[18] G. Amoussou, B. L. Agba, Z. Dziong, M. Kadoch, F. Gagnon, ″Performances Analysis of mobile ad hoc routing protocols under realistic mobility and power models″,Session 1542, OPNETWORK’06 Washington D.C., August 28 - September 1, 2006.

[19] Fan Ya-qin; Fan Wen-yong; Wang Lin-zhu, “OPNET-based Network of MANET Routing Protocols DSR Computer Simulation”, International Conference on Information Engineering (ICIE), 2010 WASE

[20] Vidhale, B.; Dorle, S.S.″ Performance Analysis of Routing Protocols in Realistic Environment for Vehicular Ad Hoc Networks″, 21st International Conference on Systems Engineering (ICSEng), 2011, Pages: 267 - 272, DOI: 10.1109/ICSEng.2011.55.

[21] Tapan K. Sarkar, Zhong Ji, Kyungjung Kim, AbdellatifMedour ″A Survey of Various Propagation Models for Mobile Communication″ IEEE Antennas and Propagation Magazine, Vol. 45, No. 3, June 2003.

[22] R. Mardeni and K. F. Kwan ″optimization of hata propagation prediction model in suburban area in Malaysia″ Progress in Electromagnetics Research C, Vol. 13, 106, 2010.

[23] S. IbenJellal, O. Cohin, S. Baranowski, U. Biaou, M. Bocquet, A. Rivencq, ″ Experimental analysis of Zigbee RF signal performances for railway application:Study on a laboratory reduced scale train″, IEEE International Conference on Advanced Logistics and Transport, 2015

[24] Vishal D. Nimavat , G. R. Kulkarni , “Simulation and Performance Evaluation of GSM propagation Channel under the Urban, Suburban and Rural Environments”, 2012 International Conference on Communication, Information & Computing Technology (ICCICT), Oct. 19-20, Mumbai, India, 2012

An Efficient and Reliable Core-Assisted Multicast Routing Protocol in Mobile Ad-Hoc Network

Faheem khan
Department of Computer Science
University of Malakand
KPK, Pakistan

Sohail Abbas
Department of Computer Science
University of Malakand
KPK, Pakistan

Samiullah Khan
Department of Computer Science
Agriculture University KPK,
Pakistan

***Abstract*—Mobile ad-hoc network is a collection of mobile nodes that are connected wirelessly forming random topology through decentralized administration. In Mobile ad-hoc networks, multicasting is one of the important mechanisms which can increase network efficiency and reliability by sending multiple copies in a single transmission without using several unicast transmissions. Receiver initiated mesh based multicasting approach provides reliability to Mobile ad-hoc network by reducing overhead.**

Receiver initiated mesh based multicast routing strongly relies on proper selection of a core node. The existing schemes suffer from two main problems. First, the core selection process is not efficient, that usually selects core in a manner that may decrease core lifetime and deteriorate network performance in the form of frequent core failures. Second, the existing schemes cause too much delay for core re-selection(s) process. The performance becomes worse in situations where frequent core failures occur due to high mobility which causes excessive flooding for reconfigurations of another core and hence delays the on-going communication and compromising the network reliability.

The objectives of the paper are as follows. First, we propose an efficient method in which the core is selected within the receiver group on the basis of multiple parameters like battery capacity and location, as a result, a more stable core is selected with minimum core failure. Second, to increase the reliability and decrease the delay, we introduce the idea of the mirror core. The mirror core takes the responsibility as a main core after the failure of the primary core and has certain advantages such as maximum reliability, minimum delay and minimizing the data collection process. We implement and evaluate the proposed solution in Network Simulator 2. The result shows that this scheme performs better than the existing benchmark schemes in terms of the packet delivery ratio, overhead and throughput.

***Keywords*—MANET; Core; Mirror core; Multicast routing; Receiver initiated; Mesh based routing; NS2**

I. Introduction

Mobile ad hoc network (MANET) is an infrastructure-less network of mobile nodes with decentralized administration and dynamic topology. Due to its infrastructure-less nature, these networks are tempted to be deployed in places where there is no pre-deployed infrastructure or where there is costly to deploy one. Hence, MANETs can be used in various situations such as jungles, mountains, deserts, nuclear disaster, in a mining field clearance, military operation and communication, battlefield, earthquake scenario, etc. where no infrastructure exist [1]. In MANETs, unlike wired networks, there are no dedicated routers for packet routing and forwarding. The MANET has a limited transmission range and all the nodes in the network strongly depend on intermediate nodes during data forwarding in multi-hop scenarios and hence, all the nodes in a network act as a host as well as router [2, 3]. Routing is an important function of the network. Routing can be of three types: unicast, broadcast and multicast routing. In unicast routing, data communication occurs in a one-to-one manner and only two nodes exchange their information with each other and in broadcasting the data communication occurs in a one-to-all fashion [4]. However, multicast routing works in one-to-many fashion and efficiently maintains the group communication by sending the similar copies of the same message to multiple nodes with a single transmission. In case of transmitting the similar data through several unicasts, multicasting minimizes channel capacity consumption, routing processes, energy utilization and end-to-end delay [5]. There are many applications of multicast routing [6, 7] such as armed forces operations and communications from one commander to a group/platoon, boss to subordinate communication, distance learning, information dissemination from air drones to a group of soldiers and presentation at the same time in different meeting rooms [8].

In MANETs, multicast routing protocols can be divided into tree based and mesh based routing protocols. In tree based multicast routing protocols, there is only one route between a sender and a receiver and is not robust against regular topology changes. However, it is well suited for environments where mobility is low [9]. Example of tree based multicasting are ad hoc multicast routing protocol utilizing increasing id numbers (AMRISs) [10], multicast ad hoc on-demand distance vector (MAODV) [11]. On the other hand, several paths are maintained from a source to the receivers in mesh based multicast routing. These multiple routes from source to all receivers give robustness, reliability and reduced latency at the cost of extra overhead as compared to the tree based multicast routing.

Mesh based multicast routing is further divided into sender initiated and receiver initiated routing protocols. In sender initiated approach every sender behaves as a core and it is the sender that initiates the mesh formation, maintains and updates the multicast paths to the receivers. Therefore, when the number of sources increases within a multicast group, the maintenance of the group becomes costly in terms of communication overhead. Example of sender initiated routing protocols are dynamic core based multicast routing protocol

(DCMP) [12] and on demand multicast routing protocol (ODMRP) [13]. Whereas, in a receiver initiated approach, one core is selected for the receiver group and it is the responsibility of the core node to maintain and update the receiver group. In situations where the number of receivers or sources increase, the receiver initiated protocols does not deteriorate performance in term of overhead as compared to the sender initiated protocol. Therefore, receiver initiated mesh based multicasting is more efficient than sender initiated mesh based multicasting. Example of receiver initiated routing protocols are preferred link based multicast (PLBM) [14], forward group multicast protocol (FGMP) [15], weight based multicast routing protocol (WBM) [16], data distribution management (DDM) [17], core assisted mesh protocol (CAMP) [18], protocol for unified multicasting through announcement (PUMA) [19], multicast for ad hoc network with swarm intelligence MANSI [20] and ODMRP.

The receiver initiated protocols suffer from two main problems. First, most of the protocols select the core node based on first come first serve basis, i.e. a node that first joins the receiver group. Therefore, the selected core may be in bad position with low battery capacity and hence not an efficient core is selected. In this protocol, we propose an efficient core selection method that elects core based on some parameters, such as battery capacity and location. As a result, the elected core would have prolonged lifetime. Second, core failures occur in the network due to various reasons, such as flat battery, out of range, or hardware fault that causes reconfiguration for re-selection of core node. As a result, this reconfiguration process will increase the overhead in the form of regular flooding of control messages and delay an ongoing communication; hence, the system will be considered as unreliable. In order to reduce the delay caused by reconfiguration and to enhance network reliability, we propose the notion of mirror core. When the core node is elected (based on some ratings), then the core will select the second topmost node as a mirror core. In case of the primary core failures, the mirror core will take the charge as a main core without causing any delay and extra overhead. Hence, the system will become more robust and the group communication will be continued without any delay.

The novelty of this research work is twofold. First, we propose a stable core that will ultimately reduce core failures. Second, we propose a solution to reduce the data collection process, delay and overhead when core failures occur. The rest of the paper is organized as follows. In Section 2, we briefly describe the literature survey of sender and receiver initiated protocols and explains their drawbacks related to the core selection and core failure. Section 3 introduces the design of ERASCA (Efficient and Reliable Core Assisted protocol in Mobile Ad-hoc Network) which builds and maintains receiver initiated mesh based multicasting with the method of core election and mirror core selection. Section 4 and 5 evaluates the proposed scheme using Network Simulator 2 (NS-2) and compared with other benchmark schemes using various metrics. The paper concludes in Section 6.

II. Related Work

The primary goal of ad-hoc multicast routing protocols is to construct and maintain a robust and efficient topology even during high network dynamics and limited bandwidth. Among these protocols, mesh based multicasting is considered more robust and reliable than tree based multicasting. The mesh based multicast routing is divided into sender initiated and receiver initiated multicast routing protocols [21].

In sender initiated approach, a sender starts the formation of the mesh. In this approach it will be the responsibility of the sender to maintain and update the multicast paths to receiver. The first problem appears when the number of receivers increases, as the number of reply packets sent back by the receivers to a sender also increases because after every successful reception of packets a receiver must reply back to the receiver, which creates a bottleneck at the sender end. Second, a sender initiated approach depends on a consistent network flooding, as every sender behaves as a core when it joins the network that leads to the problem of creating large overhead and energy consumption. Finally, the sources must be part of the multicast mesh group, even when they are not interested in a transmission. Therefore, when the source node increases, then the flooding from every source increases and will produce large overhead. Examples of sender initiated mesh based multicast protocols are DCMP, Neighbor supporting multicast protocol (NSMP) [22],ODMRP etc. On the other hand, a receiver initiated approach transfers maximum responsibility on the receivers for reliable data delivery and will solve the issues related to sender initiated routing protocols. First, in receiver initiated approach, only the receiver which didn’t receive the packet will reply to the source. Hence, receiver initiated will not be affected when the number of receivers grows as compared to its counterpart. Second, in the receiver initiated approach, the core is responsible for the maintenance and update of the receiver group as compared to sender initiated protocol where each source needs to maintain the path from each source to its corresponding receivers. Finally, this approach does not make an extra overhead as compared to the sender initiated approach because in a receiver initiated approach the sources are not forced to be a part of the group.

ODMRP is a sender initiated mesh based protocol. It uses forwarding group concept in order to transmit multicast packets through flooding. In ODMRP, the source node administers the membership of the group, maintains and updates the multicast path and the multicast group. In ODMRP, when data packets sent by the source are received by the receiver, an acknowledgment is sent by the receiver to the sender that the data is received otherwise the sender will retransmit the packet after a period of waiting. This retransmission will continue until the reception of acknowledgement from the receiver, which will create congestion and overhead as the numbers of the receiver increases. Second, in order to maintain and update the group and the paths to the receivers, ODMRP depends on the consistent network flooding from the source nodes that leads to the problem of scalability in situations where the source node increases [23].

MAODV is a receiver initiated multicast routing protocol. In MAODV, a receiver group will be established with the help of Hello messages and will make the connectivity list within the group. The first node which joins the group will be selected as a Leader (core). The Leader updates and maintains the receiver of a group with the help of Hello messages. In MAODV, when the nodes in one group find another group, they would like to merge the groups with each other. The main drawback of MAODV is the frequent link failure in high mobility because of its tree infrastructure and a single point of failure, which is the core node. Also the merging concept of one group within the other group in MAODV make it more complicated because the node will have to find the superior core within each other, which can create unnecessary delay.

CAMP is a receiver initiated on demand multicast routing protocol. It uses mesh based topology and a unicasting technique in order to establish and maintain a multicast group member to known destinations. CAMP establishes a mesh composed of shortest paths from senders to receivers and one or multiple core can be defined for each mesh. In small networks the CAMP work well, but creates a considerable amount of overhead and unreliability in large networks and high mobility [24]. Moreover, if any branch of a multicast tree fails, then all the components of the tree and its related branches must be reconnected for packet forwarding to continue the communication between the source and the destination.

PUMA is a receiver initiated mesh based protocol. It uses core node to transmit its multicast packets to the desired destination group. All the receivers are attached along the optimal path to the core, the core is selected among the receivers and therefore each and every node on the shortest path between a core and the receiver establishes the mesh. The first problem appears in PUMA, is that the first receiver in a group will be selected as the core. This first-come-first-serve based selection may cause illegitimate or inappropriate nodes to be selected as core which may have a minimum lifetime; hence, increase core failure chances in the network decreases the efficiency. Second, core failures can further cause reconfigurations and as a result reliability and network lifetime will be compromised. As, the main core fails, there is no alternative technique to prevent reconfiguration and save the existing information of the every node in the network because with core failure every node will delete all information related to the group which was achieved through communication.

All the above mentioned protocols select the core on a first-come-first-serve basis (i.e. the first node that joins the group). However, the CAMP uses the Extended Ring Search (ERS) method for another core and PUMA selected the core by election but with limited parameters. Hence the process of a core selection in all these protocols is not efficient as the selected core may be in bad position with low battery capacity and may cause frequent core failures thereby increasing overhead. Furthermore, the time and network resources required for the new core selection may cause the protocols to become inappropriate for a Quality of Service (QoS) based applications, especially the delay caused in the process. However, in ERASCA, the first receiver who joins the group is selected as a core like the above protocols, but after the failure of the first core, it does not continue the same procedure, but elects the core on battery capacity and location. This will select a resourceful core within the group and decreases the core failure. As a result, decreases the reconfiguration and increases efficiency in term of minimizing the overhead. In order to increase the reliability and reduces the delay occurs in the new core selection process, we introduce the mirror core. The mirror core acts as a primary core after the first core failure occurs and prevents the network to go into orphanage phase.

III. Protocol Description

A. Overview

ERASCA uses the IP multicast service model of permitting any source node in a network to transmit its packets to the multicast group without knowing the constituency of the group. Furthermore, the ERASCA is based on a receiver initiated approach in which the sources are not required to be part of the receiver group for the transmission of data to receiver group.

In ERASCA, if the receiver does not receive any invitation from the group then it will announce himself as a core of the group. This core node will start the formation of the receiver group through Status Declaration message (Explain in Subsection B). In ERASCA, the receivers join a multicast group using the address of a core node. As a result, a group will form and every receiver in a group will be informed from each other status (i.e. battery capacity and location).

The receiver connects to the core node through intermediate nodes with the help of the SD message, which will be flooded by the core node and form the connectivity list. All the intermediate nodes connecting receivers to a core node acting as relay nodes, collectively form the mesh. With the help of connectivity list, the sender sends a data packet towards the mesh through the best possible route. On reception of data packets through any mesh member, it is flooded within the mesh members of the group and ultimately reaches to all the receivers of the group.

In ERASCA, the receivers elect the core to become the point of contact between the mesh members and non-mesh members (these terms will be explained in Subsection E in detail) and it is the responsibility of the core to periodically broadcast the updates about entry and exit of the mesh members, group members, mirror core and about its own existence to the rest of the network by using SD messages. Hence, it is the core node that updates and maintains the mesh group.

B. Status Declaration Message

In ERASCA, the core node uses SD Message Packet Format as shown in Fig.1 to maintain and update the mesh of a group by periodically flooding the SD messages to form a connectivity list. Connectivity lists are formed throughout the network with each node, which allows the sources to send the data to the mesh of a group. Each SD message specifies a core ID, group ID, parent node, sequence number, distance to the core mesh member flag and battery capacity. With the information contained in the SD messages, define the path for sources outside a multicast group to transfer the data packets towards the group. The SD message maintains and updates the mesh of a group by informing others about the leaving of a mesh members or joining of the new receiver in the group.

0 15 31

Mesh Membership flag	Distance to Core
Group ID	
Core ID	
Mirror Core ID	
Sequence Number	
Parent ID	

Fig. 1. SD Message Packet Format

Core *ID*: Core node identifier

Mirror Core ID: Mirror Core node identifier

Group *ID*: Group *ID* of the concerned group

Sequence number: The sequence number in the best Status Declaration in which fresher sequence number is given

Mesh member flag: If the node is a part of the mesh, then the flag will be set otherwise it will not be set

Distance to core: The distance to the core in the best Status Declaration

Parent: The nearest neighbor from which it received the best Status Declaration

C. Connectivity List

A core node periodically transmits the SD messages for the concerned group due to which each node forms a connectivity list in the network. With the help of connectivity lists, nodes in the network can calculate the best path from a source to a group through parent nodes. Parent node shows the preferred neighbor (which shows the shortest path to the core) to reach the core. The source node may or may not be the group member. All nodes in the network store the information they collect from their neighbors via SD messages along with the received time into the connectivity list. Fresher SD message (one with a higher sequence number) from the neighboring nodes is preferred over the lower sequence number for the same group. Therefore, for the same group a node contain only one entry in the connectivity list for the specific neighbor with a fresher sequence number for the given core. Hence, for the same core ID, the SD message with a fresher sequence number is preferred as shown in Table 1. For the same core ID and fresher sequence number, SD message with less distance to the core is preferred. For the same core ID, when all those fields are same then SD message with higher battery capacity is valid. For the same core ID, when all those fields are same than the SD message that arrived earlier is considered valid.

Fig.2 shows the dissemination of the SD message all through the network and Table 1 shows the building of connectivity lists at node 8. The solid arrow shows the node from which it receives its best SD message. Node 8 has four neighbors in its connectivity list, i.e. 1, 7, 9 and 10. Neighbor 10 is not selected as a best entry because it has two hops distance with minimum battery capacity and a maximum delay. Neighbor 1 is not selected as a best entry because it has the minimum battery capacity and a maximum delay than node 7 and 9. However, it selects the entry it receives from neighbor 9 as the best entry, because it receives earlier than node 7. Now node 8 uses this best entry to produce its own SD message which contains all the fields as shown in Table 1.

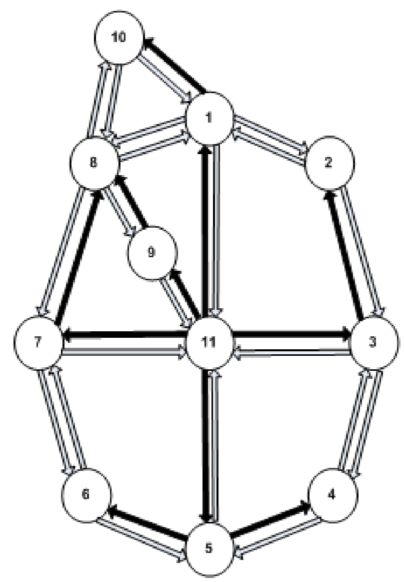


Fig. 2. Dissemination of SD Message

When a node receives a multicast data packet from a source node, it forwards it to the node from which it receives a best SD message. If the concern path is broken, then it tries next best path available, because in mesh multiple routes are available from source to group. As soon as the data packet reaches to any mesh member of the group, the mesh member floods the packet within the mesh group until the desired receivers get the data packet. Mesh members use a packet ID cache to detect and remove duplicate data packets during flooding. The routing of data packets within the network from sources to receivers are also used for the update of the connectivity lists. Because when the sender sends a data packet to the receiver through non-member then the non-member expects its parent node to forward the data packet to the mesh. As the MANET is broadcasted by nature, therefore the node also receives the packet when it is forwarded by its parent node and receives an implicit acknowledgment from the parent node that forwards its packet. But if the neighbors do not receive an implicit acknowledgment within a specific time interval from the parent node, it eliminates the parent node from its connectivity list. Therefore, connectivity lists are updated immediately as soon as it detects its parent lost.

TABLE I. CONNECTIVITY LIST AT NODE 8

Neighbor	Core ID	Group ID	Seq. no	Parent	Distance to core	BC	Time
9	11	224.0.1.2	64	11	1	90%	11132
7	11	224.0.1.2	64	11	1	90%	11138
1	11	224.0.1.2	64	11	1	87%	11144
10	11	224.0.1.2	64	1	2	80%	11159

D. Receiver Group Formation

In ERASCA two situations arise for the receiver group formation. First, a node n is interested to become a part of any receiver group, if any group exists. Hence, if the receiver group is existed, then it will have a core node. A core node is periodically transmits the SD messages for the receiver group. If node n receives SD message from any group member say m, then n will send a Join Request to the m for the receiver group membership. In reply a Join Acknowledgment is transmitted to n by m. Now it adopts the group specified in the SD message it has received and starts transmitting messages that specifies for the group. Second, if there is no receiver group then it will announce itself as a core node and start SD messages periodically to inform other receivers through SD messages to join the group.

After joining the receiver group, if receiver does not receive the SD message within 3 x SD interval after the first time, then it assumes that core has been failed. To confirm whether the core has failed or not, a receiver floods a Core Failure Announcement (CFA) Request. In CFA Request, the sequence number field is set to the highest sequence number from the old core. After sending a CFA Request, the receiver sets a CFA wait flag, as well as starting a timer with time period CFA ack timeout interval. Intermediate nodes will receive a CFA Request Reply with their fresher SD message, if they receive SD message with higher sequence number than the sequence number in CFA Request. Otherwise they will also set the CFA wait flag to TRUE, and start a timer CFA ack timeout interval and forward the CFA Request. On the other hand if the core failure is not occurred then the CFA Request will finally reach a receiver which receives a latest SD message than the sequence number in the CFA Request. The receiver then broadcast SD message with fresher sequence number in the receiver group which is forwarded back on the same route to initiator which originated the CFA Request. If a core failure has indeed occurred then the CFA Request will never reach the initiator because of the loss of the connectivity list and as a result the CFA ack timeout interval expires. Therefore, when the receiver initiating the CFA Request recognizes that it is not receiving the SD message with a higher sequence number within a specific time interval, then it will confirm and announce a CFA and will conduct an election as shown in Subsection F. CFA ack timeout interval should be set to specific time interval which is sufficient for the CFA Request to come back to a receiver which originated the CFA Request.

E. Mesh Formation

The network nodes are categorized as group member and non-group member. Non-group members *(NM)* do not belong to the mesh and are shown in black color nodes. On the other hand, group members are further divided into End Receiver *(ER)*, Intermediate Receiver *(IR)* and Group Relay *(GR)*. The nodes in white represent the End Receivers (*ERs)*. The *ERs* are terminal receivers, i.e. mesh is terminated on them, and they do not participate in the packet relay process. Whereas, the *GR* nodes can only act as intermediate nodes between the receiver and the core and we denote them by blue dots. Likewise, *IR* is the receiver node as well as the intermediate node simultaneously, denoted by red dots. As shown in the Fig.3, the intermediate node between R47 and core is R42. R42 is a receiver node but in this situation it also acts as an intermediate node and hence will be termed as *IR*.

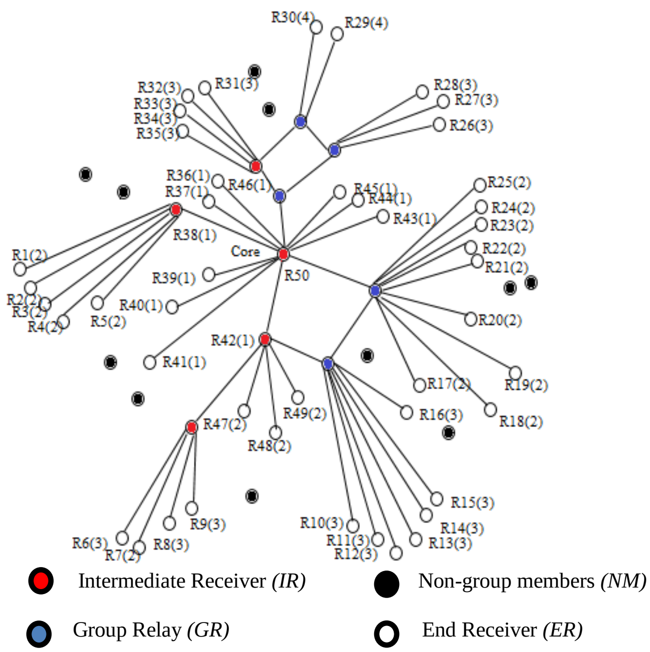


Fig. 3. Mesh Formation

Initially only receivers consider themselves as a mesh members but now GR will also consider themselves as a mesh member because they exist between core and ER and forward the packets between them and hence will be considered as part of a mesh group. A mesh group will be composed of ER, IR and GR nodes. As shown, blue nodes (GR) are the intermediates nodes that exist between the core and the receiver and having at least one or more ER node connected to it. It should be noted that flooding of SD from the core will only be carried on by the IR and GR nodes, instead by all the group members. To limit the flooding only to the IR and GR nodes, considerably reduced the overhead. In ERASCA, only the ER and IR node can be selected as a core, whereas GR cannot be selected as core node because GR acts only as an intermediate node and not a receiver.

F. Core Election

In traditional approaches, when the core node fails because of mobility or battery capacity, the group will again select the core irrespective of its position and battery capacity. With inappropriate core location, i.e. in a less populated area, the core will face a large delay with maximum link failure which decreases the efficiency of the network. Likewise, having low remaining battery capacity, the core failure occurs soon and hence a core with high battery capacity should be preferred which will possibly increase the lifetime of the network. In this approach an election is conducted, to elect a core. Thus after the failure of the first core, it does not continue the same traditional approaches, but elects the core with proper election on battery capacity and location (i.e. dense part of the network or maximum connectivity). This will select a resourceful core within the group. In order to select a resourceful core, the following steps will be performed.

Step 1: In situations when the core node fails, the group members will be aware of the core failure situation through *CFA*. The *CFA* will be flooded by the *IR* and *GR* within the group, if *IR* and *GR* do not receive 3 consecutive *SD* message within a specific time interval. Each *SD* message is announced after 3 seconds.

Step 2: After the *CFA*, an election is conducted in a receiver group. For this purpose, a *receiver n* floods the *Election Request* message to all receivers within the group as shown in Fig.4. The purpose of this message is to inform all receivers that the core has been failed. If the core is really failed, the receiver *n* will receive an *Election Reply message* from all receivers in a group; otherwise, it will receive nothing after some time interval. It would mean that the initiator may be gone out of range of the network.

Step 3: In reply all receivers in a group will flood the *Election Reply* message to receiver *n*, if the core is also recorded to be failed with all receivers in a group. The purpose of *Election Reply* is twofold. First, shows its willingness to participate in core election process. Second, each receiver will establish paths to every other receiver in a group.

Step 4: For knowing a Remaining Battery (*RB*) and number of connected neighbors of all receivers, a *Core Election Message* is flooded in a group by a *receiver t*o elect the best receiver in a group.

Step 5: All receivers will also flood a *Core Election Message (CEM)* within the group in which each receiver must include its *BC* and number of connected neighbors. As a result, every receiver will have a list of receivers. Each receiver floods its topmost receiver in a group. This will enable every receiver to have all the votes regarding topmost receiver node.

Step 6: As a result, all the receivers will know the estimated battery capacity and number of connected neighbors of each other. After exchanging information through *CEM*, receiver *n* elects its topmost receiver in a group.

Step 7: All receivers will also elect its topmost receiver in a group.

Step 8: Top most receiver is flooded by the receiver *n* as well as by the all group members.

Step 9: As a result, a topmost receiver within the receiver group with high battery capacity and maximum numbers of neighbor is elected as a core node.

Step 10: The core node will flood this news within the group about its own existence through SD message.

There are two aspects of core election process. Firstly, an efficient core is selected on the basis of battery capacity and best position, which will perform its duty as a core for a long period of time in the network. It is obvious that a good location of the receiver might be the one that is less dynamic or that the neighborhood environment is less stagnantly changing i.e. fewer changes occur in a given time interval. Also, it can be assumed that a node with maximum number of neighboring nodes will probably be in the center of the group despite at the corner of the receiver group. Secondly, the group will get rid of frequent core failures hence, overhead will be decreased. After core election, the core node sends SD message with its node ID to the whole network.

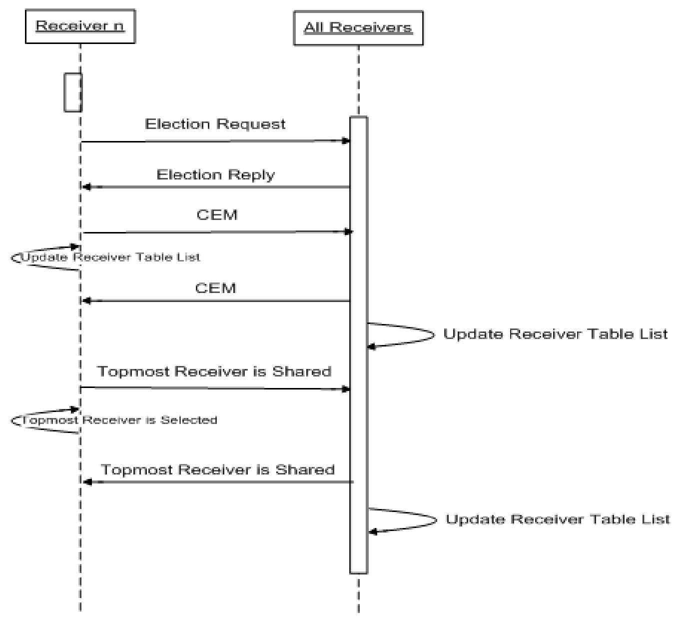


Fig. 4. Sequence Diagram

G. Mirror Core Selection

In order to resolve the issues related to core failures, we introduce the mirror core. The primary core selects the mirror core from the rating list (which should be the second topmost node). Therefore, when the main core fails the mirror core takes the responsibility as a main core and the mesh group will be maintained and updated continuously without any delay. It has certain advantages such as maximum reliability, minimum delay and minimization of the data collection process.

After the selection/election of the core, it is compulsory for the core to select the mirror core of the group. For this reason, the core selects the most suitable receiver within its broadcast range (preferably with one hop distance) as shown in a black dotted circle in Fig.5. Likewise, the mirror core can also be found within two hop distance with the help of GR in blue dots (N51, N52 and N53) and IR in red dots (R38, R46 and R42) and are explained Subsection H. Here only the ER and IR node can be selected as a mirror core and the GR nodes cannot be selected as a mirror core because GR nodes are not the member of the receiver group but only serving as an intermediate node. As soon as the mirror core becomes a primary core, it starts to transmit SD messages in the network about its status as being the core node.

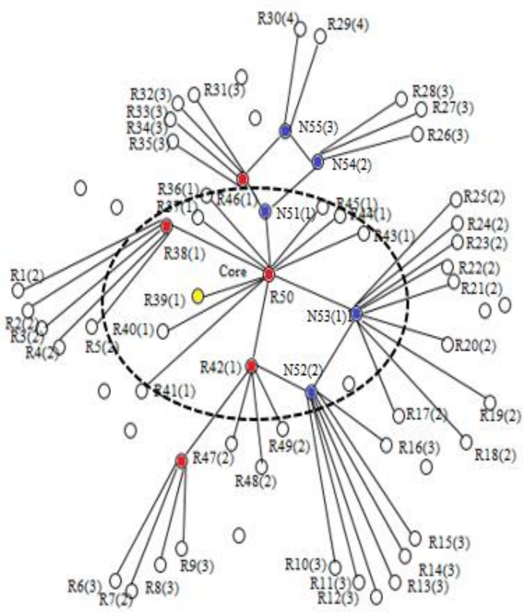


Fig. 5. Mirror Selection within neighborhood

After the core election/selection, it is the responsibility of the core node to select the mirror core. The mirror core is selected by the main core on factors like battery capacity and distance to core within a receiver group. After the core election/selection the primary core will first prefer the suitable receiver within one hop distance, if not found than prefer 2-hop distance and so on. The suitable receiver must have the highest aggregate after the core node. The aggregate depends on battery capacity and numbers of hop between the receiver and the core. For the mirror core selection the core floods the Mirror Core Selection Request (MCSReq) within the receiver group. The purpose of this message is to inform other receivers that the mirror core has not yet been selected. As a result, the core will receive a Mirror Core Selection Reply (MCSRep) from the other receivers in a group through unicasting. The purposes of MCSRep are twofold. First, it shows willingness to participate in the mirror core selection process. Second, each receiver will establish paths to the core in a group.

For knowing a Remaining Battery (RB) and distance of core (in term of hops) from each receiver, a Mirror Core Selection Message (MCSM) is flooded in a group by the core to select the best receiver in a group. In reply all the receivers in a group will send their Mirror Core Selection Message (MCSM) through unicasting to the core node. Now the core node has a table of receiver list through which the core will select the topmost receiver as a mirror core. Now the core will transmit the SD message with mirror core in its packet format and after the failure of the main core the mirror core takes the responsibility as a main core without data collection process and starts SD message without any delay. It is important to mention here, when the mirror core takes on the role of the primary core in the mesh. There are two occasions. First, when the core node depletes its resources quickly, i.e. battery capacity and explicitly announces “resource exhaustion” message. Second, when the core node has abnormally disappeared due to mobility or hardware faults, etc. In the above two situations, the mirror core takes the charge as a primary core.

H. Connectivity List of Mirror Core

If the mirror core is not found by the main core in one hop distance, then the core will select the mirror core within the group through IR and GR nodes. To select the mirror core in the group, a core should be aware from the status (battery capacity and distance to the core) of every receiver in a group. For this purpose MCSM is flooded in the group by the core through IR and GR nodes. In reply all the receivers will send their status to the core node, which gives an expanded choice to the core for the mirror core selection. Therefore, through IR and GR a suitable mirror core can be selected.

It should be noted that a mirror core should only be selected by the core node within one hop distance based on battery capacity and distance to the core, however, the receiver with low battery capacity within one hop distance is not preferred. On the other hand, a suitable receiver with high battery capacity within a two or three hop neighborhood and not more than 3-hop neighborhood will be preferred over a receiver with low battery capacity within one hop neighborhood. Because a mirror core failure will occur soon within one hop with less battery as compared to a mirror core with a high battery capacity within two or three hop neighborhood.

Table 2 shows the connectivity list of node R38. In this situation only R38 is consider, where five receivers are connected to R38 at two hop distance. R3, R4 and R5 will not be selected as a mirror core, as they have a minimum battery capacity than R1and R2. Likewise, R1 and R2 can be selected as a mirror core with high battery capacity, but priority will be given to R1, because R1 receives the SD message earlier than R2 and hence R1 will be selected as a mirror core of the primary core. Similarly, the mirror core can be selected through (N51, N52 and N53) and (R46 and R42).

TABLE II. CONNECTIVITY LIST AT NODE R38

Neighbor	Core ID	Group ID	Seq. no	Parent	d_to core	BC	Time
R1	R50	224.0.0.2	5	38	2	88%	14132
R2	R50	224.0.0.2	5	38	2	88%	14135
R3	R50	224.0.0.2	5	38	2	75%	14138
R4	R50	224.0.0.2	5	38	2	70%	14150
R5	R50	224.0.0.2	5	38	2	30%	14155

IV. Simulation Scenario

This paper implement, evaluate and compare this proposed solution in a network simulator with the benchmark schemes like PUMA and MAODV and use NS-2.35 on Ubuntu platform using Tcl/Otcl and C++ as a front and back-end languages respectively for implementing the proposed ideas. Likewise, an AWK script is developed to collect data from NS-2 trace file.

A. Metrics

In this experiment, the following metrics are used, i.e. throughput, packet delivery fraction (PDF) and overhead with the following parameters as given in Table 3.

TABLE III. Simulation Parameters

Simulator	Network simulator (NS2)
Simulator time	450 Sec
Number of nodes	50
Mobility	5
Simulation area	1000m x 1000m
Data packet size	512 bytes
IfqLen	60
MAC type	MAC802_11

Throughput: is the measurement of performance of MANET, which shows the amount of data transfer from one location to another location in the specified amount of time and depends on multiple factors like channel capacity and bandwidth etc.

Packet Delivery Fraction: can be defined as a data packet received divided by the data packet sent.

PDF = total number of packets received/ total number of packets sent

Overhead: is the total packet sent (control packet + data packet) divided by the data packets received.

Overhead= total packet sent (control packet + data packet) / data packets received

Several scenarios have been simulated in order to determine the effect of mobility, number of receivers, number of senders, ifqLen and simulation area on the performance metrics for each protocol. Five scenarios have been simulated in different environments and on the basis of these scenarios we evaluate these protocols and make the conclusion on the basis of results.

Scenario 1: Mobility changes across {0, 10, 20, 30, 40} m/s

Senders = 5, Receivers = 20, ifqLen = 60, Simulation area = 1000 x 1000

Scenario 2: Senders changes across {1, 2, 3, 4, 5}

Mobility = 5, Receivers = 20, ifqLen = 60, Simulation area = 1000 x 1000

Scenario 3: Receivers changes across {1, 2, 3, 4, 5}

Mobility = 5, Senders = 20, ifqLen = 60, Simulation area = 1000 x 1000

Scenario 4: ifqLen changes across {10, 20, 30, 40, 50, 60}

Senders = 5, Receivers = 20, Mobility = 5, Simulation area = 1000 x 1000

Scenario 5: Simulation area changes across {500 x 500, 1000 x 1000, 1500 x 1500, 2000 x 2000}

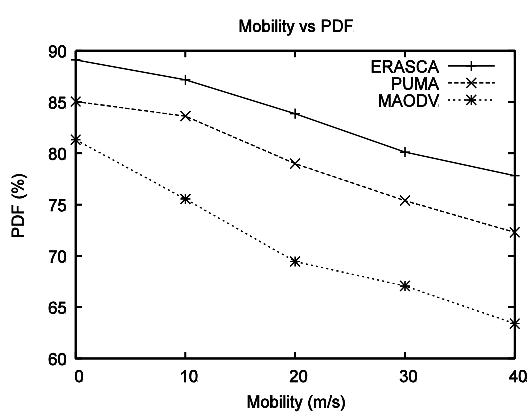
Senders = 5, Receivers = 20, Mobility = 5, ifqLen = 60

V. Protocol Comparison

In this paper, the performance of ERASCA is compared with PUMA and MAODV which are the benchmark schemes for mobile ad-hoc network using the network simulator (NS2) parameters given in Table 3. ERASCA, PUMA and MAODV are receiver initiated routing protocols. However, ERASCA and PUMA are mesh based protocols and MAODV is a tree based protocol.

A. Scenario 1

In scenario 1, the mobility is changed from 0-40 and makes all other parameters fixed as given in Table 3. On the basis of such parameters, multiple simulations are performed on protocols like PUMA and MAODV and compare their matrices like PDF, throughput and overhead with ERASCA. As shown in the Fig.6, throughput, PDF and overhead change with respect to mobility. In low mobility the packet drop decreases and PDF increases, but the opposite happens when the mobility increases. As with high mobility the link failure increases and therefore the delay is higher. As a result, throughput decreases because throughput is the packet transmission per second. In such a situation frequent flooding is used to minimize link failure and hence the overhead increases and an ongoing communication is delayed. Because of the delay, link failure and overhead the throughput and PDF decreases. Hence, the network performance decreases because of the frequent link failure and core failure. MAODV shows poor performance as compared to PUMA and ERASCA. MAODV is a tree based protocol, as tree based protocols are not resilient against mobility because of a single route between a sender and a receiver and therefore the packet delivery ratio is very less and overhead is high as compared to mesh based multicast routing protocols. It is important to mention that



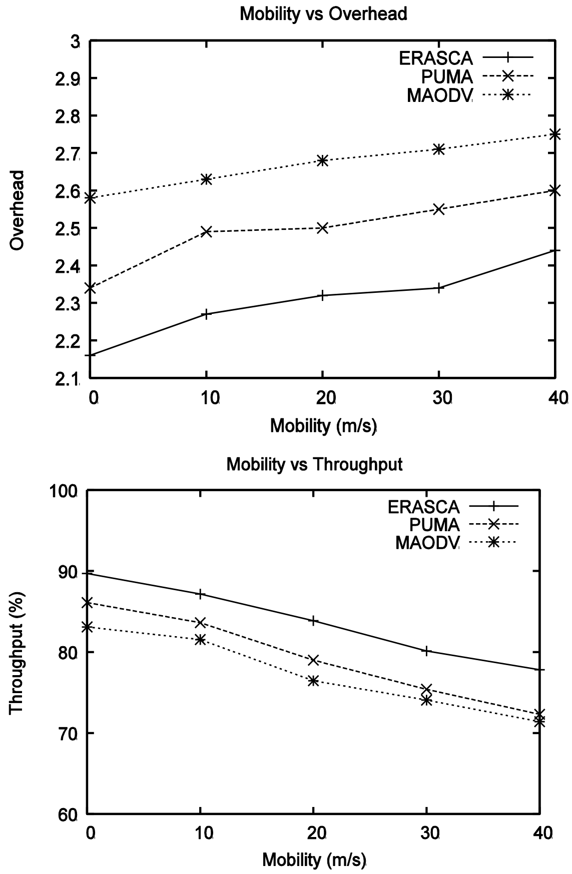
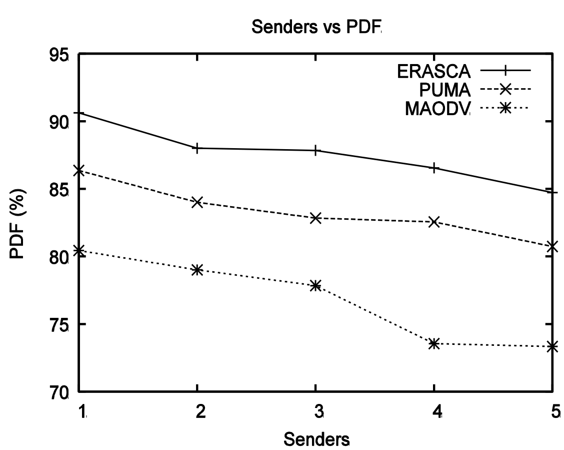


Fig. 6. Comparison of Mobility with PDF, Overhead and Throughput

When the link failure and packet drop increases then the PDF decreases. On the other hand, PUMA shows better performance than MAODV but less performance to ERASCA because of the frequent core failures. Therefore, ERASCA shows better PDF, throughput, overhead and delay as compared to MAODV and PUMA because of the stable core selection. As in ERASCA less core failure occurs and hence decreases the reconfiguration process.

B. Scenario 2



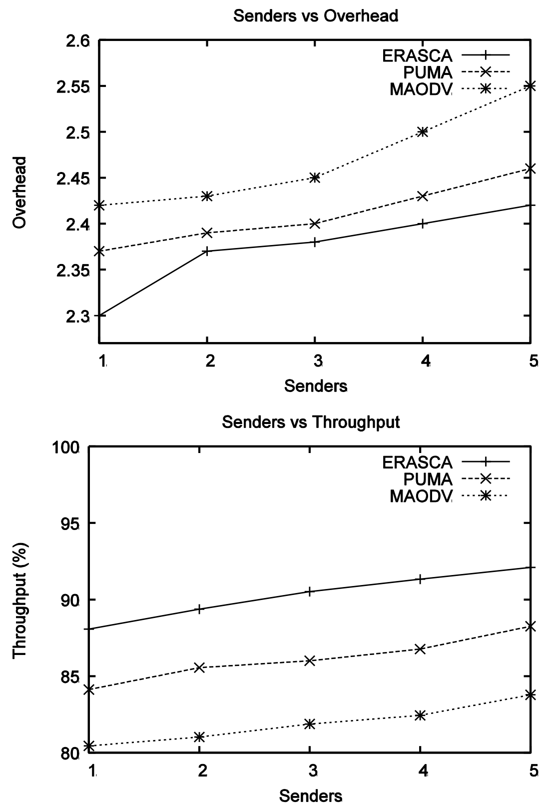


Fig. 7. Comparison of Senders with PDF, Overhead and Throughput

As shown in the Fig.7, throughput, PDF and overhead change with respect to senders. In routing when the numbers of sender increases, then the overhead and throughput also increase because of the inclusion of multiple packets from multiple senders. In Fig.7, MAODV performance is not satisfactory because of the single path between the sender and the receiver.

Since, in high mobility the possibility of link failures also increases between the source and the destination as compared to PUMA and ERASCA which are mesh based protocols. On the other hand, PUMA and ERASCA show a little difference in performance because of the redundant path availability between sender and receiver group. However, the little difference in performance is because of the frequent core failure situation in PUMA, as the inappropriate core is selected on bad location with low battery capacity.

C. Scenario 3

In Fig.8, when the number of receivers increases, then the overhead and throughput increases as it should be. The PDF increases because of the availability of multiple and short paths between the group of receivers, as well as it provide robustness to the network and decreases packet drop as compared to the long and fewer route. The maximum number of receivers also provides richer connectivity to the network; as a result, high throughput is achieved. As compared to PUMA and MAODV, ERASCA gives higher performance because of the appropriate core selection.

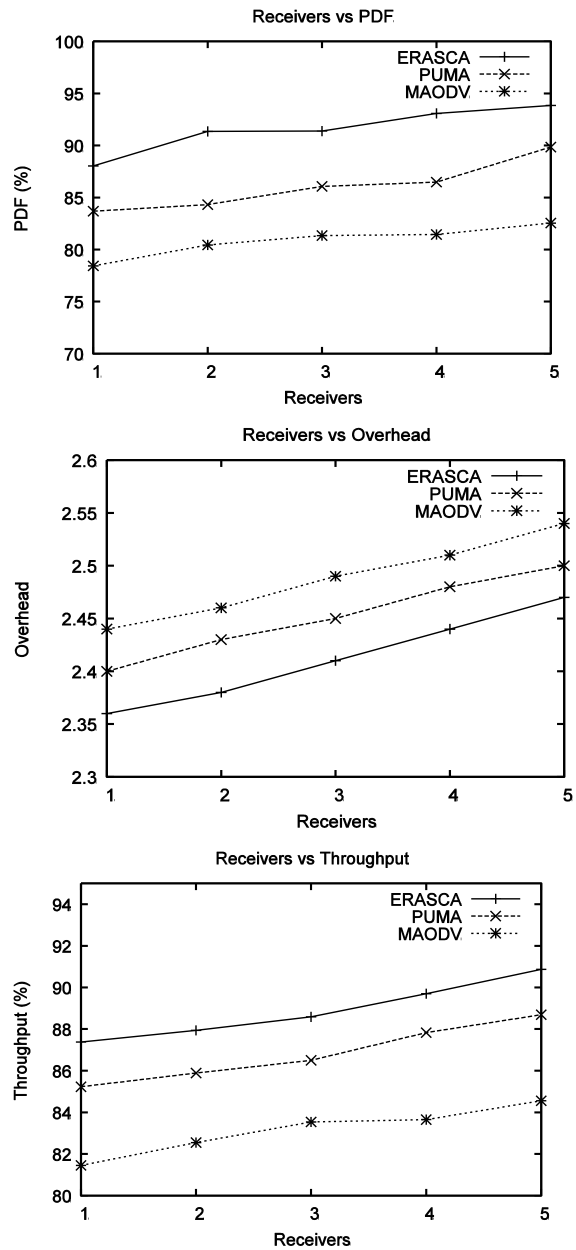


Fig. 8. Comparison of Receivers with PDF, Overhead and Throughput

The core selection is very important in MANET which affect network efficiency and lifetime of the network, but the core selection process are not efficient in PUMA and MAODV which deteriorate the performance of both protocols in the form of high overhead, as the selection of core in MAODV is not appropriate in location and energy wise. On the other hand, PUMA selects the core appropriately, but with minimum metrics and therefore it is believed that these approaches are not efficient because the selected core within a receiver may be in bad position in the network with minimum numbers of receivers and with low battery remaining. This selection increases the core failure and hence increases the packet drop and overhead. But in ERASCA, the core is selected within the best position in a receiver with high battery capacity; hence the core failure situation won't occur frequently and will improve the performance of ERASCA than PUMA and MAODV.

D. Scenario 4

Here the ifqLen is referring to the buffer size. At the start of the simulation, it is noticed that maximum packet drop occur in ERASCA, PUMA and MAODV because the smaller ifqLen represents a small buffer. Therefore, a large number of packets with small buffer ultimately increase the packet drop and hence decrease the PDF and increases the delay. Because of the maximum packet drop, sender will frequently transmit the data packet to the destination until the data is received by the destination. Thereby, increases the overhead by the frequent transmission from the source.

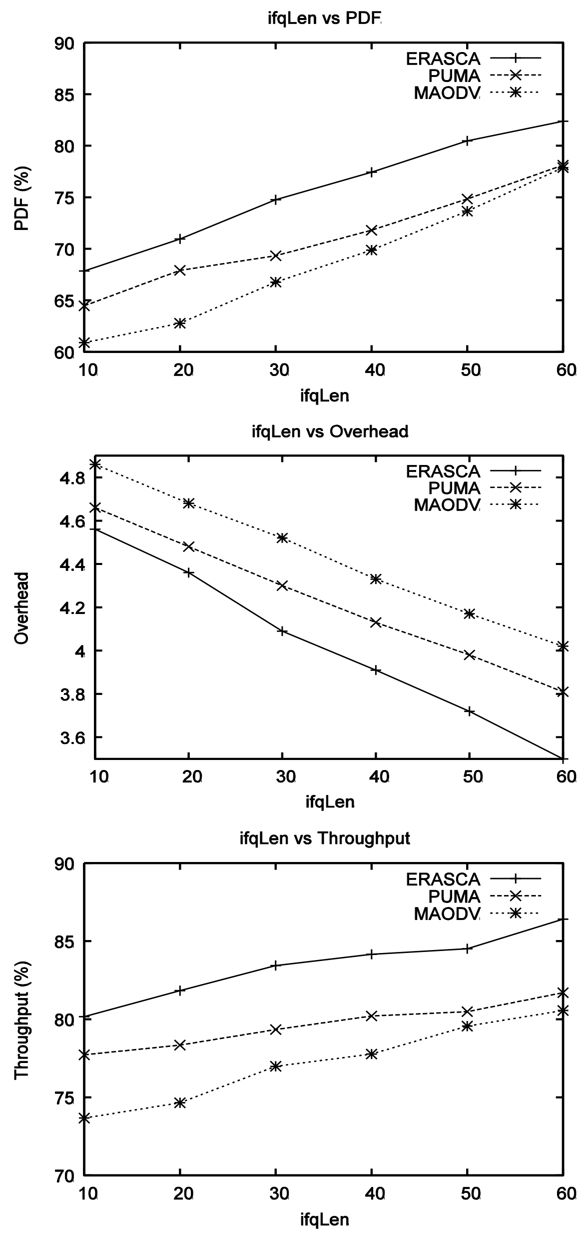
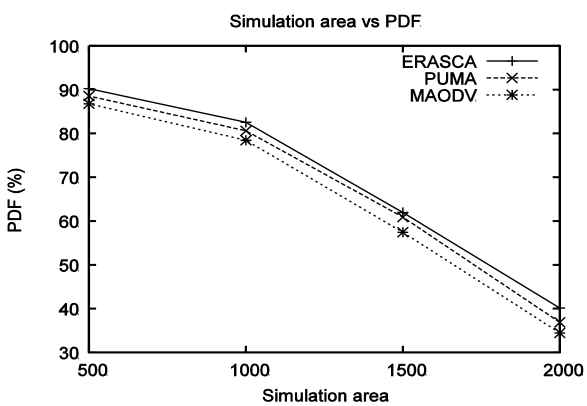


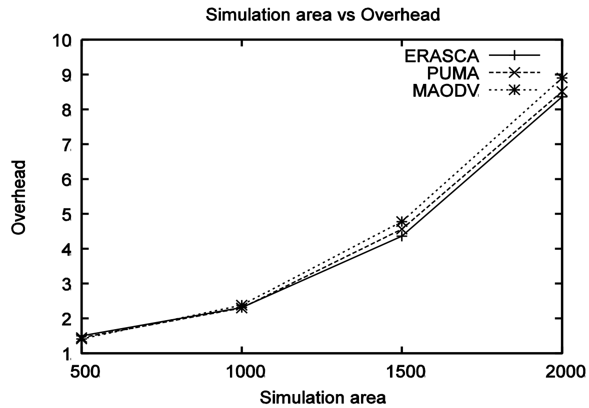
Fig. 9. Comparison of ifqLen with PDF, Overhead and Throughput

Fig.9 shows that with an increase in the ifqLen decreases the overhead, because with large ifqLen packet drop decreases and the data may successfully and frequently reaches from the sources to the group. As a result, flooding will be decreased and hence the overhead decreases. Therefore, in large queue, a large number of packets from source to destination are entertained and hence, the throughput increases. In ERASCA, the packet drop is less because of its core selection methodology. The appropriate core selection decrease core failure and hence minimize the flooding, as flooding reduces the packet drop, link failure and overhead with the increases in PDF and throughput.

E. Scenario 5

Multicast routing protocols generally show good performance within a small simulation area with shortest paths between senders and receivers, as data delivery latency and possibilities of link failure is low. Therefore, the throughput and PDF are increases and overhead decreases but in a large simulation area the throughput and PDF decreases, but increases the overhead and energy consumption with frequent packet drop, link failure and core failure. In such a situation core failure increases, as the distance between the receiver and the core increases. Hence, the regular reconfiguration for the next core node results in continuous flooding of control messages across the network, which increases the congestion, packet drop, delay and link failure. As a result, PDF and throughput decreases and increases the overhead. In Fig.10, with increase in the simulation area increases the link failure, resending of data from source to destination as well as frequent core failure. However, ERASCA shows better performance as compared to PUMA and MAODV because a stable core selection works well in large simulation area. As in the large simulation area a more stable core is required to minimize core failure because the frequent core failure in a large simulation area affects the performance of the network badly in term of link failure and delay. Hence, efficiency increases with improved lifetime of the network.





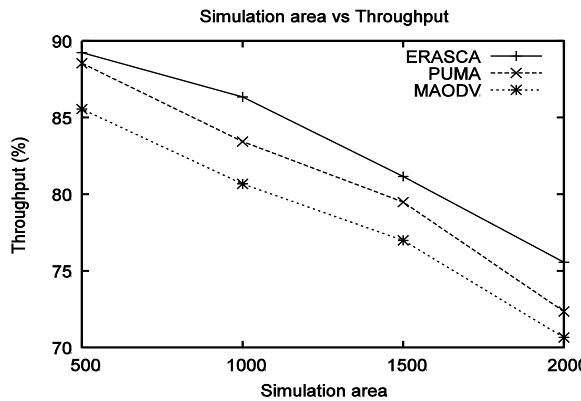


Fig. 10. Comparison of Simulation area with PDF, Overhead and Throughput

VI. CONCLUSION

Among the multicast routing protocols, ERASCA provides efficiency to MANET by reducing overhead. The ERASCA strongly relies on proper selection of core node in which the core is selected within the receiver group on the basis of multiple parameters like battery capacity and location in the network. As a result, a more stable core will be selected with high battery capacity and maximum numbers of neighbor. To increase the reliability of the network, the mirror core is introduced. Therefore, after the failure of the primary core the mirror core will take the responsibility as a primary core and will not affect the ongoing communication, hence minimizes the delay. ERASCA is compared with PUMA and MAODV; ERASCA demonstrated better performance with its core election process and in the presence of mirror core. Therefore ERASCA can be used efficiently and reliably in high mobility scenarios within large area irrespective of the number of receivers and senders with minimum packet drop and overhead with maximum reliability, throughput and PDF.

VII. Future Work

In future the ultimate plan is to secure the core election process. As the core election is an important and sensitive process and an adversary or malicious entities will always try to take over the position of core and can disrupt the core formation/ core election process by fabricating the SD messages and disseminating false data in the network/group for malicious purposes, Hence, in future a solution is propose in order to secure the core election process and counteract the malicious attacks, such as dissemination of false or fabricated information. Likewise in ERASCA, as soon as the data packet is received by any mesh member, it is flooded within the receiver group and ultimately the destined receiver will get the data soon but at the cost of overhead. The flooding is best in situation, when the mesh member (that receive the data packet) and the destined receiver are far away from each other. Therefore, the destined receiver will receive the data with minimum delay but with increase overhead. But in situation when the mesh member (that receive the data packet) and the destined receiver are near to each other, then multicasting is the better approach which will decrease the overhead. In ERASCA the flooding is preferred because of the location unpredictability between the mesh member and the destined receiver. Therefore, in future such a protocol should be design, where multicasting should be used within the mesh group with minimum delay, overhead and node prediction technique.

References

[1] A. Boukerche, "Algorithms and protocols for wireless, mobile Ad Hoc networks," John Wiley & Sons, vol. 77, 2008.

[2] D. M. N. Hemangini, "Study of Routing Protocols in Mobile Ad-Hoc Network," SSRG International Journal of Mobile Computing & Application (SSRG-IJMCA). vol. 2, 2015.

[3] M. A. Kodole and P. Agarkar, "A Survey of Routing Protocols in Mobile Ad-Hoc Networks," Multidisciplinary Journal of Research in Engineering and Technology, vol. 2, 2015.

[4] L. Junhai, X. Liu, and Y. Danxia, "Research on multicast routing protocols for mobile ad-hoc networks," Computer Networks, vol. 52, pp. 988-997, 2008.

[5] L. Junhai, Y. Danxia, X. Liu, and F. Mingyu, "A survey of multicast routing protocols for mobile ad-hoc networks," Communications Surveys & Tutorials, IEEE, vol. 11, 2009.

[6] C. E. Perkins, Ad hoc networking: Addison-Wesley Professional, 2008.

[7] P. Mohapatra, AD HOC NETWORKS: technologies and protocols: Springer Science & Business Media, 2005.

[8] M. Bouhorma, H. Bentaouit, and A. Boudhir, "Performance comparison of ad-hoc routing protocols AODV and DSR," International Conference on Multimedia Computing and Systems, pp. 511-514, 2009.

[9] Survey on Tree Based, Mesh Based and Stateless Multicast Protocols in MANET," International Journal of Innovative Research in Computer and Communication Engineering, vol. 2, 2014

[10] P. SAHU, "Disadvantage of AMRIS Protocol and its solution," International Journal of Engineering Research and Technology, 2012.

[11] C.-H. Huang, C.-T. Wu, K.-W. Ke, and H.-T. Wu, "MAODV-based multisource multicast routing with fast route recovery scheme in MANETs," in Computer Symposium (ICS), International, pp. 79-84, 2010.

[12] S. Mangai, A. Tamilarasi, and C. Venkatesh, "Dynamic core multicast routing protocol implementation using ant colony optimization in ad hoc wireless networks," International Conference on Computing, Communication and Networking,pp. 1-5, 2008.

[13] M.-A. Kharraz, H. Sarbazi-Azad, and A. Y. Zomaya, "On-demand multicast routing protocol with efficient route discovery," Journal of Network and Computer Applications, vol. 35, pp. 942-950, 2012.

[14] B. O. Reddy, C. V. Narayana, and S. S. Reddy, "Routing protocols classification for Ad Hoc Networks," International Journal of Advanced Research in Computer Science, vol. 2, 2011.

[15] P. P. M. Krishna and M. S. D. K. S. Prasad, "Mesh based and Hybrid Multicast routing protocols for MANETs: Current State of the art," Global Journal of Computer Science and Technology, vol. 12, 2012.

[16] S. K. Das, B. Manoj, and C. Murthy, "Weight based multicast routing protocol for ad hoc wireless networks," IEEE in Global Telecommunications Conference, GLOBECOM'02, pp. 117-121, 2002.

[17] A. Boukerche and K. Lu, "Optimized dynamic grid-based DDM protocol for large-scale distributed simulation systems,"19th IEEEProceedings inParallel and Distributed Processing Symposium, pp.6, 2005.

[18] O. S. Badarneh and M. Kadoch, "Multicast routing protocols in mobile ad hoc networks: a comparative survey and taxonomy," EURASIP Journal on Wireless Communications and Networking, p. 26, 2009.

[19] A. Antony, "Experimental Investigation Of Streaming Over Mobile Ad Hoc Networks Using PUMA," in International Journal of Engineering Research and Technology, 2013.

[20] S. Vasundra and B. Sathyanarayana, "Fast Recovery From Topology Changes And Communication Link Failures," i-Manager's Journal on Software Engineering, vol. 5, p. 50, 2010.

[21] M. Jahanshahi, M. Dehghan, and M. R. Meybodi, "LAMR: learning automata based multicast routing protocol for multi-channel multi-radio wireless mesh networks," Applied intelligence, vol. 38, pp. 58-77, 2013.

[22] A. J. Selvarani and P. A. Selvam, "Study of Routing Protocols in Mobile Ad Hoc Networks."

[23] A. M. A. Mo'men, H. S. Hamza, and I. Saroit, "A survey on security enhanced multicast routing protocols in Mobile Ad hoc Networks," in High-Capacity Optical Networks and Enabling Technologies (HONET), pp. 262-268, 2010.

[24] J. Garcia-Luna-Aceves and E. L. Madruga, "The core-assisted mesh protocol," Selected Areas in Communications, IEEE Journal on, vol. 17, pp.1380-1394,1999.

Application of Fuzzy Abduction Technique in Aerospace Dynamics

Sudipta Ghosh
Department of Electronics and Communication Engineering
Calcutta Institute of Engineering and Management
Kolkata, India

Binanda Kishore Mondal
Department of Instrumentation and Control Engineering
Calcutta Institute of Engineering and Management
Kolkata, India

Souvik Chatterjee
Department of Instrumentation and Control Engineering
Calcutta Institute of Engineering and Management
Kolkata, India

Debasish Kundu
Department of Electronics and Communication Engineering
Elite Polytechnic College
Kolkata, India

***Abstract*—The purpose of this paper is to apply Fuzzy Abduction Technique in aerospace dynamical problem. A model of an aeroplane is proposed for consideration at different air density level of the atmosphere and at different speed of the plane. Different air density of the atmosphere, angle of wings and speed of the plane are selected as parameters to be studied. In this paper a method is developed to determine the angle of wings of the plane with respect to its axis at different air density level of the atmosphere and at different speed of the plane. Data are given to justify our proposed method theoretically.**

***Keywords*—Fuzzy logic; Fuzzy abduction; Aerospace dynamics; Inverse Fuzzy relation**

I. Introduction

Recently a good number of researchers had introduced abduction technique in aerospace dynamical problem. Abduction is an important tool to solve various problem including diagnosis and natural language understanding [4] , [7] . Also it is applicable in high level reasoning such as hypothetical reasoning [10] and default reasoning [3] . Pople [1] bases his discussion of first – order logic and defined abduction as the procedure for derivation of hypothesis which explain a conjecture using an axiom set. Reggia [2] proposed abduction for diagnosis based on a relation between two sets . Bylander, et.al. [8] introduced plausibility (a map from the power set of hypothesis to a partially ordered set) to abduction based on relation. K.Yamade,et.al.[9] studied fuzzy abduction based on multi – valued logic and Y.Tsukamoro [11] has introduced fuzzy logic base on Lukasiewicz logic and its application to diagnosis and control . W.Pedrycz [13] has investigated numerical and applicational aspect of fuzzy relation equations and henceforth W.pedrycz [14] has introduced inverse problem in fuzzy relation equation . Afterwards Arnould, et.al.[15] have introduced "if…..then….." rule in case backward – chaining with fuzzy . Bugarin, et.al. [16] have investigated fuzzy reasoning supported by Petri nets. When an aircraft moves through the air it passes through different atmospheric layers. Density varies due to change of atmospheric layers. So the aircraft passes through different air density level. The goal of the present paper is to determine what would be the angle the wings and aircraft velocity to keep constant forward velocity at a certain height of the aircraft for a certain air density level

However, similar problems, as indicated above, in case of application of fuzzy abduction technique in aerospace dynamics have not been invested by any researcher in a similar approach. The authors considered when the air density is low then the speed of the aeroplane and angle of the wings are high and if air density is high the speed of the aeroplane and angle of the wings are low. If medium density of air is considered the speed of the aeroplane and angle of the wings are medium. Despite of importance of inference the concept has no standard definition.

In this paper authors propose a fuzzy abduction [18-23] method in aerospace dynamics. At first authors investigated modus pones and modus tolens in specific situation and showed that the inferred results were obtained as a truth value. Then the concept of derivation, fuzzy explanation and fuzzy abduction are introduced based on inference. Furthermore, the authors discussed necessary and sufficient conditions for the existence of fuzzy explanation and proposed a procedure to obtain approximate solution when the conditions are not satisfied.

In fine, the authors have illustrated this method by numerical examples.

II. Determination of Wing Angle and Speed Using Fuzzy Abduction

When an aircraft moves through the air, it propagates through different air density level because the atmospheric layers are changing. Suppose, at takeoff region the flight velocity is ω Km/m and air density level is medium. For this situation both the wings of the aircraft are at 90^{o} from the axis of its body, shown in fig.

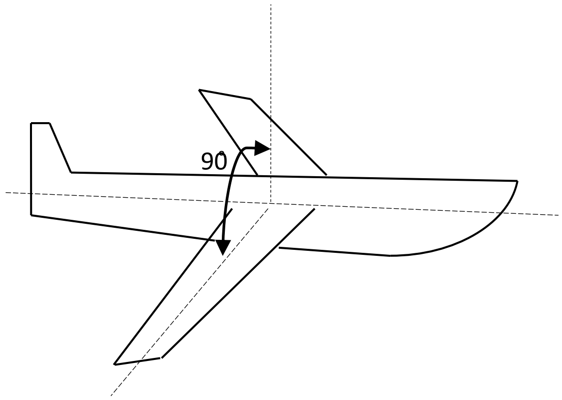


Fig. 1. Both the wings of the aircraft are perpendicular (90°) from the axis of its body

If the aircraft moves from this medium air density level to a lower air density level, it will moves down and the forward velocity will be decreased. To balance the height of the aircraft from the earth surface, the wings should be kept down to block more air and to develop a downward thrust so as to prevent the aircraft from moving down. Hence, both the wings of the aircraft should have an obtuse angle (*i.e.*, >90°) from the axis of its body, shown in fig.B

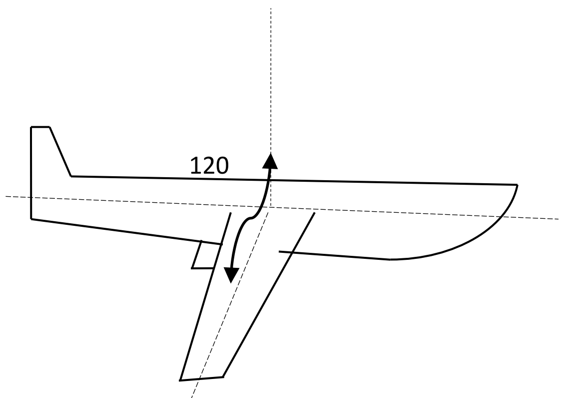


Fig. 2. Both the wings of the aircraft are at 120° angle from the axis of its body

On the other hand, if the aircraft moves from the medium air density level to a higher air density level, it will move upward and the forward velocity will be decreased. To balance the height of the aircraft from the earth surface, the wings should be kept up to pass more air and to release air thrust so as to prevent the aircraft from moving upward. Hence, both the wings of the aircraft should have an acute angle (*i.e.*, < 90°) from the axis of its body, shown in fig.C.

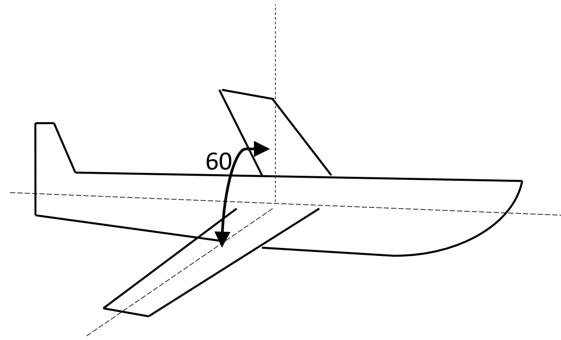


Fig. 3. Both the wings of the aircraft are at 60° angle from the axis of its body

III. RESULTS

From above discussion, we refer the problem as, what would be wings angle and aircraft velocity to keep constant forward velocity and the height of the aircraft for a certain air density level? To solve this problem it is needed to know about the relationship between the wings angle and air density level, and also the aircraft velocity and air density level. For that we set 6 fuzzy rules which are given below.

Rule 1: If wings angle is HIGH, air density will be LOW.

Rule 2: If speed is HIGH, air density will be LOW.

Rule 3: If speed is MEDIUM, air density will be MEDIUM.

Rule 4: .If speed is LOW, air density will be HIGH.

Rule 5: If wings angle is MEDIUM, air density will be MEDIUM

Rule 6: If wings angle is LOW, then air density will be HIGH.

Now, fuzzy membership curves for speed, wings angle and air density are defined.

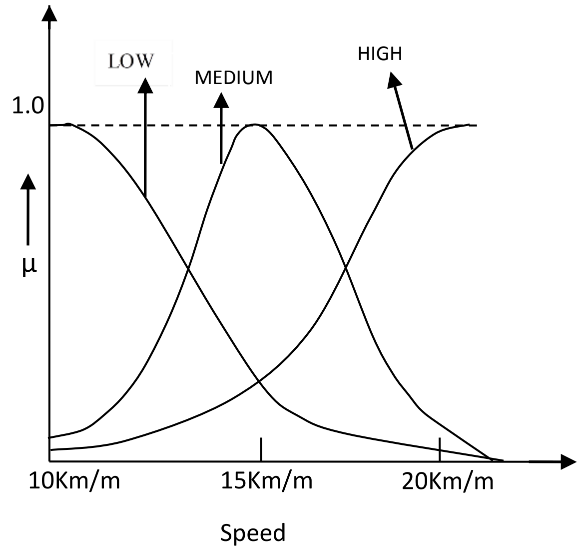


Fig. 4. Membership distribution curves of speed as LOW, MEDIUM and HIGH

From membership curve of speed we define LOW_SPEED, MEDIUM_SPEED and HIGH_SPEED as,

10 Km/m 15 Km/m 20 Km/m

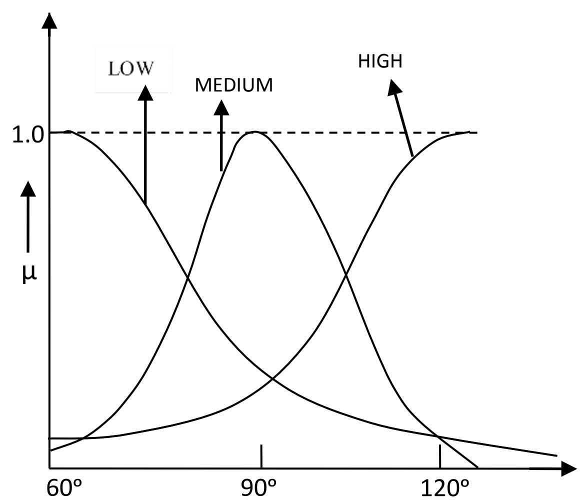
LOW_SPEED = [1.0 0.5 0.2]

10 Km/m 15 Km/m 20 Km/m

MEDIUM_SPEED = [0.1 1.0 0.2]

10Km/m 15Km/m 20Km/m

HIGH_SPEED = [0.1 0.5 1.0].



wings_angle

Fig. 5. Membership distribution curves of wings_angle as LOW, MEDIUM and HIGH

From membership curve of wings_angle we define LOW_WINGS_ANGLE, MEDIUM_WINGS_ANGLEand HIGH_WINGS_ANGLE as,

60^o 90^o 120^o

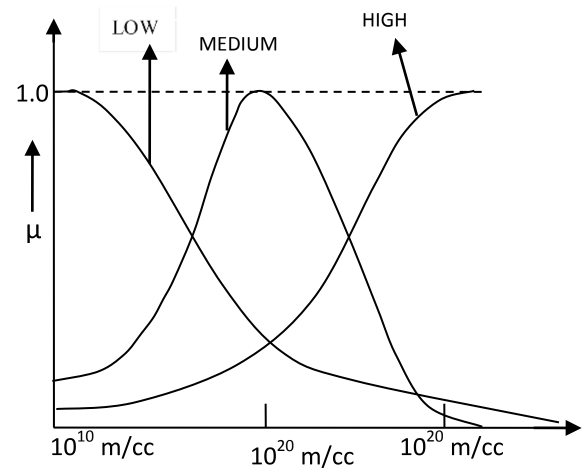
LOW_WINGS_ANGLE = [1.0 0.5 0.1]

60 90^o 120^o

MEDIUM_WINGS_ANGLE=[0.1 0.1 0.1]

60^o 90^o 120^o

HIGH_WINGS_ANGLE= [0.2 0.4 1.0].



air_density

Fig. 6. Membership distribution curves of air density as LOW, MEDIUM and HIGH

From membership curve of air_denity we define LOW_AIR_DENSITY, MEDIUM_AIR_DENSITY and HIGH_AIR_DENSITY as,

LOW_AIR_DENSITY=

10^{10}m/cc 10^{20}m/cc 10^{30}m/cc

[1.0 0.5 0.1]

MEDIUM_AIR_DENSITY=

10^{10}m/cc 10^{20}m/cc 10^{30}m/cc

[0.1 1.0 0.3]

HIGH_AIR_DENSITY=

10^{10}m/cc 10^{20}m/cc 10^{30}m/cc

[0.1 0.6 .01]

For Rule 1, using Mamdani's implication, we construct relational matrix R_2 as,

R_1=(HIGH_WINGS_ANGLE)^{T}o(LOW_AIR_DENSITY)

= $[0.2\ \ 0.4\ \ 1.0]^T$ o $[1.0\ \ 0.5\ \ 0.1]$

$$= \begin{bmatrix} 0.2 & 0.2 & 0.1 \\ 0.4 & 0.4 & 0.1 \\ 1.0 & 0.5 & 0.1 \end{bmatrix}.$$

For Rule 2, using Mamdani's implication, we construct relational matrix R_1 as,

R_2 = (HIGH_SPEED)^T o (LOW_AIR_DENSITY)

= [0.1 0.5 1.0]^T o [1.0 0.5 0.1]

$$= \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0..5 & 0.5 & 0.1 \\ 1.0 & 0.5 & 0.1 \end{bmatrix}.$$

For Rule 3, using Mamdani's implication, we construct relational matrix R_3 as,

R_3=(MEDIUM_SPEED)^{T}o(MEDIUM_AIR_DENSITY)

= [0.1 1.0 0.2]^T o [0.1 1.0 0.3]

$$= \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 1.0 & 0.3 \\ 0.1 & 0.2 & 0.2 \end{bmatrix}.$$

For Rule 4, using Mamdani's implication, we construct relational matrix R_5 as,

R_4 = (LOW_SPEED)^T o (HIGH_AIR_DENSITY)

= [1.0 0.5 0.2]^T o [0.1 0.6 1.0]

$$= \begin{bmatrix} 0.1 & 0.6 & 1.0 \\ 0.1 & 0.5 & 0.5 \\ 0.1 & 0.2 & 0.2 \end{bmatrix}.$$

For Rule 5, using Mamdani's implication, we construct relational matrix R_4 as,

R_5=(MEDIUM_WINGS_ANGLE)^{T}o(MEDIUM_AIR_DENSITY)

= [0.1 1.0 0.1]^T o [0.1 1.0 0.3]

$$= \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 1.0 & 0.3 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}.$$

For Rule 6, using Mamdani's implication, we construct relational matrix R_6 as,

R_6=(LOW_WINGS_ANGLE)^{T}o(HIGH_AIR_DENSITY)

= [1.0 0.5 0.1]^T o [0.1 0.6 1.0]

$$= \begin{bmatrix} 0.1 & 0.6 & 1.0 \\ 0.1 & 0.5 & 0.5 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}.$$

Now, say air density is 10^{25} m/cc, thus using α-cut we have

LOW_AIR_DENSITY=

(LOW_AIR_DENSITY)^0.2 = [0.2 0.2 0.1]

MEDIUM_AIR_DENSITY=

(MEDIUM_AIR_DENSITY)^0.5= [0.1 0.5 0.3]

HIGH_AIR_DENSITY= (HIGH_AIR_DENSITY)^0.8 = [0.1 0.6 0.8]

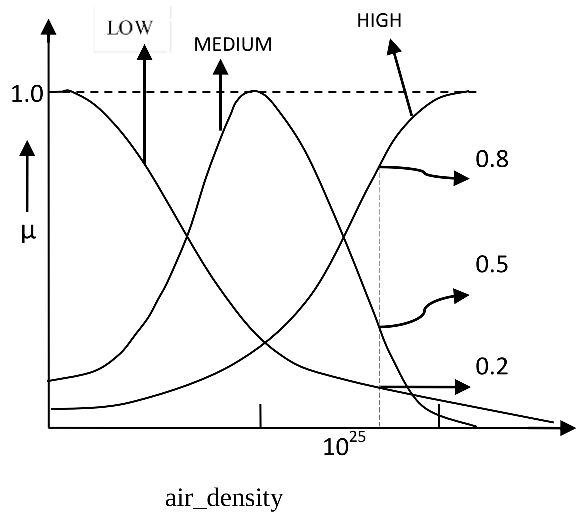


Fig. 7. α-cut for air density of 10^{25} m/cc

HIGH_SPEED = (LOW_AIR_DENSITY) o $(R_2)^{-1}$

$$= [0.2\ \ 0.2\ \ 0.2] \text{ o } \begin{bmatrix} 0.1 & 0.5 & 1.0 \\ 0.1 & 0.5 & 0.5 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$$

= [0.1 0.2 0.2].

MEDIUM_SPEED = (MEDIUM_AIR_DENSITY) o $(R_3)^{-1}$

$$= [0.2\ \ 0.5\ \ 0.1] \text{ o } \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 1.0 & 0.2 \\ 0.1 & 0.3 & 0.2 \end{bmatrix}$$

= [0.1 0.5 0.2].

LOW_SPEED = (HIGH_AIR_DENSITY) o $(R_4)^{-1}$

$$= [0.1\ \ 0.5\ \ 0.8] \text{ o } \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.6 & 0.5 & 0.2 \\ 1.0 & 0.5 & 0.2 \end{bmatrix}$$

= [0.8 0.5 0.2].

Hence, the membership distribution of speed = (HIGH_SPEED) $\cup$ (MEDIUM_SPEED) $\cup$ (LOW_SPEED)

= [0.1 0.2 0.2] $\cup$ [0.1 0.5 0.2] $\cup$ [0.8 0.5 0.2]

= [0.8 0.5 0.2].

Thus the speed

=(0.8×10+0.5×15+0.2×20)/(0.8+0.5+0.2)= 13 Km/m.

HIGH_WINGS_ANGLE = (LOW_AIR_DENSITY) o $(R_1)^{-1}$

$$= [0.2\ \ 0.2\ \ 0.2] \text{ o } \begin{bmatrix} 0.2 & 0.4 & 1.0 \\ 0.2 & 0.4 & 0.5 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$$

= [0.2 0.2 0.2].

MEDIUM_WINGS_ANGLE= (MEDIUM_AIR_DENSITY) o $(R_5)^{-1}$

$$= [0.2\ \ 0.5\ \ 0.1] \text{ o } \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 1.0 & 0.1 \\ 0.1 & 0.3 & 0.1 \end{bmatrix}$$

= [0.1 0.5 0.1].

LOW_WINGS_ANGLE = (HIGH_AIR_DENSITY) o $(R_6)^{-1}$

$$= [0.1\ \ 0.5\ \ 0.8] \text{ o } \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.6 & 0.5 & 0.1 \\ 1.0 & 0.5 & 0.1 \end{bmatrix}$$

= [0.8 0.5 0.1].

Hence, the membership distribution of wings_angle

=

(HIGH_WINGS_ANGLE) $\cup$ (MEDIUM_WINGS_ANGLE) $\cup$ (LOW_WINGS_ANGLE)

= [0.1 0.2 0.2] $\cup$ [0.1 0.5 0.1] $\cup$ [0.8 0.5 0.1]

= [0.8 0.5 0.2].

Thus the wings' angle

= $(0.8\times60^o+0.5\times90^o+0.3\times120^o)/(0.8+0.5+0.3) = 78^o$

IV. CONCLUSION

In Fuzzy Abduction technique, when we consider Aerospace Dynamics problem we faced several uncertainities.In this case,we consider when the plane goes through different layers we can see several changes have been occurred. Consider all of these problems we consider here fuzzy abduction method.

When the air density varied from low to high then the angle of wings also changed with the speed. wings angle and aircraft velocity to keep constant forward velocity and the height of the aircraft for a certain air density level we follow several rules of fuzzy abduction. From our simulation results we can calculate the wing's angle at different atmosphere.

ACKNOWLEDGEMENT

We are grateful to our college and friends at Calcutta Institute of Engineering and Management Kolkata,and sincere thanks to our principal Prof.(Dr.) Sibamay Dasgupta, who directly or indirectly is involved in the preparation of this paper.

REFERENCES

[1] D. Dubois and H. Prade, "Fuzzy sets and Systems: Theory and Applications," *Academic press*, NY, 1980.

[2] G. J. Klir and B. Yuan, "Fuzzy Sets and Fuzzy Logic: Theory and Applications," Prentice-Hall, NJ, 1995.

[3] W. Pedrycz and F. Gomide, "An Introduction to Fuzzy Sets: Analysis and Design," *MIT Press*, Cambridge, Massachusetts, 1998.

[4] T. J. Ross, "Fuzzy Logic with Engineering Applications," *McGraw-Hill*, NY, 1991.

[5] M. Togai and H. Watanabe, "Expert system on a chip: An engine for real time approximate reasoning," IEEE Expert, pp. 55-62, Fall 1986.

[6] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.

[7] H. J. Zimmermann, "Fuzzy Set Theory and Its Applications," Kluwer Academic, Dordrecht, The Netherlands, 1991.

[8] J. M. Mendel, R. I. Jhon, and F. Liu, "Interval Type-2 Fuzzy Logic System Made Simple," IEEE Trans. on Fuzzy Systems, vol. 14, pp. 808-821, no. 6, December 2006.

[9] J. M. Mendel, and H. Wu, "Type-2 Fuzzistics for Symmetric Interval Type-2 Fuzzy Sets: Part 1, Forward Problems," IEEE Trans. on Fuzzy Systems, vol. 14, pp. 781-792, no. 6, December 2006.

[10] J. M. Mendel, and H. Wu, "Type-2 Fuzzistics for Nonsymmetric Interval Type-2 Fuzzy Sets: Forward Problems," *IEEE Trans. on Fuzzy Systems*, vol. 15, pp. 916-930, no. 5, October 2007.

[11] T. Arnould, et al.: Backward-chaining withfuzzy "if... then..." rules, Proc. 2nd IEEE Inter. Conf. Fuzzy Systems, pp. 548-553 (1993).

[12] T. Arnould and S.Tano, "Interval-valued fuzzy backward reasoning," *IEEE Trans, Fuzzy Systems*, vol 3, no. 4, pp. 425-437, 1995.

[13] B. El Ayeb et al., "A New Diagnosis Approach by Deduction and Abduction," *Proc. Int'l Workshop Expert Systems in Eng.*, 1990.

[14] R. Bhatnagar, L.N. Kanal, "Structural and Probabilistic Knowledge for Abductive Reasoning**"**, IEEE Trans on Pattern Analysis and machine Intelligence**,** vol. 15 no. 3, pp. 233-245,March 1993.

[15] C. Boutillier and V. Becher, "Abduction as Belief Revision," Artificial Intelligence, vol. 77, pp. 43-94, 1995.

[16] T.Bylander,et.al.: The computational complexity of abduction, Artificial Intelligence, Vol.49, pp.25-60 (1991).

[17] Luis M. de Campos, José A. Gámez, and Serafin Moral, "Partial Abductive Inference in Bayesian Belief Networks—An Evolutionary Computation Approach by Using Problem-Specific Genetic Operators", IEEE Transactions On Evolutionary Computation**,** VOL. 6, NO. 2, APRIL 2002.

[18] S. Chakraborty, S. Gosh, A. Konar, S. Das and R. Janarthanan**, "**Extraction of Facial Features Based on Emotion using Fuzzy AbductiveReasoning",Smart Innovation,Systems andTechnologies,Volume27,pp 385-392,June 2014.

[19] S. Ghosh, G. Paul, A. Dutta and S. Ghosh, **"**Emotion Detection using Fuzzy Logic" *Journal of Mechanics of Continua And Mathematical Sciences* , vol.8,no.1,pp 1147-1165, July2013.

[20] S.Ghosh,A.Dutta,S.Roychowdhury,G.Paul, "Weather Prediction by the use of Fuzzy Logic**".***Journal of Mechanics of Continua And MathematicalSciences*,vol.8,no.2,pp 1228-1241, January.2014.

[21] S. Ghosh,A. Dutta, **"**Fault Detection Technique of Electronic Gadgets using Fuzzy Petrinet Abduction method"**.***Journal of Mechanics of Continua And Mathematical Sciences*,vol.9,no.1,pp 1264-1277, July 2014.

[22] S. Ghosh , N Das, D Kundu, G Paul **"**Fault Detection in Engineering Application using Fuzzy Petrinet and Abduction Technique**"** *Journal of Mechanics of Continua And Mathematical Sciences(ISSN 0973-8975)*,vol.9,no.2,pp 1368-1376, January 2015.

[23] S Ghosh, G Paul**, "**A Type-2 Approach in Emotion Recognition and an Extended Type-2 Approach for Emotion Detection". *Fuzzy Information and Engineering Fuzzy,vol 7,issu 4*, pp.475-498 2015.

Efficient Load Balancing Routing Technique for Mobile Ad Hoc Networks

Mahdi Abdulkader Salem
Department of Computer Sciences & Information Technology Sam Higginbottom Institute of Agriculture, Technology & Sciences, Allahabad India

Raghav Yadav
Department of Computer Science & Information Technology Sam Higginbottom Institute of Agriculture, Technology & Sciences, Allahabad India

***Abstract*—The mobile ad hoc network (MANET) is nothing but the wireless connection of mobile nodes which provides the communication and mobility among wireless nodes without the need of any physical infrastructure or centralized devices such as access point or base station. The communication in MANET is done by routing protocols. There are different categories of routing protocols introduced with different goals and objectives for MANETs such as proactive routing protocols (e.g. DSDV), reactive routing protocols (e.g. ADOV), geographic routing protocols (e.g. GRP), hybrid routing protocols etc. There are two important research problems with such routing protocols to address such as efficient load balancing and energy efficiency. In this paper, we are focusing on evaluation and analysis of efficient load balancing protocol design for MANET. Inefficient load balancing technique results in increasing routing overhead, poor packet delivery ratio, and other Quality of Service (QoS) parameters. In literature, there are a number of different methods proposed for improving the performance of routing protocols by efficient load balancing among mobile nodes communication. However, most of the methods suffer from various limitations. In this paper, we propose a novel technique for improved the QoS performance of load balancing approach as well as increasing the network lifetime. Evaluation of Network lifetime is out of scope of this paper.**

***Keywords*—AODV; MANET; Load balancing; throughput; packet delivery ratio; routing overhead**

I. Introduction

The term MANET (mobile ad hoc network) is defined as the wireless autonomous network in which nodes get connected by wireless links without using any physical infrastructure. MANET is a temporary network. Each node in MANET acts as both data sender or receiver and data forwarder in particular direction selected by routing protocol. All mobile nodes in network move randomly within a specific network area. Battlefield is one of the main area where MANET is widely used. For intercommunication purpose, such kind of networks do not need any extra support in the form of base stations or access points. Therefore, it is a total dynamic and infrastructure-free network. MANET network is nothing but a group of radio devices in which wireless communication is executed without any fixed physical foundation. The communication between the source mobile node and destination mobile node is not direct, but rather is carried by using intermediate nodes according to multi-hop communication approach. Direct communication can only be possible among neighboring nodes.

The mobile nodes in MANET are located randomly and continuously changing their positions in network. Thus, the interconnections among mobile nodes are also changing frequently. Such networks are thus self-organizing and self-configuring and one does not require central management for configuration purpose. In MANET, all nodes can communication each other using the wireless links. Due to the characteristics of MANET like allowing access to servies anywhere, anytime ubiquitously without need of any physical devices or platform, it is mainly used in crisis management services, military areas, conference halls, classrooms, etc. Development of multimedia applications like video conferencing and video on demand is possible only because of ad hoc networking developments of MANETs.

The communication in MANETs is possible due to use of routing protocols which help to discover the communication paths, select the paths, forward data on current paths, maintain the routes, and handle frequent changes in routes due to frequent nodes movements. The traditional and existing routing protocols did not addressed the issues related to QoS (Quality of Service). QoS is nothing but the level of performance of particular routing protocol of service providing to network end users. Many real time applications especially multimedia programs having the QoS requirements which must be achieved. The basic aim of QoS solutions is to get the improved deterministic behaviour of network with the objective of delivering the data carried by wireless network rightly, and utilization of network resources should be efficient. However, there is still the research challenge of maintaining the QoS solutions according to end users mobility. Many of the existing routing protocols presented so far for MANET are targeted either at minimizing the data traffic in wireless network or at reducing the number of hops taken to deliver packets.

The main reason of not providing the QoS solutions with existing routing protocols is that they are not designed with load balancing approach to cope with diver's conditions of MANET, mobility, data traffic, etc. For Mobile ad hoc networks, one of the important problem is load balancing. As we discussed earlier in this paper, the existing routing protocols do not have the processing of dealing with load balancing in MANET. Therefore, since the last two decades, a number of methods have been designed for load balancing in routing protocols. Due to emerging application requirements and also for reliable data transfer, load balancing is one of the key research areas in the field of MANETs. In MANETs, task finishing in particular is more complex if there is more traffic on mobile nodes with minimized processing capacity. There is no special technique for load sharing. Non-uniform processing or computing power of systems resulted in load imbalance in MANET. Sometimes, certain nodes in the network are overloaded and some nodes are completely idle. The mobile node with better processing capabilities can complete its tasks in quick time. Such nodes are treated that it does not load or less load all the time. Therefore, situation of nodes with less load are keep idle, and requirement of over loaded mobile nodes is objectionable.

Many routing approaches have been developed for load balancing in MANETs. Major research work is carried out by approaching the load balancing problem through congestion estimation and traffic control. Some approaches are used in energy and power metrics for making routing decision for load balancing. Clustering-based approaches also exist. Very few literatures use queue size, hop count, and bandwidth metrics for load balancing in mobile ad hoc networks. However, for MANET, there are two research challenges such as QoS improvement and energy efficiency. These two points were not efficiently addressed by existing methods. Therefore, it becomes a motivation for this research paper to present novel hybrid approach for load balancing in MANET with goal of improving QoS performance and energy efficiency performance.

The main goal of this paper is thus to present novel algorithm for efficient load balancing technique. This proposed method should address both load balancing and energy efficiency parallel. The proposed method has two important features such as method of link estimation proposed for energy efficiency improvement and another is learning of network load balancing in order to achieve the improved QoS solutions. These two contribution points are combined together in order to achieve both energy efficiency and improved QoS performance. However, in this paper, we keep the scope particularly with evaluation of QoS performance of efficient load balancing technique. Energy efficiency is out of scope of this paper. The rest of this paper is organized as follows: section II presents the related work studies over different load balancing techniques with analysis. Section III presents an overview of proposed framework, algorithm steps and design. Section IV presents the simulation studies and results achieved with different network conditions. Finally Section V presents the conclusion and discusses future work.

II. Related Work

In [1], the authors Yin and Lin presented the MALB technique which is based on multipath communication load balancing approach. For every discovered path, this protocol iteratively tracking the current traffic rate. The tracking of traffic rate is done for reducing the end-to-end delay performance in network.

In [2], authors present another technique which is based on similar approach presented in [1] for multipath communication based protocol.

In [3], authors Wu and Harms introduce the communication among the 2 node disjoint routes as the number paths among the nodes on different routes. From the practical analysis and results of this method, it can be seen that increasing the correlation results in increasing end to end delay for two numbers of paths. Therefore, to decrease this end to end delay performance, another approach is introduced in which traffic balancing is done around the least correlated routes.

In [4], [5], [6], different uni path based load balancing techniques are proposed by authors. In [4], various routing metrics are considered. In [5], packet caching approach is adopted. In [6], directional antennas approach is used.

In [7], the authors Zhu and Hassanein propose the new load balancing routing method known as LBAR. This protocol considers the nodal activity for routing metric from the total number of valid routes.

In [8], authors Lee and Riley presented approach for overloaded mobile nodes in which it is presented that such nodes would have freedom to forbidding the extra communications in order to make them load free by solving their overloaded condition. Therefore, every mobile device of MANET having the specific threshold value for making the decision on receipt of RREQ messages. There are number of other articles presented in which comparative study among multi path and single path load balancing techniques is discussed. Practically, the multipath-based load balancing methods providing the various benefits for improving the fault tolerance as well as reliability. However, it is showing that in [9], single path based techniques claim to be more efficient for load balancing.

In [9], author Pearlman, *et al.* introduced an approach for multipath-based routing method which is efficient if the alternate routes are disjoint. However, this is not simple to achieve in MANETs [10]. In [11], author Ganjali and Keshavarzian proposed that under any MANET with large number of mobile nodes the approach of multipath communication can address the load balancing more efficiently as compared to single path routing approach if there are huge number of routes utilized among all source and destination pairs.

In [12], the authors presented the performance evaluation multipath routing approach and reactive routing approach with load balancing technique.

In [13], author of this paper proposed method of load balancing in which number of realistic parameters like battery powers of every node, processing capabilities of every node, communication cost required for transferring the loads from overloaded nodes to under loaded mobile nodes.

In [14], author Saigal, *et al.* introduced another load balancing technique for MANET called as LARA (load aware routing in ad hoc). In this protocol, traffic density metric is utilized for presenting the contention degree at MAC layer. During the process of route setup, traffic density parameter is utilized for selecting the communication path with less traffic load.

In [15], the authors presented the details on selecting the right trade off among improved performance and increased routing overheads.

In [16], authors present the new technique for achieving the both improved reliability in case of path failures as well as multipath based load balancing routing in MANETs. This paper achieves the objectives through full use of multiple paths in MANETs to solve frequent paths failures problems as well as load balancing problem.

In [17], authors Chakrabarti and Kulkarni present an approach for designing the alternate paths that are maintained as well as used in the DSR protocol. This method also provides the QoS solutions by ensuring the proper bandwidth for data transfer even if mobile nodes are under the mobility.

In [18], Souinli, *et al.* introduced another technique of load-balancing which push data traffic from the network center. This approach delivered the new routing parameters which consider the mobile nodes centrality degree for reactive as well as proactive routing methods.

In [19], author Pham, *et al*. presented the multihop wireless communication networks in which IGW (internet gateway) method is used for providing the internet connectivity, wireless network linking with global Internet. However, for taking the benefits of capacity generated through the multiple gateways, routing protocol presented in [19] is required to balance the load efficiently between all the available IGWs in order to achieve the optimized network performance.

In [20], the authors Yoo, *et al.* introduced the load balancing technique called SLBA means simple load balancing approach. This method can easily added to any existing routing protocols (reactive only). This SLBA method reduces the traffic concentration by enabling every node for dropping the RREQ packets or for giving up the packet forwarding.

In [21], author Khamayseh, *et al.* presented a novel MLR (Mobility and Load aware Routing) method for reducing the impacts of broadcasting problem. Flooding process is controlled by MLR method by restricting messages of rebroadcast based on less speed as well as less loaded mobile nodes. In this method every mobile node takes decision on received RREQ message depending on number of parameters such as routing load, speed etc.

In [22], authors Cheng, *et al.* presented approach for formulating problem of dynamic load balanced clustering into the problem of dynamic optimization. To solve this problem, author presented the different types of dynamic genetic algorithms in MANETs.

III. Proposed Methodology

The flowchart in Figure 1 is showing the simulation work flow and comparative study parameters. To address the problem of achieving both efficient load balancing and energy efficiency we design and proposed novel algorithm called EELAR (energy efficient load aware routing) in which both factors traffic on mobile nodes and energy level of mobile nodes considered while communication. Algorithm 1 is our proposed algorithm.

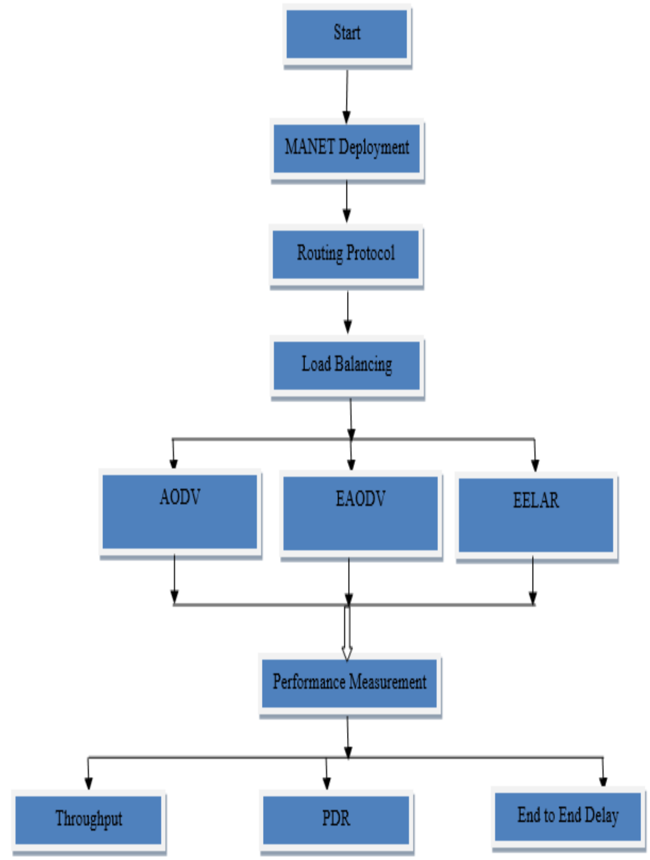


Fig. 1. Flowchart of simulation work

Algorithm 1: EELAR Method

Inputs:

Routing Table Entry,

packet p,

node ID,

Set energy threshold,

Set load threshold values.

Step 1: Extract the current packet details

Step 2: Define the routing table pointer

Step 3: Extracting the Destination Area (DA) by computing the depth from sink node

Step 4: Extracting the Forwarding Area (FA) by computing the sink node neighbours

Step 5: Finding the shortest path from source to destination

Step 6: Update routing table entries

Step 7: Apply Energy Efficiency function

7.1.: Before starting data transfer, convert all nodes except source node into the sleep state

7.2. If source node is ready to send data on selected active paths, then convert all nodes into active state from sleep state.

Step 8: Apply Load balancing function for data forwarding

Step 9: If energy level of any node goes below threshold or load on node goes particular threshold, then finds another alternate path in order to balance load or improved the network lifetime performance.

Step 10: If any node detects all its lower depth nodes below current threshold value, then it calculates new threshold and, start sending data on those paths again.

Step 11: Repeat this process still to the simulation ends.

Step 12: Stop.

IV. Experimental Results

A. Network Configurations

For practical work analysis, we used Network Simulator (NS2). In NS2, we implemented and evaluated the proposed EELAR protocol for comparative study purpose against existing AODV and EAODV routing protocols. This simulation is done on Ubuntu operating system and using NS-2.34 version. The performance of routing protocols evaluation is done based on various network scenarios and data communication approaches under the various network conditions. For this simulation study we have used two main parameters such as varying mobility speed and varying number of mobile nodes in network. We have designed two different network scenarios for evaluating the performance of proposed protocol which is named as EELAR. Tables 1 and 2 show the other configuration parameters used.

TABLE I. Simulation Configuration for Scenario 1-Varying Mobility Speed

Number of Nodes	50
Traffic Patterns	CBR (Constant Bit Rate)
Network Size (X * Y)	1000 x 1000
Simulation Time	100s
Transmission Packet Rate	10 m/s
Pause Time	1.0s
Routing Protocol	AODV/EAODV/EELAR
MAC Protocol	802.11
Channel Data Rate	11 Mbps
Mobility Speed	10 m/s to 50 m/s

TABLE II. Simulation Configuration for Scenario 2-Varying Mobile Nodes

Number of Nodes	20-100
Traffic Patterns	CBR (Constant Bit Rate)
Network Size (X * Y)	1000 x 1000
Simulation Time	25s
Transmission Packet Rate	10 m/s
Pause Time	1.0s
Routing Protocol	AODV/EAODV/EELAR
MAC Protocol	802.11
Channel Data Rate	11 Mbps
Mobility Speed	30 m/s

B. Simulation Results

We have compared the performance of three routing protocols using three performance metrics such as AODV, EAODV and proposed EELAR technique for load balancing QoS performance.

Scenario 1 Results:

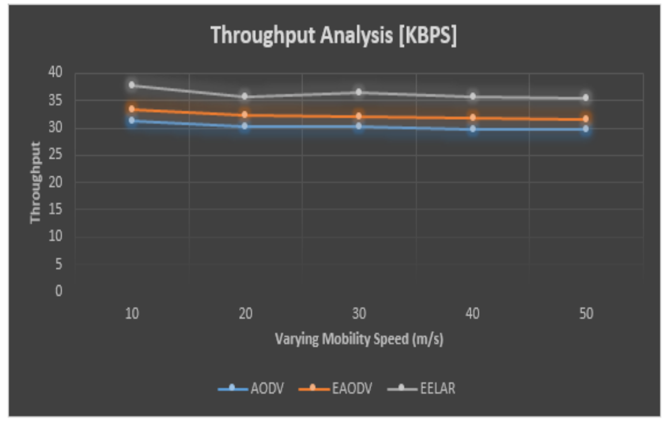


Fig. 2. Average throughput analysis for different load balancing methods of MANET

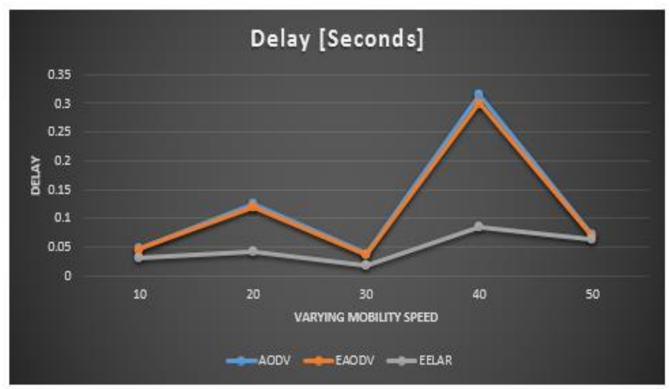


Fig. 3. Average delay analysis for different load balancing methods of MANET

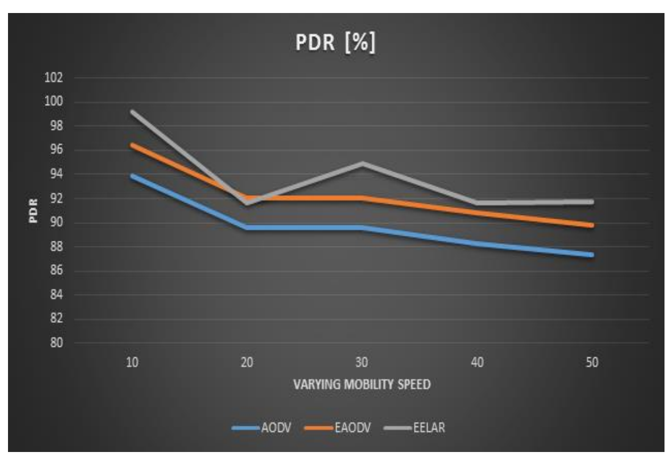


Fig. 4. Packet Delivery Ratio analysis for different load balancing methods of MANET

Figures 2, 3, and 4 show the performance analysis for average throughput, average end to end delay, and packet delivery ratio, respectively for three studied routing protocols such as AODV, EAODV, and EELAR. We vary the mobility speed of mobile nodes by keeping total number of mobile nodes 50 to each mobility speed. The results show that we have achieved better QoS performance for proposed EELAR protocol. For this first scenario. it is showing that performance of throughput and PDR is improved by 35 % as compared to EAODV protocol. The end to end delay is minimized by 15% to 18% as compared to EAODV protocol. Similarly, Figures 5, 6, and 7 show the results for throughput, delay and PDR respectively for network scenario 2. In the second scenario, the throughput shows an improvement of 38% and delay is minimized by 22% as compared to EAODV approach.

Scenario 2 Results

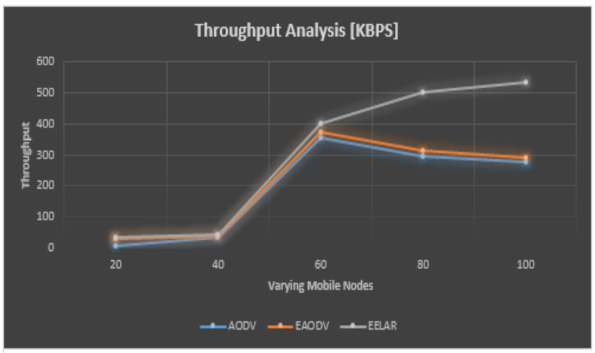


Fig. 5. Throughput analysis for different load balancing methods of MANET varying number of mobile nodes

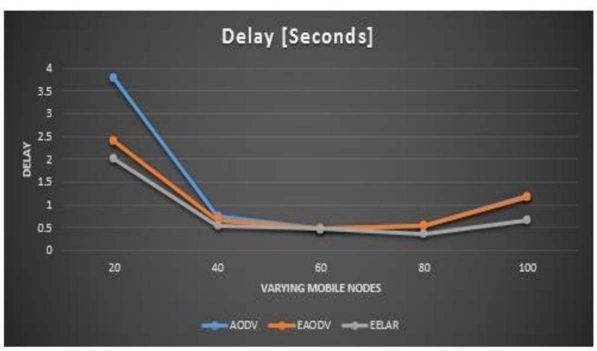


Fig. 6. Delay analysis for different load balancing methods of MANET varying number of mobile nodes

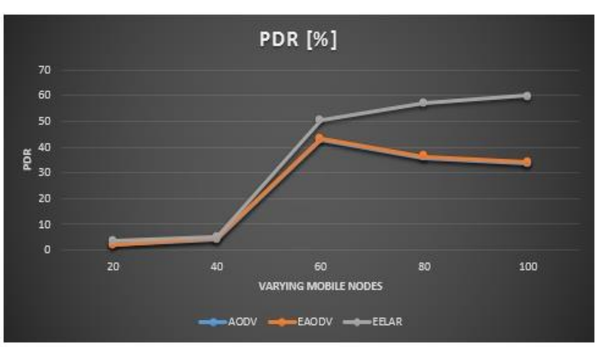


Fig. 7. Packet delivery ratio analysis for different load balancing methods of MANET varying number of mobile nodes

V. Conclusion

For MANET, load balancing technique plays a very vital role in order to achieve the QoS solutions. The traditional MANET routing protocols suffering from more routing overhead and decreased packet delivery ratio due to not addressing the load balancing in MANET communications. In this paper, we first presented the problems in MANET, then presented different solutions for load balancing techniques presented so far. We designed new load balancing technique for achieving the improved QoS performance as compared to existing EAODV and AODV routing protocols. The results section showing that we have simulated three protocols AODV, EAODV and proposed EELAR with two different network conditions. The results are compared by considering three important performance metrics of any routing protocol such as throughput, end to end delay and packet delivery ratio. In all cases, proposed load balancing approach shows improved performance when compared to existing methods.

References

[1] S. Yin, X. Lin, MALB: MANET adaptive load balancing, in: IEEE Vehicular Technology Conference (VTC2004-Fall), vol. 4, September 2004, pp. 2843–2847.

[2] L. Zhang, Z. Zhao, Y. Shu, L. Wang, O.W., W. Yang, Load balancing of multipath source routing in ad hoc networks, in: Proceeding of the IEEE International Conference on Communications (ICC 2002), May 2002.

[3] K. Wu, J. Harms, Performance study of a multipath routing method for wireless mobile ad hoc networks, in: Proceeding of the Ninth International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems (MASCOTS'01), August 2001.

[4] L. Wang, L.F. Zhang, Y.T. Shu, M. Dong, O.W.W. Yang, Multipath source routing in wireless ad hoc networks, in: Proceeding of IEEE CCECE, 2000, p. 479.

[5] A. Valera, W. Seah, S.V. Rao, Cooperative packet caching and shortest multipath routing in mobile ad hoc networks, in: Proceeding of IEEE INFOCOM, 2003.

[6] S. Roy, S. Bandyopadhyay, T. Ueda, K. Hasuike, Multipath routing in ad hoc wireless networks with omni directional and directional antenna: a comparative study, in: Proceeding of the Fourth International Workshop on Distributed Computing, Mobile and Wireless Computing (IWDC), 2002, pp. 184–191.

[7] A. Zhou, H. Hassanein, Load-balanced wireless ad hoc routing, in: IEEE Canadian Conference on Electrical and Computer Engineering, vol. 2, 2001, pp. 1157–1161.

[8] Y.J. Lee, G.F. Riley, A workload-based adaptive load-balancing technique for mobile ad hoc networks, in: IEEE Wireless Communications and Networking Conference (WCNC'2005), vol. 1, 2005, pp. 2002–2007.

[9] M. Perlman, Z. Haas, P. Scholander, S. Tabrizi, Alternate path routing for load balancing in mobile ad hoc networks, in: IEEE Military Communications Conference (MILCOM 2000), October 2000.

[10] P.Pham, S. Perreau, Multi-path routing protocol with load balancing policy in mobile ad hoc networks, in: IFIP Int'l Conference on Mobile and Wireless Communications Networks (MWCN 2002), September 2002.

[11] Y. Ganjali, A.Keshavarzian, Load balancing in ad hoc networks: single-path routing vs. multi-path routing, in: Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004), March 2004.

[12] P. Pham and S. Perreau, "Performance analysis of reactive shortest path and multi-path routing mechanism with load balance," IEEE Conference on Computer Communications (INFOCOM 2003), March 2003.

[13] Turgut, D.; Turgut, B.; Das, S.K.; Elmasri, R.; , "Balancing loads in mobile ad hoc networks," Telecommunications, 2003. ICT 2003, 10th

International Conference on , vol.1, no., pp. 490- 495 vol.1, 23 Feb.-1 March 2003.

[14] V. Saigal, A. Nayak, S. Pradhan, and R. Mall, "Load balanced routing in mobile ad hoc networks", Computer Communications, Vol.27, 2004, pp.295-305.

[15] Peter P.Pham, Sylvie Perreau , “Increasing the network performance using multi-path routing mechanism with load balance, Ad Hoc Networks, Volume 2, Issue 4, October 2004, Pages 433-459.

[16] Antonios Argyriou, Vijay Madisetti, “Using a new protocol to enhance path reliability and realize load balancing in mobile ad hoc networks”, Ad Hoc Networks, Volume 4, Issue 1, January 2006, Pages 60-74.

[17] Gautam Chakrabarti, Sandeep Kulkarni, “Load balancing and resource reservation in mobile ad hoc networks”, Ad Hoc Networks, Volume 4, Issue 2, March 2006, Pages 186-203.

[18] [Oussama Souihli, Mounir Frikha, Mahmoud Ben Hamouda, "Load-balancing in MANET shortest-path routing protocols", Ad Hoc Networks, Volume 7, Issue 2, March 2009, Pages 431-442.

[19] Vinh Pham ,Erlend Larsen ,Paal E. Engelstad, Øivind Kure, "Performance analysis of gateway load balancing in ad hoc networks with random topologies ", Proceedings of the 7th ACM international symposium on Mobility management and wireless access, 2009, pp.66-74.

[20] Younghwan Yoo, Sanghyun Ahn, Dharma P. Agrawal, "Impact of a simple load balancing approach and an incentive-based scheme on MANET performance", Journal of Parallel and Distributed Computing, Volume 70, Issue 2, February 2010, Pages 71-83.

[21] Yaser Khamayseh, Ghadeer Obiedat, Munner Bani Yassin, "Mobility and Load aware Routing protocol for ad hoc networks", Journal of King Saud University - Computer and Information Sciences, Volume 23, Issue 2, July 2011, Pages 105-113.

[22] Hui Cheng, Shengxiang Yang, Jiannong Cao, "Dynamic Genetic Algorithms for the Dynamic Load Balanced Clustering Problem in Mobile Ad Hoc Networks", Expert Systems with Applications, , 5 September 2012.

Hybrid Deep Network and Polar Transformation Features for Static Hand Gesture Recognition in Depth Data

Vo Hoai Viet, Tran Thai Son, Ly Quoc Ngoc
Department of Computer Vision and Robotics, University Of Science, VNU-HCM, Viet Nam

Abstract*—Static hand gesture recognition is an interesting and challenging problem in computer vision. It is considered a significant component of Human Computer Interaction and it has attracted many research efforts from the computer vision community in recent decades for its high potential applications, such as game interaction and sign language recognition. With the recent advent of the cost-effective Kinect, depth cameras have received a great deal of attention from researchers. It promoted interest within the vision and robotics community for its broad applications. In this paper, we propose the effective hand segmentation from the full depth image that is important step before extracting the features to represent for hand gesture. We also represent the novel hand descriptor explicitly encodes the shape and appearance information from depth maps that are significant characteristics for static hand gestures. We propose hand descriptor based on Polar Transformation coordinate is called Histogram of Polar Transformation (HPT) in order to capture both shape and appearance. Beside a robust hand descriptor, a robust classification model also plays a very important role in the hand recognition model. In order to have a high performance in recognition rate, we propose hybrid model for classification based on Sparse Auto-encoder and Deep Neural Network. We demonstrate large improvements over the state-of-the-art methods on two challenging benchmark datasets are NTU Hand Digits and ASL Finger Spelling and achieve the overall accuracy as 97.7% and 84.58%, respectively. Our experiments show that the proposed method significantly outperforms state-of-the-art techniques.

Keywords*—Hand Gesture Recognition; Deep Network; Polar Transformation; Depth Data

I. INTRODUCTION

Static hand gesture recognition which is an important component of Human Computer Interaction, has appealed many efforts invested from the research field of computer vision in recent decades for its strong potential in numerous applications, such as game interaction and sign language recognition. Hand gesture is a distinct and significant component of human action and hand gesture recognition since the information hand gestures convey is more sophisticated and linguistic than others. The goal of hand gesture recognition is to automatically analyze ongoing gesture from image. Generally speaking, hand gesture framework contains four main steps namely hand segmentation, feature extraction, gesture representation (gesture descriptor, dimension reduction …) and pattern classification. Though much progress has been made [5, 9, 16, 22, 24, 26], recognizing gesture with a high accuracy remains a challenging task due to the wide range of poses and considerable intra-class variations, e.g., rotation, scaling, viewpoint change and hand articulations. In the previous works, the authors have shown that deriving an effective gesture descriptor from image is a vital step for success of hand gesture recognition. There are two common approaches to extract gesture features [24]: appearance feature-based methods and shape feature-based methods.

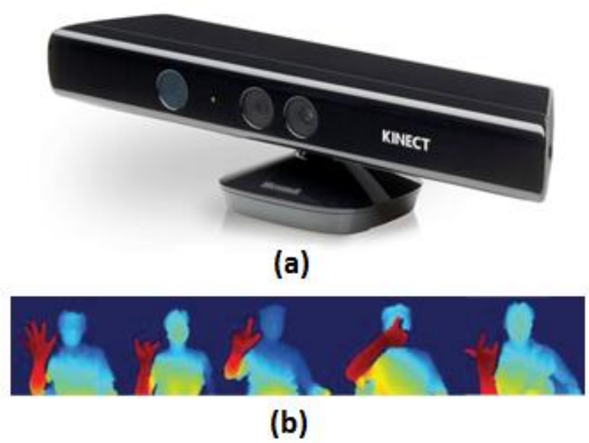


Fig. 1. Microsoft Kinect; b) Some depth images are captured by Kinect

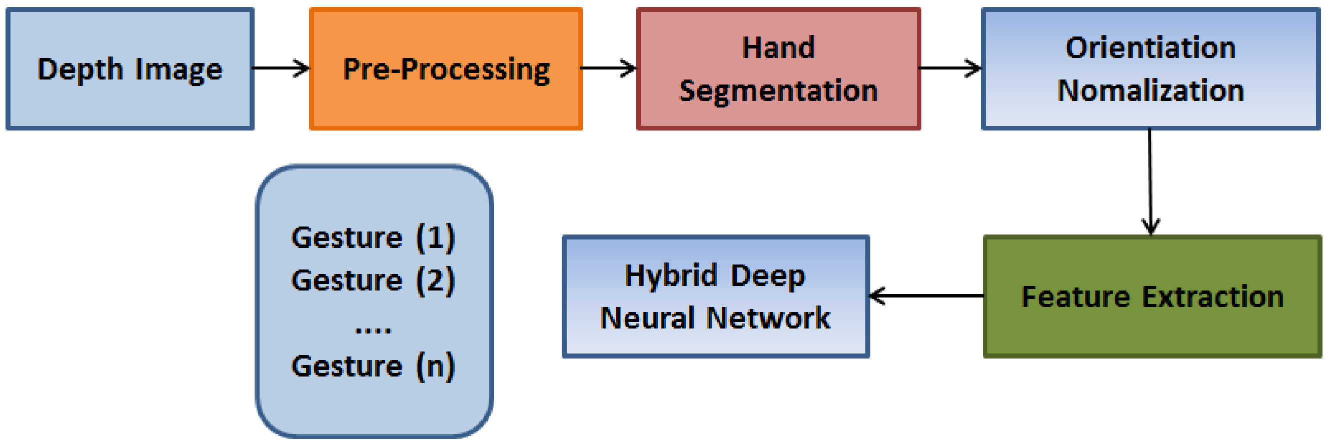


Fig. 2. Our framework of hand gesture recognition system

In all cases, moreover, it is commonly believed that in order to obtain high recognition rate, it is important to select an appropriate set of visual features that usually have to capture the particular properties of a specific domain and the distinctive characteristics of each object class. The most important aspect for any gesture recognition system is to seek an efficient feature representation. An extracted gesture feature can be considered an efficient representation if it could fulfill three criteria: firstly, it minimizes within-class variations while maximizes between-class variations; secondly, it can be easily extracted from the raw video; and thirdly, it can be described in a low-dimensional feature space to ensure computational speed during the classification step. The target of the feature extraction is to find an efficient and effective representation of the gesture which would provide robustness during recognition process.

Depth sensors have been available in many years. Though, they are used in limitation to many applications because of high cost and complexity of operations. However, the recent popularity of new 3D sensors such as Kinect [15] with low cost has alleviated the hardness of the traditional gesture recognition problem, by exploiting the depth data. With its advanced sensing techniques, this technology opens up an opportunity to significantly increase the capabilities of many automated vision-based recognition tasks [14]. And, it promoted interest within the vision and robotics community for its broad applications [14, 21]. In fact, this is a significant motivation for computer scientists to get deep in this research field to find out effective ways to utilize benefits from both the available depth and color information. Compared with conventional color data, depth maps provide several advantages, such as the ability of reflecting pure geometry and shape cues, or insensitive to changes in lighting conditions. Moreover, the range sensor provides 3D structural information of the scene, which offers more discerning information to recover postures and recognize gestures. These properties of depth data provide more natural and discriminative vision cues than color or texture. Furthermore, depth data has been demonstrated its capability to provide more information of object size, shape, and position.

However, feature extraction is just one of the significant steps to create a robust static hand gesture recognition system. Moreover, the classification is also considered as final step for the system and will determine the success of static hand gesture recognition system. This means that a powerful classifier will help to increase recognition rate of the system.

In this paper, we empirically study gesture descriptor based on Polar Transformation and projected views in depth data for gesture recognition. This descriptor combines both shape and appearance properties for gesture representation. And, hybrid deep neural network is deployed to classify gesture descriptors. The contributions of this paper are three-folds: firstly, we propose a method that is simple but effective for hand segmentation and orientation normalization. This method based on distance by horizontal and vertical region of interest. Secondly, we propose robust descriptors for hand gesture based on polar transformation. In this step, the depth map is projected onto three pre-defined orthogonal Cartesian planes and then normalized. After that, we compute the descriptors for each view and concatenate them into a feature vector. This captures appearance of gestures in creating hand descriptor. Finally, we propose hybrid deep network for gesture classification. In this model, we apply Sparse Auto-encoder (SAE) to pre-training for Deep Neural Network (DNN) so that we improve the performance of system.

The rest of this paper is organized as follows: in section II, we review related works. In section III, we introduce our approach for hand gesture recognition. In section IV, we show some results from our experiments and discussion. We conclude in section V.

II. Related Works

In recent years, sign language has been a popular topic in human behavior recognition. Many works have emerged as the American Sign Language recognition [10], the Portuguese Sign Language [19] and the Indian Sign Language [2]. Many hand gesture recognition methods which are based on visual information analysis have been proposed for hand gesture recognition [20]. The traditional approaches focused on using

RGB data. Sebastiean Marcel [19] proposed the approach based on Input-output Hidden Markov Models [23]. Moreover, the state of the art local features are also used by Chieh-Chih Wang et al. [4] and Y. Yao al. [27] such as SIFT [7] and SURF [3] with Adaboost algorithm.

Vision-based systems have been extensively researched and have been recently complemented with 3D sensors as Kinect [15]. Many research works have already used these popular sensors [5, 9, 16, 22]. One approach of recognizing hand gestures has used static depth frame as in [5, 9, 16, 22]. In [16], the authors treated each static depth frame as a regular gray scale image. They used a bank of Gabor filters to capture gradient information and solved the classification problem by random forests. In comparison to [18], the authors focused on a different type of information: contour [28] and a different application area: hand digits recognition. Without using gradients and contours, Hui Li [12] applied HOG [17] from RGB image to depth image and Zhang et al. [5] proposed a new descriptor to model hand gesture using histogram of 3D normals.

Feature extraction is just one of the significant steps to create a robust static hand gesture recognition system. Classification which is the final stage will play a very important role to the success of static hand gesture recognition system. In the classification stage, the traditional methods are used in many researches such as KNN, ANN, SVM, Adaboost and HMM... Although SVM is considered the state of art method for this stage and are used in many researches but deep learning which is an emerging trend in recent years is used in many researches with promising experimental results [6, 8, 13].

In this work, we capture both shape and appearance information to have hand descriptor. We use Polar coordinate system and depth data are projected onto three pre-defined orthogonal Cartesian planes and then they are normalized. Moreover, instead of using ANN, SVM, KNN… for classification, we apply deep network that are used in the recent years with some improvements. We apply hybrid deep network by deep neural network and pre-training it with Sparse Auto-encoder to improve the performance of system.

III. Proposed Method

The proposed hand gesture recognition system is shown in Fig. 2. Firstly, we use bilateral filter to remove noise. Secondly, we segment hand region from full depth image. Thirdly, we estimate the dominant orientation and achieve in-plane rotation invariance. Next, we extract histogram of Polar transformation to describe hand gesture. Finally, Hybrid Deep Network is used to identify the most likely class for input image.

A. Preprocessing

The 3D sensors such as Kinect based on structured light to estimate depth information, it is prone to be affected by noises due to reflection issues. These effects of noise could significantly decrease the overall performance of depth-based gesture recognition framework. Therefore, we firstly relieve the missing data and oultliers from the depth channel. As a result at [16], we adopted the bilateral filter for smoothing the depth channel. The bilateral filter [16] is a combination of a domain kernel, which gives priority to pixels that are close to the target pixel in the image plane, with a range kernel, which gives priority to the pixels which have similar labels as the target pixel. This filter is often useful to preserve edge information based on the range kernel advantages. The edge is important information to represent shape of gesture. The bilateral filter is defined as follows:

$$I^f(x) = \frac{1}{W_p}\sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|I(x_i) - I(x)\|)$$

$$W_p = \sum_{x_i \in \Omega} f_r(\|I(x_i) - I(x)\|) g_s(\|I(x_i) - I(x)\|)$$

Where I^f is the filtered image, I is the original input image , x are the coordinates of the current pixel to be filtered, Ω is the window centered in x, f_r is the range kernel for smoothing differences in intensities and g_s is the spatial kernel for smoothing differences in coordinates.In this research, f_r and g_s are supposed as Gaussian functions.

B. Hand Segmentation

The hand region extraction can be done in several ways, such as to retrieve a hand joint using a pose estimator or to filter a hand using skin color [21]. In hand detection phase, we do not use color-markers as the traditional methods. This is very important step in hand gesture recognition system. If we failed in this step, the following steps would be negatively affected and the system performance would be decreased. In this paper, the depth image generated by the camera is scaled to the range 0-255. Otsu's thresholding algorithm is applied to the depth histogram to segment the hand from the rest of the image. After thresholding the image, the pixel co-ordinates (x, y) and the corresponding un-scaled depth values (d) of the segmented hand region are extracted.

Because the depth of hand and neighbor region does not have large difference, region of interest will contain noise and unimportant region are captured. So, we propose the method to choose the interest region of hand to extract gesture descriptor and remove unimportant region. S is the interest region after segmentation with Otsu's thresholding algorithm. We calculate two distances and center of S as follows:

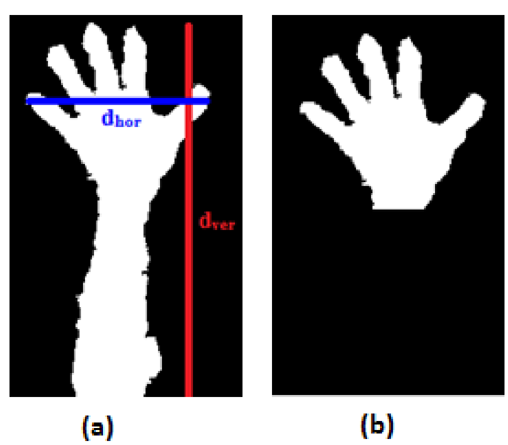


Fig. 3. Illustrate for hand region of interest segmentation: a) Hand region of Otsu's thresholding algorithm; b) Hand region of interest of our proposal

$$d_{ver} = \max(y) - \min(y)$$
$$d_{hor} = \max(x) - \min(x)$$

$$(\mathrm{x_{mean}}, \mathrm{y_{mean}}) = \frac{1}{|S|} \sum_{(\mathrm{x_i}, \mathrm{y_i}) \in S} (\mathrm{x_i}, \mathrm{y_i})$$

Hand region is segmented as follows: If d_{ver} greater than d_{hor} then we remove all the points that have y greater than y_{mean} . Otherwise, we remove all the points that have x less than x_{mean}.

C. Orientation Normalization

A big problem for static hand gesture recognition is the large intra-class variation incurred by hand rotations. The depth maps of the same gesture can significantly vary due to the in-plane rotation. Based on SIFT descriptor [7], we will assign dominant orientation for hand region. In order to estimate the dominant orientation and achieve in-plane rotation invariance, we compute the dominant depth gradient orientation as the normalization employed by SIFT descriptors [7] in 2D images. And we achieve the feature which is robust to the variety of rotation angle, scale, light conditions, viewpoints and noise.

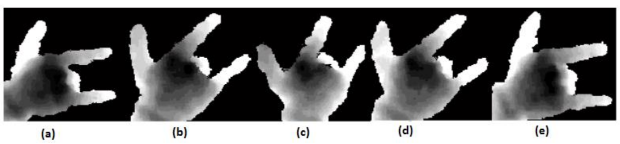


Fig. 4. Examples of one gesture have different orientations

D. Polar Feature Extraction

The shape is important property for hand gesture descriptor. A good descriptor has to contain the shape that can be described by gradient, edge… In this paper, in order to effectively describe the gesture shape, we utilize the polar coordinate system to capture the relative angles and distances between the salient points and the reference point for each gesture. This reference point is defined as the geometric center of the hand gesture and the relative distances are normalized by maximum distance on the support region, which makes the gesture descriptor insensitive to changes in scale of the hand gestures. The Cartesian coordinate system is transformed into the Polar coordinate system through the following equations:

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}$$

$$\theta_i = \tan^{-1}\left(\frac{y_i - y_c}{x_i - x_c}\right)$$

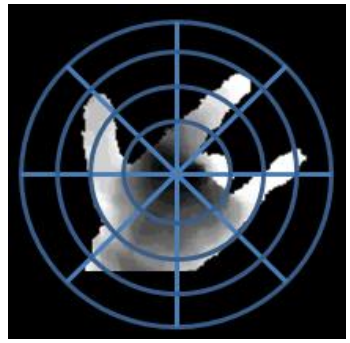


Fig. 5. Illustration of hand gesture descriptor is mapped onto polar coordinate system

where (x_i, y_i) is the coordinate of pixels in the Cartesian coordinate system. (r_i, θ_i) is the radius and the angle in the Polar coordinate system. (x_c, y_c) is the centre of the hand region. The center of the hand region can be calculated by:

$$x_c = \frac{1}{N} \sum_{i=1}^{N} x_i$$
$$y_c = \frac{1}{N} \sum_{i=1}^{N} y_i$$

where N is the total number of pixels.

In this paper, we compute the histogram of Polar coordinate system based on partition polar coordinate space into K cells by uniformly dividing each radius into R parts, and angles into A orientations such that K=A×R. Therefore, feature vector for hand gesture descriptor is K dimensions.

Moreover, in order to increase the discriminative descriptors, the depth map is projected onto three pre-defined orthogonal Cartesian planes and then normalized. After that, we transform each view into Polar coordinate and each Polar coordinate view is quantized by partitioning it into several cells with different radius and angles. This process will help capture the appearance information of hand gestures. The appearance is also importance properties to describe for hand gestures. So, this hand descriptor is extracted containing both shape and appearance information.

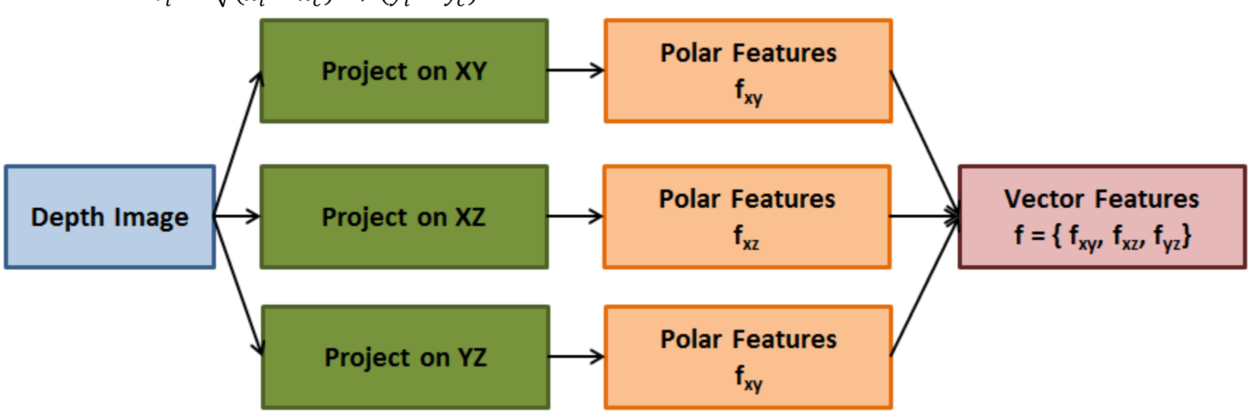


Fig. 6. Process of hand descriptor extraction on three views from hand region

E. Gesture Classification

The classification is the final step for the static hand gesture recognition system. To perform reliable recognition, there is first important problem that the features extracted from the training pattern are detectable that should have more descriptive and distinctive information. Besides, we need a good model for classifying between gestures to have a good recognition rate that accepted. The state of art method for classification is SVM have been used in many researches. However, deep learning which is an emerging trend is used in many researches with promising results in recent years. In this paper, we adopted deep neural network that is a kind of deep learning. Deep Neural Network is a neural network which has three or more hidden layers. In order to train deep neural network, a traditional way to train a deep neural network is an optimization problem by specifying a supervised cost function on the output layer with respect to the desired target. Neural Network is used to a gradient-based optimization algorithm in order to adjust the weights and biases of the network so that its output has low cost on samples in the training set. Unfortunately, deep networks trained in that manner have generally been found to perform worse than neural networks with one or two hidden layers [8, 13]. To overcome this problem, Dumitru Erhan et al. [6], answered the question "Why Does Unsupervised Pre-training Help Deep Learning?". The research indicates that pre-training is a kind of regularization mechanism, by minimizing variance and introducing a bias towards configurations of the parameter space that are useful for unsupervised learning [6, 13]. The greedy layer wise unsupervised strategy provides an initialization procedure, after which the neural network is fine-tuned to the global supervised objective.

The algorithm of the deep network training is decomposed in two steps:

- Step 1: greedily train subsets of the parameters of the network using a layer wise and unsupervised learning criterion, by repeating for each layer.
- Step 2: fine-tune all the parameters of the network using back-propagation and stochastic gradient descent.

In this paper, we adopted Sparse Auto-encoder [1] is unsupervised learning criterion to build deep neural network with 5 layers (3 hidden layers) as Fig. 7.

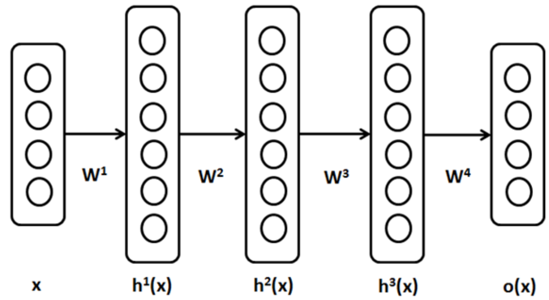


Fig. 7. Illustration of a deep neural network with 5 layers

IV. EXPERIMENTAL RESULTS

A. Data Set

We evaluate our approach on two benchmark datasets (NTU Hand Digits, ASL Finger Spelling) that we gather from the author's websites.

NTU Hand Digits dataset is the hand gesture dataset with a Kinect sensor. The dataset is collected from 10 subjects, and it contains 10 gestures. Each subject performs 10 different poses for the same gesture. Thus in total our dataset has 10 subject 10 gestures/subject 10 cases/gesture = 1000 cases, each of which consists of a color image and a depth map. Our dataset which is a very challenging real-life dataset is collected in uncontrolled environments. Besides, for each gesture, the subject poses with variations, namely the hand changes in orientation, scale, articulation, etc.

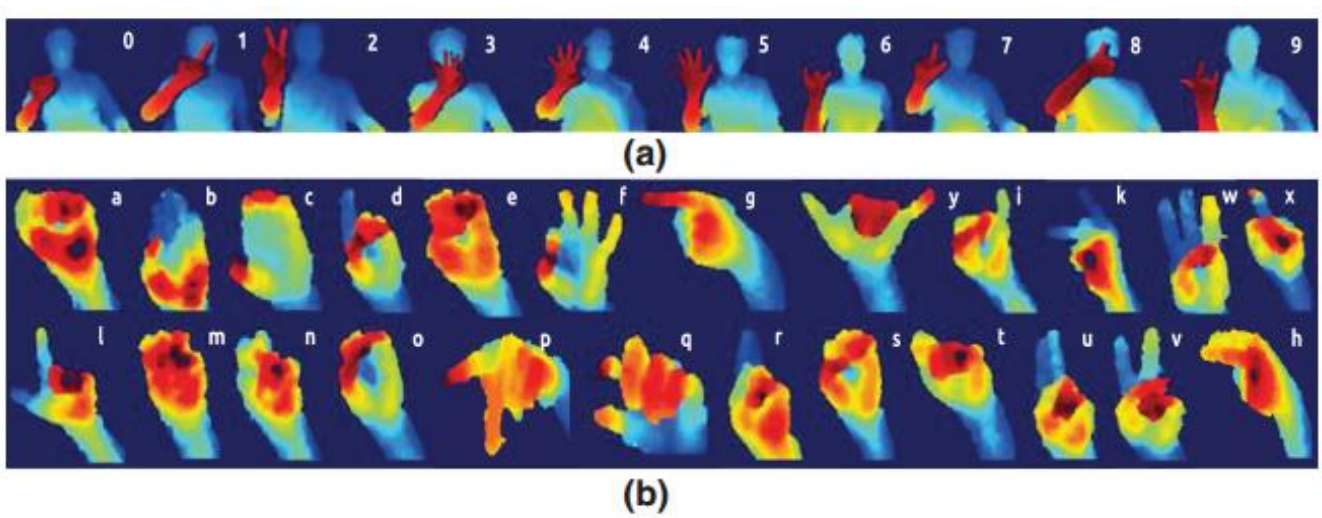


Fig. 8. (a) Some depth images from the NTU hand digits dataset. (b) Some depth images from the ASL finger spelling dataset

The ASL Finger Spelling dataset captures 60,000 hand gestures from 5 subjects. It includes 24 English letters from a to z, but with j and z discarded as these two letters in ASL are dynamic. And the dataset only provides the hand regions after segmentation. So, for the ASL Finger Spelling dataset, we skip the preprocessing step of hand segmentation as describe in section III. The dataset focus on for estimating generalization.

B. Evaluation Framework

In order to have the fair comparison with the other works, we use two experiments to compare with others. Firstly, subject independent test which uses the leave-one-out strategy, i.e., for a dataset with N subjects, N – 1 subject are used for training and the rest one for testing. This process is repeated for every subject and the average accuracy is reported. This test focuses on ability of generalization of approaches. Secondly, subject dependent test where all subjects are used in both training and testing, where the whole dataset is evenly split 50%-50% for training and testing. This test focuses the performance of approaches in the standard test in real-world the same human ability test are the things are learned then they will be tested.

For feature extraction, we adopted the number of orientations (A) and radius parts (R) for polar transformation of hand descriptor based on experiments, and a histogram with A×R bins is obtained for each polar coordinate system. So, the hand descriptor which is extracted for gesture representation from depth image is 3×A×R dimensions from three views are projected.

For classification, we use deep neural network with 5 layers (input layer, 3 hidden layers, and output layer) has the parameters such as the input layer is the number of feature vector, each hidden layer is 200 nodes, the number of output layer is the number of gesture classes in dataset (NTU Hand Digits is 10 nodes and ASL Finger Spelling is 24 nodes), the learning rate is 0.2, and the number of loop is 1000. In order to improve performance, we adopted Auto Sparse-encoder to pre-train deep neural network is proven that will have better than traditional methods without pre-training.

C. Experiment Results

We firstly evaluate our proposed approach on the two benchmark hand gesture datasets. Then we compare our experimental results to the-state-of-the-art methods to prove the effectiveness and robust of the proposed method.

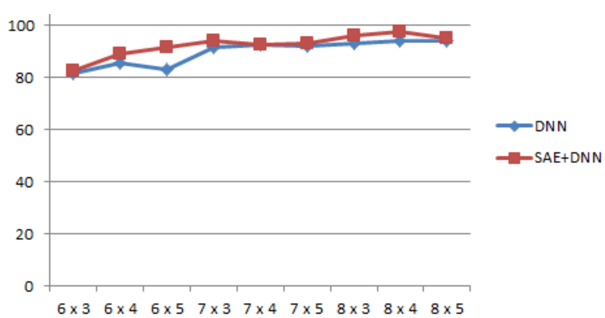


Fig. 9. Accuracies (%) of hand gesture recognition on the NTU Hand Digits dataset under different AxR of gesture descriptors, from 6x3 to 8x5

In this research, we present static hand gesture descriptor as a mixture of these two properties: 1) shape of the hand; 2) appearance of the hand. These properties are extracted frompolar coordinate system and depth image is projected on three views. The relative importance of these elements is based on the nature of the gestures that we aim to recognize. From this experimential results, we argue that no one single category of feature can deal with all kinds of static hand gesture datasets equally well. So, it is quite necessary and useful to combine different categories of features to improve the static hand gesture recognition performance. Moreover, we need a robust classification model to have a good performances. Tables I, and II give our experimental results on NTU Hand Digits, and ASL Finger Spellingdataset on both dependent and independent test. However, the same approach has the different result on the different dataset. This is the different characteristics of these datasets.The ASL Finger Spelling dataset has a larger data scale than NTU Hand Digits about the number of classes and samples.

To study the effect of the two parameters A and R in polar features, we choose the parameters from 6×3 to 8×5 on both NTU and ASL Finger Spelling datasets (see Fig. 9 and 10). The experimental results show that A=8 and R=4 are the best parameters on both datasets.

TABLE I. EXPERIMENTAL RESULTS OF OUR METHODS ON NTU DATASET

Method	Accuracy	
	Subj. Indep.	Subj. Dep.
Polar + DNN	94.2%	99.33%
Polar + SAE + DNN	97.7%	100%

TABLE II. EXPERIMENTAL RESULTS OF OUR METHODS ON ASL FINGER SPELLING DATASET

Method	Accuracy	
	Subj. Indep.	Subj. Dep.
Polar + DNN	78.17%	99.8%
Polar + SAE + DNN	84.58%	100%

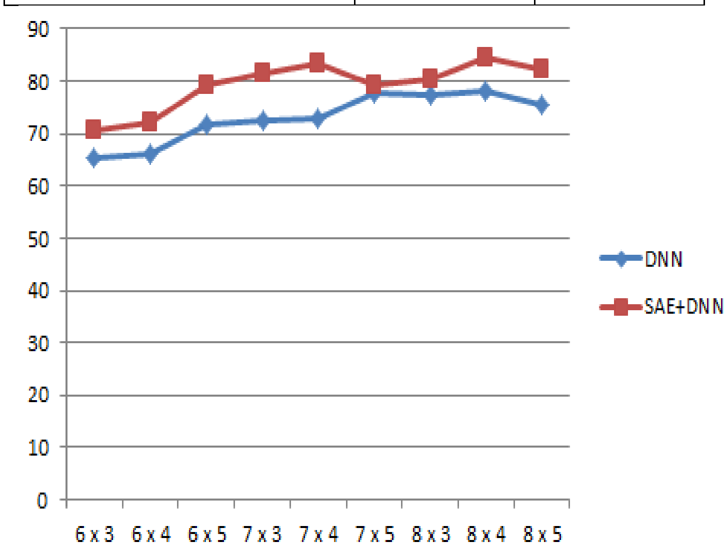


Fig. 10. Accuracies (%) of hand gesture recognition on the ASL Finger Spelling dataset under different AxR of gesture descriptors, from 6x3 to 8x5

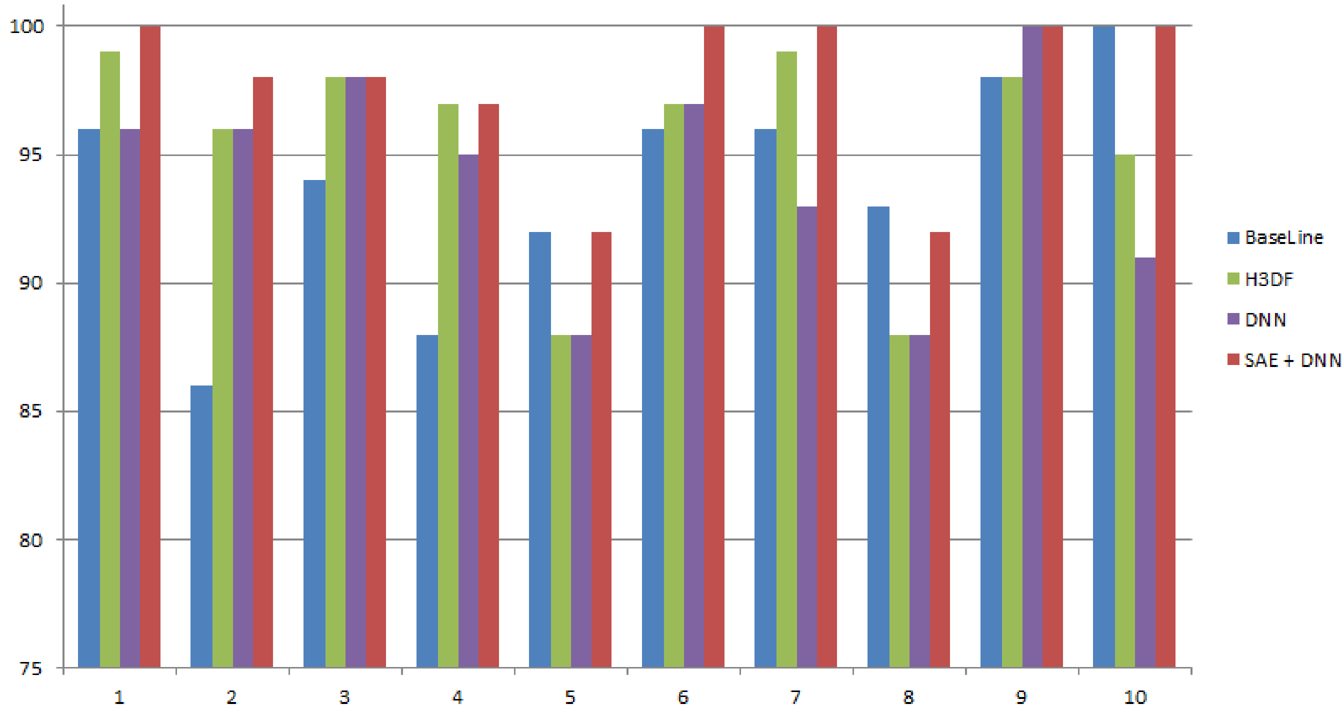


Fig. 11. Comparison of our proposed method with the baseline method in [28] and H3DF [5] on NTU Dataset

Tables III, and IV compare our experimental results with state-of-the-art results on NTU Hand Digits and ASL Finger Spelling dataset respectively. On NTU Hand Digits, our recognition rate is 97.7% on subject independent test and 100% on subject dependent test, more than the current best rate by 2.2% and 0.8%. On ASL Finger Spelling, however, our recognition rate is 84.58% and 100% more than the current best rate by 11.28% and 1.1%. Recognition rate has been improved significantly accuracy in independent test on both datasets, this shows that our approach is effective and stable on cross-dataset with the same configuration. In addition, our approach extracts gesture features based on these algorithms that is rapidly implementation and low computational cost with compact feature vector compare in comparison to existing techniques. From above experimental results, we argue that a successful gesture recognition system not only extract a robust descriptor contains both shape and appearance information but also has a robust classification model. Furthermore, the experimental results at [5, 28] and our approach also show that subject dependent tests significantly outperforms subject-independent tests and is more stable to the changes of locality. This shows a nature of training prolem in real-word that we have to get a base knowlegde about subjects in order to have a good performance when apply into complex and various subjects in real-word.

TABLE III. EXPERIMENTAL RESULTS ON NTU DATASET

Method	**Accuracy**	
	Subj. Indep.	**Subj. Dep.**
Contour Matching [28]	93.9%	N/A
HOG [5]	93.1%	96.4%
H3DF [5]	95.5%	99.2%
Polar + DNN	**94.2%**	**99.33%**
Polar + SAE + DNN	**97.7%**	**100%**

TABLE IV. EXPERIMENTAL RESULTS ON ASL FINGER SPELLING DATASET

Method	**Accuracy**	
	Subj. Indep.	**Subj. Dep.**
Contour Matching [28]	49.0%	N/A
HOG [5]	65.4%	96.0%
H3DF [5]	73.3%	98.9%
Polar + DNN	**78.17%**	**99.8%**
Polar + SAE + DNN	**84.58%**	**100%**

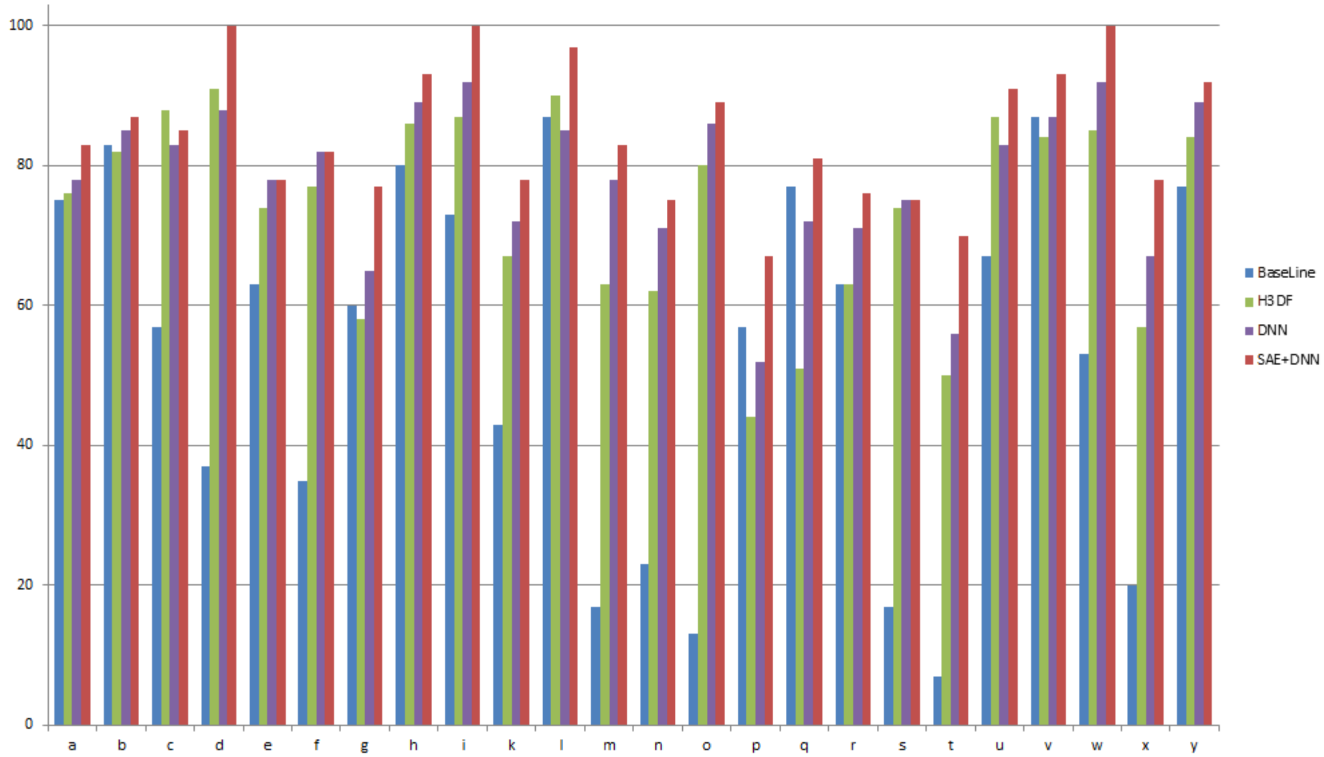


Fig. 12. Comparison of our proposed method with the baseline method in [28] and H3DF [5] on ASL Finger Spelling Dataset

V. Conclusion

In this paper, we represent a novel approach for recognizing static hand gestures based on polar transformation and deep neural network in depth data. Our proposed method consisted of steps as follows: firstly, we use bilateral filter to smooth depth data. Secondly, we segment hand region from full depth and orientation normalization based on depth gradient the same SIFT's idea. Thirdly, we represent a gesture based on using polar transformation for three views are projected from hand region and concatenate them into a feature vector. Our descriptor captures shape and appearance information that are a robust characteristics to distinguish between gestures. Finally, hybrid model is applied for gestures classification to improve the performance of system. Specify, we use deep network with sparse auto encoder for pre-training stage. In this framework, we have exploited the powerfulness of polar transformation in gesture descriptor and effectiveness of deep learning in classification stage. We have evaluated the effectiveness of our proposed on two public hand gesture recognition datasets. Our experimental results achieve superior performance to the state-of-the-art algorithm on NTU Hand Digits and ASL Finger Spelling datasets on overall accuracy as 97.7% and 84.58%, respectively. In addition, our approach is fast and compact in feature descriptor thus it is suitable for real-time hand recognition.

In the future, we will fuse with RGB features to improve the performance of system and extend this descriptor to the temporal domain to capture motion properties in recognizing dynamic hand gesture from depth videos. In addition, we also consider applying feature learning into the system.

Acknowledgment

This research is funded by Vietnam National University, Ho Chi Minh City (VNU-HCM) under grant number *B2014-18-02*.

References

[1] Andrew Ng, CS294A Lecture notes: Sparse Autoencoder, 2011.

[2] A. S. Ghotkar, R. Khatal, S. Khupase, S. Asati, and M. Hadap. Hand gesture recognition for indian sign language. In Proc. Int Computer Communication and Informatics (ICCCI) Conf, pages 1–4, 2012.

[3] Bay, H., Tuytelaars, T., and Van Gool, L .SURF:Speeded Up Robust Features", In Proceedings of the Ninth European Conference on Computer Vision, May, 2006.

[4] Chieh-Chih Wang and Ko-Chih Wang, Hand Posture Recognition Using Adaboost with SIFT for Human Robot Interaction, Recent Progress in Robotics: Viable Robotic Service to Human, pp 317-329, 2008.

[5] C. Zhang, X. Yang and Y. Tian. Histogram of 3D Facet: A Characteristic Descriptor for Hand Gesture Recognition, IEEE International Conference on Automatic Face and Gesture Recognition, 2013.

[6] Dimitru Erchan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol and Pascal Vincent, Why Does Unsupervised Pre-Training Help Deep Learning?, Journal of Machine Learning Research, 2010.

[7] D. G. Lowe. Distinctive image features from scale-in variant keypoints. International Journal of Computer V ision (IJCV), 60(2):91–110, 2004.

[8] Erhan, D., Manzagol, P.-A., Bengio, Y., Bengio, S., & Vincent, P.(2009b). The difficulty of training deep architectures and the effect of unsupervised pre-training. AISTATS'2009 (pp. 153–160), 2009.

[9] F. Domino, M. Donadeo, P. Zanuttigh, Combining multiple depth based descriptors for hand gesture recognition, Pattern Recognit. Lett. 50 (2014) 101–111, 2014.

[10] F. Ullah. American sign language recognition system for hearing impaired people using cartesian genetic programming. In Proc. 5th Int Automation, Robotics and Applications (ICARA) Conf, pages 96–99, 2011.

[11] H. Liang, J. Yuan, D. Thalmann, Parsing the hand in depth images, in: IEEE Trans. on Multimedia (T-MM), 2014.

[12] Hui Li, Lei Yang, Xiaoyu Wu, Shengmiao Xu, Youwen Wang, "Static Hand Gesture Recognition Based on HOG with Kinect", International Conference on Intelligent Human- Machine Systems and Cybernetics, 2012.

[13] Hugo Larochelle, Yoshua Bengio, Jerome Louradour and Pascal Lamblin, Exploring Strategies for Training Deep Neural Networks, Journal of Machine Learning Research, 2009.

[14] Leandro Cruz, Djalma Lucio, Luiz Velho: Kinect and RGBD Images: Challenges and Applications. SIBGRAPI Tutorials, pp 36-49, 2012.

[15] Microsoft Kinect. http://www.xbox.com/kinect, 2012.

[16] M. Camplani and L. Salgado. Efficient spatio-temporal hole filling strategy for kinect depth maps, A. M. Baskurt and R. Sitnik, Eds., vol. 8290, no. 1. SPIE, p. 82900E, 2012.

[17] N. Dalal, and B. Triggs. Histogram of Orientated Gradients for Human Detection, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886-893, 2005.

[18] N. Pugeault and R. Bowden. Spelling it out: Real-time asl finger spelling recognition, In ICCV Workshops, 2011.

[19] P. Trindade and J. Lobo. Distributed accelerometers for gesture recognition and visualization. In DoCEIS'11 - Doctoral Conference on Computing, Electrical and Industrial Systems, pages 215–223, Lisbon, Portugal, February, 2011.

[20] Pham Thanh Tung, Ly Quoc Ngoc. Elliptical Density Shape Model for Hand Gesture Recognition. The Fithth International Sysmposium on Information and Communication Technology (SoICT 2014), Hanoi, December 4th -5th , pp.186-191, 2014.

[21] Quang D Tran, Ngoc Q Ly. Sparse Spatio-Temporal Representation of Joint Shape-Motion Cues for Human Action Recognition in Depth Sequences. IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF 2013), Ha noi, Vietnam, November 10th-13th, 2013 (Best Student Running-Up Paper Award), pp.253-258, 2013.

[22] R. Munoz-Salinas, R. Medina-Carnicer, F. Madrid-Cuevas, and A. Carmona-Poyato. Depth silhouettes for gesture recognition, Pattern Recognition Letters, vol. 29, no. 3, pp. 319–329, February, 2008.

[23] Sebastian Marcel, Oliver Bernier, Jean Emmanuel Viallet and Daniel Collobert. Hand Gesture Recognition using Input – Output Hidden Markov Models, Proc. of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 456 – 461, 2000.

[24] V. Pavlovic, R. Sharma, and T. Huang. Visual interpretation of hand gestures for human-computer interaction: a review, IEEE Transaction on Pattern Analysis and Machine Intelligence, 19(7), July, 1997.

[25] Vladimir Vezhnevets, Vassili Sazonov and Alla Andreeva. A Survey on Pixel-Based Skin Color Detection Techniques, In Proceedings of the GraphiCon, 2003.

[26] Y. Wu, T. Huang, Vision-based gesture recognition: a review, Gesture-based commun. in hum. comput. interact. 103–115, 1999.

[27] Y, Yao, C.-T. Li and Y. Hu. Hand Posture Recognition Using SURF with Adaptive Boosting, British Machine Vision Conference, Guildford, UK, 3-7 September, 2012.

[28] Z. Ren, J. Yuan, and Z. Zhang. Robust hand gesture recognition based on finger-earth mover's distance with commodity depth camera. In International Conference on ACM Multimedia, 2011.

Test Case Reduction Techniques - Survey

Marwah Alian
Princess Sumaya University for Technology
Amman, Jordan

Dima Suleiman
Princess Sumaya University for Technology
Amman, Jordan

Adnan Shaout
The University of Michigan - Dearborn
Michigan - Dearborn

***Abstract*—Regression testing is considered to be the most expensive phase in software testing. Therefore, regression testing reduction eliminates the redundant test cases in the regression testing suite and saves cost of this phase. In order to validate the correctness of the new version software project that resulted from maintenance phase, Regression testing reruns the regression testing suite to ensure that the new version. Several techniques are used to deal with the problem of regression testing reduction. This research is going to classify these techniques regression testing reduction problem.**

Keywords*—*Regression testing; Test case reduction; Test Suite

I. Introduction

Regression testing is done to ensure the validity of modified software which is an essential activity related to making maintenance. The goal of this activity is to ensure that bug fixes and new functionality do not harm the correct functionality inherited from original program [1]. The size of the test-suite grows when new test cases are added to the test suite which increases the cost of regression testing [2]. Moreover, regression testing requires running a large program on a large number of test cases which means it can be expensive in both human and machine time [1].

Reducing cost is the target of researchers on the use of test-suite reduction techniques. Therefore, a number of different methods have been studied to deal with test suits such as minimization, selection and prioritization. Test suite minimization or reduction aims to reduce the number of tests to run. [3]

However, the main objective of most of proposed algorithms is to reduce the test suite size. In [4] they use the concept of Set theory to minimize the larger test suite. The set is defined as a collection of objects and entities. One set for each object type is created, and set operations such as Cartesian product, union, intersection and difference are used to reduce the size of test suite. This paper depends on the power of set functions.

While in [5] they combine the concept of software testing and Case-Based Reasoning (CBR) by assuming that test cases are treated as cases in CBR and they discuss how to maintain a number of test cases in software testing using the Case-Based Maintenance (CBM) if there is a set of test cases generated by the Path-Oriented technique. They propose a number of maintenance techniques that are used for removing unnecessary test cases and for controlling the growth of test cases. The process of maintaining CBR is called Case Based Maintenance (CBM) which is a policy of adding, updating, and deleting cases. Nicha et al. concentrate on the Deletion Policy for CBM. Their experiments show that the proposed technique can be used to reduce the number of test cases required for software testing in an efficient way.

Regression test suite (RTS) can be reduced using model-based regression which based on an Extended Finite State Machine (EFSM), for each elementary modification (EM) data and control dependences are used to capture interaction between EFSM transitions and three interaction patterns. Reduced RTS only includes test cases that one of their 3 interactive patterns is not produced for any other test case [6]. Requirements changes lead to modifications in the EFSM, Modification in model result in performing three types of model based regression testing: testing the model affection in modification, testing the modification affection on the model and testing the effect of modifications on unmodified part of the model. An extension of EFSM is the system under test (SUT) that support variable ranges over several data types and operations. In each transition level a changes occurs in EFSM. Changes can add or delete transitions [7].

The selection of most appropriate regression method is not easy since everyone has its advantages and disadvantages. Adaptive regression testing (ART) can be used to determine the most effective one which take into consideration the organization situation and testing environment [8]. ART uses certain criteria to determine which regression testing to use such as cost and benefits in addition to the organization situation such an evaluation have a problem called multiple criteria decision making (MCDM) problem, one of MCDM methods is Analytical Hierarchy Process (AHP), the results of the experiment show that techniques selected by AHB is cost effective. Decision maker must define a hierarchy that contains a description of a problem to be solved in order to be able to use AHP. Hierarchy contains the goal and the factors that may be used by decision maker.

The purpose of this survey is to collect and consider papers that deal with one of regression testing techniques that are test suite reduction. Our intention is to provide a state-of-the-art view on this field. Many different approaches have been proposed in order to reduce the cost and time of regression testing.

The rest of this research is organized as follows: Section II gives a review about other surveys in the field. Section III defines the problem of test suite reduction. In section IV a detailed outlook on the classification of test suite reduction techniques and section V presents a summary for the test suite techniques and discussion.

II. Related Work

Detecting faults in the program is the objective of software testing and it provides more assurance on the quality of the software. As the software evolves the size of the test suite grows because of the addition of new test cases to the test suite [9]. It is important to develop techniques to minimize available test suites because of the time and resource constraints for re-executing large test suites. [10] Hundreds of techniques have been proposed to solve the problem of test suite reduction, and several surveys have been introduced for such techniques.

In software testing field a survey of Chaurasia et al. [11] is given which is a literature review of test case reduction techniques. They discuss and compare three algorithms presented in [12][13][14]. All papers are about test case reduction but Control Flow Graph is used in all the methods. These algorithms worked on the basic Basis Path Testing, which is the first method worked on the low level of the code, exceptions, conditions, and loops. They discuss the algorithm advantages, disadvantages, cost, and time. In the three algorithms, time, cost and the number of test cases are reduced significantly.

In [15] they present a survey of using Genetic Algorithm in different software testing techniques. They describe how GA works and the applications of GA in different types of software testing like test planning, minimization of test cases in regression testing, model based testing and web testing. They discuss one technique for each type of software testing except for model based testing they discuss four techniques. In regression testing they introduce the technique proposed by [16] and their comparison with vector based technique in test suite reduction. Also, they compare GA Parameters that are used in different types of software testing. GA parameters are Crossover Rate, Mutation Rate, and Number of Generation.

While in [17] an empirical study of five different regression test optimization techniques is presented. They give a brief description for the three regression testing optimization techniques that are selection, prioritization, and minimization. Then they describe five implemented algorithms that are slicing algorithm, incremental algorithm, genetic algorithm, adaptive firewall approach, and simulated annealing. Also, they introduce a comparison that is based on different qualitative and quantitative criteria such as execution time of tests, precision, number of tests selected, user parameters, global variables handling, and type of testing.

A review is made to compare between Heuristics, Genetic Algorithms and Linear programming based techniques. Heuristics produced smaller size reduced set it work very good on it but it must be optimized in large scale test suite, on the other hand Heuristic GRE produced optimal representative set. Genetic Algorithm concerns about the block based test suites on software and coverage, in Genetic algorithm with time constraints the test suite becomes smaller results in minimizing running time. Smallest set is produced in Integer Linear Programming compared with other algorithms. For significant test suite reductions two or more techniques can be combined to form hybrid techniques [18].

In [19] various test case generation methods are presented such as minimizing, selection, prioritization and evaluation test cases. Also many methods that help test engineers in scheduling and ranking test cases are discussed such as test cases prioritizations and selection techniques. Some of test case generation depends on application such as those generated for web, object oriented, UML, evolutionary applications and structured based systems. System Requirements and use cases can be used to derive test cases; also test cases can be generated using UML sequence diagram, these test cases are generated for object oriented programs. In order to be able to find the sub optimal test cases we can use genetic algorithms to let test cases meet certain criteria such as if the test cases are adequate to statement, branch, and path coverage. Code based test generation generates test cases from a code, four types of inputs are taken: program to be tested, -and some information related to runtime. Dynamic path testing and evolutionary technique generate test cases by using many test values in program execution; test cases can be prioritized according to the shortest path. Test case also can be generated by traversing the graph from parent root to child node in Graph Traversal Algorithm, by using breadth or depth first in a graph tree.

There are different representations of graph models proposed for procedural programs which are discussed in [20] survey such as Control Flow Graph (CFG), Program Dependence Graph (PDG) and System Dependence Graph (SDG). Flow graph is a directed graph a set of nodes represents program statements, there are two nodes called start and stop nodes and there exists path from a start to every other node in flow diagram and from there to stop node. Data Dependence Graph is used to represents relationships between elements of a program. The relationship can be either data or control dependency.

There are different representations of graph models proposed for procedural programs that are discussed in [20] survey such as Control Flow Graph (CFG), System Dependence Graph (SDG) and Program Dependence Graph (PDG) . Flow graph is a directed graph in which a set of nodes represents program statements, there are two nodes called start and stop nodes and there exists path from a start to every other node in flow diagram and from there to stop node. Data Dependence Graph is used to represents relationships between elements of a program. The relationship can be either data or control dependency.

However, other surveys has a narrow view for the field and many types of techniques about test suit reduction requires classification into specific groups or classes represents similarities and common properties. In this survey we classify test case reduction techniques into: requirement based, coverage based, program slicing, genetic algorithm, greedy algorithm, hybrid algorithm, clustering and fuzzy logic.

III. Test Suite Reduction Problem

According to the definition given by [21], test suite reduction problem can be can be defined as:

Given a test suite T represents a set of test cases {t1, t2, t3,…, tn}, a set of test requirements R = {R1, R2, . . . , Rn} to be covered, and subsets of T, S= {S1, S2, . . . , Sn}, where each test set is associated with Ri. The objective is to find the representative subset RS ⊆ S that satisfies all of requirements and has at least one test case for each requirement R.

IV. Test Suite Reduction Techniques

Regression testing is defined as a software maintenance activity which is done to ensure the proper functionality of the software. Test suits that are developed during the development phase have a large size that it is not possible to run the entire test suite due to the time and cost. Therefore, regression test reduction process is advisable in order to reduce the test suite to minimal set of test cases that will cover all the faults in minimal time [22]. In this survey we consider the proposed test suite reduction techniques and their classification.

As shown in Fig. 1 test case reduction techniques are classified into: requirement based, coverage based, slicing, genetic algorithm, greedy algorithm, hybrid algorithm, clustering and fuzzy logic.

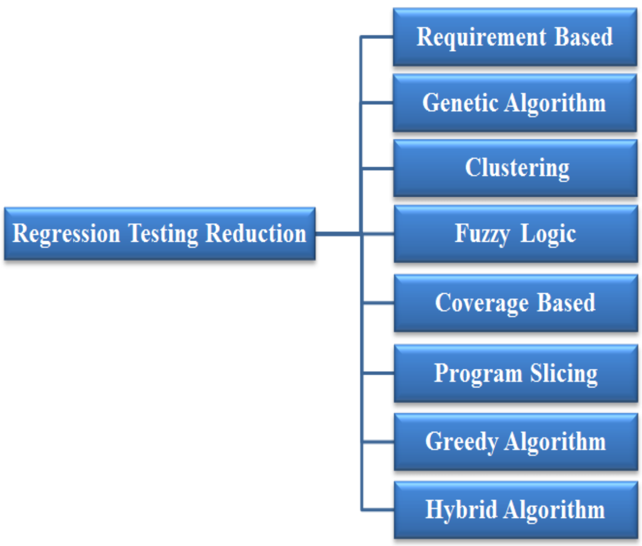


Fig. 1. Regression testing reduction techniques

A. Genetic Algorithms

There are many issues of software testing like effective generation of test cases, test case reduction, prioritization of test cases, etc. These issues demand on effort, time and cost of testing. The use of evolutionary algorithms for automatic test case generation and reduction has been an interest for researchers [15]. One such form of evolutionary algorithms is Genetic Algorithm (GA) that is computational Intelligence based approach used as a solution for the problem of test cases reduction like evolutionary computation.

In [2] they propose a genetic algorithm, for test-suite reduction which builds the initial population based on test history, it calculates the fitness value using coverage and cost information then using genetic operations it selects the successive generations. These steps will be repeated until a minimized test-suite set is found. The results show that the proposed test-suite reduction technique has cost-effectiveness and generality.

Also, the research of [23] investigates the use of genetic algorithms, for test-suite reduction. They propose a model that builds the initial population based on test history, but it calculates the fitness value depending on coverage and run time of test cases, the fit tests only allowed to be in the reduced suite. This generational process is repeated until an optimized test-suite is found.

In [16] they define the time-aware regression testing reduction problem and propose a genetic algorithm for this problem. This work describes parent selection, crossover and mutation processes of the genetic algorithm. The redundancy of test cases is removed from regression testing suite by the proposed algorithm they also minimize the total running time of the remaining test cases.

While in [24] a new approach for test case reduction is presented and implemented. This approach depends on genetic algorithm technique with varying chromosome length in order to reduce test cases in a test suite by finding a representative set of test cases that fulfill the testing criteria.

Approaches that based on Genetic algorithms need to further study the fault detection capability of block based test suite or coverage based test suite or other criteria [25].

B. Requirement Based

The purpose of test suite reduction is to satisfy all testing requirements with a minimum number of test cases. Before generating test cases it is necessary and possible to optimize testing requirements.

In [26] they solve test suite reduction by testing requirement optimization. They proposed a requirement relation graph which is proposed to minimize the requirement set by graph contraction. The experiment is designed and implemented using specification-based testing. In this work, testing requirement relation graph is introduced in order to hide the details of testing requirements and test cases and to achieve test suite reduction. Some requirement contraction methods are proposed to generate a small requirement set. Also, an empirical study on specification-based testing is performed. The study compares the relative effectiveness of testing requirement optimization to test case reduction. In order to compare with test suite reduction, all test cases of each testing requirement are generated then the greedy algorithm is applied on the constructed test suites for reduction.

The ability to detect faults is reduced as a result of reducing the size of test suite. The redundancy in test suites and size can be reduced using a model checker based techniques to create test cases [27]. The main idea of this approach is that not all test cases in test suite are important to achieve requirements, so that a subset of test cases can be chosen that fulfill the requirements. Model-checker take a finite state model and temporal logic property as input and as a result counter example will be returned if the property is not fulfilled. Counter example is a sequence of states from the initial state to the violating one. Since removing of test cases from test suite have negative impact on the ability to detect faults, this can be avoided by transforming the test cases so the redundant parts can be omitted. Common prefix of the test cases is specified then the only part after the prefix of test cases is interesting. And if there is another test case that ends with the same prefix we talked about previously then we can append the remainder of the first test case to the second test case and omit the first one. In this research the quality of test suite after removing the redundant test cases does not suffer, the experiments show that the reduction is significant and one drawback is the run time complexity.

The basic classic needs of test suite reduction is to reduce the time and cost and maintain the effectiveness of fault localization, so that a technique that keep the system free from errors is needed and at the same time keeping the efficiency by reducing the number of test cases[28].

In [28] they study three techniques and the best out of them is determined, the first one is test suite reduction and requirement optimization, if there is a test cases in input domain that satisfy testing requirements then the testing requirements is feasible. The second one is dynamic domain reduction (DDR); these algorithms based on effective testing in detection of errors, there are many testing methods such as: path testing, the purpose of this algorithm is not to test every path. Another algorithm is independent program path where at least one new processing path is considered. DDR is based on constraint based testing. The third algorithm is Ping Pong technique which uses heuristic technique to reorder the test cases that provides a good solution but not the optimal one.

The requirement optimization is good when dealing with the finite Boolean expression that classifies the requirements as true of false test cases. When dealing with arrays, loops and expression DDR succeeds. Ping Pong assures the domain coverage but it is time consuming technique, expensive and more memory is needed. It takes requirements from natural language.

Some algorithms use the decision table concepts that depend on the requirement gathered from the user to create test cases [29]. A framework of this algorithm consists of three steps: functional requirements analysis and condition determination, input generation for rule development and testing, in this step decision table is generated and with corresponding rules and test cases ,and the last step is rule development and testing using opened rules, in this step the test output will be generated. The redundancy reduction of test case is up to 30% which save cost and time; the proposed algorithm does not require the tester to have knowledge of coding.

C. Fuzzy logic

Optimization of test suites can be achieved by using fuzzy logic. It is a safe technique and can reduce the regression testing size and execution time [30]. Fuzzy logic can be used in many areas such as communication, bio informatics and experts systems. Level of testing using fuzzy logic is based on objective function quite similar to human judgment.

A combination of fuzzy logic with genetic algorithm and swarm optimization can be used to make optimization in test suite for multi-objective selection criteria. Therefore, in [31] they aim to find a test suite that is optimal for multi-objective regression testing. They propose an expert system that use a technique and level of testing based on a defined objective function, similar to human judgment using fuzzy logic based classification.

While in 2013, [32][33] use some CI based approaches in order to optimize the test suite and analyze the test suite for safe reduction which can be estimated using control flow graphs. Test cases of optimal solutions are traversed on these graphs and it is found that only fuzzy logic is safe while other approaches will be inadequate for regression testing.

D. Clustering

When designing test cases, there are many redundant test cases with no use. Such redundancy increases the testing effort and increase the cost and time of testing.

In order to reduce the time spent in testing, the number of test cases is reduced. In [34] they use the data mining approach of clustering technique in software testing to reduce the test suite.

With the help of the clustering techniques the number of test cases is reduced and the efficiency of software testing is improved. By using clustering the program can be checked with any one of the clustered test cases rather than with the entire test case that is produced by the independent paths.

A new approach is discussed by [35] which divide the test cases into clusters according to the similarity in profiling. Previous researches made partitions on execution profile by representing it as a binary vector which contains only number of times the function executed without taking into considerations the sequence of execution, so that this paper provide enhancements by making three types of profiles: file execution sequence, function call sequence and function call tree. The results show that the relation between function calls and sequential information will make enhancements in detecting the faults.

Clustering techniques based on selecting test cases of using coverage and distribution based techniques. They produce smaller representative sets of test cases but less fault detection ability [25].

E. Coverage Based

An important issue to be considered during test suite reduction is the coverage aspect. Coverage-based reduction techniques have to ensure that majority of the execution paths of the given program are exercised [9].

Regression testing must preserve the fault detection in addition to minimizing the size and the time of test. In Case Base Reasoning (CBR) there are three cases classifications: case, auxiliary and pivotal case. Case based is Artificial intelligent which search for the most similar problems to solve problems; in this case a memory is needed. Auxiliary case can be deleted without affecting competence, but only reduces the efficiency of the system. On the other hand pivotal cases have a direct effect on the competence of a system if it is deleted. As we said CBR uses path coverage criteria to reduce the redundancy test cases. Path coverage uses a control flow graph which can derived from a source code, path oriented test case generation technique used to derive a test cases from a control flow graph where each state can reveal a fault.

In [36] they use the case based reasoning (CBR) which is one of artificial intelligent concepts. This study introduces three methods. CBR and software testing have the same problems that there are many redundant test cases after reduction, the ability of reducing the faults decreases, and grow of test cases is uncontrollable. The key research issues of CBR are: continuous growing in CBR size is the existence of too many redundancy cases and the deletion all redundancy in CBR take a lot of time. In order to control number of cases there are many algorithms such as add, delete and maintenance approaches, deletion algorithms are the efficient algorithms to maintain the size of CBR system. [36] Treat the test cases as cases in CBR with the assumption that path oriented test cases method used to generate a set of test cases. There are two classifications of test cases reduction techniques proposed in this research, the first one is coverage based technique and the second one is concept analysis based technique. Their research propose three methods using CBR: The first method is Test Case Complexity for Filtering (TCCF), the proposed method that applies CBR concepts can be described in the following steps: Determine the coverage set, reachability set, auxiliary set from which we can compute the complexity for each test case and the last step is to remove test cases from the auxiliary set that have a minimum complexity value, the complexity of test case can be computed as follows: High when number of test cases are larger than the average number of test cases in test suites, when they are equals then the complexity is medium and when number of test case smaller than average it is considered as low.

The second algorithm is Test Case Impact for Filtering (TCIF) in this method the impact value is the impact of test cases according to ability to detect faults when test cases removed. The impact of test case is high when at least one fault of many times has been revealed by test cases, medium when the faults revealed for only one time and low when test case has never revealed the faults. The first three steps in this algorithm are the same as the ones in (TCCF) but the last step is different. The last step in TCIF is to remove test cases with minimum of impact value from the auxiliary.

The third algorithm is Path Coverage for Filtering (PCF) method; path testing is a structure testing since you choose test cases that determine the path to be taken within the program structure. This algorithm uses a coverage value which determines how many nodes that test case can cover. The procedure of this method has the following steps: identify the coverage set, calculate the coverage value which depends on number of test cases in each group and the last step is to remove all test cases with minimum coverage value.

The experiments randomly generate 2000 test data used in telecommunication industry. There are list of measures used in this experiment such as: number of test cases, reveals faults ability, and total reduction time. The results show that in PCF the number of test cases is minimized more than other algorithms and it consumes the least reduction time, but it is the worst in ability to reveal faults. While TCCF and TCIF are the best in term of faults detections, they are the worst that they require a lot of time in reduction process.

The main problem related to using these algorithms is that for huge systems the path coverage is ineffective since it needs a time and consumes a cost in identifying the coverage from a source code.

A technique called coverall algorithm is proposed in [37]. The test cases are generated automatically and the tester has no options to do that. This method uses algebraic conditions to give a value to variable, this variable value resulting in fewer numbers of test cases. Also this method can be used in loops and arrays, the reduction rate of test cases will reach 99%. In addition to creating test cases automatically, this method also minimizes the number of test runs.

To generate test cases automatically there are four steps: finding all constraints from start to finish node, identifying the variables with minimum and maximum values, test path to find constant values and creating a table of all possible values from the above values. As a result of this algorithm the number of generated test cases is smaller than many other algorithms such as Ping-pong, it also keeps the test cases generation to a single run. This technique is the best technique among many other techniques such as GetSplit algorithm, Ping-pong technique and Dynamic Domain Reduction (DDR) in term of reduction of test cases and other factors discussed before, however if there are more than two variables in program code the method is not applicable. This algorithm can be improved if the description of initial domain input is no longer required [37].

It is possible to reduce test suites without affecting the coverage of the states. In [38] they propose an algorithm that covers all reachable states in closed loop controller. The approach focuses on path coverage since it generate test cases from accessing the source code, a path is a sequence of statements from the beginning of program execution to its end while the sub path is a sequence from the beginning of a specified function execution to its end. In this approach test cases are generated depends on intuition that is, as long as it covers all sub-paths it is not necessary to cover all paths. There are two steps of this approach: Identify the test cases that cover all sub paths in program based on implementation of code and remove any test case that covers an already covered sub paths. For experiments purposes this approach applied to five controller programs for real world medical protocols, these programs are: PennNeuroICU, PennCardiac, PennMICU, PennHyper-Glycemia, and PennIntraoperative Experiments show that test case reduced by tens of thousands with no reduction in fault finding capability.

Reference [39] proposed a technique called TestFilter used for reducing test cases, it uses statement coverage. TestFilter choose non-redundant test case according to their weight, this process reduce the test cases storage management and execution cost.

F. Program Slicing

Program slicing is introduced by [40]. This technique is used to check a program over a specific property and to build a slice set, which is a set of statements effect to determine a statement; in many cases it is the output statement of a program, based on input values. Slicing techniques can help to show control-flow of a program for each test case and it is important to specify which statements are invoked with that test case. [55]

There are three types of slicing techniques: static slicing [40][41][42], dynamic slicing [43][44], and relevant slicing [45]

In [1] a survey is made for seven approaches that use program slicing for reducing the cost of regression testing. There are three groups of these approaches; the first group includes those using dynamic slicing, while the second group includes approaches that use program dependence graph, and the third one includes those using data flow definition of slicing.

To reduce the cost of regression testing, the study of [46] uses two algorithms: the first one generates a program called differences, this algorithm called like this since it captures the difference between certified and modified program, where certified is the previously tested program without changes and modified is the program with modification. The second algorithm uses existing test cases to test components new in modified, also it uses the test cases for which modified and certified program produced the same outputs. The idea is to avoid the cost of using new test cases and to avoid rerunning test cases that produce the same output. The second algorithm uses a context slice which is one of new type of inter procedural slice. Inter procedural slice contains a program that includes components that capture all execution statements. The slice can be defined as follows: if we have a program component *p* and variable *x* then slice includes all statements in a program that causes effect on variable *x*.

Using slicing techniques can decrease the number of required test cases and consequently the cost and time of testing will be decreased.

However, in [47] they propose a method that is intended to reduce the cost and time of testing and they investigate the effect of slicing techniques on the reduction rate of testing cost and time. This method focuses on parts of the program code that have significant impact on its output while those parts of program that have no effect on the program output are eliminated from testing process. Hence, as the size of the code is reduced, testing time and cost will be decreased. Their experiments show that a large number of program instructions, branches and paths can be covered by a small number of test cases in the sliced program.

G. Greedy Algorithm

One of well-known heuristics proposed for code-based reduction is Greedy algorithm. This algorithm can also be applied on test suites obtained from Model-based techniques. It selects the test case which satisfies the maximum number of unsatisfied requirements and an arbitrary choice is made if there is a tie situation. This process is applied repeatedly to all test cases in the test suite and produces a reduced test suite. It is stopped after all test requirements are satisfied [48][49].

In [10] inspired greedy algorithm for test suite reduction is proposed. This algorithm is based on formal concept analysis of the relation between testing requirements and test cases. This analysis is used for objects that have discrete properties. They consider test cases as objects and requirements as their attributes. Using concept analysis framework, maximum grouping of objects and attributes are identified and called contexts. Reduction rules are used for reducing objects and attributes. This greedy algorithm differs than classical greedy heuristic which uses object implications without considering attribute implication. In their algorithm, context table is constructed initially then the size of context table is reduced by applying the objects reduction rules, attribute reduction and owner reduction. The reduction of objects and attributes will reduce the size of the context table by removing redundant attributes and objects from further consideration. While the owner reduction not only removes redundant attributes and objects but also it selects a test case which will be added to the reduced suite and the requirements covered by these test cases are not considered in further iterations. In each iteration, interference among test cases is also removed using greedy heuristic. The size of the reduced test case suite generated by their algorithm has the same or smaller size than that generated by the traditional heuristic algorithms.

Also, Weighted greedy algorithm is proposed by [50] for test suite reduction also called Weighted Set Covering Technique. This work starts by determining test cases which can satisfy all the requirements. If the test case does not satisfy requirements then the algorithm repeatedly eliminate redundant test cases then update the test suite and the remaining requirements that are uncovered. The essential test cases that are selected are added to the reduced set. In order to handle the remaining uncovered requirements, prioritization and sorting take place for test cases. Then, a selection for test cases depends on decreasing order of priority is repeated until all the requirements are satisfied. The optimized test suite had a higher efficiency. Their experiment is made on the test suite of Student Achievement Retrieval Navigation Model. The algorithm reduces the size of the test suites and minimizes test cost.

Real world java programs are used to implement four test suite reduction techniques for JUnit test suites [51]. The study of [51] cares about benefits and the cost of test suite reduction. The results show that JUnit suite is reduced without affecting the fault detection capability. The first technique involved is the greedy technique where a test case that satisfies maximum number of unsatisfied requirements is selected. The second technique is Harrold et al.'s Heuristic [21]; the main idea of this technique is to select test cases according to their essential. The next technique is GRE Heuristic brings together characteristics of essential test cases and 1 to 1 redundant test cases into greedy strategy, the heuristic terminates when all requirements are satisfied. The last algorithms is ILP Technique, the first ILP is single objective while the second ILP is multiple objective test suit reduction. The purpose of first ILP model is to minimize then number of test suite selected [51].

Coverage Based Test Suite Reduction (CBTSR) is a new algorithm for Test Suite Reduction that is proposed by [9]. They identify an optimal representative test set for test cases that are related to the given testing objective. Then they Apply data flow testing to generate test cases as well as requirements in order to examine the physical structure of the program and to locate sub-paths. After that they use the proposed CBTSR algorithm for test suite reduction. Also, they perform a set of empirical studies on ten subject programs.

The reduction process in CBTSR algorithm starts with the construction of test case requirement matrix which maps the testing requirements with test cases. An association between a test case and requirement is represented by one or zero otherwise. Each row in the matrix denotes the requirement coverage and each column denotes the test case which overlaps with the requirement. Then the generation of the reduced test suite is made through simple mathematical operations. The results show that CBTRS algorithm selects near optimal test cases which satisfy a maximum number of testing requirements. Thus, it reduces the size of the test suite by retaining test cases that offer maximum percentage of Requirement Coverage.

Test suite reduction approaches that based on Greedy algorithm provide significant reduction in test suite but need to be optimized in large scale test suites [25].

H. Hybrid Algorithm

Some algorithms try to reduce the number of test cases using hybrid techniques such as genetic algorithms and bee colony [52]. Bee colony consists of three groups of bees: employed, onlookers and scouts. Using bees as agents the algorithm can explore the minimum set of test cases. This paper suggests using ant colony with genetic algorithms.

Moreover, in [53] they formulate three hybrid combinations, Rank, Merge, and Choice, and describe their usefulness. They produce a uniform representation for hybrid criteria and suggest that the hybrid criteria of others can be described using Merge and Rank formulations, and that the hybrid criteria outperform the constituent individual criteria.

While in [54] they propose multi-objective test suite reduction. They introduce a hybrid multi-objective genetic algorithm.

Their algorithm combine the efficient approximation of the genetic algorithm with the greedy approach to produce high quality Pareto fronts in order to achieve multiple objectives. Objective functions are considered as a mathematical description of test criterion. A cost cognizant version of the greedy algorithm is implemented for two objective optimization that are computational cost and statement coverage.

Three optimization objectives are also considered for fault detection history such as code coverage, fault coverage and execution time. These objectives are combined using the classical weighted-sum approach by taking the weighted sum of fault coverage per unit time and code coverage per unit time. The testing decisions that are taken by their technique have been more efficient.

V. SUMMARY

Regression testing is made when changes are performed to existing software in order to provide confidence that the new changes that are introduced do not affect the behavior of the existing, unchanged part of the software. As software evolves, the test suite tends to grow which implies that it may be expensive to execute the entire test suite [3].

Hundreds of techniques have been proposed in order to reduce the number of test cases, but there are still many researches in this field. In this research, a review for what have been proposed by researchers to solve the problem of test case reduction is presented and a classification for test case reduction is introduced. The reviewed test case reduction techniques are classified into: requirement based, coverage based, slicing, genetic algorithm, greedy algorithm, hybrid algorithm, clustering and fuzzy logic. More details about these techniques and their classification are demonstrated in Table II.

However, Greedy algorithm based techniques provide significant reduction in the number of test cases, but they need to be optimized in large scale test suites. While genetic algorithm based techniques need to be examined on the fault detection capability and other criteria.

While hybrid techniques are introduced for significant reduction in test case suites, they provide high complexity. Clustering techniques select test cases based on coverage and distribution techniques, they produce smaller representative sets but less fault detection ability [25].

The main problem related to coverage based techniques is that for large systems the path coverage is ineffective since it consumes time and cost in identifying the coverage from a source code [36]. The advantages and disadvantages of the proposed classifications are given in Table I. Many techniques can be incorporated with existing hybrid techniques and with genetic algorithms more mutation strategies can be introduced.

TABLE I. CLASSIFICATTION OF TEST SUITE REDUCTION- ADVANTAGES AND DISADVANTAGES

Classification	Advantages	disadvantages
Program slicing	decrease the number of required test cases and consequently the cost and time of testing will be decreased.	need to be examined on the fault detection capability and larger generated data
Genetic algorithm	Reduce the number of test cases and also decreases total running time.	need to be examined on the fault detection capability and other criteria
Greedy algorithm	provide significant reduction in the number of test cases	involve random selection of test case in a tie situation.

		need to be optimized in large scale test suites
Fuzzy logic	a safe technique and reduce the regression testing size and execution time	Need more experiments and studies
Requirement base	Provide a good percentage of redundancy reduction of test cases.	Some of them are time consuming and need more memory depending on how to represent the requirements
Coverage Base	reduction rate of test cases is very high and it reduce time	for large systems the path coverage is ineffective since it consumes time and cost in identifying the coverage from a source code
Hybrid algorithm	provide significant reduction in the number of test cases and multi-objective optimization	high complexity
Clustering	produce smaller representative sets of test cases	less fault detection ability

TABLE II. TEST SUITE REDUCTION TECHNIQUES

Year	List of Authors	Classification	Paper Title	Technique	Enhancement
1995	David Binkley	Program Slicing	Semantics Guided Regression Test Cost Reduction,	Two algorithms are used: the first one generates a program called differences, this algorithm called like this since it captures the difference between certified and modified program, were certified is the previously tested program without changes and modified is the program with modification. The second algorithm uses a context slice which is one of new type of inter procedural slice.	decrease the number of required test cases
2005	Xue-ying Ma, Bin-kui Sheng, and Cheng-qing Ye,	Genetic Algorithm	Test-Suite Reduction Using Genetic Algorithm	Builds the initial population based on test history, it calculates the fitness value using coverage and cost information.	it has cost-effectiveness and generality
2005	Sriraman Tallam, Neelam Gupta	Greedy Algorithm	A Concept Analysis Inspired Greedy Algorithm for Test Suite minimization	Inspired greedy algorithm for test suite reduction. Test cases are considered as objects and requirements as their attributes. Using concept analysis framework, maximum grouping of objects and attributes are identified and called contexts. Reduction rules are used for reducing the size of context table by applying the objects reduction rules, attribute reduction and owner reduction.	The size of the reduced test case suite has the same or smaller size than that generated by the traditional heuristic algorithms
2006	Preeyavis Pringsulaka and Jirapun Daengdej,	Coverage Based	Coverall Algorithm for Test Case Reduction	Coverall algorithm uses algebraic conditions to give a value to variable, this variable value resulting in fewer numbers of test cases. This method can be used in loops and arrays.	The reduction rate of test cases reach 99%.
2006	Saif-ur-Rehman Khan, Aamer Nadeem,	Coverage Based	TestFilter: A Statement-Coverage Based Test Case Reduction Technique	TestFilter uses statement coverage by choosing non-redundant test cases according to their weight.	This process reduce the test cases storage management and execution cost.
2006	B. Suri, I. Mangal, and V. Srivastava,	Hybrid Algorithm	Regression Test Suite Reduction using an Hybrid Technique Based on BCO And Genetic Algorithm	Combine genetic algorithms and bee colony. Bee colony consists of three groups of bees: employed, onlookers and scouts. Using bees as agents the algorithm can explore the minimum set of test cases.	Explore the minimum set of test cases.

2007	Gordon Fraser and Franz Wotawa ,	Requirement Based	Redundancy Based Test-Suite Reduction	Subsets of test cases that fulfill the requirements are chosen. Model-checker take a finite state model and temporal logic property as input and as a result counter example will be returned if the property is no not fulfilled.	The reduction is significant but one drawback is the run time complexity
2008	Zhenyu chen, baowen xu, xiaofang zhang, changhai nie.	Requirement Based	A novel approach for test suite reduction based on Requirement relation contraction.	a requirement relation graph is proposed to minimize the requirement set by graph contraction. An empirical study on specification-based testing is performed.	The study compares the relative effectiveness of testing requirement optimization to test case reduction
2010	Shin Yoo , Mark Harman.	Hybrid Algorithm	Using hybrid algorithm for Pareto efficient multi-objective test suite minimization	introduce a hybrid multiobjective genetic algorithm. Their algorithm combine the efficient approximation of the greedy approach with the genetic algorithm to produce high quality Pareto fronts in order th achieve multiple objectives.	The testing decisions that are taken by their technique have been more efficient.
2010	Siripong Roongruangsuwan and Jirapun Daengdej,	Coverage Based	Test Case Reduction Methods by Using CBR,	use an artificial intelligent concept of case based reasoning (CBR). propose three methods using CBR: Test Case Complexity for Filtering (TCCF), Test Case Impact for Filtering (TCIF) and Path Coverage for Filtering (PCF) Method	In PCF the number of test cases is minimized more than other algorithms and it consumes the least reduction time
2010	S. Nachiyappan, A. Vimaladevi and C.B. SelvaLakshmi,	Genetic Algorithm	An Evolutionary Algorithm for Regression Test Suite Reduction	initial population is built based on test history, but it calculates the fitness value depending on coverage and run time of test cases,	Reduce test case suite size.
2011	Lingming Zhang, Darko Marinov, Lu Zhang, Sarfraz Khurshid,	Greedy ALgorithm	An Empirical Study of JUnit Test-Suite Reduction	The study cares about benefits and the cost of test suite reduction by testing four techniques on JUnit test suites; greedy technique, Harrold Heuristic, GRE Heuristic and ILP.	The results show that JUnit suite reduced with affection the fault detection capability.
2012	Shengwei Xu, Huaikou Miao, Honghao Gao,	Greedy Algorithm	Test Suite Reduction Using Weighted Set Covering Techniques.	Weighted greedy algorithm is used for test suite reduction also called Weighted Set Covering Technique. It starts by determining test cases which can satisfy all the requirements. If the test case does not satisfy requirements then the algorithm repeatedly eliminate redundant test cases then update the test suite and the remaining requirements that are uncovered. The experiment is made on a test suite of Student Achievement Retrieval Navigation Model.	The optimized test suite had a higher efficiency. The algorithm reduces the size of the test suites and minimizes test cost.
2012	Liang You, Yansheng Lu	Genetic Algorithm	A genetic algorithm for the time-aware regression testing reduction problem	removes redundant test cases in the regression testing suite minimizes running time of the remaining test cases	Reduce the size of test suite and running time.
2012	Haider, A.A.; Rafiq, S.; Nadeem	Fuzzy logic	Test suite optimization using fuzzy logic	an expert system that use a technique and level of testing based on a defined objective function, similar to human judgment using fuzzy logic based classification	Optimize test suite.
2013	Christian Murphy, Zoher Zoomkawalla, Koichiro Narita,	Coverage Based	Automatic Test Case Generation and Test Suite Reduction for Closed-Loop Controller Software	An algorithm that covers all reachable states in closed loop controller. It focuses on path coverage since it generates test cases from accessing the source code. Two steps of this approach: Identify the test cases that cover all sub paths in program based on implementation of code and remove any test case that covers already covered sub paths. For experiments purposes this approach applied to five controller programs for real world medical protocols.	The number of test cases is reduced by tens of thousands with no reduction in fault finding capability.
2013	A. Ali Haider, A.	Fuzzy logic	Computational Intelligence	Use CI based approach and analyses the	fuzzy logic is a

	Nadeem, S. Rafiq,		and Safe Reduction of Test Suite	test suite for safe reduction which can be estimated using control flow graphs. Test cases of optimal solutions are traversed on these graphs and it is found that only fuzzy logic is safe while other approaches will be inadequate for regression testing	safe approach and it is adequate for regression testing
2013	Sampath, S.; Bryce, R.; Memon, A.M.	Hybrid Technique	A Uniform Representation of Hybrid Criteria for Regression Testing	Three hybrid combinations are formulated, Rank, Merge, and Choice, and describe their usefulness. They produce a uniform representation for hybrid criteria and suggest that hybrid criteria of others can be described using Merge and Rank formulations, and that the hybrid criteria outperform the constituent individual criteria.	Use Rank, Merge, and Choice operations to perform hybrid criteria that outperform the constituent individual criteria.
2013	Asghar Mohammadian , Bahman Arasteh.	Program Slicing	Using Program Slicing Technique to Reduce the Cost of Software Testing.	A method that focuses on parts of the program code that have significant impact on its output while those parts of program that have no effect on the program output are eliminated from testing process. It shows that a large number of program instructions, branches and paths can be covered by a small number of test cases in the sliced program.	As code size reduced, testing time and cost are decreased.
2014	Gupta, A. ; MNNIT Allahabad, Allahabad, India ; Mishra, N. ; Kushwaha, D.S	Requirement Based	Rule Based Test Case Reduction Technique Using Decision Table	A framework consists of three steps: functional requirements, analysis and condition determination. the proposed algorithm does not require the tester to have knowledge of coding.	The redundancy reduction of test case is up to 30% which save cost and time;
2014	B.subashini , d.jeyamala..	Clustering	Reduction of test cases using clustering Technique	Use data mining approach of clustering technique to reduce the test suite. By using clustering the program can be checked with any one of the clustered test cases rather than with the entire test case that is produced by the independent paths.	the number of test cases is reduced and the efficiency of software testing is improved.
2015	Sudhir Kumar Mohapatra, Srinivas Prasad	Genetic Algorithm	Finding Representative Test Case for Test Case Reduction in Regression Testing	Reduce test cases in a test suite by finding a representative set of test cases that fulfill the testing criteria.	Reduce the number of test cases.
2015	R. Wang, B. Qu, Y. Lu,	Clustering	Empirical study of the effects of different profiles on regression test case reduction	divide the test cases into clusters according to the similarity in profiling. provide enhancements by making three types of profiles: file execution sequence, function call sequence and function call tree.	The relation between function calls and sequential information will improve detecting the faults.
2015	Preethi Harris and Nedunchezhian Raju.	Greedy Algorithm	A Greedy Approach for Coverage-Based Test Suite Reduction.	The reduction process starts with the construction of test case requirement matrix which maps the test cases with the testing requirements. An association between a test case and requirement is represented by one or zero otherwise. Then the generation of the reduced test suite is made through simple mathematical operations.	It selects near optimal test cases that satisfy the maximum percentage of Requirement Coverage. and it reduce the test suite size

REFERENCES

[1] D. Binkley, The application of program slicing to regression testing, Information and Software Technology, 40(11-12), pp. 583-594.

[2] X. Ma, B. Sheng, and C. Ye, Test-Suite reduction using genetic algorithm,Vol. 3756 of the series Lecture Notes in Computer Science, 2005, pp. 253-262, springer.

[3] S. Yoo, and M. Harman, Regression testing minimization, selection and prioritization: a survey, Software Testing, Verification and Reliability. 22(2012), pp.67–120, DOI: 10.1002/stvr.430

[4] I. Mangal, D. Bajaj and P. Gupta, Regression Test Suite Minimization using Set Theory, International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4, No. 5, 2014, pp. 502-506

[5] N. Kosindrdecha, S. Roongruangsuwan and J. Daengdej, Reducing test cases created by path oriented test case generation, American Institute of Aeronautics and Astronautics, Inc. AIAA Infotech@Aerospace 2007 Conference and Exhibit, California, USA.

[6] Y. Chen, R. Probert, H. Ural, Regression test suite reduction using extended dependence analysis, Proceedings of the 4th international workshop on Software quality assurance, in conjunction with the 6th ESEC/FSE, ACM Press, 2007, pp.62-69.

[7] B. Guo, M. Subramaniam and H. Guo, An approach to regression test selection of adaptive EFSM Tests, Fifth IEEE International Conference on Theoretical Aspects of Software Engineering.2011. pp. 217 – 220.

[8] Md. Arafeen and H. Do., Adaptive regression testing strategy: An empirical study, 22nd IEEE International Symposium on Software Reliability Engineering, 2011.

[9] P. Harris and N. Raju, A greedy approach for coverage-based Test Suite reduction, The International Arab Journal of Information Technology, Vol. 12, No.1, 2015 pp. 17-23.

[10] S. Tallam, N. Gupta, A concept analysis inspired greedy algorithm for test suite minimization, 2005 ACM 1595932399/05/0009

[11] V. Chaurasia, Y. Chauhan and T. K, A survey on test case reduction techniques, International Journal of Science and Research (IJSR), 2012

[12] J. Offutt, Z. Jin and J. Pan, “The dynamic domain reduction procedure for test data generation,” Software Practice and Experience, Vol. 29, No. 2, 1999, pp. 167-193.

[13] Q. Wang, S. Jiang and Y. Zhang, “An approach to generate basis path for programs with exception-handling constructs”, In IACSIT Press, 2012, International Conference on Computer Science and Information Technology (ICCSIT), Singapore.

[14] Dr. R.P. Mahapatra, M. Mohan and A. Kulothungan, “Effective tool for test case Execution time reduction,” In IACSIT, International Symposium on Computing, Communication and Control (CSIT), Singapore, 2011.

[15] C. Sharma, S. Sabharwal, R. Sibal, A survey on software testing techniques using genetic algorithm, IJCSI International Journal of Computer Science Issues, 2013, Vol. 10, No. 1, No 1.

[16] L. You, Y. Lu, "A genetic algorithm for the time-aware regression testing reduction problem", International conference on natural computation, IEEE, 2012, pp. 596 – 599.

[17] Jyoti and K. Solanki, A comparative study of five regression testing techniques : A Survey , INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH, Vol. 3, No. 8, 2014, pp. 76-80.

[18] R. Singh and M. Santosh, Test case minimization techniques: A review, International Journal of Engineering Research & Technology (IJERT), Vol. 2, No. 12, 2013. Pp.1048- 1056

[19] I. Hooda and R. Chhillar, A review: study of test case generation techniques, International Journal of Computer Applications. Vol. 107, No.16, 2014, pp. 33-37

[20] S. Biswas and R. Mall, Regression test selection techniques: A survey, Informatica 35. 2011, pp. 289–321

[21] M. Harrold, R. Gupt and M. Soffa, “A methodology for controlling the size of a test suite,” ACM Transactions in Software Engineering and Methodology, Vol. 2, No. 3, 1993, pp. 270-285.

[22] Isha Mangal Deepali Bajaj Priyanka Gupta, Regression Test Suite Minimization using Set Theory , International Journal of Advanced Research in Computer Science and Software Engineering 4(5), May - 2014, pp. 502-506

[23] S. Nachiyappan, A. Vimaladevi and C.B. SelvaLakshmi, &ldquo, An evolutionary algorithm for regression test suite reduction, Proc. Int',l Conf. Comm. and Computational Intelligence, 2010, pp. 503-508.

[24] S. Mohapatra, S. Prasad, "Finding representative test case for test case reduction in regression testing", IJISA, vol.7, no.11, , 2015, pp.60-65. DOI: 10.5815/ijisa.2015.11.08

[25] R. Singh and M. Santosh, Test case minimization techniques: A review, International Journal of Engineering Research & Technology (IJERT). Vol. 2, No. 12. 1048 -1056.

[26] Z. chen, b. xu, x. zhang, c. nie, A novel approach for test suite reduction based on Requirement relation contraction, ACM. 390-394.

[27] G. Fraser and F. Wotawa, Redundancy based test-suite reduction, In Proceedings of the 10th International Conference on Fundamental Approaches to Software Engineering, Springer. Vol. 4422, 2007, pp. 291-305.

[28] R. SALWAN , R. SEHGAL, Test cases reduction technique considering the time and cost as evaluation standards, International Journal of Computer Science and its Applications, 2013, pp.347-351.

[29] A. Gupta, N. Mishraa and D. Kushwaha, Rule based test case reduction technique using decision table, 2014, pp. 1398 – 1405.

[30] Z. Anwar and A. Ahsan, Multi-objective regression test suite optimization with Fuzzy logic, IEEE. INMIC 2013.

[31] Haider, A.A.; Rafiq, S.; Nadeem, A. "Test suite optimization using fuzzy -logic", Emerging Technologies (ICET), International Conference on, 2012, pp. 1 – 6

[32] Haider, A.A.; Nadeem, A.; Rafiq, S. "Computational intelligence and safe reduction of test suite", Emerging Technologies (ICET), IEEE 9th International Conference on, 2013, pp. 1 – 6

[33] Haider, A.A.; Nadeem, A.; Rafiq, S. "On the Fly Test Suite Optimization with FuzzyOptimizer", Frontiers of Information Technology (FIT), 11th International Conference on, 2013, pp.101 – 106

[34] B.subashini, d.jeyamala, Reduction of test cases using clustering Technique. International Journal of Innovative Research in Science, Engineering and Technology Vol 3, Special Issue 3, 2014, International Conference on Innovations in Engineering and Technology (ICIET’14). 1992-1995

[35] R. Wang, B. Qu, Y. Lu, Empirical study of the effects of different profiles on regression test case reduction, IET Softw., 2015, Vol. 9, No. 2, pp. 29–38

[36] S. Roongruangsuwan and J. Daengdej, Test case reduction methods by using CBR, Assumption University, ceur-ws.org,Vol-646,2010.

[37] P. Pringsulaka and J. Daengdej.2006., Coverall algorithm for test case reduction. In Aerospace Conference, 2006 IEEE.IEEE.

[38] C. Murphy, Z. Zoomkawalla and K. Narita, Automatic test case generation and test suite reduction for closed-loop controller software, Technical Report, 2013.

[39] S. Khan, A. Nadeem, TestFilter, a Statement-coverage based test case reduction technique. Proc. 10th IEEE Int, Multitopic Conf. 2006, pp. 275-280, doi:10. 1109/INMIC.

[40] M. Weiser, Program slicing. In Proceedings of the 5th international conference on Software engineering, ICSE ’81, 1981, pp. 439–449, Piscataway, NJ, USA, IEEE Pres0s.

[41] M. Weiser, Program slicing.1984. IEEE Trans, Softw. Eng. Vol. 10, No. 4, pp. 352–357.

[42] M. David Weiser, Program slices: formal, psychological, and practical investigations of an automatic program abstraction method. PhD thesis, Ann Arbor, MI, USA, 1979. AAI8007856.

[43] Bogdan Korel and Janusz Laski. Dynamic program slicing. In Information Processing Letters, 1988.

[44] Bogdan Korel and Janusz Laski. Dynamic slicing of computer programs. J. Syst. Softw., 13(3):187–195, December 1990.

[45] T. Gyim´othy, A. Besz´edes, and I. Forg´acs, An efficient relevant slicing ´ method for debugging. SIGSOFT Softw. Eng. Notes, Vol. 24, No. 6, 1999, pp. 303–321.

[46] D. Binkley, Semantics Guided Regression Test Cost Reduction. IEEE International Conference on Software Maintenance, Vol. 23, No. 8, 1997, pp. 498516.

[47] A. Mohammadian , B. Arasteh, Using program slicing technique to reduce the cost of software testing. Journal of Artificial Intelligence in Electrical Engineering, Vol. 2, No.7, 2013, pp.24-33.

[48] Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2001). Introduction to algorithms. MIT Press, Cambridge, MA

[49] Ana Emília V. B. Coutinho, Emanuela G. Cartaxo1, Patrícia D. L. Machado. "Test suite reduction based on similarity of test cases." 7st Brazilian workshop on systematic and automated software testing—CBSoft 2013.

[50] S. Xu, H. Miao and H.Gao, Test suite reduction using weighted set covering techniques, 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. IEEE, 2012.

[51] L. Zhang, D. Marinov, L. Zhang and S. Khurshid, An empirical study of JUnit test-ruite Reduction, 22nd IEEE International Symposium on Software Reliability Engineering, 2011, pp.170-179.

[52] B. Suri, I. Mangal, and V. Srivastava, Regression test suite reduction using an hybrid technique based on BCO and genetic algorithm, Special Issue of International Journal of Computer Science & Informatics (IJCSI), ISSN (PRINT) : 2006, 2231–5292, Vol.- II, No-1, 2

[53] Sampath, S.; Bryce, R.; Memon, A.M, "A Uniform representation of hybrid criteria for regression testing", Software Engineering, IEEE Transactions on. Vol. 39, No. 10. , 2013, pp. 1326 – 1344.

[54] S. Yoo, M. Harman, "Using hybrid algorithm for Pareto efficient multi-objective test suite minimization", The Journal of Systems and Software .83, 2010, pp. 689–70.

[55] Nguyen Huu Phat. 2013. Slicing-based test case generation 2013, University of Bordeaux. Internship Report

Energy Provisioning Technique to Balance Energy Depletion and Maximize the Lifetime of Wireless Sensor Networks

Hassan Hamid Ekal[1*]
Department of Communication Engineering, Faculty Of Electrical and Electronic Engineering, University Tun Hussein Onn Malaysia, Johor, Malaysia

Jiwa Bin Abdullah[2]
Department Of Communication Engineering, Faculty Of Electrical and Electronic Engineering, University Tun Hussein Onn Malaysia, Johor, Malaysia

***Abstract*—With the promising technology of Wireless Sensor Networks (WSNs), lots of applications have been developed for monitoring and tracking in military, commercial, and educational environments. Characteristics of WSNs and resource limitation impose negative impacts on the performance and effectiveness of these applications. Imbalanced energy consumption among sensor nodes can significantly reduce the performance and lifetime of the network. In multi-hop corona WSN, the traffic imbalance among sensor nodes will make nodes located near to the sink consume more energy and finish their energy faster than those distant from the sink. This would cause what is called "energy hole", which prevents the network from performing the intended tasks properly. The objective of the work in this paper is to balance energy consumption to help improving the lifetime of corona-based WSNs. To maximize the lifetime of the network, an innovative energy provisioning technique is proposed for harmonizing the energy consumption among coronas by computing the extra needed energy in every corona. Experimental results of the evaluation revealed that the proposed technique could improve the network lifetime noticeably via fair balancing of energy consumption ratio among coronas.**

***Keywords*—Wireless Sensor Network (WSN); Lifetime; Node deployment; Energy provisioning**

I. INTRODUCTION

The high potential for utilization in numerous applications has made wireless sensor networks (WSNs) widely popular. A lot of applications regarding event and activity measurements have been developed for monitoring and tracking in different environments, where human access is dangerous. These applications include intelligence traffic and supply chain management, health care and habitat monitoring, gas and temperature detection, military and national security domains, and many others [1, 2]. WSNs are considered as dynamic multi-hop routing networks, where sensor nodes are connected to centralized powerful machine, called Base Station (BS) or sink. WSNs characteristics impose serious challenges on the design of WSNs [3]. Due to the size and nature, sensors have restricted power capacity, and thus, they have limits in processing and computing capacities.

In most scenarios, sensor nodes are immobile and do not change location after deployment. However, sensors activities usually result in power dissipation in some regions, which causes a dynamic change in the network topology. Also, sensor nodes are programmed to alternatively be in active mode and in sleeping mode to save energy. In sleeping mode, sensors are disconnected from the network as they turn off their transceivers. With such dynamic changes in the network topology, providing connectivity, and at the same time, minimizing the energy consumption is difficult.

Thus, reliable energy-efficient mechanisms for WSNs can help prolong the lifetime of the network considerably [4]. Figure 1 shows an example of WSN architecture in operation.

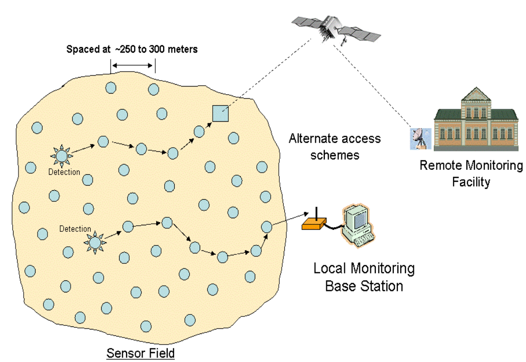


Fig. 1. Example of WSN architecture

With the inherited limitations of sensor nodes, WSNs therefore have networking resource limitation such as communication range and bandwidth, frequent topology changes, high density and node unreliability [5]. In addition to the aforementioned characteristics and limitations, there are several challenges in deploying WSNs effectively. These challenges include scalability, routing, energy consumption, as well as self-organizing and fault tolerance [6].

II. REQUIREMENTS AND CHALLENGES OF SENSING APPLICATIONS

Sensor nodes should be capable of managing, controlling and organizing themselves into a network in order to effectively perform their tasks [7]. This would help in facilitating the network management and satisfying the application requirements. However, the limited resources of sensor nodes make it necessary to have reliable communication protocols and efficient management services. Improving communications and reducing energy consumption in the network require communications protocols to be implemented at different layers that present the protocol stack.

This can be noticeably operative with regard to network efficiency, as well as latency and energy consumption. Nonetheless, communication protocols developed for traditional networking are not appropriate for WSNs, as they are not projected to operate under such resources limitation. Therefore, numerous energy-efficient protocols have been proposed to operate in all of the five layers in the protocol stack. Communications among the layers is supported by utilizing the concept of cross-layering. In other words, the information of protocol in a certain layer is shared across the rest of the layers, to satisfy some requirements of WSNs [1] [8-13].

As sensor nodes operate with a limited power, their energy usage is very important matter in designing and deploying WSNs. Consequently, much research work has been done on energy harvesting and moderating energy consumption [14-21] [23]. Once a sensor node finishes its energy, it will be disconnected from the network; this may have significant effect on the performance of the application. The network lifetime basically relies on the connectivity and number of active nodes in the network. Therefore, energy should be utilized resourcefully in order to make the best use of the lifetime of the network operation. Energy or power harvesting (also known as energy scavenging) is the process where sensor nodes derive their energy from external sources including solar cells [6] [22], fuel cells, wind energy, salinity gradients, mobile supplier, and acoustic noise [23].

These energy source techniques are utilized to overcome the energy limitation and also to provide satisfactory level of quality of service (QoS) to the application. Reliable communications are required to support the data transmission to the intended destination. Thus, buffer monitoring, congestion control and packet-loss recovery mechanisms are needed. Communication strength relies on the sensors placement. Sparse deployment of sensors may result in higher energy usage due to distant transmission. On the other hand, dense deployment of sensors may lead to more local communications due to short-range transmission. Coverage is of vital concern when it comes to sensor deployment. The coverage area of the network is specified by the number of sensors deployed [6].

III. Energy Issue in Corona-Based WSNs

In many-to-one multi-hop network architecture, energy consumption of the sensor nodes is often imbalanced. Nodes placed close to the sink (or BS) certainly consume more energy than those located far from the sink [24], and accordingly, they finish their energy (die) earlier than the distant ones which are unable to send data to the sink even though they have much unused energy. Consequently, energy holes (hot spots) are created in the network [25-27]. In contrast, without multi-hopping, distant sensor nodes have to transmit data to the sink directly. Thus, due to long transmission distance, they finish their energy faster than the ones closer to the sink. The network is partitioned into islands when energy holes phenomenon happens. This will reduce the network lifetime considerably, and the network dies before performing the intended tasks completely. In multi-hop corona WSNs (Figure 2), energy holes occur at the inner coronas (rings) which are close to the sink.

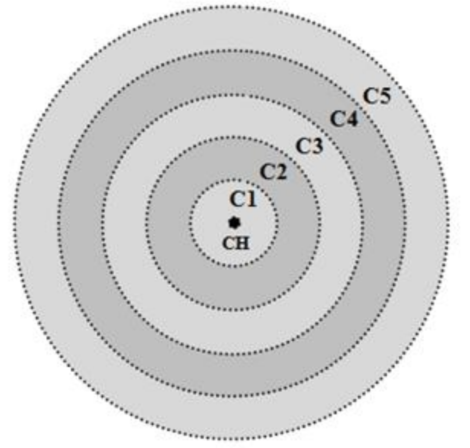


Fig. 2. Many-to-one Multi-hop corona-based WSN

There have been several solutions proposed to alleviate this problem through the use of non-uniform node deployment, dynamic methods (through mobile sink or clustering algorithm), transmission range (control of multi-level ranges), and heterogeneous nodes (nodes with capabilities for relaying data and/or energy provisioning)[28-34]. Figure 3 shows some of the solutions, which can be demonstrated by using multi-hop corona network model, to overcome the energy holes problem.

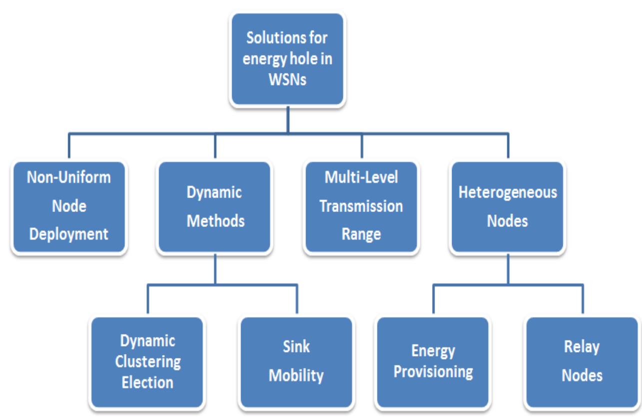


Fig. 3. Solutions developed to overcome energy hole problem

The lifetime of the network can be prolonged by eliminating (or limiting the development of) the energy holes. This can be achieved by designing energy-efficient solutions for the five layers that form WSN, namely physical layer, data link layer, network layer, transport layer, and application layer. In designing those solutions, several issues such as security, storage, synchronization, localization, coverage, data compression and aggregation should be considered. Reliable algorithms of those solutions may improve the network lifetime remarkably.

Improving the lifetime is one of the challenges when designing WSNs, as it relies on a number of parameters such as network architecture, protocols, energy model, channel characteristics, techniques for data collection, and also on how the lifetime of the network is defined [35]. Tackling and mitigating the factors (such as area under monitoring, number of coronas, corona width, node deployment strategy, and node transmission range) that bring about or precipitate the energy hole may also enhance the lifetime.

The aim of the research presented in this paper is to propose an energy-balanced provisioning technique to alleviate the energy hole problem in multi-hop corona WSNs to overcome the limitations of the previous research works and provide practicable enhancement to WSNs in terms of network lifetime, energy consumption, and connectivity. The rest of this paper is organized as follows: Section 4 presents the design of the proposed energy provisioning technique. Section 5 describes the simulation scenarios for performance evaluation of the proposed technique and presents the analytical results gained from simulations and the comparison to the traditional uniform technique from energy distribution point of view. Finally, Section 6 concludes this paper and suggests some related future work.

IV. Proposed Energy Provisioning Technique

Energy depletion of the coronas can be balanced by distributing the energy to some extent, but requires the energy in each corona to be identified beforehand. Also, to achieve the maximum network lifetime, it is necessary to consider the additional energy needed to balance the energy consumption between coronas by defining the initial energy of each node with regard to the corona location where the node resides. This can be achieved by using a new energy provisioning technique that is based on innovative mathematical system.

This section presents theoretical design of the proposed energy provisioning technique, which is based on the calculation of the energy increment ratio in each corona, aiming at balancing the energy consumption. For energy provisioning, the mathematical relationship between network lifetime and the additional energy that is necessary to balance energy depletion among coronas will be estimated. The procedure starts by calculating the network lifetime in each corona. Then, the relationship between the energy consumption of each corona and the energy needed to relay packets received from the nodes of the outer coronas is determined. Next, the extra energy required for nodes in each corona to support the transmission of data packets from outer coronas is estimated. After that, the relationship between increasing the network lifetime and the extra energy required is formulated. Here, for energy increment computation, the extra energy required to balance the energy in each corona is computed.

A. *Corona model and the corresponding assumptions & definitions*

The research work presented in this paper considers a circular multi-hop corona WSN with immobile sink placed at the center of the circular area. The network radius is rad_R and each node in the network operates with a transmission range of *R*. As the corona network model can be represented by $(k+1)$-tuple $(k, w_1, \dots, w_k)$, the corona model covering the circular area can be divided into k number of coronas *C* of a specific width w. C_{i} indicates the *i*th corona where C_1 represents the first corona (innermost corona). w_1 denotes the radius of the inmost corona, while $w_i = 2, 3, \dots, k$ signifies the width of any other corona *i*. Corona network model can be categorized into uniform-width and non-uniform-width, based on the corona width/radius. In the uniform-width corona model, $w_i = s_i$ for all $i = 1, 2, 3, \dots, k$, meaning that width of every corona is the same as the sensing range of the node. On the other hand, in the non-uniform-width corona model, there is only one corona with a width (or radius) that is different from the rest of coronas, either shorter or longer. In this paper, only the uniform-width corona model was considered to avoid associated complexity and high cost of estimating the width of non-uniform corona model. Every node was to create and transmit l bits per unit of time via the multi-hop communication, through which, sensor nodes which belonged to corona C_i would forward incoming data to next corona$(C_{i-1} | \, i \geq 2)$. The area under monitoring was assumed within an environment with PLE of 4 (that is, *n*=4), and the communication environment was considered contented and error-free. The proposed energy provisioning technique was considered for constant monitoring applications only, as event-driven applications are out of the scope of the research. The commonly used q-switch routing algorithm in WSN was utilized in this study, while data aggregation was not considered as it might cause imbalanced energy consumption.

For corona-based WSNs, the following definitions regarding the lifetime are applicable:

Definition 4.1. The lifetime of a corona C_i (lC_i)(in unit time) is determined by the ratio:

$$lC_i = \frac{T\varepsilon_{C_i}}{EC_i} \tag{1}$$

where $T\varepsilon_{C_i}$ is the total initial energy of C_i, and EC_i is the energy consumption in corona C_i. $T\varepsilon_{C_i}$ can be obtained from εN_i , where ε represents the node initial energy and N_i, represents the number of nodes in corona C_i.

Definition 4.2. The lifetime of the network lt_{net} is the total time from the point at which the network starts operating, to the point when the first node finishes its energy (dies). In other words, the lifetime of the network is represented by the lifetime of first node. In case the energy consumption of nodes inside corona C_i being uniform, the energy consumption of each node (EN_i) in corona $C_i$is:

$$EN_i = \frac{EC_i}{N_i} \tag{2}$$

Accordingly, the lifetime of every node (lt_{N_i}) in corona C_i is estimated using:

$$lt_{N_i} = \frac{\varepsilon N_i}{EC_i} \quad (3)$$

The lifetime of each sensor node (lt_{N_i}) should be similar to the lifetime of C_i (lC_i) as determined in Definition 4.1. Hence, the lifetime of the network is defined by the shortest lifetime of any corona, as follows:

$$lt_{net} = min_{C_i} \frac{\varepsilon N_i}{EC_i}, \forall C_i \quad (4)$$

Definition 4.3. For corona model $(k, w_1, \ldots, w_k)$ with a sensing range s_i of the node for $i = 1, 2, 3, \ldots, k$ in corona C_i where $\frac{w_i}{2} \leq s_i\sqrt{2}\,(\sum_{j=1}^{i} w_j)$, the maximum corona coverage can be achieved if the sensor nodes are deployed with distance d_i^{dep} from the center of corona C_i. d_i^{dep} can be computed as follows:

$$d_i^{dep} = \sqrt{\frac{\left(\sum_{j=1}^{i-1} w_j\right)^2 + \left(\sum_{j=1}^{i} w_j\right)^2 - 2s_i^2}{2}}, \quad i = 2, 3, \ldots, k \quad (4)$$

B. Computation of the ratio of energy increment

This part of the research work aims to compute the amount of the extra energy required in each corona i to balance the energy. This is done by estimating the increment ratio of energy V needed in any corona C_i, which helps in balancing the energy depletion among the sensor nodes, based on the geometric area of coronas in the area of monitoring.

To relay a packet from outer coronas towards the sink, each sensor node located in corona i would need E_{N_i} amount of energy, which can be obtained as in Equation (5) below:

$$E_{S_i} = (\mathrm{E_{rx}} + \mathrm{E_{tx}})/\mathrm{E_{tx}} \quad (5)$$

where E_{tx} denotes the energy needed to transmit one bit of data over a definite distance d, which can be computed as in Equation (6) below:

$$E_{tx} = E_{elec} + \alpha d^n \quad (6)$$

where α denotes the energy depletion over the operational amplifier throughout the data transmission, d which indicates the transmitter-receiver distance, and n indicates the Path Loss Exponent (PLE) depending on the system environment.

E_{rx} in Equation (5) denotes the energy needed for data reception, which can be achieved as in Equation (7).

$$E_{rx} = E_{elec} \quad (7)$$

where E_{elec} represents the electronic energy. Accordingly, if by letting $\mathbb{R}^{(i+1),k}$ to denote the region covered from corona (i+1) to corona k (i.e. $[(i+1), k] = \{C_i \in \mathbb{R}^{(i+1),k} \mid i+1 \leq C_i \leq k\}$, and $\mathbb{R}_C^i$ to denote the region of corona C_i, then the increment ratio of energy of any corona C_i for conveying the packets from outer coronas to the innermost corona is estimated as in Equation (8):

$$V_i = \frac{\mathbb{R}^{(i+1),k}}{\mathbb{R}_C^i} \times E_{N_i} \quad (8)$$

From Equation (8), Equation (9) is derived as:

$$V_i = \frac{\int_0^{2\pi}\int_{i\times \mathrm{w}}^{rad_R} \mathrm{r}d\mathrm{r}d\theta}{\int_0^{2\pi}\int_{(i-1)\times w}^{i\times w} \mathrm{r}d\mathrm{r}d\theta} \times E_{N_i} \quad (9)$$

Further simplification yields Equation (10) as follows:

$$V_i = \frac{{rad_R}^2 - i^2 w^2}{2iw^2 - \mathrm{w}^2} \times \frac{\mathrm{E_{rx}} + \mathrm{E_{tx}}}{\mathrm{E_{tx}}} \quad (10)$$

where rad_R, w, E_{rx}, and E_{tx} denote the radius of the network, width of each corona, reception energy and transmission energy of one bit, respectively. By letting $rad_R = k \times w$, then the increment ratio of energy of corona C_i needed to achieve energy balancing is calculated as:

$$V_i = \frac{k^2 - i^2}{2i - 1} \times E_{N_i} \quad (11)$$

In this case, the packets generated in the external coronas should be relayed to the inner ones via corona C_i, thus, the amount of energy needed in corona C_i in order to relay the received packets must be supplied to corona C_i.

C. Sensing, relaying, and idle scheduling

To achieve improved efficiency of the network, it is assumed that each sensor node can be in one of two sates, either in the sensing state or in the idle state. Sensor nodes which are selected to sense the monitoring area can perform sensing only or along with data relaying, whereas those not selected for sensing may remain idle or relay data received from other coronas.

Energy consumption of a sensor node can be controlled by proper scheduling between the sensing and relaying operations or move to idle state, by utilizing efficient and simple turn-based scheduling algorithm. Counter (*count*) is used to regulate whether a sensor node should sense/relay by that time or not. The algorithm is presented in Figure 4.

```
/*as the sensor node still has energy, it is enabled to sense the monitoring area*/
whlie (nodeEnergy > 0)
    Sense = true;

    if (count_i mod count_i^max = 0) then
    /* the sensor node is enabled to sense and also relay data from other corona C_i*/
        if (Relay = true) then
            Send(ownData + relayData);
        else
        /* the sensor node sends its own data only*/
            Send(ownData);
        end if

    else if (count_i mod count_i^max ≠ 0) then
    /* the sensor node is not enabled for sensing, but enabled to relay data*/
        if (Relay = true) then
            Send(relayData);
        end if
    end if

    /* Change the count value for the sensor node*/
    count_i = (count_i + 1) mod count_i^max;
end whlie
```

Fig. 4. Pseudo-code of the scheduling algorithm

While nodes in corona$C_i (i = 1, 2, 3, \ldots, k)$ are placed with uniform space of $2\pi d_i^{dep}/N_i$ to each other, sensor nodes in corona C_1 are placed in the middle of the corona width (at point $w_1/2$)with uniform space of $\pi w_1/N_1$ to each other.

Before being in either state, sensor nodes should be assigned a to $count_i$ value that varies from 1 to $count_i^{max} = N_i/{N_i}^{min}$, where ${N_i}^{min}$ is the minimum number of deployed nodes that can cover corona C_i . The *count* value of $count_i^{max}$ is assigned to the first sensor node and the value $count_i^{max} - 1$is assigned to the following sensor node. This procedure is repeated until *count* of value 1 is reached, and until all nodes are assigned with *count* values.

Nodes operate in rounds, where they perform sensing, transmitting, and receiving data in a fixed time interval. When the network operates, the sensor node starts sensing (and might forward data as well) if it's $count_i \bmod count_i^{max} = 0$. In the subsequent round, *count* value will be increased by 1 for every sensor node in the network. This is done in a cyclic routine to define which node should sense the monitoring area and relay data; if the current value for *count* is $count_i^{max}$, the next value for *count* will be 1, and so on.

V. Performance Evaluation

To validate the proposed energy provisioning technique aimed at balancing energy consumption and maximizing the lifetime in WSNs, evaluation scenarios (described in the following subsection) had been simulated using MATHLAB on a machine running the CentOS 5.5 version of the Linux operating system.

A. Evaluation scenarios

The corresponding parameters setup for the evaluation scenarios of the proposed technique is presented in Table 1.

TABLE I. Scenario Parameters Setup

Parameter	Value
Initial energy of each node ε	0.5J
Electronic energy consumption E_{elec}	50nJ/bit
Network radius rad_R	500m
Geometric progression q	2
Energy dissipate α	0.0013pJ/bit/m^4
Packet length l	300bits
Path loss exponent n	4
Number of coronas k	10
Uniform corona width w	100m
Transmission range R	100m

For the evaluation scenarios, the sensor transmission range was made equal to the sensing range and also to the width of the corona it resided in. Having a corona network with radius rad_R, the total number of coronas k was achieved by using:

$$k = \frac{rad_R}{R} \tag{12}$$

That is, for a given network radius rad_R of 500m and Transmission range R of 100m, the number of coronas k should be 5.With each round of 1ms period, every sensing node would generate data of 300bits.

For comparison purpose, the minimum number of sensor nodes ${N_i}^{min}$ for covering a uniform coronaC_i width$w$ of 100 was set to 12, 36, 64, 84, and 120, for N_1, N_2, N_3, N_4, N_5, respectively, based on Cosine Rule. The performance experiments of the evaluation scenarios had been conducted 20 times as the minimum test requirement in computer sciences to achieve non-overlapping confidence interval. The evaluation experiments were done for both the energy distribution based on the proposed technique and normal (uniform) energy distribution under similar parameters setup described earlier.

B. Experimental results

Figure 5 shows the ratio of energy needed to be added to each of the four coronas C_1, C_2, C_3, C_4 (as C_4 was the last corona and its sensor nodes did not relay data from other coronas) in order to improve the network lifetime with balanced energy consumption. From the figure, to maximize the lifetime of a network of radius rad_R of 500m where the width of each corona was 100m, the ratios of required additional energy to the corresponding coronas C_1, C_2, C_3, C_4 were approximately 65.50, 21.40, 12.19 and 9.00 times more than the normal (uniform) case without the use of the proposed technique.

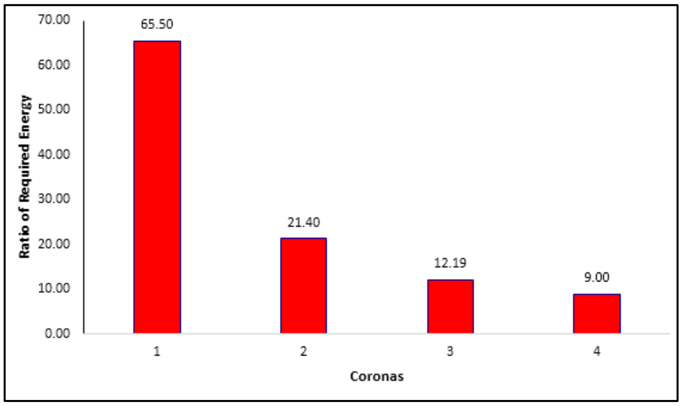


Fig. 5. Ratio of energy increment in each corona for balanced energy consumption

With the evaluation of the proposed technique by increasing the energy in the network and distributing the extra energy in the coronas accordingly, the results also showed that the network lifetime had been increased considerably. Table 2 presents the percentage of the lifetime increase for coronas C_1, C_2, C_3 and C_4 when the total network energy was increased using the proposed technique compared to the use of normal (uniform) energy distribution.

TABLE II. LIFETIME RATIO GAINED THROUGH INCREASING ENERGY USING THE PROPOSED TECHNIQUE VS. NORMAL ENERGY DISTRIBUTION

Corona Number	Lifetime before energy increase	Lifetime after energy increase using normal energy distribution	Lifetime after energy increase using the proposed technique	Ratio of lifetime increase using normal energy distribution	Ratio of lifetime increase using the proposed technique
1	192.3005	384.6	384.6005	2	2
2	768.405	1536.815	3513.395	2	4.6
3	1720.645	3441.29	11846.65	2	6.9
4	3019.33	6038.66	28126.65	2	9.4

From the data as shown in Table 2, it is clear that, with the use of the proposed technique, the total lifetime could be increased 9.4 ($> 2^3$) times, whereas it could be increased only by 2 ($=2^0$) times when using normal energy distribution. It was found that, by increasing the total energy according to derived Equation (11) in the proposed energy provisioning technique, the network lifetime could be increased remarkably. The results showed that the lifetime had been improved by about 40% when using the proposed technique compared to the uniform technique for energy distribution.

VI. CONCLUSION AND FUTURE WORK

In multi-hop corona Wireless Sensor Networks (WSNs) with uniformly distributed nodes, when an energy hole appears due to the death of some nodes in critical location, data packets cannot be sent from distant nodes to the sink. This means that the network lifetime finishes early, resulting in wastage of a significant amount of energy. This paper presents the theoretical design and development of an innovative energy provisioning technique to balance energy depletion and maximize network lifetime. The experimental evaluation results reveal that, when using the proposed technique for efficient energy distribution by computing the extra needed energy, the lifetime can be enhanced by about 40% compared to that of the uniform technique. With the proposed technique, the network can be considered homogeneous if the sensor nodes capabilities are the same, and that the extra added (required) energy can be provided by adding more sensors to meet the total energy requirements in specific coronas. Otherwise, sensor nodes in those coronas could be supplied by different initial energy to fulfill the energy needed. The network in such case is heterogeneous, where defining the initial energy in every node with regard to its corona can be beneficial to enable a proper utilization of the proposed technique in heterogeneous WSN. This objective is considered in the future work.

ACKNOWLEDGEMENT

The authors would like to thank the management and the Centre For Graduate Studies of University Tun Hussein Onn Malaysia (UTHM) for sponsoring the research. The authors would like also to appreciate the assistance from the Office for Research, Innovation, Commercialization and Consultancy Management (ORICC).

REFERENCES

[1] Holger Karl, A. W., Protocols and architectures for wireless sensor networks, Book, ISBN: 0-470-09510-5, 2005.

[2] Shi Lan; Miao Qilong; Jinglin Du, "Architecture of Wireless Sensor Networks for Environmental Monitoring," Education Technology and Training, 2008. and 2008 International Workshop on Geoscience and Remote Sensing. ETT and GRS 2008, vol.1, no., pp.579, 582, 21-22 Dec. 2008.

[3] Islam, A. K M M; Baharun, S.; Wada, K., "An overview on Dynamic Wireless Sensor Network Architectures," Informatics, Electronics & Vision (ICIEV), 2012 International Conference on , vol., no., pp.464,468, 18-19 May 2012.

[4] Akyildiz, I. F., Wireless Sensor Networks, Series in Communications and Networking. John Wiley amp; Sons Ltd., 2010.

[5] Houda Labiod, Wireless Ad Hoc and Sensor Networks ISTE Ltd, John Wiley and Sons, Inc. ISTE Ltd, 2008

[6] Yick, J., Mukherjee, B., & Ghosal, D., Wireless sensor network survey. Computer Networks, 52, 2292-2330, 2008.

[7] Jiong Jin; Yee Wei Law; Wei-Hua Wang; Palaniswami, M., "A hierarchical transport architecture for wireless sensor networks," International Conference on Intelligent Sensors, Sensor Networks and Information Processing, 2008. ISSNIP 2008, vol., no., pp.145,150, 15-18 Dec. 2008

[8] Berfield, A. & Mosse, D., (2006), Efficient Scheduling for Sensor Networks, 3rd Annual International Conference on Mobile and Ubiquitous Systems - Workshops, 2006.

[9] Berfield A., Chrysanthis P.K & Labrinidis A., (2006), Efficient handling of sensor failures, Proceedings of the 3rd workshop on Data management for sensor networks: in conjunction with VLDB 2006, ACM.

[10] Zabin, F. , Misra, S., Woungang, I. , Rashvand, H.F. , Ma, N.-W. & Ali, M. A., (2008), REEP: data-centric, energy-efficient and reliable routing protocol for wireless sensor networks, Communications, IET, pages 995 – 1008.

[11] Casari, P. , Nati, M. ; Petrioli, C. & Zorzi, M. (2007), Efficient Non-Planar Routing around Dead Ends in Sparse Topologies using Random Forwarding, ICC '07. IEEE International Conference on Communications, Pages 3122 – 3129.

[12] Chamam A. & Pierre S., (2009), A distributed energy-efficient clustering protocol for wireless sensor networks, Computers & Electrical Engineering, Volume 36, Issue 2, March 2010, Pages 303–312

[13] Baghyalakshmi, D., Ebenezer, J. & Satyamurty, S.A.V., (2010), Low latency and energy efficient routing protocols for wireless sensor networks, International Conference on Wireless Communication and Sensor Computing, ICWCSC 2010. Pages 1 – 6.

[14] Azad, A. K. M., & Kamruzzaman, J. (2011). Energy-Balanced Transmission Policies for Wireless Sensor Networks. Mobile Computing, IEEE Transactions on, 10, 927-940.

[15] Ammari, H. M., & Das, S. K., Promoting Heterogeneity, Mobility, and Energy-Aware Voronoi Diagram in Wireless Sensor Networks. IEEE Transactions on Parallel and Distributed Systems, 19, 995-1008, 2008.

[16] Charilaos, E., Sotiris, N., & Jose, R. (2006). Energy balanced data propagation in wireless sensor networks. Wirel. Netw., 12, 691-707.

[17] Chen, Y., Li, Q., Fei, L., & Gao, Q. (2012, 9-12 Sept. 2012). Mitigating energy holes in wireless sensor networks using cooperative communication. Paper presented at the Personal Indoor and Mobile Radio Communications (PIMRC).

[18] Cheng, P., Chuah, C. N., & Liu, X. (2004). Energy-aware node placement in wireless sensor networks. Paper presented at the Global Telecommunications Conference, 2004, GLOBECOM'04.

[19] Guo, W., Liu, Z., & Wu, G. (2003). An energy-balanced transmission scheme for sensor networks. Paper presented at the 1st international conference on Embedded Networked Sensor Systems.

[20] Haibo, Z., & Hong, S., Balancing Energy Consumption to Maximize Network Lifetime in Data-Gathering Sensor Networks, IEEE Transactions on Parallel and Distributed Systems, 20, 1526-1539, 2009.

[21] Esseghir, M., Bouabdallah, N., & Pujolle, G. (2007). Energy provisioning model for maximizing wireless sensor network lifetime. Paper presented at the Global Information Infrastructure Symposium, GIIS 2007.

[22] Raghunathan, V., Kansai, A., Hse, J., Friedman, J., & Srivastava, M. (2005). Design considerations for solar energy harvesting wireless embedded systems. Information Processing in Sensor Networks, IPSN 2005, 1, 457-462.

[23] Rahimi, M.; Shah, H.; Sukhatme, G.; Heideman, J.; Estrin, D., "Studying the feasibility of energy harvesting in a mobile sensor network," IEEE International Conference on Robotics and Automation, ICRA '03, vol.1, no., pp.19,24 vol.1, 14-19 Sept. 2003.

[24] Ramos, H.S.; Oliveira, E.M.R.; Boukerche, A.; Frery, A.C.; Loureiro, A.A.F., "Characterization and mitigation of the energy hole problem of many-to-one communication in Wireless Sensor Networks," International Conference on Computing, Networking and Communications (ICNC), vol., no., pp.954,958, Jan. 30 2012-Feb. 2 2012.

[25] Perillo, M., Cheng, Z., & Heinzelman, W. (2005, 17-21 July 2005). An analysis of strategies for mitigating the sensor network hot spot problem. Paper presented at the Mobile and Ubiquitous Systems: Networking and Services, MobiQuitous 2005.

[26] Meicheng, L., Jie, Z., Ming, L., Yuming, B., "A novel solution for energy hole of Wireless Sensor Network," 33rd Chinese Control Conference (CCC), vol., no., pp.456,460, 28-30 July 2014

[27] Pathak, A.; Zaheeruddin, Z.; Tiwari, M.K., "Minimizing the Energy Hole Problem in Wireless Sensor Networks by Normal Distribution of Nodes and Relaying Range Regulation," Fourth International Conference on Computational Intelligence and Communication Networks (CICN), vol., no., pp.154,157, 3-5 Nov. 2012

[28] Nazir, B.; Hasbullah, H., "Mobile Sink based Routing Protocol (MSRP) for Prolonging Network Lifetime in Clustered Wireless Sensor Network," International Conference on Computer Applications and Industrial Electronics (ICCAIE), vol., no., pp.624,629, 5-8 Dec. 2010

[29] Xiaobing, W., Guihai, C., & Das, S. K., Avoiding Energy Holes in Wireless Sensor Networks with Nonuniform Node Distribution. IEEE Transactions on Parallel and Distributed Systems, 19, 710-720, 2008.

[30] Qian Zhao; Nakamoto, Y., "Routing Algorithms for Preventing Energy Holes and Improving Fault Tolerance in Wireless Sensor Networks," *Computing and Networking (CANDAR), 2014 Second International Symposium on* , vol., no., pp.278,283, 10-12 Dec. 2014.

[31] Yu Xue; Xiangmao Chang; Shuiming Zhong; Yi Zhuang, "An efficient energy hole alleviating algorithm for wireless sensor networks," *Consumer Electronics, IEEE Transactions on* , vol.60, no.3, pp.347,355, Aug. 2014.

[32] Liu Meicheng; Zhang Jie; Lyu Ming; Bo Yuming, "A novel solution for energy hole of Wireless Sensor Network," *Control Conference (CCC), 2014 33rd Chinese* , vol., no., pp.456,460, 28-30 July 2014.

[33] Saleem, F.; Javaid, N.; Moeen, Y.; Akbar, M.; Khan, Z.A.; Qasim, U., "MEET: Multi-hop Energy Efficient Protocol for Energy Hole Avoidance Using Variable Transmission Range in Wireless Sensor Networks," *Broadband and Wireless Computing, Communication and Applications (BWCCA), 2014 Ninth International Conference on* , vol., no., pp.478,484, 8-10 Nov. 2014.

[34] Kim, Min-Gon; Young-Tae Han; Hong-Shik Park, "Energy-Aware Hybrid Data Aggregation Mechanism Considering the Energy Hole Problem in Asynchronous MAC-Based WSNs," *Communications Letters, IEEE* , vol.15, no.11, pp.1169,1171, November 2011.

[35] Chen, Y., & Zhao, Q. (2005). On the lifetime of wireless sensor networks. *Communications Letters, IEEE, 9*, 976-978.

Learning on High Frequency Stock Market Data Using Misclassified Instances in Ensemble

Meenakshi A.Thalor
Research Scholar,Computer Engineering
Vishwakarma Institute of Technology
Savitribai Phule Pune University
Pune, India

Dr. S.T.Patil
Professor, Computer Engineering
Vishwakarma Institute of Technology
Savitribai Phule Pune University
Pune, India

***Abstract*—Learning on non-stationary distribution has been shown to be a very challenging problem in machine learning and data mining, because the joint probability distribution between the data and classes changes over time. Many real time problems suffer concept drift as they changes with time. For example, in stock market, the customer's behavior may change depending on the season of the year and on the inflation. Concept drift can occurs in the stock market for a number of reasons for example, trader's preference for stocks change over time, increases in a stock's value may be followed by decreases. The objective of this paper is to develop an ensemble based classification algorithm for non-stationary data stream which would consider misclassified instances during learning process. In addition, we are presenting here an exhaustive comparison of proposed algorithms with state-of-the-art classification approaches using different evaluation measures like recall, f-measure and g-mean.**

Keywords*—*Classifiers; Concept drift; Data stream; Ensemble; Non-stationary Environment

I. Introduction

Nowadays most of the applications are online applications, where huge amount of data increasingly arrives at every time stamp which is generated from different sources. So it is very important to train classifiers incrementally over the time so that they can learn different concepts of non-stationary data streams.

Conventional data mining algorithms assumes that each dataset is produced from a single, static and hidden function i.e. the function (model/classifier) generating data at training time is the same as that of testing time. Whereas in non-stationary data stream, data is continuously coming and the function which is generating instances at time *t* need not be the same function at time *t+1*. This difference in the underlying function is called as concept drift [1].

The concept drift problem is studied in literature with different terminology as "concept shift" , "concept drift", dataset shift ,"change of classification", "changing environments", "non-stationary environment" etc. Concept drift in data stream happens when the relationship between the input and class variables changes over time and this can happen because of change in the following:

1) The class priors, $P(c_i)$, $i = 1, 2, 3, \dots k$, where k is the number of classes;

2) The distribution of the classes, $P(X| c_i)$, where $i = 1, 2, 3, \dots k$ and X is a vector of labeled instances; and

3) The posterior distribution of the class membership $P(c_i |X)$, $i = 1, 2, 3, \dots k$

For providing training to classifiers incrementally over the time so that they can learn different concepts of non-stationary data streams we are using ensemble based approach [2] as shown in fig 1.

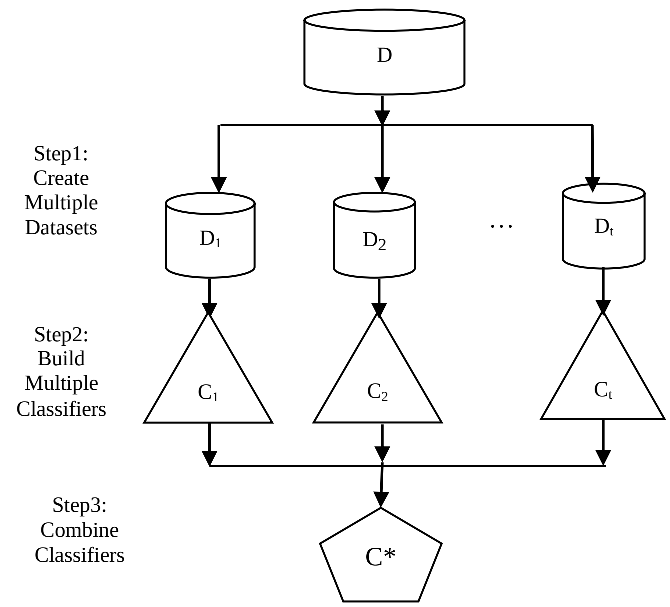


Fig. 1. Ensemble Based Learning

In data mining, the ensemble is a pool of classifiers whose individual classifications/predictions are combined in some way to classify unseen examples. The strategy in ensemble systems [3] is to create subsets of incoming data stream and for each subset a classifier is trained and tested and then these classifiers collectively would do decision making and predict the label for unseen data

Performance of learning algorithms dependent upon the size of the data chunks (block/batch). Bigger blocks [4] can results in accurate classifiers as classifiers are getting more data for training, but can contain too many different concept drifts. Whereas smaller blocks are better for drifted data stream, but usually lead to poorer classifiers as training data is less.

In this paper, state of the art Learn^{++}.NSE algorithm is evaluated for handling non-stationary data with our proposed approach. This paper is organized as follows: Section 2 offers an overview of related work. Section 3 presents proposed algorithm i.e. ENSDS_P and Section 4 provides detail of proposed algorithm with its pseudo code. Section 5 provides a rigorous evaluation of the proposed algorithm with one of the existing algorithms. Section 6 concludes the paper.

II. Related Work

The first experiment of ensembles in data streams was the one proposed by Street and Kim with their Streaming Ensemble Algorithm [5] (SEA) where a chunk of d instances is read from the data stream and used to build a classifier. As fixed size of ensemble was used, so they compare new generated classifier against a pool of previously trained classifiers (from previous chunk), and if its current classifier improves the quality of ensemble it is included at the cost of the worst classifier. SEA uses a simple majority vote and may not be able to perform in recurring environments.

Wang et al. proposed Accuracy Weighted Ensemble [6] (AWE) of classifiers on each incoming data chunk and use that chunk to evaluate the performance of all existing classifiers in the ensemble. The weight of each classifier is the difference of error rate of a random classifier and the mean square error of the classifier for the current chunk. The mean square errors of old classifiers are high, and thus the weights of old classifiers are small.

Brzezinski and Stefanowski proposed the Accuracy Updated Ensemble [7] (AUE) which is derived from AWE.It uses the same principles of chunk-based ensembles, but with incremental base components/classifiers. It not only builds new classifiers, but also conditionally updates existing classifiers on new chunks rather than just adjusting their weights. The updation of existing base classifiers makes AUE better than AWE in case of gradual drift but conditionally updating of base classifiers is less accurate for sudden drift.

Robi Polikar et al. proposed Learn^{++}.NSE [8], [9], [10], [11], [12] (Nonstationary Environment) which generates classifiers sequentially using batches of examples/instances (Not true online learner as it converts the online data stream into a series of chunks of a fixed size). At each time step, one new classifier is trained on recent distribution, using an instance weighting distribution. In Learn^{++}.NSE each classifier's weight is computed using a weighted average of its prediction error on old and current batch and finally uses weighted majority voting to obtain ensemble's output.

Most recently, Brzezinski and Stefanowski proposed AUE2 [13] introduces a new weighting function, does not require cross-validation on the existing classifiers, does not keep a classifier buffer, prunes its base learners, and always unconditionally updates its components. Classifiers are updated after every chunk, so they can react to gradual drifts. It can react to sudden drifts and gradual drifts but not for reoccurring concepts. Compared to Learn^{++}.NSE, AUE2 incrementally trains existing component classifiers, retains only k of all the created components, and uses a different weighting mechanism which ensures that components will have non-zero weights.

III. Proposed Algorithm

Fig. 2 depicts the flow diagram of ensemble for non-stationary data stream with propagation (ENSDS_P). ENSDS_P is our proposed algorithm, which is an ensemble of classifiers, where the classifiers are generated from data arrived at time t and evaluated on recent data. All generated classifiers are combined by using weighted majority voting to provide the predictions of unseen data.

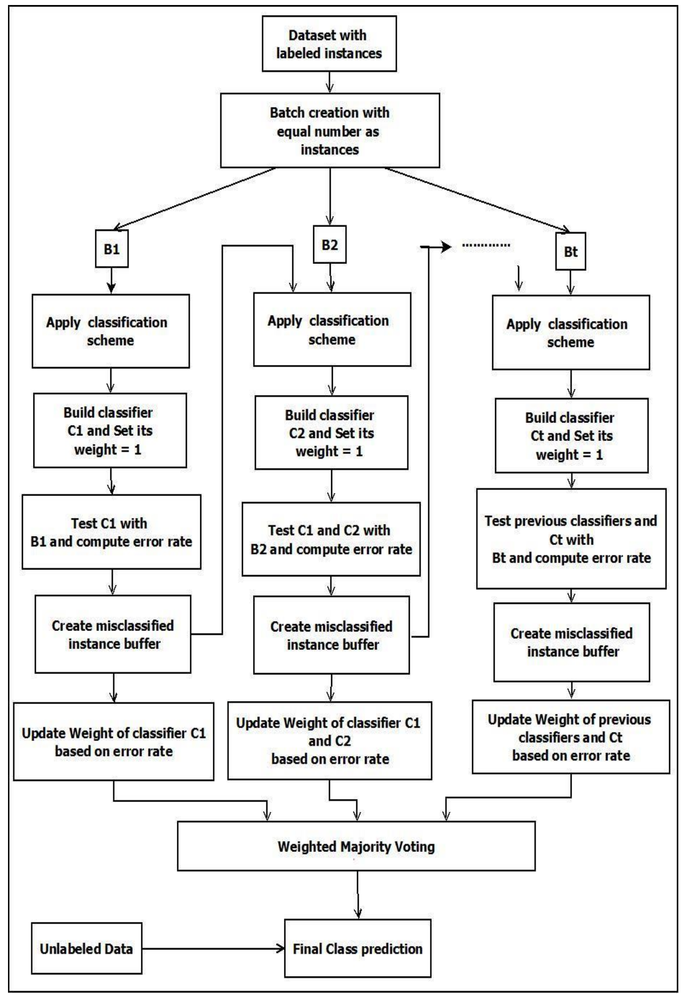


Fig. 2. Flow Diagram of ENSDS_P

One of the major differences in ENSDS_P as compared to existing approaches is, we are not updating a set of weights for each instance rather we believe all instances are equally important while they are using in training so uniform weight is considered and secondly we are propagating the misclassified instances of a classifier to subsequent classifier for improving the performance.

In this system, data is continuously arriving in non-stationary manner. For learning purpose, we take dataset containing labeled instances. Divide this incoming data into number of batches where each batch contains equal number of instances. First apply any suitable classification scheme to create a classifier. The performance of classifier is then evaluated with same batch of instances. If the error rate of classifier is more than 50% i.e. half of the predictions are wrong then delete that recently generated classifier and again repeat the classification process till we get a classifier having an error rate less than 50%.

After creation of first classifier, a misclassified instance buffer is used to store hard to classify instances. All hard to classify instances are propagated to next classifier so that next subsequent classifier can learn them with their training chunk and overall system performance can be improved.

When the next batch of data get available, the incorrect classified instances of previous classifier would be combined with labeled instances of current batch and then apply classification scheme. From this step we get next classifier. This process is continued till we get classifier for all batches and all these classifiers are combined using weighted majority voting scheme. When unlabeled data is arrived, it is predicted by created ensemble using weighted majority voting.

Two variations of ENSDS_P are developed and analyzed, first approach named as ENSDS_P_F where we are propagating misclassified instances, but preserving fix batch size for all classifiers while other approach named as ENSDS_P_D where we are propagating misclassified instances and dynamic chunk size is used for training of classifiers.

IV. ALGORITHMIC DESCRIPTION

Fig. 3 presents the pseudo code of proposed algorithm. For each t, a new classifier generates on current training chunk D^t, and the performance of all previously generated classifiers would be evaluated on current data chunk by ε_k^t parameter and misclassified instances would be saved in a buffer M_k^t .The misclassified instances will be propagated to next subsequent classifier with their training chunk.

In step 1, a uniform weight is assigned to all instances of current data chunk .Step 2 is only different in ENSDS_P_D and ENSDS_P_F rest of algorithm will remain same for both. ENSDS_P_D achieves D^t by eq. 1 where total no. of instances in current data chunk would be union of D^t and misclassified instances of previous data chunk.

$$D^t = D^t \cup M_k^{t-1} \tag{1}$$

ENSDS_P_F achieves D^t by eq. 2 and 3 where ND^t represent new dataset whose size equal to size of current data chunk minus size of misclassified instance buffer.

$$Size(ND^t) = Size(D^t) - Size(M_k^{t-1}) \tag{2}$$

$$D^t = ND^t \cup M_k^{t-1} \tag{3}$$

After formation of k^{th} classifier as in step 3, the performance of existing classifiers will be evaluated over the current training dataset D^t and we will get ε_k^t which is error of k^{th} classifier on current D^t. If error generated by current classifier is more than .5 that is half of the predictions are wrong then generate a new classifier for the current distribution. If error generated by one of theprevious classifier is more than .5 then set its $\varepsilon_k^t = 0.5$ as in step 4.We are not normalizing the ε_k^t as its value remains between 0 to 0.5 and voting power of a classifier having $\varepsilon_k^t = .5$ will remain low.

Algorithm: ENSDS_P
Input: For each dataset D^t where t=1, 2, ….
Training Data: $\{X_i^t \in X;\ y_i^t \in Y = \{1, \dots, c\}\}$, i=1,…, m instances
Description: Supervised learning algorithm to handle non-stationary data stream
Do for t = 1, 2, ...

1. Initialize $D^t(i) = 1/m, \forall i$,
2. If t = 1 then
   Goto step 3
   Else
   Refer Eq. 1 for ENSDS_P_D
   or
   Refer Eq. 2,3 for ENSDS_P_F
3. Call base classifier with D^t, obtain $h_k: X \rightarrow Y$ where k = 1, 2, ...t
4. Evaluate all existing classifiers (h_k) on D^t
   $$\varepsilon_k^t = \sum_{i=1}^{m} D^t(i).\ [|h_k(X_i) \neq y_i|]$$
   If $\varepsilon_{k=t}^t > \frac{1}{2}$ generate a new h_k
   If $\varepsilon_{k<t}^t > \frac{1}{2}$ Set $\varepsilon_k^t = \frac{1}{2}$
5. $M_k^t = \forall$i where $[|h_k(x_i) \neq y_i|]$ is true
6. Compute the weight for kth classifierh_k
   $$Sig_k^t = \frac{1}{1 + e^{-a(t-k-b)}}$$
   $$w_k^t = \begin{cases} 1 & t = k \\ \dfrac{Sig_k^t}{Sig_k^t + \sum_{j=1}^{t-1} w_k^{t-j}}, & \text{otherwise} \end{cases}$$
7. Calculate classifier voting weights
   Voting $w_k^t = \ln\left(\frac{1}{\sum_{j=1}^{t} w_k^j \varepsilon_k^j}\right)$ for k = 1, …, t
8. Obtain the final hypothesis
   $$H^t(X_i) = arg\ \max_c \sum_k w_{k*}^t\ [|h_k(X_i) = c|]$$

Fig. 3. The Pseudo code of the algorithm ENSDS_P

In step 5, we are creating M_k^t parameter which represents a buffer to hold misclassified instances. These misclassified instances would be propagated to next classifier before its formation with its training chunk. A nonlinear sigmoid function is used to set weight of a classifier. Because of this, if a classifier will be evaluated more than once then its sigmoid weight will get increased.

The weight to a classifier is assigned based on its performance on previous distributions as well as on recent distribution so weighted average of classifier is computed in step 6. When a classifier is generated it's $w_k^t = 1$, after its evaluation on recent environment its w_k^t gets keep updated. If a classifier does not performs well on recent environment, then its weightederror $(w_k^t.\varepsilon_k^t)$ will gets increased. In step 7 the weight error average is computed to determine the voting weight of classifiers. The voting power of each classifier is computed using logarithm of the inverse of its weighted error average. If weighted error average is high a classifier will get less power of voting.

The time complexity of ENSDS_P is O(t*k*O(x*m)+k*t*m) where O(x*m) is the time complexity of Naïve Bayes classifier, x is number of features and m is number of instances in training set, k indicates number of classifiers, t indicates number of data chunks to be predicted.

V. Comparative Evaluation and Analysis

In the following subsections; we describe the tested datasets, experimental setup, and comparative analysis of experimental results.

A. Datasets

For doing the comparison of ENSDS_P and existing algorithm (Learn^{++}.NSE) we are using different datasets with different batch sizes. The proposed algorithm is tested over real time datasets.

1) IBM_EOD_Direction: The IBM_EOD_Direction dataset contains stock data of IBM Company, where we are considering open, high, low, close, volume and rate of change in closing price to find out the stock index movement (Up, Down) for classification task. For training purpose data from period 2-Jan-2000 to 13-April-2016 (3999 examples) is fetched and for testing purpose data from period 2-Jan-2001 to 13-April-2016 (3841 examples) is fetched using Google finance.

2) IBM_EOD_Trading: The IBM_EOD_ Trading dataset contains stock data of IBM Company, where we are considering open, high, low, close, volume and rate of change in closing price to find out Buy or Sell class for Stock data. For training purpose data from period 2-Jan-2000 to 13-April-2016 (3999 examples) is fetched and for testing purpose data from period 2-Jan-2001 to 13-April-2016 (3841 examples) is fetched using Google finance.

The purpose of considering stock data is as we know that stock market data is high frequency data which is complex, non-stationary, chaotic and non-linear and suites our research topic .Concept drift can occurs in the stock market for a number of reasons for example, traders preference for stocks change over time, increases in a stock's value may be followed by decreases. Stock market data can possess sudden, gradual and recurring drift at any moment of time.

The analysis over IBM_EOD_Direction dataset would help trader to know the position of stock market index at next moment of time and analysis over IBM_EOD_Trading would help trader to take decision whether its right time to sell or purchase the stock.

B. Experimental Setup

For experiment analysis, the proposed algorithm is implemented in Java using MOA and WEKA framework. The source code of Learn^{++}.NSE is obtained from MOA extensions for comparison purpose. The experiments were conducted on a machine equipped with Processor Intel(R) Core(TM) i3-2120 CPU @ 3.30GHz, 2 Core(s), 4 Logical Processor(s) and 4 GB of RAM. Here we have used different batch size for comparison purpose. However, the optimal batch size is different for each stream. For rigorous evaluation, we are considering different evaluation measures [14] like P=precision, R=recall, A=accuracy, F-M=f-measure, and G-M=g-mean.

C. Results

Table 1 depicts the performance of Learn^{++}.NSE and both the versions of ENSDS_P respectively to classify the stock index movement(Up, Down) over IBM_EOD dataset where we are considering Naïve Bayes as base classifiers, different batch size and no pruning strategy is used.

TABLE I. Comparison Over IBM_EOD_Direction Dataset

Batch Size	Algorithms	P	R	A	F-M	G-M
500	Learn++.NSE	58.77	94.13	91.07	72.36	92.37
	ENSDS_P_D	87.83	84.94	94.48	86.36	90.75
	ENSDS_P_F	88.61	76.84	92.42	82.31	86.36
400	Learn++.NSE	57.33	98.21	91.30	72.40	94.22
	ENSDS_P_D	79.32	94.98	95.05	86.45	95.03
	ENSDS_P_F	82.20	86.86	93.99	84.47	91.14
300	Learn++.NSE	29.84	87.02	85.16	44.44	86.02
	ENSDS_P_D	88.61	56.32	84.07	68.87	73.80
	ENSDS_P_F	91.23	63.71	87.92	75.03	78.84

It is clear from fig. 4 that for each batch size we are retrieving high true positives and low false positives hence precision is higher. The results shows precision of both the versions of ENSDS_P is significantly high and recall is approximately equal.

Generally, there always remains a tradeoff between precision and recall. F-measure is appropriate evaluation measure which gives the balance between precision and recall. As compare to Learn^{++}.NSE we are able to maintain a good balance between precision and recall so proposed algorithm can also be used with imbalanced data. The values of evaluation measures proved the validity of proposed work hence evaluation results shows that proposed algorithms effectively provides incremental learning over high frequency stock market data.

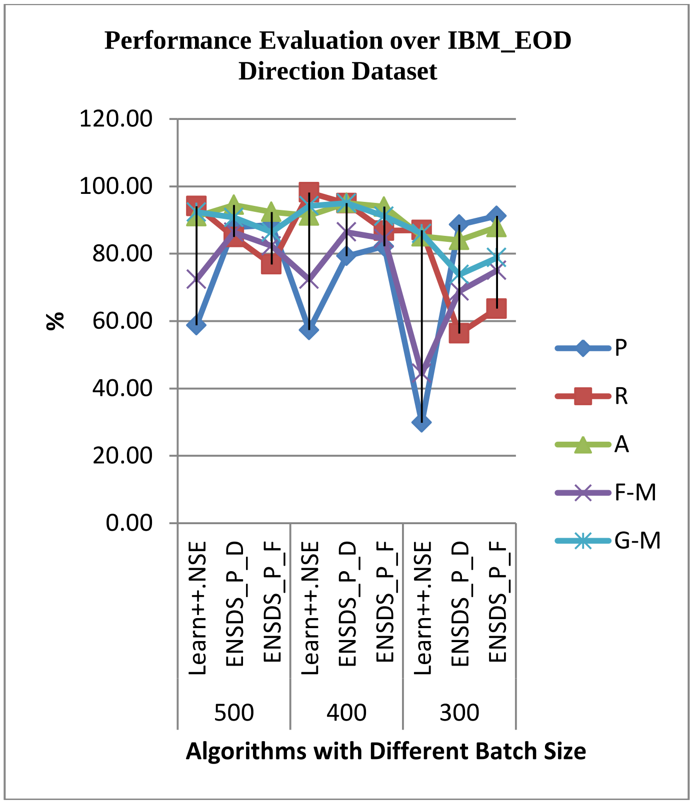


Fig. 4. Performance Analysis of Learn++.NSE and ENSDS_P over IBM_EOD_Direction dataset

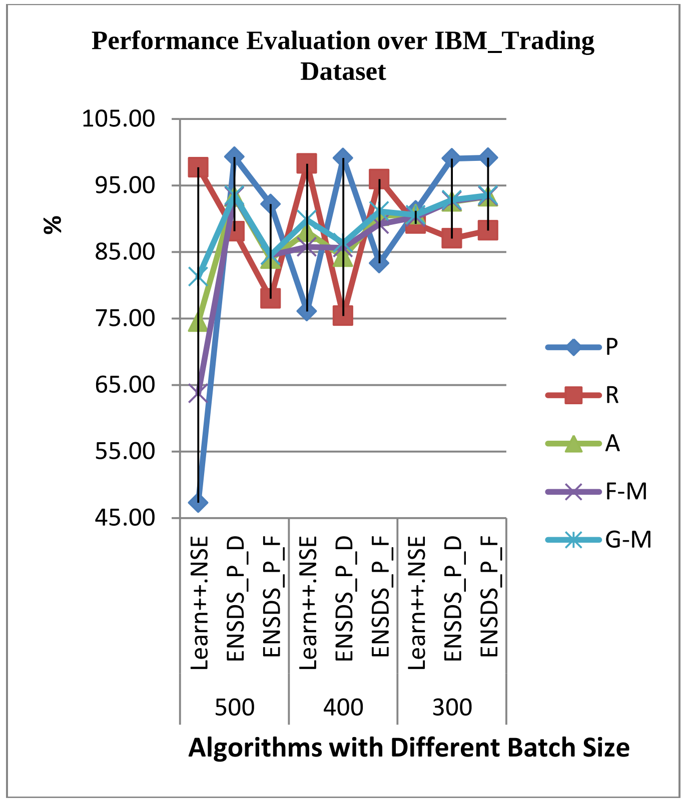


Fig. 5. Performance Analysis of Learn^{++}.NSE and ENSDS_P over IBM_EOD_Trading Stock dataset

Table 2 depicts the performance of Learn^{++}.NSE and versions of ENSDS_P respectively to classify the stock trading (Buy, Sell) over IBM_EOD_Trading dataset where we are considering Naïve Bayes as base classifiers, different batch size and no pruning strategy is used.

TABLE II. COMPARISON OVER IBM_EOD_TRADING DATASET

Batch Size	Algorithms	P	R	A	F-M	G-M
500	Learn++.NSE	47.28	97.73	74.49	63.73	81.26
	ENSDS_P_D	99.29	88.11	93.31	93.36	93.52
	ENSDS_P_F	92.20	77.98	83.96	84.50	84.51
400	Learn++.NSE	76.06	98.30	88.02	85.76	89.82
	ENSDS_P_D	99.12	75.40	84.25	85.65	86.35
	ENSDS_P_F	83.31	95.95	90.42	89.18	91.13
300	Learn++.NSE	91.21	89.25	90.63	90.22	90.58
	ENSDS_P_D	99.07	87.02	92.55	92.66	92.84
	ENSDS_P_F	99.18	88.23	93.34	93.38	93.54

Fig. 5 represents that as compared to Learn^{++}.NSE we have achieved high precision, accuracy, f-measure and g-mean for all batches on IBM_EOD_Trading dataset.

After testing proposed algorithm on different datasets, evaluation measures confirm the validity and excellence of proposed algorithm. The non-stationary data can have class imbalanced problem so result can be biased toward the majority class; thus the classifier tends to misclassify the minority class instances. In imbalanced application area, proposed algorithm can be used and can provide a balance between majority and minority instances.

VI. CONCLUSION

From the implementation and analysis of ENSDS_P we can conclude that the performance of ENSDS_P is better as compared to Learn^{++}.NSE on different datasets. Evaluation measures also confirm the validity of proposed algorithm's scores. The selection of optimal batch size varies from dataset to datasets. The non-stationary data can have class imbalanced problem so result can be biased toward the majority class; thus the classifier tends to misclassify the minority class instances. If dataset is highly imbalanced then there is need to add some balancing mechanism in proposed algorithm to achieve high performance.

REFERENCES

[1] Moreno-Torres, J., Raeder, T., Alaiz-Rodríguez, R., Chawla, N.V.,Herrera, F., "A unifying view on dataset shift in classification", Pattern Recogniion, 45, 521–530 , 2011.

[2] R. Polikar , "Ensemble based systems in decision making", IEEE Circuits and Systems Magazine, Vol. 6, No. 3 ,pp. 21-45, 2006

[3] Meenakshi A.Thalor ,Dr.S.T.Patil, "Review of ensemble based classification algorithms for nonstationary and imbalanced data" ,IOSR Journal of Computer Engineering,e-ISSN: 2278-0661, Vol. 16, pp. 103-107, Feb 2014.

[4] Read, J., Bifet, A., Pfahringer, B. & Holmes, G. "Batch-incremental versus instance-incremental learning in dynamic and evolving data ", IDA 2012, pp. 313-323, Helsinki, Finland, October 25-27 2012

[5] W. N. Street and Y. Kim, "A streaming ensemble algorithm (SEA) for large-scale classification," Intellegent Conference on Knowledge Discovery & Data Mining, pp. 377-382, 2001.

[6] H. Wang, W. Fan, P. Yu, and J. Han, "Mining concept-drifting data streams using ensemble classifiers," in Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Min., pp. 226–235, 2003.

[7] Dariusz Brzezinski and Jerzy Stefanowski , "Accuracy updated ensemble for data streams with concept drift,"Proceedings of the 6th international conference on Hybrid artificial intelligent systems - Volume Part II,2011,pp. 155-163.

[8] Elwell R. and Polikar R., "Incremental learning of concept drift in non-stationary environments, "IEEE Trans. on Neural Networks, vol. 22, 2011,pp. 1517-1531.

[9] R. Elwell and R. Polikar, "Incremental learning of variable rate concept drift, "International Workshop on Multiple Classifier Systems (MCS 2009) in Lecture Notes in Computer Science, vol. 5519, pp. 142-151, 2009.

[10] M. Karnick, M. Ahiskali, M. Muhlbaier, and R. Polikar, "Learning concept drift in nonstationary environments using an ensemble of classifiers based approach, "International Joint Conerence.on Neural Network,2008, pp. 3455-3462.

[11] M. Muhlbaier and R. Polikar, "An ensemble approach for incremental learning in nonstationary environments, "Multiple Classifier Systems, pp. 490-500, 2007.

[12] Michael D. Muhlbaier and RobiPolikar, "Multiple classifiers based incremental learning algorithm for learning in nonstationary environments", Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, vol. 6, 2007,pp. 3618–3623.

[13] Brzezinski, D.; Stefanowski, J., "Reacting to different types of concept drift: the accuracy updated ensemble algorithm, " , IEEE Transactions on Neural Networks and Learning Systems Vol. 25(1),2014, 81-94.

[14] Jesse Davis, Mark Goadrich , "The Relationship Between Precision-Recall and ROC Curves," In Proceedings of the 23rd international conference on Machine learning , pp. 233-240,2006.

Multi-Objective Task Scheduling in Cloud Computing Using an Imperialist Competitive Algorithm

Majid Habibi
Department of Computer Engineering,
Tabriz Branch, Islamic Azad University,
Tabriz, Iran

Nima Jafari Navimipour
Department of Computer Engineering,
Tabriz Branch, Islamic Azad University,
Tabriz, Iran

***Abstract*—Cloud computing is being welcomed as a new basis to manage and provide services on the internet. One of the reasons for increased efficiency of this environment is the appropriate structure of the tasks scheduler. Since the tasks scheduling in the cloud computing environment and distributed systems is an NP-hard problem, in most cases to optimize the scheduling issues, the meta-heuristic methods inspired by nature are used rather than traditional or greedy methods. One of the most powerful meta-heuristic methods of optimization in the complex problems is an Imperialist Competitive Algorithm (ICA). Thus, in this paper, a meta-heuristic method based on ICA is provided to optimize the scheduling issue in the cloud environment. Simulation results in MATLAB environment show the amount of 0.7 percent improvement in execution time compared with a Genetic Algorithm(GA).**

Keywords—Cloud Computing; Tasks scheduling; Imperialist Competitive Algorithm

I. Introduction

Today, data centers are composed of thousands of computers that are distributed worldwide and users use the computers frequently to send e-mail, read, write, search, etc. Users perform the mentioned operations using the available browser and as a service receiver attached to the servant. Also, cloud computing [1, 2, 3, 4, 5] combines parallel concepts and computing systems to provide for computers and other machines shared resources, hardware, software and information [6, 7, 8, 9]. In the cloud, users use the provided services according to their needs. The National Institute of Standards that is the Information Technology Laboratory, provides a comprehensive definition of cloud computing: Cloud computing is a pay-per-use model for enabling access, and easy and on-demand network access to a shared resource of computing resources configurable as networks, servers, saving resources, programs, and services that can be prepared with a high-speed and be published with little effort of management or interaction of servant. There are many challenges in a cloud environment. However, task scheduling is an important challenge that has continued to remain a challenge despite many efforts made recently in the field . The services provided in cloud computing are grouped into 4 categories, including SaaS (Software as a Service) [10, 11, 12, 13, 14], IaaS (Infrastructures as a Service) [15, 16, 17, 18, 19], PaaS (Platforms as a Service) [20, 21, 22, 23, 24, 25] and EaaS (Expert as a Service) [8, 9, 26, 27, 28, 29].

In distributed systems, task scheduling should dedicate subtasks to resources that enhance system performance and that is the scheduling methods which determine the performance order of these tasks [30]. As the allocation of processor and resource to tasks is a complex issue in heterogeneous distributed systems such as cloud environment, many methods and algorithms have been provided to reduce time complexity and simultaneous function of subtasks. Some of the available challenges in the field of scheduling tasks in heterogeneous distributing systems include heterogeneous resources, total running time, runtime and productivity convergence speed in meta-heuristic methods and efficiency of scheduling method. According to the importance of this issue, in this paper, research has been conducted in the field of scheduling computing tasks on cloud computing and heterogeneous systems.

In the second section of this paper, a review and background of research have been provided. In section 3, we define the proposed method. In section 4, the results of the simulation and comparing of the proposed algorithm with other algorithms have been discussed and assessed. Finally, in the last section the conclusion and future work are presented.

II. Related Work

In this section, the function of a number of related procedures in the field of scheduling tasks on cloud systems has been studied and discussed. Sung-Soo Kim, *et al.* [31] in 2014, using optimization algorithm based on biogeography (BBO) which is a new method of meta-heuristic algorithms, optimized task scheduling in cloud computing. Several criteria have been used to assess the task scheduling algorithms and the runtime and the task circle duration have been considered as important criteria which are the main aim of this algorithm. The advantage of this algorithm is to optimize the duration of performance of functions.

Two task scheduling methods named Heterogeneous Earliest Finish Time (HEFT) and Critical-Path-On-a-Processor (CPOP) based on the list have been presented with limited processors to achieve high efficiency and fast timing [32]. HEFT method in every procedure selects a task that has an upward rank and maps via using the method according to processors rank that reduces the earliest start time. However,in CPOP method to indicate priority the sum of the values of upward rank and downward rank of tasks is used, as well as these two algorithms are different in choosing the processor of tasks, so that CPOP algorithm assigns tasks on the critical path to the processors that minimize the makespan of all tasks in critical path.

In [32], a scheduling algorithm called Longest Dynamic Critical Path (LDPC) is proposed with a high-efficiency. Provided algorithm is used on distributed computing systems with a limited number of processors. LDCP scheduling method is based on a list which uses a new method to select tasks for scheduling in a heterogeneous environment of distributed computing systems which enables this algorithm to schedule tasks with a better quality. In this paper, the proposed algorithm has been compared with two DLS with HEFT scheduling algorithm; the survey results show that the proposed method of mentioned algorithms is better in acceleration and time of scheduling. Also, the efficiency of the proposed algorithm with the increase of communications costs between tasks in the graphic is increased compared to other two algorithms. As a result, this algorithm provides a practical solution for scheduling programs with a parallel power with more communication costs in heterogeneous environments of computing systems.

In [33], a two-step algorithm called Hybrid Heuristic–Genetic Scheduling (H2GS) has been suggested for tasks scheduling in heterogeneous distributed computing systems. The first stage of the algorithm is based on LDCP list for scheduling with high quality and in second stage scheduling obtained from the first stage is injected to the initial population of a GA called Genetic Algorithm for Scheduling (GAS). The function of the second stage leads to a shorter schedule. The efficiency of H2GS is compared with two methods of DLS and HEFT and the results show that the proposed algorithm improves the mentioned methods. In [34], a task scheduling method in heterogeneous computing systems has been presented by employing GAs of multiple priority offers. The main idea of this method that is called Multiple priority memetic algorithm (MPQGA) is the use of simultaneous advantageous of heuristic and evolutionary algorithm and avoidance of their objections. In order to assign priorities to tasks, a GA is used and heuristic method of Earliest Finish Time (EFT) for mapping tasks to the processor is applied. MPQGA algorithm of combining processes has used a mutation and strong fit for tasks scheduling. Experimental and comparative test results conducted on the Real world graphs and random revealed that MPQGA algorithm is optimized two heuristic algorithms and a meta-heuristic method called Basic Genetic Algorithm (BGA).

Verma, *et al.* [35] used an improved GA for tasks scheduling in the cloud computing environment. The initial population in this algorithm unlike the standard GA is not chosen randomly and the results of the two algorithms are used in the production of the initial population.

Ting, *et al.* [36] have provided an optimized algorithm for tasks scheduling in the cloud computing environment. After applying the selecting operators, combination and mutation of simulated GAs are used; in this case, the new generation will be closer to the optimal solution. Meanwhile, in this method the parameter of quality of service is composed of five parameters included completion time, bandwidth, cost, distance, and reliability, which due to the type of task different values are assigned to each of these parameters.

Vey, *et al.* [37] have used GAs to minimize the time to complete tasks in a cloud computing environment. A matrix that is the anticipated time for implementation of each task from any source is used to calculate the completion time. Also, a parameter is used that is the indicator of time which finishes the allocated processor in the previous schedule; with this parameter, the workload of the source will also be used to find the optimal sequence.

III. THE PROPOSED METHOD

One of the weaknesses of the ICA is its rapid convergence to the local optimum points. To improve this problem and increase the scanning ability of algorithm, a process similar to the fixed mixing process which is used in GAs is applied in the policy of absorption. This means that after applying the absorption policies on the empire to increase the seeking range of issue a uniform recombination is used in the production of the new position of colonies. Here, one of the parent recombination operators is the position of the colonizer and the position of colonized. In order to prevent random seek and mutation to inappropriate responses, a condition is used in the body of algorithms that is when the resulted position is better compared to the previous position of each colony, the current position is replaced. Formula (1) shows how this process functions in which Colonyik is the coloniali of empirek and Imperialistk shows colonial empire k. Also, α is a coincidence brother that is composed of zero and one.

$$\mathrm{Colony}_{\mathrm{k}} = a(Colony_{ik}) + (1-a)*(imperialist_k) \quad (1)$$

In Figure 1, an overall stage of the proposed algorithm for pseudo-code has been shown.

```
Input:npop(Population-size),problem-size,ep,α, β, pr
For i=1 to npop do
      C_iposition ← RandomPosition(problem-size)
      If i<=ep then
         EmpriresPopulation ← C_iposition
      Else
         C_w← GetWorstSolution(EmpiresPopulation)
          If Cost(C_iposition) < Cost(C_wiposition) then
                  Replace(EmpiresPopulation,C_i,C_w)
          Else
                C_iempire ← ssignAnEmpire(EmpiresPopulation)
          End
       End
       Populaton ← C_i
  End
EvaluatePopulaton(Population)
EvaluateEmpiresPopulation(EmpiresPopulation,Population)
ImperialisticCompetition(EmpiresPoplution,Population)
EliminiateWeakestEmpire(EmpiresPoplution,Population)
```

End
EvalutePopulation(Population)
BestSol ← GetBestSolution(Population)
Return BestSol

Fig. 1. The proposed algorithm for pseudo-code

IV. Simulation Results

In order to assess the efficiency, the proposed method is compared with GA regarding the time to complete the tasks and the productivity of resources via using MATLAB. In order to assess the efficiency of the proposed algorithm, the data in Table 1 is used and Table 2 shows the parameters of the ICA.

TABLE I. Simulation Parameters

Parameter	Values
Number of work	60~30
Virtual machine	12~6
The time of tasks	6000~1000 (MI)
Tasks size	200~50 (MB)
Bandwidth	500~100 (Mbps)
Processor speed	500~100 (MIPS)

TABLE II. Imperialist Competitive Parameters

Parameter	Values
Maximum frequency	100
Population size	50
Number of empires	10
(The absorption coefficient) β	2
The possibility of revolution	0.1
Rate of revolution	0.05
(Colonies the cost factor) δ	0.1
(Selection pressure) η	10

Due to the nature of meta-heuristic algorithm and random initial positions, each algorithm has been implemented 10 times on average and the obtained average results are considered as a final answer and criteria for comparison. Figure 2 shows the comparison of the proposed method with GA in terms of standard completion time, regardless of the distribution time.

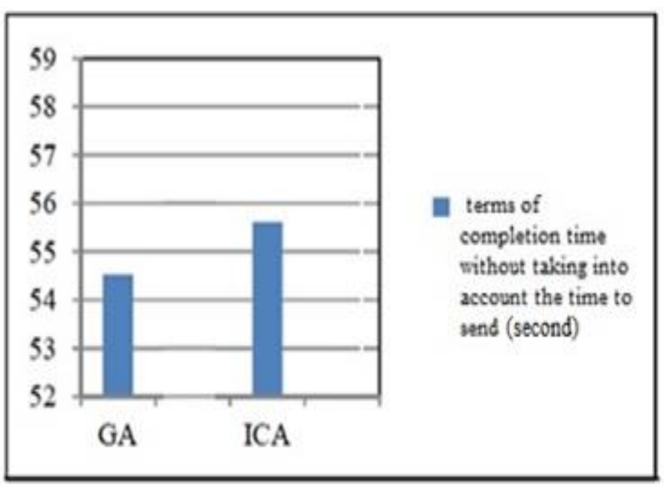


Fig. 2. Chart of methods comparison in terms of completion time without taking into account the time to send

The comparison chart of ICA with a GA is presented in a cloud environment, based on the funding function formula (1) is also shown in Figure 3, Time to send tasks from the scheduler to sources is considered according to bandwidth parameters, in the proposed cost function.

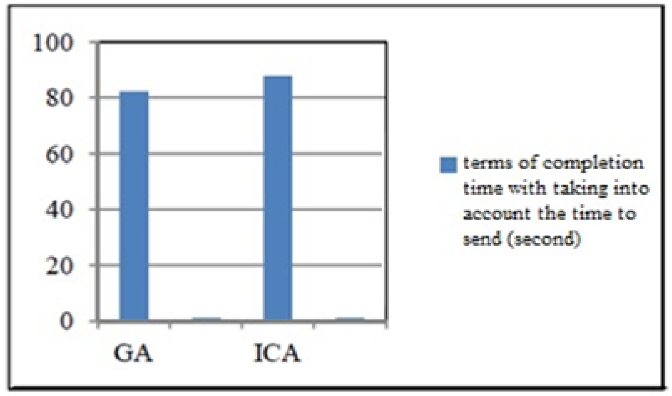


Fig. 3. The chart of comparing algorithms based on the cost function

Figure 4 also shows the efficiency of sources in the ICA comparing to the GA. The results showed an improvement in performance of the proposed algorithm comparing to the other methods proposed for tasks scheduling in the cloud environment such as GAs, particle swarm, and the standard ICA.

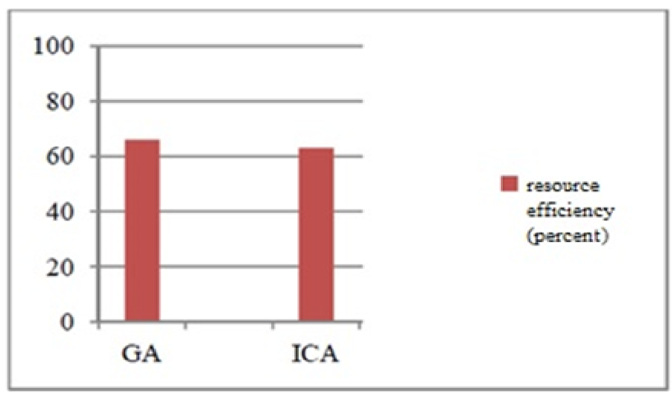


Fig. 4. The chart of comparing algorithms based on resource efficiency

In this section, an improved version of the ICA is proposed for tasks scheduling in a cloud environment. In the most proposed algorithms for mapping tasks to resources in the cloud environment, only the time of performing tasks is noted and features of sources bandwidth and time is not considered as an effective parameter in producing the final answer. In this paper, a cost function aware of bandwidth is used for optimal scheduling and enhanced efficiency of tasks scheduling resources in the cloud computing environment. As well as in order to avoid algorithm to be trapped in the local optimization and enhancing scanning ability of algorithm, after the policy of absorption a process similar to the uniform mixture of GAs has been applied to empires. Results of implementation and comparison with other proposed methods for tasks scheduling in a cloud environment, such as GA, particle swarm, and competition algorithm also represent achieving better responses and suitable colonialism standard tasks scheduling.

V. Conclusion

In most of the algorithms provided for mapping the tasks to resources in cloud environment, only attention is given to the time of performing tasks and bandwidth features of resources and time of sending tasks are not considered as effective parameters in producing the final answer. In this article for optimal timing and increasing resource efficiency, a cost function aware of bandwidth is used for tasks scheduling in cloud computing environment. Also, in order to avoid algorithms to be trapped in local optimization and increasing the scanning ability of algorithms, the same process with uniform recombination available in GAs has been applied to empires after the policy of absorption. The results of issue implementation and its comparison with other proposed methods for tasks scheduling in the cloud environment, such as GAs and standard ICA represents achieving better response and more convenient tasks scheduling. In this paper, improved version of ICA for tasks scheduling in the cloud environment is proposed. For this reason, further studies in tasks scheduling in the cloud computing environment, such as MPQMA meta-heuristic methods is used to improve tasks scheduling.

References

[1] M. C. Chu-Carroll, "Code in the Cloud," Programming Google App Engine, Pragmatic Bookshelf, 306 pages, 2011.

[2] S. Asghari and N. J. Navimipour, "Review and Comparison of Meta-Heuristic Algorithms for Service Composition in Cloud Computing", Majlesi Journal of Multimedia Processing, Vol. 4, No. 4, 2016.

[3] M. Chiregi, N. Jafari Navimipour, "Trusted services identification in the cloud environment using the topological metrics", Karbala International Journal of Modern Science, 2016. (in press)

[4] M. Chiregi, N. J. Navimipour, "A new method for trust and reputation evaluation in the cloud environments using the recommendations of opinion leaders' entities and removing the effect of troll entities", Computers in Human Behavior, Vol. 60, pp. 280-292, 2016.

[5] B. Keshanchi, N. Jafari Navimipour, "Priority-based task scheduling in the cloud systems using a memetic algorithm", Journal of Circuits, Systems and Computers, 2016. (in press)

[6] B. A. Milani and N. J. Navimipour, "A comprehensive review of the data replication techniques in the cloud environments: Major trends and future directions", Journal of Network and Computer Applications, Vol. 64, pp. 229-238, 2016.

[7] N. J. Navimipour, "Task scheduling in the Cloud Environments based on an Artificial Bee Colony Algorithm", International Conference on Image Processing, Production and Computer Science (ICIPCS'2015), Istanbul,Turkey, pp. 38-44, 2015.

[8] N. J. Navimipour, F. S. Milani, "Task scheduling in the cloud computing based on the cuckoo search algorithm", International Journal of Modeling and Optimization, Vol. 5, No. 1, p. 44-47, 2015.

[9] N. J. Navimipour, A. M. Rahmani, A. H. Navin, M. Hosseinzadeh, "Expert Cloud: A Cloud-based framework to share the knowledge and skills of human resources", Computers in Human Behavior, Vol. 46, pp. 57-74, 2015.

[10] J. Lin, D. Fu, J. Zhu, "What is Cloud Computing?", IT as a Service, Vol. 11, No. 2, pp. 10-13, 2009.

[11] P. Buxmann, T. Hess, S. Lehmann, "Software as a Service", Wirtschaftsinformatik, Vol. 50, pp. 500-503, 2008.

[12] V. Choudhary, "Software as a service: Implications for investment in software development",40th Annual Hawaii International Conference on in System Sciences(HICSS 2007), pp. 209a-209a, 2007.

[13] M. Almorsy, J. Grundy, A. S. Ibrahim, "Adaptable, model-driven security engineering for SaaS cloud-based applications", Automated Software Engineering, Vol. 21, pp. 187-224, 2014.

[14] Z. Zeng, B. Veeravalli, "Optimal metadata replications and request balancing strategy on cloud data centers", Journal of Parallel and Distributed Computing, Vol. 74, pp. 2934-2940, 2014.

[15] S. Bhardwaj, L. Jain, S. Jain, "Cloud computing: A study of infrastructure as a service (IAAS)", International Journal of engineering and information Technology, Vol. 2, pp. 60-63, 2010.

[16] A. Khajeh-Hosseini, D. Greenwood, I. Sommerville, "Cloud migration: A case study of migrating an enterprise it system to iaas", 3rd International Conference on in Cloud Computing (CLOUD), pp. 450-457, 2010.

[17] A. Nathani, S. Chaudhary, G. Somani, "Policy based resource allocation in IaaS cloud", Future Generation Computer Systems, Vol. 28, pp. 94-103, 2012.

[18] A. Iosup, R. Prodan, D. Epema, "Iaas cloud benchmarking: approaches, challenges, and experience", in Cloud Computing for Data-Intensive Applications, ed: Springer, pp. 83-104, 2014.

[19] W. Wang, B. Liang, B. Li, "Revenue maximization with dynamic auctions in iaas cloud markets", IEEE/ACM 21st International Symposium on in Quality of Service (IWQoS), pp. 1-6, 2013.

[20] P. Mell and T. Grance, "Draft NIST working definition of cloud computing", Referenced on June. 3rd, Vol. 15, 2009.

[21] M. Miller, J. Lei, "Cloud computing", M. Beijing: Machine Industry Publication, Vol. 4, 2009.

[22] S. Eludiora, O. Abiona, A. Oluwatope, A. Oluwaranti, C. Onime, L. Kehinde, "A user identity management protocol for cloud computing paradigm," Int'l J. of Communications, Network and System Sciences, Vol. 4, pp. 152-163, 2011.

[23] H. Dinesha and V. Agrawal, "Multi-level authentication technique for accessing cloud services", International Conference on, in Computing, Communication and Applications (ICCCA), pp. 1-4, 2012.

[24] D. Zeginis, F. D'Andria, S. Bocconi, J. Gorronogoitia Cruz, O. Collell Martin, P. Gouvas, et al., "A user-centric multi-PaaS application management solution for hybrid multi-cloud scenarios", Scalable Computing: Practice and Experience, Vol. 14, 2013.

[25] M. Sellami, S. Yangui, M. Mohamed, S. Tata, "PaaS-independent Provisioning and Management of Applications in the Cloud," Sixth International Conference on in Cloud Computing (CLOUD), IEEE, pp. 693-700, 2013.

[26] N. J. Navimipour, A. H. Navin, A. M. Rahmani, M. Hosseinzadeh, "Behavioral modeling and automated verification of a Cloud-based framework to share the knowledge and skills of human resources", Computers in Industry,Vol.68, pp. 65-77, 2015.

[27] N. J. Navimipour, "A formal approach for the specification and verification of a trustworthy human resource discovery mechanism in the Expert Cloud", Expert Systems with Applications, Vol.42, pp. 6112–6131, 2015.

[28] M. Oussalah, D. Professor Ali Hessami, N. Jafari Navimipour, A. Masoud Rahmani, A. Habibizad Navin, M. Hosseinzadeh, "Job scheduling in the Expert Cloud based on genetic algorithms", Kybernetes, Vol. 43, pp. 1262-1275, 2014.

[29] M. Ashouraie, N. Jafari Navimipour, M. Ramage, and P. Wong, "Priority-based task scheduling on heterogeneous resources in the Expert Cloud", Kybernetes, Vol. 44, 2015.

[30] M. A. Khan, "Scheduling for heterogeneous systems using constrained critical paths", Parallel Computing, Vol. 38, pp. 175-193, 2012.

[31] S. S. Kim, J.-H. Byeon, H. Yu, H. Liu, "Biogeography-based optimization for optimal job scheduling in cloud computing," Applied Mathematics and Computation, Vol. 247, pp. 266-280, 2014.

[32] M. I. Daoud, N. Kharma, "A high performance algorithm for static task scheduling in heterogeneous distributed computing systems", Journal of Parallel and distributed computing, Vol. 68, pp. 399-409, 2008.

[33] M. I. Daoud, N. Kharma, "A hybrid heuristic–genetic algorithm for task scheduling in heterogeneous processor networks," Journal of Parallel and Distributed Computing, Vol. 71, pp. 1518-1531, 2011.

[34] Y. Xu, K. Li, T. T. Khac, M. Qiu, "A multiple priority queueing genetic algorithm for task scheduling on heterogeneous computing systems", 9th International Conference on in High Performance Computing and

Communication, 14th International Conference on Embedded Software and Systems (HPCC-ICESS), IEEE, Liverpool, pp. 639-646, 2012.

[35] S. Kaur, A. Verma, "An efficient approach to genetic algorithm for task scheduling in cloud computing environment", International Journal of Information Technology and Computer Science (IJITCS), Vol. 4, pp. 74-79, 2012.

[36] G. Guo-ning, H. Ting-lei, G. Shuai, "Genetic simulated annealing algorithm for task scheduling based on cloud computing environment", International Conference on Intelligent Computing and Integrated Systems, Guilin, pp. 60-63, 2010.

Big Data Classification Using the SVM Classifiers with the Modified Particle Swarm Optimization and the SVM Ensembles

Liliya Demidova
Moscow Technological Institute
Ryazan State Radio Engineering University
Moscow, Russia

Evgeny Nikulchev
Moscow Technological Institute
Moscow,
Russia

Yulia Sokolova
Ryazan State Radio Engineering University
Ryazan, Russia

***Abstract*—The problem with development of the support vector machine (SVM) classifiers using modified particle swarm optimization (PSO) algorithm and their ensembles has been considered. Solving this problem would allow fulfilling the high-precision data classification, especially Big Data classification, with the acceptable time expenditures. The modified PSO algorithm conducts a simultaneous search of the type of kernel functions, the parameters of the kernel function and the value of the regularization parameter for the SVM classifier. The idea of particles' «regeneration» served as the basis for the modified PSO algorithm. In the implementation of this algorithm, some particles change the type of their kernel function to the one which corresponds to the particle with the best value of the classification accuracy. The offered PSO algorithm allows reducing the time expenditures for the developed SVM classifiers, which is very important for Big Data classification problem. In most cases such SVM classifier provides the high quality of data classification. In exceptional cases the SVM ensembles based on the decorrelation maximization algorithm for the different strategies of the decision-making on the data classification and the majority vote rule can be used. Also, the two-level SVM classifier has been offered. This classifier works as the group of the SVM classifiers at the first level and as the SVM classifier on the base of the modified PSO algorithm at the second level. The results of experimental studies confirm the efficiency of the offered approaches for Big Data classification.**

Keywords—Big Data; classification; ensemble; SVM classifier; kernel function type; kernel function parameters; particle swarm optimization algorithm; regularization parameter; support vectors

I. Introduction

Big Data is a term for data sets that are so large and/or complex that traditional data processing technologies are inadequate. They require technologies that can be used to store and process the exponentially increasing data sets which contain structured, semi structured and unstructured data. Volume, variety and velocity are three defining characteristics of Big Data. Volume refers to the huge amount of data, variety refers to the number of data types and velocity refers to the speed of data processing. The problems of the Big Data management result from the expansion of all three characteristics. The Big Data does not consist of only numbers and strings but also geospatial data, audio, video, web data, social files, etc. obtained from various sources such as sensors, mobile phones, cameras and so on.

The main purpose of the Big Data technologies is to provide the high quality of data processing and data analysis. Nowadays the Big Data technologies have been applied in many fields of science and engineering, including physical, biological and biomedical sciences. Also, they have been used in government agencies, financial corporations, large enterprises, etc.

The high volume of storage space, in particular, the cloud storage is needed to manage and reuse Big Data which can be useful for many purposes, for example, for hardware and software maintenances. It is therefore necessary to perform the analytical, retrieval and process operations, which are very complex and time consuming ones. In order to overcome these difficulties new Big Data technologies have been getting a lot of attention over the last few years. The Big Data processing improves the transfer speed of the data sets in comparison to the speed of the simple data exchanges. The Big Data mining tools are very useful to the end users when they solve their own actual problems.

Currently many efficient approaches must be implemented when dealing with the Big Data. In particular, the feature selection, clustering and classification plays an important role in the Big Data analysis, when it is necessary to retrieve, search or classify a data, using the Big Data sets. These approaches are useful for such spheres as pattern recognition, machine learning, bio-informatics, data mining, semantic ontology and so on. As there are many algorithms available for the feature selection, clustering and classification, it is necessary to find the appropriate algorithms which must be chosen properly for the problem of the Big Data analysis.

The machine learning algorithms can be considered along a spectrum of the supervised and unsupervised learning algorithms. In the strictly unsupervised learning, the problem is to find the structure such as clusters in the unlabeled data set. The supervised learning uses the training set of the classified data to construct classifier, which can be used to classify new data. In both cases, the Big Data applications demonstrate the growing number of features and the growing volume of the input data.

The Support Vector Machine (SVM) algorithm is the supervised machine learning algorithm. Currently, the SVM algorithm (one of the boundary classification algorithms [1, 2]) is used for different classification problems in various applications with great success.

The SVM classifiers based on the SVM algorithm have been applied for credit risk analysis [3], medical diagnostics [4], handwritten character recognition [5], text categorization [6], information extraction [7], pedestrian detection [8], face detection [9], Earth remote sensing [10], etc.

SVM classifier uses special kernel function to construct a hyperplane separating the classes of data. An example of the separating hyperplane in the 2D space is shown in Fig. 1.

The SVM classifier is used for training, testing, and classification. Satisfactory quality of training and testing allows using the resulting SVM classifier in the classification of new objects.

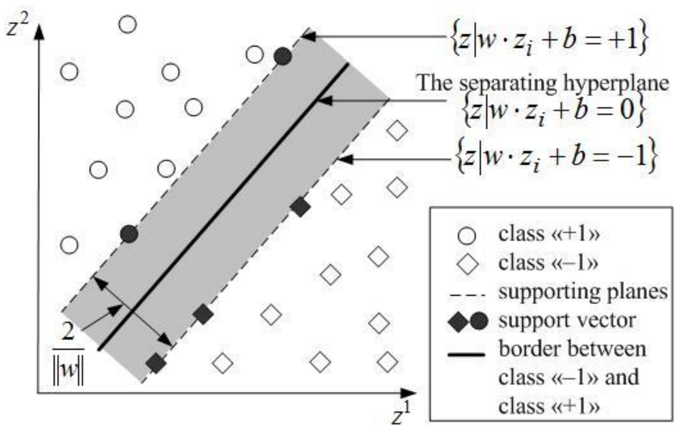


Fig. 1. Linear separation for two classes by the SVM classifier in the 2D space

SVM algorithms are well-known for their excellent performance in the sphere of the statistical classification. Still, the high computational cost due to the cubic runtime complexity is problematic for the Big Data sets: the training of the SVM classifier requires solving a quadratic optimization problem [1, 3]. Using a standard quadratic problem solver for the SVM classifier training would involve solving a big quadratic programming problem even for a moderate sized data set. This can limit the size of problems which can be solved with the application of the SVM classifier. Nowadays methods like SMO [11], chunking [12] and simple SVM [13], Pegasos [14] exist that iteratively compute the required solution and have a linear space complexity [15].

In recent years to mitigate the problem of the high computational cost the cascade SVM algorithm had been proposed [16]. In this algorithm the SVM classifier is iteratively trained on subsets of the original data set, acquired support vectors of the resulting models are combined to create new training sets. The general idea is to bound the sizes of all considered training sets and therefore obtain a significant speedup. This algorithm can easily be parallelized because the number of independent models has to be fitted during each stage of the cascade [17].

In the millennium of Big Data it is necessary to develop data mining algorithms which are suitable for the Big Data analysis. Several parallel algorithms have been developed using threads, MPI, MapReduce and so on [18]. Among all these techniques MapReduce is practically well suited for the Big Data analysis. One of the last trends in the Big Data processing and analysis is using the Hadoop framework for the SVM classifiers development [18]. Hadoop is an open-source software framework for the distributed storage and distributed processing of very large data sets on the computer clusters built from the commodity hardware. The Hadoop cluster is a special type of computational cluster designed specifically for storing and analyzing huge amounts of unstructured data in the distributed computing environment.

Therefore the use of the SVM algorithm is very perspective for the Big Data classification [19, 21].

Choosing optimal parameters for the SMV classifier is a significant problem at the moment. It is necessary to find the kernel function type, values of the kernel function parameters and value of the regularization parameter, which must be set by a user and shouldn't be changed [1, 2]. It is impossible to provide implementing of high-accuracy data classification with the use of the SVM classifier without adequate solution to this problem.

In the simplest case solution to this problem can be found by a search of the kernel function types, values of the kernel function parameters and value of the regularization parameter that demands significant computational expenses. For an assessment of classification quality, the indicators of classification accuracy, classification completeness, etc. can be used [3].

Usually, developing the binary classifiers requires working with the complex, multiextreme function, multi-parameter objective function.

Gradient methods are not suitable for search of the optimum of this objective function, but search algorithms of stochastic optimization, such as the genetic algorithm [22, 24], the artificial bee colony algorithm [25], the particle swarm algorithm [26-29], etc., have been used. Each of the optimal decision is carried out at once in all space of possible decisions.

The particle swarm algorithm (Particle Swarm Optimization, PSO algorithm), which is based on an idea of possibility to solve the optimization problems using modeling of animals' groups' behavior is the simplest algorithm of evolutionary programming because for its implementation it is necessary to be able to determine only value of the optimized function [26-29].

The traditional approach to application of the PSO algorithm consists of the repeated applications of the PSO algorithm for the fixed type of the kernel functions to choose optimal values of the kernel function parameters and value of the regularization parameter with the subsequent choice of the best type of the kernel function and values of the kernel function parameters and value of the regularization parameter corresponding to this kernel function type.

In a traditional approach to the application of the PSO algorithm it applied repeatedly to the fixed type of the kernel functions to find the optimal parameters. Whereas with a new approach the algorithm uses simultaneous search for the best type of the kernel function, values of the kernel function parameters and value of the regularization parameter. Hereafter, particle swarm algorithms corresponding to traditional and modified approaches will be called as the traditional PSO algorithm and the modified PSO algorithm respectively [30, 31].

It is necessary to say that the PSO algorithm and other nature inspired swarm optimization algorithms are very well suited for the distributed architecture and handling of high volume unstructured data in the Big Data analytics.

In recent years, much attention is paid to the question of increasing the accuracy of the models based on the machine learning algorithms. Therefore approaches dealing with the creation of the classifiers' ensembles for the accuracy increase of the classification solution have been investigated [3–5]. The training of the SVM ensemble is the training procedure of the finite set of the base (individual) classifiers: the individual solutions are combined to form the resulting classification decisions, based on the aggregated classifier. There are different approaches to choose the combination rules of the individual classifiers in the ensemble and the strategies for the creation of the resulting classification decisions [2].

The main purposes of this paper are the following: to create the modified PSO algorithm and compare it with the traditional one using the time required to find the optimal parameters of the SVM classifier and the classification accuracy of data; to improve the accuracy of the classification decisions using the SVM ensemble based on the decorrelation maximization algorithm for the different strategies of the decision-making on the data classification and the majority vote rule; to improve the accuracy of the classification decisions using the two-level SVM classifier.

The rest of this paper is structured as follows. Section II presents the main stages of the SVM classifier development. Section III details the proposed new approach for solving the problem of the simultaneous search of the kernel function type, values of the kernel function parameters and value of the regularization parameter for the SVM classifier. This approach is based on the application of the modified PSO algorithm, the main idea of which is the «regeneration» of particles: some particles change their kernel function type to the one which corresponds to the particle with the best value of the classification accuracy. Section IV is devoted to the problems of the development of the SVM ensembles on the base of the decorrelation maximization algorithm for the different strategies of the decision-making on the data classification and the majority vote rule. Section V details the two-level SVM classifier. This classifier works as the group of the SVM classifiers at the first level and as the SVM classifier on the base of the modified PSO algorithm at the second level. Experimental results follow in Section VI. Finally, conclusions are drawn in Section VII.

II. THE SUPPORT VECTOR MACHINE CLASSIFIER

Let the experimental data set be a set in the form of $\{(z_1,y_1),\ldots,(z_s,y_s)\}$, in which each object $z_i \in Z$ ($i=\overline{1,s}$; s is the number of objects) is assigned to a number $y_i \in Y=\{+1;\ -1\}$ having a value of $+1$ or -1 depending on the class of the object z_i. It is assumed that every object z_i is mapped to q -dimensional vector of numerical values of characteristics $z_i=(z_i^1,z_i^2,\ldots,z_i^q)$ (typically normalized by values from the interval [0, 1]) where z_i^l is the numeric value of the l -th characteristic for the i -th object ($i=\overline{1,s}$, $l=\overline{1,q}$) [30], [31]. It is necessary to use the special function $\kappa(z_i,z_\tau)$, which is called the kernel, to build the classifier $F:\ \ Z \to Y$, which compares the class to the number from the set $Y=\{+1;\ -1\}$ or some object from the set Z.

To build «the best» SVM classifier it is necessary to implement the numerous repeated training (for the training data set with S elements) and testing (for the test data set $s-S$ elements, $S<s$) on the different randomly generated training and test sets with following determination of the best SVM classifier in terms of the highest possible classification quality provision. The test set contains the part of data from the experimental data set. The size of the test set must be equal to 1/10 – 1/3 of the experimental data set. The test set doesn't participate in controlling the parameters of the SVM-classifier. This set is used to measure classifier's accuracy. The SVM classifier with satisfactory training and testing results can be used to classify new objects [1–3].

The separating hyperplane for the objects from the training set can be represented by equation $\langle w,z\rangle+b=0$, where w is a vector-perpendicular to the separating hyperplane; b is a parameter which corresponds to the shortest distance from the origin of coordinates to the hyperplane; $\langle w,z\rangle$ is a scalar product of vectors w and z [1–3]. The condition $-1<\langle w,z\rangle+b<1$ specifies a strip that separates the classes. The wider the strip, the more confidently we can classify objects. The objects closest to the separating hyperplane, are exactly on the bounders of the strip.

If classes can be separated by the straight line, a hyperplane can be chosen so that no objects from the training set get between them and then maximize the distance between the hyperplanes (width of the strip) $2/\langle w,w\rangle$, which will solve the problem of quadratic optimization [1, 2]:

$$\begin{cases} \langle w,w\rangle \to \min, \\ y_i\cdot(\langle w,z_i\rangle+b)\ge 1, \quad i=\overline{1,S}. \end{cases} \tag{1}$$

Finding the separating hyperplane is basically the dual problem of searching a saddle point of the Lagrange function, which reduces to the problem of quadratic programming, containing only dual variables [1, 2]:

$$\begin{cases} -L(\lambda) = -\sum_{i=1}^{S} \lambda_i + \\ \quad + \frac{1}{2} \cdot \sum_{i=1}^{S} \sum_{\tau=1}^{S} \lambda_i \cdot \lambda_\tau \cdot y_i \cdot y_\tau \cdot \kappa(z_i, z_\tau) \to \min_{\lambda}, \\ \sum_{i=1}^{S} \lambda_i \cdot y_i = 0, \\ 0 \le \lambda_i \le C,\ i = \overline{1,S}, \end{cases} \quad (2)$$

where λ_i is a dual variable; z_i is the object of the training set; y_i is a number (+1 or −1), which characterize the class of the object z_i from the experimental data set; $\kappa(z_i, z_\tau)$ is a kernel function; C is a regularization parameter ($C > 0$); S is a quantity of objects in the experimental data set; $i = \overline{1,S}$.

In training of the SVM classifier it is necessary to determine the kernel function type $\kappa(z_i, z_\tau)$, values of the kernel parameters and value of the regularization parameter C, which allows finding a compromise between maximizing of the gap separating the classes and minimizing of the total error. A herewith typically one of the following functions is used as the kernel function $\kappa(z_i, z_\tau)$ [1, 3, 32]:

– linear function: $\kappa(z_i, z_\tau) = < z_i, z_\tau >$;

– polynomial function: $\kappa(z_i, z_\tau) = (< z_i, z_\tau > +1)^d$;

– radial basis function:

$\kappa(z_i, z_\tau) = exp(- < z_i - z_\tau, z_i - z_\tau > /(2 \cdot \sigma^2))$;

– sigmoid function: $\kappa(z_i, z_\tau) = th(k_2 + k_1 \cdot < z_i, z_\tau >)$,

where $< z_i, z_\tau >$ is a scalar product of vectors z_i and z_τ; d [$d \in N$ (by default $d = 3$)]; σ [$\sigma > 0$ (by default $\sigma^2 = 1$)]; k_1 [$k_1 > 0$ (by default $k_1 = 1$)] and k_2 [$k_2 < 0$ (by default $k_2 = -1$)] are some of parameters; th is a hyperbolic tangent.

These kernel functions allow dividing the objects from different classes.

As a result of the training, the classification function is determined in the following form [1], [3]:

$$f(z) = \sum_{i=1}^{S} \lambda_i \cdot y_i \cdot \kappa(z_i, z) + b\ . \quad (3)$$

The classification decision, associating the object z to the class −1 or +1, is adopted in accordance with the rule [1], [3]:

$$F(z) = sign(f(z)) = sign\left(\sum_{i=1}^{S} \lambda_i \cdot y_i \cdot \kappa(z_i, z) + b \right). \quad (4)$$

The SVM classifier training results in determining the support vectors. Using the PSO algorithm provides better accuracy of classification by choosing the kernel function type, values of the kernel function parameters and value of the regularization parameter.

Quality of the SVM classifier can be measured by different classification quality indicators [3]. There are cross validation data indicator, accuracy indicator, classification completeness indicator and ROC curve analysis based indicator, etc.

III. The Modified Particle Swarm Optimization Algorithm

In the traditional PSO algorithm the n-dimensional search space (n is the number of parameters which are subject to optimization) is inhabited by a swarm of m agents-particles (elementary solutions). Position (location) of the i-th particle is determined by vector $x_i = (x_i^1, x_i^2, \ldots, x_i^n)$, which defines a set of values of optimization parameters. These parameters can be presented in an explicit form or even absent in the analytical record of the objective function $u(x) = u(x^1, x^2, \ldots, x^n)$ of the optimization algorithm (for example, the optimum is the minimum which must be achieved).

The particles must be situated randomly in the search space during the process of initialization. Each i-th particle ($i = \overline{1,m}$) has its own vector of speed $v_i \in R^n$ which influence i-th particle ($i = \overline{1,m}$) coordinates' values in every single moment of time corresponding to some iteration of the PSO algorithm.

The coordinates of the i-th particle ($i = \overline{1,m}$) in the n-dimensional search space uniquely determine the value of the objective function $u(x) = u(x^1, x^2, \ldots, x^n)$ which is a certain solution of the optimization problem [26–29].

For each position of the n-dimensional search space where the i-th particle ($i = \overline{1,m}$) was placed, the calculation of value of the objective function $u(x_i)$ is performed. A herewith each i-th particle remembers the best value of the objective function found personally as well as the coordinates of the position in the n-dimensional space corresponding to the value of the objective function. Moreover each i-th particle ($i = \overline{1,m}$) «knows» the best position (in terms of achieving the optimum of the objective function) among all positions that had been «explored» by particles (due to it the immediate exchange of information is replicated by all the particles). At each iteration particles correct their velocity to, on the one hand, move closer to the best position which was found by the particle independently and, on the other hand, to get closer to the position which is the best globally at the current moment. After a number of iterations particles must come close to the best position (globally the best for all iterations). However, it is possible that some particles will stay somewhere in the relatively good local optimum.

Convergence of the PSO algorithm depends on how velocity vector correction is performed. There are different approaches to implementation of velocity vector v_i correction for the i-th particle ($i = \overline{1,m}$) [26]. In the classical version of the PSO algorithm correction of each j-th coordinate of velocity vector ($j=\overline{1,n}$) of the i-th particle ($i=\overline{1,m}$) is made in accordance with formula [26]:

$$v_i^j = v_i^j + \hat{\varphi}\cdot\hat{r}\cdot(\hat{x}_i^j - x_i^j) + \tilde{\varphi}\cdot\tilde{r}\cdot(\tilde{x}^j - x_i^j)\,, \qquad (5)$$

where v_i^j is the j-th coordinate of velocity vector of the i-th particle; x_i^j is the j-th coordinate of vector x_i, defining the position of the i-th particle; $\hat{x}_i^j$ is the j-th coordinate of the best position vector found by the i-th particle during its existence; $\tilde{x}^j$ is the j-th coordinate of the globally best position within the particles swarm in which the objective function has the optimal value; $\hat{r}$ and $\tilde{r}$ are random numbers in interval (0, 1), which introduce an element of randomness in the search process; $\hat{\varphi}$ and $\tilde{\varphi}$ are personal and global coefficients for particle acceleration which are constant and determine behavior and effectiveness of the PSO algorithm in general.

With personal and global acceleration coefficients in (5) random numbers $\hat{r}$ and $\tilde{r}$ must be scaled; the global acceleration coefficient $\tilde{\varphi}$ operates by the impact of the global best position on the speeds of all particles and the personal acceleration coefficient $\hat{\varphi}$ operates by the impact of the personal best position on the velocity of some particle.

Currently different versions of the traditional PSO algorithm are known. In one of the most known canonical version it is supposed to undertake the normalization of the acceleration coefficients $\hat{\varphi}$ and $\tilde{\varphi}$ to make the convergence of the algorithm not so much dependent on the choice of their values [26].

Correction of each j-th coordinate of the velocity vector ($j=\overline{1,n}$) of the i-th particle ($i=\overline{1,m}$) is performed in accordance with formula [26]:

$$v_i^j = \chi\cdot[v_i^j + \hat{\varphi}\cdot\hat{r}\cdot(\hat{x}_i^j - x_i^j) + \tilde{\varphi}\cdot\tilde{r}\cdot(\tilde{x}^j - x_i^j)]\,, \qquad (6)$$

where χ is a compression ratio;

$$\chi = 2\cdot K \,/\, \left|2-\varphi-\sqrt{\varphi^2-4\cdot\varphi}\right|; \qquad (7)$$

$$\varphi = \hat{\varphi}+\tilde{\varphi} \;\; (\varphi > 4); \qquad (8)$$

K is some scaling coefficient, which takes values from the interval (0, 1).

When using formula (6) for correction of velocity vector the convergence of the PSO algorithm is guaranteed and there is no need to control the particle velocity explicitly [26].

Let the correction of velocity vector of the i-th particle ($i=\overline{1,m}$) is executed in accordance with one of the formulas (5) or (6). The correction of the j-th coordinate of the i-th particle ($i=\overline{1,m}$) can be executed in accordance with the formula:

$$x_i^j = x_i^j + v_i^j\,. \qquad (9)$$

Then for each i-th particle ($i=\overline{1,m}$) the new value of the objective function $u(x_i)$ can be calculated and the following check must be performed: whether a new position with coordinates vector x_i became the best among all positions in which the i-th particle has previously been placed. If new position of the i-th particle is recognized to be the best at the current moment the information about it must be stored in a vector $\hat{x}_i$ ($i=\overline{1,m}$).

Value of the objective function $u(x_i)$ for this position must be remembered. Then among all new positions of the swarm particles the check of the globally best position must be carried out. If some new position is recognized as the best globally at the current moment, the information about it must be stored in vector $\tilde{x}$. Value of the objective function $u(x_i)$ for this position must be remembered.

In the case of the SVM classifier's development with the use of the PSO algorithm the swarm particles can be defined by vectors declaring their position in the search space and corded by the kernel function parameters and the regularization parameter: (x_i^1, x_i^2, C_i), where i is a number of particle ($i=\overline{1,m}$); x_i^1, x_i^2 are the kernel function parameters of the i-th particle, [parameter x_i^1 is equal to the kernel function parameters d, σ or k_2 (depending on the kernel function type which corresponds to a swamp particle); parameter x_i^2 is equal to the kernel function parameter k_1, if the swamp particle corresponds to the sigmoid type of the kernel function, otherwise this parameter is assumed to be zero]; C_i is the regularization parameter [30, 31].

After that to choose the optimal parameter values of the kernel function and the regularization parameter traditional approach to the application of the PSO algorithm is concluded numerous times for the fixed kernel function’s type.

As a result for each type T of the kernel function, participating in the search, the particle with the optimal combination of the parameters values ($\tilde{x}^1, \tilde{x}^2, \tilde{C}$) providing high quality of classification will be defined [30, 31].

The best type and the best values of the required parameters are found using the results of the comparative analysis of the best particles received at realization of the PSO algorithm with the fixed kernel function type.

Along with the traditional approach to the application of the PSO algorithm in the development of the SVM classifier there is a new approach that implements a simultaneous search for the best kernel function type $\tilde{T}$, parameters’ values $\tilde{x}^1$ and $\tilde{x}^2$ of the kernel function and value of the regularization parameter $\tilde{C}$ [30, 31]. At such approach each i-th particle in a swamp ($i=\overline{1,m}$) defined by a vector which describes particle's position in the search space: (T_i, x_i^1, x_i^2, C_i), where T_i is the number of the kernel function type (for example, 1, 2, 3 – for polynomial, radial basis and sigmoid functions accordingly); parameters x_i^1, x_i^2, C_i are defined as in the previous case. It is possible to «regenerate» particle through changing its coordinate T_i on number of that kernel function type, for which particles show the highest quality of classification. In the case of particles' «regeneration» the parameters' values change so that they corresponded to new type of the kernel function (taking into account ranges of change of their values). Particles which didn't undergo «regeneration», carry out the movement in own space of search of some dimension.

The number of particles taking part in «regeneration» must be determined before start of algorithm. This number must be equal to 15% – 25% of the initial number of particles. It will allow particles to investigate the space of search. A herewith they won't be located in it for a long time if their indicators of accuracy are the worst.

The offered modified PSO algorithm can be presented by the following consequence of steps [30].

Step 1. To determine parameters of the PSO algorithm: number m of particles in a swamp, velocity coefficient K, personal and global velocity coefficients $\hat{\varphi}$ and $\tilde{\varphi}$, maximum iterations number $N_{\max}$ of the PSO algorithm. To determine types T of kernel functions, which take part in the search ($T=1$ – polynomial function, $T=2$ – radial basis function, $T=3$ – sigmoid function) and ranges boundaries of the kernel function parameters and the regularization parameter C for the chosen kernel functions' types T: $x_{\min}^{1T}$, $x_{\max}^{1T}$, $x_{\min}^{2T}$, $x_{\max}^{2T}$, $C_{\min}^{T}$, $C_{\max}^{T}$ ($x_{\min}^{2T}=0$ and $x_{\max}^{2T}=0$ for $T=1$ and $T=2$). To determine the particles' «regeneration» percentage p.

Step 2. To define equal number of particles for each kernel type function T, included in search, to initialize coordinate T_i for each i-th particle ($i=\overline{1,m}$) (every kernel function type must be corresponded by equal number of particles), other coordinates of the i-th particle ($i=\overline{1,m}$) must be generated randomly from the corresponding ranges: $x_i^1 \in [x_{\min}^{1T},\ x_{\max}^{1T}]$, $x_i^2 \in [x_{\min}^{2T},\ x_{\max}^{2T}]$ ($x_i^2=0$ under $T=1$ and $T=2$), $C_i \in [C_{\min}^{T},\ C_{\max}^{T}]$. To initialize random velocity vector $v_i(v_i^1, v_i^2, v_i^3)$ of the i-th particle ($i=\overline{1,m}$) ($v_i^2=0$ under $T=1$ and $T=2$). To establish initial position of the i-th particle ($i=\overline{1,m}$) as its best known position $(\hat{T}_i, \hat{x}_i^1, \hat{x}_i^2, \hat{C}_i)$, to determine the best particle with coordinates' vector $(\tilde{T}, \tilde{x}^1, \tilde{x}^2, \tilde{C})$ from all the m particles, and to determine the best particle for each kernel function type T, including in a search, with coordinates' vector $(\overline{T}, \overline{x}^{1T}, \overline{x}^{2T}, \overline{C}^{T})$. Number of executed iterations must be considered as 1.

Step 3. To execute while the number of iterations is less than the fixed number $N_{\max}$:

- «regeneration» of particles: to choose p% of particles which represent the lowest quality of classification from particles with coordinate $T_i \neq \tilde{T}$ ($i=\overline{1,m}$); to change coordinate T_i (with the kernel function type) on $\tilde{T}$; to change values of the parameters x_i^1, x_i^2, C_i of «regenerated» particles to let them correspond to a new kernel function type $\tilde{T}$ (within the scope of the corresponding ranges);
- correction of velocity vector $v_i(v_i^1, v_i^2, v_i^3)$ and position (x_i^1, x_i^2, C_i) of the i-th particle ($i=\overline{1,m}$) using formulas:

$$v_i^j = \begin{cases} \chi \cdot [v_i^j + \hat{\varphi} \cdot \hat{r} \cdot (\hat{x}_i^j - x_i^j) + \tilde{\varphi} \cdot \tilde{r} \cdot (\overline{x}^{jT} - x_i^j)], & j=1,\ 2, \\ \chi \cdot [v_i^j + \hat{\varphi} \cdot \hat{r} \cdot (\hat{C}_i - C_i) + \tilde{\varphi} \cdot \tilde{r} \cdot (\overline{C}^{T} - C_i)], & j=3, \end{cases} \quad (10)$$

$$x_i^j = x_i^j + v_i^j \text{ for } j=1,\ 2, \quad (11)$$

$$C_i = C_i + v_i^3, \quad (12)$$

- where $\hat{r}$ and $\tilde{r}$ are random numbers in interval (0, 1), χ is a compression ratio calculated using the formula (7); formula (10) is the modification of formula (6): the coordinates' values $\overline{x}^{1T}, \overline{x}^{2T}, \overline{C}^{T}$ are used instead of the coordinates' values $\tilde{x}^1, \tilde{x}^2, \tilde{C}$ of the globally best particle;
- accuracy calculation of the SVM classifier with parameters' values (T_i, x_i^1, x_i^2, C_i) ($i=\overline{1,m}$) with aim to find the optimal combination $(\tilde{T}, \tilde{x}^1, \tilde{x}^2, \tilde{C})$, which will provide high quality of classification;
- increase of iterations number on 1.

The particle with the optimal combination of the parameters' values $(\tilde{T}, \tilde{x}^1, \tilde{x}^2, \tilde{C})$ which provides the highest quality of classification on chosen the function types will be defined after execution of the offered algorithm.

After executing of the modified PSO algorithm it can be found out that all particles will be situated in the search space which corresponds to the kernel function with the highest classification quality because some particles in the modified PSO algorithm changed their coordinate, which is responsible for number of the kernel function. A herewith all other search spaces will turn out to be empty because all particles will «regenerate» their coordinate with number of the kernel function type. In some cases (for small value $N_{\max}$ and for small value p) some particles will not «regenerate» their kernel function type and will stay in their initial search space.

The modified PSO algorithm allows reducing the time expenditures for development of the SVM classifier.

IV. The Support Vector Machine Ensemble

In most cases SVM classifier based on the modified PSO algorithm provides high quality of data classification. In exceptional cases the SVM ensembles can be used to increase the classification accuracy. The using of the SVM ensemble allows fulfilling the high-precision data classification, especially Big Data classification, with the acceptable time expenditures.

After training, each classifier generates its own (individual) classification decisions, same or different from the actual results of classification. Accordingly, the different individual SVM classifiers correspond to the different classification accuracy. The quality of the received classification decisions can be improved on the base of ensembles of the SVM classifiers [3], [33–36]. In this case, the finite set of individually trained classifiers must be learned. Then the classification decisions of these classifiers are combined. The resulting solution is based on the aggregated classifier. The majority vote method and the vote method based on the degree of reliability can be used as the rules (strategies) of the definition of the aggregated solutions.

The majority vote method is one of the most common and frequently used methods for combining decisions in the ensemble of classifiers. But this method does not fully use the information about the reliability of each individual SVM classifier. For example, suppose that the SVM classifier ensemble aggregates the results of five individual SVM classifiers, where values of the function $f(z)$ of the object z (3) obtained from the three individual SVM classifiers, are negative (class –1), but very close to the neutral position, and values of the function $f(z)$ of the other two SVM classifiers are strongly positive (class +1), i.e. very far away from the neutral position. Then the result of the aggregated decision of the ensemble on the basis of «one classifier – one vote» is following: the object z belongs to the negative class (majority vote), although it is obvious, that the best and more appropriate choice for the object z is a positive class. Despite the good potential of the majority vote method for combining of the group of decisions, it is recommended to use other methods to increase the accuracy of classification.

Vote method based on the degree of reliability uses value of the function $f(z)$ for the object z obtained by each individual SVM classifier. The greater the positive value of $f(z)$ in (3) returned by the SVM classifier, the more precisely the object z is determined in class +1, and the less negative value $f(z)$, the more precisely the object z is defined in class –1. Values «–1» and «+1» for $f(z)$ indicate that the object z is situated on the boundary of the negative and positive classes, respectively.

When using an ensemble of classifiers for solving classification problems special attention should be paid to the methods of forming a set of individual classifiers, which can later be used in the development of the final SVM classifier. It is experimentally confirmed [3], [33–37], that the ensemble of classifiers shows better accuracy than any of its individual members, if individual classifiers are accurate and varied. Therefore, the formation of the set of the individual SVM classifiers is required: 1) to use the various kernel functions; 2) to build classifiers in the different ranges of change of the kernel parameters and regularization parameter; 3) to use various sets of training and test data. To select the appropriate members of the ensemble in the set of the trained SVM classifiers it is recommended to use the principle of maximum decorrelation. In this case the correlation between the selected classifications should be as small as possible. After training, each private j-th classifier from the k trained classifier will correspond to a certain array of errors: $e_{ij} = | y_{ij} - \tilde{y}_{ij} |$, where e_{ij} is the error of j-th classifier at i-th row of the experimental data set ($i = \overline{1,s}$; $j = \overline{1,k}$); y_{ij} is the classification decision (−1 or +1) of j-th classifier at i-th row of the experimental data set; $\tilde{y}_{ij}$ is the real meaning of a class (−1 or +1), for which the i-th object is belong to.

The SVM classifiers not permitting an error on the experimental data set should be excluded from further consideration and from the remaining quantity of the SVM classifiersIt is necessary to select an appropriate number of individual SVM classifiers with maximal variety. To solve this problem decorrelation maximization algorithm can be used. This algorithm provides a variety of individual SVM classifiers, being used in the construction of the ensemble [3]. If the correlation between the selected classifiers is small, then the decorrelation is maximal.

Let there be an error matrix E of set of individual SVM classifiers with size $s \times k$:

$$E = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1k} \\ e_{21} & e_{22} & \dots & e_{2k} \\ \dots & \dots & \dots & \dots \\ e_{s1} & e_{s2} & \dots & e_{sk} \end{bmatrix}, \tag{13}$$

where e_{ij} is the error of the j-th classifier at the i-th row of the experimental data set ($i = \overline{1,s}$; $j = \overline{1,k}$).

On the basis of the error matrix E (13) the following assessments can be calculated [3]:

– mean:

$$\bar{e}_j = \frac{1}{s} \sum_{i=1}^{s} e_{ij} \quad (j = \overline{1,k}); \tag{14}$$

– variance:

$$\mathrm{V}_{jj} = \frac{1}{s} \sum_{i=1}^{s} (e_{ij} - \bar{e}_j)^2 \quad (j = \overline{1,k}); \tag{15}$$

– covariance:

$$\mathrm{V}_{tj} = \frac{1}{s} \sum_{i=1}^{s} (e_{ij} - \bar{e}_j) \cdot (e_{it} - \bar{e}_t) \quad (j = \overline{1,k}, t = \overline{1,k}); \tag{16}$$

Then the elements r_{tj} of the correlation matrix with size $k \times k$ are calculated as:

$$r_{tj} = V_{tj} \Big/ \sqrt{V_{tt} \cdot V_{jj}} \; ; \tag{17}$$

where r_{tj} is the correlation coefficient, representing the degree of correlation of t -th and j -th classifiers ($j=\overline{1,k}$; $t=\overline{1,k}$); $r_{jj}=1$ ($j=\overline{1,k}$).

Using the correlation matrix R it is possible for each individual j -th classifier to calculate the plural-correlation coefficient ρ_j , which characterizes the degree of correlation of j -th and all other $(k-1)$ classifiers with numbers t ($t=\overline{1,k}$; $t \neq j$) [3]:

$$\rho_j=\sqrt{1-|R|/R_{jj}} \quad (j=\overline{1,k}), \tag{18}$$

where $|R|$ is the determinant of the correlation matrix R ; R_{jj} is the cofactor of the element r_{jj} of the correlation matrix R .

A quantity ρ_j^2 called the coefficient of determination. It shows the proportion of the variation of the analyzed variable, which is explained by variation of the other variables. The coefficient of determination ρ_j^2 can take values from 0 to 1. The closer the coefficient to 1, the stronger the relationship between the analyzed variables (in this case, between individual classifiers) [3]. It is believed that there is a dependency, if the coefficient of determination is not less than 0.5. If the coefficient of determination greater than 0.8, it is assumed that high dependence exists.

For selection of individual SVM classifiers for integration into the ensemble it is necessary to determine the threshold θ. Thus, the j -th individual classifier must be removed from the list of classifiers if the coefficient of determination ρ_j^2 satisfies to condition $\rho_j^2>\theta$ ($j=\overline{1,k}$). If it is necessary to identify the most various classifiers, generating decisions with the most different arrays of errors on the experimental data set, thresholds θ , satisfying to condition $\theta<0.7$ should be selected. The additional considerations can be also taken into account to avoid the exclusion of insufficient or excessive number of individual SVM classifiers.

The decorrelation maximization algorithm can be summarized into the following steps [3].

Step 1. To calculate the matrix V and the correlation matrix R with formulas (15), (16) and (17) respectively.

Step 2. To calculate the multiple correlation coefficients ρ_j ($j=\overline{1,k}$) with (18) for all classifiers.

Step 3. To remove classifiers, for which $\rho_j^2>\theta$ ($j=\overline{1,k}$), from the list of classifiers.

Step 4. To repeat iteratively steps 1 – 3 for the remaining classifiers in the list until for all classifiers the condition $\rho_j^2 \le \theta$ ($j=\overline{1,k}$) will not satisfied.

As a result, the list of classifiers used to form the ensemble will consist of m ($m \le k$) individual classifiers.

For classifiers selected in the ensemble, it is necessary to carry out:

– the rationing of degrees of the reliability;

– the strategy search for the integration of members of the ensemble;

– the calculation of the aggregated decision of the ensemble.

Value of the reliability $f_j(z)$, which is defined for the object z by the j -th classifier, falls into the interval (-∞, +∞). The main drawback of such values is that in the ensemble the individual classifiers with large absolute value are often dominated in the final decision of the ensemble. To overcome this drawback, the rationing is carried out: the transformation of values of degrees of reliability in the interval [0; 1] is fulfilled. In the case of binary classification in the rationalization for the object z the values of the reliability of its membership to positive class (labeled +1) $g_j^+(z)$ and to negative class $g_j^-(z)$ are determined. These values can be determined by the formulas [3]:

$$g_j^+(z)=\frac{1}{1+e^{-f_j(z)}}, \tag{19}$$

$$g_j^-(z)=1-g_j^+(z). \tag{20}$$

The selected individual classifiers are combined into the ensemble using $g_j^+(z)$ and $g_j^-(z)$ ($j=\overline{1,m}$) in accordance with one of the following five strategies [3].

1) *Maximum strategy:*

$$A(z)=\begin{cases} 1, & \text{if} \quad \max\limits_{j=\overline{1,m}} g_j^+(z) \ge \max\limits_{j=\overline{1,m}} g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \tag{21}$$

2) *Minimum strategy:*

$$A(z)=\begin{cases} 1, & \text{if} \quad \min\limits_{j=\overline{1,m}} g_j^+(z) \ge \min\limits_{j=\overline{1,m}} g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \tag{22}$$

3) *Median strategy:*

$$A(z)=\begin{cases} 1, & \text{if} \quad \frac{1}{m}\sum\limits_{j=1}^{m} g_j^+(z) \ge \frac{1}{m}\sum\limits_{j=1}^{m} g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \tag{23}$$

4) *Mean strategy:*

$$A(z)=\begin{cases} 1, & \text{if} \quad \sum\limits_{j=1}^{m} g_j^+(z) \ge \sum\limits_{j=1}^{m} g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \tag{24}$$

5) *Product strategy:*

$$A(z)=\begin{cases} 1, & \text{if} \quad \prod\limits_{j=1}^{m} g_j^+(z) \ge \prod\limits_{j=1}^{m} g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \tag{25}$$

The value $A(z)$ is an aggregated measure of the reliability's value of the SVM classifier ensemble. It can be used to integrate the members of the ensemble [3].

The learning algorithm of the ensemble of the SVM classifiers can be summarized into the following steps.

Step 1. To divide the experimental data set into k training data sets: TR_1, …, TR_k.

Step 2. To learn k individual SVM classifiers with the different training data sets TR_1, …, TR_k and to obtain k individual SVM classifiers (ensemble members).

Step 3. To select m ($m \le k$) SVM classifiers from k classifiers using the decorrelation maximization algorithm.

Step 4. To determine values of m classification functions for each selected individual SVM classifier: $f_1(z)$,…, $f_m(z)$.

Step 5. To transform values of degrees of reliability, using (19) and (20), for the positive class $g_1^+(z)$,…, $g_m^+(z)$ and for the negative class $g_1^-(z)$, … , $g_m^-(z)$.

Step 6. To determine the aggregated value $A(z)$ of the reliability of the SVM classifier ensemble using (21) – (25).

This algorithm, used for the weak SVM classifiers, will provide a better quality of the classification accuracy than accuracy of any single individual classifier used for aggregation.

The problem of choosing of the threshold θ is very important. Value θ for which all five rules of classification (21) – (25) show stable improvement of the classification quality must be chosen as the threshold value θ^* ($\theta^* < 0.7$). Thus the use of each of the five rules leads to improvement of the classification quality resulting in the reduction of the number of erroneous decisions, when the smaller number of individual classifiers, corresponding to the threshold value θ^*, is applied. Such stable improvement of the classification quality isn't observed for all examined values θ' (for which $\theta' > \theta^*$).

It should be noted, that the majority vote rule may be used for decisions, obtained using the classification rules (21) – (25), to determine the required threshold value θ^*.

V. Two-Level SVM Classifier

The main problem which limits the use of the PSO algorithm is associated with quite a lot of time required to search for the optimal parameters of the SVM classifier (the kernel function type, the values of the kernel function parameters and the value of the regularization parameter). The search time can be partly reduced by using a small number of particles in the swarm and a small number of iterations of the PSO algorithm. But in this case we limit the number of the generated and compared SVM classifiers, and will probably find the worst decision.

One approach to reducing the search time is associated with the reduction in the size of the training data set. A herewith those objects that will not affect the classification results shouldn't be considered. This approach is based on the following theoretical fact of the development of the SVM classifier: the classification function (3) performs the summation only for the support vectors for which $\lambda_i \ne 0$. These vectors contain all the information about the objects division and play the main role in the construction of the hyperplanes separating the classes.

Therefore the two-level SVM classifier has been developed. This SVM classifier works as the group of the SVM classifiers at the first level and as the SVM classifier on the base of the modified PSO algorithm at the second level.

This two-level SVM classifier is iteratively trained on subsets of the original experimental data set at the the first level. Then support vectors of the obtained SVM classifiers are combined to create the new training set for the SVM classifier on the base of the modified PSO algorithm.

The proposed approach to can be described by the following consequence of steps.

1) To train k SVM classifiers on the original experimental data set using different training data sets TR_1, TR_2, …, TR_k at the first level of the two-level SVM classifier. A herewith it is necessary to use the different kernel functions types, the different values of the kernel function parameters and the regularization parameter.

2) To obtain the support vectors sets SV_1, SV_2, …, SV_k from the trained SVM classifiers and form the set SV as the union of the support vectors sets SV_1, SV_2, …, SV_k. Let this set SV consists of t objects ($t \le s$, where s is the number of objects in the experimental data set).

3) To select from the set SV the subset SV^+, consisting of T ($T \le t$) objects (support vectors), which have been correctly classified by the SVM classifiers. It is necessary to ensure that false data is not participated in the training of the SVM classifier on the base of the modified PSO algorithm at the second level of the two-level SVM classifier. The rest objects (support vectors) from the set SV will form the subset SV^- with $t-T$ objects. The subset SV^+ will be used for the training and the subset SV^- will be used for the testing of the SVM classifier on the base of the modified PSO algorithm.

4) To develop the SVM classifier on the base of the modified PSO algorithm.

5) To classify objects (from the original experimental data set) which not included in the sets SV^+ and SV^-.

The using of the two-level SVM classifier also allows carrying out the high-precision data classification, especially Big Data classification, with the acceptable time expenditures.

VI. Experimental Studies

The assessment of the offered approaches for the development of the SVM classifiers and their ensembles has been carried out by test and real data.

In the first experiments for the particular data set the traditional PSO algorithm and the modified PSO algorithm have been applied. Comparison between these algorithms was executed using the found optimal parameters values of the SVM classifier, classification accuracy and spent time. All data sets used in the experimental researches we taken from the Statlog project and from the UCI library for machine learning.

Particularly, we used two data sets for medical diagnostics, two data sets for credit scoring and one data set for the creation of the predictive model of the spam recognition on the base of the e-mails' data set:

- Breast cancer data set of The Department of Surgery at the University of Wisconsin, in which the total number of instances is 569 including 212 cases with the diagnosed cancer (class 1) and 357 cases without such diagnosis (class 2); a herewith each patient is described by 30 characteristics ($q = 30$) and all information was obtained with the use of digital images (WDBC data set in the Tabl. I, the source is http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/);
- Heart disease data set, in which the total number of instances is 270 including 150 cases with the diagnosed heart disease (class 1) and 120 cases without such diagnosis (class 2); a herewith each patient is described by 13 characteristics ($q = 13$) (Heart data set in the Tabl. I, the source is http://archive.ics.uci.edu/ml/machine-learning-databases/statlog/heart/; a herewith desease was found for 150 patients (class 1) and desease was not found for 120 patients (class 2));
- Australian consumer credit data set, in which the total number of instances is 690 including 382 creditworthy cases (class 1) and 308 default cases (class 2); a herewith each applicant is described by 14 characteristics ($q = 14$) (Australian data set in the Tabl. I, the source is http://archive.ics.uci.edu/ ml/machine-learning-databases/ statlog/australian/);
- German credit data set, in which the total number of instances is 1000 including 700 creditworthy cases (class 1) and 300 default cases (class 2); a herewith each applicant is described by 24 characteristics ($q = 24$) (German data set in the Tabl. I; the source is http://archive.ics.uci.edu/ml/machine-learning-databases/statlog/german/);
- Spam data set, in which in which the total number of instances is 4601 including 1813 cases with the spam (class 1), that is equal to 39.4% of the data set size, and 2788 cases without the spam (class 2); a herewith each e-mail is described by 57 characteristics ($q = 57$) (Spam data set in the Tabl. I; the source is https://archive.ics.uci.edu/ml/machine-learning-databases/spambase/).

The Spam data set we consider as an example of the Big Data. It is logical, especially, if we plan to use the developed SVM-classifier for the identification of the new spam patterns in a data flow.

Also, we used two test data sets Test1 (Test1 data set in the Tabl. I) and Test2 (Test2 data set in the Tabl. I; the source is http://machinelearning.ru/wiki/images/b/b2/MOTP12_svm_example.rar) [26].

For all data sets binary classification has been performed.

Experimental calculations were made on the base of PC under the Microsoft Windows 7 for x64-based Operating System with the random access memory of 3 GB and the four-nuclear Intel® Core™ i3 processor with the kernels' clock frequency of 2.53 GHz. The SVM algorithm from the software package MATLAB 7.12.0.635 was applied for the modeling.

For development of the SVM classifier the traditional and the modified PSO algorithms were used, meaning that the choice of the optimal values of the SVM classifier parameters was conducted. The kernels with polynomial (#1), radial basis (#2) and sigmoid (#3) functions were included in the search and the identical values of the PSO algorithm parameters and the identical ranges of values' change of the required SVM classifier parameters were established.

The short description of characteristics of each data set is provided in the Table I. Here the search results of the optimal values of parameters of the SVM classifier with the application of the traditional PSO algorithm and the modified PSO algorithm are presented (in the identical ranges of parameters' change and at the identical PSO algorithm parameters), number of error made during the training and testing of the SVM classifier and search time.

TABLE I. THE SEARCH RESULTS BY MEANS OF THE TRADITIONAL PSO ALGORITHM AND THE MODIFIED PSO ALGORITHM

Data set	**Number of objects**	**Number of characteristics**	**PSO algorithm type**	**Found parameters**				**Errors**		**Number of the support vectors**	**Accuracy (%)**	**Search time (sec.)**
				Kernel number	C	x_1	x_2	***At the training***	***At the testing***			
Test1	300	2	traditional	1	7.97	3	-	0 of 240	0 of 60	6	100	1478
			modified	1	2.83	8	-	0 of 240	0 of 60	6	100	382
Heart	270	13	traditional	2	9.6	3.34	-	7 of 230	7 of 40	106	94.81	2276
			modified	2	6.01	2.99		6 of 230	3 of 40	131	96.67	876
WDBC	569	30	traditional	2	9.36	2.89	-	0 of 427	3 of 142	113	99.47	3919
			modified	2	9.84	3.99	-	0 of 427	2 of 142	79	99.65	1464
Australian	690	14	traditional	2	9.28	2.51	-	23 of 518	25 of 172	248	93.04	9086
			modified	2	4.45	2.22		26 of 518	20 of 172	258	93.33	2745
German	1000	24	traditional	1	1.58	3	-	0 of 850	42 of 150	438	95.8	14779
			modified	1	5.53	4	-	0 of 850	42 of 150	546	95.8	5766

Test2	400	2	traditional	2	6.31	0.19	-	7 of 340	8 of 60	156	96.25	15632
			modified	2	5.69	0.22	-	11 of 340	4 of 60	146	96.25	7146
Spam	4601	57	traditional	2	7.82	2.47	-	40 of 3681	56 of 920	1634	97.91	92645
			modified	2	8.57	2.45	-	36 of 3681	60 of 920	1659	97.91	44933

For example, for the WDBC data set with the use of the traditional and the modified PSO algorithms the kernel with radial basis function (#2) was determined as the optimal. For the traditional PSO algorithm the optimal values of the kernel parameter and the regularization parameter are equal to $\sigma = 2.89$ and $C = 9.36$ accordingly. For the modified PSO algorithm the optimal values of the kernel parameter and the regularization parameter are equal to $\sigma = 3.99$ and $C = 9.84$ accordingly.

The classification accuracy by the traditional PSO algorithm is equal to 99.47%, and the classification accuracy by the modified PSO algorithm is equal to 99.65%. The search time came to 3919 and 1464 seconds accordingly.

For the Spam data set with the use of the traditional and the modified PSO algorithms the kernel with radial basis function (#2) also was determined as the optimal. For the traditional PSO algorithm the optimal values of the kernel parameter and the regularization parameter are equal to $\sigma = 2.47$ and $C = 7.82$ accordingly. For the modified PSO algorithm the optimal values of the kernel parameter and the regularization parameter are equal to $\sigma = 2.45$ and $C = 8.57$ accordingly.

The classification accuracy by the traditional and modified PSO algorithm is equal to 97.91%. The search time came to 92645 and 44933 seconds accordingly.

Fig. 2–4 show for the Spam data set the location examples of the particles swarm in the D-2 search spaces and in the D-3 search space at the initialization, at the 3-rd iteration and at the 12-th iteration. These locations of the particles in the swamp were obtained with the use of the modified PSO algorithm.

The kernels with polynomial, radial basis and sigmoid functions were included in the search. A herewith the following change ranges of values' parameters were set: $3 \le d \le 8$, $d \in N$ (for the polynomial function); $0.1 \le \sigma \le 10$ (for the radial basis function); $-10 \le k_2 \le -0.1$ and $0.1 \le k_1 \le 10$ (for the sigmoid function).

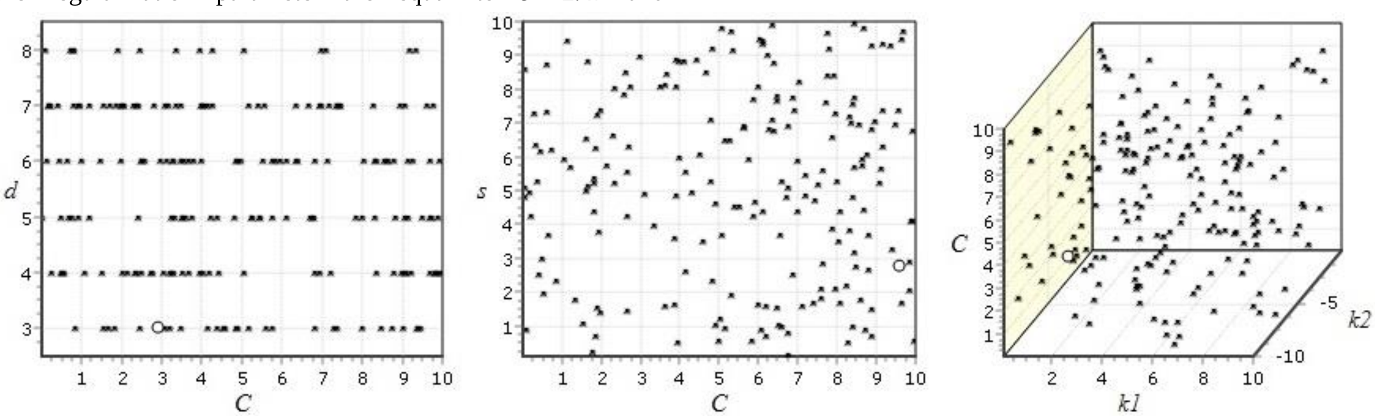


Fig. 2. Location of the particles in the swamp at the initialization (polynomial kernel function is on the left, radial basis is in the middle, sigmoid is on the right)

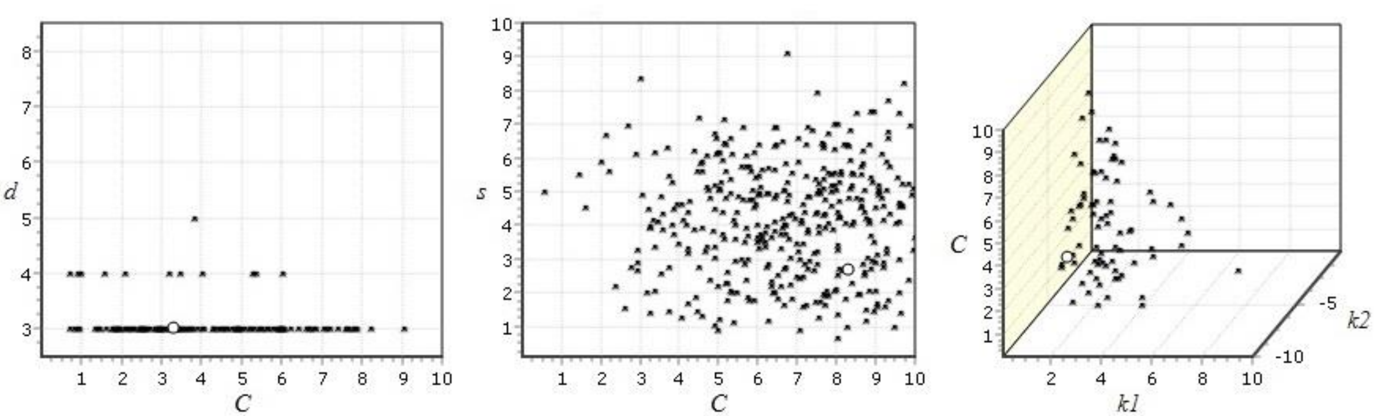


Fig. 3. Location of the particles in the swamp at the 3-rd iteration (polynomial kernel function is on the left, radial basis is in the middle, sigmoid is on the right)

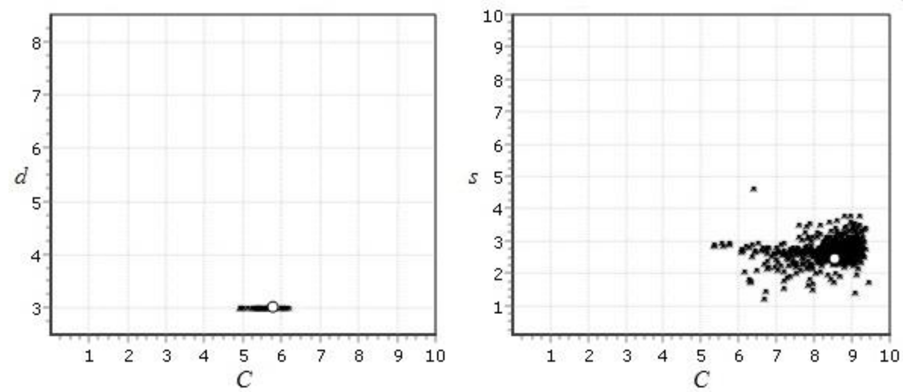


Fig. 4. Location of the particles in the swamp at the 12-th iteration (polynomial kernel function is on the left, radial basis is on the right)

Change range for the regularization parameter C was determined as: $0.1 \le C \le 10$. Moreover, the following values of parameters of the PSO algorithm were set: number m of particles in a swarm equal to 600 (200 per each kernel function type); iterations' number $N_{\max} = 20$; personal and global velocity coefficients equal to $\hat{\varphi} = 2$ and $\tilde{\varphi} = 5$ accordingly; the scaling coefficient $K = 0.3$; «regeneration» coefficient of particles $p = 20\%$.

Particles are marked by asterisk bullets in the search spaces and the best position from the search space is marked by white round bullet. During realization of the modified PSO algorithm the swamp particles moves towards the best (optimal) position for the current iteration in the search space and demonstrate collective search of the optimal position. A herewith velocity and direction of each particle are corrected. Moreover «regeneration» of particles takes place: some particles change own search space to space, in which particles show the best quality of classification.

Thus, at the realization of the modified PSO algorithm there is a change of the particles' coordinates, which are responsible for the parameters of the kernel function $\kappa(z_i, z_\tau)$ and the regularization parameter C. Besides, the type of the kernel function also changes. As a result the particles moves towards the united search space (in this case – the space corresponding to the radial basis kernel function) leaving the space where they were initialized.

In the reviewed example only 7 particles didn't change their kernel function type after 20 iterations. Other particles situated near the best position responsible for the optimal solution in the search space (Fig. 5).

Fig. 6 shows the location examples of the particles swarm in the D-2 search space at the initialization and at the 2-nd, the 7-th and the 10-th iterations of the traditional PSO algorithm for the radial basis kernel function. The best particle has been found at the 8-th iteration, though 20 iterations have been executed.

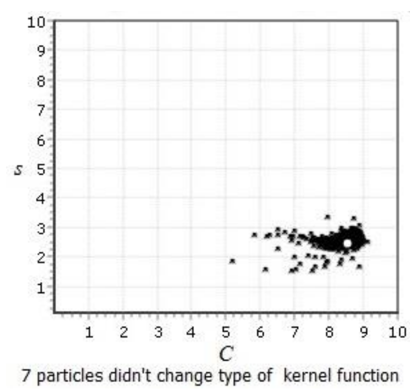
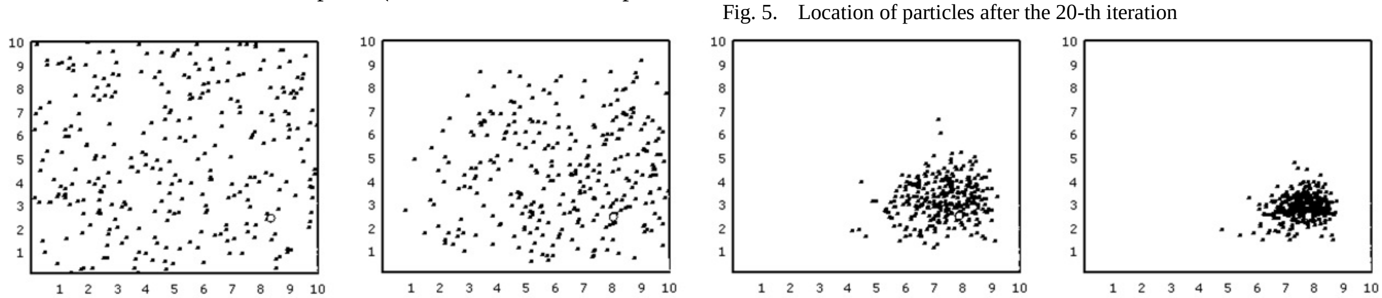


Fig. 5. Location of particles after the 20-th iteration



(a) (b) (c) (d)

Fig. 6. Location of the particles in the swamp: a) at the initialization; b) at the 2-nd iteration; c) at the 7-th iteration; d) at the 10-th iteration (the horizontal axis corresponds to the regularization parameter C, the vertical axis corresponds to the parameter σ)

TABLE II. THE CHARACTERISTICS OF THE BEST CLASSIFIER AT THE REALIZATION OF THE PSO ALGORITHM

Traditional PSO algorithm (with the radial basis kernel function)					**Modified PSO algorithm**				
Performance stage	**Errors**		**Total quantity of error**	**Number of the support vectors**	**Performance stage**	**Errors**		**Total quantity of error**	**Number of the support vectors**
	at the training	***at the testing***				***at the training***	***at the testing***		
Initialization of the swarm	45	85	130	1483	Initialization, the 1-st and 2-nd iterations	45	56	101	1494
The 1-st iteration	41	88	129	1648	The 3-d iterations	43	57	100	1582
The 2-nd iteration	56	73	129	1264	The 4-th iterations	38	61	99	1592
The 3-d and the 4-th iterations	54	73	127	1240	The 5-th and the 6-th iterations	38	61	99	1584
The 5-th iteration	54	73	127	1240	The 7-th iteration	37	61	98	1671
The 6-th iteration	40	58	98	1645	The 8-th iteration	37	61	98	1589
The 7-th iteration	35	61	96	1686	The 9-th , the 10-th and 11-th iterations	38	59	97	1602
From the 8-th to 20-th iteration	40	56	96	1634	From the 12-th to 20-th iteration	36	60	96	1659

The kernels with polynomial, radial basis and sigmoid functions were included in the search. The following change ranges of values' parameters were set: $3 \le d \le 8$, $d \in N$ (for the polynomial function); $0.1 \le \sigma \le 10$ (for the radial basis function); $-10 \le k_2 \le -0.1$ and $0.1 \le k_1 \le 10$ (for the sigmoid function.

Table II shows the information on the best SVM classifier at the different iterations of the traditional PSO (for the radial basis kernel function, which was defined as the best kernel function) and modified PSO algorithm (for three kernel functions) for the Spam data set.

It is visible from the Table I, that as a result of the search for the reviewed data sets both algorithms determined identical kernel function type as the optimal, similar values of the kernel function parameter and the regularization parameter, and also similar accuracy values of training and testing of the SVM classifier.

But the modified PSO algorithm is more effective, because it took less (more than 2 – 3 times) time for search compared to the traditional one.

At the determination of the optimal parameters' values of the SVM classifier with use of the traditional or modified PSO algorithm in the chosen search space we must create the huge number of the SVM classifiers to figure out, which shows the maximum classification accuracy under the minimum number of the support vectors. Therefore at the implementation of the PSO algorithm with 600 particles under 20 iterations of the PSO algorithm it is necessary to build and compare 12000 SVM classifiers.

If the average time of the training and testing of the SVM classifier equal to 5 seconds, then the time expenditures for the search of the optimal parameters' values of the SVM classifier will be $600 \times 20 \times 5 = 60000$ seconds or about 16.67 hours, that considerably surpasses the time expenditures for the development of 18 SVM classifiers (90 seconds) under the SVM ensemble development.

The experimental studies show, that the search time is defined: a) by the own parameters of the PSO algorithm (the speed coefficients, the quantity of the kernel functions and the types of the kernel functions, the search ranges, etc.); b) by properties of the experimental data set used for the training and testing of the SVM classifier (in particular, by the size of the data set and the number of characteristics). The lesser search time of the modified PSO algorithm in comparison with the search time of the traditional PSO algorithm is explained by the fact that some particles "regenerate" from the one search space (with the one kernel function type) to another search space (with the another kernel function type). The time expenditures for the SVM classifier development for the first kernel function type are more expensive than for the second kernel function type (in particular, the most expensive on time is the polynomial kernel function).

It should be noted that the SVM classifier for the German data set doesn't have really good classification accuracy assessment (in comparison with the SVM classifiers for other data sets). The attempt of the SVM classifier training in this case leads to the SVM classifier with not really high classification accuracy or to the retraining of the SVM classifier when the number of error for the test set is significantly more, than for the training set (with the acceptable classification accuracy for the experimental data set in general). Therefore, it is expedient to try to use other approaches to the classifier development, in particular, the approach based on the SVM ensemble development.

One more reason to use SVM ensemble is the realization of the PSO algorithm, which deals with the high time expenditures: to increase the classification accuracy we need to increase the number of the PSO algorithm iterations or/and number of particles in the swarm, however it doesn't guarantee that the expected high classification accuracy will be obtained. Therefore it is necessary to try to develop the SVM ensemble on the base of the individual SVM classifiers with not really high classification accuracy. The classification accuracy of the SVM ensemble should have higher classification accuracy than the classification accuracy of the used individual SVM classifiers.

In the last experiments the usefulness of the SVM ensembles was confirmed with application of test and real data sets.

Several individual SVM classifiers using different types of the kernel function, different values of the kernel function functions of the kernel parameters and different values of the regularization parameter were learned in experiments for the particular data sets. The different training and test sets randomly generated from the original data set were used. Then the decorrelation maximization algorithm for the different strategies of the decision-making on the data classification and the majority vote rule were applied.

For example, for the German data set we developed 18 individual SVM classifiers with use of various input parameters.

At the testing it was found, that the individual classifiers indicate the classification accuracy in range from 83.5% to 93.2%, and the initial values of the determination coefficient (if $\theta^* = 1$), calculated for all 18 individual classifiers, are in the range from 0.049 to 0.534. As a result, the threshold values θ were examined from the range $[0.1;\ 0.55]$ with step 0.05. Values of the classification parameters corresponding to the different threshold values θ are given in the Table III.

The optimal threshold value θ^* for the reviewed example is 0.3, since for $\theta^* = 0.3$ all five classification rules (strategies) (21) – (25) give the stable improvement of the classification quality when the number of classifiers reduces to the number corresponding to the threshold value $\theta^* = 0.3$. The finite number of classifiers in the SVM ensemble proved equal to 6. A further decrease in the number of classifiers is not feasible (due to a further sharp decrease in their number and a substantial reduction of their variety).

The use of the median strategy (or sum strategy) with $\theta^* = 0.3$ allowed classifying correctly 98.29% of the objects of the initial data set. At the same time, the maximum classification accuracy of one of the individual SVM classifiers, used in the SVM ensemble, was equal to 93.2%, and the accuracy reached with use of the majority vote rule was equal to 96.8%.

Thus, the use of the SVM ensemble allowed increasing the classification accuracy by more than 5% compared to the maximum classification accuracy of one of the individual classifiers in the SVM ensemble.

TABLE III. VALUES OF CLASSIFICATION PARAMETERS AT THE DIFFERENT THRESHOLD VALUES OF THE DETERMINATION COEFFICIENT (GERMAN DATA SET)

Value of classification	**Strategy**	**The threshold value of the determination coefficient**							
		0.55	**0.5**	**0.45/0.4**	**0.35**	**0.3**	**0.25**	**0.2**	**0.15/0.1**
Overall accuracy (%)	Majority vote	96.80	96.80	96.80	96.80	96.80	96.80	96.80	96.80
	Maximum and minimum	79.90	80.10	81.10	83.80	90.30	91.20	92.30	93.50
	Median and sum	95.20	96.20	95.60	97.20	98.29	97.70	97.80	97.60
	Product	87.40	89.10	89.00	90.90	97.10	97.30	97.00	96.10
Sensitivity (%)	Majority vote	98.00	98.00	98.00	98.00	98.00	98.00	98.00	98.00
	Maximum and minimum	84.71	84.86	85.86	88.00	94.57	95.57	96.29	96.71
	Median and sum	96.14	97.43	96.86	98.57	99.29	99.14	99.43	98.57
	Product	89.14	90.86	90.86	92.29	99.00	99.00	98.86	97.71
Specificity (%)	Majority vote	94.00	94.00	94.00	94.00	94.00	94.00	94.00	94.00
	Maximum and minimum	68.67	69.00	70.00	74.00	80.33	81.00	83.00	86.00
	Median and sum	93.00	93.33	92.67	94.00	95.67	94.33	94.00	95.33
	Product	83.33	85.00	84.67	87.67	92.67	93.33	92.67	92.33
Number of errors of the 1-st type	Majority vote	14	14	14	14	14	14	14	14
	Maximum and minimum	107	106	99	84	38	31	26	23
	Median and sum	27	18	22	10	5	6	4	10
	Product	76	64	64	54	7	7	8	16
Number of errors of the 2-nd type	Majority vote	18	18	18	18	18	18	18	18
	Maximum and minimum	94	93	90	78	59	57	51	42
	Median and sum	21	20	22	18	13	17	18	14
	Product	50	45	46	37	22	20	22	23
Number of classifiers in the ensemble		18	15	13	8	6	5	4	3

TABLE IV. THE PARAMETERS AND CHARACTERISTICS OF THE INDIVIDUAL CLASSIFIERS (SPAM DATA SET)

Number of the individual SVM classifier	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Kernel function type	1	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3	4
Regularization parameter	1.2	3.1	2	1.3	2.5	4	0.5	4	10	15	12	10	0.25	0.5	0.5	0.2	3.4	1.2
Kernel parameters	3	3	4	4	3	0.2	0.4	0.6	5	10	10	15	0.5	0.15/ -1.00	0.50/ -1.50	0.20/ -2.20	0.30/ -0.80	-
Overall accuracy (%)	83.33	91.33	86.72	85.59	90.70	93.44	93.68	94.04	83.89	94.63	92.98	92.87	90.33	87.00	83.53	82.29	82.59	91.72
Number of the support vectors	578	606	618	549	516	3099	3156	2891	646	686	754	836	3444	585	385	1160	628	874
Size of the training set	3681	3681	3911	3681	3451	3451	3451	3451	3681	3221	3681	3681	3681	3911	3911	3681	3451	3911
Quantity of errors at the training	574	286	482	509	273	5	13	11	581	151	253	263	18	510	652	659	613	316
Size of the test set	920	920	690	920	1150	1150	1150	1150	920	1380	920	920	920	690	690	920	1150	690
Quantity of errors at the testing	193	113	129	154	155	297	278	263	160	96	70	65	427	88	106	156	188	65
Sensitivity of the classifier (%)	58.52	79.54	69.39	66.30	78.60	83.45	84.11	85.11	59.85	90.95	85.22	86.05	98.95	81.69	68.34	74.30	81.69	84.28

Specificity of the classifier (%)	99.46	99.00	97.99	98.13	98.57	99.93	99.89	99.86	99.53	97.02	98.03	97.31	84.72	90.46	93.40	87.48	83.18	96.56
Number of errors of the 1-st type	752	371	555	611	388	300	288	270	728	164	268	253	19	332	574	466	332	285
Number of errors of the 2-nd type	15	28	56	52	40	2	3	4	13	83	55	75	426	266	184	349	469	96
Initial determination coefficient	0.474	0.385	0.312	0.311	0.377	0.106	0.134	0.121	0.383	0.63	0.711	0.757	0.025	0.473	0.486	0.276	0.308	0.572
Development time, s	5	6	16	5	5	7	8	7	3	2	3	3	8	2	2	4	2	2

For the Spam data set we also developed 18 individual SVM classifiers with use of various input parameters.

The parameters and characteristics of 18 individual classifiers have been shown in the Table IV. The kernels with polynomial (#1), radial basis (#2), sigmoid (#3) and linear (#4) functions were included in the search. In the Table IV for the sigmoid kernel function the first number is k_1, the second number is k_2.

Also Table IV shows information on the time expenditures for the training of each individual SVM-classifier. The total time of the training is 90 seconds. At the training for each individual SVM-classifier the training set was formed in a random way on the base of the initial experimental data set of the e-mails. The number of instances in the test set was equal to 10%–25% of the initial number of instances in the initial experimental data set.

At the testing it was found, that the individual classifiers indicate the classification accuracy ranged from 82.29% to 94.63%, and the initial values of the determination coefficient (if $\theta^* = 1$), calculated for all 18 individual classifiers, are in the range from 0.025 to 0.757. As a result, the threshold values θ were examined from the range $[0.15;\ 0.8]$ with step 0.05. Values of the classification parameters corresponding to the different threshold values θ are given in the Table V.

The optimal threshold value θ^* for the reviewed example belongs to the range $[0.15;\ 0.25]$, since for the threshold values from this range all five classification rules (21) – (25) give the stable improvement of the classification quality when the number of classifiers reduces to the number corresponding to the threshold value θ^* from the range $[0.15;\ 0.25]$. The finite number of classifiers in the SVM ensemble proved is equal to 4. A further decrease in the number of classifiers is not feasible (due to a further sharp decrease in their number and a substantial reduction of their variety).

Table VI shows information on the characteristics of the individual SVM-classifier, which take a part in the SVM ensemble. This ensemble was created on the base of the strategies (21) – (25) for the threshold values θ^* from the range $[0.15;\ 0.25]$. Also Table VI shows information on the characteristics of the best SVM-classifier on the base of the traditional PSO algorithm and the modified PSO algorithm.

Use of the maximum (minimum) strategy allowed classifying correctly 98.59% of the objects in the initial data set. At the same time, the maximum classification accuracy of one of the individual SVM classifiers, used in the SVM ensemble, was equal to 94.04% (for the 13-th SVM classifier), and the accuracy reached with use of the majority vote rule was equal to 96.8%.

The application of other strategies also leads to increasing of the classification accuracy in comparison to the classification accuracy of the individual SVM classifiers, the classification accuracy on the base of the majority vote rule and the classification accuracy of the SVM classifier on the base of the PSO algorithm.

TABLE V. VALUES OF CLASSIFICATION PARAMETERS AT THE DIFFERENT THRESHOLD VALUES OF THE DETERMINATION COEFFICIENT (SPAM DATA SET)

Value of classification	**Strategy**	**The threshold value of the determination coefficient**							
		0.8	**0.7**	**0.6**	**0.5**	**0.4**	**0.35**	**0.3**	**0.25-0.15**
Overall accuracy (%)	Majority vote	95.44	95.44	95.44	95.44	95.44	95.44	95.44	95.44
	Maximum and minimum	84.35	84.35	84.35	84.35	84.22	84.13	88.05	98.59
	Median and sum	94.31	94.28	94.02	93.89	95.26	95.31	97.09	98.44
	Product	86.55	86.50	86.39	86.29	85.87	85.09	93.26	98.44
Sensitivity (%)	Majority vote	89.30	89.30	89.30	89.30	89.30	89.30	89.30	89.30
	Maximum and minimum	83.78	83.78	83.78	83.78	83.95	83.89	82.24	96.47
	Median and sum	86.82	86.43	85.77	85.22	88.36	89.41	94.04	96.14
	Product	84.28	84.34	84.17	84.00	84.56	84.17	91.01	96.14
Specificity (%)	Majority vote	99.43	99.43	99.43	99.43	99.43	99.43	99.43	99.43
	Maximum and minimum	84.72	84.72	84.72	84.72	84.44	84.29	91.82	99.96
	Median and sum	99.18	99.39	99.39	99.53	99.75	99.14	99.07	99.93
	Product	88.02	87.91	87.84	87.77	86.73	85.69	94.73	99.93
Number of errors of the 1-st type	Majority vote	194	194	194	194	194	194	194	194
	Maximum and minimum	294	294	294	294	291	292	322	64
	Median and sum	239	246	258	268	211	192	108	70
	Product	285	284	287	290	280	287	163	70
Number of errors of the 2-nd type	Majority vote	16	16	16	16	16	16	16	16
	Maximum and minimum	426	426	426	426	435	438	228	1
	Median and sum	23	17	17	13	7	24	26	2
	Product	334	337	339	341	370	399	147	2

Number of classifiers in the ensemble	18	16	15	14	11	8	5	4

TABLE VI. THE CLASSIFICATION RESULTS ON THE BASE OF THE INDIVIDUAL SVM CLASSIFIERS AND THEIR SVM ENSEMBLE

The classifier characteristics	**The SVM classifier on the base**		**The classifier number, which take a part in the SVM ensemble**				**Strategy**			
	of the traditional PSO algorithm	***of the modified PSO algorithm***	***6***	***7***	***8***	***13***	***Maximum and minimum***	***Median and sum***	***Product***	***Majority vote***
Overall accuracy (%)	97.91	97.91	93.44	93.68	94.04	90.33	98.59	98.44	98.44	95.44
Sensitivity of the classifier (%)	96.47	96.14	83.45	84.11	85.11	98.95	96.47	96.14	96.14	89.30
Specificity of the classifier (%)	98.85	99.07	99.93	99.89	99.86	84.72	99.96	99.93	99.93	99.43
Number of errors of the 1-st type	64	70	300	288	270	19	64	70	70	194
Number of errors of the 2-nd type	32	26	2	3	4	426	1	2	2	16
Determination coefficient (θ)	-	-	0.078	0.104	0.091	0.004	$\theta^* \in [0.15;\ \ 0.25]$			-

Thus, the use of the SVM ensemble allowed increasing the classification accuracy almost by 5% compared to the maximum classification accuracy of one of the individual classifiers in the SVM ensemble.

The SVM ensemble with 98.59% classification accuracy doesn't concede to the SVM classifier on the base of the modified PSO algorithm with 97.91% classification accuracy and strongly surpasses it at the minimization of the time expenditures.

The proposed two-level SVM classifier was used for the Test2 data set classification (Table I). It is evident that, despite the small volume ($s = 400$) and the number of characteristics ($q = 2$), the PSO algorithm finds the optimal parameters for the SVM classifier in quite a long time (longer than, for example, for the WDBC data set of 569 objects with 30 characteristics). This is due to the data being hard to separate. Fig. 7 shows the location of the data in the 2D space and its division into two classes. Objects of the first class are marked by asterisk bullets, objects of the second class are marked by plus bullets. It is obviously that it is very difficultly to draw the curve separating the classes.

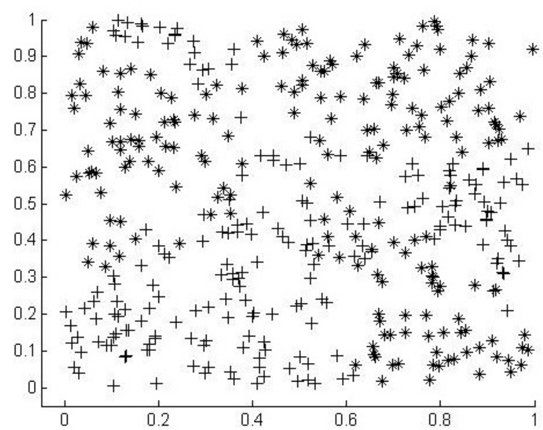


Fig. 7. Representation of the data set Test2 in 2D space

TABLE VII. THE PARAMETERS AND CHARACTERISTICS OF THE INDIVIDUAL CLASSIFIERS (TEST2 DATA SET)

Number of the individual SVM classifier	1	2	3	4	5	6	7	8	9	**Result**
Kernel function type	1	1	1	2	2	2	3	3	3	**2**
Regularization parameter	1	0.2	1.2	1.2	1.7	4.5	8.4	9	6.6	**8.5**
Kernel parameters	3	3	3	0.5	0.6	0.6	0.80; -3.00	0.50; -2.20	0.90; -2.50	**0.25**
Overall accuracy (%)	90.25	89.75	90.75	91.5	90.5	91	86.25	88.5	85.75	**96.75**
Number of the support vectors	84	112	96	121	106	87	150	150	103	**101**
Size of the training set	300	320	340	340	320	300	340	320	300	**204**
Quantity of errors at the training	27	29	30	25	27	23	46	35	40	**9**
Size of the test set	100	80	60	60	80	100	60	80	100	**11**
Quantity of errors at the testing	12	12	7	9	11	13	9	11	17	**4**
Size of the classified set	-	-	-	-	-	-	-	-	-	**185**
Quantity of errors at the classification	-	-	-	-	-	-	-	-	-	**0**
Sensitivity of the classifier (%)	84.39	85.37	86.34	90.24	91.22	91.22	79.51	83.9	79.02	**97.07**
Specificity of the classifier (%)	96.41	94.36	95.38	92.82	89.74	90.77	93.33	93.33	92.82	**96.41**
Number of errors of the 1-st type	7	11	9	14	20	18	13	13	14	**7**
Number of errors of the 2-nd type	32	30	28	20	18	18	42	33	43	**5**
Development time, s	<1	<1	<1	1	1	1	1	1	1	**2465**

For this data set the group of 9 SVM classifiers was trained (Table VII). Three kernel functions were included in the search: polynomial (# 1), radial basis (# 2) and sigmoidal (# 3). In the Table VII for the sigmoid kernel function the first number is k_1, the second number is k_2.

At the first level of the two-level SVM classifier 215 objects were selected from the initial 400 objects. These 215 objects have been identified by the group of the SVM classifiers as the support vectors. Noteworthy, 204 objects were classified correctly and entered in the training set SV^+, and 11 objects were incorrectly classified and entered in the test set SV^-. The time used for the development of one individual SVM classifier is on average less than 1 second.

At the second level of the two-level SVM classifier the SVM classifier on the base of the modified PSO algorithm has been created. A herewith we used the training set SV^+ and the test set SV^-. The search time for optimal parameters amounted to 2465 seconds, that almost 3 times less than the search time for the original experimental data set (7146 seconds).

The remaining 185 objects (more than 46%) were not used in the development of the SVM classifier and compiled the classifying data set. These objects were correctly classified by the two-level SVM classifier.

Fig. 8 shows the classification results of the Test2 data set: on the left – the part of the objects (support vectors) and their separating curve; on the right – the original experimental data set (after the classification of the remaining 185 objects).

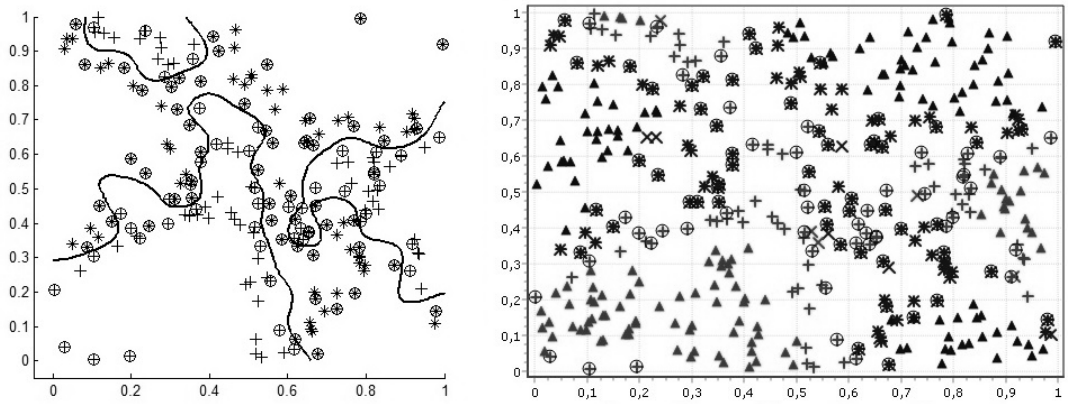


Fig. 8. The classification results of the Test2 data set

TABLE VIII. THE PARAMETERS AND CHARACTERISTICS OF THE INDIVIDUAL CLASSIFIERS (SPAM DATA SET)

Number of the individual SVM classifier	1	2	3	4	5	6	7	8	9	10	**Result**
Kernel function type	1	1	1	2	2	2	2	3	3	3	**1**
Regularization parameter	1	2	2.5	10	12	10	10	0.5	0.5	3.4	**0.1**
Kernel parameters	3	4	3	15	10	15	8.00	0.15; -1.00	0.50; -1.50	0.30; -0.80	**3**
Overall accuracy (%)	89.15	86.66	84.7	94.09	94.39	93.96	91.28	86.09	84.24	81.85	**97.26**
Number of the support vectors	579	561	576	752	760	859	723	515	584	638	**717**
Size of the training set	3681	3911	3451	3221	3681	3681	3681	3451	3911	3451	**1834**
Quantity of errors at the training	364	497	505	178	189	214	315	484	624	625	**27**
Size of the test set	920	690	1150	1380	920	920	920	1150	690	1150	**221**
Quantity of errors at the testing	135	117	199	94	69	64	86	156	101	210	**26**
Size of the classified set	-	-	-	-	-	-	-	-	-	-	**2546**
Quantity of errors at the classification	-	-	-	-	-	-	-	-	-	-	**73**
Sensitivity of the classifier (%)	73.97	69.11	62.82	92.11	89.74	91.89	80.14	76.78	77.88	81.96	**95.37**
Specificity of the classifier (%)	99.03	98.06	98.92	95.37	97.42	95.3	98.53	92.14	88.38	81.78	**98.49**
Number of errors of the 1-st type	27	54	30	129	72	131	41	219	324	508	**42**
Number of errors of the 2-nd type	472	560	674	143	186	147	360	421	401	327	**84**
Development time, s	4	8	7	4	3	3	3	2	2	3	**19599**

During the experiments it was found that the individual classifiers show the accuracy of ranging from 85.75% to 91.5%. The accuracy of the two-level SVM classifier amounted to 96.75%. Thus, using the two-level SVM classifier has improved the classification accuracy by more than 5% compared to the maximum precision of one of the SVM classifiers. The number of objects used in the training and testing of the SVM classifier was reduced from 400 to 215.

Besides, the offered two-level SVM classifier has been used for Spam data set classification. For this data set the group of 10 SVM classifiers was trained (Table VIII). Three kernel functions were included in the search: polynomial (# 1), radial basis (# 2) and sigmoidal (# 3). In the Table VIII for the sigmoid kernel function the first number is k_1 , the second number is k_2 . A herewith we used the SVM classifiers which show the acceptable classification accuracy (more than 80%) under the small number of the support vectors (till 1000).

At the first level of the two-level SVM classifier 2055 objects (that is equal to about 45% of the original experimental data set) were selected from the initial 4601 objects. These 2055 objects have been identified by the group of the SVM classifiers as the support vectors. Noteworthy, 1834 objects were classified correctly and entered in the training set SV^+ , and 221 objects were incorrectly classified and entered in the test set SV^- . The time used for the development of one individual SVM classifier is on average less than 4 second.

At the second level of the two-level SVM classifier we found the best SVM classifier on the base of the modified PSO algorithm with the polynomial kernel function. A herewith $d = 3$ and $C = 0.1$. The search time for optimal parameters amounted to 19566 seconds, that almost 2 times less than the search time for the original experimental data set (44933 seconds).

During the experiments it was found that the individual classifiers show the accuracy of ranging from 85.75% to

91.5%. The accuracy of the two-level SVM classifier amounted to 97.26%. Thus, the two-level SVM classifier improved the classification accuracy by almost 3% compared to the maximum accuracy of one of the SVM classifiers. The number of objects used in the training and testing of the SVM classifier was reduced from 4601 to 2055 (i.e. more than twice).

Thus, the results of experimental studies confirm the efficiency of the offered approaches for Big Data classification.

VII. Conclusion

The efficiency of the suggested approaches has been confirmed by the results of experimental studies.

The SVM classifiers on the base of the modified PSO algorithm allow classifying data with the high classification accuracy.

The modified PSO algorithm allows choosing the best kernel function type, values of the kernel function parameters and value of the regularization parameter within appropriate time expenditures, which turned out to be significantly less than when using the traditional PSO algorithm. The main feature of the modified PSO algorithm is using the «regeneration» of the particles.

The SVM ensembles based on the decorrelation maximization algorithm for the different strategies of the decision-making on the data classification and the majority vote rule allow reducing the accident classification decision received by one classifier, and help to improve the classification accuracy. The shortcomings of some classifiers are compensated by strengths of others classifiers thanks to combination of their results. Classifiers counterbalance the results' accident of each other, finding the most plausible output classification decision. It allows finding the best classification result with minimum classification error.

The two-level SVM classifier also allows improving the classification accuracy within appropriate time expenditures.

Further researches will be devoted to the development of recommendations on the application of the SVM classifiers based on the modified PSO algorithm and their ensembles for the solution of the practical problems, especially for the Big Data classification problems. It is necessary to say that the PSO algorithm and other nature inspired swarm optimization algorithms are very well suited for the distributed architecture and handling of high volume unstructured data in the Big Data analytics.

References

[1] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing Multiple Parameters for Support Vector Machine," Machine Learning, vol. 46 (1–3), pp. 131–159, 2002. http://dx.doi.org/10.1023/A:1012450327387

[2] V. Vapnik, "Statistical Learning Theory," Wiley, New York, 1998.

[3] Lean Yu, Shouyang Wang, Kin Keung Lai, and Ligang Zhou, "Bio-Inspired Credit Risk Analysis. Computational Intelligence with Support Vector Machines," Springer-Verlag Berlin Heidelberg, p. 244, 2008.

[4] J.S. Raikwal, and K. Saxena, "Performance Evaluation of SVM and K-Nearest Neighbor Algorithm over Medical Data set," International Journal of Computer Applications, vol. 50, no. 14, pp. 35–39, 2012. http://research.ijcaonline.org/volume50/number14/pxc3881055.pdf

[5] Y. LeCun, L.D. Jackel, L. Bottou, C. Cortes, at al., "Learning Algorithms for Classification: A Comparison on Handwritten Digit Recognition," Neural Networks: The Statistical Mechanics Perspective, Oh, J. H., Kwon, C. and Cho, S. (Ed.), World Scientific, pp. 261–276, 1995.

[6] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," Lecture Notes in Computer Science, vol. 1398, pp. 137–142, 2005. http://dx.doi.org/10.1007/BFb0026683

[7] Y. Li, K. Bontcheva, and H. Cunningham, "SVM Based Learning System For Information Extraction," Lecture Notes in Computer Science, vol. 3635, pp. 319–339, 2005. http://dx.doi.org/10.1007/11559887_19

[8] M. Oren, C. Papageorgious, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian Detection Using Wavelet Templates," 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 193–199, 1997. http://dx.doi.org/10.1109/CVPR.1997.609319

[9] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An Application to Face Detection," 1977 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 130–136, 1997. http://dx.doi.org/10.1109/CVPR.1997.609310

[10] I. Saha, U. Maulik, S. Bandyopadhyay, and D. Plewczynski. SVMeFC: SVM Ensemble Fuzzy Clustering for Satellite Image Segmentation // IEEE Geoscience and Remote Sensing Letters, 2012, vol. 9, no. 1, pp. 52–55.

[11] S.K. Shevade, S.S. Keerthi, C. Bhattacharyya, and K.R.K. Murthy, "Improvements to the SMO Algorithm for SVM Regression," IEEE Transactions on Neural Networks, vol. 11, no. 5, pp. 1188–1193, 2000. http://dx.doi.org/10.1109/72.870050

[12] E. Osuna, R. Freund, and F. Girosi, "Improved Training Algorithm for Support Vector Machines," 1997 IEEE Workshop Neural Networks for Signal Processing, pp. 24–26, 1997. http://dx.doi.org/10.1109/NNSP.1997.622408

[13] S.V.N. Vishwanathan, A. Smola, and N. Murty, "SSVM: a simple SVM algorithm," Proceedings of the 2002 International Joint Conference on Neural Networks, vol. 3, pp. 2393-2398, 2002. http://dx.doi.org/10.1109/IJCNN.2002.1007516

[14] S. Shalev-Shwartz, Y. Singer, N. Srebro, and A. Cotter, "Pegasos: Primal Estimated sub-Gradient Solver for SVM," Mathematical Programming, vol. 127, no. 1, pp. 3–30, 2011. http://dx.doi.org/10.1007/s10107-010-0420-4

[15] L. Bottou, and C.-J. Lin, "Support Vector Machine Solvers," MIT Press, pp. 1–28, 2007.

[16] H.P. Graf, E. Cosatto, L. Bottou, I. Durdanovic, and V. Vapnik, "Parallel Support Vector Machines: The Cascade SVM," Advances in Neural Information Processing Systems, 17, 521–528, 2005.

[17] O. Meyer, B. Bischl, and C. Weihs, "Support Vector Machines on Large Data Sets: Simple Parallel Approaches," In M. Spiliopoulou, L. Schmidt-Thieme, and R. Janning, editors, Data Analysis, Machine Learning and Knowledge Discovery, Studies in Classification, Data Analysis, and Knowledge Organization, pp. 87–95, 2014.

[18] A. Priyadarshini, and S. Agarwal, "A Map Reduce based Support Vector Machine for Big Data Classification," International Journal of Database Theory and Application, vol.8, no.5 (2015), pp. 77–98. doi: 10.14257/ijdta.2015.8.5.07

[19] G. Cavallaro, M. Riedel, M. Richerzhagen, J. A. Benediktsson, and Antonio Plaza, "On Understanding Big Data Impacts in Remotely Sensed Image Classification Using Support Vector Machine Methods," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8 , issue: 10, pp. 4634–4646, 2015. doi: 10.1109/JSTARS.2015.2458855

[20] P. Yasodha, and N.R. Anathanarayanan, "Analysing Big Data to Build Knowledge Based System for Early Detection of Ovarian Cancer," Indian Journal of Science and Technology, vol 8(14), 2015. doi: 10.17485/ijst/2015/v8i14/65745

[21] P. Rebentrost, M. Masoud, and L. Seth, "Quantum Support Vector Machine for Big Data Classification," Phys. Rev. Lett. 113, 130503, 2014. doi: 10.1103/PhysRevLett.113.130503

[22] D.E. Goldberg, B. Korb, and K. Deb, “Messy genetic algorithms: Motivation, analysis, and first results,” Complex Systems, vol. 3, no. 5, pp. 493–530, 1989. http://www.complex-systems.com/abstracts/v03_i05_a05.html

[23] D.R. Eads, D. Hill, S. Davis, S.J. Perkins, J. Ma, R.B. Porter, and J.P Theiler, “Genetic algorithms and support vector machines for time series classification,” Proc. SPIE 4787, Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation, p. 74, 2002. http://dx.doi.org/10.1117/12.453526

[24] S. Lessmann, R. Stahlbock, and S.F. Crone, “Genetic algorithms for support vector machine model selection,” 2006 IJCNN'06. International Joint Conference on Neural Networks, pp. 3063–3069, 2006. http://dx.doi.org/10.1109/IJCNN.2006.247266

[25] D. Karaboga, and B. Basturk, “Artificial Bee Colony (ABC) Optimization Algorithm for Solving Constrained Optimization Problems,” Proceeding IFSA '07 Proceedings of the 12th international Fuzzy Systems Association world congress on Foundations of Fuzzy Logic and Soft Computing, pp. 789–798, 2007.

[26] Jun Sun, Choi-Hong Lai, and Xiao-Jun Wu, “Particle Swarm Optimisation: Classical and Quantum Perspectives,” CRC Press, p. 419, 2011.

[27] R. Poli, J. Kennedy, and T. Blackwell, “Particle swarm optimization,” Swarm Intelligence, vol. 1, no. 1, pp. 33–57, 2007. http://dx.doi.org/10.1007/s11721-007-0002-0

[28] W. Xun, Y.B. An, and R. Jie, “Application of Parallel Particle Swarm Optimize Support Vector Machine Model Based on Hadoop Framework in the Analysis of Railway Passenger Flow Data In China,” Chemical Engineering Transactions, vol. 46, pp. 367–372, 2015. doi: 10.3303/CET1546062

[29] P.S. Duggal, S. Paul, and P. Tiwari, “Analytics for the Quality of Fertility Data using Particle Swarm Optimization,” International Journal of Bio-Science and Bio-Technology, vol. 7, no.1, pp. 39-50, 2015. doi: 10.14257/ijbsbt.2015.7.1.05

[30] L. Demidova, and Yu. Sokolova, “Modification Of Particle Swarm Algorithm For The Problem Of The SVM Classifier Development,” 2015 International Conference "Stability and Control Processes" in Memory of V.I. Zubov (SCP). pp. 623–627, 2015.

[31] L. Demidova, E. Nikulchev, and Yu. Sokolova, “The SVM Classifier Based on the Modified Particle Swarm Optimization,” International Journal of Advanced Computer Science and Applications, vol. 7, no. 2, pp. 16-24, 2016.

[32] A. Karatzoglou, D. Meyer, and K. Hornik, “Support vector machines in R,” Research Report, WU Vienna University of Economics and Business, Vienna, 2005. http://epub.wu.ac.at/id/eprint/1500

[33] L. Demidova, Yu. Sokolova and E. Nikulchev, “Use of Fuzzy Clustering Algorithms’ Ensemble for SVM classifier Development,” International Review on Modelling and Simulations (IREMOS), vol. 8, no. 4, pp. 446–457, 2015. http://dx.doi.org/10.15866/iremos.v8i4.6825

[34] L. Demidova, and Yu. Sokolova, “SVM-Classifier Development With Use Of Fuzzy Clustering Algorithms' Ensemble On The Base Of Clusters' Tags' Vectors' Similarity Matrixes,” 16th International Symposium on Advanced Intelligent Systems, pp. 889–906, 2015.

[35] L. Demidova, and Yu. Sokolova, “Development of the SVM Classifier Ensemble for the Classification Accuracy Increase,” 6-th Seminar on Industrial Control Systems: Analisys, Modeling and Computing, 2016

[36] L. Demidova, and Yu. Sokolova, “Training Set Forming For SVM Algorithm With Use Of The Fuzzy Clustering Algorithms Ensemble On Base Of Cluster Tags Vectors Similarity Matrices,” 2015 International Conference "Stability and Control Processes" in Memory of V.I. Zubov (SCP), pp. 619–622, 2015.

[37] M.S. Eastaff., and P. Premalatha, “Analysis of Big Data Based On Ensemble Classification,” Proceedings of the UGC Sponsored National Conference on Advanced Networking and Applications (27th March 2015), pp. 191–193, 2015.

Nonlinear Condition Tolerancing Using Monte Carlo Simulation

JOUILEL Naima
High School of engineering ENSAM Meknes
Research Laboratory in Mechanical, Mechatronic & Control – L2MC. Morocco

RADOUANI Mohammed
High School of engineering ENSAM Meknes
Research Laboratory in Mechanical, Mechatronic & Control – L2MC. Morocco

ELGADARI M'hammed
High School of engineering ENSAM Meknes
Research Laboratory in Mechanical, Mechatronic & Control – L2MC. Morocco

EL FAHIME Benaissa
High School of engineering ENSAM Meknes
Research Laboratory in Mechanical, Mechatronic & Control – L2MC. Morocco

***Abstract*—To ensure accuracy and performance of the products, designers tend to hug the tolerances. While, manufacturers prefer to increase them in order to reduce costs and ensure competition. The analysis and synthesis of tolerances aim on studying their influence on conformity with functional requirements. This study may be conducted in the case of the most unfavorable configurations with the "worst case" method, or "in all cases" using the statistical approach. However, having a nonlinear condition make it difficult to analyse the influence of parameters on the functional condition.**

In this work, we are interested in the tolerance analysis of a mechanism presenting a nonlinear functional condition (slider crank mechanism). To do this we'll develop an approach of tolerances analysis combining the worst case and the statistical methods.

Keywords—Worst case tolerancing; statistical tolerancing; Monte Carlo simulation; nonlinear condition; slider crank system

I. INTRODUCTION

The industrial challenges lead to develop models and design support tools to meet customer needs by optimizing the triptych: cost quality and time. Several works have been conducted in order to analyze the impact of deviating component surfaces on functions, properties and assemblability of the product [1, 2, 3].

Mechanical products tolerancing consist on specifying limits of dimensional characteristics variations (often unidirectional) with a tolerance interval. There are mainly two tolerancing approaches:

- "worst case", ensures the assembly and functionality of the mechanical system, but leads to a higher production comparing to the second approach;
- "statistic", ensures a low cost of production but accepts mechanisms whose functionality is not respected.

The tolerance analysis aims to study the influence of dimensional characteristics variations on the respect of functional requirements. This analysis may be conducted in the case of the most unfavorable configurations "worst case" or "in all cases" using the statistical approach. The analysis is so difficult to be performed in the case of the nonlinear functional conditions, because we cannot write them as an algebraic combination [4]. Our work aims to propose an approach based on the combining of the worst case and the statistical methods that allows the tolerance analysis and synthesis in the case of a nonlinear functional conditions.

This paper is organized as follows: In section 2 we treat tolerance analysis at worst case and we present its limits on tolerancing nonlinear condition. Section 3 describes the statistical tolerancing method using Monte Carlo Simulation. In section 4, we apply our approach on a slider crank system that presenting a nonlinear functional condition.

II. TOLERANCE ANALYSIS "AT WORST CASE"

The analysis at "Worst Case" (WC) is mainly used to operate the system in the most adverse conditions [5, 6, 7]. It allows finding extreme values (maximum and minimum) that can be reached by the resulting operating condition for any combination of the real initial dimensions.

In the case of mono-domain systems, the relationship between the resulting functional requirement and basic characteristics is often translated into a linear additive relationship. Overall, it resulted in a unidirectional relationship as follows:

$$Y = \sum_{k=1}^{n} \alpha_k X_k = \sum_{k=1}^{n} \alpha_k \left(X_{kn} \pm \frac{t_k}{2}\right) \quad (1)$$

Where:

- Y is the functional condition which must be between y_{min} and y_{max};
- X_{kn} is the nominal value of X_k;
- α_k is the influence coefficient ($a_k = \pm 1$);
- t_k is tolerance.

In this case, the functional condition often reaches its limit values only for a linear combination of limit values (minimum and / or maximum) of partial components dimensions such as:

$$y_{min} \leq \left(\sum_{k=1}^{n} \alpha_k X_k\right)_{min} = \sum_{k=1}^{n} \alpha_k X_{kn} - \frac{1}{2}\sum_{k=1}^{n} |\alpha_k|\, t_k$$

$$\left(\sum_{k=1}^{n} \alpha_k X_k\right)_{max} = \sum_{k=1}^{n} \alpha_k X_{kn} + \frac{1}{2}\sum_{k=1}^{n} |\alpha_k|\, t_k \leq y_{max} \qquad (2)$$

The functional condition is then satisfied if:

$$y_{max} - y_{min} = t_y \leq \sum_{k=1}^{n} |\alpha_k|\, t_k \qquad (3)$$

Practically, the relationship between the resulting functional requirement and basic characteristics is more complex than the linear additive relationship.

III. TOLERANCES ANALYSIS WITH MONTE CARLO SIMULATION

Monte Carlo simulation (MCS) is a method for predicting errors manufacturing [8]. Its core idea is to use random samples of parameters or inputs to explore the behavior of a complex system or process. For this, the Monte Carlo simulation uses pseudo-random generators with numbers corresponding to different types of statistical distributions. By using Monte Carlo Simulation, the results obtained are more realistic than those obtained by conventional methods of calculation.

The user must define the random distribution of input variables. The number of experiments generated must be large enough to reliably determine the statistical parameters of output variables.

The simulation defines a statistical data generally described by the mean dimension:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} Z_i \qquad (4)$$

And standard deviation:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_i - \mu)^2} \qquad (5)$$

Where:

- Z_i is the value of the resulting dimension to the 'i_{th}' simulation cycle;
- N is the total number of simulation cycles;

The general approach of applying the MCS method is presented in analyzing functional condition tolerances is presented in Figure 1.

For the evaluation of the realization frequency of the functional condition, the statistical group is converted into a histogram (Figure 2).

The standard deviation 'σ' is a parameter characterizing the dispersion or variation of the values distribution around an average. Higher are the values concentrated around the average, lower is the standard deviation. In a normal distribution, the standard deviation 'σ' is used to establish confidence intervals for desired confidence levels. The production process is often considered satisfactory at ±3σ. So 99.73% of assemblies are in the interval $[\mu_J - 3\,\sigma_J\,;\ \mu_J + 3\,\sigma_J]$ (Figure 3). For centric distribution, the functional requirement will be respected for 99.73% of assemblies if:

Tolerance on the requirement = 6 σ

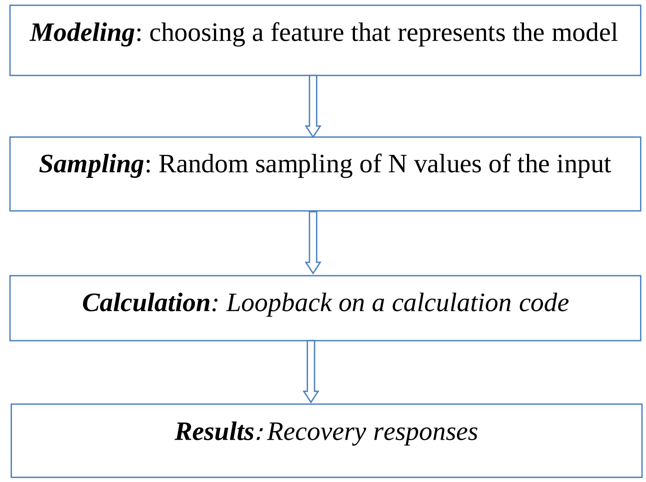


Fig. 1. General approach of applying the MCS

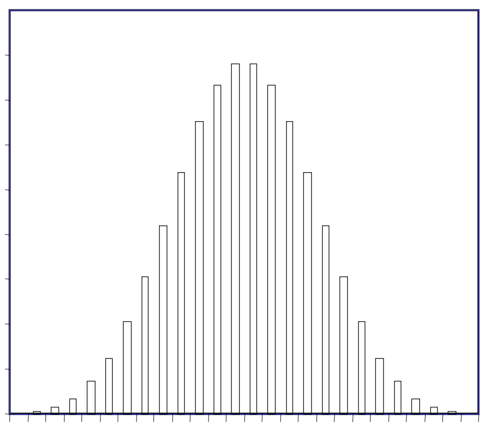


Fig. 2. View as histogram

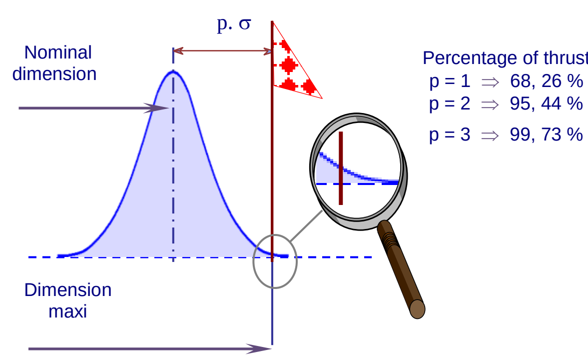


Fig. 3. Percentage of trust in an interval

IV. APPROACH OF TOLERANCING

To solve the problem of the functional condition non-linearity,the following approach is adopted based on both:

Worst Case tolerancing and Monte Carlo Simulation. This combination will ensure system's functionality with large tolerance intervals (figure 4).

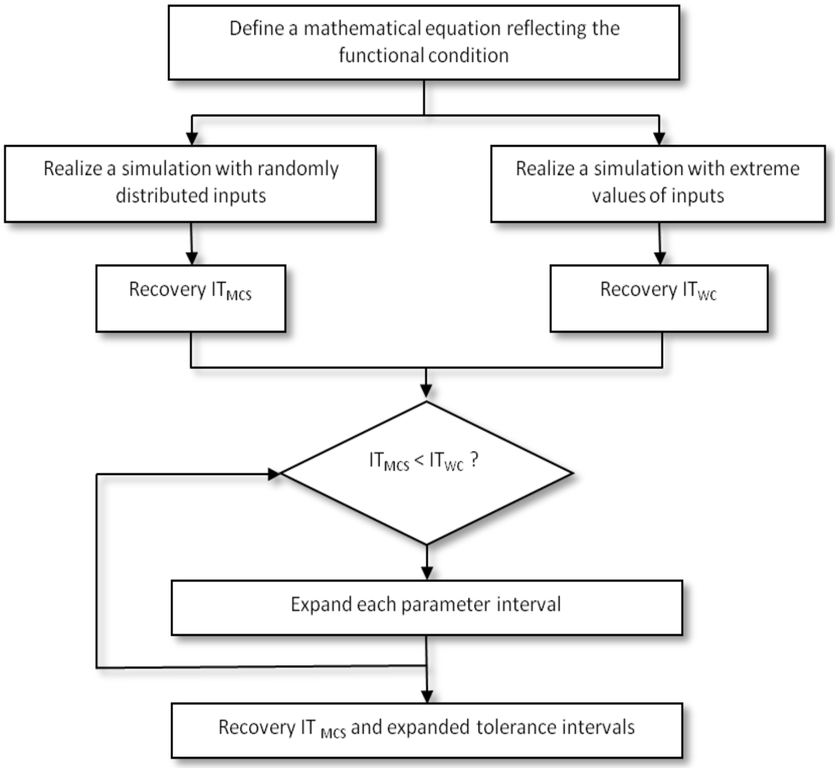


Fig. 4. flowchart of analyzing nonlinear condition tolerance

V. Case of Study

A. System Overview

The case of slider crank system (Figure 4), the position X is given by:

$$X = r\cos(\alpha) + \sqrt{L^2 - (r\sin(\alpha) + A)^2} \quad (6)$$

The maximum 'X_{max}' position of the piston relative to the axis of the crankshaft is illustrated in the simplified diagram below:

$$X_{max} = \sqrt{(r + L)^2 - A^2} \quad (7)$$

Note that the equation reflecting the functional condition ($X_f = X_{max}$) is not linear.

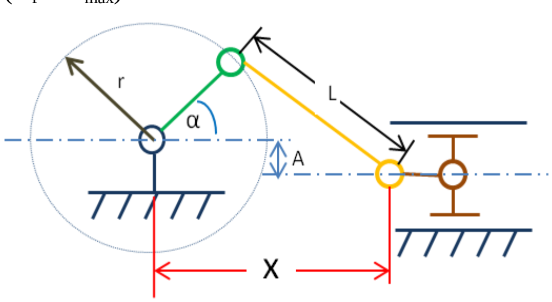


Fig. 5. slider crank system

B. Tolerances analysis «worst case"

We implement the mathematical equation reflecting the functional condition in the 20-sim software [9]. The program translating the mathematical model is:

parameters

```
real L {mm};
real r  {mm};
real A {mm};
real omega {rad/s};
variables
real X {mm};
real Cf {mm};
```

equations

```
X=r*cos(omega*time)+sqrt ((L)^2-(r*sin(omega*time)-  A)^2);
Cf = sqrt ((L+r)^2-A^2);
```

The simulation is launched using the extreme values and then we recover 'L_i' limits (lower limit) and 'L_s' (upper limit) of the functional condition, Then the tolerance interval at "worst case"(IT_{wc}) is deduced.

TABLE I. PARAMETERS AND NOMINAL DISPERSION

Parameters	Nominal values	Li	Ls	IT	σ
L	178	177,300	178,700	±0,70	0,233
r	39	38,300	39,700	±0,70	0,233
A	13	12,300	13,700	±0,70	0,233

TABLE II. RESULTS OF DIMENSIONAL TOLERANCING AT WORST CASE

Parameters	Nominal values	L_i	L_s	IT_{wc}
L	178	177,3	178,7	±0,7
r	39	38,3	39,7	±0,7
A	13	12,3	13,7	±0,7
C_f	216,6103	215,1642861	218,053365	±1,**444539**

It is noted from the results of the simulation that:

$$\mathrm{Li}(\mathbf{C_f}) = \sqrt{\left(Li(\mathbf{L}) + Li(\mathbf{r})\right)^2 - \left(Ls(\mathbf{A})\right)^2} \quad (8)$$

$$\mathrm{Ls}(\mathbf{C_f}) = \sqrt{\left(Ls(\mathbf{L}) + Ls(\mathbf{r})\right)^2 - \left(Li(\mathbf{A})\right)^2} \quad (9)$$

The presented method has made it possible to conduct an analysis of the tolerances at "worst case" where the relationship between the resulting functional requirement and basic characteristics is nonlinear. These results guarantee the mechanism operation with parameters within their limits configurations.

C. Tolerance Analysis by Monte Carlo: statistical approach

The result of the analysis with Monte Carlo simulation depends on the sample size 'N'. The number of draws 'N' has to be high in order to get accurate results. Indeed, the accuracy of this statistical analysis increases as a proportion to '√N' [10].

If the value of N is sufficiently large, the result of the method reaches a stable value, substantially independent of N. In fact, if the difference between two results of simulation in different draws is greater than 5%, then the number of draws N is not sufficient to reach stability [11].

These tests determine the extent of the result variations and calculate its standard deviation (Figure 6). We will use a sample size of N = 10000 to predict statistically the form of 'X_{max}'.

The implementation of the MCS method is to first generate the histograms of Figure 8 based on the values of the table II. We chose tolerance intervals (IT) of ± 0.70 and homogeneous distribution of ± 3σ for each parameter.

The histograms in figure 8 show the variation of each parameter and that of 'C_f' in their tolerance range. They essentially provide information on the behavior of each parameter within its variation area.

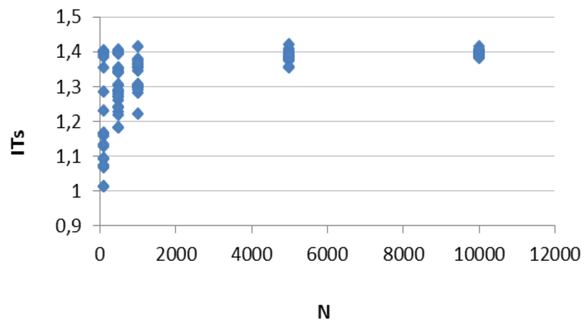


Fig. 6. Evolution of the estimated values of tolerance intervals versus "N"

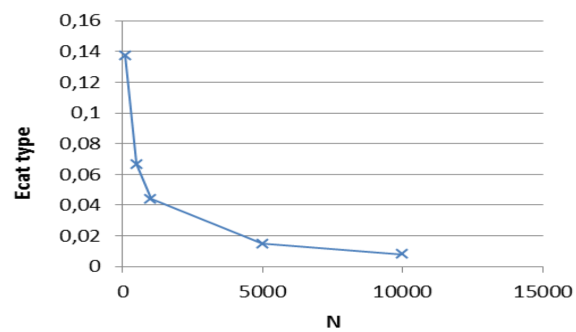
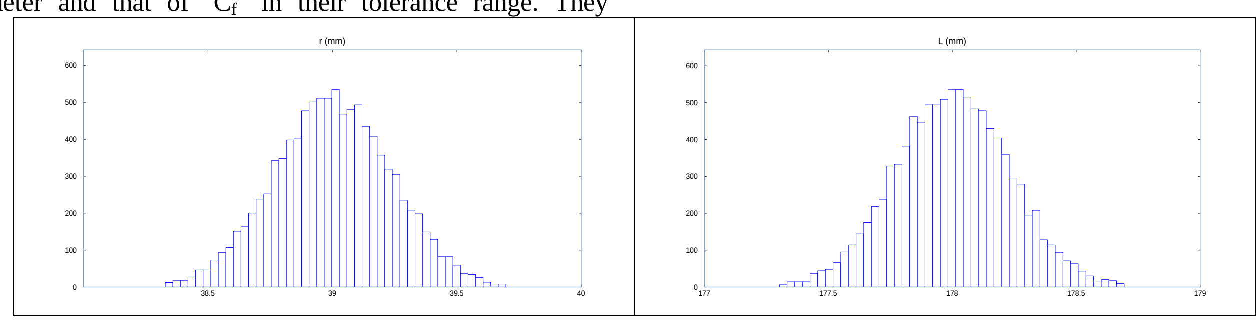


Fig. 7. Evolution of the standard deviation of the estimated tolerance intervals versus "N"

The result of the MCS method is also a random variable. To characterize the random variable, twenty Monte Carlo simulations are launched for each value of "N" (figure8).

The effects of changes in basic parameters on the kinematic behavior of the system are presented in Figure 9.



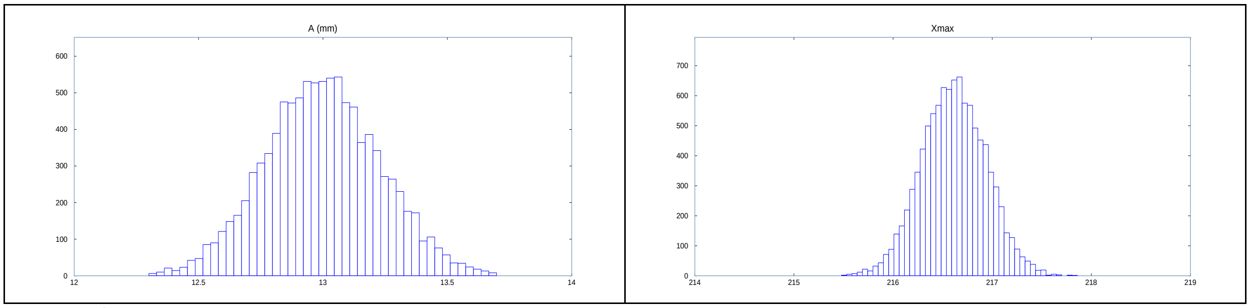


Fig. 8. Variation of the dimensions L, A,r and C_f in their tolerance bands

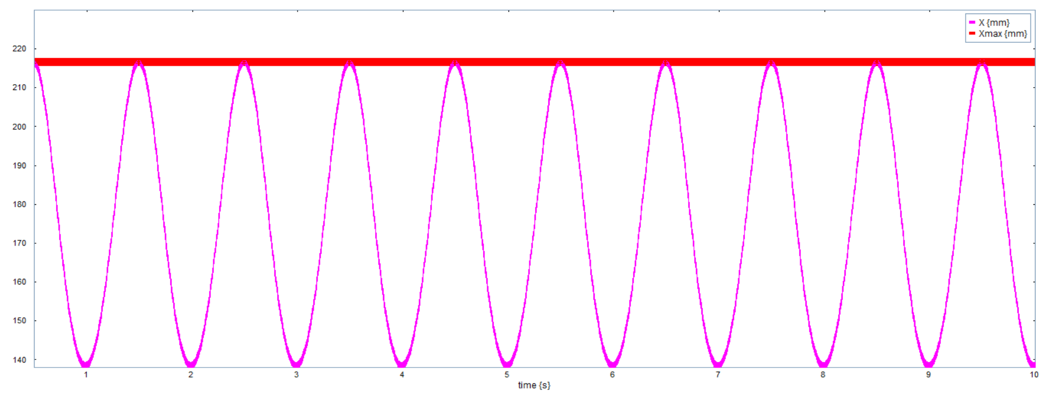


Fig. 9. The effects of changes in parameters X and X_{max}

TABLE III. RESULTS OF DIMENSIONAL TOLERANCING WITH MCS

Parameters	Nominal values	L_i	L_s	IT_{MCS}	σ
L	178	177,300	178,700	±0,70	0,233
R	39	38,300	39,700	±0,70	0,233
A	13	12,300	13,700	±0,70	0,233
X_{max}	216,610249	215,41491	217,75478	**±1,16993**	0,390

By comparing the results obtained by the worst case method to those found with MCS, (Table IV), we find that the statistical variation zone is within the area determined by the arithmetic method WC ($IT_{MCS} < IT_{WC}$).

TABLE IV. COMPARISON OF IT_{WC} AND IT_{MCS}

C_f	L_i (Cf)	L_s (Cf)	IT
$IT_{\_WC}$	215,1643	218,0534	±1,4445
$IT_{\_MCS}$	215,4149	217,7548	±1,1699

D. Tolerances optimization

Since the statistical variation area is included in the arithmetic area, the optimization of parameters IT_{MCS} intervals is possible. For this, we proceed to the expansion of each parameter tolerances intervals, until the new value of the interval IT_{MCS} is acceptable ($IT_{MCS} < IT_{wc}$). The optimization cycle is stopped. For the studied mechanism, the optimization cycle yielded the values shown in the table V.

TABLE V. IT'S OPTIMIZED WITH MCS

Parameters	Nominal values	L_i	L_s	IT_{opt}	σ
L	178	177,15	178,85	± 0,85	0,283
r	39	38,15	39,85	± 0,85	0,283
A	13	12,15	13,85	± 0,85	0,283
C_{f_opti}	216,610249	215,11767	217,978238	±1,43028	0,477

VI. CONCLUSION

In this work, a new method based on Monte Carlo simulation and the Worst Case was proposed to analyze nonlinear condition tolerance interval. Tolerance intervals were expanded for different dimensions without affecting the whole system functionality. This will reduce manufacturing costs while maintaining the overall system functionality.

REFERENCES

[1] F. Charpentier, " Maîtrise du processus de modélisation géométrique et physique en conception mécanique," Thesis, University of Bordeaux, 2014.

[2] P. A. Adragana, « tolérancement des systèmes assemblés, une approche par le tolérancement inertiel et modal », Thesis, University of Savoie, 2007.

[3] V. Moradinaftchali , L. Song, X. Wang "Improvement in quality and productivity of an assembled product: A riskless approach", Computers and Industrial Engineering, ELSEVIER, Vol. 94, pp. 74–82, April 2016.

[4] J. U. Turner, "Tolerances in Computer-Aided Geometric Design," Thesis, Faculty of Rensselaer Polytechnique Institute, 1987.

[5] L. Joskowicz, E. Sacks,V. Srinivasan, " Kinematic tolerance analysis.," Computer-Aided Design, vol. 29, pp. 147-157, 1997.

[6] Y. Ostrovsky-berman,L. Joskowicz, "Tolerance envelopes of planar mechanical parts with parametric tolerances," Computer-Aided Design, vol. 37, pp. 531-544, 2005.

[7] M. Temmreman, "Analyse et synthèse au pire des cas et statistique dans l'environnement CFAO," Thèse de doctorat, LISMMA-ISMACM, Ecole Centrale de Paris, Paris, 2001.

[8] J. U. Turner, " A feasability space approach for automated tolerancing," Journal of Engineering for industry, pp. 341-346, 1993.

[9] 20-sim® home page: http://www.20sim.com/product/bondgraphs.html

[10] R. CVETKO, CHASE, K. W., MALEBY, S. P., "New metrics for evaluating Monte Carlo tolerance analysis of assemblies," Proceedings of the ASME International Mechanical Engineering Conference and Exposition, 1998

[11] N. JOUILEL, M. RADOUANI, M. EL GADARI, B. EL FAHIME, "Mechatronic tolerancing: Bond Graph approach". COMPUSOFT, An International Journal of Advanced Computer Technology, Vol. 5, Issue no 2, pp. 2063-2070, February 2016.

AES Inspired Hex Symbols Steganography for Anti-Forensic Artifacts on Android Devices

Somyia M. Abu Asbeh
Software Engineering Dept.
Princess Sumaya University for Technology
Amman, Jordan

Sarah M. Hammoudeh
Faculty of Medical and Human Sciences
University of Manchester
Manchester, UK

Arab M. Hammoudeh
College of Medicine
University of Sharjah
Sharjah, UAE

***Abstract*—Mobile phones technology has become one of the most common and important technologies that started as a communication tool and then evolved into key reservoirs of personal information and smart applications. With this increased level of complications, increased dangers and increased levels of countermeasures and opposing countermeasures have emerged, such as Mobile Forensics and anti-forensics. One of these anti-forensics tools is steganography, which introduced higher levels of complexity and security against hackers' attacks but simultaneously create obstacles to forensic investigations. In this paper we proposed a new data hiding approach, the AES Inspired Steganography (AIS), which utilizes some AES data encryption concepts while hiding the data using the concept of hex symbols steganography. As the approach is based on the use of multiple encryption steps, the resulting carrier files would be unfathomable without the use of the cipher key agreed upon by the communicating parties. These carrier files can be exchanged amongst android devices and/or computers. Assessments of the proposed approach have proven it to be advantageous over the currently existing steganography approaches in terms of character frequency, security, robustness, length of key, and Compatibility.**

Keywords—Mobile Forensics; Anti-Forensics; Artifact Wiping; Data Hiding; Steganography; AES

I. Introduction

As mobile phones rapidly evolved from communication means to reservoirs of personal information and smart applications [1], they allowed their users to be exposed to increasing dangers and complexities. Consequently, many fields and technologies have been developed as counter-measures to such dangers. One of these fields is the Mobile Forensics, which aims at collecting and analyzing digital evidence to resolve mobile issues. However, on the other side, opposing measures such as Anti-Forensics technologies have been developed to hinder the use of mobile forensics [2]. One of these anti-forensics tools is steganography.

Steganography systems are utilized to embed secret message in hex symbols, image, audio and video files that can only be discovered by the parties informed of the secret key of the steganography chosen algorithm. Thus, steganography introduces a higher level of complexity that would protect against attacks but at the same time create an obstacle for forensic investigations [3].

This paper will be proposing a new steganography approach inspired by the Advanced Encryption Standard (AES) process, a formal encryption method adopted by the National Institute of Standards and Technology of the US Government, and is accepted worldwide. This encryption method was developed and adopted as a replacement of the Data Encryption Standard (DES) method due to the disadvantages it presented.

The AES encryption method is a 128 bits or 16 bytes block cipher that processes a single block of data at a time and encrypts data through several rounds with the aid of an encryption key. During these ten to fourteen rounds, the data is continuously mixed-up and re-encrypted leading to the increase in the security of the hidden data. A single encryption key is used in the AES method with a length of 128 bits (16 bytes), 192 bits (24 bytes), or 256 bits (32 bytes). The same key would be used for both the encryption and decryption processes which known as symmetric encryption, the opposite of the asymmetric encryption observed in other methods as they utilize two different keys, a public and a private key, in the encryption process [4].

This paper will be introducing some of the currently existing anti-forensics approaches and techniques. Thenceforth, the paper will be presenting the new, AES Inspired Steganography (AIS) approach. This method will be utilizing hex symbols for the embedding of the secret message, similarly to our previously proposed HAS approach [5]. This approach would be applied to purposefully created hex symbol carrier files (using HxD for example) and viewed and edited using the WinHex software.

The AIS approach is proposed to have advantages over the currently existing steganography approaches in its capacity, security and robustness. Capacity refers to the maximum amount of that the stego-medium can contain. Security refers to the ability of the approach and stego-medium to maintain the secrecy of the data by eliminating chances of discovery by third parties. Robustness signifies the ability of the stego-medium to withstand modifications without the loss or compromise of its secretly hidden content [6].

The paper will be presenting background information and related work on anti-forensic techniques, artifact wiping, data hiding, and steganography tools and approaches in sections 2 and 3. Then the paper will be elaborating further on anti-forensics steganography in section 4. The description of the newly proposed AES Inspired Steganography (AIS) is presented in section 5 accompanied by the explanation of the implementation process in section 6. Finally the new approach is analyzed and discussed in section 7.

II. Related Work

Data hiding embeds information in carrier files without changing the general content and format of the file. However, encryption leads to general changes observable by eye to the carrier files reducing the security of the file. Therefore, although encryption increases the difficulty of deciphering the secret messages, the evidence of its existence leaves it prone to attacks. Therefore, combining it with steganography could reduce the vulnerability of this method. AES inspired approached have been developed previously and incorporated into the currently used multimedia steganography methods such as image steganography.

In their paper [7] “A Novel Steganographic Scheme Based on Hash Function Coupled with AES Encryption ” (2014), Rinu et al. presented an AES inspired steganography approach in which the textual data to be hidden is encrypted using the AES approach and embedded in a coloured image using hash based algorithm.

Singh and Attri (2015) proposed another AES inspired steganography approach in their paper [8] “Dual Layer Security of data using LSB Image Steganography Method and AES Encryption Algorithm”. In their work they propose an approach in which data would be embedded in carrier files using LSB image Steganography and encrypted using AES-128 bits encryption resulting in a 2 layered protection of the hidden data. However, this approach has found to result in the invalidation of the stego image used as the carrier file.

Another approach utilizing the concepts of the AES and steganography was presented by Goyal and Sharma in their paper [9] “Proposed AES for Image Steganography in Different Medias” (2014). Their approach utilizes the process of the modified AES (consisting of key expansion, sub bytes modification, shift of rows and mix up of columns) in image-audio steganography. However, the same issue of image invalidation is observed to occur upon the application of this method.

Ramaiya et al. (2013) proposed an image steganography technique based on the AES method in their paper [10] “Secured Steganography Approach Using AES”. The text to be hidden was converted into binary representation in their approach and then embedded into the cover image. The method allows for the use of 128 bit block size of text & 128 bits of Secrete key.

III. Anti-Forensics

Anti-forensics (AF) techniques are used to avoid and eliminate the possibility of evidence detection by the mobile forensics tools [1]. AF techniques and tools are continuously and rapidly evolving. Two major types of Anti-Forensic techniques, artifact wiping and data hiding, will be briefly presented next.

A. Artifact Wiping

Artifact wiping, also known as sanitation, overwrites data files from digital devices permanently erasing them. Some artifact wiping tools, including Binary Code (BC) wipe, Eraser, and Pretty Good Privacy (PGP) wipe, target empty and unallocated spaces [11].

B. Data Hiding

Data hiding tools have been developed to secretly embed and hide undiscoverable data through multiple approaches. These approaches include transferring data to other portable storage devices and then wiping the data from the phone; making data “invisible” and concealing their existence; embedding data in multimedia (hex symbols, image, audio and video) files; and altering file extensions.

IV. The Anti-Forensic Steganography

According to [3], "Steganography is the art and science of hiding information in plain sight". Thus, through steganography, a stego-system unknown to third, uninvolved parties can be created to allow for data exchange under extremely secure conditions. Digitally, data hiding techniques are important tools for the utilization of steganography. Through these tools, hex symbols, image, audio and video steganography can be applied. Steganography techniques are generally categorized into Spatial domain and frequency domain.

A spatial domain technique embeds the information to be concealed in the intensity pixels of the carrier multimedia file. The advantage of this category of techniques is their use of the Least Significant Bit (LSB) algorithms to embed the load of data. However, the drawback is that the majority of the LSB techniques are susceptible to attacks. In frequency domain techniques, on the other hand, images are transformed to frequency components by using some techniques, such as Fast Fourier Transform (FFT), Discrete Cosine Transformation (DCT) or Discrete Wavelet Transform (DWT). Thenceforth, the messages are planted and hidden in some or all of the transformed coefficients [12].

In brief, the process of steganography is commenced through an agreement of two parties on a stego-system and a secret key for the embedding algorithm. The accordingly chosen embedding algorithm would be responsible for allocating the carrier files according to their hexadecimal content. The hexadecimals are modified and replaced with the hexadecimals of the secret message to be exchanged by parties involved. This process prevents any third party lacking the knowledge of the secret key and the chosen embedding algorithm from discovering the embedded data or breaching the carrier file contents [3].

In cryptography, sensitive and secret message is stored and transmitted accorss insecure networks while protected from intruding parties access. Created with a secret key, the encrypted data can only be accessed by the meant parties possessing this key which aids in the deciphering the data [13].

In this paper, we developed an approach that combines concepts from steganography and encryption. The secret message in encrypted using an AES-like process and embedded using the hex symbols algorithm steganography proposed in our previous paper [5]. Subsequent encryption steps are applied to the carrier file as well further to increase the security of the hidden data. We call this approach the AES Inspired Steganography (AIS) (figure 1).

V. The Proposed AES Ispired Steganography (AIS)

A. AES Inspired Steganography (AIS) Design

In this paper, we will be introducing a new data hiding and encryption method that we call AES Inspired Steganography (AIS). Through this method we aim at overcoming the problem of changing and invalidating the carrier file observed in traditional encryption methods.

In general, this method consists of multiple steps of encryption applied to the secretly hidden message. The message on the other hand is embedded into a hex symbols carrier file, which is divided into embedding matrices and cipher key matrices according to varied patterns chosen by the communicating parties. Furthermore, encryptions and rearrangements are applied to the hidden data before and after being embedded into the carrier file. The use of such variations in hiding and encrypting the data allows for increasing the security measures of the approach.

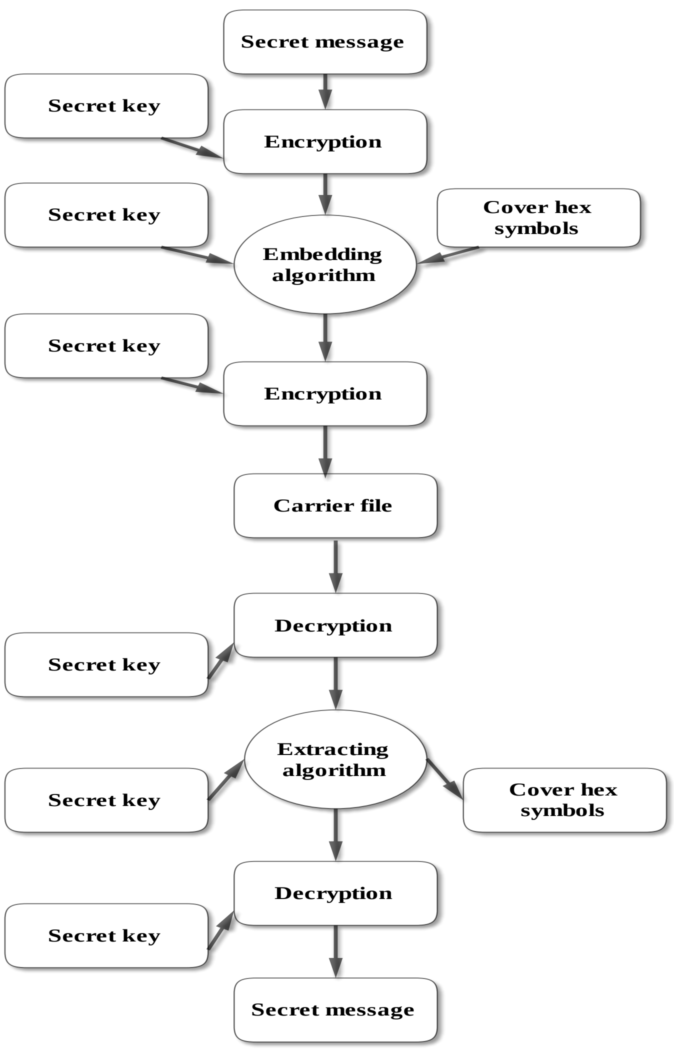


Fig. 1. The general AES Inspired Steganography (AIS) process

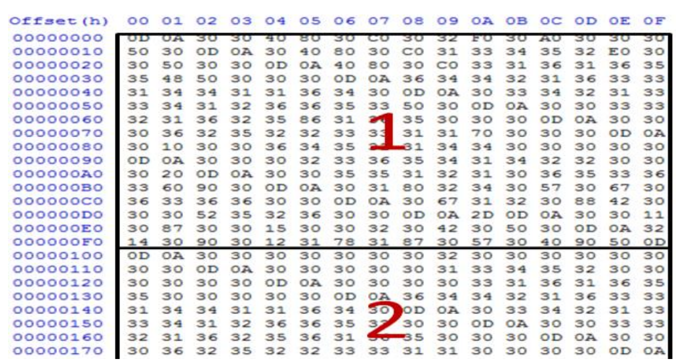
More specifically, the encryption process includes inverting the hexadecimal representation of the of the secret message characters before embedding the message in the carrier file. The embedding process was designed as well to have varying patterns, such that, different combinations of different choices of embedding matrices would be identified to contain the secret message. Furthermore, more specific patterns will be used to identify which characters of the segments would be replaced by the characters of the secret message.

Besides the embedding matrices, cipher key matrices and black segments are included in the matrices divisions causing the decipher process to be even more difficult without a secret key.

After the embedding process, rearrangement of the segments and application of the XOR operation between the cipher key and the embedding matrices further masks the hidden message. These steps are followed by random rearrangements and switching of the rows and columns locations of the matrices. The output version of the carrier file in this case will maintain its validity and integrity. Furthermore, the original content of the carrier file is unfathomable as it mainly consists of random hex symbols. Therefore, the steganography process will not alarm and attract the attention of intruding parties. Additionally, as the file is specifically created for the steganography process, unlike the multimedia files, the possibility of comparisons to originals copy of the file to identify changes is eliminated.

B. AES Inspired Steganography Algorithm (AIS)

First of all the communicating parties (i.e. sender and receiver) agree upon certain patterns that will be used as keys for embedding and extracting of the secret message content. These patterns are created as follows. A carrier file in the form of symbols is created, for example using HxD software, and converted into hexadecimal using WinHex program. The resulting hex symbols file is segmented into 16x16 matrices and numbered as shown in fig.2.



Offset (h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F
00000000	0D	0A	30	30	40	80	30	C0	30	32	F0	30	A0	30	30	30
00000010	50	30	0D	0A	30	40	80	30	C0	31	33	34	35	32	E0	30
00000020	30	50	30	30	0D	0A	40	80	30	C0	33	31	36	31	36	35
00000030	35	48	50	30	30	30	0D	0A	36	34	34	32	31	36	33	33
00000040	31	34	34	31	31	36	34	30	0D	0A	30	33	34	32	31	33
00000050	33	34	31	32	36	36	35	33	50	30	0D	0A	30	30	33	33
00000060	32	31	36	32	35	86	31	[illegible]	35	30	30	30	0D	0A	30	30
00000070	30	36	32	35	32	32	33	[illegible]	31	31	70	30	30	30	0D	0A
00000080	30	10	30	30	36	34	35	[illegible]	31	34	34	30	30	30	30	30
00000090	0D	0A	30	30	30	32	33	36	35	34	31	34	32	32	30	30
000000A0	30	20	0D	0A	30	30	35	35	31	32	31	30	36	35	33	36
000000B0	33	60	90	30	0D	0A	30	31	80	32	34	30	57	30	67	30
000000C0	36	33	36	36	30	30	0D	0A	30	67	31	32	30	88	42	30
000000D0	30	30	52	35	32	36	30	30	0D	0A	2D	0D	0A	30	30	11
000000E0	30	87	30	30	15	30	30	32	30	42	30	50	30	0D	0A	32
000000F0	14	30	90	30	12	31	78	31	87	30	57	30	40	90	50	0D
00000100	0D	0A	30	30	30	30	30	30	30	32	30	30	30	30	30	30
00000110	30	30	0D	0A	30	30	30	30	30	31	33	34	35	32	30	30
00000120	30	30	30	30	0D	0A	30	30	30	30	33	31	36	31	36	35
00000130	35	30	30	30	30	30	0D	0A	36	34	34	32	31	36	33	33
00000140	31	34	34	31	31	36	34	30	0D	0A	30	33	34	32	31	33
00000150	33	34	31	32	36	36	35	[illegible]	30	30	0D	0A	30	30	33	33
00000160	32	31	36	32	35	36	31	[illegible]	35	30	30	30	0D	0A	30	30
00000170	30	36	32	35	32	32	33	33	31	31	30	30	30	30	0D	0A

Fig. 2. The segmented and numbered hex symbols of the carrier file

The segments are then sorted according to the chosen pattern into segments for embedding and segments for the cipher key. Each of these segments are coupled such as each segment used for embedding would be accompanied by a segment for a cipher key. Fig.3 provides an example on the arrangement of the segments in the hex symbols matrices.

Pattern-1	Pattern-2	Pattern-3	Pattern-4	Pattern ...
Cipher keys-C1	Embedding-E1	Cipher keys -C1	Embedding-E1	...
Embedding-E2	Cipher keys-C2	Embedding -E2	Cipher keys-C2	...
Cipher keys-C3	Embedding-E3	Cipher keys-C3	Embedding-E3	...
Cipher keys-C4	Cipher keys-C4	Embedding-E4	Embedding-E4	...
Embedding-E5	Embedding-E5	Cipher keys-C5	Cipher keys-C5	...
Embedding-E6	Cipher keys-C6	Embedding-E6	Cipher keys-C6	...
P1=C1E2C3C4E5E6	P2=E1C2E3C4E5C6	P3=C1E2C3E4C5E6	P4=E1C2E3E4C5C6	

Fig. 3. The division and pattern specification of the embedding and cipher key matrices in the carrier file

From the embedding matrices, specific segments would be chosen to conceal the secret message. Cipher key segments will be as well allocated to each of the chosen embedding segments. The allocation pattern of these chosen segments will be specified in the hiding keys shared between the communicating parties. This paper will be applying pattern 1 (P1) = C1E2C3C4E5E6 as an example. The hex symbol representation of the carrier file is divided into 16×16 matrices which are further divided into16 segments (each forming a 4×4 matrix). These segments are then numbers from 1 to 16.

The segments to be eliminated from the embedding process would be specified as black segments. In our example, matrix E2 includes 4 black segments, 1, 7, 10, and 16. Each of the segments is then to be given a pattern which would be indicated using an alphabet as shown in figure 4.

0D	0A	30	30	30	30	30	30	30	32	30	30	30	30	30	30
30	30	0D	0A	30	30	30	30	30	31	33	34	35	32	30	30
30	30 1	30	30	0D	0A 2A	30	30	30	30 3B	33	31	36	31 4C	36	35
35	30	30	30	30	30	0D	0A	36	34	34	32	31	36	33	33
31	34	34	31	31	36	34	30	0D	0A	30	33	34	32	31	33
33	34	31	32	36	36	35	33	30	30	0D	0A	30	30	33	33
32	31 5D	36	32	35	36 6E	31	36	35	30 7	30	30	0D	0A 8F	30	30
30	36	32	35	32	32	33	33	31	31	30	30	30	30	0D	0A
30	30	30	30	36	34	35	31	31	34	34	30	30	30	30	30
0D	0A	30	30	30	32	33	36	35	34	31	34	32	32	30	30
30	30 9G	0D	0A	30	30 10	35	35	31	32 11H	31	30	36	35 12I	33	36
33	30	30	30	0D	0A	30	31	35	32	34	30	30	30	30	30
36	33	36	36	30	30	0D	0A	30	31	31	32	30	30	30	30
30	30	30	35	32	36	30	30	0D	0A	2D	0D	0A	30	30	30
30	30 13J	30	30	31	30 14K	30	30	30	30 15L	30	30	30	0D 16	0A	30
30	30	30	30	31	31	31	31	31	30	30	30	30	30	30	0D

Fig. 4. A demontration of a chosen pattern (pattern 1_E2) applied to an embedding metrix of the carrier file. The figure indicates the chosen black segments indicated with only numbers while other segments are indicated by a combination of numbers and alphabets

The input secret message is then converted into a hexadecimal representation; therefore each character of its content will be in the form of a two digits hex symbol. These digits representing each character of the secret message are then inverted. For example if the letter 'n' was to be hidden, it will be first converted to the hex code number 63, then inverted to become 36. The resulting inverted hex representations will then be embedded into matrices segments of the carrier file that will have a unique pattern as shown in Fig.4.

Similarly the cipher key matrices will be divided to 16 segments each formed of a 4x4 matrix, which will be numbered from 1 to 16 (figures 9 and 10). Black segments will be chosen as well, in this case, 3, 8, 10, and 15.

The black segments in the cipher key matrix will be relocated to match the locations of those in the embedding matrix to which the cipher key matrix is coupled. For example segment number 3 will be moved forward and relocated prior to segment number 1. Segments number 1 and 2 will then be shifted 1 block foreword each and so on (fig.5).

30	32	F0	30	0D	0A	30	30	40	80	30	C0	A0	30	30	30
C0	31	33	34	50	30	0D	0A	30	40	80	30	35	32	E0	30
30	C0 3	33	31	30	50 1	30	30	0D	0A 2	40	80	36	31 4	36	35
36	34	34	32	35	48	50	30	30	30	0D	0A	31	36	33	33
31	34	34	31	31	36	34	30	34	32	31	33	0D	0A	30	33
33	34	31	32	36	36	35	33	30	30	33	33	50	30	0D	0A
32	31 5	36	32	35	86 6	31	36	0D	0A 8	30	30	35	30 7	30	30
30	36	32	35	32	32	33	33	30	30	0D	0A	31	31	70	30
30	10	30	30	36	34	35	31	31	34	34	30	30	30	30	30
0D	0A	30	30	30	32	33	36	35	34	31	34	32	32	30	30
30	20 9	0D	0A	30	30 10	35	35	31	32 11	31	30	36	35 12	33	36
33	60	90	30	0D	0A	30	31	80	32	34	30	57	30	67	30
36	33	36	36	30	30	0D	0A	30	88	42	30	30	67	31	32
30	30	52	35	32	36	30	30	0A	30	30	11	0D	0A	2D	0D
30	87 13	30	30	15	30 14	30	32	30	0D 16	0A	32	30	42 15	30	50
14	30	90	30	12	31	78	31	40	90	50	0D	87	30	57	30

Fig. 5. Illustration of the re-arrangement of the cipher key matrix to allow the black segments to have superimposing location as thise of the embedding metrix

The hex decimals are then converted to binary representation both the cipher key and the embedding matrices, specifically the decimals representing the secret message content in the embedding segments and those opposing them in the cipher key segments. After the conversion to the binary representations, the decimals are processed using the XOR operation and the resulting binary representation is then converted back to hexadecimal representation. The process is illustrated in table I.

TABLE I. THE XOR OPERATION

Input (hexadecimal)	0A	XOR	47
Convert to binary	00001010	X	01000111
Output(binary)	01001101		
Convert to hexadecimal	4d		

The 16 segments are then randomly rearranged and the order would be described in a table and allocated a certain code such as Random(S) E2 =3,4,1,2,7,8,5,6,11,12,9,10,15,16,13,14.

The contents of these matrices are relocated by exchanging rows with columns in order to increase the difficulty for hackers.

Similarly, the segments of the cipher key matrices would be rearranged randomly such as Random(S) C1=1,3,4,2,8,5,6,7,11,9,12,10,16,13,15,14. The contents of these matrices are relocated by exchanging rows and columns. Then the hexadecimal representation of each of the characters would be inverted.

The hex symbols book is a table like reference shared between the communicating parties with the key information to decipher the encrypted message. This book would include the arrangement of the embedding and cipher key matrices (E# / C#) for a collection of chosen pattern. Furthermore each pattern would be accompanied with information about the black segments, the original allocations of each of the segments before the randomization processes and the patterns (indicated by the alphabet representations) used in each of the segments to indicate the characters containing the concealed hex decimals as shown in table II.

TABLE II. EXAMPLE OF THE SHARED KEY HEX SYMBOL CODEBOOK

Symbols	S	O	M
Pattern - number	P1=C1E2C3C4E5E6	….	…
Embedding segment block number	E2-B = 1-7-10-16 E5-B =3-8-10-15 E6-B =2-5-11-13	….	…
Cipher key segment block number	C1-B=3-8-10-15 C3-B=2-5-11-13 C4-B=1-7-15-1	….	…
Pattern allocated to 4x4 segments	E2-W-1 =ABCDEFGHIJKL E5-W-2 =JKLGHIDEFABC E6-W-3 =DGJAEKBHLCFI	….	…
Random key	C1=1,3,4,2,8,5,6,7,11, 9,12,10,16,13,15,14 . .	……	… …
	E2=3,4,1,2,7,8,5,6,11, 12,9,10,15,16,13,14 . .	…… .	… … …

VI. IMPLEMENTATION

An example of the proposed AIS (AES Inspired Steganography) Approach will be presented in this section in which the secret message "Steganography is the art and science of hiding information in plain sight. EAS is a symmetric encryption" will be embedded in a hex symbols carrier file. The characters of the message are converted to into the hexadecimal representation to begin with. Each letter of the message would be represented by two hexadecimal character components. The two hexadecimals forming each character are then inverted as shown in fig.6, for eample 73 would become 37.

s	t	e	g	a	n	o	g	r	a	p	h	y
37	47	56	76	16	e6	f6	76	27	16	07	86	97
	i	s		t	h	e		a	r	t		a
02	96	37	02	47	86	56	02	16	27	47	02	16
n	d		s	c	i	e	n	c	e		o	f
e6	46	02	37	36	96	56	e6	36	56	02	f6	66
	h	i	d	i	n	g		i	n	f	o	r
02	86	96	46	96	e6	76	02	96	e6	66	f6	27
m	a	t	i	o	n		i	n		p	l	a
d6	16	47	96	f6	e6	02	96	e6	02	07	c6	16
i	n		s	i	g	h	t	E	A	S		i
96	e6	02	37	96	76	86	47	54	14	35	02	96
s		a		s	y	m	m	e	t	r	i	c
37	02	16	02	37	97	d6	d6	56	47	27	96	36
	e	n	c	r	y	p	t	i	o	n		
02	56	e6	36	27	97	07	47	96	f6	e6		

Fig. 6. Secret message hex symbols after inversion

The resulting inverted, hexadecimal representation of secret message's characters is embedded into embedding segment 2 (E2) according to our choice of 'S', which represents pattern-1 (Fig.7). As explained before, the secret message will be embedded after the choice of the black segments and the embedding pattern of the secret message in each of the segments of the embedding matrix (shown in green).

0D	0A	30	30	37	47	56	76	30	32	27	97	30	30	30	30
30	30	0D	0A	16	e6	f6	76	30	31	16	02	35	32	30	30
30	30	30	30	0D	0A	30	30	30	30	07	96	02	47	86	56
35	30	30	30	30	30	0D	0A	36	34	86	37	02	16	27	47
02	16	e6	46	36	56	02	f6	0D	0A	30	33	34	32	31	33
33	02	37	36	36	66	02	33	30	30	0D	0A	30	30	33	33
32	31	96	56	35	36	31	36	35	30	30	30	0D	86	96	30
30	36	32	e6	32	32	33	33	31	31	30	30	46	96	e6	76
02	30	30	30	36	34	35	31	f6	e6	34	30	30	30	30	30
96	e6	30	30	30	32	33	36	e6	02	31	34	32	16	e6	30
66	f6	27	0A	30	30	35	35	02	07	31	30	36	96	02	36
d6	16	47	96	0D	0A	30	31	96	c6	34	30	30	30	30	30
36	33	36	36	30	30	0D	0A	30	27	56	32	30	30	30	30
37	96	76	86	16	02	37	97	0D	96	e6	0D	0A	30	30	30
47	54	14	35	d6	d6	56	47	30	36	36	30	30	0D	0A	30
02	96	37	02	31	31	31	31	31	02	27	30	30	30	30	0D

Fig. 7. Secret message hex symbols after inversion. Illustration of the embedding matrix (E2) after the embedding of the secret message according to 'S' (pattern-1). the embedded secret message is represented by bold green characters while the black segments are shaded with light grey

Similarly, the cipher key matrices are prepared by choosing the black segments and rearranging the segments of the matrix (figure 8). Accordingly, the locations of the black segments in both the embedding and cipher key matrices would be superimposing as explained earlier.

30	32	F0	30	0D	0A	30	30	40	80	30	C0	A0	30	30	30
C0	31	33	34	50	30	0D	0A	30	40	80	30	35	32	E0	30
30	C0	33	31	30	50	30	30	0D	0A	40	80	36	31	36	35
36	34	34	32	35	48	50	30	30	30	0D	0A	31	36	33	33
31	34	34	31	31	36	34	30	34	32	31	33	0D	0A	30	33
33	34	31	32	36	36	35	33	30	30	33	33	50	30	0D	0A
32	31	36	32	35	86	31	36	0D	0A	30	30	35	30	30	30
30	36	32	35	32	32	33	33	30	30	0D	0A	31	31	70	30
30	10	30	30	36	34	35	31	31	34	34	30	30	30	30	30
0D	0A	30	30	30	32	33	36	35	34	31	34	32	32	30	30
30	20	0D	0A	30	30	35	35	31	32	31	30	36	35	33	36
33	60	90	30	0D	0A	30	31	80	32	34	30	57	30	67	30
36	33	36	36	30	30	0D	0A	30	88	42	30	30	67	31	32
30	30	52	35	32	36	30	30	0A	30	30	11	0D	0A	2D	0D
30	87	30	30	15	30	30	32	30	0D	0A	32	30	42	30	50
14	30	90	30	12	31	78	31	40	90	50	0D	87	30	57	30

Fig. 8. Rearrangement of the cipher key matrix in order for the black segments (shaded with light grey) to have superimposing location with those of the embedding matrix

Subsequently, the XOR operation is applied to the matrices as shown in figure 9 to produce the newly encrypted form of the embedding matrices (figure 10).

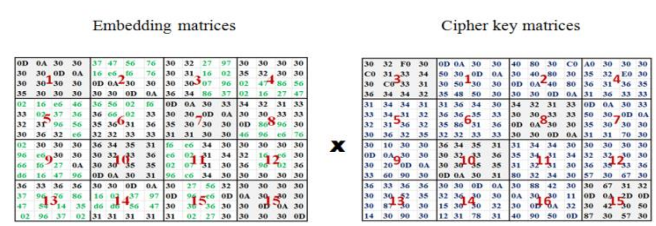


Fig. 9. The application of the XOR operation to the secret message containing characters of the embedding matrix and their opposing characters on the cipher key metrix

0D	0A	30	30	3a	4d	66	46	30	32	17	57	30	30	30
30	30	0D	0A	46	d6	f6	7c	30	31	96	32	35	32	30
30	30	30	30	0D	0A	30	30	30	30	47	16	34	76	b0
35	30	30	30	30	30	0D	0A	36	34	86	3d	33	20	14
33	22	d2	77	07	60	36	c6	0D	0A	30	33	34	32	31
33	36	06	04	36	50	37	33	30	30	0D	0A	30	30	33
32	31	a0	64	35	36	31	36	35	30	30	30	0D	b6	a6
30	36	32	d3	32	32	33	33	31	31	30	30	77	a7	96
32	30	30	30	36	34	35	31	c7	d2	34	30	30	30	30
9b	ec	30	30	30	32	33	36	d3	36	31	34	32	24	d6
56	d6	2a	0A	30	30	35	35	33	35	31	30	36	a3	31
e5	76	d7	a6	0D	0A	30	31	16	f4	34	30	30	30	30
36	33	36	36	30	30	0D	0A	30	af	14	32	30	30	30
07	a6	24	b3	24	34	07	a7	0D	a6	96	0D	0A	30	30
77	d3	24	05	c3	e6	66	75	30	3b	3c	30	30	0D	0A
16	a6	a7	32	31	31	31	31	31	92	77	30	30	30	30

Fig. 10.

Fig. 11. The resulting embedding matrix from the XOR operation

The segments of the new embedding matrix are rearranged randomly (figure 11).

30	32	17	57	30	30	30	30	0D	0A	30	30	3a	4d	66	46
30	31	96	32	35	32	30	30	30	30	0D	0A	46	d6	f6	7c
30	30	47	16	34	76	b0	63	30	30	30	30	0D	0A	30	30
36	34	86	3d	33	20	14	74	35	30	30	30	30	30	0D	0A
0D	0A	30	33	34	32	31	33	33	22	d2	77	07	60	36	c6
30	30	0D	0A	30	30	33	33	33	36	06	04	36	50	37	33
35	30	30	30	0D	b6	a6	30	32	31	a0	64	35	36	31	36
31	31	30	30	77	a7	96	46	30	36	32	d3	32	32	33	33
c7	d2	34	30	30	30	30	30	32	30	30	30	36	34	35	31
d3	36	31	34	32	24	d6	30	9b	ec	30	30	30	32	33	36
33	35	31	30	36	a3	31	36	56	d6	2a	0A	30	30	35	35
16	f4	34	30	30	30	30	30	e5	76	d7	a6	0D	0A	30	31
30	af	14	32	30	30	30	30	36	33	36	36	30	30	0D	0A
0D	a6	96	0D	0A	30	30	30	07	a6	24	b3	24	34	07	a7
30	3b	3c	30	30	0D	0A	30	77	d3	24	05	c3	e6	66	75
31	92	77	30	30	30	30	0D	16	a6	a7	32	31	31	31	31

Fig. 12. Random rearrangement of the embedding matrix

Finally, one more rearrangement is applied to the embedding matrix by interchanging the positions of the rows and columns of the embedding matrix as shown in Fig.12. This is achieved by flipping the whole segment elements around the diagonal as shown in equation 1:

$$x'_{ij} = x_{ji} \tag{1}$$

Where x'_{ij} are the new matrix elements of the stego-file and x_{ji} are the old matrix elements.

30	30	30	36	0D	30	35	31	c7	d3	33	16	30	0D	30	31
32	31	30	34	0A	30	30	31	d2	36	35	f4	af	a6	3b	92
17	96	47	86	30	0D	30	30	34	31	31	34	14	96	3c	77
57	32	16	3d	33	0A	30	30	30	34	30	30	32	0D	30	30
30	35	34	33	34	30	0D	77	30	32	36	30	30	0A	30	30
30	32	76	20	32	30	b6	a7	30	24	a3	30	30	30	0D	30
30	30	b0	14	31	33	a6	96	30	d6	31	30	30	30	0A	30
30	30	63	74	33	33	30	46	30	30	36	30	30	30	30	0D
0D	30	30	35	33	33	32	30	32	9B	56	e5	36	07	77	16
0A	30	30	30	22	36	31	36	30	ec	d6	76	33	a6	d3	a6
30	0D	30	30	d2	06	a0	32	30	30	2a	d7	36	24	24	a7
30	0A	30	30	77	04	64	d3	30	30	0A	a6	36	b3	05	32
3a	46	0D	30	07	36	35	32	36	30	30	0D	30	24	c3	31
4d	d6	0A	30	60	50	36	32	34	32	30	0A	30	34	e6	31
66	f6	30	0D	36	37	31	33	35	33	35	30	0D	07	66	31
46	7c	30	0A	c6	33	36	33	31	36	35	31	0A	a7	75	31

Fig. 13. The final form of the embedding segment after the exhange of the locations of the rows and columns

Rearrangements are applied to the cipher key matrices as well. Initially, a random arrangement is applied (figure 13), followed by interchanging the locations of the rows and columns.

0D	0A	30	30	30	32	F0	30	A0	30	30	30	40	80	30	C0
50	30	0D	0A	C0	31	33	34	35	32	E0	30	30	40	80	30
30	50	30	30	30	C0	33	31	36	31	36	35	0D	0A	40	80
35	48	50	30	36	34	34	32	31	36	33	33	30	30	0D	0A
34	32	31	33	31	34	34	31	31	36	34	30	0D	0A	30	33
30	30	33	33	33	34	31	32	36	36	35	33	50	30	0D	0A
0D	0A	30	30	32	31	36	32	35	86	31	36	35	30	30	30
30	30	0D	0A	30	36	32	35	32	32	33	33	31	31	70	30
31	34	34	30	30	10	30	30	30	30	30	30	36	34	35	31
35	34	31	34	0D	0A	30	30	32	32	30	30	30	32	33	36
31	32	31	30	30	20	0D	0A	36	35	33	36	30	30	35	35
80	32	34	30	33	60	90	30	57	30	67	30	0D	0A	30	31
30	88	42	30	36	33	36	36	30	67	31	32	30	30	0D	0A
0A	30	30	11	30	30	52	35	0D	0A	2D	0D	32	36	30	30
30	0D	0A	32	30	87	30	30	30	42	30	50	15	30	30	32
40	90	50	0D	14	30	90	30	87	30	57	30	12	31	78	31

Fig. 14. Random rearrangement of the cipher key matrix

Finally the hexadecimal representation of each of the characters in the cipher key matrix is inverted (figure 14).

D0	05	03	53	43	03	D0	03	13	53	13	08	03	A0	03	04
A0	03	05	84	23	03	A0	03	43	43	23	23	88	03	D0	09
03	D0	03	05	13	33	03	D0	43	13	13	43	24	03	A0	05
03	A0	03	03	33	33	03	A0	03	43	03	03	03	11	23	D0
03	0C	03	63	13	33	23	03	03	D0	03	33	63	03	03	41
23	13	0C	43	43	43	13	63	01	A0	02	06	33	03	78	03
0F	33	33	43	43	13	63	23	03	03	D0	09	63	25	03	09
03	43	13	23	13	23	23	53	03	03	A0	03	63	53	03	03
0A	53	63	13	13	63	53	23	03	23	63	75	03	D0	03	78
03	23	13	63	63	63	68	23	03	23	53	03	76	A0	24	03
03	0E	63	33	43	53	13	33	03	03	33	76	13	D2	03	75
03	03	53	33	03	33	63	33	03	03	63	03	23	D0	05	03
04	03	D0	03	0D	05	53	13	63	03	03	D0	03	23	51	21
08	04	A0	30	A0	03	03	13	43	23	03	A0	03	63	03	13
03	08	04	D0	03	D0	03	07	53	33	53	03	D0	03	03	87
0C	03	08	A0	33	A0	03	03	13	63	53	13	A0	03	23	13

Fig. 15. The inversion of the hexadecimal representaions of the cipher key matrix characters

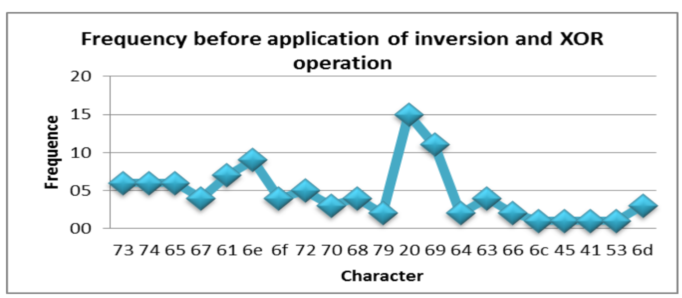
With these final rearrangements the carrier files would be ready for the safe exchange between the communicating parties. With the aid of the secret keys and codebooks shared by the communicating parties, the steps would be retraced to recover the key message.

VII. Analysis and Discussion

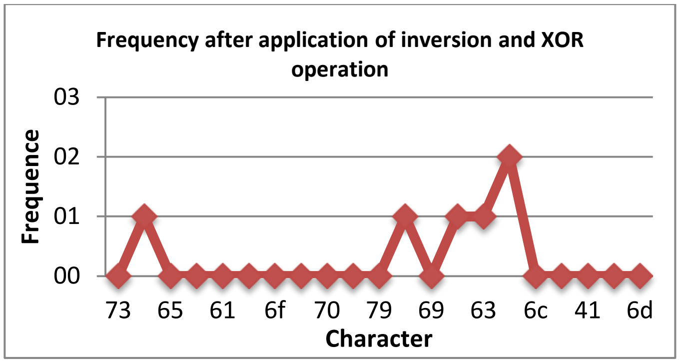
The frequency, degree of safety, robustness, key length, compatibility and capacity of the proposed AIS scheme were analyzed as follows.

A. Frequency

As the two hexadecimal character components of the hex symbol are inverted and processes according to the XOR operation with the cipher key, the calculated frequency of occurrence before and after the application of these changes will differ (fig. 15-a and 1-b).



(a)



(b)

Fig. 16. Character frequency assessment of the embedded message: (a) before and (b) after the inversion of the characters and application of the XOR operation with the superimposing characters of the cipher key matrix

From the compiled frequencies analysis in table III and fig. 16, clear differences were observed in the frequencies of the characters before and after the application of the changes. This observation positively indictes the high level of security against third attacking parties, which is expected to be even further increased with the increase in the length of the secret message.

TABLE III. CHARACTER FREQUENCY BEOFRE (F.B) AND AFTER (F.A) INVERSION AND APPLICATION OF THE XOR WITH THEIR COUPLED CIPHER KEY SEGMENTS

Character	F.B	F.A
73	06	00
74	06	01
65	06	00
67	04	00
61	07	00
6e	09	00
6f	04	00
72	05	00
70	03	00
68	04	00
79	02	00
20	15	01
69	11	00
64	02	01
63	04	01
66	02	02
6c	01	00
45	01	00
41	01	00
53	01	00
6d	03	00

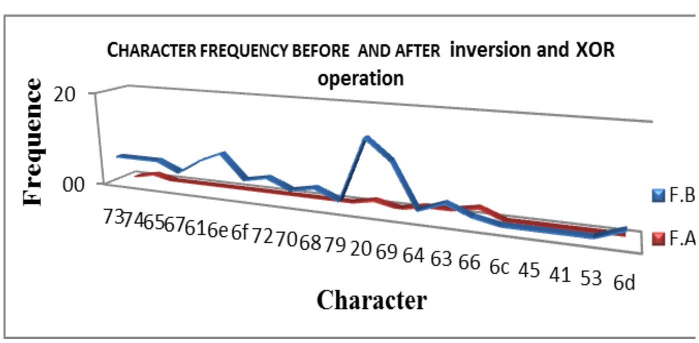


Fig. 17. Character frequency assessment of the embedded message: (a) before (F.B) and (b) after (F.A) the inversion of the characters and application of the XOR operation with the superimposing characters of the cipher key matrix

B. Using WinHex

The use of WinHex to formulate the hex symbols during the hiding process is advantageous as the content will be difficult to trace and compare with previous versions. This advantage becomes more critical as frequent rearrangements of the hex symbols are applied throughout the steganograohy procedure.

A comparison was conducted between the hex symbols in the carrier file (viewed using WinHex) prior to and after the encryption of the secret message. As shown fig.17, a complete change occurred in the hex symbols content of the carrier file, making it impossible for a third party to detect any traces of the message.

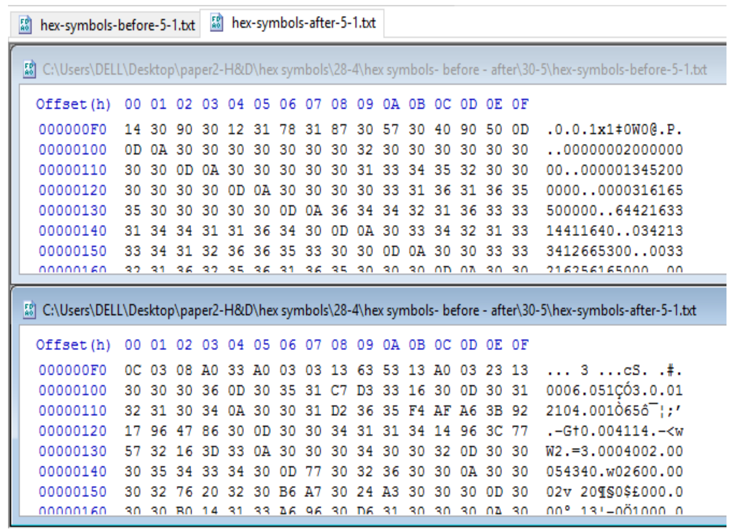
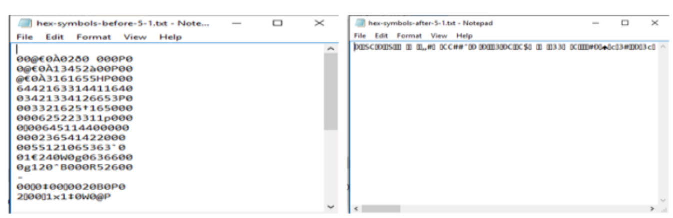


Fig. 18. Comparison between the Hex Symbols before and after the encryption of the secret message

The carrier file content was additionally compared without being viewed using WinHex (Figure 18). The comparison has shown the maintenance of the integrity of the file before and after the encryption process. Such as, the encryption would not cause the invalidity of the file after the encryption as seen in other approaches [8][9].



(a) (b)

Fig. 19. Hex symbols content of the carrier file (a) before and (b) after the inversion of the characters and application of the XOR operation with the superimposing characters of the cipher key matrix

Moreover the alteration of the hex symbols by the inversion of each character element doesn't increase the original file size, leaving it stable and unchanging in terms of elements number. Furthermore, the use of random numbers to select the segments provides an extra complication against deciphering the hidden text. A comparison between available steganalysis tools and hex symbols is presented in Table IV.

TABLE IV. COMPARISON BETWEEN STEGANALYSIS AND HEX SYMBOLS

STEGANALYSIS TOOLS		HEX SYMBOLES
OurSecret OmniHide BDV DataHider Max file encryption Masker StegoStick	These tools work by embedding information within videos by attaching it bluntly to the end of the file EOF [3].	Hex symbols substitutes the hexadecimal precisely on the same position.
OurSecret	This signature can be found after the last byte of the authentic unmodified file.	A valid signature similar to OurSecret does not appear.
OmniHide Pro	White space characters tailing the initial sequence of bytes.	Hex symbols do not show the name of the embedded file.

C. The safety and security against encryption traces

The proposed approach does not require the use of any external encryption and data hiding tools, but rather utilizes tools embedded in usually commonly software such as Microsoft Excel. Therefore, traces of steganography and encryption specified tools on the communicating parties' personal devices would be hard to detect. Therefore, the elimination of such a potent traces source would reduce the risk of alarming the attacking parties. Moreover, the hex symbol file extension can be changed to mislead hackers and investigators.

D. Compression

When tested under compression options such as WinRAR and ZIP file formats, the carrier file has been found to resist changes in size and content. This resistance indicates the robustness of the proposed approach against modification that could be applied to the file which steadily maintaining the file's integrity and content safety.

E. Key length

As shown in table V, in our example, the size of the hex symbols was 1536 bytes and the size of the secret message used was 336 bytes while the size of the cipher key was equally 336 bytes as suggested earlier with regards to the pattern size required for embedding the secret message. Any increase in the required pattern size for embedding the secret message will result in a simultaneous and equal increase in the size of the secret message and cipher key. Therefore, in our approach we have been able to include all the methods of ciphering texts; the transposition (permutation), the substitution and the one-time pad. Achieving the one-time pad is a very significant strength of our new approach, as it was developed to have an equal size for both the cipher key and the secret message regardless of the size of the secret message. Furthermore, the longer the cipher key and the message are, the harder it is to identify the cipher key by intruding parties. Moreover, the use of the hex symbols carrier file and hexadecimal representation of its content allows for a higher embedding capacity in comparison with the use of binary representation.

TABLE V. LENGTHS OF THE EMBEDDING AND CIPHER KEY SEGMENTS, THE HEX SYMBOLS IN THE CARRIER FILE, THE SECRET MESSAGE AND THE CIPHER KEY

Segment for embedding		**embedding**		**Segment for cipher key**		**Cipher key**	
name	Size/ byte	byte	bit	name	Size/ byte	byte	bit
E1	256	96	768	C1	256	96	768
E2	256	120	960	C2	256	120	960
E3	256	120	960	C3	256	120	960
	768	336	2688		768	336	2688

Hex symbols		**embedding**		**Cipher key**	
byte	bit	byte	bit	byte	bit
1536	12288	336	2688	336	2688

F. Compatibility & Capacity

The use of Hex symbols was very compatible with the use of the AES concept as both approaches are based on the use of the hexadecimal representation of the content or target text. Such compatibility allows for coherence and flexibility as well as high storage capacity due to the use of the hexadecimal representation in comparison with the methods utilizing binary representations. The use of the hexadecimal representation has the advantage of the higher robustness in comparison to the binary representation as the writing and modification of the hexadecimal representation is relatively easier. Furthermore, basing the approach on the hexadecimal representation reduced the length of the code needed in comparison to the binary representation.

In this proposed approach we have added several steps to increase the complexity degree of the hidden message. First, the approach chooses certain segments and hexadecimal characters to embed in while leaving some without embedded character. Therefore, upon the continuous rearrangement of the matrices, the location of the secret message characters would be hard to identify. Second, we added the idea of using random black segments to increase the complexity of the encryption. Finally, we inverted the hexadecimal representations of the characters at several occasions. The combination of these modifications with the multiple steps of encryption and the steganography process resulted in a very complex system that would only be deciphered through the use of the secret key book.

VIII. CONCLUSION & FUTURE WORK

The AES Inspired Steganography (AIS) approach that we propose in this paper represents a modified, improved version of both the AES and the steganography approaches, as it overcomes the weaknesses of each of the techniques through the strength of the other. The approach utilizes the multi-step encryption idea of the AES in combination with the safe data hiding concept of the steganography to conceal secret messages in hex symbols carrier files. This approach has been proven to have advantages over the currently existing steganography approaches in terms of capacity, safety and robustness.

In the future, this approach can be developed further to increase its complexity. The length of the secret message as well as the cipher key could as well be further modified and increased to increase the capacity of the approach. Furthermore, the approach could be developed to be incorporated into other applications and techniques.

ACKNOWLEDGMENT

We thank prof. Nihad Yusuf – Dean of King Abdullah I Faculty of Graduate Studies and Scientific Research, prof. Arafat Awajan – Dean of King Hussein Faculty for Computing Sciences, and Dr. Saqer Abdel Rahim – Head of the computer science department in King Hussein Faculty for computing Sciences, for their continuous support of this work.

REFERENCES

[1] A. Distefano, G. Me and F. Pace, "Android anti-forensics through a local paradigm," Digital Investigation, vol. 7, pp. S83-S94, August 2010.

[2] K. Dahbur and B. Mohammad “The Anti-Forensics Challenge,” Proceedings of the 2011 International Conference on Intelligent Semantic Web-Services and Applications - ISWSA '11, ACM Press, April 2011.

[3] T. Sloan and J. Hernandez-Castro, “Forensic analysis of video steganography tools,” PeerJ Computer Science, vol. 1, pp. e7, May 2015.

[4] "Introduction to AES Encryption,"2016, Townsend Security. [online]. available: https://townsendsecurity.com/sites/default/files/AES_Introduction.pdf

[5] S. M. Abu Asbeh, H. A. Al-Sewadi, S. M. Hammoudeh, A. M. Hammoudeh, “Hex Symbols Algorithm for Anti-Forensic Artifacts On Android Devices,” International Journal of Advanced Computer Science and Applications (IJACSA), vol. 7, no. 4, April 2016.

[6] S. Sirsikar and A. Deshpande, “Steganographic Tools for BMP Image Format,” International Journal of Computer Science & Emerging Technologies (IJCSET), vol. 2, pp. 200-204, February 2011.

[7] Manoj gowtham.G.V, Senthur.T, Sivasankaran.M, Vikram.M,Bharatha Sreeja.G, “AES BASED STEGANOGRAPHY,” International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 1, January 2013.

[8] Satwinder Singh and Varinder Kaur Attri, “Dual Layer Security of data using LSB Image Steganography Method and AES Encryption Algorithm,” International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 8, No. 5 (2015).

[9] Yojna Goyal , Manmohan Sharma, “PROPOSED AES FOR IMAGE STEGANOGRAPHY IN DIFFERENT MEDIAS,” IJRET: International Journal of Research in Engineering and Technology, Volume: 03 Issue: 10 | Oct-2014.

[10] MANOJ RAMAIYA, NAVEEN HEMRAJANI and ANIL KISHORE SAXENA, “SECURED STEGANOGRAPHY APPROACH USING AES,” International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), Vol. 3, Issue 3, Aug 2013.

[11] P. A. Kotsopoulos and Y. C. Stamatiou, “Uncovering Mobile Phone Users’ Malicious Activities Using Open Source Tools,” Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on, Istanbul, pp. 927-933, August 2012.

[12] K.Dasgupta1, J.K. Mandal and P.Dutta “HASH BASED LEAST SIGNIFICANT BIT TECHNIQUE FOR VIDEO STEGANOGRAPHY(HLSB) “ International Journal of Security, Privacy and Trust Management (IJSPTM), Vol. 1, No 2, April 2012.

[13] Obaida Mohammad Awad Al-Hazaimeh, “A NEW APPROACH FOR COMPLEX ENCRYPTING AND DECRYPTING DATA,” International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.2, March 2013.

Implementation of Novel Medical Image Compression Using Artificial Intelligence

Mohammad Al-Rababah
Department of computer sciences
Community College, Northern Border University, Saudi Arabia

Abdusamad Al-Marghirani
Department of computer sciences
Faculty of Computing and Information Technology,
Northern Border University, Saudi Arabia

***Abstract*—The medical image processing process is one of the most important areas of research in medical applications in digitized medical information. A medical images have a large sizes. Since the coming of digital medical information, the important challenge is to care for the conduction and requirements of huge data, including medical images. Compression is considered as one of the necessary algorithm to explain this problem. A large amount of medical images must be compressed using lossless compression. This paper proposes a new medical image compression algorithm founded on lifting wavelet transform CDF 9/7 joined with SPIHT coding algorithm, this algorithm applied the lifting composition to confirm the benefit of the wavelet transform. To develop the proposed algorithm, the outcomes compared with other compression algorithm like JPEG codec. Experimental results proves that the anticipated algorithm is superior to another algorithm in both lossy and lossless compression for all medical images tested. The Wavelet-SPIHT algorithm provides PSNR very important values for MRI images.**

***Keywords*—Medical image; lossless Compression; lifting wavelets; CDF9/7; Lifting scheme; SPIHT coding**

I. INTRODUCTION

The current trend is the increasing use of digitized medical images [1,2] . Most of modern techniques of medical imaging produce 3D data (MRI, CT , ultrasound, positron emission tomography) and even 4D (functional MRI, 3D echocardiography dynamic). Some images are intrinsically volume while others correspond on the contrary, a succession of images 2D (also called image stack), to which is added a extra dimension , namely the difference between two successive cuts . In fact , the majority of medical images produced nowadays can be shown as images in at least three dimensions[3}.

Medical images are constantly develop by giving a representation increasingly precise parts of the human body . However more image is more accurately the amount of data generated is large[1]. Medical images such as functional MRI , dynamic and tomography image dynamic 3D echocardiography is increasingly used as they are considered among the most effective techniques in medical imaging, but they produce the most voluminous data , hence the need for their compression for storage and / or transport through networks of telecommunication[4]..

Today, the large sizes use of numerical schemes in medical imaging (MRI, X scanner, nuclear medicine, etc.) generates massive volumes of data. However ,Medical imaging play an important role on the diagnosis of diseases and surgical planning. These mega data need efficient transmission and long-term storage and [2]. So it is necessary to use algorithms of compression images in order to reduce the amount of data to be stored and transmitted.

For this reason, this paper is decomposed into three parts: the first part will present a representation of the Lifting scheme, then the second part present the biorthgonal wavelet CDF 9/7 and finally and finally this paper present a SPIHT algorithm for medical image coding. In order to evaluate medical image compression by our algorithm, The PSNR results obtained are compared with the existing techniques namely JPEG codec[6].

This paper is organized as follows: Sect. 2 provides an overview of related works for medical image compression. Section 3 describes the proposed algorithm for medical image compression. In Sect. 4 experimental results are furnished and are compared with other algorithms. Finally, conclusions are drawn in Sect. 6.

II. RELATED WORK

In the medical image compression community [2],many technique for compression used like Lossless JPEG (Joint Photographic Experts Group) and lossless Wavelet [7]. JPEG formats is adopted by the Digital Imaging and Communications in Medicine (DICOM) group in their widely file format, but in the last years the wavelet compression algorithm give more performances than JPEG codec[4. In spite of this, a many image compression researches examine the use of compression for applying to medical images.

Currently, the compression in a radiology department is always performed without loss when it exists. It is performed by standard as lossless JPEG syntactically provided in the medical image DICOM1 format standard. This type of compression with a exact reconstruction of the original image, ensuring data integrity remains the preferred practitioners for obvious reasons of diagnosis. However it provides poor performance in terms of bit rate. The compression ratio (TC) potential vary about 2 to 8 next the information content of the image and the method applied [2,3]. The origin of the preference of doctors for lossless versus lossy compression is, as we said, avoid medical errors associated with poor image reconstruction. Indeed, the main problem of lossy compression for the medical images is because details important could disappear (others might possibly appear). These details are usually difficult to discern structures because they cause small changes in contrast

The standards of this family are known as JPEG and JPEG2000 names. The technique used for encoding still images can be reduced to two main families: method transform and structural engineering (seeking uniformity in the image as : textures, contours , histogram) . The structural approach uses techniques manipulate the intensity value of the pixel in the image. The two large families may be used together to a coding system. In other words, there is no border between Transform approach and structural approach. compression ratio is called the ratio between the sizes of the raw image and the compressed image[2].

III. Proposed Algorithm for Medical Image Compression

In this paper proposed image compression algorithm is simulated using Color transform system and lifting wavelet CDF 9/7 where the compression is done using wavelet decomposition. This is analyzed to get the horizontal, vertical, approximation and diagonal details. The result is a lossless compression image. Wavelets coefficients are coded by the SPIHT algorithm coding[8] .In SPIHT algorithm it requires few bits to capture the same amount of information.

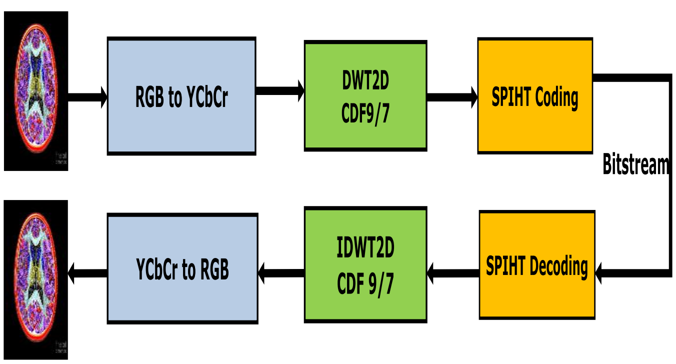


Fig. 1. Medical image Compression scheme

A. Color Space Conversion

According to the color space the image in a transformation of the components necessary to change color representation . The standard offers two modes to transform spend an RGB - type representation (Red, Green , Blue) representation luminance YCbCr chrominance: transformed reversible or irreversible color. The irreversible transform ICT (Irreversible Color Transform) is used in the case of a lossy compression . RCT reversible color transform (Reversible Color Transform) used in the case of lossless compression or lossy , is a approximation of the previous transformed [6].The RGB to YCbCr is performed respecting to eq (1).

$$\begin{aligned} &Y=0.257R+0.504G+0.098B+16 \\ &Cb=-0.148R-0.291G+0.439B+128 \qquad (1) \\ &Cr=0.439R-0.368G-0.071B+128 \end{aligned}$$

However, the inverse transformation is simply expressed by eq (2).

$$\begin{aligned} &R=1.164(Y-16)+1.596(Cr-128) \\ &G=1.164(Y-16)-0.813(Cr-128)-0.391(Cb-128) \quad (2) \\ &B=1.164(Y-16)+2.018(Cb-128) \end{aligned}$$

B. Discrete Wavelet Transforms

The discrete wavelet transform (DWT) is commonly used in the case of images 2D by application of separable filters in both directions [12] . The transformation Two-dimensional discrete wavelet (2 -D DWT) is performed following separately the line order , then the order of the columns. We obtain four sub-bands by level of resolution : an approximation sub- band called LL and three sub-bands details LH , HH and HL respectively representing the horizontal details , and diagonal vertical image . Once the first level of decomposition performed , the transformed may be applied iteratively at each approximation LL sub band obtained for each level of resolution.

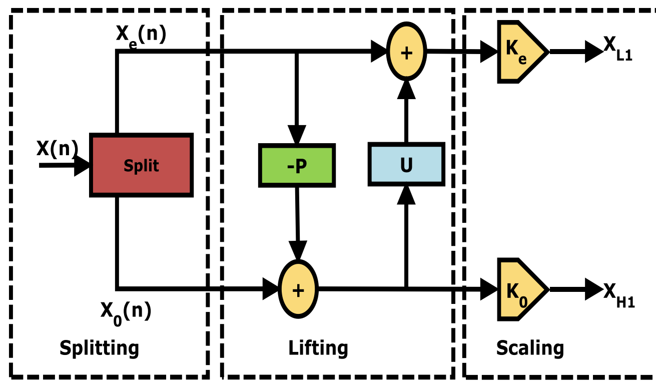


Fig. 2. The Lifting -based wavelet

*1) **Splitting**:* In this state, the original image X is decomposed into two different parts, Xe(n)=X(2n) and $X_0(n)$ =X(2n+1) that present respectively all even-indexed and odd-indexed pixels of image X [6].

*2) **Lifting:*** In this state, using the prediction operation P, X0(n) esteemed from Xe(n) and leads to an wrong signal d(n) that is the exact part of the original signal. Next, To update d(n) we applyed it to the update operation U , and the obtained signal is united with Xe(n) to s(n) estimate, that represents the sleek part of the original signal [5]

*3) **Scaling:*** In this component, a normalization parameter is functional to d(n)and s(n) , correspondingly. In the similar-indexed part s(n) is improved by a normalization parameter Ke to produce the wavelet subband XL1 . In addition, in the odd-indicator part the fault signal d(n) is enhance by K0 to obtain the wavelet subband XH1[14].

C. Biorthogonal Wavelets CDF 9/7

Biorthogonal wavelets 9/7 part of the family symmetrical biorthogonal wavelets CDF. The low-pass filters associated with 9/7 wavelet thus have p = 9 coefficients for the analysis, p = 7 coefficients to the synthesis and are described in Table I. The 9/7 biorthogonal wavelets are shown in Table 1.

TABLE I. A: THE ANALYSIS FILTER COEFFICIENTS

Analysis filter coefficients		
	Low-pass filter	**High-pass filter**
0	0.6029490182363579 0.6029490182363579	+1.115087052457000
±1	+0.266864118442875	+0.591271763114250
±2	-0.078223266528990	-0.057543526228500
±3	-0.016864118442875	-0.091271753114250
±4	+0.026748757410810	

TABLE I. B : THE SYNTHESIS FILTER COEFFICIENTS

Synthesis filter coefficients		
	Low-pass filter	**High-pass filter**
0	+1.115087052457000	0.6029490182363579
±1	-0.591271763114250	-0.266864118442875
±2	-0.057543526228500	-0.078223266528990
±3	+0.091271763114250	+0.016864118442875
±4		+0.026748757410810

Four states are used by the Lifting scheme of the biorthogonal transform 9/7. Two prediction operators and two update operators as presented in Fig.3. [13, 14].

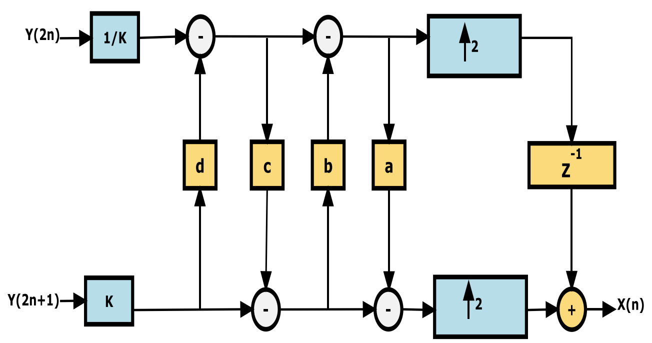


Fig. 3. Lifting implementation of the analysis side of the CDF 9/7 wavelet

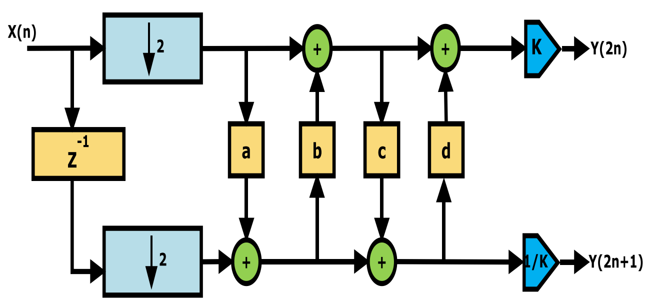


Fig. 4. Lifting implementation of the synthesis side of the CDF 9/7 filter bank

D. SPIHT Coding Scheme

When wavelets coefficients for the decomposition image are obtained, the goal is to find a good algorithm to code the wavelet coefficients into an efficient result. The algorithm adopted in this paper is SPIHT algorithm[7]. The coder SPIHT (Set Partitioning In Hierarchical to Trees) is an efficient algorithm for coding with and without loss the image.The SPIHT algorithm is composed by three steps:initialization, sorting pass and refinement pass. And it contain three lists: LIS list of insignificant sets (LIS), list of insignificant pixels(LIP) and list of significant pixels (LSP). All nodes of the lowest frequency sub band are initialized in LIP stage[14]. In LIP each pixel is compared with the current threshold and a bit (0 or 1) is generated to indicate wich pixel it is significant or not[11]. If a pixel in the offspring set is significant then it is moved to LSP and if it is insignificant moved to LIP and finally the bitsream of bits is generated[7].

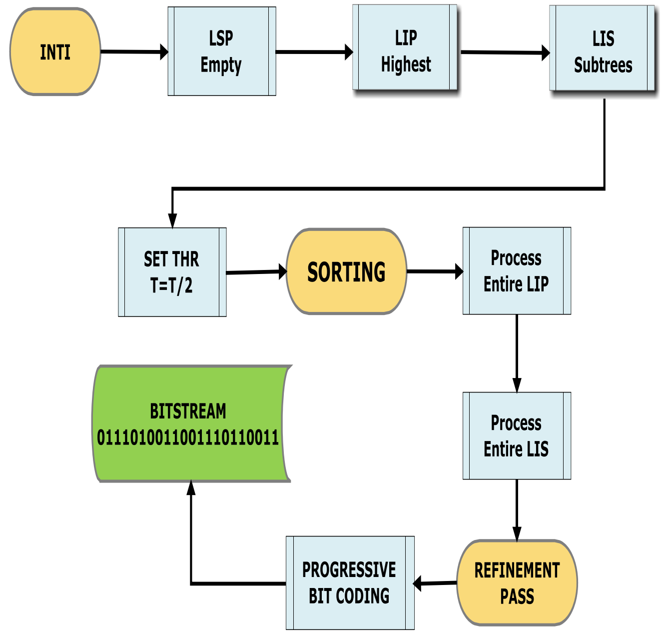


Fig. 5. Flowchart of SPIHT Algorithm

The SPIHT algorithm [2] is based on the principles whileproposing to partition recursively treescoefficients (Figure 5). SPIHT performs partitioningrecursive of the shaft so as to determine the position ofsignificant coefficients in the offspring of the coefficientconsidered. Their sign is sent as soon as they are identifiedas signifiers and they are added to the list ofcoefficients to refine. This algorithm also worksby bit planes . It offers outstanding performance ,EZW without reaching those of entropy coding . By addingan entropy coding significance information , aadditional gain between 0.3 and 0.6 dB is obtained . Bitssent when the significance of the match passesprogram executed in the encoder during the execution ofthe ranking algorithm and significant coefficientsinsignificant.

IV. RESULTS AND DISCUSSION

A. Medical Image standard Test

In this paper for lossless compression methods based on 2D lifting wavelet transforms and SPIHT coding. The simulation is doing on medical image, the spatial location and frequency are important [3, 4]. We applied our proposed algorithm on the tests color medicals images encoded by 24 bits per pixels(bpp).

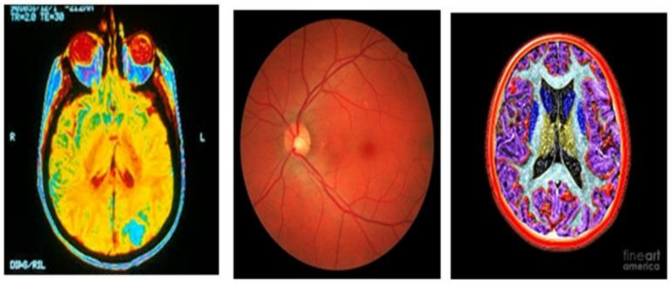


Fig. 6. Originals colors images

The goal of this simulationis to reducing the rates for which the medicals images quality remains acceptable. compressed image quality performances and judgments are given by the PSNR evaluation parameters.

B. Compression Quality Evaluation

PSNR is the most commonly used measure for performance testing methods compression of images with loss. It evaluates the image degradation rebuilt after compression by measuring the error between the original image and the image X rebuilt Y [1] . The unity of PSNR is the decibel dB , the higher it is and the less distortion. PSNR is given by:

$$PSNR = 10\log 10(\frac{(Dynamic \quad of \quad \mathrm{Im}\,ages)^2}{MSE}) \quad (3)$$

For two M × N grayscale images F and $\hat{F}$,Mean Square Error (MSE) which requires, where one of the images is considered as a compression of the other is defined as:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{i=M} \sum_{j=1}^{j=N} (F(i,j) - \hat{F}(i,j))^2 \quad (4)$$

Results

The proposed codec is implemented using MATLAB 2013A for differents color images. This algorithm is simulated on different tests medical images.This image is taken from the GE Medical System database.

This algorithm is implemented with three level of decomposition for Discrete Wavelet transform(DWT). The results of three such images are shown here. Figure 6 present some results for simulation of different ratio (RC) values. According to the quality parameter PSNR,decoded imagegive a good PSNRfrom 0. 5bpp, As a result improvement in performance is observed at lower bit rates.

Fig. 7. PSNR results for medical images test

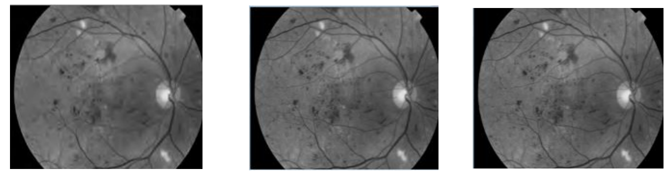


Fig. 8. PSNR results for medical images test

To show the performance of the proposed method, we make a comparison between these different types of transform CDF9/7 (lifting scheme)) coupled with SPIHT coding and JPEG coder, which we calculate the various parameters (PSNR) in order to study the influence of coding choices on color medical images. For each application we vary the bit-rate from 0.05 to 2, and we calculate the PSNR . The results obtained are given in Table II.

TABLE II. COMPARISON WITH EXISTING METHOD (JPEG CODER)

JPEG			**Proposed Algorithm**		
bpp	**CR**	**PSNR(dB)**	**bpp**	**CR**	**PSNR(dB)**
0.05	**160**	**23.5**	**0.05**	**160**	**33**
0.3	**26**	**32.3**	**0.3**	**26**	**34.4**
0.75	**10**	**34.6**	**0.75**	**10**	**42.5**
1	**8**	**41.3**	**1**	**8**	**44.7**
2	**4**	**44.5**	**2**	**4**	**47.9**

V. CONCLUSION

This paper present a fresh algorithm for medical image compression. This algorithm is based on lifting wavelet transforms and SPIHT code. The algorithm provides awfully essential PSNR values for MRI images and it is more appropriate for this type of images. Accordingly, we bring to a close that the results obtained by our compression codec are incredibly agreeable in terms of compressed image quality (PSNR) and compression ratio . In perception, we desire to revise our algorithm for applaying to compress medical image 3D and 4D.

References

[1] T. H. Oh, H. S. Lim, and S. Y. Pang, "Medical image processing: from lossless to lossy compression," in Proceedings of the IEEE Region 10 Conference (TENCON '05), pp. 1–6, November 2005.

[2] M kaur and Wasson V , Region of Interest based Compression techniques for Telemedicine Application, International Journal of Research in Electronics and Computer Engineering , 2015:Vol. 3, pp. 63-67.

[3] BairagiVinayakK., and Ashok Sapkal M. ROI-based DICOM image compression for telemedicine. Sadhana, 2013; Vol. No.38, pp.123- 131.

[4] Devi S. S. &VidhyaK.. Development of medical image compression techniques. In IEEEConference on Computational Intelligence and Multimedia Applications, 2007; Vol. 3, pp. 97-101.

[5] Daubechies, W. Sweldens: Factoring Wavelet Transforms into Lifting Steps, Journal of Fourier Analysis and Applications, Vol.4, No. 3, May 1998, pp. 247 – 269.

[6] P. Wu P., Xie K., Yu H., Zheng Y. and Mao Yu W.A New Preprocessing Algorithm Used in Color Image Compression..Advances in Future Computer and Control Systems. Springer Berlin Heidelberg, 2012;pp.465-471.

[7] S.G. Miaou, S.T. Chen, S.N. Chao: Wavelet-based Lossy-to-lossless Medical Image Compression using Dynamic VQ and SPIHT Coding, Biomedical Engineering: Applications, Basis & Communications, Vol. 15, No. 6, Dec. 2003, pp. 235 – 242.

[8] D.S. Taubman, M.W. Marcellin: JPEG2000: Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers, London, 2002.

[9] G. Pau: Advanced Wavelets and Space-time Decompositions: Application to Video Coding Scalable, Phd thesis, National School of Telecommunications, Paris, 2006.

[10] Savakis, R. Carbone: Discrete Wavelet Transform Core for Image Processing Applications, Real-time Imaging IX, SPIE-IS&T Electronic Imaging, San Jose, CA, USA, SPIE Vol. 5671, Jan. 2005, pp. 142 – 151.

[11] A.M. Lakhdar, R. Méliani, M. Kandouci: Robust Image Transmission Performed by SPIHT and Turbo-codes, Serbian Journal of Electrical Engineering Vol. 5, No. 2, Nov. 2008, pp. 353 – 360.

[12] Sure. Srikanth, Sukadev Meher, "Compression Efficiency for Combining Different Embedded Image Compression Techniques with Huffman Encoding", International conference on Communication and Signal Processing, April 3-5, 2013, India,IEEE.

[13] Salija.P , Manimekalai.M.A.P, N.A Vasanthi, PhD. "ROI and SeamSPIHT based Efficient Image Compression for Mobile Multimedia and Medical Applications", International Journal of Computer Applications (0975 – 8887) Volume 64– No.12, February 2013.

[14] K. Kannan, S.A. Perumal, K. Arulmozhi: Optimal Decomposition Level of Discrete, Stationary and Dual Tree Complex Wavelet Transform for Pixel based Fusion of Multi-focused Images, Serbian Journal of Electrical Engineering Vol. 7, No. 1, May 2010, pp. 81 – 93.

EDAC: A Novel Energy-Aware Clustering Algorithm for Wireless Sensor Networks

Ahmad A. Ababneh
Electrical Engineering Department
Jordan University of Science & Technology
Irbid, Jordan

Ebtessam Al-Zboun
Electrical Engineering Department
Jordan University of Science & Technology
Irbid, Jordan

***Abstract*—Clustering is a useful technique for reducing energy consumption in wireless sensor networks (WSN). To achieve a better network lifetime performance, different clustering algorithms use various parameters for cluster head (CH) selection. For example, the sensor's own residual energy as well as the network's total residual energy are used. In this paper, we propose an energy-distance aware clustering (EDAC) algorithm that incorporates both the residual energy levels of sensors within a cluster radius as well as the distances. To achieve this, we define a metric that is calculated at each sensor based on local information within its neighborhood. This metric is incorporated within the CH selection probability. Using this metric, one can choose the sensors with low residual energy levels to have the greatest impact on CH selection which results in CH selection being biased to be close to these sensors. This results in reducing their communication energy cost to the CH. Simulation results indicate that our proposed EDAC algorithm outperforms both the LEACH and the energy-efficient DEEC protocols in terms of network lifetime.**

***Keywords*—Clustering algorithms; Sensor networks**

I. Introduction

Wireless sensor networks (WSNs) have garnered much attention in the last decade. This is as a result of advances in networking, wireless communication, micro-fabrications, micro-processors, and the wide range of applications [1]. WSN consists of a number of sensor nodes deployed in an area of interest. Sensors collect data and send data to a central processor (e.g. Base Station (BS)). WSNs have important applications such as remote environmental monitoring [3], military applications (e.g., battlefield surveillance) [2], and industrial monitoring applications which include machine health monitoring, and industrial control applications[4]. One of the challenges faces the use of WSN is energy efficiency because it can be difficult (e.g., hazardous areas) to replace the batteries, so the design of energy- efficient network protocols becomes an important issue with respect to network lifetime extension [5]. A major source of energy dissipation is communication between sensors and the base station. To guarantee a good balanced distribution of the energy load between sensor nodes, clustering communication protocols have been designed and implemented.

In WSN, the clustering protocol is a key factor in achieving energy efficiency, so the design of an energy-efficient clustering protocol for WSN is very important. In WSNs the sensor nodes are energy constrained. Therefore, it is very important to find some solutions to offer high scalability and satisfy high energy efficiency to prolong network lifetime. One solution solution is by grouping sensor nodes into sets called clusters. Clustering achieves better lifetime of the sensor network by breaking the sensor network into groups of sensors to conserve communication energy. As a result, saving the energy and increasing the overall lifetime of the network is achieved.

Adopting clustering scheme produces two-level hierarchy; the higher level and the lower level. The higher level is formed by the nodes that are responsible for aggregating and fusing the received data from sensor nodes in the sensing area and then transmit it to a central processor; such nodes are called the Cluster Head (CH) nodes. The lower level of the hierarchy is formed by the nodes that are responsible for detecting the required data from the sensing region and then sending it to the corresponding CH. Each cluster includes number of sensor nodes and one cluster head (CH) [6]. CH selection can be centralized performed by the BS or the end user based on some criterion. It can also be distributed in nature and performed by the sensors themselves on a localized level. The BS is responsible for processing data received from sensor nodes to be used by the end user.

In this paper, we propose a novel distributed energy-efficient cluster head selection algorithm in which two factors are incorporated: the sensors' residual energy levels and the distances between sensors and the CH.

The rest of the paper is arranged as follows. In section 2, a literature review about several clustering algorithms is introduced. In section 3, the network model and the energy expenditure model are adopted. In section 4, we show how the proposed protocol will be used in the process of cluster head selection, and simulation results are explained in section 5.

II. Literature Review

The most important and widely used probabilistic clustering protocols are LEACH [7], HEED [8] and DEEC [9]. In LEACH, the CH is selected using rotation. The selection of cluster heads is based on setting a predefined percentage of CHs for the network. In the LEACH algorithm, each sensor locally calculates a random number and compares it to some threshold that depends on the percentage of CH needed. LEACH performance in homogeneous network (i.e., of the same energy level) is good, whereas in heterogeneous network it is not. HEED is a hierarchical, distributed, clustering algorithm, this algorithm uses both of the remaining energy of the sensor node and the intra-cluster communication periodically in the scheme of cluster head selection. Another algorithm is DEEC which is tailored for use in heterogeneous networks. DEEC uses a probability that is implemented as a ratio between the sensor remaining energy and the network average energy. Nodes with high initial and residual energy will be elected to be CHs with higher probability than that with low initial and residual energy in the network. In [10], SEP is proposed for the two-level heterogeneous networks, where the two-level heterogeneous network includes two types of sensor nodes; the normal nodes and the advanced nodes, in this protocol the process of cluster head selection consists of rounds, the decision of being a cluster head or not is made by the sensor node at each round based on its initial energy relative to that of other nodes.

III. System Model and Energy Expediture Model

In this section, we will introduce the system model and then the energy expenditure model will be explained.

A. System model

For simplicity, we model the sensing area as a 2-D square area of dimension M. A group of N static sensor nodes is dispersed randomly as shown in Figure (1), [7]. The network is divided into clusters, and the CHs transmit the fused and the aggregated data to the BS which is located at the center of the sensing region.

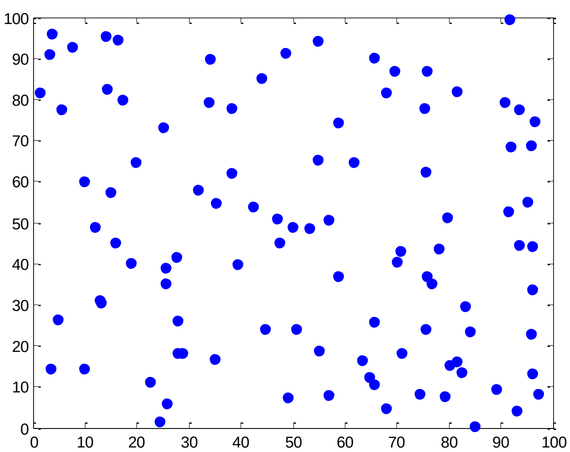


Fig. 1. 100-Nodes Random Networks

For the two-level heterogeneous networks, the sensor nodes are classified into advanced nodes and normal nodes, in which two different levels of energy are assigned. Let m denote the fraction of advanced nodes in the network with initial energy E_i, and is provided with a times excessive energy compared to the normal nodes then the network will contain mN advanced nodes that are supplied with $E_i(1+a)$ initial energy, and $(1-m)N$ normal nodes supplied with E_i initial energy. Thus, the total initial energy of the network in this case can be obtained as:

$$E_{tot} = N(1-m)E_{ini} + Nm(1+a)E_{ini} = NE_{ini}(1+am) \quad (1)$$

On the other hand, it is possible to equip the sensor nodes with multi-level of energy; in this case, the advanced node s_i is provided with initial energy of $E_i(1+a_i)$ which is a_i times more energy compared to the initial energy of the normal nodes. We note that a_i can be a random quantity. Thus, the total initial energy of the network can be written as:

$$E_{tot} = \sum_{i=1}^{N} E_i(1+a_i) = E_i\left(N + \sum_{i=1}^{N} a_i\right) \quad (2)$$

The following assumptions are held:

- Sensor nodes are aware of their locations.
- Communication channel is symmetric, between CH and corresponding sensors
- Single hop communication between sensors and their CH.

B. Energy Expenditure Model

Communication between sensor nodes dissipates most of its energy depending on the distance between the sending and receiving sensor nodes. We use the 1st order radio frequency energy consumption model to describe the energy consumption for sensor nodes [11] which incorporates both free-space and multi-path energy loss. According to this model, shown in Figure (2) [7,12], the energy spent for transmitting $l-bit$ data message to a sensor at distance d is given as

$$E_{TX}(l,d) = \begin{cases} lE_{elec} + \epsilon_{fs}d^2 & d < d_0 \\ lE_{elec} + \epsilon_{mp}d^4 & d > d_0 \end{cases} \quad (3)$$

Where lE_{elec} is the electronic energy which is the energy dissipated to operate the transmitter or the receiver circuit to process one bit, ϵ_{fs} and ϵ_{mp} are the amplifier parameters of transmission corresponding to the free-space model and multi-path respectively, $d_{0=}\sqrt{\epsilon_{fs}/\epsilon_{mp}}$ denotes the threshold distance, and d is the distance between sensor node s_i and sensor node s_i and given as:

$$\left|d_{\{i,j\}}\right| = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \quad (4)$$

Similarly, the energy consumed to receive this message is given as

$$E_{RX}(l) = l \times E_{elec} \quad (5)$$

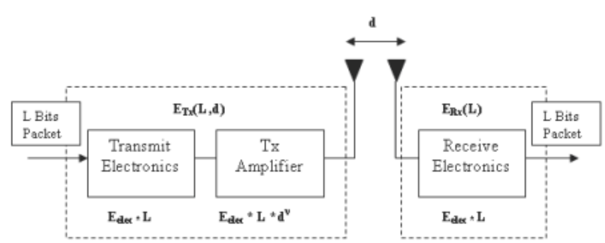


Fig. 2. Radio Energy Dissipation Model

IV. Proposed Clustering Protocol

In this section, we will introduce our proposed clustering protocol called the EDAC. The goal of this protocol is to incorporate the residual energy of sensor node with the cluster head selection process. In contrast to other clustering protocols that only incorporate the residual energy of sensor nodes, we also incorporate on energy distance measure in the selection process.

We first define our distance-energy metric and then explain the steps of our protocol.

A. Distance-Energy Metric

It is necessary to propose a metric that quantifies how good a sensor node could be as a CH. This metric needs to take into account both the residual energy of sensor nodes in addition to the energy expenditure in transmitting data in intra-cluster communication. As noted from energy model, the energy expended in intra-cluster communication is proportional with distance, therefore it is preferred for a CH to be as close as possible to the sensor nodes in its cluster radius. Moreover, sensor nodes with a low residual energy should have more impact in the CH selection process.

The sensor nodes residual energy can be calculated during a single frame/epoch as

$$E_{res(i)} = E_i - E_{diss(i)} \tag{6}$$

Where E_i is the sensor node energy for the current round, and $E_{diss(i)}$ is the energy dissipated in the n^{th} sensor node. For simplicity, assume that the maximum distance between any sensor node and the BS is $< d_0$, the CH dissipated energy will be:

$$E_{CH} = LE_{elec}\left(\frac{N}{k}\right) + LE_{DA}\left(\frac{N}{k}\right) + LE_{elec} + L\epsilon_{fs}d_{to\ BS}^2 \tag{7}$$

Where k is the number clusters, E_{DA} is the consumed energy when processing a bit per signal and $d_{to\ BS}^2$ represents the distance from the CH to the BS. Where the dissipated energy in the non-CH node is given by the following formula:

$$E_{non-CH} = LE_{elec} + L\epsilon_{fs}d_{to\ CH}^2 \tag{8}$$

Where $d_{to\ CH}^2$ represents the distance from the sensor node to its CH. CH needs energy to receive the sensed data from the sensor nodes within its cluster, to aggregate and to transmit these data to the BS. Sensor node dissipates energy only when transmitting its sensed data to its CH. Thus, the total energy dissipated in the cluster during a round can be obtained as:

$$E_{Cluster} = E_{CH} + \left(\frac{N}{k}\right)E_{non-CH} \tag{9}$$

And the total dissipated energy in the network is equal to:

$$E_{round} = L\left(2NE_{elec} + NE_{DA} + k\epsilon_{mp}d_{to\ BS}^4 + N\epsilon_{fs}d_{to\ CH}^2\right) \tag{10}$$

Taking into account the uniform distribution of nodes in the network, the following can be obtained [11], [13]:

$$d_{to\ CH} = \frac{M}{\sqrt{2\pi k}}\ ,\ \ d_{to\ BS} = 0.765\frac{M}{2} \tag{11}$$

Now, the optimal number of cluster heads can be found by differentiating E_{round} with respect to k and equating to zero.

$$k_{opt} = \sqrt{\frac{N}{2\pi}}\sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}\frac{M}{d_{to\ BS}^2} \tag{12}$$

B. The Proposed Protocol(EDAC)

The main idea behind our proposed protocol is to incorporate not only the sensor's own residual energy but also the residual energy levels of sensors within its clustering radius. Moreover, we incorporate the distance between a sensor and the nearby sensors. The lower the residual energy level of a sensor, the more important it becomes to reduce communication energy by placing the CH close to it. This will have the effect of reducing overall energy consumption in the network and extending its lifetime.

The first step in the proposed protocol is to calculate the sensor node residual energy and the distance to the sensor nodes, then the i^{th} sensor node calculates the clustering weight as

$$w_s(i) = \sum_{j\in I_i}\left(\alpha\, R\_E_{res(j)}d_{im} + (1-\alpha)\frac{(1 - R\_E_{res(j)})}{d_{ij}}\right) \tag{14}$$

Where, summation is taken over all sensors in the set I_i which is the set of sensors within the clustering radius of the i-th node and d_{ij} is the distance between sensor node i and sensor node j. The quantity $R\_E_{res(j)}$ denotes the relative residual energy of the node and is given as

$$R\_E_{res(j)} = \frac{E_{res(j)}}{E_{ini(j)}}$$

Where $E_{ini(j)}$ and $E_{res(j)}$ are the initial and residual energy levels of the j-th node.

We note that depending on the value of the parameter $\alpha \in [0,1]$ in Eqn.(14) we can determine the weight $w_s(i)$ value. For example, a large value of α would give more importance to how much residual energy is left in the node whereas a larger value would give more importance to how much energy has the node already spent.

After calculating the weight of the sensor node, the sensor nodes transmit these data to all other sensor nodes in the cluster. In the next step, all sensor nodes use the received data to calculate the mean weight within their cluster radius as

$$w_c(i) = \sum_{j\in I_i} w_s(j) \tag{15}$$

where, $|I_i|$ is the cardinality of the set I_i.

We then propose modifying the CH selection probability associated with the i-th sensor to be as follows

$$p_{si} = p_s\frac{w_s(i)}{w_c(i)} \tag{16}$$

Then in our EDAC protocol, we apply a similar approach as in other probabilistic clustering methods where the i-th node calculates a threshold $T(i)$ given as

$$T(i) = \begin{cases} \dfrac{p_{si}}{1 - p_{si}\left(r \bmod \dfrac{1}{p_{si}}\right)}, & i \in G \\ 0 \quad , & else \end{cases} \qquad (17)$$

where, G denotes the set of eligible cluster heads. The i-th node then generates a random number between [0,1] and compares the generated number against $T(i)$. If the number is less than (i) , then the node becomes a cluster head. The steps of our proposed algorithm are explained in Table 1.

TABLE I. THE PROPOSED CLUSTERING PROTOCOL (EDAC)

The proposed clustering protocol (EDAC)
• For $r = 1 : r_{max}$; r_{max} : maximum number of rounds
• For $i = 1 : N$, N is the index for sensor node
• Find set I_i (sensors in cluster radius)
• Calculate weight from Eqn.(14)
For $i = 1 : N$
• Calculate cluster weight from Eqn.(15)
• Calculate p_{si} from Eqn.(16)
• Perform CH selection using Eqn.(17)
• Inform all sensor nodes in the cluster
• Cluster formation will begin
• End of current round condition
• Restart new round condition

V. SIMULATION RESULTS

In the following experiments, we compare the performance of our proposed algorithm versus that of both the LEACH and DEEC algorithms with emphasis on the network lifetime.

We consider a wireless sensor network with $N = 100$ sensor nodes randomly distributed in an area with dimensions of $100m \times 100m$. We assume the BS is in the center of the sensing region. Table (2) summarizes the radio parameters used in simulations. The proposed protocol is compared with LEACH and DEEC protocols.

TABLE II. PARAMETERS USED SIMULATIONS

Parameter	value
E_{elec}	500nJ/bit
ϵ_{fs}	10PJ/bit/m$_2$
ϵ_{mp}	0.0013 PJ/bit/m$_4$
E_0	0.5J
d_0	70m
Message size	4000 bit
p	0.1

In the first experiment, we consider a two level heterogeneous network with normal nodes having an initial energy level of E_0 and advanced nodes having an energy level of $2E_0$. We note that we set $\alpha = 0.5$ in our proposed algorithm. We set the percentage of advanced nodes to 0.3. Results are depicted in Fig.(3) and (4) below.

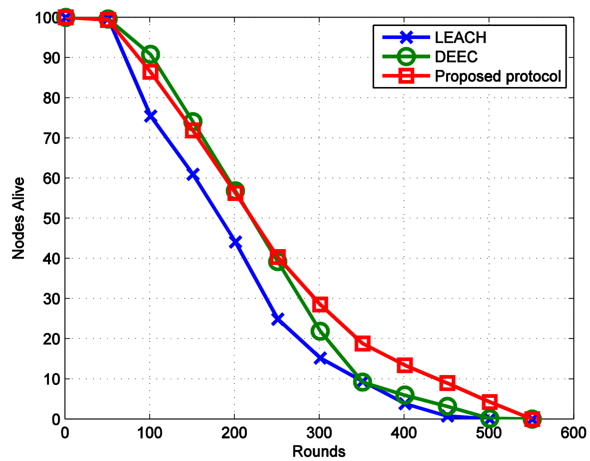


Fig. 3. Alive nodes performance of the LEACH, DEEC and EDAC proposed protocol for a two-level heterogeneous network

From Fig.(3), we can observe the differences in performance between protocols which comes from the different metrics that are used by these protocols in the process of CH selection. It is clear that the stable time of the proposed protocol is comparable to that of the LEACH and DEEC protocols. Both of DEEC and the proposed protocol achieve higher performance than LEACH protocol. Also, we can note that the unstable region of our proposed protocol is larger than that of LEACH and DEEC protocols, which means more rounds and longer network lifetime. The first sensor node dies after approximately 60 rounds, we expect that this sensor node is a normal node; because the probability of a normal node to die is greater than that of the advanced node. Furthermore, we expect that during the last rounds only the advanced nodes will be alive.

Figure (4) shows that the messages delivered by our proposed protocol are more than that of LEACH and DEEC protocols, the comparison between protocols is made with $\alpha = 0.5$. So, our proposed protocol achieves a better throughput. This is a result of the improvement in network lifetime.

We next investigate the performance of our EDAC protocol for two-level heterogeneous networks but for different percentages of advanced nodes (i.e., varying m). With parameters are the same as used in Table.(2), number of alive nodes and sent packets are depicted in Figs.(5) and (6). One notes that as the number of advanced nodes increases, so does the network's lifetime as evident in Fig.(6).

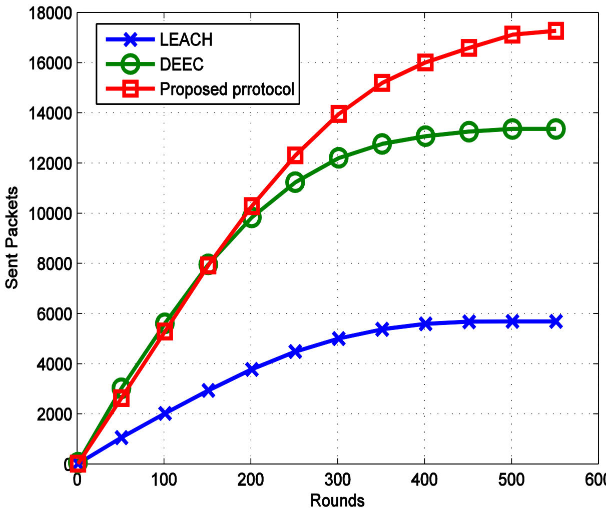


Fig. 4. Number of sent packets for LEACH, DEEC and the proposed protocol

As a result, the accumulative number of sent packets increases as m is increased. This agrees with what we expect from an energy efficient clustering method.

We next investigate the performance of our proposed protocol versus that of the LEACH and DEEC protocols in terms of the network lifetime as we vary the range of initial energies. More precisely, sensors are equipped with random initial energy levels that fall in the range $[E_0,\ E_0(1+a)]$. Thus, we have in effect a multilevel network. Results in Fig.(7) show the first round when 10% of the nodes die (i.e., consumed their energy) as we vary a. It is noted that both the DEEC and our proposed EDAC protocol have a comparable

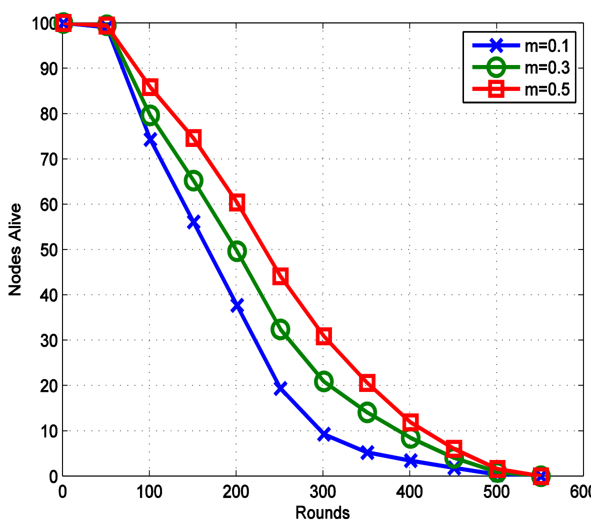


Fig. 5. Number of alive nodes using the EDAC algorithm for various m values

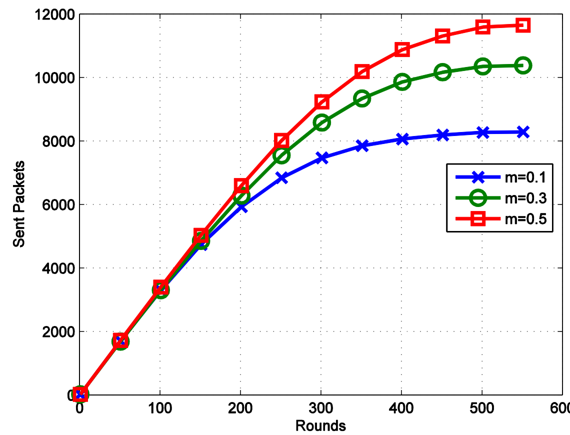


Fig. 6. Number of sent packets using the EDAC algorithm for various m values

performance and they both outperform the LEACH algorithm. On the other hand, Fig. (8) depicts the first round when all sensors have died out. Results show that our proposed method outperforms both the LEACH and DEEC algorithms. In particular, the proposed EDAC extend the network's lifetime by almost 11% and 37% with comparison to the DEEC and LEACH protocols, respectively.

VI. Conclusion

In this paper, we propose a probabilistic energy aware clustering algorithm which we call the EDAC. In addition to a sensor's residual energy, the sensor incorporates the residual energy levels of sensors within the cluster radius in addition to the distances between sensors. Using this metric allows selection of cluster head so that it can be close to sensors with varying levels of residual energies (e.g., close to sensors with low residual energy levels). The weighted metric is used in constructing a cluster head selection probability for each sensor. Simulation results, indicate that the proposed algorithm is applicable for both two-level and multi-level networks. In addition, the EDAC performance in terms of network lifetime and sent packets has been shown to outperform both those of the LEACH and DEEC protocols under different setups.

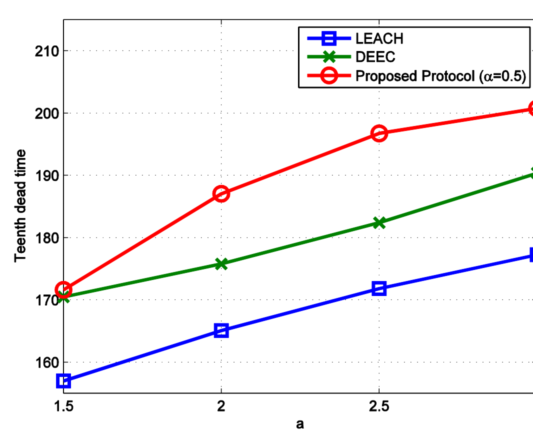


Fig. 7. Round which the time when the first 10% of the nodes die for different energy ranges

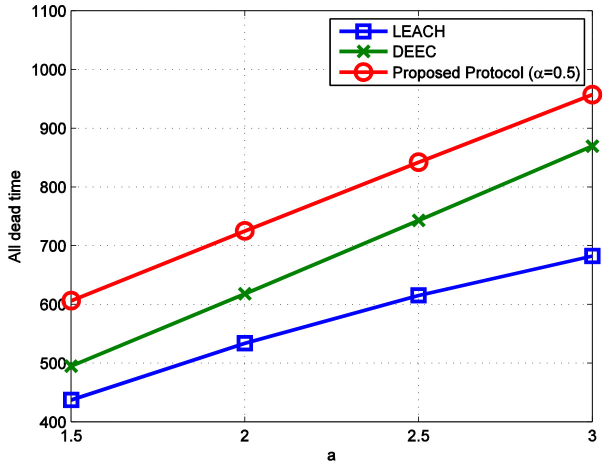


Fig. 8. Round all nodes die for different energy ranges

REFERENCES

[1] Chong, Chee-Yee and Kumar, Srikanta P, Sensor Networks: Evolution, Opportunities, and Challenges, Proceedings of the IEEE,VOL. 91, NO. 8, (2003), 1247-1256.

[2] M. P. urii; Z. Tafa ; G. Dimi; V. Milutinovi, A Survey of Military Applications of Wireless Sensor Networks, Embedded Computing (MECO), 2012 Mediterranean Conference on,(2011), 196-199.

[3] L. M. Oliveira and J. J. Rodrigues, Wireless Sensor Networks: A Survey on Environmental Monitoring, Journal of communications,VOL. 6, NO. 2,(2003), 143-151.

[4] Garc´ıa-Hern´andez, Carlos F and Ibarguengoytia-Gonzalez, Pablo H and Garc´ıa-Hern´andez, Joaqu´ın and P´erez-D´ıaz, Jes´us A, Wireless Sensor Networks and Applications: A Survey, IJCSNS International Journal of Computer Science and Network Security,VOL. 7, NO. 3,(2007), 264– 273.

[5] Arampatzis, Th and Lygeros, John and Manesis, Stamatis, A Survey of Applications of Wireless Sensors And Wireless Sensor Networks, Intellegent Control, 2005. Proceedings of the 2005 IEEE International Sympsium on, Mediterrean Conference on Control and Automation,(2005), 719-724.

[6] Liu, Xuxun, A Survey on Clustering Routing Protocols inWireless Sensor Networks, Sensors,VOL. 12, NO. 8,(2012), 11113-11153.

[7] Heinzelman, Wendi Rabiner and Chandrakasan, Anantha and Balakrishnan, Hari, Energy-Efficient Communication Protocol for Wireless Microsensor Networks, System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on, IEEE,VOL. 12, NO. 8,(2000), 10–pp

[8] Younis, Ossama and Fahmy, Sonia, HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks, Mobile Computing, IEEE Transactions on,VOL. 3, NO. 4,(2004), 366–379

[9] Qing, Li and Zhu, Qingxin and Wang, Mingwen, Design of a Distributed Energy-Efficient Clustering Algorithm for Heterogeneous Wireless Sensor Networks, Computer communications,VOL. 29, NO. 12,(2006), 2230– 2237

[10] Smaragdakis, Georgios and Matta, Ibrahim and Bestavros, Azer and others, SEP: A Stable Election Protocol for Clustered Heterogeneous Wireless Sensor Networks, Second international workshop on sensor and actor network protocols and applications (SANPA 2004),VOL. 3, NO. 8,(2004), 11113-11153.

[11] Bandyopadhyay, Seema and Coyle, Edward J, An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks, INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies,VOL. 3, NO. 8,(2003), 1713–1723.

[12] Abo-Zahhad, Mohammed and Ahmed, Sabah M and Sabor, Nabil and Sasaki, Shigenobu, A New Energy-Efficient Adaptive Clustering Protocol Based on Genetic Algorithm for Improving the Lifetime and the Stable Period of Wireless Sensor Networks, International Journal of Energy Information and Communications,VOL. 5, NO. 3,(2014).

[13] Heinzelman, Wendi B and Chandrakasan, Anantha P and Balakrishnan Hari, An Application-Specific Protocol Architecture for Wireless Microsensor Networks, Wireless Communications, IEEE Transactions on,VOL. 1, NO. 4,(2002), 660–670.

Artificial Neural Networks and Support Vector Machine for Voice Disorders Identification

Nawel SOUISSI, Adnane CHERIF
University of Tunis El-Manar, Faculty of sciences of Tunis
Innov'COM Laboratory, 2092, Tunis
Tunis, Tunisia

***Abstract*—The diagnosis of voice diseases through the invasive medical techniques is an efficient way but it is often uncomfortable for patients, therefore, the automatic speech recognition methods have attracted more and more interest recent years and have known a real success in the identification of voice impairments. In this context, this paper proposes a reliable algorithm for voice disorders identification based on two classification algorithms; the Artificial Neural Networks (ANN) and the Support Vector Machine (SVM). The feature extraction task is performed by the Mel Frequency Cepstral Coefficients (MFCC) and their first and second derivatives. In addition, the Linear Discriminant Analysis (LDA) is proposed as feature selection procedure in order to enhance the discriminative ability of the algorithm and minimize its complexity. The proposed voice disorders identification system is evaluated based on a widespread performance measures such as the accuracy, sensitivity, specificity, precision and Area Under Curve (AUC).**

Keywords—Automatic Speech Recognition (ASR); Pathological voices; Artificial Neural Networks (ANN); Support Vector Machine (SVM); Linear Discriminant Analysis (LDA); Mel Frequency Cepstral Coefficients (MFCC)

I. Introduction

When the mechanism of voice production is affected, the voice becomes pathological and sometimes intelligible which causes many problems and difficulties to integrate the social environment and to have an easy exchange between members of the same community. Therefore, the diagnosis of voice impairments is imperative to avoid so many issues. Voice disorders can be classified into three main categories: organic, functional or combination of both [1]. This study is designed for organic voice disorders. Indeed, a voice disorder is organic if it is caused by structural (anatomic) or physiologic disease, either a disease of the larynx itself or by remote systemic or neurologic diseases that alter larungeal structure or function [2]. In this research, we have worked on both structural and neurogenic disorders. Four types of pathologies are examined: Chronical laryngitis, Cyst, Reinke edema and Spasmodic dysphonia since they are widespread diseases and their medical analysis is a bit tricky to date. Among many techniques to identify voice diseases, the automatic acoustic analysis has proven its efficiency last years and has attracted more and more success. The advantage of acoustic analysis is its nonintrusive nature and its potential for providing quantitative data with reasonable expenditure of analysis time [3]. Therefore, several techniques and methods have been introduced and many studies have been conducted in the literature. Some of these researches indicate that voice disorders identification can be done by the exploitation of Mel Frequency Cepstral Coefficients (MFCC) with the harmonics-to-noise ratio, normalized noise energy and glottal-to-noise excitation ratio, Gaussian mixture model was used as classifier [4]. Also, Daubechies' discrete wavelet transform, linear prediction coefficient, and least-square Support Vector Machine (LS-SVM) were investigated in [5]. In addition, a voice recognition algorithm was proposed in [6] based on the MFCC coefficients, their first and second derivatives, performance of F-ratio and Fisher's discriminant ratio as feature reduction methods and Gaussian Mixture Model (GMM) as classifier; the main idea, here, consists in demonstrating that the detection of voice impairments can be performed using both mel cepstral vectors and their first derivative, ignoring the second derivative. In this paper, we will prove that the contribution of the first and second derivatives of the MFCC features mainly depends on the classifier. Indeed, the Artificial Neural Networks (ANN) and the Support Vector Machine (SVM) as classifiers are investigated in this work and a comparative study between their respective performances is conducted. In addition, three combinations of the MFCC features, their first and second derivatives are proposed for the feature extraction task. In order to select the most relevant parameters from the resulting feature vector, the Linear Discriminant Analysis (LDA) is suggested as feature selection procedure. Furthermore, the system performance is assessed in terms of the accuracy, sensitivity, specificity, precision and Area Under Curve (AUC). In the next section, the methodology and database used in this work are described as well as the performance measures. Then, section 3 presents the experimental results and section 4 discusses these obtained results. Finally, we conclude this paper with section 5.

II. Materials and Methods

A. Database

In this research, we have selected the voice samples from the 'Saarbrucken Voice Database' (SVD) [7], [8] which is a German disorders voice database collected in collaboration with the Department of Phonetics and ENT at the Caritas clinic St. Theresia in Saarbrucken and the Institute of Phonetics of the University of the Saarland. It contains 2225 voice samples with a sampling rate of 50 kHz and with a 16 bit amplitude resolution. Subjects have sustained the vowels [i], [a] and [u] for 1s long. In this study, the continuous vowel [a] phonation produced by 50 normal people and 70 patients were examined. Four types of pathologies are investigated: Chronical laryngitis

(24), Cyst (6), Reinke's edema (19) and Spasmodic dysphonia (21).

B. The Proposed Algorithm

In this paper, the extraction of the acoustical features from the speech signal is performed by the MFCC parameterization method. In addition, the first and second derivatives which provide information about the dynamics of the time-variation in MFCC original features were investigated to verify their contribution to the proposed algorithm. In order to optimize the voice disorders detection, a projection based Linear Discriminant Analysis (LDA) as feature selection method is suggested and a comparative study is elaborated between optimized and non-optimized features for every tested combination. As regards the classification task, the Artificial Neural Networks (ANN) are used as unconventional approach in addition to the Support Vector Machine as a new method successfully exploited in recent years, Fig. 1.

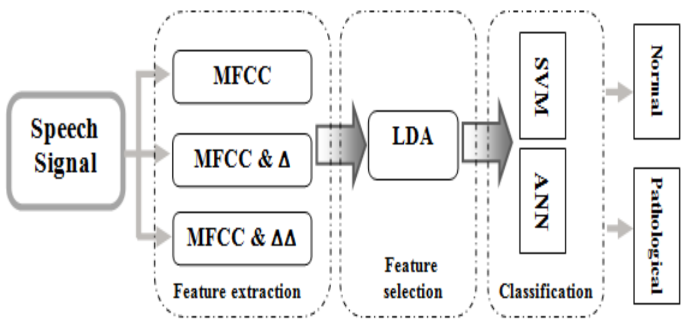


Fig. 1. Block diagram of the proposed system

C. Feature Extraction Method

Feature extraction is obviously the most crucial task in speech recognition process. In this research, the Mel Frequency Cepstral Coefficients (MFCC) procedure is chosen as a robust technique commonly used and has proven its efficiency in speech recognition.

The Mel Frequency Cepstral Coefficients (MFCC) is a nonparametric frequency domain approach which is based on human auditory perception system.

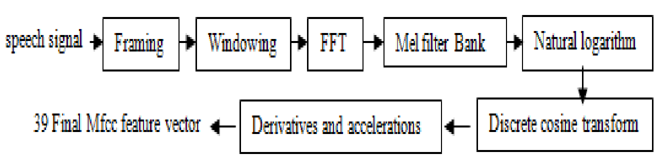


Fig. 2. Block diagram of the MFCC procedure

As presented in Fig. 2, the procedure of the MFCC features extraction starts by the decomposition of the speech signal into small frames since it is slowly time varying and can be treated as a stationary random process when considered under a short time frame [9], Then, windowed (a 30 ms. Hamming window was used) with no preemphasis. The frames were extracted with a 50% frame shift. The spectral coefficients of the speech frames are estimated using the nonparametric fast Fourier transform (FFT)-based approach. On the other hand, the human auditory system perceives sound in a nonlinear frequency binning. Therefore, Mel filtering process has to be performed. Thus, the obtained speech signal spectrum is filtered by a group of triangle bandpass filters that simulate the characteristics of human's ear [9], [10]. The following equation is used to compute the Mel frequency f_{Mel} for a given linear frequency f_{Hz} in Hz.

$$f_{Mel} = 2595 * \log(1 + f_{Hz} / 700) \qquad (1)$$

The nonlinear characteristic of human auditory system in frequency is approximated by the Mel filtering procedure. At this stage, a natural logarithm is applied on each output spectrum from Mel bank. Finally, The Discrete Cosine Transform (DCT) is performed to convert the log Mel spectrum into time domain; thus the Mel Frequency Cepstrum Coefficients (MFCC) are obtained. Besides, there are several ways to approximate the first derivative of a cepstral coefficient. In this research, we use the following formula [11]:

$$\frac{dx(t)}{dt} = \Delta x(t) \approx x\sum_{m=-M}^{M} m(t+m) \qquad (2)$$

Where x(t) is the cepstral coefficient, t is the frame number and 2M + 1 is the number of frames considered in the evaluation. The same formula can be applied to the first derivative to produce the acceleration.

For each time frame, the MFCC feature vector is composing of N original cepstral features, N delta cepstral coefficients and N delta- delta coefficients. Where N is the number of MFCC features chosen for a simulation. In this work, several experiments were conducted using 13 original MFCC features, their derivatives and accelerations in order to perform a comparative study between the different proposed combinations.

D. Feature Selection Procedure

In this research, Linear Discriminant Analysis (LDA) is suggested as a feature selection procedure which is a supervised subspace learning method based on Fisher Criterion [12]. Indeed, it aims to estimate the parameters of a projection matrix in order to map features from an *h*-dimensional space to a k-dimensional space (k<h) in which the between class scatter is maximized while the within-class scatter is minimized. The within-class scatter calculates the average variance of the data within each class, while the between-class scatter represents the average distance between the means of the data in each class and the global mean [13]. Linear Discriminant Analysis is investigated in this research in order to optimize the proposed identification algorithm since it is able to select the most relevant parameters from a feature vector in order to minimise the complexity of the system while improving recognition rates.

E. Classification Algorithms

Two classification algorithms are proposed in this work and a comparative study is established between their performance rates in order to conclude the most effective classifier for the identification of voice disorders.

1) Support Vector Machine:

Support Vector Machines are a class of learning techniques introduced by Vladimir Vapnik in the early 90s [14], [15]. The binary classification is where the training data comes only from two different classes (+1 or -1). The idea of SVM is to find a hyperplane that best separates the two classes with maximum margin. If the data is linearly separable, it is called « Hard-margin SVM ». If the data is non-linearly separable, it is called "Soft-margin SVM". In this case, the data are mapped into a higher-dimensional space where the function becomes linear. This transformation space is often performed using a "Kernel Mapping function" and the new space is called "Features space". The most widely used SVM kernel functions are linear kernel, polynomial kernel and Radial Basis function (RBF) as Gaussian kernel.

The training phase of the SVM classifier involves searching the hyperplane that maximizes the margin. Such hyperplane is called « hyperplane optimal separation ».

In this research, the proposed algorithm was trained with the « Radial Basis Function » (RBF) as a Gaussian SVM kernel and LIBSVM which is a SVM library [16].

2) Artificial Neural Networks:

Artificial Neural Networks are absolutely one of the most effective approaches for speech recognition thanks to their numerous architectures and learning algorithms, In this paper, the architecture of the proposed neural networks is composed of three layers, an input layer for the transmission of the input features without distortion, a hidden layer containing 250 neurons (sigmoid is applied as activation function) and an output layer containing a linear function neuron. Each layer is completely connected to the next one. The proposed neural network learning is performed based on the principles of the Bayesian regularization algorithms. Indeed, the network weight values are adjusted successively at every step of learning in order to achieve an output as close as possible to the considered data [17].

Concerning the Bayesian approach, it is based on the exploitation of a random distribution of the network weight probabilities. The neural network learning consists in determining the distribution knowing the training data. Indeed, after the examination of the training data, the initial probability attributed to weights, before performing the learning, is transformed into a final distribution through the application of the Bayes theorem [17].

F. Evaluation Process

In order to judge the effectiveness and the robustness of the proposed algorithm, it has to be assessed according to different performance measures. In this research, five performance measures were used: accuracy, sensitivity, specificity, precision and the Area Under Curve (AUC) from the Receiver Operating Characteristic Curve (ROC). Indeed, sensitivity measures the ability of the algorithm to recognise pathological samples. It opposes specificity which evaluates the ability of the algorithm to identify normal samples. Precision represents the proportion of well-classified pathological samples from the pathological class. Furthermore, Accuracy measures the algorithm correct classification rate and the AUC which is an important statistical property for evaluating the discriminability between the two classes of normal and pathological samples. Therefore, the AUC provides another way to measure the accuracy of the proposed system. These measures are based on the following notions:

TP : True Positive : identified as pathological when pathological samples are actually present

TN : True Negative : identified as normal when normal samples are actually present

FP : False Positive : identified as pathological when normal samples are actually present

FN : False Negative : identified as normal when pathological samples are actually present

These measures can be calculated as follows:

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}}$$

$$\boldsymbol{Sensitivity} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}}$$

$$\boldsymbol{Specificity} = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}}$$

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}}$$

$$\boldsymbol{Area\ Under\ Curve\ (AUC)} = \frac{1}{2}\times\left(\frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}}+\frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}}\right)$$

III. Experimental Results

In this research, the dataset was divided into two parts: 70% of the data were used for training and 30% for validation. All simulations were conducted in MATLAB 2013a with Intel Core-i7, 2.20 GHz CPU and 4 GB RAM.

A. Evaluation Based on the SVM Performance

In this part of the article, we present the SVM performance rates for different combinations of the MFCC coefficients before and after applying the LDA feature selection procedure. Table 1 shows the SVM performance in terms of accuracy (Acc %), sensitivity (Sens %), specificity (Spec %), precision (Prec %) and AUC (%) for the different MFCC feature vectors.

The experimental results show that there is a slight increase, in the SVM performance rates between the MFCC and MFCC_Delta1 combinations, of 0.04% in the accuracy rate, 0.03% in the AUC rate, 0.04% in the sensitivity rate, 0.05% in the specificity rate and 0.07% in the precision rate. Whereas, the system performances are exactly equal for the combinations of MFCC_Delta1 and MFCC_Deltas1&2 with an accuracy rate of 80.4%, sensitivity of 87.83%, specificity of 73.58%, AUC of 80.7% and precision of 72.29%. Therefore, we can note that the first and the second derivatives don’t provide a significant improvement in the system performances when the SVM is used as classifier which demonstrates that the

SVM algorithm is not sensible to the information provided by these features about the dynamics of the time-variation in the MFCC original vector. Besides, after applying the LDA procedure, the SVM performance rates are certainly less close but not enough distant to change the whole analysis about the contribution of the first and the second derivatives in the proposed algorithm when the SVM is applied as classifier.

TABLE I. THE SVM PERFORMANCE BASED ON THE MFCC COMBINATIONS BEFORE APPLYING THE LDA PROCEDURE (TABLE 1-1) AND AFTER INCLUDING LDA PROCEDURE (TABLE 1-2)

	Acc%	Sens%	Spec%	AUC%	Prec%
Table 1-1					
MFCC	80.36	87.81	73.53	80.67	72.22
MFCC+Delta 1	80.4	87.83	73.58	80.7	72.29
MFCC+Delta 1&2	80.4	87.83	73.58	80.7	72.29
Table 1-2					
MFCC+LDA	86.28	98.23	76.83	87.53	74.1
MFCC+Delta1+LDA	86.07	97.96	76.65	87.31	73.91
MFCC+Deltas1&2+LDA	**86.44**	**98.24**	**77.04**	**87.64**	**74.42**

In the literature, previous results found by Godino-Llorente et al. [6] demonstrate that the detection of voice impairments can be performed using both mel cepstral vectors and their first derivative, ignoring the second derivative when the Gaussian Mixture Models are applied as classifier. However, our findings prove that even the first derivative can be ignored in the detection of voice impairment and only the original Mel Frequency Cepstral Coefficients are significant with the SVM classifier.

On the other hand, the LDA feature selection method was applied considering the different MFCC feature vectors. The experimental results show a significant improvement in the system performance. Thus, Fig. 3 exposes an optimization of 5.92% for the MFCC features which leads to an accuracy rate of 86.28%.

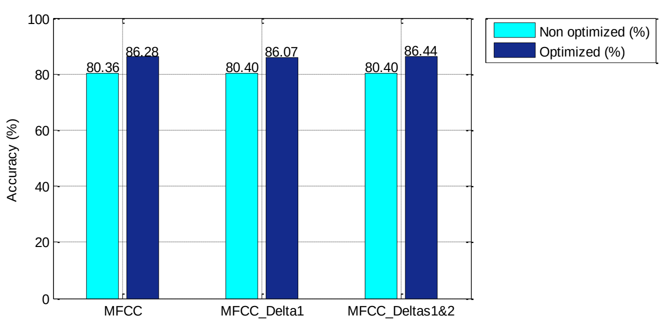


Fig. 3. Comparison between the SVM accuracy rates of the optimized and non-optimized MFCC features

Similarly, the optimized MFCC_Delta1 and MFCC_Delta1&2 combinations provide the accuracy rates of 86.07% and 86.44% representing an increase of 5.67% and 6.04%, respectively.

The AUC rates for the different MFCC combinations are presented in Fig. 4. It is observed that the improvement is important between optimized and non-optimized features such as the increase of 6.86% for the MFCC combination and 6.61% for the MFCC_Delta1 and 6.94% between the optimized and non-optimized MFCC_Delta1&2 features.

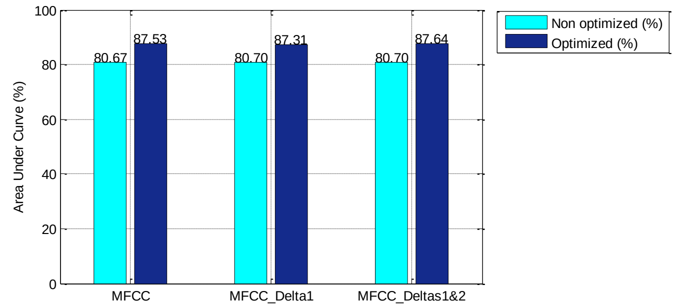


Fig. 4. Comparison between the SVM AUC rates of the optimized and non-optimized MFCC features

Hence, the LDA procedure can be considered efficient in the selection of the most relevant parameters in order to obtain the optimized feature vector able to achieve best performance rates. Thus, the best performances were achieved by the optimized MFCC_Delta1&2 with a slight increase comparing to the other optimized features as mentioned in Table 1.

B. Evaluation Based on the ANN Performance

Table 2 shows the ANN performance rates for different combinations of the MFCC coefficients before and after applying the LDA feature selection procedure. The system performances are presented in terms of accuracy (Acc%), sensitivity (Sens%), specificity (Spec%), precision (Prec%) and AUC(%) for the different MFCC feature lengths.

TABLE II. THE ANN PERFORMANCE BASED ON THE MFCC COMBINATIONS BEFORE APPLYING THE LDA PROCEDURE (TABLE 2-1) AND AFTER INCLUDING LDA PROCEDURE (TABLE 2-2)

	Acc%	Sens%	Spec%	AUC%	Prec%
Table 2-1					
MFCC	75.13	78.7	71.33	75.02	72.89
MFCC+Delta 1	81.19	89.85	73.62	81.74	71.67
MFCC+Delta 1&2	85.2	91.07	79.34	85.21	79.33
Table 2-2					
MFCC+LDA	80.25	84.16	72.48	81.87	72.97
MFCC+Delta1+LDA	84.06	98.78	75.39	85.59	72.07
MFCC+Deltas1&2+LDA	**87.82**	**99.12**	**80.31**	**87.96**	**81.42**

It is obvious that the ANN performance is increasingly better after integrating the first and second derivatives of the MFCC features. In fact, the accuracy and AUC rates are about 75.13% and 75.02%, respectively, for the combination of the original MFCC features whereas these rates are about 81.19% and 81.74%, respectively, when the first MFCC derivatives are associated with the original ones. This improvement is enhanced for the combination of the MFCC features with their first and second derivatives since a significant increase in the system performance measurements is observed. Indeed, this combination offers an accuracy of 85.2% and AUC of 85.21%. Therefore, the first and second derivatives of the MFCC coefficients can be considered significant when the ANN is applied as classifier since they offer a great improvement in the system performance compared to the results of the original MFCC features. In fact, this variation between the different combinations is observed before and after applying the LDA transformation.

As regards the LDA method, it was applied to the different MFCC combinations in order to select the most significant parameters from the feature extraction task to be the input vector of the ANN architecture. This strategy leads to an optimization in the system performance. Indeed, the experimental results show an improvement in the ANN performance measurements for all the optimized MFCC feature combinations. Fig. 5 compares the ANN accuracy rates of the optimized and non-optimized MFCC vectors.

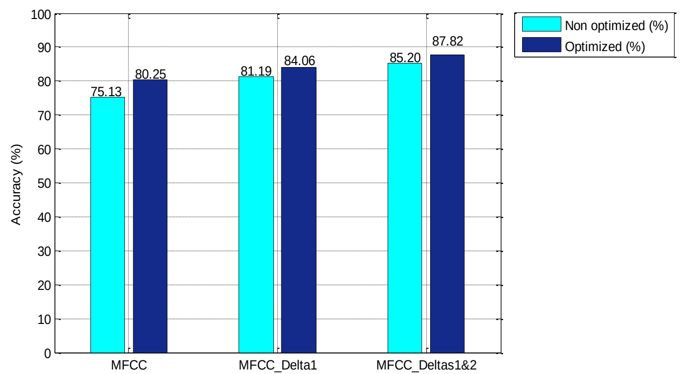


Fig. 5. Comparison between the ANN accuracy rates of the optimized and non-optimized MFCC features

The experimental results exposed in Fig. 5 show an optimization of 5.12% in the accuracy rate of the non-optimized MFCC features, while the improvement is about 2.87% for the combination of the MFCC features and their first derivatives. Also the optimization procedure provides a 2.62% increase in the accuracy rate of the MFCC features associated with their first and second derivatives. In fact, the improvement was observed for all performance measures namely the AUC rates which were improved to reach 81.87% for the MFCC combination with an optimization of 6.85% while 3.85% and 2.75% were the improvement rates for the combination of MFCC_Delta1 and MFCC_Delta1&2, respectively, Fig. 6.

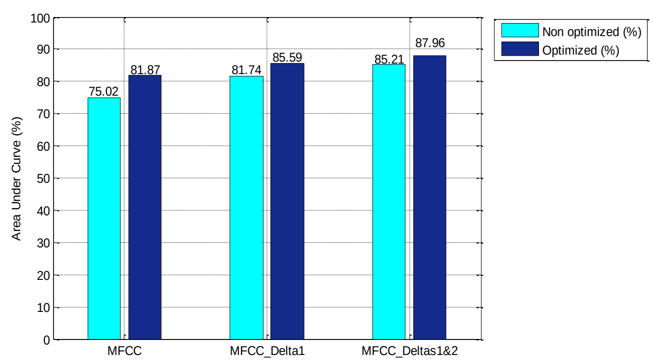


Fig. 6. Comparison between the ANN AUC rates of the optimized and non-optimized MFCC features

Finally, the optimized MFCC_Delta1&2 combination reached the best ANN performance rates with an accuracy rate of 87.82%, sensitivity of 99.12%, specificity of 80.31%, AUC of 87.96% and a precision of 81.42% as mentioned in Table 2.

IV. Discussion

In this paper, the ANN is proposed as unconventional approach in addition to the SVM as a new method successfully exploited in speech recognition. The main motivation for conducting this research was to investigate the efficiency of each of those classifiers in the identification of voice disorders. In addition, it was interesting to scrutinize the contribution of the first and second derivatives of the MFCC features for every classifier. The experimental results demonstrate that the effect of these derivative features depends on the classifier. Indeed, when the SVM is used as classifier, the first and second derivatives do not provide any improvement to the system performance comparing to the original MFCC features. However, when the ANN is used as classifier, these derivative features can be considered important since they contribute in the improvement of the system performance. In this case, there is an average improvement about 4% between the combination of the MFCC, MFCC_Delta1 and the MFCC_Delta1&2.

Besides, the LDA procedure is used to select the most relevant parameters from a resulting feature vector in order to reduce the system dimensionality without affecting its performance. Indeed, our findings show that the LDA method minimizes the system complexity while improving the performance rates for every feature combination; therefore it can be considered as an optimization procedure.

Table 3 compares the proposed algorithms with previous significant works. It is observed that the proposed algorithm appears competitive for the detection of voice disorders from the Saarbrucken Voice Database (SVD).

TABLE III. Comparative Table Between Proposed Algorithm and Previous Works

Authors	Database	Measurements	Classifier	Best Acc	Ref.
D.Martinez et al	SVD	MFCC, harmonics to noise ratio, normalized noise energy and glottal to noise excitation ratio	GMM	87.9%	[4]
Ahmed al Nasheri et al	SVD	autocorrelation of the filter band's output	GMM	72%	[18]
			SVM	71.12%	
El Emary et al	SVD	MFCC, Jitter and Shimmer	GMM	82.37%	[19]
Current paper	SVD	MFCC, First and second derivatives, LDA	SVM	86.44%	
			ANN	87.82%	

Finally, with an accuracy rate of 86.44%, sensitivity of 98.24%, specificity of 77.04%, AUC of 87.64% and precision of 74.42%, the SVM classifier can be judged efficient for voice disorders identification. Also, the ANN classifier offers an accuracy rate of 87.82%, sensitivity of 99.12%, specificity of

80.31%, AUC of 87.96% and precision of 81.42% which are slightly better than those of the SVM classifier which leads to conclude that the ANN classifier is likewise effective for voice impairment identification. With these performance rates, the proposed algorithm can be considered reliable for the identification of pathological voices from normal ones.

V. Conclusion

This paper proposes an optimized voice disorders identification algorithm based on short-term cepstral parameters and the Linear Discriminant Analysis as feature selection method. As regards the classification task, it is performed by the Artificial Neural Networks and the Support Vector Machine. The three combinations of MFCC, MFCC_Delta1 and MFCC_Delta1&2 are examined in order to conclude the role of the derivative features. Indeed, experimental results demonstrate that the contribution of the first and second derivative of the MFCC features varies according to the classifier. In addition, the LDA transformation can be considered as optimization procedure since it improves the system performance while reducing its dimensionality. The accuracy rates of 86.44% and 87.82% were obtained by the SVM and the ANN, respectively. Therefore, we can conclude that ANN and SVM are efficient for voice disorders identification with a slight advantage to the ANN. Many future improvements can be proposed such as including other feature extraction methods in a hybrid schema in order to improve performance rates. For instance, we can suggest the Discrete Wavelet Transform to be integrated with the proposed MFCC features. In addition, the real time implementation of the proposed algorithm may be envisaged.

References

[1] A. Akbari and M. K. Arjmandi, “An efficient voice pathology classification scheme based on applying multi-layer linear discriminant analysis to wavelet packet-based features,” Biomedical Signal Processing and Control, vol. 10, pp. 209-223, 2014.

[2] A. E. Aronson and D. M. Bless, Clinical voice disorders, Fourth ed., New York: Thieme, 2009.

[3] Lions Voice Clinic, University of Minnesota, Department of Otolaryngology, P. O. Box 487, 420 Delaware St., SE, Minneapolis, MN55455,USA.

[4] D. Martinez, E. Lleida, A. Ortega, A. Miguel and J. Villalba, “Voice Pathology Detection on the Saarbruecken Voice Database with Calibration and Fusion of Scores Using MultiFocal Toolkit,” Advances in Speech and Language Technologies for Iberian Languages, vol. 328, pp. 99-109, 2012.

[5] E. F. Fonseca, R. C. Guido, P. R. Scalassara, C. D. Macciel and J. C. Pereria, “Wavelet time frequency analysis and least square support vector machine for the identification of voice disorders,” Comp. Bio. Med., vol. 37, pp. 571-578, 2007.

[6] J. I. Godino-Llorente, P. Gomez-Vilda and M. Blanco-Velasco, “Dimensionality reduction of a pathological voice quality assessment system based on Gaussian mixture models and short-term cepstral parameters,” IEEE Trans. Biomed. Eng., vol. 53, pp. 1943- 1953, 2006.

[7] W. J. Barry and M. Putzer, Saarbrucken Voice Database, Institute of Phonetics, Univ. of Saarland.

[8] M. Putzer and J. Koreman, “A German database of patterns of pathological vocal fold vibration,” Phonus 3, Institute of Phonetics, University of the Saarland, pp. 143-153, 1997.

[9] X. Xiong, “Robust speech features and acoustic models for speech recognition,” PhD Dissertation, School of computer engineering, Nanyang Technological University, 2009.

[10] V. Tiwari, “MFCC and its applications in speaker recognition,” International Journal on Emerging Technologies, vol. 1, pp. 19-22, 2010.

[11] J. W. Picone, “Signal modeling techniques in speech recognition,” in Proc. of the IEEE, vol. 81, pp. 1215-1247, 1993.

[12] G. Quanquan, L. Zhenhui and H. Jiawei, Linear Discriminant Dimensionality Reduction, ser. Lecture Notes in Computer Science, Machine Learning and Knowledge Discovery. Germany: Springer, 2011, pp. 549-564.

[13] V. S. Tomar, “Discriminant feature space transformations for automatic speech recognition,” Department of Electrical and Computer Engineering, McGill University, Montreal, 2012.

[14] I. Guyon, B. Boser and V. Vapnik, “Automatic capacity tuning of very large VC-dimension classifiers,” Advances in Neural Information Processing Systems, pp. 147-155, 1993.

[15] B. E. Boser, I. M. Guyon and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” in Proc. WCLT'92, New York, 1992.

[16] C. C. Chang and C. J. Lin, “LIBSVM: a library for support vector machines,” ACM Trans. Intell. Syst. Technol., vol. 27, pp. 1-27, 2011.

[17] R.M. Neal, Bayesian learning for neural networks, New York : Spring Verlag, 1996.

[18] A. Al-nasheri, A Zulfiqar, M Ghulam and A Mansour, “Voice pathology detection using auto-correlation of different filters bank,” in Proc. AICCSA'14, Doha, Qatar, 2014.

[19] I. M. M. El Emary, M. Fezari and F. Amara, “Towards developing a voice pathologies detection system,” Journal of Communications Technology and Electronics, vol. 59, pp. 1280-1288, 2014.

Evaluating Damage Potential in Security Risk Scoring Models

Eli Weintraub
Department of Industrial Engineering and Management
Afeka Tel Aviv Academic College of Engineering
Tel Aviv, Israel

***Abstract*—A Continuous Monitoring System (CMS) model is presented, having new improved capabilities. The system is based on the actual real-time configuration of the system. Existing risk scoring models assume damage potential is estimated by systems' owner, thus rejecting the information relying in the technological configuration. The assumption underlying this research is based on users' ability to estimate business impacts relating to systems' external interfaces which they use regularly in their business activities, but are unable to assess business impacts relating to internal technological components. According to the proposed model systems' damage potential is calculated using technical information on systems' components using a directed graph. The graph is incorporated into the Common Vulnerability Scoring Systems' (CVSS) algorithm to produce risk scoring measures. Framework presentation includes system design, damage potential scoring algorithm design and an illustration of scoring computations.**

Keywords—CVSS; security; risk management; configuration; Continuous Monitoring; vulnerability; damage potential; risk scoring

I. Introduction

Cyber-attackers cause damage to organizations and personal computers by stealing their business or private data and by making changes in their software and hardware [1]. The damages are usually categorized by security experts to three kinds: loss of confidentiality, integrity or availability. Vulnerabilities are software weaknesses or exposures. An attack is performed by exploiting software vulnerabilities in the target system. Attackers make use of vulnerabilities stemming from bugs that are potential causes to security failures. Exploits are planned to attack certain components having specific vulnerabilities. Users' computers might be damaged by exploited vulnerabilities. Defending computers depends on the amount of knowledge an organization has of it's computing systems' vulnerabilities. This work focuses on gaining accurate knowledge of computers' configuration, thus enabling improved organizational risk mitigation activities, to defend computers from threats caused by attackers. Accurate knowledge of computers' risks assists security managers to adopt security measures effectively. Reference [2] states that Stuxnet worm included a process of checking hardware models and configuration details before launching an attack. Both, attackers and security managers are interested in gaining accurate and detailed information of the system. Risk managers make decisions on activities actions they have to perform in order to limit their exposure to risks according to the amount of potential damage and vulnerability characteristics [3].

Risk has many definitions in research publications. In this research we use the definition of [4]: "An event where the outcome is uncertain". According to this definition, this work is aimed at lessening risk uncertainty. The proposed model focuses on an improved collateral damage potential evaluation process which is based on the real-time information on systems' configuration components, and on system interfaces with users.

Several software products are used to defend computers from cyber attackers. Antivirus software, antispyware and firewalls are examples to some of these tools based on periodic assessment of the target computer by comparing computers' software to the known published vulnerabilities. Those tools are effective only against known threats and not against new unpublished threats. CMS monitor computer systems in a near real time process aimed at detecting vulnerabilities and notifying security managers. Contemporary systems use vulnerabilities databases which are continually updated as new vulnerabilities are detected and a scoring algorithm which predicts potential business damages. This work focuses on the impacts of the components incorporating the configuration about potential damages. The CMS evaluates damage potential relating to the actual configuration. Each time changes are performed to components damage potential is evaluated and updated. CMS's are useful tools for limiting the time-frames organizations are exposed to risks.

Computers are at risk to known threats until the time a patch is prepared for defending the vulnerable software, an activity that may last weeks or months. Even after a patch is prepared by the software vendor a computer might still be at risk until the moment the new patch is loaded to the system. Loading patches to computer systems is usually performed as a periodical process, not continuously to avoid too many interrupts required for uploading the patch on organizations' computers. Other software tools are based on heuristic algorithms which are planned to detect irregular suspicious activities of the software running on the computers. In today's environment of zero-day exploits, conventional systems updating for security mitigation activities has become a cumbersome process. There is an urgent need for a solution that can rapidly evaluate system vulnerabilities' potential damages for immediate risk mitigation [5].

Security Continuous Monitoring (SCM) tools use techniques for monitoring, detecting and notifying of security threats in real time. After identifying these risks, the tools evaluate the potential impacts on the organization. Reference [6] states that SCM systems which are running on computers continuously try to detect systems' vulnerabilities, are aimed at closing the gap between the zero-day of identifying the vulnerability, until the moment the computer is loaded by a patch. The time gap may be considerably long.

This paper describes the mechanisms of a new SCM framework of a system that will produce better risks scoring than current known systems. The framework bases processes on two grounds: 1) knowledge concerning real computers' configuration of the target system, and 2) a prediction algorithm which runs continuously and computes damage potential estimates for use of risk scoring models.

The rest of the paper is organized as follows: In section 2 a description of current known existing solutions. In section 3 a presentation of the proposed framework including systems architecture. In section 4 a description of the risk scoring algorithm which computes risk scores. In section 5 presentation of the results. In section 6 conclusions and future research directions.

II. Existing Solutions

SCM systems are using external vulnerabilities databases for evaluation of the target computers' risk. There are several owners of vulnerability databases [5] for example The Sans Internet Storm Center services and The National Vulnerability Database (NVD). Vulnerability Identification Systems (VIS) aimed to identify vulnerabilities according to three categories: code, design, and architecture. Examples for VIS systems are The Common Vulnerabilities and Exposures (CVE), and The Common Weakness Enumeration (CWE).

This work uses NVD vulnerabilities database as an illustration of the proposed model.

Risk evaluation uses scoring systems which enable parameters estimation for estimating vulnerabilities' impacts on the organization. The Common Vulnerability Scoring System (CVSS) is a framework that enables user organizations receive IT vulnerabilities characteristics [1].

CVSS uses three groups of parameters to score potential risks: basic parameters, temporal parameters and environmental parameters. Each group is represented by score compound parameters ordered as a vector which is used to compute the score. Basic parameters represent the intrinsic specifications of the vulnerability. Temporal parameters represent the specifications of a vulnerability that might change over time due to technical changes. Environmental parameters represent the specifications of vulnerabilities derived from the local IT specific environment used by users' organization. CVSS enables omitting the environmental metrics from score calculations in cases that users' environment has no effect on the score and in cases the users do not specify the detailed description of environment and it's components.

CVSS is a common framework for characterizing vulnerabilities and predicting risks, used by IT managers, risk managers, researchers and IT vendors. It uses an open framework which enables managers to deal with organizations' risks based on facts rather than evaluations. Organizations adopting CVSS framework may gain the following benefits:

- A standard scale for characterizing vulnerabilities and scoring risks.
- Normalizing vulnerabilities according to specific IT platforms. The computed scores enable users make decisions according to vulnerability risks.
- CVSS uses an open framework. Organizations can see the characteristics of vulnerabilities and the logical process of scoring evaluation.
- Environmental scores. Organizations using the environmental parameters benefit by considering changes in its IT environment according to predicted risk scores.

There are few other vulnerability scoring systems besides CVSS differing by the parameters' specifications and scoring scales. CERT/CC emphasizes internet infrastructure risks. SANS vulnerability system considers users' IT configuration. Microsoft emphasizes attack vectors and vulnerabilities' impacts.

Using CVSS scoring system, basic and temporal parameters are specified and published by products' vendors who have the best knowledge of their product. Environmental parameters are specified by the users who have the best knowledge of their environments and business impacts.

This paper focuses mainly on environmental metrics.

Business damages caused by a vulnerability are influenced by the IT exploited component. CVSS environmental parameters specify the characteristics of a vulnerability that is associated with user’s IT configurations' components. Environmental parameters are of three groups:

1) Collateral Damage Potential (CDP).

A group of parameters which measure the economic potential damage caused by a vulnerability.

2) Target Distribution (TD).

Parameters indicating the percentage of vulnerable components in users' environment.

3) Security Requirements (CR, IR, AR).

Parameters indicating security importance measures in users' organization. Those parameters are subdivided to parameters indicating the confidentiality (CR), integrity (IR), and availability (AR). Higher security requirements may cause higher security damages on the organization.

Categorization of IT components according to security requirement measures should encompass all assets to raise the possibility of predicting organizational damages. Federal

Information Processing Standards (FIPS) requirements demands implementation of a categorization [6], but does not require using any particular scale, thus risk comparison of users' systems is difficult.

III. The Proposed Framework

Federal organizations are moving from periodic to continuous monitoring implementing SCM's which will improve national cyber security posture [7]. The proposed framework includes two capabilities which are not found in current practices. First, the environmental parameters are based on the components of the system as updated in the systems' Configuration Management Data Base (CMDB) [8]. This capability enables basing the scoring models to predict organizational damage potential relating to actual IT configuration rather than relying on user's estimates. According to [9] it is impossible for organizations to make precise estimates of the economic damages caused by an attack without having full knowledge of users' IT environment. Reference [10] [11] states that network configuration should be monitored continually and available vulnerabilities must be analyzed in order to provide the necessary security level.

Several researchers tried to simulate IT configuration processes using graphs. Researchers studied the impacts of component dependency graphs [12] [13] [14]. [15] Claims that CVSS does not take into consideration component dependencies, which impacts dramatically the exploitability of a vulnerability. [15] States that current CVSS do not reveal the fact that vulnerabilities on highly depended packages usually bring larger attack surfaces compared to those detected on a client application, even when they have the same CVSS scores. [15] Studied the impacts of components dependencies which refer to a code reuse by a component from the library packages that it relies upon. [16] Presents a risk estimation model that makes use of CVSS to produce security risk levels implemented as a Bayesian Belief Network (BBN) topology.

This research models the configuration using a visual directed graph to represent network structure describing network components and component' messages relationships. Visualization tools are used to model network structure or attack paths. Modeling attack paths enables analyzing network security to predict future attacks. Attack graphs can represent potential attack paths which an attacker can take to reach the system. According to [17] attack graphs act as a tool in finding critical paths in large networks based on the threats and vulnerabilities identified. The layout of an attack graph can be adjusted to represent the real enterprise network. [18] Proposes an attack graph-based probabilistic metric for network security. According to the model this research proposes, knowledge concerning the environmental components is represented as a directed graph which includes information on systems' components, the links which represent data reads/writes between components and systems' impacts on external users such as an error caused to a users' interface or errors in transactions routed to other interfacing systems. Each link is assigned a probability which resembles the occurrence probability of the specific link between the two components. Occurrence probabilities are computed regularly by monitoring the daily system' processes at production activities capturing all message passing among components. In the past such automated systems were not advanced, but according to [19], there are currently automated tools to generate visualization maps of systems activities (specifically for attack vectors monitoring), with the inputs from the system and its environment. Components' collateral damage potential scores are computed by activation of a rollback algorithm using the directed graph. Graph design represents all software activities processed by the system. The activities are initiated by external inputs which belong to the attack surface. Those external input components pass messages to internal components of the system which pass further messages to other components, ending at generation of external interfaces. This research assumption is that users are capable of estimation business damages relating to external interfaces only. They have no capability of estimating business damages relating to internal software components or to external systems' inputs. The focus of this work is in evaluating the potential damages to all input and internal components. The damage caused by an exploit to a vulnerable surface input component is computed by evaluating all message passing and impacts on neighboring components to all internal components according to their occurrence probabilities, ending at the generation of a wrong output, which is delivered to a certain user. The user is capable of estimating business damages caused by wrong information written on users' interfaces.

The proposed CMS model examines a database of published asset vulnerabilities, compares in real time computers' assets for existing exposures and calculates computers' potential damages, based on the directed graph. Risk scoring is performed by considering vulnerabilities even before patches are prepared and loaded on the computers' system. The CMS proposed architecture presented in Fig. 1. Following, a description of systems' components and processes.

- Continuous Monitoring System (CMS)

The system runs continuously and starts computing potential damages in two cases: first is whenever a new vulnerability is publishes and indicated in the NVD, second is whenever a change is made in a systems' component or in systems' interfacing component. Following a description of systems' modules.

- Vulnerabilities database (NVD).

Vulnerabilities database includes all known vulnerabilities and their specification as published by database owners. Examples of vulnerability specifications used by NVD are: vulnerability category, vendor name, product name, published start and end dates, vulnerability update dates, vulnerability severity, access vector, and access complexity [6].

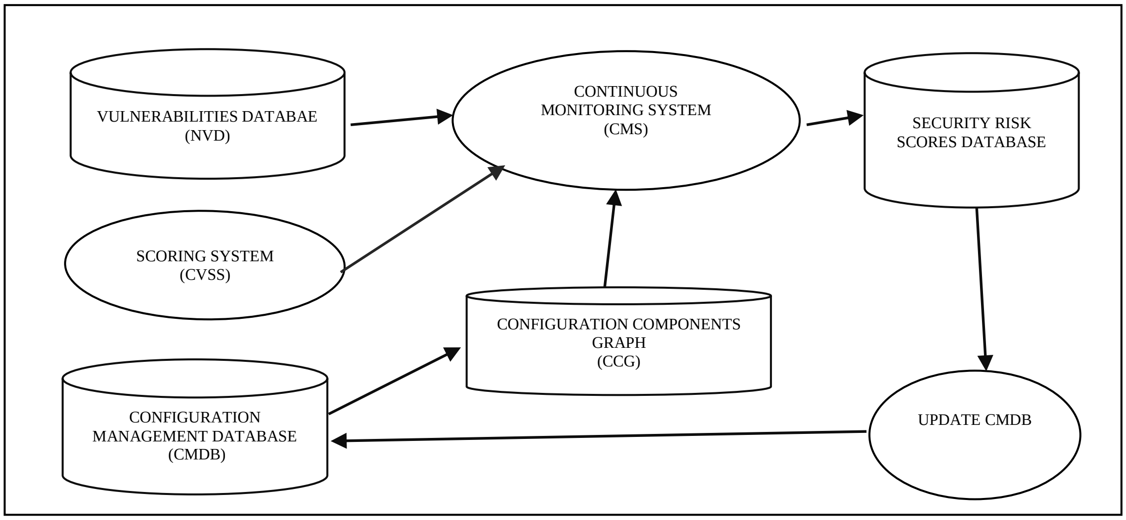


Fig. 1. Continuous Monitoring System architecture

- Scoring system (CVSS)

Scoring system (CVSS) is the algorithm this research uses for illustration of the proposed model. CVSS computes security risk scores according to parameter groups: basic, temporal and environmental. There are also other known scoring algorithms, some of them for public use other commercial.

- Configuration Management Database (CMDB).

CMDB is a database which includes all hardware and software components of the target system. According the proposed model the CMDB includes detailed information of the hardware and software. The CMDB contains detailed information of each module, systems' components and relationships among the components. Software is dealt in the resolution of programs, services and parameters. Data is handled in the resolution of databases, tables and data items. Input/output interfaces are handled using screen-names, reports and messages. The target system might be one computer or a group of organizations' computers. The CMDB includes all components in computers' configuration, components which interface with the target system directly or indirectly up to external and end-users' interfaces. The CMDB includes also the security requirements (CR, IR, AR) of each component in the resolution of data items' security requirements. Security requirements are specified by systems' owners according to business potential damages. CMDB includes also all interfaces among components. For each interface an indication of the direction of messages passing between the components and the probability of messages passing occurrence.

- Configuration Components Graph (CCG).

CCG is a directed graph including all the components in the CMDB organized as a directed graph which enables operating a rollback process which is aimed to compute components' collateral damages. The Rollback process starts at output external components, which potential damages has been assigned by users, continues backward to connected internal components, ending in input surface components [see Fig 2]. At the end of the rollback process all collateral damage potential values of systems' components are calculated. The graph includes three kinds of nodes: external inputs (IN), External output (OUT), and internal components (INTERNAL). Arcs between nodes represent message passing from one component to other components. Each arc is assigned a real number between [0, 1] representing the occurrence probability of the link between the connected nodes. External inputs represent all kinds of inputs to the system such as user interfaces, e-mails etc'. A subgroup of the external input components are surface attack components which might be of the following types [20].

1) Services available in the firewall which handles incoming messages

2) Systems code that processes incoming data, email, XML, office documents, industry-specific custom data exchange formats (EDI)

3) Interfaces, SQL, web forms

4) Employees accessing sensitive information

Messages are forwarded from the external inputs to internal components following to external outputs. Damage potential score evaluation goes in the backward direction: from external output scores (which are estimated by the users), back to their corresponding input components, and finally back to the corresponding external input components.

- Security Risk Scores Database.

The database includes all computed damage potential scores as computed by the CMS. The scores are then updated in the CMDB by the UPDATE CMDB process. The CMDB is used for retrieval purposes by business managers and analysts. Regular requests for damage potential score will be supplied by the CMDB. In cases of updates to systems' components or to NVD records, CMS will initiate an activation of the rollback process using the CCG. CMDB scores represent the damage potential evaluated updated scores for all systems' components including internal and external input components. This update process is needed to prevent unnecessary risk score heavy computations which were already evaluated and has been written in the past in the CMDB.

IV. The Risk Scoring Algorithm

CVSS's framework is based on three kinds of parameters: basic and temporal parameters are specified and published by products' vendors who have the best knowledge of their product. Environmental parameters are specified by the users who have the best knowledge of their environments and vulnerability business impacts. This work deals with the environmental parameters. According to [6], in many organizations IT resources are labeled with criticality ratings based on network location, business function, and potential for loss of revenue or life. For example, the U.S. government assigns every unclassified IT asset to a grouping of assets called a system. Every system must be assigned three “potential impact” ratings to show the potential impact on the organization if the system is compromised according to three security objectives: confidentiality, integrity, and availability. Thus, every unclassified IT asset in the U.S. government has a potential impact rating of low, moderate, or high with respect to the security objectives of confidentiality, integrity, and availability. This rating system is described within Federal Information Processing Standards (FIPS) 199.5 [21]. CVSS follows this general model of FIPS 199, but does not require organizations to use any particular system for assigning the low, medium, and high impact ratings. Reference [22] states that organizations should define the specifications of security risks of their specific environment, but does not define the ways organizations have to specify that information. The Department of State (State) has implemented an application called iPost and a risk scoring program that is intended to provide continuous monitoring capabilities of information security risk to elements of its information technology (IT) infrastructure. According to [23] the iPOST scoring model does not refine the base scores of CVSS to reflect the unique characteristics of its environment. Instead, it applied a mathematical formula to the base scores to provide greater separation between the scores for higher-risk vulnerabilities and the scores for lower-risk vulnerabilities. This work is targeted to fill-in this vacuum.

The CMDB defined in this work handles configurations' information of the system including the following entities: database tables, software components, system components such as operating system, database management systems, utility programs, development components, UI screens, etc. Each component is describes including knowledge relating to security requirements needed for operation of the risk scoring algorithm. The CMDB includes also all relationships among components, for example message passing to/from two components and function calls. The CMDB manages five kinds of environmental information for every system component. Table I includes information concerning the characteristics assigned to systems' components. Characteristic values are based on [21] definitions. The information is categorized according to its security type which is defined as a specific category of information (e.g., privacy, medical, proprietary, financial, investigative, contractor sensitive, security management). Reference [21] states that the potential impact is low if the loss of confidentiality, integrity, or availability could be expected to have a limited adverse effect on organizational operations, organizational assets, or individuals. The potential impact is moderate if the loss of confidentiality, integrity, or availability could be expected to have a serious adverse effect on organizational operations, organizational assets, or individuals. The potential impact is high if the loss of confidentiality, integrity, or availability could be expected to have a severe or catastrophic adverse effect on organizational operations, organizational assets, or individuals.

TABLE I. CMDB – Components Table

Column ID	Column Name	Column Description	Values (*)
COMPONENT ID	Software or Hardware	Value is equal to component ID in NVD	unique
COMPONENT TYPE	According to external or internal entities	I = Input external O = Output external INT = Internal	I, O, INT
CDP	Collateral Damage Potential	This metric measures the potential for loss of life or physical assets through damage or theft of property. The metric may also measure economic loss of productivity or revenue.	N, L, M, MH, H
TD	Target Distribution	This metric measures the proportion of vulnerable systems.	N,L,M,H
CR	Confidentiality Requirement	The importance of the affected IT asset to a user’s organization, measured in terms of confidentiality.	L,M,H
IR	Integrity Requirement	Guarding against improper information modification or destruction.	L,M,H
AR	Availability Requirement	“Ensuring timely and reliable access to and use of information…”.	L,M,H

(*) N=none, L=low, LM=low medium, M=medium, MH=medium high, H=high

Table II describes the relationships among couples of components which were described in Table I. A relationship between two components represents certain activities performed between the components for example read from an external input component, write to an external output component, function calls from one to another component. This table is used for generation of the directed graph. Each component will be represented by a node in the graph, and each relationship will be represented as an arc. This table includes description of all the relationships among systems' components. Each row represents one link between two components. Each link between two components is assigned a link probability which represents the occurrence probability of the specific activity, meaning the statistical probability that the input component will activate the output component relating to all the activities generated by that specific input component.

TABLE II. CMDB – LINKS TABLE

Column ID	Column Name	Column Description	Values
COMPONENT ID	Value is equal to component ID in NVD and CMDB – Components table. This in the input of the link	ID of the component which performs a certain activity on the component on the second end component of the link	I, INT (cannot be an output external component)
COMPONENT ID	Value is equal to component ID in NVD and CMDB – Components table. This in the output of the link	ID of the component which is impacted by certain activity of the components which is the input of the link	O, INT (cannot be an input external component)
LINK PROBABILITY		Probabilities distribution of all links from one input component to other components. The probability is calculated by monitoring the operational system.	A real number between [0.1] The sum of all the probabilities outgoing from one component us equal to 1.

The components directed graph is outlined using Tables I and II as described in Fig. 2.

Two external input nodes are component no' 1 and 2 which belong to the attack surface. Components 3,4,5,6 are internal. Components 7, 8 are external outputs. Arrows represent links among components. Following the graph structure formalism.

Let i be a component which activates components j and k. (The presented algorithm enables a varying number of linked components). Each link from i is assigned a value which indicates occurrence probability of the event that component i activates components j and k. For example the probability that component 2 activates component 3 is equal 0.2, while the probability of i activating component 4 is 0.8.

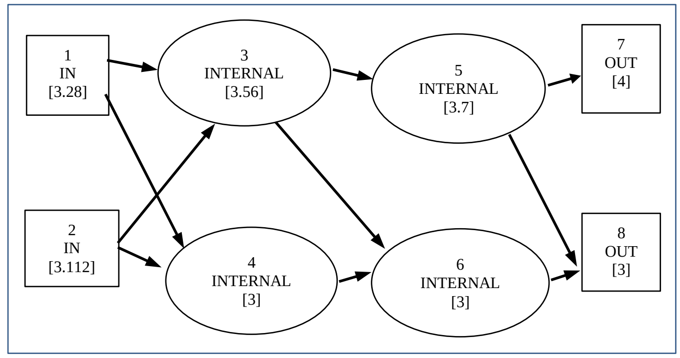


Fig. 2. Configuration Components Graph

In order to illustrate the scoring algorithm we use the graph of Fig 2 which is generated from the CMDB defined in Tables I and II. The contents of the CMDB follows in Tables III and IV.

TABLE III. CMDB – COMPONNETS TABLE EXAMPLE

Component ID	Component Type (*)	CDP	TD	CR	IR	AR
1	I	3.28	L	L	L	L
2	I	3.112	M	M	M	M
3	INT	3.56	M	M	M	M
4	INT	3	H	H	H	H
5	INT	3.7	L	L	L	L
6	INT	3	H	H	H	H
7	O	4	L	L	L	L
8	O	3	M	M	M	M

(*) I = INPUT, O= OUTPUT, INT=INTERNAL

TABLE IV. CMDB – LINKS TABLE EXAMPLE

Component ID	Component ID	Link Probability
1	3	0.5
1	4	0.5
2	3	0.2
2	4	0.8
3	5	0.8
3	6	0.2
4	6	1
5	7	0.7
5	8	0.3
6	8	1

Following an illustration of the rollback scoring process. The underlying assumption is that systems' owner is capable of estimating CDP values to external outputs since only those components have direct impact on business users. Business users are unable to assign CDP to internal component, nor to input components since they have no knowledge of the technical relationships between an internal software component on other components, nor impacts on his business. This research underlying assumption is contrary to [24] assumption who state that according to their scoring model the temporal and environmental metrics including CDP need to be specified by users with no differentiation between user interfaces and internal components. The rollback algorithm presented in this work calculates the CDP of all systems' component, based on two types of information: the CDP of the external components, and occurrence probabilities of links among all systems 'components. The rollback algorithm is performed according to the following formalization outlined in Fig. 3:

The algorithm activates a function running on graph nodes computing CDP's for all system components. The algorithm starts by computing the CDP's of the components which generate the external outputs and continues backwards until ending by computing the CDP's of the internal and external inputs. Following are the notations used by the algorithm and algorithms' logic.

Graph G represents all system components. Graph nodes represent components, arcs represent occurrence probabilities of links between couples of components.

Ci indicates component number i.

CDP (Ci) is the CDP of component Ci. CDP's are represented as real number between [0, 5] according to CVSS definitions. CDP value 0 indicates no potential damage, CDP value 5 indicates the maximal damage potential.

n indicates the number of components in graph G.

Input: A directed graph G including all Pr (i, j) assigned for all the nodes in G.
Input: CDP values assigned by system users to all external outputs. Internal and input CDP's are assigned to null.
Output: A set of computed CDP C(i) assigned to all nodes of graph G.

Method:

1. For all components of graph G. i running from n to 1.
2. While all CDP C(j) connected to C(i) as output are not equal null

$$\text{Compute } \mathrm{CDP(Ci)} = \sum_{j=1}^{m(i)} \mathrm{CDP(Cj)} * \mathrm{Pr(i,j)}$$

3. Return the set of computed CDP C(i) of graph G.

Fig. 3. Algorithm for computing CDP values for graph nodes

m (i) indicates the number of components linked to component i as output nodes. Each component i might be linked to a varying number of output components.

Pr (i, j) indicates the occurrence probability of activities performed by component i to component j. The sum of all the probabilities of activities performed by component i to all output connected components equals 1.

For illustration, following the computation of CDP's of all systems components according to the algorithm. The computations according to the iterations are shown in Table V. The computed CDP's are also presented inside brackets in each component of Fig 2.

TABLE V. CDP COMPONENTS COMPUTATIONS

Step Number	Component ID	CDP	Remarks
1	7	4	Estimated by systems' owner
2	8	3	Estimated by systems' owner
3	5	3.7	4 * 0.7 + 3 * 0.3
4	6	3	3 * 1.0
5	3	3.56	3.7 * 0.8 + 3 * 0.2
6	4	3	3 * 1.0
7	1	3.28	3 * 0.5 + 3.56 * 0.5
8	2	3.112	3.56 * 0.2 + 3 * 0.8

V. RESULTS

AS presented in Table V the algorithm calculated all CDP's starting from the external output users' estimates, continuing to all internal CDP's, ending with external input components. The user estimates only two CDP's of the external outputs 7 and 8. In steps 3 and 4 the algorithm calculates the CDP's of components 5 and 6. THE calculated CDP of component 5 is 3.7 higher than component 6 which is 3. The rational is that component 6 impacts are less harmful to external component number 8, which is CDP 3, whereas component 5 has higher impacts due to impacting on component 7 which has a higher CDP of 4 according to users' estimates.

Second example of the rational implemented by the algorithm is the CDP's calculated for input components 1 and 2. Component number 1 is an external input with calculated CDP of 3.28 while component number 2 with a calculated CDP of only 3.112. The rational is that input component 2 has less impacts on external outputs since it impacts more on external output component 8 than on component 7, together with the fact that component 8 CDP is less harmful to the user than output component 7 having a higher CDP of 4.

It was illustrated that all CDP's are calculated basing on users' CDP's estimates of external outputs only, while all internal and input components CDP's are calculated by the algorithm. This illustrates the advantage of the algorithm compared to other algorithms which are based on user's estimates, making to use of the technological characteristics of the specific environment and all relationships among components.

Questions remaining for future improvements include adding more information to the CCG. Present information includes occurrence probabilities of links, but it is reasonable to assume that varying external inputs causing varying probabilities. In such cases it might be logical to define a graph in which link probabilities depend on several external inputs, instead of relying on their average. Such a solution may be more accurate.

The algorithm computes CDP using the expected CDP's according to their corresponding occurrence probabilities. It should be said that with a minor change in the logic, CDP could be calculated according to the maximal damage potential instead of the expected potential. Such decision should be taken by business risk manager. Incorporating the CDP computed values in CVSS scoring model needs a minor modification to CVSS algorithm: using the calculated CDP's instead of the estimated CDP's for all systems' components. Using CVSS model needs no other modifications.

VI. Conclusions

This work presents a new framework of a Security Continuous Monitoring System, structure and mechanisms. The CMS uses CVSS scoring model for risk scoring operating in real time. According to the proposed model CVSS uses CDP's environmental parameters which are evaluated by the suggested algorithm, based on the technological configuration of the system, instead of CDP figures which are currently estimated using users' personal knowledge. A structure of a directed graph and scoring algorithm described and illustrated.

The model helps risk managers in estimating the organizational damages related to security risks, basing their estimates on the specific technological structure by using the algorithm. Using this model will bring more accurate estimates to vulnerability risks, thus enabling efficient risk mitigation plans and improved defense to organizations.

Further research of the model is incorporating more information in the scoring model such as detailed specifications of the configuration such as certain types of components (operation system, browsers, application, development languages etc') and modeling several relationship types among components such as varying kinds of links indicating varying relationships such as read activities, write activities and function calls. It might be reasonable to research the impacts of the varying component types on the evaluated CDP's and eventually on business risk scores.

More research is needed in supplying quantitative measures to the CVSS model. In our view CVSS model uses too many qualitative measures. At present most measures are based on users' estimates. Parameters such as TD – Target distribution may use the technological aspects of the configuration instead of users' rough estimates. Other environmental parameters such as confidentiality, integrity and availability requirements might also be based on models relating to quantifiable business damages to technological components.

References

[1] P. Mell, K. Scarfone, and S. Romanosky, “CVSS – A complete guide to the common vulnerability scoring system, version 2.0”, 2007.

[2] L. Langer, "Stuxnet: dissecting a cyber warfare weapon, security and privacy", IEEE, Volume 9 Issue 3, pages 49-51, NJ, USA, 2011.

[3] S. Tom and D. Berrett, "Recommended practice for patch management of control systems", DHS National Cyber Security Division Control Systems Security Program, 2008.

[4] A. Terje and R. Ortwin, "On risk defined as an event where the outcome is uncertain", Journal of Risk Research Vol. 12, 2009.

[5] Y. F. Nũez, "Maximizing an organizations' security posture by distributedly assessing and remeding system vulnerabilities", IEEE – International Conference on Networking, Sensing and Control, China, April 6-8, 2008.

[6] K. Dempsey, N. S. Chawia, A. Johnson, R. Johnson, A. C. Jones, A. Orebaugh, M. Scholl and K. Stine, "Information security continuous monitoring (ISCM) for federal information systems and organizations", NIST, 2011.

[7] M. G. Hardy, "Beyond continuous monitoring: threat modeling for real-time response", SANS Institute, 2012.

[8] A. Keller and S. Subramanianm, "Best practices for deploying a CMDB in large-scale environments", Proceedings of the IFIP/IEEE International conference and Symposium on Integrated Network Management, pages 732-745, NJ, IEEE Press Piscataway, 2009.

[9] M. R. Grimalia, L. W. Fortson and J. L. Sutton, "Design considerations for a cyber incident mission impact assessment process", Proceedings of the Intrnational Conference on Security and Management (SAM09), Las Vegas, 2009.

[10] I. Kotenko and A. Chechulin, "Fast network attack modeling and security evaluation based on attack graphs", Journal of Cyber Security and Mobility Vol. 3 No. 1 pp 27-46, 2014.

[11] Weintraub E., "Security Risk Scoring Incorporating Computers' Environment", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 4, 2016.

[12] I. Chowdhury and M. Zulkernine, "Can complexity, coupling, and cohesion metrics be used as early indicators of vulnerabilities?", In Proceedings of the 2010 ACM Symposium on Applied Computing, ACM, 2010.

[13] R.J. Ellison, J.B. Goodenough, C.B. Weinstock, and C. Woody, "Evaluating and mitigating software supply chain security risks", Technical report, DTIC Document, 2010.

[14] V.H. Nguyen and L.M.S. Tran, "Predicting Vulnerable Software Components with Dependency Graphs", Proceedings of the 6th International Workshop on Security Measurements and Metrics, NY, USA, 2010.

[15] S. Zhang, X. Zhang, X. Ou, L. Chen, N. Edwards, and J. Jin, "Assessing Attack Surface with Component-based Package Dependency", 9TH International Conference on network and system security, 2015, USA.

[16] S.H. Houmb, V.N.L. Franqueira, and E.A. Engum, "Quantifying security risk level from CVSS estimates of frequency and impact", The Journal of Systems and Software 83 (2010).

[17] V. Viduto, W. Huang, and C. Maple, "Toward optimal multi-objective models of network security", Survey. In: 17th International Conference on Automation and Computing, 10th, September 2011, University of Huddersfield, Huddersfield, United Kingdom, http://eprints.hud.ac.uk/22831/ retrieved April, 16, 2016.

[18] L. Wang, T. Islam, T. Long, A. Singhal, and S. Jajodia, "An Attack graph-Based Probabilistic Security Metric", IFIP International Federation for Information Processing 2008.

[19] V. Shandilya, and C. B. Simmons, and S. Shiva, "Use of Attack Graphs in Security Systems", Journal of Computer Networks and Communications, Vol. 2014.

[20] S. Northcutt, "The Attack Surface Problem", SANS Technology Institute-Security Laboratory – Defense in Depth Series, http://www.sans.edu/research/security-laboratory/article/did-attack-surface, retrieved April, 03, 2016.

[21] FIPS Publication 199 - Federal Information proccessing standards publication, "Standards for security categorization of federal information

and information systems", Department of Commerce, USA, February, 2004.

[22] E. Weintraub and Y. Cohen, "Continuous monitoring system based on systems' environment", ADFSL - Conference on Digital Forensics, Security and Law, May 19, 2015, Florida, USA.

[23] GAO – United States Government Accountability Office Report to Congressional Request, "Information security – state has taken steps to implement a continuous monitoring application but key challenges remain", July, 2011.

[24] J. A. Wang, M. Guo, H. Wang, M. Xia, & L. Zhou, "Ontology-based Security Assessment for Software Products", CSIIRW '09 Proceedings of the 5th Annual Workshop on Cyber Security and Information Intelligence Research: Cyber Security and Information Intelligence Challenges and Strategies, 2009.

Conservative Noise Filters

Mona M. Jamjoom
Department of Computer Sciences
College of Computer & Information Sciences
King Saud University, Riyadh, Saudi Arabia

Khalil El Hindi
Department of Computer Sciences
College of Computer & Information Sciences
King Saud University, Riyadh, Saudi Arabia

***Abstract*—Noisy training data have a huge negative impact on machine learning algorithms. Noise-filtering algorithms have been proposed to eliminate such noisy instances. In this work, we empirically show that the most popular noise-filtering algorithms have a large False Positive (FP) error rate. In other words, these noise filters mistakenly identify genuine instances as outliers and eliminate them. Therefore, we propose more conservative outlier identification criteria that improve the FP error rate and, thus, the performance of the noise filters. With the new filter, an instance is eliminated if and only if it is misclassified by a mutual decision of Naïve Bayesian (NB) classifier and the original filtering criteria being used. The number of genuine instances that are incorrectly eliminated is reduced as a result, thereby improving the classification accuracy.**

Keywords—component; Instance Reduction Techniques; Instance-Based Learning; Class noise; Noise Filter; Naive Bayesian; Outlier; False Positive

I. Introduction

In Machine Learning (ML), the quality of the training data can have a huge impact on the induced classifier. Noise can cause a learning algorithm to overfit the training data [1], which harms the classifier's performance. ML algorithms use different methods to mitigate the effect of noise. For example, decision-tree learning algorithms use pruning techniques [2] whereas neural networks use validation datasets to determine when to stop the training process [1]. In Instance-Based Learning (IBL), the effect of noise is mitigated by using a large number of several similar instances instead of just one as done by the k Nearest Neighbor (kNN) algorithm, where k (the number of neighbors) is usually set to 3. Another general approach that can mitigate the effect of noise is to use a noise-filtering algorithm that determines and eliminates the outlier instances [3], [4], [5], [6], [7], [8]. Although, most of these methods were designed for IBL methods, they can also be used to preprocess the training data before using them with other ML approaches, such as decision trees [9] and neural networks [10], [11]. The efficacy of ML methods, specifically kNN, is highly influenced by the quality of training data. This is most obvious when the number of neighbors, k, is set to one [12]; the problem is less severe when k is set to larger values (e.g., 3). Some Instance Reduction (IR) techniques [13], [14], [15] have been developed as noise filters specifically to tackle the noise problem.

In this work, we empirically show that these filters may eliminate some genuine instances because they mistakenly identify them as outliers due to the noise effect [13]. In other words, their FP (i.e., incorrectly eliminated noisy instances) error rate in identifying outliers is relatively high. This is because they use a relaxed outlier identification criterion, which is especially bad when the available training data are limited in size. This work proposes more conservative identification criteria to replace the outlier identification criteria of ENN [14], RENN, and All-kNN [15] noise filters. The proposed method uses the decision of NB classifier and the decision of the noise filter being used to determine whether to discard the instance or keep it.

The empirical results using 50 benchmark datasets obtained from UCI machine learning repository [16] show that the new method improves the performance of these noise filters at different noise ratios. The proposed conservative methods proved to be effective at minimizing the FP error rate. In other words, the methods managed to save more genuine instances and improved the classification accuracy. We present a comprehensive comparison between the methods in terms of the average classification accuracy, number of datasets in which the methods achieve better results and significantly better results, the average percentage of eliminated FP instances, the average ratio of data reduction, and the average percentage of True Positive (TP) instances (i.e., correctly eliminated noisy instances) eliminated by each algorithm.

The paper is organized as follows: Section 2 presents an overview of noise-filtering techniques used in the paper and NB classifier, section 3 presents the conservative criteria and empirically discusses and analyses the results, and section 4 concludes the paper.

II. Related Work and Background Material

In this section, we review the noise-filtering material that we modify as well as the Naïve Bayesian algorithm.

A. IR Techniques for Noise Filtering

The problem of noise in classification has been an active research area for many decades, and most machine learning algorithms focus on this issue. The problem has continued to be a major challenge due to the uncertainty property of the noise [17]. Various approaches to dealing with the problem of noise have been integrated with the learning algorithms to mitigate its effect and improve the learning capabilities. [18] categorized the techniques that handle noise into three main groups: robust, polishing, and filtering.

Robust techniques leave the noise in the dataset and use an embedded pruning phase to mitigate its effect. Such techniques are used in decision trees and rule learning. In decision trees some branches are pruned C4.5 [19], while in rule learning some preconditions of rules are pruned CN2 [20]. However, the classifier built from noisy a dataset may be less representative and less predictive if the noise ratio is very high. [21] preferred to handle the noise as a pre-processing phase, so that the constructed classifier is not affected by the noise.

The other two techniques (i.e., polishing and filtering) use more pure training data as the noise is preprocessed before any classifier is constructed; therefore, most studies tend to use such techniques [17]. Polishing techniques try to repair noisy instances by replacing the suspected attribute values with other appropriate values [22], [23]. The new values are determined based on the class of an instance and some additional attribute values. Correcting or relabeling an instance is a risky process because it can replace a noisy value with another [24], [18], [25]. Meanwhile, filtering techniques use an independent noise-filtering stage in which noisy instances that meet certain criteria are determined and discarded. Noise filtering has been implemented in different forms with different types of classifiers [26], [6], [27], [28], [4], [9], [10], [3], [5], [8], [29] and has been proven to be effective in improving the classification accuracy [25].

IR techniques were developed to speed up and reduce the storage requirements for IBL while preserving the classification accuracy [13]. Some such techniques are designed specifically to work as noise filters (e.g., ENN [14], RENN, and All-kNN [15]). The retained set of instances is purer and better represents the underlying instance space. The filtering techniques used in this paper include the following:

- The **ENN** [14] is a decremental algorithm that starts with the complete training set and eliminates an instance if it is misclassified by its k nearest neighbor(s). We set k to 3 in this work. The algorithm smooths the decision boundaries by removing the noisy instances, which are typically closer to the border. The pseudo code for the ENN algorithm is shown in Figure 1.
- The **RENN** [15] is a repeated form of ENN until no more instances can be removed. This will increase the gap between classes.
- **All-kNN** [15] is a batch algorithm that starts with a complete training set. It marks all instances misclassified by its i neighbors for all i = 1 to k. The elimination is done once, after all the instances in the training set are examined. Internal noisy instances within a class as well as odd instances on the border will be removed. The pseudo code for the All-kNN algorithm is shown in Figure 2.

```
T is a training set contains all training instances

For each instance (i)
If (instance(i).class<> majority class of k neighbors)
        Remove instance(i) from T
```

Fig. 1. Pseudo code for ENN algorithm

```
T is a training set contains all training instances
oldk=k
For each instance (i)
 For k=1 to oldk
   If (instance(i).class<> majority class of k neighbors)
          mark instance(i) to be removed
Remove all marked instances
```

Fig. 2. Pseudo code for All-kNN algorithm

B. *Naïve Bayesian (NB)*

The NB classifier is a simple form of Bayesian Network (BN) with one parent and several children [30]. It is a probabilistic classifier based on Bayes' theorem with strong (naïve) independence assumptions between the features, given the class [31]. To classify an instance, NB calculates the conditional probability for each instance class value and considers the class with the maximum probability as the predicted class, as shown in Equation 1.

$$\text{Class}_{\,predicted} = \underset{c \in C}{\operatorname{argmax}}\, p(c) \cdot \prod_j p(a_j|c) \qquad (1)$$

where, C is a vector of all class attribute values, $p(c)$ is the probability of class c, and $\prod_j p(a_j|c)$ is the naïve assumption that all the attribute values are conditionally independent given the class value. Domingos et al. [32] found that NB performance is competitive with more sophisticated ML methods, such as DT, IBL, and rule induction, even if the features' dependency is very strong. Moreover, NB is a strongly noise-tolerant algorithm [33], [34]. Nettleton et al. [33] performed a systematic analysis of robustness of many ML algorithms to noise—namely, NB, C4.5, IBk, and SMO. They determined that NB is the most noise-robust ML algorithm. An extended NB structure obtained from noisy data was presented in [35], in which the NB is constructed from noisy data and is incorporated with the NB model constructed from real data using linear equations and optimization methods. The proposed method was effective in handling noisy data and achieving classification accuracy. El Hindi [36] enhanced the performance of NB classification by using a fine-tuning stage to improve the probabilities estimation, but this degraded the sensitivity of NB toward noise. Therefore, several modifications for the fine-tuning process were proposed by [34]. They simply assign smaller weights for noisy instances during the fine-tuning process instead of eliminating or correcting them.

III. Conservative Noise Filters

In this section, we study and compare the performance of the reviewed noise filters and suggest more conservative criteria to identify outliers. To study their robustness to noise, we performed several experiments with different noise ratios. We paid special attention to the FP error rate [37] of each algorithm, because we believe that these noise filters mistakenly classify many genuine instances as outliers and, consequently, eliminate them.

TABLE I. CHARACTERISTICS OF DATA SETS USED IN THE EXPERIMENTS

Data set	# of Classes	# of Attributes	# of Instances
anneal	6	39	798
anneal.ORIG	6	39	798
arrhythmia	16	280	452
audiology	22	70	226
autos	7	26	205
Breast-cancer	2	10	286
Breast-w	2	10	699
bridges_version1	6	13	108
bridges_version2	6	13	108
car	4	7	1728
colic	2	23	368
colic.ORIG	2	28	368
congress-voting-1984	2	17	435
credit-a	2	16	690
credit-g	2	21	1000
cylinder-bands	2	40	512
dermatology	6	35	366
diabetes	2	9	768
ecoli	8	8	336
flags	8	30	194
glass	7	10	214
heart-c	5	14	303
heart-h	5	14	294
heart-statlog	2	14	270
hepatitis	2	20	155
hypothyroid	4	30	3772
ionosphere	2	35	351
iris	3	5	150
labor	2	17	57
lung-cancer	2	57	32
lymph	4	19	148
monk2	2	7	169
musk1	2	167	476
postoperative-patient-data	3	9	90
primary-tumor	22	18	339
segment	7	20	2310
sick	2	30	3772
solar-flare_1	2	13	323
solar-flare_2	3	13	1066
sonar	2	61	208
Soybean	19	36	683
spect_train	2	23	80
splice	3	62	3190
sponge	3	46	72
tennis	2	5	14
trains	2	33	10
Vehicle	4	19	846
Vote	2	17	435
Vowel	11	14	990
Zoo	7	18	101

In this work, the FP rate refers to the rate of mistakenly classifying genuine instances as outliers. Of course, a noise filter with a low FP rate is better than a noise filter with a high FP rate because the former helps retain more instances that are genuine. We must also take into account the False Negative (FN) rate, which refers to the rate of mistakenly classifying instances as genuine instances when they are in fact outliers. A noise filter with a high FN rate is bad because it would fail to eliminate all outliers.

We used 50 benchmark datasets obtained from the UCI Repository for Machine Learning [16] to experimentally test the FP and FN rates of the noise filters. We deliberately inserted class noise in the training sets by replacing the class values of some randomly selected instances with other random class values. We kept the class values in the test datasets unchanged and used different noise ratios of 0%, 5%, 10%, 15%, and 20%. Each noisy experiment was repeated five times. We used the kNN algorithm, with k = 3, and the discretized VDM (DVDM) as a distance function [38]. Ten-fold cross-validation and a paired t-test with a confidence level of 95% were used in all experiments. We compared the different methods with respect to the following criteria: average classification accuracy, the number of datasets in which each method achieved better, and significantly better results. We also calculated the average reduction size and the FP rate for each noise filter. The FP rate was calculated according to Equation 2 [39],

$$\text{The False Positive Rate (FPR)} = \frac{FP}{FP+TN} \quad (2)$$

where TN is the number of correctly retained instances (i.e., True Negative). We implemented kNN and the classical IR noise filters (i.e., ENN, RENN, and All-kNN) in the Weka work frame [40]. Table 1 lists the main characteristics of these datasets in terms of the number of class values, the number of attributes, and the number of instances.

A. The Prerforamnce of the IR Noise Filters

We simply applied the selected noise-filtering algorithms to the datasets and compared the results of the kNN before and after filtering. Table 2 summarizes the results. As shown in Table 2, at 0% noise, kNN outperforms all noise filters. In addition, All-kNN outperforms RENN, which is consistent with the results reported by [13] and [15]. However, the advantage of using noise filters is obvious when used with noisy training sets, especially when the noise ratio increases. At 5% noise, the noise filters start to influence the results in a positive way.

Although, at 5%, noise ENN shows better results than the rest of the filters, at 10%, 15%, and 20% noise, RENN emerges as the best noise filter; ENN is more conservative and thus more suitable when we have a low ratio of noise, yet when we have a large ratio of noise, we need a more aggressive algorithm, such as RENN. In general, applying a noise-filtering algorithm to noisy data before classification improves the classification accuracy, which is consistent with the findings of [13].

Table 3 shows the average reduction in the size of datasets as a result of applying each noise-filtering algorithm. RENN has the largest reduction in size rate among all algorithms due to its natural repeated elimination process.

It is obvious that the number of instances eliminated by the original noise-filtering algorithms is much greater than the number of outliers. For example, at 15% noise, the ENN, RENN, and All-kNN eliminate 27.4%, 30.65%, and 21.26% of the datasets, respectively, which is much higher than the noise ratio. Thus, these noise filters eliminate some genuine instances that are incorrectly classified by the filter as an outlier, resulting in the average percentage of eliminated FPs listed in Table 6 for these filters being considered high.

TABLE II. THE RESULTS OF THE KNN ALGORITHM BEFORE AND AFTER APPLYING EACH NOISE FILTER

Class Noise %	Criteria	kNN	ENN	kNN	RENN	kNN	All_kNN
0	Average Accuracy	**78.79**	78.07	**78.79**	77.10	**78.79**	78.06
	# Better datasets	**30**	16	**30**	16	**20**	12
	# Sign. better	**3**	1	**3**	1	**3**	2
5	Average Accuracy	76.66	**77.60**	**76.66**	76.50	**76.66**	76.32
	# Better datasets	9	**40**	16	**33**	14	**35**
	# Sign. better	2	**22**	7	**23**	8	**22**
10	Average Accuracy	75.68	**76.01**	**75.68**	75.52	**75.68**	75.64
	# Better datasets	14	**33**	15	**34**	16	**34**
	# Sign. better	2	**18**	9	**24**	12	**23**
15	Average Accuracy	73.61	**74.36**	73.61	**73.93**	73.61	**73.80**
	# Better datasets	14	**36**	16	**34**	19	**31**
	# Sign. better	2	**18**	6	**24**	8	**18**
20	Average Accuracy	71.84	**72.55**	71.84	**72.86**	71.84	**71.86**
	# Better datasets	14	**36**	14	**35**	22	**27**
	# Sign. better	2	**14**	5	**25**	9	**19**

TABLE III. THE AVERAGE REDUCTION IN SIZE OF THE NOISE FILTER AT DIFFERENT NOISE RATIOS

Class Noise %	ENN%	RENN%	All_kNN%
0	13.25	15.74	9.26
5	18.26	20.97	13.81
10	23.37	25.92	17.89
15	27.40	30.65	21.26
20	31.49	34.92	24.30

B. Hybrid Outlier-Identification Criteria

The fact that these noise filters have a high FP rate suggests that the noise filters are too relaxed or loose in determining the outliers. To improve their performance, we believe that their criteria for determining outliers need to be more conservative and restricted. As the NB learning algorithm is robust to noise, we use it to make the noise identification criteria more conservative. We consider an instance an outlier if the NB classifier misclassifies it; the condition is added to the original noise identification criteria of ENN, RENN, and All-kNN. As a result, we ultimately double check if an instance is really an outlier before discarding it. The new hybrid outlier-identification criteria aim to sift through the suspicious instances to determine the actual outliers. We expect that this approach will reduce the number of mistakenly eliminated FP instances. Mutual decision of more than one condition for elimination can be helpful in this case.

Mutual decision has been used with instance reduction techniques in [9]. The authors applied the DROP5 algorithm on instances marked by the All-kNN algorithm (i.e., AllKnnDROP5) as a pre-pruning phase before applying the rule induction on the reduced training set. The hybrid reduction techniques gave better results in terms of average accuracy when compared to ENN, AllKnn, and DROP5 techniques individually. Mutual decision looks relatively similar to the ensembles' voting concept [25], in which considering different methods can lead to better decisions and thus better performance, as shown in [9].

To evaluate the performance of the conservative noise filters, we re-performed the previously described experiments using these algorithms. We prefixed the name of the noise filters with NB (e.g., "NB_ENN") to distinguish the conservative algorithms from the original algorithms. Table 4 shows the results of proposed conservative filters and their original counterparts. When noise-free, the conservative filters achieved better than all other filters used. This is expected as they eliminate fewer genuine instances.

In the presence of noise, the conservative filter with ENN and All-kNN gave better results than their corresponding regular ENN and All-kNN filters. The improved accuracy during noise increases is not less on average than 0.4% and 0.7% for NB_ENN and NB_All-kNN, respectively. Combining NB with RENN does not enhance the original RENN; this will be explained shortly.

Of course, the conservative filters eliminate fewer instances than the original filters. This is obvious and has been confirmed by the average reduction size of each algorithm reported in Table 5. The average reduction in size of the noise filter algorithms is larger than the noise ratio.

On the other hand, this average is close to the noise ratio when we used the conservative noise filters. Comparing Tables 3 and 5 shows a big difference in reduction between the conservative filter and the original one.

Table 6 shows the average percentage of eliminated FP instances by each algorithm. As we can see, the All-kNN filter eliminated more FPs among other techniques' reach by an average of 12.08%, followed by RENN with an average elimination of 9.43% and ENN with an average of 5.57%.

Clearly, the number of FPs eliminated by the noise filters is high when compared with the number eliminated by the conservative filters, and this number increases as the noise ratio increases in most cases. For example, at the 20% noise ratio, ENN, RENN, and All-kNN eliminated 6.24%, 10.59%, and 15.43% of FP instances, respectively, compared to 5.41%, 7.3%, and 4.96% eliminated by the corresponding conservative filters. NB_All-kNN had the lowest average percentage of eliminated FPs among all conservative filters. The conservative filters save more FPs than the regular filter, thus the performance of NB_ENN and NB_All-kNN improved accordingly, as demonstrated in Table 4. This means that these retained instances are important and contribute positively in the classification as they increase the number of correctly retained instances (i.e., TN instances).

TABLE IV. SUMMARY OF RESULTS COMPARING THE CONSERVATIVE FILTER WITH THEIR COUNTERPARTS

Class Noise %	Criteria	NB_ENN	ENN	NB_RENN	RENN	NB_All_kNN	All_kNN
0	Average Accuracy	**79.14**	78.07	**78.17**	77.10	**78.96**	78.06
	# Better datasets	**27**	19	**28**	14	**26**	19
	# Sign. better	**6**	3	**3**	2	**2**	**2**
5	Average Accuracy	77.58	**77.60**	75.77	**76.50**	**77.08**	76.32
	# Better datasets	**28**	17	16	**31**	**26**	15
	# Sign. better	**9**	8	10	**16**	**15**	4
10	Average Accuracy	**76.42**	76.01	74.06	**75.52**	**76.38**	75.64
	# Better datasets	**27**	17	19	**29**	**30**	15
	# Sign. better	**12**	3	8	**18**	**12**	2
15	Average Accuracy	**74.78**	74.36	72.91	**73.93**	**74.78**	73.80
	# Better datasets	**29**	14	18	**30**	**30**	14
	# Sign. better	**17**	3	5	**19**	**15**	2
20	Average Accuracy	**73.07**	72.55	71.49	**72.86**	**72.60**	71.86
	# Better datasets	**29**	13	15	**32**	**30**	13
	# Sign. better	**16**	2	3	**21**	**22**	1

TABLE V. THE AVERAGE REDUCTION IN SIZE OF CONSERVATIVE FILTERS AT DIFFERENT NOISE RATIOS

Class Noise %	NB_ENN%	NB_RENN%	NB_ All_kNN%
0	5.23	8.35	6.29
5	7.63	11.03	8.33
10	9.79	13.51	10.37
15	11.89	15.88	12.35
20	13.68	18.02	14.09

TABLE VI. THE FPR FOR EACH ALGORITHM

Class Noise%	The average percentage of eliminated FPs					
	NB_ENN %	ENN %	NB_RENN %	RENN %	NB_All_kNN %	All_kNN %
5	4.28	5.07	6.59	8.43	4.46	8.76
10	4.66	5.29	6.84	8.98	4.58	10.93
15	5.1	5.68	7.03	9.7	4.82	13.19
20	5.41	6.24	7.3	10.59	4.96	15.43
Average	**4.86**	**5.57**	**6.94**	**9.43**	**4.71**	**12.08**

TABLE VII. THE FNR OF THE NOISE FILTERS

Class Noise%	The False Negative Rate					
	NB_ENN %	ENN %	NB_RENN %	RENN %	NB_All_kNN %	All_kNN %
5	22.32	19.67	27.11	16.31	34.62	32.19
10	25.2	22.23	27.82	16.81	35.89	33.38
15	28.15	24.63	29	17.87	38.31	35.34
20	31.11	27.38	30.46	18.98	40.03	36.84
Average	**26.7**	**23.48**	**28.6**	**17.49**	**37.21**	**34.44**

As Table 4 shows, NB_RENN does not improve the regular filter despite the number of FPs eliminated by RENN being high compared to the conservative filter. The reason must be related to the FN rate of the algorithms. In other words, NB_RENN must have a higher FN rate than RENN, which simply means it is less effective at identifying and eliminating all outliers. Therefore, we calculated the FN rate, which is defined as Equation 3 [39],

$$FNR=1-TPR \qquad (3)$$

where TPR is defined as Equation 4 [39],

$$\text{The True Positive Rate (TPR)} = \frac{TP}{TP+FN} \qquad (4)$$

where FN is the number of incorrectly retained instances (i.e., False Negative).

Table 6 and Table 7 show the FPR and FNR for each algorithm at different noise ratio. As expected, Table 7 shows that RENN has the lowest FNR at all noise ratios; making it more conservative by combining it with NB substantially increases the FNR. For example, at 0% noise, RENN has an FNR of 16.31% while NB_RENN has 27.11%. Meanwhile, at 20% noise, RENN has an FNR of 18.98% while NB_RENN has 30.46%. Thus, although making RENN more conservative reduces its FPR (see Table 6), it also substantially reduces its ability to identify and eliminate all outliers. The problem is less severe for the other two algorithms because making them more conservative slightly increases their FNR.

Table 8 shows the results of the kNN before and after applying the conservative noise-filtering algorithm. Table 2 and Table 8 indicate that the best conservative algorithm is the NB_ENN whereas the best non-conservative algorithm is RENN. Therefore, it is logical to compare the effect of each compared to the kNN algorithm. Table 9 shows the result of this comparison, demonstrating that the NB_ENN's performance is much better than or equal to that of RENN on all noise for all ratios using all comparison criteria. The only exception is at 20% noise, where RENN achieves significantly better results than kNN for 20 datasets whereas NB_ENN achieves significantly better results than kNN for 13 datasets. However, at the same noise ratio, NB_ENN still achieves better average accuracy than RENN and achieves better results (but not significantly better) than kNN for 28 datasets whereas RENN achieves better results for 21 datasets. Thus, using the conservative ENN (i.e., NB_ENN) is probably better than using RENN, especially for small noise ratios ranging from 0% to 15%.

TABLE VIII. THE RESULTS OF KNN BEFORE AND AFTER APPLYING THE CONSERVATIVE NOISE FILTERS

Class Noise %	Criteria	kNN	NB_ENN	kNN	NB_RENN	kNN	NB_All_kNN
0	Average Accuracy	78.79	**79.14**	**78.79**	78.17	78.79	**78.96**
	# Better datasets	21	**25**	**26**	21	20	**23**
	# Sign. better	**4**	1	**3**	1	**3**	2
5	Average Accuracy	76.66	**77.58**	**76.66**	75.77	76.66	**77.08**
	# Better datasets	7	**42**	18	**31**	11	**38**
	# Sign. better	2	**27**	7	**20**	4	**23**
10	Average Accuracy	75.68	**76.42**	**75.68**	74.06	75.68	**76.38**
	# Better datasets	11	**38**	18	**32**	14	**34**
	# Sign. better	3	**22**	13	**19**	9	**24**
15	Average Accuracy	73.61	**74.78**	**73.61**	72.91	73.61	**74.78**
	# Better datasets	15	**35**	20	**30**	17	**33**
	# Sign. better	1	**19**	8	**19**	8	**24**
20	Average Accuracy	71.84	**73.07**	**71.84**	71.49	71.84	**72.56**
	# Better datasets	11	**39**	15	**35**	22	**27**
	# Sign. better	2	**15**	9	**21**	9	**21**

TABLE IX. COMPARING THE BEST CONSERVATIVE (NB_ENN) AND THE BEST NON-CONSERVATIVE (RENN) ALGORITHMS

Class Noise %	Criteria	kNN	NB_ENN	Difference= NB_ENN-KNN	kNN	RENN	Difference= RENN-KNN
0	Average Accuracy	78.79	**79.14**	**0.35**	**78.79**	77.1	**-1.69**
	# Better datasets	21	**25**	**4**	**30**	16	**-14**
	# Sign. better	**4**	1	**-3**	**3**	1	**-2**
5	Average Accuracy	76.66	**77.58**	**0.92**	**76.66**	76.5	**-0.16**
	# Better datasets	7	**42**	**35**	16	**33**	**17**
	# Sign. better	2	**27**	**25**	7	**23**	**16**
10	Average Accuracy	75.68	**76.42**	**0.74**	**75.68**	75.52	**-0.16**
	# Better datasets	11	**38**	**27**	15	**34**	**19**
	# Sign. better	3	**22**	**19**	9	**24**	**15**
15	Average Accuracy	73.61	**74.78**	**1.17**	73.61	**73.93**	**0.32**
	# Better datasets	15	**35**	**20**	16	**34**	**18**
	# Sign. better	1	**19**	**18**	6	**24**	**18**
20	Average Accuracy	71.84	**73.07**	**1.23**	71.84	**72.86**	**1.02**
	# Better datasets	11	**39**	**28**	14	**35**	**21**
	# Sign. better	2	**15**	**13**	5	**25**	**20**

IV. CONCLUSION

The problem of classical noise filters is that they eliminate large numbers of good instances when such instances are incorrectly identified as outliers and consequently eliminated. These good instances are useful for improving the classification accuracy of the induced classifier. Therefore, this work proposed a simple modification for these algorithms to make them more conservative. We proposed using hybrid outlier-identification criteria by combining an NB classifier with the original filtering criteria used by the algorithm. This work empirically shows that the conservative filters outperform the original filters because they have a smaller false positive rate (i.e., eliminate fewer genuine instances). The only exception is the conservative RENN (i.e., NB_RENN), which performs poorly compared to RENN despite the fact that NB_RENN has a smaller false positive rate, but a much higher false negative rate than RENN. Consequently, NB_RENN performs poorly compared to RENN, especially at large noise ratios. However, the conservative ENN (i.e., NB_ENN) outperforms RENN, especially at small noise ratios (e.g., from 0% to 15%). Future work should develop and investigate more hybrid noise-filtering criteria.

REFERENCES

[1] T. Mitchell, Machine Learning, vol. 4, no. 1. McGraw Hill, 1997.

[2] J. Quinlan, “C4. 5: programs for machine learning,” Mach. Learn., vol. 240, p. 302, 1993.

[3] H. Liu and S. Zhang, “Noisy data elimination using mutual k-nearest neighbor for classification mining,” J. Syst. Softw., vol. 85, no. 5, pp. 1067–1074, May 2012.

[4] F. Muhlenbach, S. T. Ephane, and D. A. Zighed, “Identifying and Handling Mislabelled Instances,” J. Intell. Inf. Syst., vol. 22, no. 1, pp. 89–109, 2004.

[5] F. Pasquier, S. Delany, and P. Cunningham, “Blame-based noise reduction: An alternative perspective on noise reduction for lazy learning,” Dublin, Trinity Coll. Dublin, Dep. Comput. Sci., pp. 1–17, 2005.

[6] N. Segata and E. Blanzieri, “Noise reduction for instance-based learning with a local maximal margin approach,” J. Intell. Inf. Syst. 35, no. October, 2010.

[7] N. Segata, E. Blanzieri, and P. Cunningham, “A scalable noise reduction technique for large case-based systems,” 8th Int. Conf. Case-based Reason. (ICCBR 09), vol. volume 565, pp. 328–342, 2009.

[8] X. Zeng and T. Martinez, “A Noise Filtering Method Using Neural Networks,” in Soft Computing Techniques in Instrumentation, Measurement and Related Applications, 2003, pp. 26–31.

[9] O. Othman and C. H. Bryant, “Preceding Rule Induction with Instance Reduction Methods,” Proc. 9th Int. Conf. Mach. Learn. Data Min.

Pattern Recognition. Lect. Notes Comput. Sci. , Springer-Verlag, Berlin, pp. 209–218, 2013.

[10] M. El Hindi, KAL-Akhras, "Smoothing decision boundaries to avoid overfitting in neural network training," Neural Netw. World, vol. 21, no. 4, pp. 311–326, 2011.

[11] K. El Hindi and M. Alakhras, "Eliminating border instance to avoid overfitting," Antonio Palma dos Reis (Ed.). Proceeding Intell. Syst. Agents, pp. 93–99 . IADIS press Algarve, Portugal, 2009.

[12] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algorithms," Mach. Learn., vol. 6, no. 1, pp. 37–66, Jan. 1991.

[13] D. R. Wilson and T. R. Martinez, "Reduction techniques for instance-based learning algorithms," Mach. Learn., vol. 38, no. 3, pp. 257–286, 2000.

[14] D. L. Wilson, "Asymptotic Properties of Nearest Neighbor Rules Using Edited Data," IEEE Trans. Syst. Man. Cybern., vol. 2, no. 3, pp. 408–421, 1972.

[15] I. Tomek, "An experiment with the edited nearest-neighbor rule," IEEE Trans. Syst. Man. Cybern., vol. 6, no. 6, pp. 448–452, 1976.

[16] "UCI Machine Learning Repository." [Online]. Available: http://archive.ics.uci.edu/ml/machine-learningdatabases/.

[17] H. Yin and H. Dong, "The Problem of Noise in Classification :Past,Current and Future work 1 2 1," in Communication Software and Networks (ICCSN), 2011, pp. 412–416.

[18] C. M. Teng, "A Comparison of Noise Handling Techniques," Proc. Fourteenth Int. Florida Artif. Intell. Res. Soc. Conf., pp. 269–273, 2001.

[19] J. R. Quinlan, "Induction of Decision Trees," Mach. Learn., pp. 81–106, 1986.

[20] P. Clark and T. Niblett, "The CN2 rule induction algorithm," Mach. Learn., vol. 3, no. 4, pp. 261–284, 1989.

[21] D. Gamberger, N. Lavrac, and S. Dzeroski, "Noise detection and elimination in data preprocessing: Experiments in medical domains," Appl. Artif. Intell. An Int. J., vol. 14, no. 2, pp. 205–223, 2000.

[22] C. M. Teng, "Correcting noisy data," Proc.\ Intl.\ Conf.\ Mach. Learn., 1999.

[23] C. M. Teng, "Polishing blemishes: issues in data correction," IEEE Intell. Syst., vol. 19, no. 2, pp. 34–39, 2004.

[24] Y. Yang, X. Wu, and X. Zhu, "Dealing with predictive-but-unpredictable attributes in noisy data sources," Proc. 8th Eur. Conf. Princ. Pract. Knowl. Discov. databases, Pisa, Italy, 2004.

[25] C. E. Brodley and M. A. Friedl, "Identifying Mislabeled Training Data," J. Artif. Intell. Res., vol. 11, pp. 131–167, 1999.

[26] S. Massie, S. Craw, and N. Wiratunga, "When Similar Problems Don ' t Have Similar Solutions," Proc. 7th Int. Conf. Case-Based Reason. (ICCBR 07), Springer-Verlag, Berlin, Heidelb., pp. 92–106, 2007.

[27] S. J. Delany, N. Segata, and B. Mac Namee, "Profiling instances in noise reduction," Knowledge-Based Syst., vol. 31, pp. 28–40, Jul. 2012.

[28] B. Frénay and M. Verleysen, "Classification in the Presence of Label Noise : a Survey," Neural Networks Learn. Syst. IEEE Trans., vol. 25, no. 5, pp. 845–869, 2014.

[29] S. Verbaeten and A. Van Assche, "Ensemble Methods for Noise Elimination in Classification Problems," Proc. 4th Int. Conf. Mult. Classif. Syst., vol. 2709, pp. 317–325, 2003.

[30] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," Artif. Intell. Rev., vol. 26, no. 3, pp. 159–190, Nov. 2007.

[31] G. H. G. John and P. Langley, "Estimating Continuous Distributions in Bayesian Classifiers," in In Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence. Montreal, Quebec, Canada, 1995, vol. 1, pp. 338–345.

[32] P. Domingos and M. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Los," Mach. Learn., vol. 29, no. 2–3, pp. 103–130, 1997.

[33] D. F. Nettleton, A. Orriols-Puig, and A. Fornells, "A study of the effect of different types of noise on the precision of supervised learning techniques," Artif. Intell. Rev., vol. 33, no. 4, pp. 275–306, Jan. 2010.

[34] K. El Hindi, "A noise tolerant fine tuning algorithm for the Naïve Bayesian learning algorithm," J. King Saud Univ., vol. 26, pp. 237–246, 2014.

[35] Y. Yang, Y. Xia, Y. Chi, and R. Muntz, "Learning naive Bayes classifier from noisy data," … Calif. Los Angeles, Dep. …, no. 030056, pp. 1–19, 2003.

[36] K. El Hindi, "Fine tuning the Naïve Bayesian learning algorithm," AI Communcations, vol. 27, pp. 133–141, 2014.

[37] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," Inf. Process. Manag., vol. 45, no. 4, pp. 427–437, 2009.

[38] D. R. Wilson and T. R. Martinez, "Improved Heterogeneous Distance Functions," J. Artif. Intell. Res., vol. 6, pp. 1–34, 1997.

[39] T. Fawcett, "ROC Graphs: Notes and Practical Considerations for Data Mining Researchers," 2003.

[40] U. of Waikato, "WEKA: The Waikato Environment for Knowledge Acquisition." [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka/.

Awareness Training Transfer and Information Security Content Development for Healthcare Industry

Arash Ghazvini
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia,
43600 UKM, Bangi Selangor, Malaysia

Zarina Shukur
Faculty of Information Science and Technology,
Universiti Kebangsaan Malaysia,
43600 UKM, Bangi Selangor, Malaysia

***Abstract*—Electronic Health Record (EHR) becomes increasingly pervasive and the need to safeguard EHR becomes more vital for healthcare organizations. Human error is known as the biggest threat to information security in Electronic Health Systems that can be minimized through awareness training programs. There are various techniques available for awareness of information security. However, research is scant regarding effective information security awareness delivery methods. It is essential that effective awareness training delivery method is selected, designed, and executed to ensure the appropriate protection of organizational assets. This study adapts Holton's transfer of training model to develop a framework for effective information security awareness training program. The framework provides guidelines for organizations to select an effective delivery method based on the organizations' needs and success factor, and to create information security content from a selected healthcare's internal information security policy and related international standards. Organizations should make continual efforts to ensure that content of policy is effectively communicated to the employees.**

Keywords*—information security; human error; awareness training program; training content; security policy; electronic health record

I. Introduction

The general objective of this paper is to enhance effectiveness of information security awareness training programs. More specifically, this paper aims to develop a framework that works as a guideline for organizations to select the right training delivery method that produces desirable outcomes. An effective training method must fulfil organization's needs and requirements while taking into account employees preferences. The paper also offers recommendations on how to augment internal information security policy document in healthcare sectors.

This study is influenced by importance of personal data protection to encourage researchers to study factors influencing users' behavior and attitudes toward information security, which impacts integrity of healthcare organizations. Security breaches are inevitable threats that always challenges organizations distinctively. Thus, organizations need to safeguard vital information and assets to prevent organization's integrity from being compromise. Information security breaches result in both direct cost (e.g. loss of intellectual property) and indirect cost (e.g. loses of reputation and potential loss in market share).

Human error is considered as the biggest threat to information security effectiveness [3]. Lack of employees' attention to security policy and standards is the real threat to information system. According to IT security practitioners survey conducted by [17], minimum of 78% of security breaches experienced by organizations are as a result of employees' negligence (Fig 1). Nevertheless, human error can be minimized through awareness training programs [4]. The significance of information security is best defined as the level of user comprehension on Information security awareness. In every organization, employees have varying knowledge of information security awareness [20].

Human errors are categorized into normal human errors and abnormal human errors. Normal human errors refer to individual honest mistakes that are already recognized and can be prevented in advance [5]. These kind of errors can be corrected through training programs with an intention to promote behaviors of individuals toward organizational policy. Education and training programs in organizations can help to improve employees' awareness toward security of e-health system and help them to adhere to appropriate behaviors that do not compromise the security of the system.

In what follows, the background of the study is presented. Section three illustrates the research design. The conceptual framework is presented in section four of this paper. Section five and six discus training content and information security document, respectively. Section seven focuses on training delivery methods. Conclusion is the last section of the paper.

II. Background

Even though the number of information security awareness training programs are growing progressively, there is inadequate evidence to verify their effectiveness and impact on daily activities in a work environment [21]. Literature [6][13] has stated that some of the information security awareness training programs are not effective enough. For instance, number of awareness training programs tends to be more informative without integrating into employees' daily activities that leads to disciplinary actions. Some other awareness training programs are only provided as one-time session that cannot truly change users' behavior toward information system. Awareness training programs should be a regular activity and reinforced periodically.

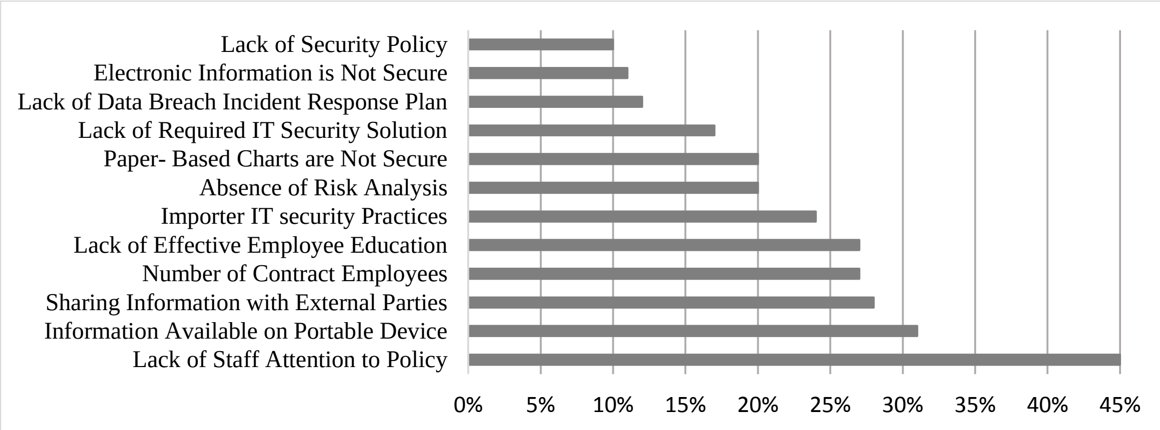


Fig. 1. Factor that Mostly Put Data at Risk

Source: [17]

Training and awareness programs are an effective approach to reduce the risk of individual contribution in electronic health system. However, routine, traditional training programs have been failed because they do not involve critical thinking and do not require users to think about security concepts [19]. On the other hand, there are new interactive training approaches that have been successful to engage employees with training activities including computer games, web-based sessions, E-learning, teleconferencing, and crossword puzzles. The key to impart a concept is to hold users engaged sufficiently long so he or she will absorb that concept, especially when the training program is mandatory.

Another reason for awareness training failure is some of training programs are too advanced for trainees. Employees, particularly those with no advance computer skills such as ordinary staff working in healthcare clinics, have different level of computer knowledge, and thus they require to be trained differently. It was also observed that most of trained employees do not attempt to apply the learned skills in work environment [6][13]. Moreover, many awareness training programs do not measure users' performance before and after the training, and therefore, it is not possible to evaluate the training's outcome. Additionally, number of employees are not motivated to contribute in awareness training program as the program do not promote creative activates [14]. Similarly, [19] argued that traditional and routine information security awareness training programs have been failed because they do not involve critical thinking and do not reinforce users to think about security concepts.

Effective awareness training techniques should be differentiated from ineffective ones [15]. Literature [21] stated that it is essential to increase the effectiveness of information security awareness training programs by encouraging employees to make effort in transferring the skills learned to their daily job activities. It is important to understand and emphasize the factors that differentiate effective trainings from ineffective trainings. Consequently, the existing gap of information security awareness training programs should be bridged to refine and improve the effectiveness of training programs [16][9].

Furthermore, Content of information security awareness training program should be developed from organization's policy based on the selected training technique. Each training approach requires specific content structure to be fitted in the program. The main objective of training content is to enforce information security policy document. Professional and complicated training material makes employees confused or bored towards subject matters. Hence, the training content should be easy to comprehend to motivate the trainees to learn as well as ensuring the delivery of selected content. Exceptionally, in the domain of information security, developers of most training programs are the experts who do not take audience's profiles into account.

III. Research Design

Training is ineffective unless translated into individual performance [23]. Effectiveness of information security awareness is often an overlooked element of an organization's security program. There is a broad range of awareness training delivery methods. However, research is insufficient regarding the effectiveness of delivery methods [1]. Similarly, a side-by-side comparison of different awareness delivery methods of information security is lacking [1][6]. Training programs can truly make a difference in employees' performance, and hence, it is important to understand effective transfer of training in organizations.

The aim of this research is, firstly, to develop a conceptual framework for effective transfer of training, and an opt-in framework for selecting an effective awareness training delivery method. Secondly, to provide a side-by-side comparison of different information security awareness delivery method. This guideline will help organizations to effectively select a delivery method and design a training program based on the organizations' needs and success factors. Lastly, to offer insights on augmentation of internal information security policy document to be used by healthcare organizations.

Hospital Universiti Kebangsaan Malaysia (HUKM) is one of the leading healthcare organizations in Malaysia that has adopted electronic healthcare systems. HUKM is selected as primary healthcare to collect necessary information to conduct this study. A series of semi-structured interviewees were conducted with HUKM decision makers in order to obtain necessary information to design the framework. The collected data is significant to create information security content for the selected healthcare and vital in the process of developing the framework. The framework can be used as a guideline to adapt an effective technique for information security awareness training programs. This guideline will be used to help decision makers to measure strength and weaknesses of each awareness training technique with respect to the organization's need. The developed framework can be used by any organization to select a successful awareness training program.

Even though HUKM is ISO certified, nevertheless, there are insufficient details or outdated sections in the internal information security policy document. To create appropriate information security content for awareness training program at HUKM, this study attempts to augment hospital's internal information policy document based on relevant international policies, as explained in next sections. The purpose of augmentation is to encourage HUKM's policy makers to update their internal information security policy. However, it shows the process of content creation to be followed by other healthcare organizations to enhance their internal security policy document. This is an inevitable stage before creating content or selecting a proper awareness training technique.

A. *Transfer of Training*

Holton (1996) developed a training conceptual model that focuses on individual performance. Learning, individual performance, and organizational results are three primary outcomes of the model for training intervention. These outcomes are described, respectively, as the learning outcome achievement desired by an organization, learning being applied on the job as a result of change in individual performance, and results at the organizational level as a consequence of change in individual performance. Fig 2 demonstrates the Holton's transfer of training model. Holton's model suggests that transfer of training is affected by three crucial factors including motivation to transfer, transfer climate, and transfer design. Only when the three primary influences on transfer behavior are at their appropriate levels, learning is expected to lead to change individual performance [9].

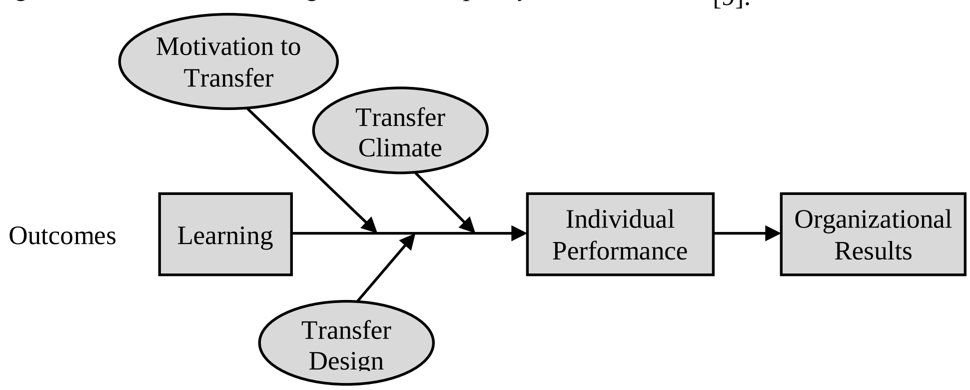


Fig. 2. Holton's Factors Affecting Transfer of Training, (1996)

IV. Conceptual Framework

The conceptual framework of this study is influenced by Holton's model (1996) (Fig 3). New components are added to enhance the pre-existing model. Holton proposes motivation to transfer, transfer design, and transfer climate as three major components of transfer of training that directly influence individuals' performance. This paper suggests that primarily effect of these components is on the choice of training delivery method [9]. Holton stated that learning is expected to lead to individual performance change. However, this study argues that the impact of learning on individuals' performance is affected by training delivery method. Therefore, it is important to put extra attention in selecting an effective method, especially when it comes to information security. Rapid changes, new trends, and security concerns requires constant update of learning outcomes. Demand for awareness is increasing every day and new training methods are introduced and utilized by organizations. If awareness training is effective, then it will significantly enhance individuals' performance. Hence, it is important to select an effective awareness training method that can fulfill organizations' need and requirements. It is also vital to consider training success factors and organizations' culture in order to succeed.

Holton stated that for training program to yield organizations results, it should have the ability to achieve results and motivate the organization and individuals to practice. Again, when it comes to motivation, training method plays a significant role. Effective training methods promote individual willingness not only to participate in training program but also to apply lesson learnt in their daily tasks. Moreover, Information security awareness training program is not a onetime task and it needs to be repeated frequently in order to improve the result and measure its effectiveness in enhancing employees' awareness. The aim of awareness training programs is to better safeguard organizations' assets and information security infrastructure.

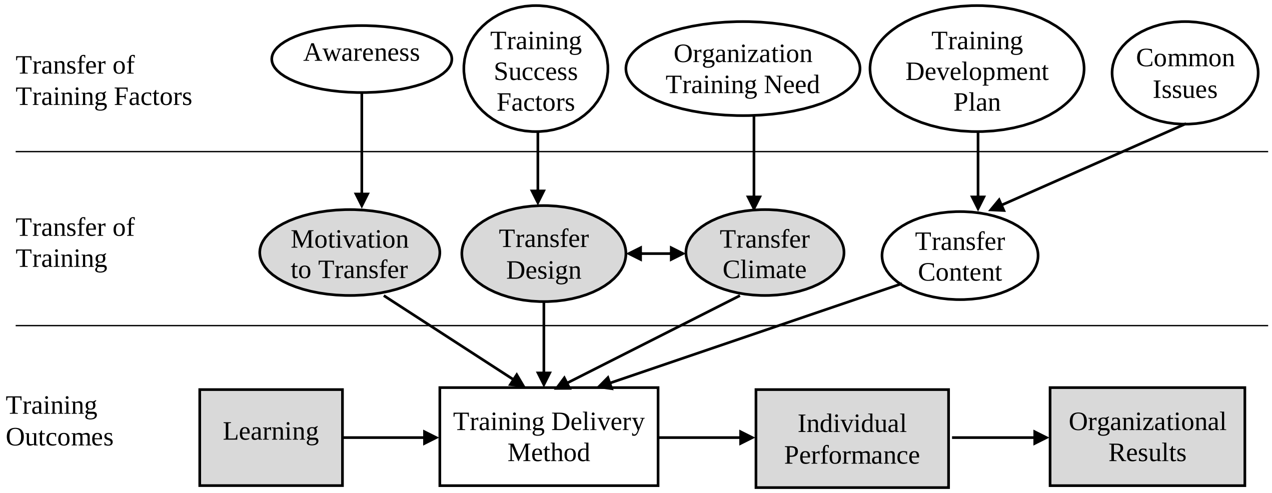


Fig. 3. Factors Affecting Transfer of Information Security Awareness Training Program

A. *Motivation to Transfer*

Hilton defined Motivation to transfer as individual's desire to transfer the necessary knowledge and skills in the training program on the job. This paper argues that awareness can motivate employees to learning. Employees are not aware of their roles to mitigate security issues. Survey conducted by [17] titled *Human Factor in Data Protection* revealed that most of risks in security breaches are driven from a lack of attention by staff to the security policy of an organization. Employees need to be aware of their important role in protecting organization's vital information to avoid compromising the system by rookie mistakes. Understanding the importance of their role in security effectiveness will motivate employees to attend training programs and incorporate their learning into their job performance.

B. *Transfer Climate*

The transfer climate arises from employees' perception of their work environment. It influences the degree in which employees apply the learned skills on the job. Holton defined transfer climate as a mediating variable in the relationship between the organizational context and employees' job attitudes and work behavior. Similarly, an effective awareness program must be based on the characteristics of an organization including organization size, business requirement, budget, target audience, and organization mission and culture. A properly designed awareness program that is in line with organization's need will effectively influence employees' attitude and work behavior.

It is crucial to constantly enhance the information security awareness culture in organizations and transform this culture into actual behavior. One way to improve training climate is to distinguish how different organizations have different needs. A more efficient and cost effective approach to implementing an employee security awareness model is to use a specific program that addresses the specific needs of the organization. Awareness programs must be designed with intention of creating organizational-wide security-minded cultures so that people work in a more secure manner and protect the assets of their organization [1].

C. *Transfer Design*

Holton's model did not provide guidelines to explain what constitutes appropriate transfer designs. According to Holton, the main failure to transfer is training design. In the context of information security, there are various types of security awareness delivery methods adapted by organizations. However, as stated by [2][6], many programs are not effective enough to change employees' attitude and work behavior. Awareness programs often seem less likely to improve employees' performance and many programs fail to enhance expected outcomes. The training itself has a direct influence on transfer of training and the key is to design an effective awareness program. Even though there are some researches on the efficiency of various information security delivery methods, but research is scant regarding effective delivery method of information security awareness.

This study provides guidelines for organizations on how to decide on an effective awareness delivery method for their organization. Enhancing an effective awareness program requires decision makers to make critical decisions about training delivery method as well as *training success factors*. Although it is important for a security awareness program to ensure that the appropriate topics are covered, it is vital to select the right delivery methods [20]. As with any program, the success of information security awareness program will rely heavily on how the awareness information is delivered [1]. As stated by [7], "*The lecture as a teaching tool is dead. Current programs don't work because we rely on old models of teaching. People learn in different ways. Some people are visual learners, while others learn better from reading or discussing. We need to move away from canned* web-*delivered training to interactive, hands-on learning to build more effective security awareness programs*" [7].

D. *Transfer Content*

This study suggests transfer content to be the fourth factor in transfer of training in Holton's model. Holton stated that the three crucial factors affect transfer of training are motivation to transfer, transfer climate, and transfer design. However, the importance of training content in selecting an effective delivery method is disregarded. Training content must fit to the selected delivery method. The structure of content affects the choice of delivery method. The approach by which content is presented will affect the choice of delivery method and, consequently, it affects individuals' performance. For instance, MCQ structure cannot be delivered in conference or brown bag seminars, instead web-based training or computer game-based training are proper choice for MCQ structure.

Holton stated "trainer judge training content to reflect job requirement'. It is essential to pointing out that training content must fit targeted users as well. Content must be relevant and user specific. Different users with different background require different approaches. Many training programs fail due to the complexity or inadequacy of the training materials. For instance, Training programs with too professional and complicated contents make employees bored and confused. There must a development plan to create proper content for targeted audience.

Literature [8] proposed guidelines for healthcare organizations to develop information security training content. As stated by the authors, the content of an information security awareness training program must be driven from organization information security policy.

V. Training Content

As a preliminary step, organizations should identify the information security mistakes commonly made by employees to be used in developing training content. In other words, training content should cover common mistakes occurring in organization. Moreover, training content should be tailor made to organizations' internal information security policy while consistent with international standards [18]. It is also important to note that information security mistakes made by junior employees may be different from those made by senior employees. Therefore, training content should cover all target employees with different level of awareness knowledge [4].

VI. Information Security Policy Document

There is no specific information security policy document tailored for Hospital Universiti Kebangsaan Malaysia (HUKM), therefore, they operate on Universiti Kebangsaan Malaysia (UKM) security policy document. However, once comparing with international standards, it was recognized that the information security policy document is outdated and it needs to be augmented. Hence, new policies are proposed to augment the current information security policy document (Table 1). The international standards consider for augmentation of HUKM information security policy document include:

- *ISO 27002* which provides guidelines for organizational information security standards and information security management practices including the selection, implementation and management of controls taking into consideration the organization's information security risk environment.
- *SANS* (The System Administration, Networking, and Security) institute, which provides information security policy and standards as a guideline for organizations to develop and implement security policies.
- *HIPAA* (The Health Insurance Portability and Accountability Act of 1996) that is designed to protect confidential healthcare information through improved security standards and federal privacy legislation. It defines requirements for storing patient information before, during and after electronic transmission.

Table 1 demonstrates overview of steps were involved in augmenting the HUKM's information security policy document. First, HUKM's information security policy topics were acknowledged among the three international sources to identify relevant clauses and controls (specified by **√**). Second, strength and quality of policy statement provided by each source is evaluated, and then, compared them with HUKM's policy statements. Next, indispensable information were extracted from the sources (specified by bracket) to be added to relevant part of policy document. Supplementary sections are proposed when required to enhance HUKM's policy document. As shown in the table below, two sections have been added to the augmented document and existing sections were updated in comparison with other relevant international standards.

VII. Training Delivery Method

As mentioned earlier, there are number of awareness training delivery methods. Organizations need to select an effective training delivery method based that can fulfill training needs of both organization and employees. Selection of a training method should be based on the information obtained from interviews with management and the pre-developed training program plan. The selected technique should fulfill the need of both organization and employees.

A guide to selection of awareness training program framework (Fig 4) and selection of awareness training program guidelines (Table 3) aim to provide a guideline for decision makers to select an effective awareness training method to deliver information security content. The framework is developed based on insights from healthcare decision makers coupled with extensive literature study. The designed framework is implemented in the selected healthcare for further evaluation and to recognize which awareness training technique can best fit that particular organization.

TABLE I. POLICY AUGMENTATION

Policy	UKM	ISO 27002 2005	SANS	HIPAA
Software Application Security				
Control of Application Software	[√]	√	√	
Control of Unlicensed Software	[√]	[√]	[√]	
Control of Source Code Storage	[√]	[√]		
Control of Malicious and Defective Code Software	[√]	√		
Control of Malicious Code	[√]	[√]	[√]	√
Control of Version Changes	[√]	[√]		
Server Security				
Physical Security Control	[√]	√		
Control of Database	[√]	√		
Control of Logical Access	[√]	√		
User Identification	[√]	[√]	√	√
User Authentication	[√]	[√]	[√]	√
Information Back-up	[√]	[√]		
Maintenance	[√]	[√]		
Workstation Use			√	[√]
Network Tools Security				
Control of Network Tools Installation Security	[√]	√		
Control of Network Tool Configuration	[√]	√		
Control of Physical Security	[√]	√		
Control of Logical Access	[√]	√		
Wired Network	[√]	√		
Wireless Network (WIFI)	[√]	√		
Control of Network Equipment Maintenance	[√]	√		
Network Security				
User Accessibility	[√]	[√]		
Local Area Network	[√]	√		
Wireless Network (WIFI)	[√]	√		
Connection with other Networks	[√]	√	[√]	
Not Promoted Access	[√]	√		
Firewall	[√]	√		
Internet Policy			[√]	
E-mail Accounts Security				
Controls on the Use of E-mail Accounts	[√]	√	√	
Control of Mailbox Maintenance	[√]	√		
Controls on the Use of E-Mail Software	[√]	√		
Audit Trail	[√]	√	√	
ICT Security Incident Management	[√]	[√]		√

LEGEND:	
√	Available source
[√]	Selected source for augmentation

Literature [8] developed a framework to select an effective training delivery method. The framework is further improved and presented in Fig 4. The authors discussed that an effective awareness training program should be designed by giving significant thoughts to organizations' need as well as training success factors. The content of any training program must be tied to individual organization's need. Therefore, organizations should conduct a training need assessment survey prior to designing their awareness training program. Moreover, an effective information security awareness training program should be based on the training success factors defined by the organization.

A. Organization Training Need Assessment

Literature [22][6] stated that effective awareness training program cannot be developed without giving significant attentions to specific characteristics of an organization. Awareness raining programs should be tied to organizations' training need and requirements. Hence, it is important to conduct an organization training need assessment survey to obtain necessary information to develop training program. The results of the survey will provide justification to convince management for allocating adequate resources to meet the identified awareness and training needs. Literature [22][6] stated important training criteria to be investigated during the survey that include *organization size*, *business requirement*, *funding*, *target audience*, *organization mission and culture*, *organization rule and responsibility*.

B. Training Development Plan

The next step is to clarify the format of training program such as learning outcomes, length of training, target learners, overall format of training, overall description of the training, participant requirements, instructional material and aids needed, logistical issues. Training programs are developed with regards to the capability and requirement of an organization. For example, large organizations are likely to allocate more budgets on training program or they require more employees to participate in the program [11].

C. Training Success Factors

Literature [12][18] Foundation Report, 2010) introduced training success factors that are critical in effectiveness of a program. The success factors include *defining goal*, *motivation*, *fun*, *creativity*, *duration*, *enforce policies*, *organizational culture*, *audience profile*, *easy to understand*, *feedback and evaluation*, and *management support*. In order to obtain satisfactory training outcomes organizations should collect adequate information regarding the training success factors. An effective awareness training program promotes employees' willingness to participate in training activities. Satisfied employees are more encouraged to apply learned skills in their daily job activities.

Furthermore, this study brings together critical training successes factors suggested by previous surveys [12], Rudolph et al., 2002). The training success factors include learning process, time frame, population, training cost, coverage of topics, availability, fun, motivation, challenge, feedback and measurability, effectiveness, updatability, customization, and supervision. The illustrations of these critical factors are provided in Table 2. Although it may not be realistic to expect a training program to satisfy all the success factors, but decision makers should consider important elements related to their organization's need and the learning objectives.

Table 3 provides a side-by-side comparison of the commonly used techniques as suggested in [1]. This table is designed for easy utilization of the above technique selection for awareness training program framework. Some of the boxes in the table are marked by ¤ sign because some of the criteria are subjected to design and implantation approaches. For instance, there are two primary models of web-based instruction namely synchronous (instructor-facilitated) and asynchronous (self-directed, self-paced). Instruction can be delivered by a combination of static method (learning portals, hyperlinked pages, screen cam tutorials, streaming audio/video, and live Web broadcasts) that is categorized as passive learning process. Instruction can also be delivered by interactive method (threaded discussions, chats, and desk-top video conferencing) that is categorized as active learning process. Based on a guide to selection of awareness training program framework (Fig 4) and selection of awareness training program guidelines (Table 3), Organizations should be able to decide on the best awareness training program that best fit the organization.

VIII. Conclusion

Human errors are known as the most serious threats to information security in Electronic Health Record systems. Employees who interact with EHR systems need to be educated about the risks and hazards associated with information security. There is a wide range of information security awareness techniques. However, research is insufficient on effective information security awareness delivery methods. It is essential that effective awareness training delivery method is selected, designed, and implemented to ensure proper protection of the organizational assets. It could be of a great help for organizations to have a step-by-step guideline that provides them with the necessary information on how to select an effective training technique, which fulfills the organization's need and requirement. The authors developed a framework for effective transfer of training. The aim is empowering healthcare decision makers to easily select an effective awareness training program to deliver information security content. Nevertheless, other industries might benefit from the guidelines by applying minor modifications. In this paper training success factors are discussed as critical factors in selecting an effective delivery method. It also explains the process of augmentation of information security content based on internal policy and international standards that can be used as a guideline for healthcare organizations.

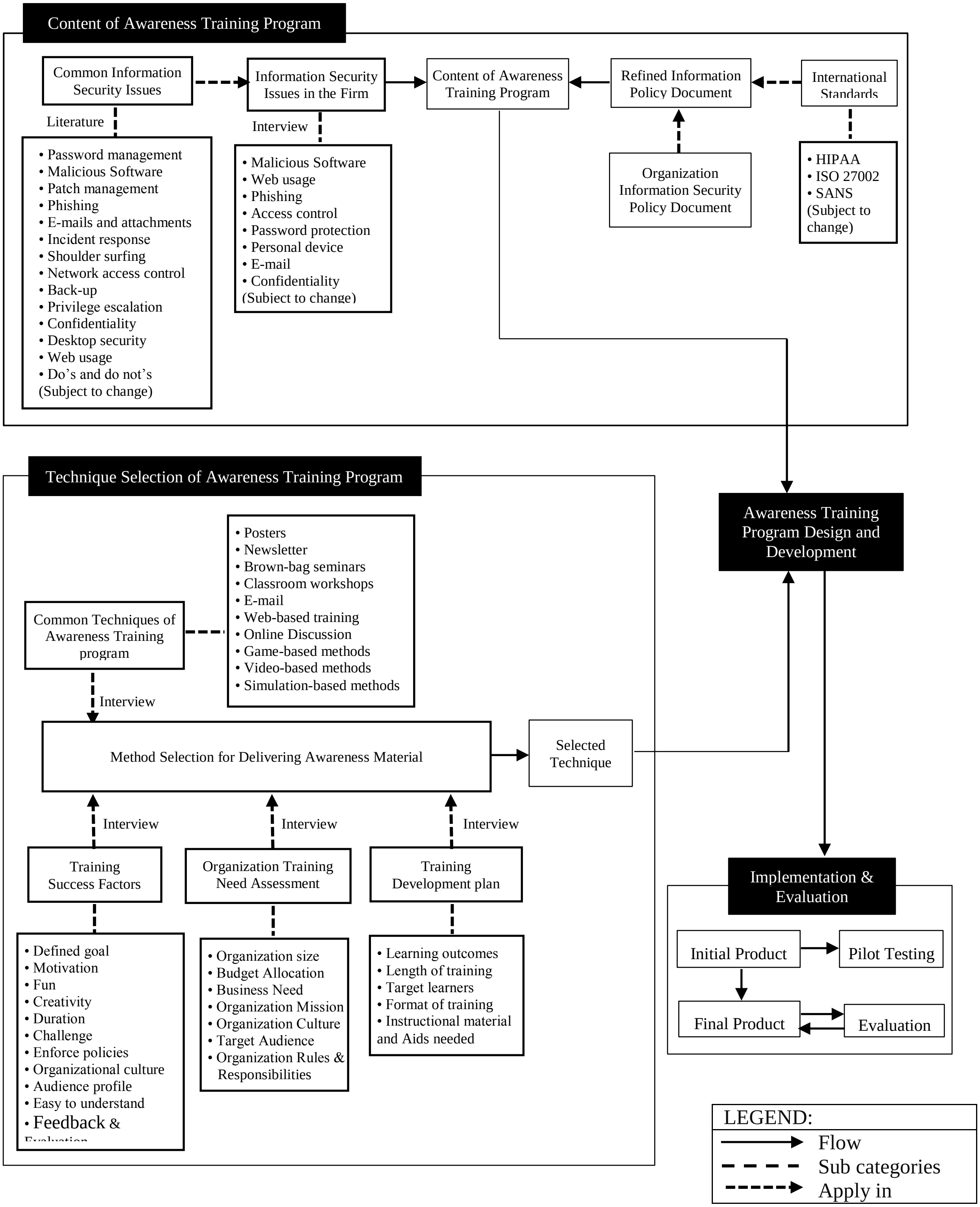


Fig. 4. A Guide to Selection of Awareness Training Program Framework

Source: Adapted from [8]

TABLE II. TRAINING SUCCESS FACTORS

Success Factor	Description
Learning Process	Training methods fit into two categories; active and passive. In Active learning the responsibility of learning lies with the learner. It covers all methods of training where the participants are involved and active in the learning process. In passive learning knowledge is directly transferred from one entity to another. It is normally a one way transfer from entity with more knowledge of the topic towards an entity with less knowledge.
Time Frame	Employee’s daily job responsibilities require flexible schedule. each individual should be given sufficient time to participate in the training program and follow the process without worrying about affecting their performance at the workplace.
Population Coverage	Number of audience covered by an awareness training program.
Training Cost	Design and implementation cost of a training program.
Coverage of Topics	Some of the training techniques are suitable for disseminating of a single message, whereas others can be used for delivering a number of messages.
Content Updatability	Training content should be developed in a way that it allows trainers to update and modify the content if necessary.
Customization	Training contents should be tailor made to each organization based on their specific needs and requirements.
Fun	The amount of entertaining is directly related to individual learning. Participants should be given opportunities to have fun and enjoy what they are doing when engaged with training activities.
Motivation	Motivational factors are needed to encourage individuals to change the way they used to behave and operate.
Challenge	An effective training technique must challenge and engage the participants. Many programs fail to challenge the user which could lead to privation of self-motivation that may encumber successful delivery of the materials.
Supervision	In some training programs trainer directly lead and supervise the program. Whereas, some training programs are run without any supervision.
Feedback & Measurability	Every successful training program provides feedback to both trainees and instructors. Feedback and evaluation are the strengths of each training program and an easy way to distinguish effective trainings form non-effective ones. Trainers must measure and evaluate employees performance before and after training sessions.
Easily Accessible	It refers to the availability of training programs to the organizations and users. For instance, the availability of experienced trainers, training materials, location, and etc.

TABLE III. SELECTION OF AWARENESS TRAINING PROGRAM GUIDELINES

Training Delivery Method		Awareness Training Program Development												
		Success Factors									Content		Implementation & Evaluation	
		Active Learning Process	Flexible Time Frame	Large Population Coverage	Low Cost	Multiple Topic Coverage	Easily Accessible	Fun	Motivation	Challenge	Content Updatability	Customization	Supervision	Feedback & Measurability
Paper-based Methods	Posters	×	√	√	√	×	×	×	×	×	×	×	×	×
	Newsletter	×	√	√	√	×	¤	×	×	×	×	×	×	×
Instructor-led methods	Brown-bag seminars	×	×	×	√	×	×	×	¤	×	√	√	√	×
	Classroom workshops	×	×	×	¤	√	×	×	×	×	√	√	√	×
Online methods	E-mail	×	√	√	√	×	√	×	×	×	√	×	×	√
	Web-based training	¤	¤	¤	¤	√	√	¤	¤	¤	√	√	×	¤
	Online Discussion	√	×	×	√	√	√	√	√	√	√	√	√	¤
Game-based methods		√	√	√	×	√	√	√	√	¤	¤	¤	×	√
Video-based methods		×	√	√	×	√	√	×	×	×	×	×	×	×
Simulation-based methods		√	×	¤	×	×	×	√	√	√	¤	√	×	√

Source: [7], [1]

LEGEND:	
√	Applicable
×	Not applicable
¤	Subject to change

REFERENCES

[1] J. Abawajy, “User preference of cyber security awareness delivery methods,” Behavior & Information Technology, Vol. 33, No. 3, pp 237–248, 2012.

[2] L. A. Annetta, “The “I’s” have it: a framework for serious educational game design,” Review of General Psychology, 14 (2), 105-112, 2010.

[3] T. Asai, and J. L. C. Perez, “Human-related problems in information security faced by Japanese, British and American overseas companies because of cultural differences,” China-USA Business Review, Vol. 11, No. 1, Pp 86-101, 2012.

[4] A. Bakhtyari Shahri, Z. Ismail, and N. Z. A. B Rahim, “Security effectiveness in health information system: through improving the human factors by education and training,” Australian Journal of Basic and Applied Sciences, 6, 226-233, 2012

[5] X. Y. Cheng, Y. M. Wang, and Z. L. Xu, “Risk assessment of human error in information security,” Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 2006.

[6] B. D. Cone, C. E. Irvine, M. E. Thompson, and T. D. Nguyen, “A video game for cyber security training and awareness,” Computers & Security, 26, 63-72, 2007

[7] B. Gardner, V. Thomas, “Building an information security awareness program: defending against social engineering and technical threats,” Elsevier 2014.

[8] A. Ghazvini, Z. Shukur, “An effective awareness training program for information security in Hospital Universiti Kebangsaan Malaysia (HUKM),” Journal of Next Generation Information Technology. Vol. 6, No. 3, pp. 1-12, 2015.

[9] HIMSS Analytics. The 2010 HIMSS Analytics Report: Security of Patient Data. Technical Report. Available: http://www.mmc.com/views/Kroll_HIMSS_Study_April2010.pdf. 25 February 2014.

[10] E. F. Holton, “The flawed four-level evaluation model,” Human Resource Development Quarterly, vol. 7, no. 1, pp 5-21, 1996.

[11] I-TECH (International Training and Education Center for Health) (2012). Department of Global Health. University of Washington. Available: http://globalhealth.washington.edu/organization/international-training-education-center-health-i-tech

[12] S. Manke, and I. Winker, “The habits of highly effective security awareness program: a cross-computer comparison. internet security advisors group,” 2012.

[13] J. Mohan, and R. R. R. Yaacob, “The malaysian telehealth flagship application: a national approach to health data protection and utilization and consumer rights,” 2014.

[14] T. Monk, J. Niekerk, and R. Solms, “concealing the medicine: information security education through game play,” Institute for ICT Advancement, Nelson Mandela Metropolitan University, 2010.

[15] A. Nagarajan, J. M. Allbeck, and A. Sood, “Exploring game design for cybersecurity training,” Proceedings of the 2012 IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems, Bangkok, Thailand, 2012.

[16] R. Parks, C. H. Chu, and H. Xu, “Healthcare information privacy research: issues, gaps and what next?,” 17th American Conference on Information Systems (AMCIS 2011), 4-8 August 2011 Detroit.

[17] Ponemon Institute. 2012. The human factor in data protection. Ponemon Institute. Available Online athttp://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/reports/rpt_trend-micro_ponemon-survey-2012.pdf

[18] Rockefeller Foundation Report. (2010). From Silos to Systems: An overview of ehealth’s transformative power. Available: http://www.rockefellerfoundation.org/uploads/files/e331d255-059f-4fc6-b814-5938f8ee017e-rf.silos_1-13.pdf

[19] G. N. Samy, and R. Ahmad, “Threats to health information security,” The Fifth International Conference on Information Assurance and Security, 2009, Universiti Teknologi Malaysia (UTM), Malaysia.

[20] R. S. Shaw, C. C. Chen, L. A. Harris, H. J. Huang, “The impact of information richness on information security awareness training effectiveness,” Computers & Education, Vol. 52, No. 1, pp 92–100, 2009.

[21] N. Waly, R. Tassabehji, and M. Kamala, “Improving Organizational Information Security Management: The Impact of Training and Awareness,” IEEE 14th International Conference on High Performance Computing and Communications. Bradford University. United Kingdom, 2012.

[22] M. Wilson, J. Hash, “Building an information technology security awareness and training program,” National Institute of Standards and Technology (NIST), 2013. Available: http://csrc.nist.gov/publications/nistpubs/800-50/NIST-SP800-50.pdf

[23] S. Yamnill, G. N. McLean, “Theories supporting transfer of training,” Human Resource Development Quarterly, vol. 12, no. 2, pp 195-208, 2001.

Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance

Nawaf Hazim Barnouti
Al-Mansour University College
Baghdad, Iraq

Wael Esam Matti
Al-Mansour University College
Baghdad, Iraq

Sinan Sameer Mahmood Al-Dabbagh
Al-Mansour University College
Baghdad, Iraq

Mustafa Abdul Sahib Naser
Al-Mansour University College
Baghdad, Iraq

***Abstract*—In this paper, an automatic face recognition system is proposed based on appearance-based features that focus on the entire face image rather than local facial features. The first step in face recognition system is face detection. Viola-Jones face detection method that capable of processing images extremely while achieving high detection rates is used. This method has the most impact in the 2000's and known as the first object detection framework to provide relevant object detection that can run in real time. Feature extraction and dimension reduction method will be applied after face detection. Principal Component Analysis (PCA) method is widely used in pattern recognition. Linear Discriminant Analysis (LDA) method that used to overcome drawback the PCA has been successfully applied to face recognition. It is achieved by projecting the image onto the Eigenface space by PCA after that implementing pure LDA over it. Square Euclidean Distance (SED) is used. The distance between two images is a major concern in pattern recognition. The distance between the vectors of two images leads to image similarity. The proposed method is tested on three databases (MUCT, Face94, and Grimace). Different number of training and testing images are used to evaluate the system performance and it show that increasing the number of training images will increase the recognition rate.**

***Keywords*—Face Detection; Face Recognition; PCA; LDA; Viola-Jones; Feature Extraction; Distance Measurement; MATLAB; MUCT; Face94; Grimace**

I. INTRODUCTION

Face detection is among the important advanced topics in computer vision and pattern recognition communities and it is the first important step for facial analysis methods and among the most important issues in computer vision like face recognition, facial expression, head tracking, face verification. With the arrival of the internet and low price digital cameras, in addition to impressive image editing software such as adobe Photoshop, average users have more access to the tools of digital doctoring than in the past. The objective of face detection would be to determine if there are any faces in the image, then return the location and the bounding box of each face in the image regardless of illuminations, occlusions, facial pose, orientation and expression. Automatic human detection and tracking is an essential and challenging field of research and offers many application areas [1]. Tracking is regarded as a challenging step of tracking system, which localizes and associates the feature across a series of frames. Face recognition has attached much more attention because of its great potential in numerous applications (security, criminal justice system, surveillance, human-computer interactions, image database investigation, smart card application, multimedia environments with adaptive human-computer interface, video indexing and civilian applications) [2] [17].

The expanding use of computer vision in replacing human beings, surveillance, has started the research in the field of face detection. Earlier research is biased to human recognition rather than tracking. Tracking the movement of human beings raised the requirements for tracking. Tracking movements are of high interest in identifying the activities of individual and knowing the attention of individual [1]. The performance of different faces based applications, from standard face recognition and verification to the latest face clustering, retrieval and tagging, depends on efficient and accurate face detection. Face detection is an important part of face recognition system simply because it has the ability to focus computational resources on the important part of an image containing face.

Face recognition involves recognizing individuals with their intrinsic facial characteristic. Compared to other biometrics, face recognition is more natural, non-intrusive and can be used without the cooperation of the individual. Face recognition system can be used in two modes: verification and identification. Face verification system (one-to-one matching) involves confirming or denying the identity claimed by a specific individual. Face identification system (one-to-many matching) attempts to find the identity of a given individual against all image templates in face individual database [3].

Face recognition methods can be divided into appearance-based or model-based methods. Appearance-based (Holistic) face recognition legally attempt to identify faces using global representations based on the entire image rather than local facial features. An image is considered as a high dimensional vector. Statistical methods are frequently used to gain a feature space from the image distribution. The sample image is compared to the training set. Appearance-based methods can be classified as either linear or non-linear. Linear appearance-based methods perform a linear dimension reduction [4]. The face vectors are projected to the basis vectors; the projection coefficient are used as the feature representation for each face image through the projection of the face image vector onto the basis vectors. Linear methods are Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA). Non-linear appearance-based methods are usually more complicated than linear methods. Direct non-linear manifold schemes are explored to learn this non-linear manifold. Linear subspace analysis is an approximation of a non-linear manifold. Kernel PCA (KPCA) is widely used. Model-based face recognition scheme aims to construct a model of the human face that can capture facial variations. It can be either 2-Dimensional or 3-Dimensional. These models are frequently morphable. Morphable model make it possible for classifying faces even when pose changes are present. Model-Based methods are Elastic Bunch Graph Matching (EBGM) or 3D Morphable Models [5]. Hybrid method is a combination of appearance-based and model-based methods. Regardless of the method, the most important concern in face recognition is dimensionality. Suitable methods are needed to reduce the dimension of the studded space. Working on higher dimensions causes overfitting, where the system starts to memorize. Computational complexity is also an important problem when working on large database.

II. PROPOSED METHODOLOGY

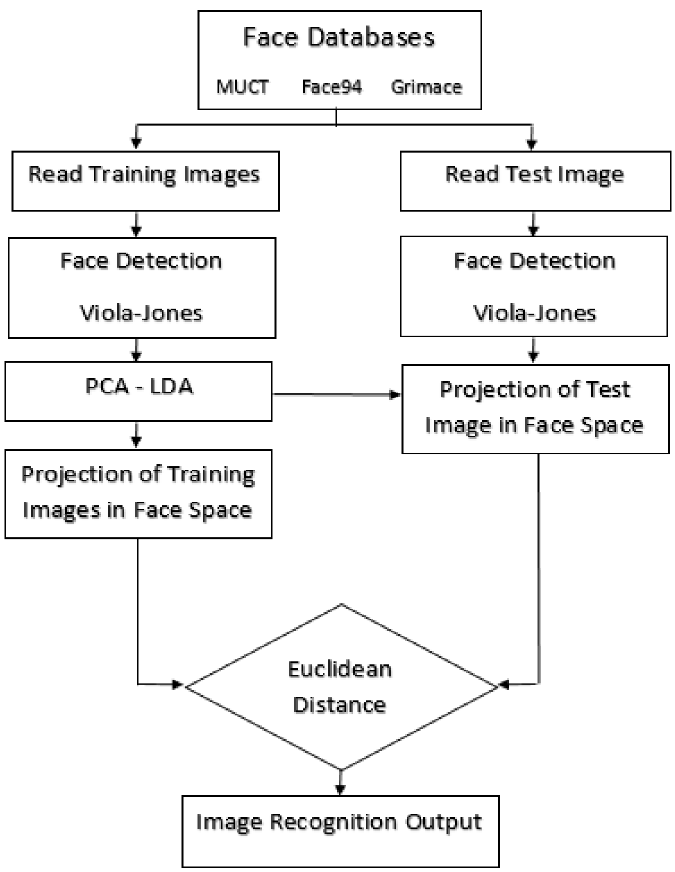


Fig. 1. Face Recognition Proposed Methodology Process

Face recognition is a complex image processing problem in real world applications. In this work, details are provided for the method and training process of the proposed face detection and recognition system. Technologies characteristics and features make face recognition important and better performer depending on the application. Face recognition basically divided into three steps begins with face detection continue with feature extraction and end up with distance measurement process. Three benchmark databases (MUCT, Face94 and Grimace) are used to test the system performance. MUCT database contains 3.755 faces with 76 manual landmarks, Face94 contains 153 images each with a resolution of 180*200 pixel, and Grimace database contains face images with 20 individuals each having 20 images. Viola-Jones face detection method is used to detect and crop face region in each face database. The linear appearance-based method PCA-LDA is used for feature extraction and dimension reduction. Finally, Square Euclidean Distance measurement is used. The distance between two images is a major concern in pattern recognition. Image similarity is the distance between the vectors of two images. Figure 1 shows the proposed methodology process.

III. FACE DATABASE

Numerous face databases are available for face recognition researchers. These databases differ in size, scope and purpose. It is recommendable to use a standard test face recognition database for researchers to be able to directly compare the final results. The photographs in many of these databases are acquired by small research teams specifically to study face recognition. MUCT, Face94, and Grimace databases are used in this work. Table I shows the important features of different face recognition databases.

TABLE I. FACE DATABASES FEATURES

Database	**Format**	**Individuals**	**Image Size**
MUCT	RGB	276	480 * 640
Face94	RGB	153	180 * 200
Grimace	RGB	20	180 * 200
ORL	Gray	40	92 * 112
FERET	Gray/RGB	1199	256 * 384
UMIST	Gray	20	92 * 112
Indian	RGB	40	640 * 480

A. MUCT Database

The Milborrow/University of Cape Town (MUCT) database contains 3.755 faces with 76 manual landmarks. The database is created to diversity lighting, age, and ethnicity. In this database, all images captured in December 2008 are from the individuals around the University of Cape Town campus. The individuals in this database are university students, parents, high school teachers, and employees, each individual is photographed using five webcams, which makes the database useful for applications that require multiple occurring views of the individual [6]. Figure 2 shows a sample images of MUCT database.

Fig. 2. MUCT Database Sample Images

B. *Face94 Database*

The Face94 contains 153 images each with a resolution of 180*200 pixel and the directories comprise images of male and female individuals in separate directories (20 females, 113 males, and 20 male staffs). The images are mainly from first year underground students. The majority of the individuals are between 18 and 20 years old. The lighting is artificial and some of the images are captured with glasses, and a mixture of tungsten and fluorescent overhead [7]. Figure 3 shows a sample images of Face94 database.

Fig. 3. Face94 Database Sample Images

C. *Grimace Database*

The Grimace database contains face images with 20 individuals each having 20 images with a resolution of 180*200 pixel with a small head scale variation. This database contains images of both female and male individuals. The lighting of the images minimally varies. The images of the individuals from various facial origins have major expression variations with breads and glasses [8]. Figure 4 shows a sample images of Grimace database.

Fig. 4. Grimace Database Sample Images

IV. FACE DETECTION

Face detection is generally considered as a certain case of object-class detection and it's a popular topic in biometrics research. Face detection is the first step of face recognition system. Objects can be detected using one of the face detection methods. Then feature extraction and distance measurement methods can be applied to the system. In object-class detection, the task is to find the location of all objects in an image that belong to a given object [3]. Face detection is not simple because it carries lots of variations of appearance in images, such as facial expression, pose variation, image orientation, occlusion and illuminating condition. In this work, Viola-Jones face detection method is used.

The Viola-jones object detection method suggested by Paul Viola and Michael Jones in 2001. This method has the most impact in the 2000's and known as the first object detection framework to provide relevant object detection that can run in real time. Viola-Jones requires full view frontal upright faces [9]. At a high level, the method read an input image with a window looking for human face features. When enough features are found, then this window type of the image is reported to be a face [10]. In order to bring different size faces, the window must be scaled and the process is repeated. For each window scale involves through the method separately of the other scales. This method happens to be rather time consuming resulting from the calculation of the different images size. To decrease the number of features each window have to check and each window is passed through levels. Early levels include less features to check and are much easier to pass but later levels end up having more features and are more demanding. At each level, the evaluation of features for that levels are collected and whether if the collected value does not pass the threshold, the level is failed and this window will be not recognized as a face. The Viola-Jones face detection method is divided into three main parts (Integral image, classifier learning with AdaBoost and attentional cascade structure) that make it possible to build a successful face detection that can be used on real time application.

A. *Creating an Integral Image*

An image representation called the integral image. Integral image also known as a summed area table. Integral image is computed as a pre-processing step. The first step of Viola-Jones method is to convert the input face image into an integral image. This can be done by making each pixel equal to the entire summation of all pixels above and to the left of the concerned pixel [9]. The integral image can be calculated as shown in the equation below:

$$I\,(x, y) = \sum_{x' \leq x, Y' \leq y} O\,(x', y') \tag{1}$$

Where I is the integral image and O is the original image.

To complete the summation of any rectangular area by using the integral image is extremely efficient. The summation of pixels in rectangle area $Z = [A1, A2, A3, A4]$ can be calculated as shown in the equation below:

$$I\ (x,y) = \sum_{(x,y)\,\in\, Z} O(A4) + O(A1) + O(A2) + O(A3) \qquad (2)$$

Features will be calculated in constant time considering that the summation of the pixels can be computed in the constituent rectangles in constant time. Viola-Jones have noticed that a detector with a basic resolution of 24*24 pixels offers positive results [10].

B. AdaBoost Training

AdaBoost is a machine learning boosting method capable of finding a highly accurate hypothesis by combining many week hypothesis each with average accuracy. The AdaBoost method is generally viewed as the first step straight into more practical boosting methods [9].

C. Cascade Structure

Cascade of gradually more complex classifiers achieves even better detection rates. The concept of the Viola-Jones face detection method is to scan the detector frequently by the same image each time with a new size. Regardless of whether an image should contain one or more faces, there is no doubt that an excessive large amount of the evaluated sub windows might still be non-faces [10]. The Cascade classifier consist of levels each containing a strong classifier. The responsibility of each level is to evaluate if a given sub-window is actually non-face or maybe a face. The implementation contains 22 levels with early levels containing much less features and later levels containing more in depth detailed features. Typically, early levels are passed more frequently with later levels being more demanding.

V. Feature Extraction

Feature extraction involves reducing the amount of resources required to describe a large amount of data. Feature extraction from given data is a critical problem for the successful application of machine learning. In this work PCA and LDA are used as feature extraction and dimension reduction method from the original face images. PCA and LDA produce feature vectors in a reduced dimension.

A. Principal Component Analysis (PCA)

Principal component analysis (PCA) is one of the most important methods used in pattern recognition and compression. PCA is feature extraction and dimension reduction method [11]. PCA is a common statistical method using a holistic approach to find patterns in high dimensional data. The purpose of PCA is derived from the information theory approach, which break down facial images into small sets of characteristic feature images called Eigenfaces which used to represent both existing and new faces [12]. In PCA method, the 2-Dimensional face image matrices must be transformed into a 1-Dimensional vector. The 1-Dimensional vector can be either row or column vector. As a result, the image representation leads to a high dimensional space [13].

PCA Method Steps are as Follows:

1) Training set of total M images are used to compute the Average Mean as shown in the equation below:

$$average = \frac{1}{M}\sum_{n=1}^{M} TrainingImages\ (n) \qquad (3)$$

2) Original image will be subtracted from the Average Mean as shown in the equation below:

$$Sub = TrainingImages - average \qquad (4)$$

3) Calculate the Covariance Matrix as shown in the equation bellow:

$$Covariance = \sum_{n=1}^{M} Sub(n)\ Sub^{T}(n) \qquad (5)$$

4) Calculate the Eigenvalues and Eigenvectors of the Covariance Matrix.

5) Sort and choose the best Eigenvalues. The highest Eigenvalues that belong to a group of Eigenvectors is chosen, these M Eigenvectors describe the Eigenfaces. Given that new faces are encountered, the Eigenfaces can be updated or recalculated accordingly.

6) Project the training samples onto Eigenfaces.

B. Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is also known as fischerface method used to overcome drawback the PCA of its application kept in small image database. It is achieved by projecting the image onto the Eigenface space by PCA after that implementing pure LDA over it to classify the Eigenface space projected data [12]. LDA searches for those vectors in the underlying space that best discriminate among classes. LDA group images of the same class and separates images of different classes. As Mathematically two measures are defined (within-class scatter matrix and between class scatter matrix) [14]. For all samples of all classes the between-class scatter matrix SB and the within-class scatter matrix SW are defined as shown in the equations below:

$$SB = \sum_{n=1}^{N}\sum_{m=1}^{M_n} (S_m^n - u_n)\ (S_m^n - u_n)^T \qquad (6)$$

$$SW = \sum_{n=1}^{N} (u_n - u)\ (u_n - u)^T \qquad (7)$$

Where S_m^n is the m^{th} sample of class n, u_n is the mean of class n, N is the number of classes, M_n is the number of samples in class n and u is the mean of all classes. Then subspace for LDA is spanned by a set of vectors $W = [W1, W2, W3, \dots, W_M]$.

$$W = \arg\max = mod\ \ (W^T * SB\ W / W^T * SW\ W) \qquad (8)$$

The goal is to maximize the between class measure while minimizing the within class measure. Figure 5 shows that when maximize the ratio of between class variance to within class variance will find a good class separation. To do this we maximize ratio $\det|SW|/\det|SB|$ to prove that if SW is non-singular matrix. The with class scatter matrix represent how face images are distributed closely with-in classes and between class scatter matrix describe how classes are separated from each other. When face images are projected into the discriminant vector W. Face Images should be distributed closely with-in classes and should be separated between classes as much as possible. In other words, these discriminant vectors minimize the denominator and maximize the numerator [15].

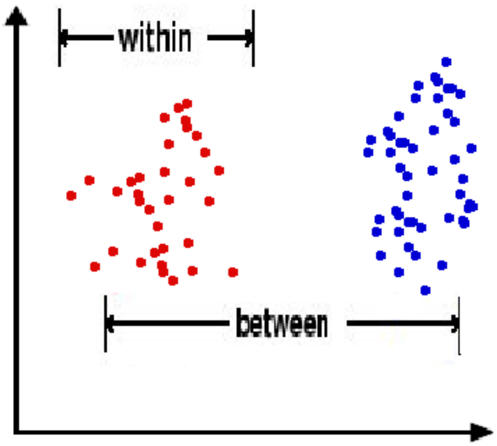


Fig. 5. Class Separation in Linear Discriminant Analysis

VI. Distance Measurement

Once the features are extracted and selected using PCA-LDA, the next step is to measure the distance between images. Most face recognition methods from the last decade help in deciding according to the distance measurement. The distance between two images is a major concern in image recognition and computer vision. The final step of face recognition is measuring the distance between two images. Image similarity is the distance between the vectors of two images. The distance among feature space representations are used as the basis for recognition decisions [16]. One way or another, distance measurement has a big impact in face recognition area. Distance measurement methods are used in many areas like finance, data mining, voice recognition and signal decoding.

Euclidean distance is used for distance measurement between images. Euclidean Distance is defined as the straight line distance between two points, which examines the root of square differences between the coordinates of a pair of objects [16]. Euclidean Distance can be calculated using the equation below:

$$EuclideanDistance(X,Y) = \sqrt{\sum_{n=1}^{No.of\ Images} (X_n - Y_n)^2} \quad (9)$$

Without the square roots, we can obtain the Square Euclidean Distance (SED) measurement. The standard Euclidean Distance can be squared in order to place progressively greater weight on objects that are farther apart. In this case, the equation becomes as shown below:

$$SquareEuclidean(X,Y) = \sum_{n=1}^{No.of\ Images} (X_n - Y_n)^2 \quad (10)$$

VII. Result and Discussion

In this analysis, three databases (MUCT, Face94, and Grimace) are used to evaluate the system performance. In MUCT database, 8 individuals with 1 to 3 training and testing images for each individual is used. While, in Face94 and Grimace databases, 8 individuals with 1 to 4 training and testing images for each individual are used. The simulation of the proposed methodology was performed using MATLAB software package The analysis shows that increasing the number of training images will increase the recognition rate. Viola-Jones method is used for face detection on each database. This method achieved high detection rate and all images are detected and cropped in the three databases. Figure 6 shows a sample image detection and cropping using Viola-Jones method.

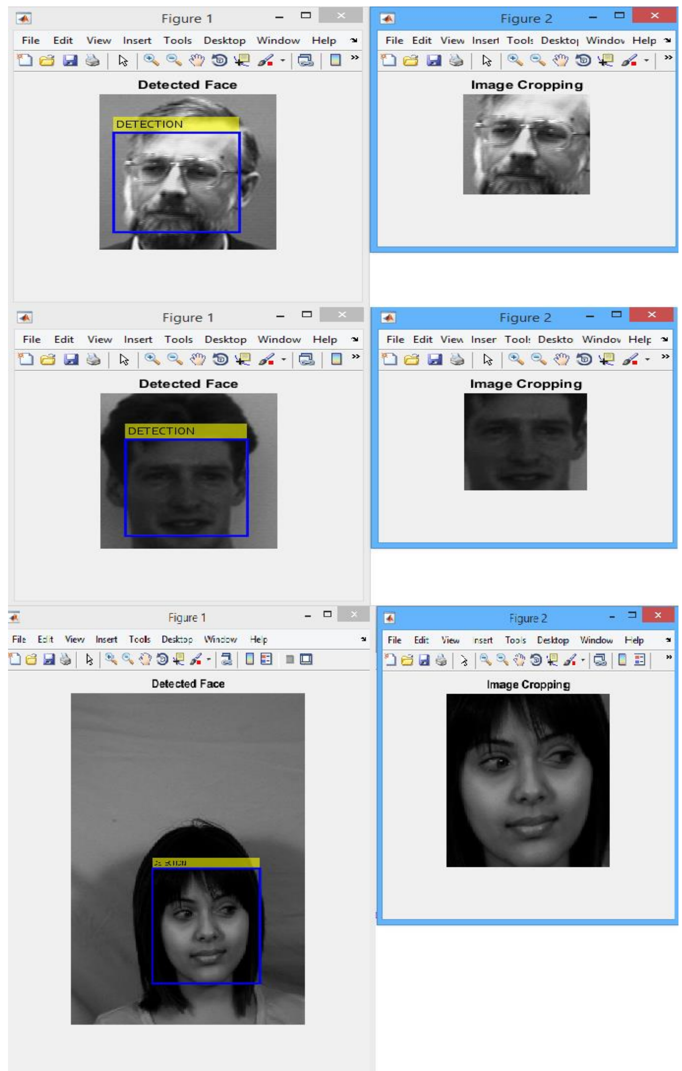


Fig. 6. Image Detection and Cropping Using Viola-Jones Method

Figure 7 shows MUCT images database after detection and cropping. Figure 8 shows Face94 images database after detection and cropping. Figure 9 shows Grimace images database after detection and cropping. PCA-LDA are applied on the detected cropped images for feature extraction and dimension reduction. Different number of training and testing images are used in each database. Square Euclidean distance is used to measure the distance between two images. Face94 and Grimace databases with 1 to 4 images shows high recognition rates, while MUCT database with 1 to 3 images shows low recognition rates. To avoid this problem, the number of images in the database must be increased to become 1 to 8 images for each individual. Table II shows the recognition rate of MUCT with 1 to 8 images. Table III shows the recognition rates of Face94 and Grimace databases.

Fig. 7. Detection and Cropping on MUCT Database

Fig. 8. Detection and Cropping on Face94 Database

Fig. 9. Detection and Cropping on Grimace Database

TABLE II. THE MUCT DATABASE RECOGNITION RATES

Test Image	Train Image	Recognition Rates MUCT
7	1	60.71 %
6	2	66.67 %
5	3	82.5 %
4	4	84.38 %
3	5	85 %
2	6	87.5 %
1	7	87.5 %

TABLE III. FACE94 AND GRIMACE DATABASE RECOGNITION RATES

Test Image	Train Image	Recognition Rates Face94	Recognition Rates Grimace
3	1	100 %	91.67 %
2	2	100 %	100 %
1	3	100 %	100 %

VIII. CONCLUSION

The purpose of this work was to implement an automatic face recognition system based on appearance-based methods. Face detection using Viola-Jones method is used to detect and crop faces in each database. Viola-Jones method show high detection rates. MUCT, Face94, and Grimace databases are used, each with 8 individuals and 1 to 3 images are choosing for each individual in MUCT database, 1 to 4 images are choosing for each individual in Face94 and Grimace databases. PCA-LDA is used for feature extraction and dimension reduction. PCA-LDA implementation was successful. Square Euclidean Distance is used to measure the distance between two images, which leads to find image similarity. Face94 and Grimace databases using different number of testing and training images shows high recognition rates, while MUCT database shows low recognition rates. In MUCT database, increasing the number of images to become 1 to 8 images for each individual shows increasing the recognition rates. The recognition time was acceptable and takes few seconds. The results show increasing in recognition rates when increase the number of training images.

REFERENCES

[1] Hatem, H., Beiji, Z., Majeed, R., Lutf, M. and Waleed, J., 2015. Face Detection and Pose Estimation Based on Evaluating Facial Feature Selection. *International Journal of Hybrid Information Technology*, *8*(2), pp.109-120.

[2] Bakshi, U. and Singhal, R., 2014. A survey on face detection methods and feature extraction techniques of face recognition. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, *3*(3), pp.233-237.

[3] Rath, S.K. and Rautaray, S.S., 2014. A Survey on Face Detection and Recognition Techniques in Different Application Domain. *International Journal of Modern Education and Computer Science*, *6*(8), pp.34-44.

[4] Fernandes, S. and Bala, J., 2013. Performance Analysis of PCA-based and LDA-based Algorithms for Face Recognition. *International Journal of Signal Processing Systems*, *1*(1), pp.1-6.

[5] Gayathri, S., Mary Jeya priya, R., and Dr.Valarmathy, S., 2014. Face Recognition by Using Distance Classifier Based On PCA and LDA,

International Journal of Innovative Research in Science, Engineering and Technology, 3(3), pp.1121-1126.

[6] Milborrow, S., Morkel, J. and Nicolls, F., 2010. The MUCT landmarked face database. *Pattern Recognition Association of South Africa*, *201*(0).

[7] Abdullah, M., Wazzan, M. and Bo-saeed, S., 2012. Optimizing face recognition using PCA. *International Journal of Artificial Intelligence & Applications,* 3(2), pp.23-31.

[8] Mandal, T. and Wu, Q.J., 2008, December. Face recognition using curvelet based PCA. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on* (pp. 1-4). IEEE.

[9] Cha, Z. and Zhengyou, Z., 2010. A survey of recent advances in face detection. *Microsoft Research, Microsoft Corporation*.

[10] Hefenbrock, D., Oberg, J., Thanh, N.T.N., Kastner, R. and Baden, S.B., 2010, May. Accelerating viola-jones face detection to fpga-level using gpus. In *2010 18th IEEE Annual International Symposium on Field-Programmable Custom Computing Machines* (pp. 11-18). IEEE.

[11] Kaushik, Sangeeta, Dubey, R. B., and Abhimanyu Madan. 2014. Study of Face Recognition Techniques, *International Journal of Advanced Computer Research*, 4(17), pp.939-949.

[12] Singh, A., Singh, S.K. and Tiwari, S., 2012. Comparison of face Recognition Algorithms on Dummy Faces. *The International Journal of Multimedia & Its Applications*, *4*(4), pp.121-135.

[13] Barnouti, Nawaf Hazim. 2016. Face Recognition Using Eigen-Face Implemented On DSP Processor, *International Journal of Engineering Research and General Science*, 4(2), pp.107-113.

[14] Bhattacharyya, S.K. and Rahul, K., 2013. Face Recognition by Linear Discriminant Analysis. *International Journal of Communication Network Security, ISSN*, 2(2), pp.2231-1882.

[15] Kamerikar, Umesh Ashok, and Dr. M.S. Chavan. 2014. Experimental Assessment of LDA and KLDA for Face Recognition, *International Journal of Advance Research in Computer Science and Management Studies*, 2(2), pp.137-146.

[16] Gawande, M.P. and Agrawal, D.G., Face recognition using PCA and different distance classifiers. *IOSR Journal of Electronics and Communication Engineering*, 9(1), pp.1-5.

[17] Barnouti, Nawaf Hazim. 2016. Improve Face Recognition Rate Using Different Image Pre-Processing Techniques, *American Journal of Engineering Research (AJER)*, 5(4), pp.46-53.

Data Mining Framework for Generating Sales Decision Making Information Using Association Rules

Md. Humayun Kabir
Dept. of Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka-1342, Bangladesh

***Abstract*—The rapid technological development in the field of information and communication technology (ICT) has enabled the databases of super shops to be organized under a countrywide sales decision making network to develop intelligent business systems by generating enriched business policies. This paper presents a data mining framework for generating sales decision making information from sales data using association rules generated from valid user input item set with respect to the sales data under analysis. The proposed framework includes super shop's raw database storing sales data collected through sales application systems at different Point of Sale (POS) terminals. Apriori algorithm is famous for association rule discovery from the transactional database. The proposed technique using customized association rule generation and analysis checks the input items with sales data for validation of the input items. The support and confidence of each rule are computed. Sales decision making information about input items is generated by analyzing each of the generated association rules, which can be used to improve sales decision making policy to attract customers in order to increase sales. It is hoped that this approach for generating sales decision making information by analyzing sales data using association rules is more specific decision and application oriented as the business decision makers are not usually interested to all of the items of the sales database for making a specific sales decision.**

Keywords—databases; data mining framework; Apriori algorithm; association rule; sales decision making information

I. Introduction

With the huge growing size of data and information in our modern technology world, organizational decisions are largely becoming dependent on computerized systems. Various sectors, e.g., education, agriculture, garments, stock exchange, finance and banking, super shops and many other sectors have potential applications of data mining systems in generating business intelligence (BI) for effective operations and efficient decision making. In recent days, massive data are collected from the customer's purchase records using application software system in the super shops, which are stored in databases using relational database management systems (RDBMSs), e.g., Oracle, MySQL.

Interesting patterns can be discovered from the customer's purchase records by analyzing the super shop's sales database using the data mining techniques, which may reflect customer's buying patterns [1] [2]. The business authority may use this buying pattern to predict customer's buying habit and frequency of buying a particular item in the super shop. This information can be used productively by the business decision makers for efficient stock management and customer attraction to promote sales [1].

In data mining, intelligent algorithms are applied to organizational operational data stored in RDBMS or consolidated historic data stored in data warehouses (DWs) to extract hidden knowledge, interesting patterns, missing values, and new rules to speed up organizational decision making process [1]. Characterization, classification, association rule discovery and analysis, and clustering are the major data mining tasks, which are now widely used in various organizational applications. Each type of tasks has one or more algorithms which have been developed to perform data mining on transactional databases. The algorithm for mining association rules has been introduced in [3] in 1993. Apriori [4] is the famous algorithm for discovering association rules from the transactional database. Various research works [5] [6] [7] have been performed on this algorithm in various application domains. In association rule mining [4] [5] [6] [7] [8], association rules are generated by discovering relationship between the items of the transactions stored in the operational databases. Millions of sales transactions are performed in the supermarkets which are stored in operational databases for future processing. The management information system (MIS) and decision support system (DSS) personnel of a business organization must have the information about the customers buying habits, the best selling items, the items which make the highest profit, and the items not sold at all to develop attractive business policies to improve customer service and to increase sales. Infrequent items can be placed on offer based on the sales decision making knowledge extracted by analyzing sales data for each of the generated association rules to avoid business loss [1]. Association rule analysis [6] [7] [8] can be applied to the sales records of the supermarkets or sales centers to discover the buying patterns of customers. Based on this buying pattern, certain items can be placed on offer on discount to attract customers of a particular buying habit to buy a particular item which he was not used to buying. It is possible to generate association rules using only input item set rather than all items contained in the transactional database, so

This research work is based on the work presented in [1] which was funded by the University Grants Commission (UGC) of Bangladesh in 2011-2012 in the Faculty of Mathematical and Physical Sciences, Jahangirnagar University, Savar, Dhaka, Bangladesh.

that the generated rules can be useful in business decision making to promote sales on the specific item(s).

In business intelligence [9] [10], data of organizations is transformed into information to extract knowledge for effective decision making. Some recent research works on association rules are presented in [11] [12] [13]. Efficient algorithms have been introduced in [11] to generate candidate itemsets from transactional databases. These algorithms reduce candidates and also improve runtime for long transactions. An efficient approach for mining useful association rules from large transactional database using a clustering method to classify database and soft sets have been presented in [12]. A new association rule mining algorithm CMARM based on Confabulation-MapReduce is presented in [13] for analyzing medical data. This algorithm is also useful for infrequent items. In this paper, an approach for discovering business intelligence from the transactional database is presented for data mining using association rule generation and analysis technique [1]. Association rules generated from user input item set can be analyzed using sales data to generate information about sales items which can be used to improve sales decision making process. This approach first verifies whether the user input items exist within the transactional database, and then generates association rules from the valid input items only. Next, each of these generated association rules is analyzed based on the operational data records stored in the sales relation of the transactional database. The support and confidence [4] [5] [6] [7] of each rule is computed. A table of the generated association rules with their support and confidence is provided in each analysis case, which can be used to further analyze sales data to generate sales decision making information to obtain improved sales policy.

The remainder of this paper is organized as follows. Section II describes how data mining technique can be employed in the super shop's sales system using association rules. Section III presents a framework for data mining from relational databases as well as data warehouses. Section IV presents experimentation and results to demonstrate how the customized association rule generation and analysis technique can be used to generate sales decision making information. Finally, Section V concludes with a guideline to the future work.

II. Application of Data Mining Technique in Sales System Using Association Rules

Today's business organizations are facing challenges on how they will attract customers. They have to change their business policies by adopting information and communication technology (ICT) incorporating intelligent software system within their automated business management system to speed up the business activities, quality of services and efficient policy making. Millions and millions of sales data records are collected through daily product sales at supermarkets and business organizations [1] [2], which are stored in relational databases using RDBMSs. This massive customer data may reflect customer's buying habits and frequency of buying a particular item.

Data mining can be performed on data records stored in a single relation or multiple relations. A super shop database may contain some relations containing data about the super shop, item details, supplier details, purchase, sales, members, sales man, customer, gifts, bill etc. [2]. In the scope of this paper, the purpose is to extract knowledge about customer's buying pattern from a sales database using association rules, and hence, a single relation called *itemssold* is used for a sales system. In this section, we consider this relation with the following attributes which store super shop's sales data of some customers.

itemssold(*salesdate, item1, item2, item3, …, itemn*)

An association rule [3] [4] [5] [6] [7] [8] is a form of expressing the relationship between two data item sets. Association rules can be generated from items representing daily sales records stored in a super shop's operational database, which can be used to identify frequently purchased item sets of the customers [1]. The rule discovery usually considers data items stored in a single relation. The information about data items contained in the operational database can be expressed by the frequency of occurrence of the data items of a rule by measuring support and confidence of each rule. If X and Y are two sets of items, and I is the set of items in a super shop's sales transaction, then an association rule can be expressed in the form $X \Rightarrow Y$ [3] [4] [5] [6] [7] where $Y = I \setminus X$, $X \cap Y = \{\}$, *i.e.*, empty set, X is the rule antecedent, Y is the rule consequent, X and Y are non-empty subsets of I, and $\Rightarrow$ is an implication operator. TABLE I shows some customer data records stored in the *itemssold* relation of a transactional database.

TABLE I. CUSTOMER'S PURCHASE RECORDS STORED IN ITEMSSOLD RELATION [1]

Date	**Items**				
Sales date	***item1***	***item2***	***Item3***	***item4***	***item5***
04/09/2012	Beef	Ruhi	Fine Rice	Milk	Soft Drink
20/10/2012	Onion	Hilsha	Carrots	Bread	Milk
10/11/2012	Beef	Hilsha	Fine Rice	Milk	Soft Drink
25/11/2012	Onion	Rice	Banana	Bread	Dal
05/12/2012	Beef	Shrimp	Fine Rice	Milk	Cooking Oil

An association rule *r1* can be defined using the items {Beef, Fine Rice, Milk, Soft Drink} shown in TABLE I in the following form [1]:

{Beef, Fine Rice, Milk} $\Rightarrow$ {Soft Drink}

The support *s* and confidence *c* of an association rule $X \Rightarrow Y$ can be defined in terms of the number of occurrences of the item set in the antecedent X and consequent Y of a rule within a database relation containing N transactions. The support and confidence of an association rule are defined in [3] [4] [5] [6] [7] [8]. It can be computed as follows [4] [5] [6] [7] [8].

Support $s(X \Rightarrow Y)$

$= ((\text{Number of records containing items } X \cup Y) / N) \times 100$

The support *s* is 40% for rule *r1*. In this case, 2 records contain the items of *r1* and $N = 5$ for the transaction records shown in TABLE I.

Confidence $c(X \Rightarrow Y)$

$= ((\text{support of } X \cup Y) / \text{support of } X) \times 100$

where

support of $X \cup Y$

$= ((\text{Number of records containing items } X \cup Y) / N) \times 100$

support of X

$= (\text{Number of records containing items } X / N) \times 100$

The confidence c of rule *r1* is 66.66% for the transaction records shown in TABLE I.

III. System Architecture for Sales Data Mining Framework

The system architecture of a sales data mining framework for the application of data mining technique to the sales database of a typical super shop is depicted in Fig. 1. The framework includes two data storage components for mining knowledge: one is relational database component which stores daily sales data of the super shop, and the other is a data warehouse (DW) component which is constructed from the operational databases.

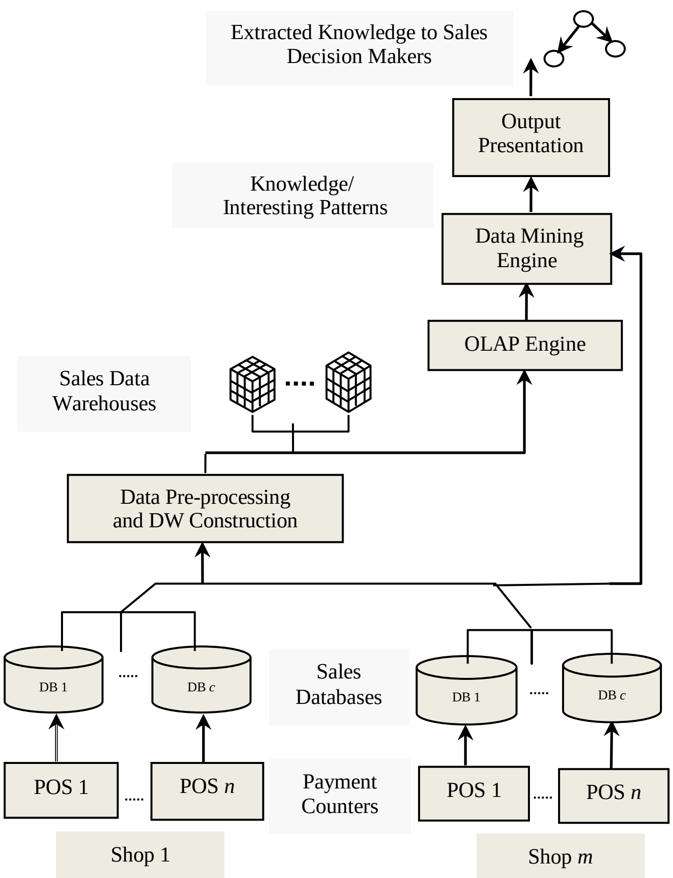


Fig. 1. System Architecture of a Sales Data Mining Framework

The point of sale (POS) terminals of the super shops are used to collect daily sales data through sales application software system, and the sales data is stored in the relational databases DB_1, ..., DB_c maintained by the database servers. Data mining algorithms for classification, association rule analysis, and clustering can be employed in the data mining engine as required by the sales system to extract sales decision making knowledge. The framework shows database(s) which may consist of multiple relations, and in that case, multi-relational data mining techniques may need to be employed. For mining knowledge from DWs, required data may be selected from the DW(s) using OLAP engine. Each of the input data sources is treated differently by the mining modules within the data mining engine, as the DWs may be constructed using different schema structures [8] and may have different data content than that of the relational databases. The mined output knowledge is represented graphically to the sales decision makers through the output presentation component to be used in efficient decision making. It is a 7 layer architecture for knowledge discovery from databases. Among the layers, only 4 layers are used in data mining from sales databases whereas all of the 7 layers are required for data mining from sales data warehouses.

IV. Experimentation and Results

The performance of association rule generation technique and the application of these association rules in mining knowledge from sales database for using in sales decision making are experimented in two different approaches which are explained below.

A. Approach 1

In this approach, Weka 3.4.3 Associator[1] and Weka 3.7.12 Associator[2] are applied to the purchase records of some customers shown in TABLE II after removing the *salesdate* attribute from the *itemssold* relation to discover association rules using Apriori algorithm.

TABLE II. Purchase Records [1]

Items				
item1	***item2***	***item3***	***item4***	***item5***
Beef	Ruhi	Fine rice	Milk	Soft drink
Onion	Hilsha	Carrots	Bread	Milk
Beef	Hilsha	Fine rice	Milk	Soft drink
Onion	Rice	Banana	Bread	Dal
Beef	Shrimp	Fine rice	Milk	Cooking oil

The best association rules obtained by Weka 3.4.3 Associator using the customer's purchase records shown in TABLE II are listed below in TABLE III. In TABLE II, each record consists of a subset of the item set I = {Beef, Ruhi, Fine rice, Milk, Soft drink, Onion, Hilsha, Carrots, Bread, Rice, Banana, Dal, Shrimp, Cooking oil} of 14 items in the set contained in the purchase records of customers. Only 3 items forming the subset I_s = {Fine rice, Beef, Milk} out of these 14 items is used to generate 10 best rules as indicated in TABLE III using Weka 3.4.3 Associator. These rules can be classified into two categories: rules containing i) 2 items and ii) 3 items.

[1] url: http://www.cs.waikato.ac.nz/ml/weka/
[2] url: http://www.cs.waikato.ac.nz/ml/weka/

TABLE III. THE BEST RULES GENERATED USING WEKA 3.4.3 ASSOCIATOR [1]

Rule No.	Rule
1.	item3=Fine rice 3 ==> item1=Beef 3 conf: (1)
2.	item1=Beef 3 ==> item3=Fine rice 3 conf: (1)
3.	item4=Milk 3 ==> item1=Beef 3 conf: (1)
4.	item1=Beef 3 ==> item4=Milk 3 conf: (1)
5.	item4=Milk 3 ==> item3=Fine rice 3 conf: (1)
6.	item3=Fine rice 3 ==> item4=Milk 3 conf: (1)
7.	item3=Fine rice 3 item4=Milk 3 ==> item1=Beef 3 conf: (1)
8.	item1=Beef 3 item4=Milk 3 ==> item3=Fine rice 3 conf: (1)
9.	item1=Beef 3 item3=Fine rice 3 ==> item4=Milk 3 conf: (1)
10.	item4=Milk 3 ==> item1=Beef 3 item3=Fine rice 3 conf: (1)

The first 6 rules shown in TABLE III are generated using only any 2 items of the item subset I_s. Each of these 6 rules contains any two items of I_s. All of these two items of I_s are found in each of the 3 records out of the 5 purchase records shown in TABLE II. These 2-item rules contain items which have high occurrences (60%) within the purchase records. The remaining 4 rules contain all of the 3 items contained in the item subset I_s which are found in each of the 3 records out of the 5 purchase records shown in TABLE II. These 3-item rules contain items which also have high occurrences (60%) within the purchase records. The association rules which can be constructed using the items $I - I_s$ are ignored as these rules will contain low occurring items of the purchase records. The recent version of Weka 3.7.12 also produces the same association rules from the purchase records of TABLE II with some additional parameter values which are skipped here for the purpose of simplicity.

B. Approach 2

In this approach, customized association rule generation and analysis technique [1] is applied on user input item set to generate association rules which are then analyzed using the purchase records of customers to extract sales decision making information. In this analysis, all of the possible association rules which can be generated from the items of the input item set are only considered. The association rules generated from the input items not contained in the sales database are rejected. The association rule generation process starts by taking an item set as input, and generates the antecedent and consequent item sets taking all of the possible non-empty subsets of the input item set. In this customized approach, association rules are generated from item set generated only from the valid user input item set containing sales items rather than generating a full set of association rules containing all of the items or a subset of the items contained in the sales relation. This approach generates a small number of association rules for analyzing a query in relation to a specific business decision. In this case, the generated association rules are analyzed using a single sales relation called *itemssold* with the sales records shown in TABLE IV [1] to generate sales decision making information. In this approach, 4 analysis cases are performed for 4 different input item sets to generate sales decision making information from the generated association rules. In each analysis case, a table represents the number of the occurrences of the antecedent item set X and the consequent item set Y within the sales relation by X_N and Y_N respectively along with the support *s* and the confidence *c* of each rule.

TABLE IV. PURCHASE RECORDS OF CUSTOMERS IN ITEMSSOLD RELATION [1]

ID	Items				
CustomerID	***item1***	***item2***	***item3***	***item4***	***Item5***
0001	Beef	Ruhi	Fine rice	Milk	Soft drink
0002	Beef	Hilsha	Fine rice	Milk	Soft drink
0003	Beef	Hilsha	Fine rice	Milk	Soft drink
0004	Onion	Rice	Banana	Bread	Dal
0005	Beef	Hilsha	Fine rice	Milk	Soft drink
0006	Beef	Hilsha	Fine rice	Milk	Soft drink
0007	Onion	Hilsha	Carrots	Bread	Milk
0008	Beef	Fine rice	Hilsha	Milk	Soft drink
0009	Beef	Hilsha	Fine rice	Milk	Soft drink
0010	Beef	Shrimp	Fine rice	Milk	Cooking oil
0011	Beef	Ruhi	Fine rice	Milk	Soft drink
0012	Beef	Hilsha	Fine rice	Milk	Soft drink
0013	Beef	Hilsha	Carrots	Milk	Soft drink
0014	Onion	Rice	Banana	Bread	Fruit
0015	Beef	Shrimp	Fine rice	Milk	Cooking oil
0016	Beef	Katla	Fine rice	Milk	Pasta
0017	Onion	Hilsha	Carrots	Bread	Milk
0018	Beef	Hilsha	Fine rice	Milk	Fruit
0019	Beef	Hilsha	Fine rice	Milk	Soft drink
0020	Beef	Shrimp	Fine rice	Milk	Cooking oil
0021	Beef	Shrimp	Fine rice	Milk	Soft drink

The *customer* relation is used to inform sales offer information to the appropriate customer. In this case, the attributes of these two relations are considered as shown below.

customer(*customerID*, *customerName*, *address*, *mobilePhone*, *email*)
itemssold(*customerID, item1, item2, item3, …, itemn*)

The customized association rule generation and analysis technique is used to analyze the customer's purchase records shown in TABLE IV for each of the generated association rules to generate sales decision making information. A number of analysis cases are completed for testing [14] [15] the customized association rule generation and analysis technique to generate sales decision making information represented with (customers, missing items) pair.

Definition 1: (*customers*, *missingitems*) pair. In this definition, for an association rule, a (*customers*, *missingitems*) pair is discovered where *customers* represent those customers who did not buy the items which are *missingitems* in the current purchase record for a successful matching of the rule antecedent X. Missing items in a matched record are those which are frequent items in other records, but missing in the currently matched record, so it may be placed on offer. The following symbols are used to determine a (*customers*, *missingitems*) pair.

$I_R \rightarrow$ rule item set

$I_X \rightarrow$ rule antecedent item set

$I_Y \rightarrow$ rule consequent item set

$I_{R'} \rightarrow$ matched record item subset which is matched with a subset of the rule items

$I_M \rightarrow$ missing item set to be placed on offer

customers $\rightarrow$ List of customers, initially {}

missingitems $\rightarrow$ List of missing items, initially {}

If *match*(I_X, $I_{R'}$) = False, then (*customers, missing items*) = ({}, {}). For matching to be successful, $I_R = I_X \cup I_Y$, $I_{R'} \subseteq I_R$, $I_X \subseteq I_{R'}$. In the case of a successful matching, (*customers*, *missingitems*) = (*customers* $\cup$ *customerID*, *missingitems* $\cup$ $item_i$) for i = 1 to n, where n is the number of items. For a particular association rule, the items of the missing item set to be placed on offer will be $I_M = I_R \setminus I_{R'}$.

1) Analysis Case 1: The item set {Beef, Fine rice, Soft drink} is used as input to the customized association rule generation and analysis technique [1] to generate association rules consisting of all of these 3 input items. Each of these generated rules is analyzed using the customer's purchase records shown in TABLE IV. The generated association rules with their corresponding support and confidence with the floating point values rounded up to 2 precision points are provided in TABLE V. Within the purchase records, only 11 records contain all of the rule items.

TABLE V. GENERATED ASSOCIATION RULES WITH TTHE SUPPORT AND CONFIDENCE USING CUSTOMIZED ASSOCIATION RULE GENERATION AND ANALYSIS TECHNIQUE [1]

Rule No.	Generated 3-item Association Rules	X_N	Y_N	s %	c %
1.	{Beef} $\Rightarrow$ {Fine rice, Soft drink}	17	11	52.38	64.71
2.	{Fine rice} $\Rightarrow$ {Beef, Soft drink}	16	12	52.38	68.75
3.	{Soft drink} $\Rightarrow$ {Beef, Fine rice}	12	16	52.38	91.67
4.	{Beef, Fine rice} $\Rightarrow$ {Soft drink}	16	12	52.38	68.75
5.	{Beef, Soft drink} $\Rightarrow$ {Fine rice}	12	16	52.38	91.67
6.	{Fine rice, Soft drink} $\Rightarrow$ {Beef}	11	17	52.38	100.0

The analysis of the generated association rules shown in TABLE V demonstrates that the rules consist of only input item set {Beef, Fine rice, Soft drink} among the 17 items I_T = {Beef, Ruhi, Fine rice, Milk, Soft drink, Hilsha, Onion, Rice, Banana, Bread, Dal, Carrots, Shrimp, Cooking oil, Fruit, Katla, Pasta} contained in the *itemssold* relation shown in TABLE IV. Thus, only a small number of rules are analyzed to make any sales decision regarding these 3 input items forming a item group. This technique eliminates the need for analyzing a large number of rules that could be generated from all of the 17 items found within 21 data records shown in TABLE IV. The extraction of any sales decision making information usually relates to a single item, a pair or a group of 3 to 4 items. Customers often purchase items in pairs or groups, and the number of items in the group usually does not exceed 4. Some examples of such pairs of items or item groups which go together are {Computer, Software}, {Beef, Fine rice, Soft drink}, {Bread, Milk}, {Bread, Egg}, {Bread, Banana}, {Bread, Milk, Banana, Egg} etc. which may help to determine customers buying pattern. The missing of an item in a group within the customer's purchase records may be considered as a data mining problem, and the frequency of occurrence of this item in the pair or item group within the purchase records may be used in sales decision making. A large number of rules can be generated from the items of I_T taking a minimum of 2 items in the antecedent and consequent parts of a rule. Thus, the customized approach avoids the need for analyzing all of the association rules which can be generated from the 17 items. To make a sales decision in relation to any subset of the item set I_T, analyzing irrelevant rules may be cumbersome, time consuming and unrealistic also.

a) Generated Sales Decision Making Information: In analysis case 1, only 3-item rules shown in TABLE V are analyzed to generate sales decision making information by the customized association rule generation and analysis technique [1] with minimum support = 50% and minimum confidence = 60%. To assist in sales decision making, a list of customers is generated who may be interested in the corresponding items as listed in TABLE VI by analyzing each of the generated association rules shown in TABLE V using the purchase records of TABLE IV.

TABLE VI. GENERATED SALES DECISION MAKING INFORMATION FOR ITEM SET {BEEF, FINE RICE, SOFT DRINK} [1]

Rule No.	Customers	Items on Offer
1.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
1.	{0013}	{Fine rice}
2.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
3.	{0013}	{Fine rice}
4.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
5.	{0013}	{Fine rice}
6.	{}	{}

In the above analysis, it is assumed that the higher the support and the higher the confidence of an association rule over the minimum support and minimum confidence, it is more likely that the items within the rule under analysis would be more frequent in the customers buying pattern. Based on this assumption, the sales decision making information about customers and the items of interest, *i.e.*, missing item(s) in the purchase record(s), to be placed on offer is generated as shown in TABLE VI.

From the above analysis, it can be predicted that for any rule satisfying the minimum support and minimum confidence, the item(s) missing in a customer's purchase record which contains a subset of the valid input items, all of the items of the rule antecedent X, and at least one item of the rule consequent Y, may attract some customers if the item(s) is placed on offer. From this analysis, we can also decide that the stock of these rule items should be increased.

2) Analysis Case 2: In this analysis case, the input item set {Beef, Fine rice, Milk, Soft drink} is used by the customized association rule generation and analysis technique [1] to generate 4-item association rules. Each of these generated rules is analyzed using the customer's purchase records of *itemssold* relation shown in TABLE IV to generate sales decision making information. TABLE VII provides the generated association rules consisting of all of the 4 input items with their corresponding support and confidence with the floating point values rounded up to 2 precision points. Within the purchase records, only 11 records contain all of the rule items.

TABLE VII. GENERATED ASSOCIATION RULES WITH THE SUPPORT AND CONFIDENCE USING CUSTOMIZED ASSOCIATION RULE GENERATION AND ANALYSIS TECHNIQUE

Rule No.	Generated 4-item Association Rules	X_N	Y_N	s %	c %
1.	{Beef} ⇒ {Fine rice, Milk, Soft drink}	17	11	52.38	64.71
2.	{Fine rice} ⇒ {Beef, Milk, Soft drink}	16	12	52.38	68.75
3.	{Milk} ⇒ {Beef, Fine rice, Soft drink}	19	11	52.38	57.89
4.	{Soft drink} ⇒ {Beef, Fine rice, Milk}	12	16	52.38	91.67
5.	{Beef, Fine rice} ⇒ {Milk, Soft drink}	16	12	52.38	68.75
6.	{Beef, Milk} ⇒ {Fine rice, Soft drink}	17	11	52.38	64.71
7.	{Beef, Soft drink} ⇒ {Fine rice, Milk}	12	16	52.38	91.67
8.	{Fine rice, Milk} ⇒ {Beef, Soft drink}	16	12	52.38	68.75
9.	{Fine rice, Soft drink} ⇒ {Beef, Milk}	11	17	52.38	100.0
10.	{Milk, Soft drink} ⇒ {Beef, Fine rice}	12	16	52.38	91.67
11.	{Beef, Fine rice, Milk} ⇒ {Soft drink}	16	12	52.38	68.75
12.	{Fine rice, Milk, Soft drink} ⇒ {Beef}	11	17	52.38	100.0
13.	{Milk, Soft drink, Beef} ⇒ {Fine rice}	12	16	52.38	91.67
14.	{Soft drink, Beef, Fine rice} ⇒ {Milk}	11	19	52.38	100.0

a) Generated Sales Decision Making Information: In this analysis case, only 14 association rules consisting of 4 items shown in TABLE VII are analyzed by applying the customized association rule generation and analysis technique [1] for generating sales decision making information as provided in TABLE VIII considering minimum support = 50% and minimum confidence = 60%.

TABLE VIII. GENERATED SALES DECISION MAKING INFORMATION FOR ITEM SET {BEEF, FINE RICE, MILK, SOFT DRINK}

Rule No.	Customers	Items on Offer
1.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
1.	{0013}	{Fine rice}
2.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
4.	{0013}	{Fine rice}
5.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
6.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
6.	{0013}	{Fine rice}
7.	{0013}	{Fine rice}
8.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
9.	{}	{}
10.	{0013}	{Fine rice}
11.	{0010, 0015, 0016, 0018, 0020}	{Soft drink}
12.	{}	{}
13.	{0013}	{Fine rice}
14.	{}	{}

The sales decision making information about customers and the items of interest, i.e., missing item(s) in their purchase records, to be placed on offer is generated. In the above analysis, all of the association rules except rule 3 shown in TABLE VII satisfy the minimum support and minimum confidence. Rule 3 is excluded from sales decision making information generation as shown in TABLE VIII as it's confidence does not satisfy the minimum confidence level. Using the prediction of analysis case 1, it is expected that the items shown in TABLE VIII may attract some customers if the items are placed on offer. From this analysis, it may also be decided that the stock of these rule items should be increased. In TABLE VIII, the customer list and the items of interest to be placed on offer both are {} for the association rules 9, 12 and 14. In this generation process, the items contained in the antecedent X of the rule under analysis must be fully matched with any of the purchase record items, and also the rule consequent $Y \neq \{\}$ such that $\exists i \in Y$ and $i \in I_T$ where i is any item contained in the purchase records.

3) Analysis Case 3: The input item set {Onion, Hilsha, Bread, Milk} is used by the customized association rule generation and analysis technique [1] to generate 4-item association rules. Each of these generated rules is analyzed using the purchase records of the itemssold relation to generate information for the valid rules only. TABLE IX provides the generated association rules consisting of all of the input items with their corresponding support and confidence with the floating point values rounded up to 2 precision points.

TABLE IX. GENERATED ASSOCIATION RULES WITH THE SUPPORT AND CONFIDENCE USING CUSTOMIZED ASSOCIATION RULE GENERATION AND ANALYSIS TECHNIQUE

Rule No.	Generated 4-item Association Rules	X_N	Y_N	s %	c %
1.	{Onion} ⇒ {Hilsha, Bread, Milk}	4	2	9.52	50.0
2.	{Hilsha} ⇒ {Onion, Bread, Milk}	12	2	9.52	16.67
3.	{Bread} ⇒ {Onion, Hilsha, Milk}	4	2	9.52	50.0
4.	{Milk} ⇒ {Onion, Hilsha, Bread}	19	2	9.52	10.53
5.	{Onion, Hilsha} ⇒ {Bread, Milk}	2	2	9.52	100.0
6.	{Onion, Bread} ⇒ {Hilsha, Milk}	4	12	9.52	50.0
7.	{Onion, Milk} ⇒ {Hilsha, Bread}	2	2	9.52	100.0
8.	{Hilsha, Bread} ⇒ {Onion, Milk}	2	2	9.52	100.0
9.	{Hilsha, Milk} ⇒ {Onion, Bread}	12	4	9.52	16.67
10.	{Bread, Milk} ⇒ {Onion, Hilsha}	2	2	9.52	100.0
11.	{Onion, Hilsha, Bread} ⇒ {Milk}	2	19	9.52	100.0
12.	{Hilsha, Bread, Milk} ⇒ {Onion}	2	4	9.52	100.0
13.	{Bread, Milk, Onion} ⇒ {Hilsha}	2	12	9.52	100.0
14.	{Milk, Onion, Hilsha} ⇒ {Bread}	2	4	9.52	100.0

Within the purchase records, only 2 records contain all of the rule items. In this analysis case, only the generated 4-item association rules are analyzed for generating sales decision making information using the customized association rule generation and analysis technique [1] considering minimum support = 50% and minimum confidence = 60%. In this case, no sales decision making information is generated as the support of each association rule is 9.52% which is less than the minimum support. We can identify the infrequent items from TABLE IX. The business authority must have to make efficient policies to increase the sales of these infrequent items and also make necessary steps for proper stock management.

4) Analysis Case 4: In this case, the input item set {Beef, Shrimp, Fine rice, Milk} is used by the customized association rule generation and analysis technique [1] to generate 4-item association rules as shown in TABLE X. Each of these generated association rules is analyzed using the customer's purchase records to generate sales decision making information. TABLE X provides the generated association rules with their corresponding support and confidence with the floating point values rounded up to 2 precision points. Within the purchase records, only 4 records contain all of the rule items. In this analysis, only 4-item rules are used for generating sales decision making information by the customized association rule generation and analysis technique [1]. Considering the minimum support = 50% and minimum confidence = 60%, no sales decision making information is generated as the support 19.05% of the generated rules does not satisfy the minimum support value.

TABLE X. GENERATED ASSOCIATION RULES WITH THE SUPPORT AND CONFIDENCE USING CUSTOMIZED ASSOCIATION RULE GENERATION AND ANALYSIS TECHNIQUE

Rule No.	Association Rules Consisting of 4 Input Items	X_N	Y_N	s %	c %
1.	{Beef} ⇒ {Shrimp, Fine rice, Milk}	17	4	19.05	23.53
2.	{Shrimp} ⇒ {Beef, Fine rice, Milk}	4	16	19.05	100.0
3.	{Fine rice} ⇒ {Beef, Shrimp, Milk}	16	4	19.05	25.0
4.	{Milk} ⇒ {Beef, Shrimp, Fine rice}	19	4	19.05	21.05
5.	{Beef, Shrimp}⇒ {Fine rice, Milk}	4	16	19.05	100.0
6.	{Beef, Fine rice} ⇒ {Shrimp, Milk}	16	4	19.05	25.0
7.	{Beef, Milk} ⇒ {Shrimp, Fine rice}	17	4	19.05	23.53
8.	{Shrimp, Fine rice}⇒ {Beef, Milk}	4	17	19.05	100.0
9.	{Shrimp, Milk} ⇒ {Beef, Fine rice}	4	16	19.05	100.0
10.	{Fine rice, Milk} ⇒ {Beef, Shrimp}	16	4	19.05	25.0
11.	{Beef, Shrimp, Fine rice} ⇒ {Milk}	4	19	19.05	100.0
12.	{Shrimp, Fine rice, Milk} ⇒ {Beef}	4	17	19.05	100.0
13.	{Fine rice, Milk, Beef} ⇒ {Shrimp}	16	4	19.05	25.0
14.	{Milk, Beef, Shrimp} ⇒ {Fine rice}	4	16	19.05	100.0

We can identify the infrequent items from TABLE X. The business authority must have to make efficient policies to increase the sales of these items and also make necessary steps for proper stock management.

V. CONCLUSION AND FUTURE WORK

In this paper, the system architecture of a data mining framework for the application of data mining technique to the sales database of a super shop has been presented. It is shown that association rules can be generated from valid input items with respect to the sales data under analysis contained in the sales database. The outcome of the application of the customized association rule generation and analysis technique on sales database has been presented in this paper. The success of this technique in generating sales decision making information usually depends on the frequency of occurrences of the items within the sales data records, and the nature of customer's buying pattern of purchasing particular item groups which go together. The research work demonstrates how the generated association rules can be used to produce sales decision making information by analyzing sales data. Each of the generated association rules is verified to see whether the correct association rules are generated consisting of only user input items contained in the sales database. A table representing the number of occurrences of the rule antecedent, consequent, and the support and confidence of each rule for a given user input item set is provided. This work clarifies how the application of the data mining technique using association rules can help to improve decision making in a sales system using the generated sales decision making information. In this work, a single relation of sales data is used for generating sales decision making information by analyzing each of the generated association rules. To improve customer service using the generated sales decision making information, a customer database relation can be used to inform sales offer information to the appropriate customer. Future work may include multiple relations for association rule mining and analysis to generate sales decision making business intelligence. More research is needed for the analysis of less frequent items in sales systems. The improvement in the current implementation of the customized association rule generation and analysis technique may also be done in the near future.

ACKNOWLEDGMENT

The author would like to thank all of his colleagues and experts who have made valuable comments to this work. He also thanks all of the people who helped in any matter related to this research work.

REFERENCES

[1] Md. Humayun Kabir, Integration and Testing of Association Rule Discovery Algorithm for Developing Masranga Data Mining Software System, Research Project Report, 2012, Faculty of Mathematical and Physical Sciences, Jahangirnagar University, Savar, Dhaka, Bangladesh.

[2] Md. Humayun Kabir, Development of a Data Mining Software for Data Warehouse Construction and Mining Interesting Patterns from Organizational Databases, Research Project Report, 2010, Faculty of Mathematical and Physical Sciences, Jahangirnagar University, Savar, Dhaka, Bangladesh.

[3] R. Agrawal, T. Imielimski and A. Swami, "Mining Association Rules between Sets of Items in Large Databases", Proceedings of the ACM SIGMOD Conference on Management of Data, Washington DC, pp. 207-216.

[4] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules", Proceedings of the 20th VLDB Conference, 1994, Santiago, Chile, pp. 487-499.

[5] R. Srikant and R. Agrawal, "Mining generalized association rules", Proceedings of the 21st VLDB Conference, 1995, Zurich, Swizerland, pp. 407-419.

[6] M. Karaolis, J. A. Moutiris, L. Papconstantinou, and C. S. Pattichis, "Association Rule Analysis for the Assessment of the Risk of the Coronary Heart Events", in the Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, 2009 (EMBC 2009), pp. 6238-6241. IEEE Conference Publications.

[7] Z. Zhu and J.-Y. Wang, "Book recommendation service by improved association rule mining algorithm", Proceedings of the 6th International

Conference on Machine Learning and Cybernetics, Hong Kong, IEEE Conference Publications, 2007, pp 3864-3869.

[8] S. Nestorov and N. Jukić, "Ad-Hoc association rule mining within the data warehouse", Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03), IEEE Computer Society 2002.

[9] A. Martin, T. M. Lakshmi, and V. P. Venkatesan, "A Business Intelligence Framework for Business Performance using Data Mining Techniques", in the International Conference on Emerging Trends in Science, Engineering and Technology 2012, IEEE publications, pp. 373-380.

[10] R. T. Hans and E. MnKandla, "Modeling software engineering projects as a Business: A Business Intelligence Perspective", AFRICON 2013, IEEE Conference Publications, pp. 1-5.

[11] V. S. Tseng, B.-E. Shie, C.-W. Wu and P. S. Yu, "Efficient Algorithms for Mining High Utility Itemsets from Transactional Databases", IEEE Transactions on Knowledge and Data Engineering, Vol. 25, No. 8, August 2013, IEEE Computer Society, pp. 1772-1786.

[12] B. Li, Z. Pei, K. Qin, "Association Rules Mining based on Clustering Analysis and Soft Sets", IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing. 2015, IEEE Computer Society, pp. 675-681.

[13] J. Gautam and N. Srivastava, "Analysis of medical domain using CMARM: Confabulation Mapreduce association rule mining algorithm for frequent and rare itemsets", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 11, 2015, pp. 224-228.

[14] J. Collofello and K. Vehathiri, "An Environment for Training Computer Science Students on Software Testing", Proceedings of the 35th Annual Conference on Frontiers in Education FIE '05, 2005, pp. T3E-6.

[15] Y. Labiche, "Integration testing object-oriented software systems: An experiment driven research approach", 24th Canadian Conference on Electrical and Computer Engineering (CCECE) 2011, IEEE Conference Publications, pp. 652-655.

A Reversible Data Hiding Scheme for BTC-Compressed Images

Ching-Chiuan Lin
Department of Information Technology
Overseas Chinese University
Taichung, Taiwan

Kuo Feng Hwang
Department of Information Technology
Overseas Chinese University
Taichung, Taiwan

Shih-Chieh Chen
Department of Multimedia and Game Design
Overseas Chinese University
Taichung, Taiwan

Chi-Ming Yao
Department of Information Technology
Overseas Chinese University
Taichung, Taiwan

***Abstract*—This paper proposes a reversible data hiding scheme for BTC-compressed images. A block in the BTC-compressed image consists of a larger block-mean pixel and a smaller block-mean pixel. Two message bits are embedded into a pair of neighboring blocks. One is embedded by expanding the difference between the two larger block-mean pixels and the other is embedded by expanding the one between the two smaller block-mean pixels. Experimental results show that the embedding strategy may decrease the modification of images. The proposed scheme may obtain a stego-image with high visual quality and a payload capacity of one bit per block, approximately.**

***Keywords*—Block Truncation Coding; Reversible Data Hiding; Difference Expansion**

I. Introduction

Transmitting secret data over the Internet is a popular application. To prevent secret data from malicious attack, users usually encrypt important data before transmission. Encrypting data is a complicated computation which coverts data into a meaningless format, which may attract hackers' attention and result in undesired attack. Hiding data in an image is an alternative way for secret communication. It embeds important data by slightly modifying the image. Hackers may not percept that important data are embedded in the normal image. Therefore, an undesired attack may be avoided.

To speed up transmission or decrease required storage, images are usually converted into smaller ones with compressed format. JPEG[1] is a popular image format which applying Discrete Cosine Transform (DCT) to compress an image. It needs complicated computation for image compression and decompression. Another technique for image compression is Vector Quantization (VQ) [2] which records image blocks in a code book and uses an index of code word to encode a block of image. A popular block-oriented image compression method is Block Truncation Coding (BTC) [3]. Compared to JPEG and VQ, BTC is a simple and efficient encoding method for image compression.

Many data hiding schemes for BTC-compressed images were proposed [4–9]. However, most of them were irreversible after secret data were extracted. Namely, the original image may not be completely recovered. This may degrade the image and decrease user's motivation for hiding data in the image. Therefore, a reversible data hiding scheme, which can completely recover the original image, is required.

Reversible data hiding schemes embed data in redundant space of an image [10–13]. Most of the schemes belong to the two families: shifting histogram and difference expansion. The former shifts histogram of pixels or differences to get redundant space located in the peak point of histogram for embedding a message. The latter expands the difference between a pair of pixels, i.e. doubles the difference. As a result, the expanded difference would be an even number and the least significant bit of each expanded difference value is equal to 0 which is the available embedding space.

TABLE I. An Example of Difference Expansion

(y_1, y_2)	Δy	ΔY	Embed a bit of 0		Embed a bit of 1	
			(y_1', y_2')	$\Delta Y'$	(y_1', y_2')	$\Delta Y'$
(95, 95)	0	0	(95, 95)	0	(95, 96)	1
(95, 96)	1	2	(94, 96)	2	(94, 97)	3
(94, 96)	2	4	(93, 97)	4	(93, 98)	5
(94, 97)	3	6	(92, 98)	6	(92, 99)	7

Table I is an example illustrating the embedding process of difference expansion, where (y_1, y_2) are pixel values, Δy and ΔY are original and expanded differences, respectively. After embedding a bit of 0 or 1 into the pair of pixels, their stego-pixel values and difference would become (y_1', y_2') and $\Delta Y'$, respectively, as shown in the table. If $\Delta y = 1$, expanding a difference may be implemented by either increasing the larger pixel or decreasing the smaller by one. If $\Delta y > 1$, increasing the larger one and decreasing the smaller one at the same time is a better option, since it may result in less modification of pixel values and get a benefit of smaller perception of image distortion by human vision.

In the decoding process, a bit of $s = 0$ or $s = 1$ is extracted if $\Delta Y'$ is equal to an even or odd number, respectively, and the original difference may be obtained by calculating $\Delta y = \lfloor \Delta Y'/2 \rfloor$. Then $y_1 = y_1' + \lfloor (\Delta y + 1)/2 \rfloor$ and $y_2 = y_2' -$

$\lfloor \Delta y/2 \rfloor - s$ are completely recovered. For a 256-level grayscale image, if blocks with $y_1' < 0$ or $y_2' > 255$, these blocks would not be candidates for embedding a message. These exceptions are recorded in the overhead information for identifying if a block embeds a message.

This paper proposes a reversible data hiding scheme for BTC-compressed images. Embedding space is gotten from expanding a difference between mean pixel values. A block in the BTC-compressed image consists of a larger block-mean pixel and a smaller block-mean pixel. Two message bits are embedded into a pair of neighboring blocks. One is embedded by expanding the difference between the two larger block-mean pixels and the other is embedded by expanding the one between the two smaller block-mean pixels. Experimental results show that the embedding strategy may decrease the modification of images. The proposed scheme may obtain a stego-image with high visual quality and a payload capacity of one bit per block, approximately.

The rest of this paper is organized as follows. The BTC algorithm is briefly reviewed in Section II. Section III introduces the proposed scheme including embedding and extraction processes. To help readers understand the proposed scheme, an embedding example is also given in this section. Section IV demonstrates our experimental results in terms of image visual quality and payload capacity. Finally, conclusions are given in Section V.

II. Block Truncation Coding

In the following, the BTC encoding algorithm is briefly reviewed. Notations are defined as follows:

- X_i is a block of image, with $n = k \times k$ pixels in the block,
- $x_i(p,q)$ is the pixel value of block i where p and q are indexes of pixels and $1 \le p, q \le k$,
- $\bar{x}_i = \frac{\sum_{p=1}^{k}\sum_{q=1}^{k} x_i(p,q)}{k \times k}$ is the average pixel value of block i,
- m_i is the number of pixels in block i with $x_i(p,q) \ge \bar{x}_i$,
- $\bar{x}_i^{MAX} = \left\lfloor \frac{\sum x_i}{m_i} |_{x_i(p,q) \ge \bar{x}_i} \right\rfloor$ is the larger block-mean pixel value, for $x_i(p,q) \ge \bar{x}_i$, in block i,
- $\bar{x}_i^{min} = \left\lfloor \frac{\sum x_i}{n - m_i} |_{x_i(p,q) < \bar{x}_i} \right\rfloor$ is the smaller block-mean pixel value, for $x_i(p,q) < \bar{x}_i$, in block i, and
- X_i' is the decoded block of BTC-compressed image.

The BTC encoding algorithm is introduced as follows:

1) Select an image and divide it into non-overlapping blocks each of them contains $k \times k$ pixels.

2) For each block i, calculate $\bar{x}_i$, and then m_i, $\bar{x}_i^{MAX}$, $\bar{x}_i^{min}$.

3) Encode block i denoted by $\hat{B}_i = (\bar{x}_i^{MAX}, \bar{x}_i^{min}, B_i)$, where $B_i = \{b_i(p,q) | b_i(p,q) = 0 \ or \ 1\}$ is a binary block with $b_i(p,q) = 0$ and $b_i(p,q) = 1$ if $x_i(p,q) < \bar{x}_i$ and $x_i(p,q) \ge \bar{x}_i$, respectively.

4) Repeat step 3 until all blocks are encoded.

The following example gives an illustration for the process of BTC. Given, for a 256-level grayscale image, a block $X_i = \{x_i(p,q) | 0 \le x_i(p,q) \le 255\}$ is as follows:

136	132	133	134
135	134	137	138
132	132	131	132
133	134	135	136

We have $\bar{x}_i = 134$, $m_i = 9$, $\bar{x}_i^{MAX} = 135$, $\bar{x}_i^{min} = 132$ and $B_i = \{b_i(p,q) | b_i(p,q) = 0 \text{ or } 1\}$ as shown below.

1	0	0	1
1	1	1	1
0	0	0	0
0	1	1	1

Note that $\bar{x}_i$ may not be exactly equal to $(\bar{x}_i^{MAX} + \bar{x}_i^{min})/2$. For convenience, we use $\hat{B}_i = (\bar{x}_i^{MAX}, \bar{x}_i^{min}, B_i)$ to denote an encoded block. In the example, $\bar{x}_i^{MAX}$, $\bar{x}_i^{min}$ and B_i need a memory space of 8, 8 and 16 bits, respectively. Compared to its uncompressed block, the compression rate of BTC is $(8+8+16)/(8*16) = 0.25$. This may significantly decrease the required space for storing an image without complicated computation.

The encoded image is decoded by replacing $b_i(p,q)$ with $\bar{x}_i^{MAX}$ or $\bar{x}_i^{min}$ if $b_i(p,q) = 1$ or $b_i(p,q) = 0$, respectively. In the above example, X_i' is decoded as follows:

135	132	132	135
135	135	135	135
132	132	132	132
132	135	135	135

III. Proposed Scheme

The proposed scheme is designated to embed a message into a BTC-compressed image and extract the message from the stego-image. Therefore, the BTC encoding procedure in Section II must be applied to a grayscale cover image if it is not compressed by the BTC algorithm. We will introduce the proposed scheme, including the embedding and extraction procedures in the following sections. Note that the stego-image would be completely recovered in the proposed scheme.

A. Embedding procedure

The embedding procedure is used to embed a binary bit string into a BTC-compressed image *T*. Required embedding space is obtained from expanding the difference between block-mean values in blocks. Details are listed as follows:

1) Convert a message into a binary bit string $S = s_1 s_2 \ldots s_i \ldots$, where $s_i \in \{0,1\}$. For example, a message $(23)_{16}$ is converted into $(00100011)_2$ and $s_1 = 0, s_2 = 0, s_3 = 1$, etc.

2) Sequentially scan the BTC-compressed image T in an order which was negotiated with the decoder. For each pair of blocks $\hat{B}_{2i-1} = (\bar{x}_{2i-1}^{MAX}, \bar{x}_{2i-1}^{min}, B_{2i-1})$ and $\hat{B}_{2i} = (\bar{x}_{2i}^{MAX}, \bar{x}_{2i}^{min}, B_{2i})$, $i = 1,2,3,\ldots$ in the image T, calculate

$$\Delta x_i^{MAX} = \left| \bar{x}_{2i-1}^{MAX} - \bar{x}_{2i}^{MAX} \right|,$$

$$\Delta x_i^{min} = \left|\bar{x}_{2i-1}^{min} - \bar{x}_{2i}^{min}\right|,$$

$$\begin{pmatrix}\bar{\bar{x}}_{2i-1}^{MAX}\\ \bar{\bar{x}}_{2i}^{MAX}\end{pmatrix} = \begin{cases}\begin{pmatrix}\bar{x}_{2i-1}^{MAX} + \lfloor\Delta x_i^{MAX}/2\rfloor + s_{2i-1}\\ \bar{x}_{2i}^{MAX} - \lfloor(\Delta x_i^{MAX}+1)/2\rfloor\end{pmatrix} & \text{if } \bar{x}_{2i-1}^{MAX} > \bar{x}_{2i}^{MAX},\\ \begin{pmatrix}\bar{x}_{2i-1}^{MAX} - \lfloor(\Delta x_i^{MAX}+1)/2\rfloor\\ \bar{x}_{2i}^{MAX} + \lfloor\Delta x_i^{MAX}/2\rfloor + s_{2i-1}\end{pmatrix} & \text{otherwise,}\end{cases}$$

$$\begin{pmatrix}\bar{\bar{x}}_{2i-1}^{min}\\ \bar{\bar{x}}_{2i}^{min}\end{pmatrix} = \begin{cases}\begin{pmatrix}\bar{x}_{2i-1}^{min} + \lfloor\Delta x_i^{min}/2\rfloor + s_{2i}\\ \bar{x}_{2i}^{min} - \lfloor(\Delta x_i^{min}+1)/2\rfloor\end{pmatrix} & \text{if } \bar{x}_{2i-1}^{min} > \bar{x}_{2i}^{min},\\ \begin{pmatrix}\bar{x}_{2i-1}^{min} - \lfloor(\Delta x_i^{min}+1)/2\rfloor\\ \bar{x}_{2i}^{min} + \lfloor\Delta x_i^{min}/2\rfloor + s_{2i}\end{pmatrix} & \text{otherwise.}\end{cases}$$

3) If $0 \le \bar{\bar{x}}_{2i-1}^{min}, \bar{\bar{x}}_{2i-1}^{MAX}, \bar{\bar{x}}_{2i}^{min}, \bar{\bar{x}}_{2i}^{MAX} \le 255$, *encode blocks* $2i-1$ *and* $2i$ *as* $\ddot{B}_{2i-1} = (\bar{\bar{x}}_{2i-1}^{MAX}, \bar{\bar{x}}_{2i-1}^{min}, B_{2i-1})$ *and* $\ddot{B}_{2i} = (\bar{\bar{x}}_{2i}^{MAX}, \bar{\bar{x}}_{2i}^{min}, B_{2i})$.

4) If $\bar{\bar{x}}_{2i-1}^{min} < 0$, $\bar{\bar{x}}_{2i}^{min} < 0$, $\bar{\bar{x}}_{2i-1}^{MAX} > 255$, or $\bar{\bar{x}}_{2i}^{MAX} > 255$, let the pair of blocks be unchanged, i.e. $\ddot{B}_{2i-1} = (\bar{x}_{2i-1}^{MAX}, \bar{x}_{2i-1}^{min}, B_{2i-1})$ and $\ddot{B}_{2i} = (\bar{x}_{2i}^{MAX}, \bar{x}_{2i}^{min}, B_{2i})$, and record the pair of blocks index $2i-1$ as an overhead information. In other words, this pair of blocks embeds nothing, and s_{2i-1} and s_{2i} are embedded into the next pair of blocks.

5) Obtain the stego-image $T' = \{\ddot{B}_i | i = 1,2,3,\ldots\}$.

Note that if $\Delta x_i^{MAX} = 0$, $\Delta x_i^{min} = 0$, $s_{2i-1} = 0$, and $s_{2i} = 0$, the pair of blocks are also unchanged and they embed two bits of 0. This means an unchanged pair of blocks is not equivalent to embedding nothing.

The proposed scheme embeds two bits in a pair of blocks, instead of embedding one bit in a block. Our embedding strategy is to decrease the modification of an image. According to our experiments, for most blocks, Δx_i^{MAX} or Δx_i^{min} is usually less than $\left|\bar{x}_{2i-1}^{MAX} - \bar{x}_{2i-1}^{min}\right|$ or $\left|\bar{x}_{2i}^{MAX} - \bar{x}_{2i}^{min}\right|$. Namely, expanding a smaller difference can make slighter image modification than expanding a larger one.

B. Extraction procedure

Whenever a decoder would like to extract the embedded message from a stego-image and recover it to its original BTC-compressed image *T*, the extraction procedure would be applied. Details of the procedure are listed as follows:

1) Block by block scan the stego-image T' as the order in the embedding procedure.

2) For each pair of blocks $\ddot{B}_{2i-1} = (\bar{\bar{x}}_{2i-1}^{MAX}, \bar{\bar{x}}_{2i-1}^{min}, B_{2i-1})$ and $\ddot{B}_{2i} = (\bar{\bar{x}}_{2i}^{MAX}, \bar{\bar{x}}_{2i}^{min}, B_{2i})$ $i = 1,2,3,\ldots$, if index $2i-1$ is not recorded in the overhead information, do step 3, otherwise skip blocks $\ddot{B}_{2i-1}$ and $\ddot{B}_{2i}$ since they embed nothing and are not needed to be recovered.

3) Calculate

$$\Delta\ddot{x}_i^{MAX} = \left|\bar{\bar{x}}_{2i-1}^{MAX} - \bar{\bar{x}}_{2i}^{MAX}\right|,$$

$$\Delta\ddot{x}_i^{min} = \left|\bar{\bar{x}}_{2i-1}^{min} - \bar{\bar{x}}_{2i}^{min}\right|,$$

$$\Delta x_i^{MAX} = \lfloor\Delta\ddot{x}_i^{MAX}/2\rfloor,$$

$$\Delta x_i^{min} = \lfloor\Delta\ddot{x}_i^{min}/2\rfloor$$

and extract

$$s_{2i-1} = \Delta\ddot{x}_i^{MAX} \bmod 2,$$

$$s_{2i} = \Delta\ddot{x}_i^{min} \bmod 2.$$

Then calculate

$$\begin{pmatrix}\bar{x}_{2i-1}^{MAX}\\ \bar{x}_{2i}^{MAX}\end{pmatrix} = \begin{cases}\begin{pmatrix}\bar{\bar{x}}_{2i-1}^{MAX} - \lfloor\Delta x_i^{MAX}/2\rfloor - s_{2i-1}\\ \bar{\bar{x}}_{2i}^{MAX} + \lfloor(\Delta x_i^{MAX}+1)/2\rfloor\end{pmatrix} & \text{if } \bar{\bar{x}}_{2i-1}^{MAX} > \bar{\bar{x}}_{2i}^{MAX},\\ \begin{pmatrix}\bar{\bar{x}}_{2i-1}^{MAX} + \lfloor(\Delta x_i^{MAX}+1)/2\rfloor\\ \bar{\bar{x}}_{2i}^{MAX} - \lfloor\Delta x_i^{MAX}/2\rfloor - s_{2i-1}\end{pmatrix} & \text{otherwise,}\end{cases}$$

$$\begin{pmatrix}\bar{x}_{2i-1}^{min}\\ \bar{x}_{2i}^{min}\end{pmatrix} = \begin{cases}\begin{pmatrix}\bar{\bar{x}}_{2i-1}^{min} - \lfloor\Delta x_i^{min}/2\rfloor - s_{2i}\\ \bar{\bar{x}}_{2i}^{min} + \lfloor(\Delta x_i^{min}+1)/2\rfloor\end{pmatrix} & \text{if } \bar{\bar{x}}_{2i-1}^{min} > \bar{\bar{x}}_{2i}^{min},\\ \begin{pmatrix}\bar{\bar{x}}_{2i-1}^{min} + \lfloor(\Delta x_i^{min}+1)/2\rfloor\\ \bar{\bar{x}}_{2i}^{min} - \lfloor\Delta x_i^{min}/2\rfloor - s_{2i}\end{pmatrix} & \text{otherwise,}\end{cases}$$

and recover the pair of blocks to

$$\hat{B}_{2i-1} = (\bar{x}_{2i-1}^{MAX}, \bar{x}_{2i-1}^{min}, B_{2i-1}) \text{ and}$$

$$\hat{B}_{2i} = (\bar{x}_{2i}^{MAX}, \bar{x}_{2i}^{min}, B_{2i}).$$

4) Obtain the original BTC-compressed image T.

C. An example illustrating the proposed scheme

This section gives an example to illustrate the proposed scheme. Figure 1(a) is a BTC-compressed cover image with 8 blocks. Since B_{2i-1} and B_{2i} are binary arrays and they remain unchanged during the embedding procedure, their contents would not be shown in the example for simplicity. Let the message to be embedded be a character “A” and it’s ASCII code is $S = (41)_{16} = (00100001)_2$.

The encoder calculates $\Delta x_1^{MAX} = |\bar{x}_1^{MAX} - \bar{x}_2^{MAX}| = |135 - 134| = 1$, $\Delta x_1^{min} = \left|\bar{x}_1^{min} - \bar{x}_2^{min}\right| = |130 - 132| = 2$, and

$$\begin{pmatrix}\bar{\bar{x}}_1^{MAX}\\ \bar{\bar{x}}_2^{MAX}\end{pmatrix} = \begin{pmatrix}\bar{x}_1^{MAX} + \lfloor\Delta x_1^{MAX}/2\rfloor + s_1\\ \bar{x}_2^{MAX} - \lfloor(\Delta x_2^{MAX}+1)/2\rfloor\end{pmatrix} = \begin{pmatrix}135 + \lfloor 1/2\rfloor + 0\\ 134 - \lfloor(2+1)/2\rfloor\end{pmatrix} = \begin{pmatrix}135\\ 133\end{pmatrix},$$

$$\begin{pmatrix}\bar{\bar{x}}_1^{min}\\ \bar{\bar{x}}_2^{min}\end{pmatrix} = \begin{pmatrix}\bar{x}_1^{min} - \lfloor(\Delta x_1^{min}+1)/2\rfloor\\ \bar{x}_2^{min} + \lfloor\Delta x_2^{min}/2\rfloor + s_2\end{pmatrix} = \begin{pmatrix}130 - \lfloor(2+1)/2\rfloor\\ 132 + \lfloor 2/2\rfloor + 0\end{pmatrix} = \begin{pmatrix}129\\ 133\end{pmatrix}.$$

Since $0 \le \bar{\bar{x}}_1^{min}, \bar{\bar{x}}_1^{MAX}, \bar{\bar{x}}_2^{min}, \bar{\bar{x}}_2^{MAX} \le 255$, we have stego-blocks $\ddot{B}_1 = (135, 129, B_1)$ and $\ddot{B}_2 = (133, 133, B_2)$ as shown in Figure 1(b). Blocks 3 and 4 embed a bit of 1 and 0 in the larger and smaller block-mean pixels, respectively. Their embedding results would be $\ddot{B}_3 = (137, 130, B_3)$ and $\ddot{B}_4 = (134, 134, B_4)$. The embedding results of remaining four blocks are shown in Figure 1(b).

i	$\hat{B}_{2i-1}$	$\hat{B}_{2i}$
1	(135,130, B_1)	(134,132, B_2)
2	(136,131, B_3)	(135,133, B_4)
3	(134,132, B_5)	(135,131, B_6)
4	(133,130, B_7)	(135,130, B_8)

(a) Cover image

i	$\ddot{B}_{2i-1}$	$\ddot{B}_{2i}$
1	(135,129, B_1)	(133,133, B_2)
2	(137,130, B_3)	(134,134, B_4)
3	(133,132, B_5)	(135,130, B_6)
4	(132,130, B_7)	(136,131, B_8)

(b) Stego-image

Fig. 1. An embedding example

To extract the embedded message from the stego-image in Figure 1(b), the decoder scans the image as the order in the embedding procedure and calculates

$\Delta\ddot{x}_1^{MAX} = |\bar{\bar{x}}_1^{MAX} - \bar{\bar{x}}_2^{MAX}| = |135 - 133| = 2$,

$\Delta\ddot{x}_1^{min} = \left|\bar{\bar{x}}_1^{min} - \bar{\bar{x}}_2^{min}\right| = |129 - 133| = 4$,

$\Delta x_1^{MAX} = \lfloor \Delta\ddot{x}_1^{MAX}/2 \rfloor = 1$,

$\Delta x_1^{min} = \left\lfloor \Delta\ddot{x}_1^{min}/2 \right\rfloor = 2$ and

extracts

$s_1 = \Delta\ddot{x}_1^{MAX} \bmod 2 = 2 \bmod 2 = 0$,

$s_2 = \Delta\ddot{x}_1^{min} \bmod 2 = 4 \bmod 2 = 0$.

Then calculate

$$\begin{pmatrix} \bar{x}_1^{MAX} \\ \bar{x}_2^{MAX} \end{pmatrix} = \begin{pmatrix} \bar{\bar{x}}_1^{MAX} - \lfloor \Delta x_1^{MAX}/2 \rfloor - s_1 \\ \bar{\bar{x}}_2^{MAX} + \lfloor (\Delta x_1^{MAX} + 1)/2 \rfloor \end{pmatrix}$$
$$= \begin{pmatrix} 135 - \lfloor 1/2 \rfloor - 0 \\ 133 + \lfloor (1+1)/2 \rfloor \end{pmatrix}$$
$$= \begin{pmatrix} 135 \\ 134 \end{pmatrix}, \text{ and}$$
$$\begin{pmatrix} \bar{x}_1^{min} \\ \bar{x}_2^{min} \end{pmatrix} = \begin{pmatrix} \bar{\bar{x}}_1^{min} + \left\lfloor (\Delta x_1^{min} + 1)/2 \right\rfloor \\ \bar{\bar{x}}_2^{min} - \left\lfloor \Delta x_1^{min}/2 \right\rfloor - s_2 \end{pmatrix}$$
$$= \begin{pmatrix} 129 + \lfloor (2+1)/2 \rfloor \\ 133 - \lfloor 2/2 \rfloor - 0 \end{pmatrix}$$
$$= \begin{pmatrix} 130 \\ 132 \end{pmatrix}.$$

Finally, the pair of blocks are recovered to $\hat{B}_1 = (135, 130, B_1)$ and $\hat{B}_2 = (134, 132, B_2)$. A bit of 1 is extracted from $s_3 = \Delta\ddot{x}_3^{MAX} \bmod 2 = |\bar{\bar{x}}_3^{MAX} - \bar{\bar{x}}_4^{MAX}| \bmod 2 = |137 - 134| \bmod 2 = 1$. Similarly, a bit of 0 is extracted from $s_4 = |130 - 134| \bmod 2 = 0$. The process continues until the remaining messages are extracted and stego-blocks are recovered. Note that, in the example, none of block index is recorded in the overhead information.

IV. Experimental Results

To show the feasibility and performance of the proposed scheme, we implemented the proposed scheme on a personal computer with Java. The implementation included compressing a grayscale image into a BTC-format image as that in Section II. Test images are shown in Figure 2 and their dimension is 512×512. The block size of BTC-format image is 4×4 pixels. First, a randomly generated message, a binary bit string, was generated and embedded into a test cover image, i.e. the BTC-compressed image. Then we extracted the embedded message from the stego-image and recovered the stego-image to its cover image. The experimental results show that the extracted message is exactly the same as the embedded message and the cover image can be completely recovered. This means our proposed scheme can reversibly embed a message into a BTC-compressed image.

The peak signal noise ratio (PSNR) was used to evaluate the performance of the proposed scheme. It was defined as follows,

$$\text{PSNR} = 10\log_{10}\frac{255^2}{MSE}\text{dB},$$

where MSE is the mean square error. For an image with N pixels, MSE is computed as

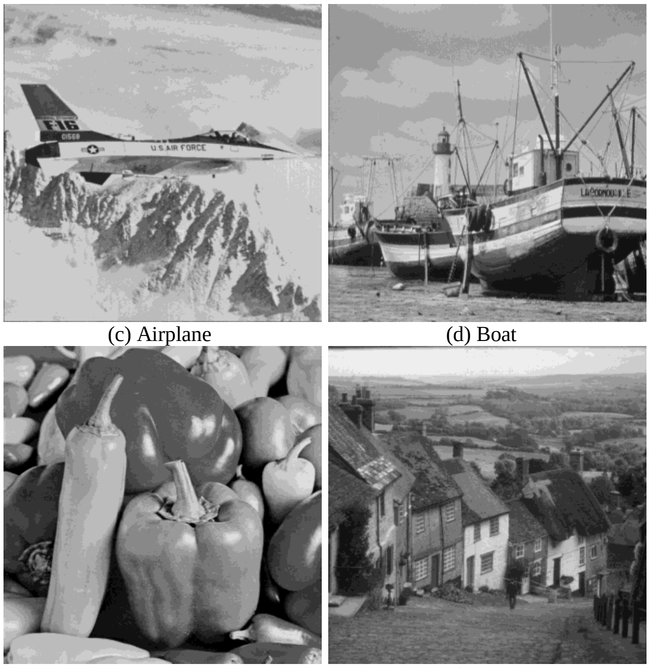
$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - x_i')^2,$$

where x_i and x_i' are cover and stego-pixels, respectively.

A larger PSNR implies that a stego-image is more similar to its cover image than a smaller one. It also implies that the visual quality of a stego-image with a higher PSNR is better than that with a smaller one. Researchers usually would like to get an embedding scheme which can obtain a higher PSNR.

Table II shows the visual quality, in terms of PSNR, of a stego-image applying the proposed scheme. When a message is embedded into an image, the image may be distorted by the modification of pixels. The more the messages are embedded, the more the image will be distorted. A good embedding scheme may provide enough embedding space and keep image quality as high as possible. Table II shows that payload capacity is approximately equal to the number of blocks, which means most blocks may embed a message. In addition, the image quality, i.e. PSNR, is more than 28 dB. This shows that a stego-image is similar to its cover image and they may not be distinguishable by human vision.

(a) Lena (b) Baboon



(c) Airplane (d) Boat

(e) Pepper (f) Gold

Fig. 2. Test images

TABLE II. EMBEDDING PERFORMANCE OF THE PROPOSED SCHEME

Images	PSNR(dB)	Payload(bits)
Lena	28.25	16,272
Baboon	28.03	16,366
Airplane	28.38	16,356
Boat	28.67	16,308
Pepper	29.42	16,016
Gold	29.72	16,368

V. CONCLUSIONS

A reversible data hiding scheme for BTC-compressed image has been proposed. In the proposed scheme, an image is divided into non-overlapping blocks and the BTC algorithm is applied to compress the image. Then a message with two bits is embedded into two blocks by expanding the difference between larger block-mean pair and the one between the smaller block-mean pair. The original BTC-format image may be completely reconstructed after the embedded message is extracted. Experimental results show that the proposed scheme may obtain a stego-image with high visual quality and a payload capacity of one bit per block, approximately. The proposed scheme is a good encoder for applications which need a reversible embedding scheme for BTC-compressed images without complicated computations.

REFERENCES

[1] G. K. Wallace, "The JPEG still picture compression standard," IEEE Transactions on Consumer Electronics, 38(1), pp. xviii–xxxiv, 1992.

[2] R. Gray, "Vector quantization," IEEE ASSP Magazine, 1(2), pp.4–29, 1984.

[3] E. Delp and O. Mitchell, "Image compression using block truncation coding," IEEE Transactions on Communications, 27(9), pp. 1335–1342, 1979.

[4] J. Chen, W. Hong, T.-S. Chen, and C.-W. Shiu, "Steganography for BTC compressed images using no distortion technique," The Imaging Science Journal, 58(4), pp. 177–185, 2010.

[5] J.-M. Guo and Y.-F. Liu, "High capacity data hiding for error-diffused block truncation coding," IEEE Transactions on Image Processing, 21(12), pp. 4808–4818, 2012.

[6] D. Ou and W. Sun, "High payload image steganography with minimum distortion based on absolute moment block truncation coding," Multimedia Tools and Applications, 74(21), pp 9117–9139, 2015.

[7] Y.-C. Chou and H.-H. Chang, "A high payload data hiding scheme for color image based on BTC compression technique," 2010 Fourth International Conference on Genetic and Evolutionary Computing (ICGEC), pp. 626–629, 2010.

[8] H. Luo, Z. Zhao, and Z.-M. Lu, "Joint secret sharing and data hiding for block truncation coding compressed image transmission," Information Technology Journal, 10(3), pp.681–685, 2011.

[9] C.-C. Chang, Y.-H. Chen, and C.-C. Lin, "A data embedding scheme for color images based on genetic algorithm and absolute moment block truncation coding," Soft Computing, 13(4), pp 321–331, 2009.

[10] J. Tian, "Reversible data embedding using a difference expansion," IEEE Transactions on Circuits Systems for Video Technolgy, 13 (8), pp. 890–896, 2003.

[11] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," IEEE Transactions on Circuits and Systems for Video Technology, 16(3), pp. 354–362, 2006.

[12] C.-C. Lin and N.-L. Hsueh, "A lossless data hiding scheme based on three-pixel block differences," Pattern Recognition, 41(4), pp. 1415–1425, 2008.

[13] C.-C. Chang, C.-C. Lin, C.-S. Tseng, and W.-L. Tai, "Reversible hiding in DCT-based compressed images," Information Sciences, 177(13), pp. 2768–2786, 2007.

Delay-Decomposition Stability Approach of Nonlinear Neutral Systems with Mixed Time-Varying Delays

Ilyes MAZHOUD
University of Tuins El Manar, National Engineering School of Tunis, Laboratory of research in Automatic Control, BP 37, Belvédère, 1002 Tunis, Tunisia

Issam AMRI
University of Tuins El Manar, National Engineering School of Tunis, Laboratory of research in Automatic Control, BP 37, Belvédère, 1002 Tunis, Tunisia

Dhaou SOUDANI
University of Tuins El Manar, National Engineering School of Tunis, Laboratory of research in Automatic Control, BP 37, Belvédère, 1002 Tunis, Tunisia

***Abstract*—This paper deals with the asymptotic stability of neutral systems with mixed time-varying delays and nonlinear perturbations. Based on the Lyapunov–Krasovskii functional including the triple integral terms and free weighting matrices approach, a novel delay-decomposition stability criterion is obtained. The main idea of the proposed method is to divide each delay interval into two equal segments. Then, the Lyapunov–Krasovskii functional is used to split the bounds of integral terms of each subinterval. In order to reduce the stability criterion conservatism, delay-dependent sufficient conditions are performed in terms of Linear Matrix Inequalities (LMIs) technique. Finally, numerical simulations are given to show the effectiveness of the proposed stability approach.**

Keywords—Neutral systems; Lyapunov–Krasovskii approach; asymptotic stability; mixed time-varying delays; nonlinear perturbations; Linear Matrix Inequalities (LMIs)

I. Introduction

Neutral time-delay appears in many fields of sciences and engineering, including neural networks, industrial, economy, chemical processes and population models. In fact, the presence of time-delay causes the instability, the oscillation, and performances' degradation of dynamical systems. Neutral systems are a part of a specific class of infinite dimensions. Their stability study can be a complex issue. Recently, the stability problem of neutral systems has been the subject of considerable research [1-23]. Thus, several approaches of delay-dependent stability criteria have been developed for this problem.

The stability criteria of neutral systems with mixed time-varying delays can be classified into two concepts. Firstly, the delay-dependent stability which is based on the size of time-delay and it gives the upper bound of delay in the formulation. Secondly, the delay-independent stability class doesn't include any information about the size of the time-delay. Indeed, the delay-dependent is often less conservative than the delay-independent.

In order to reduce the conservatism, many researchers studied the nonlinear neutral systems stability with mixed time-varying delays such as in [1] where authors consider the delay-dependent robust stability of uncertain neutral systems with mixed time-varying delays. In [2], I. Amri et al. have been studied a delay-dependent exponential stability condition for nonlinear neutral systems with mixed delays. They employ a delay-decomposition approach and the known free weighting matrices method.In [3], novel delay-decomposition condition of neutral systems with time-varying delays is proposed and new stability results were derived. In [4], the authors have been presented a new asymptotic stability results for nonlinear neutral system with mixed delays by using the delay-dividing approach. In [5], the exponential stability of neutral delay differential systems with nonlinear uncertainties is used. The problem of the delay-dependent robust stability criteria for neutral systems with mixed time-varying delays and nonlinear perturbations has been studied in [6]. In [7], new less conservative robust stability criteria of neutral systems with mixed time-varying delays and nonlinear perturbations are derived by using the delay method.

In this paper, the problem of asymptotical delay-decomposition stability for nonlinear neutral systems with mixed time-varying delays is investigated. By using a new augmented Lyapunov–Krasovskii functional including the triple integral terms for interval time-varying delays as well as the free-weighting matrices technique and Jensen integral inequality, new sufficient delay-dependent stability conditions have been proposed and expressed in terms of LMIs. These stability conditions can be easily solved by various convex optimization algorithms.

The remainder of this paper is organized as follows. In Section 2, the stability problem of nonlinear neutral systems is described. Some related preliminaries are also given. The main result of this paper is presented in Section 3. Numerical examples are carried out in Section 4 in order to illustrate the proposed results. Section 5 concludes this paper.

II. Problem Description and Preliminaries

This paper considers the nonlinear neutral systems with mixed time-varying delays of Equation (1):

$$\begin{cases} \dot{x}(t) - A_2\,\dot{x}(t-\tau(t)) = A\,x(t) + A_1\,x(t-h(t)) \\ \qquad\qquad + f_0(t, x(t)) \\ \qquad\qquad + f_1(t, x(t-h(t))) \\ \qquad\qquad + f_2(t, \dot{x}(t-\tau(t))) \\ x(t) = \varphi(t),\ \dot{x}(t) = \phi(t),\ \forall t \in \left[-\max\left\{\tau_M, h_M\right\}, 0\right] \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector $A, A_1, A_2 \in \mathbb{R}^{n\times n}$ are constant matrices with appropriate dimensions. $\tau(t),\ h(t)$ are neutral and discrete time-varying delays satisfying the following equations:

$$0 < h_m \le h(t) \le h_M, \quad \dot{h}(t) \le \mu < 1, \tag{2}$$

$$0 < \tau_m \le \tau(t) \le \tau_M, \quad \dot{\tau}(t) \le \eta < 1, \tag{3}$$

The initial conditions functions $\varphi(t)$, $\phi(t)$ are continuously differentiable on $[-\max\{\tau_M, h_M\}, 0]$. The functions $f_0(t, x(t))$,

$f_1(t, x(t-h(t)))$ and $f_2(t, \dot{x}(t-\tau(t)))$ are unknown nonlinear uncertainties satisfying $f_0(t,0)=0,\ f_1(t,0)=0,\ f_2(t,0)=0$ and

$$\begin{cases} \| f_0(t, x(t))\| \le \beta_0 \|x(t)\|, \\ \| f_1(t, x(t-h(t)))\| \le \beta_1 \|x(t-h(t))\|, \\ \| f_2(t, \dot{x}(t-\tau(t)))\| \le \beta_2 \|\dot{x}(t-\tau(t))\|, \end{cases} \tag{4}$$

where $\beta_0 \ge 0, \beta_1 \ge 0$, $\beta_2 \ge 0$ are given constants.

Constraint (4) can be rewritten as follows:

$$\begin{cases} f_0^T(t,x(t))\ f_0(t,x(t)) \le \beta_0^2\, x^T(t)\ x(t) \\ f_1^T(t,x(t-h(t)))\ f_1(t,x(t-h(t))) \le \beta_1^2\, x^T(t-h(t))\ x(t-h(t)) \\ f_2^T(t,\dot{x}(t-\tau(t)))\ f_2(t,\dot{x}(t-\tau(t))) \le \beta_2^2\, \dot{x}^T(t-\tau(t))\ \dot{x}(t-\tau(t)) \end{cases} \tag{5}$$

For simplicity, note that:

$$f_0 := f_0(t,\, x(t))),\ \ f_1 := f_1(t,\, x(t-h(t))),\ \ f_2 := f_2(t,\, \dot{x}(t-\tau(t)))$$

Moreover, for dividing the each interval time-varying delay into two equal subintervals $\left[h_m, \alpha_1 h_M\right]$ $or\left[\alpha_1 h_M, h_M\right]$ and $\left[\tau_m, \alpha_2\tau_M\right]$ $or\left[\alpha_2\tau_M, \tau_M\right]$, two different cases for time-varying delays have been presented.

Case I: $h(t), \tau(t)$ are differentiable functions, satisfying for all $t \ge 0$:

$$\begin{aligned} &h_m \le h(t) \le h_M \quad \text{and} \ \ \dot{h}(t) \le \mu < 1, \\ &\tau_m \le \tau(t) \le \tau_M \quad \text{and} \ \ \dot{\tau}(t) \le \eta < 1. \end{aligned} \tag{6}$$

Case II: $h(t)$ is not differentiable or the upper bound of the derivative of $h(t)$ and $\tau(t)$ is a differentiable function, and $h(t), \tau(t)$ satisfying:

$$\begin{aligned} &h_m \le h(t) \le h_M\ , \\ &\tau_m \le \tau(t) \le \tau_M,\ \ \dot{\tau}(t) \le \eta < 1. \end{aligned} \tag{7}$$

where $h_m, h_M, \tau_m, \tau_M, d$ and μ are positive scalars.

This paper is devoted to investigate the delay-dependent stability analysis of time-varying delays system (1) satisfying (2) and (3) equations and under nonlinear perturbations inequalities (4) and (5). It aims to formulate a less conservative stability technique to estimate the upper bound for the delay interval. Before deriving the proposed stability criteria, the following lemmas are needed.

Lemma 1. [8]

For any constant matrix $R \in \mathbb{R}^{n\times n}, R = R^T > 0,$ a scalar function $h := h(t) > 0,$ and a vector valued function $\dot{x} : [-h, 0] \to \mathbb{R}^n$ such that the following integrations are well defined, then:

$$-h \int_{t-h}^{t} \dot{x}^T(s)\, R\, \dot{x}(s)\, ds \le \psi_1^T(t) \begin{bmatrix} -R & R \\ R & -R \end{bmatrix} \psi_1(t)$$

$$-\frac{h^2}{2} \int_{-h}^{0}\int_{t+\theta}^{t} \dot{x}^T(s)\, R\, \dot{x}(s)\, ds\, d\theta \le \psi_2^T(t) \begin{bmatrix} -R & R \\ R & -R \end{bmatrix} \psi_2(t)$$

where $\psi_1^T(t) = \left[x^T(t)\ x^T(t-h)\right]$ and $\psi_2^T(t) = \left[h\, x^T(t) \int_{t-h}^{t} x^T(s)\, ds\right].$

Lemma 2. [9]: The following matrix inequality

$$\begin{pmatrix} Q(x) & S(x) \\ S^T(x) & R(x) \end{pmatrix} < 0,$$

where $Q(x) = Q^T(x)$, $R(x) = R^T(x)$ and $S(x)$ depend on affine on x, is equivalent to $R(x) < 0$, $Q(x) < 0$ and $Q(x) - S(x) R^{-1}(x) S^T(x) < 0.$

Lemma 3. [10]: For any scalar $\tau(t) \ge 0$ and any constant matrix $R \in \mathbb{R}^{n\times n}, R = R^T > 0,$ the following inequality holds:

$$-\int_{t-\alpha_1 h_M}^{t-h(t)} \dot{x}^T(s)\, R\, \dot{x}(s)\, ds \le (\alpha_1 h_M - h(t))\, \xi^T(t)\, F\, R^{-1}\, F^T \xi(t)$$

$$+ 2\,\xi^T(t)\, F\left[x(t-h(t)) - x(t-\alpha_1 h_M)\right],$$

where

$$\xi^T(t) = \begin{bmatrix} x^T(t) & x^T(t-h_m) & x^T(t-h(t)) & x^T(t-\alpha_1 h_M) & x^T(t-\tau_m) & x^T(t-\tau(t)) & x^T(t-\alpha_2\tau_M) \\ \dot{x}^T(t) & \dot{x}^T(t-\tau(t)) & \left(\int_{t-\alpha_2\tau_M}^{t} x(s)\, ds\right)^T & \left(\int_{t-\alpha_1 h_M}^{t-h_m} x(s)\, ds\right)^T & \left(\int_{t-\alpha_2\tau_M}^{t-\tau_m} x(s)\, ds\right)^T & f_0^T\ f_1^T & f_2^T \end{bmatrix}$$

and F is free-weighting matrix with appropriate dimensions.

III. Main Results

In order to obtain some less conservative conditions, new delay-decomposition method for nonlinear neutral system (1) is developed. The first delay-interval $\left[h_m, h_M\right]$ is divided into two segments $\left[h_m, \alpha_1 h_M\right]$ and $\left[\alpha_1 h_M, h_M\right]$. The second delay-interval $\left[\tau_m, \tau_M\right]$ is decomposed into two subintervals $\left[\tau_m, \alpha_2\tau_M\right]$ and $\left[\alpha_2\tau_M, \tau_M\right]$. The following theorem presents new stability criteria for interval time-varying delay system (1).

Theorem1. In Case I, if $h_m \le h(t) \le \alpha_1 h_M \ (0<\alpha_1<1)$ and $\tau_m \le \tau(t) \le \alpha_2 \tau_M \ (0<\alpha_2<1)$, for given positive scalars $h_m, h_M, \tau_m, \tau_M, \eta, \mu, \beta_0, \beta_1$ and β_2, the system (1) with uncertainty (5) and mixed time-varying delays satisfying(2) and (3) is asymptotically stable if there exist symmetric positive definite $n\times n$ matrices $P, Q_i(i=1,..,7), R_j(j=1,..,7)$, for any free matrix variables $T_a, Y_a, W_a, N_a, X_a, F_a\ (a=1,2)$ and scalars $\varepsilon_i \ge 0\ (i=0,1,2)$ such that the following symmetric LMI holds:

$$\begin{bmatrix} \Omega & \sqrt{\alpha_1 h_M - h_m}\,T & \sqrt{\alpha_1 h_M}\,Y & \sqrt{\alpha_1 h_M - h_m}\,W & \sqrt{\alpha_2\tau_M-\tau_m}\,N & \sqrt{\alpha_2\tau_M}\,X \\ * & -(R_1+R_2) & 0 & 0 & 0 & 0 \\ * & * & -R_1 & 0 & 0 & 0 \\ * & * & * & -R_2 & 0 & 0 \\ * & * & * & * & -R_3 & 0 \\ * & * & * & * & * & -R_3 \end{bmatrix} < 0, \tag{8}$$

where $\Omega = (\Omega_{i,j})_{15\times 15}$ and:

$\Omega_{1,1} = Q_1+Q_2+Q_3+Q_4+Q_6+Q_7+Y_1+Y_1^T+X_1+X_1^T$
$\quad -2R_5-2R_6-2R_7+F_1A+A^TF_1^T+\varepsilon_0\beta_0^2\mathrm{I}$

$\Omega_{1,3} = -Y_1+Y_2^T+F_1A_1$

$\Omega_{1,6} = -X_1+X_2^T$

$\Omega_{1,8} = P-F_1+A^TF_2^T$

$\Omega_{1,9} = F_1A_2$

$\Omega_{1,10} = \dfrac{2}{\alpha_2\tau_M}R_5$

$\Omega_{1,11} = \dfrac{2}{(\alpha_1 h_M + h_m)}R_6$

$\Omega_{1,12} = \dfrac{2}{(\alpha_2\tau_M+\tau_m)}R_7$

$\Omega_{1,13} = \Omega_{1,14} = \Omega_{1,15} = F_1$

$\Omega_{2,2} = -Q_3+W_1+W_1^T$

$\Omega_{2,3} = -W_1+W_2^T$

$\Omega_{3,3} = -(1-\mu)Q_2+T_1+T_1^T-Y_2-Y_2^T-W_2-W_2^T+\varepsilon_1\beta_1^2\mathrm{I}$

$\Omega_{3,4} = -T_1+T_2^T$

$\Omega_{3,8} = A_1^TF_2^T$

$\Omega_{4,4} = -Q_1-T_2-T_2^T$

$\Omega_{5,5} = -R_4-Q_7$

$\Omega_{5,7} = R_4$

$\Omega_{6,6} = -(1-\eta)Q_4+N_1+N_1^T-X_2-X_2^T$

$\Omega_{7,7} = -R_4-Q_6-N_2-N_2^T$

$\Omega_{8,8} = M-F_2-F_2^T$

$\Omega_{8,9} = F_2A_2$

$\Omega_{8,13} = \Omega_{8,14} = \Omega_{8,15} = F_2$

$\Omega_{9,9} = -(1-\eta)Q_5+\varepsilon_2\beta_2^2\mathrm{I}$

$\Omega_{10,10} = -\dfrac{2}{\alpha_2^2\tau_M^2}R_5$

$\Omega_{11,11} = -\dfrac{2}{(\alpha_1^2h_M^2-h_m^2)}R_6$

$\Omega_{12,12} = -\dfrac{2}{(\alpha_2^2\tau_M^2-\tau_m^2)}R_7$

$\Omega_{13,13} = -\varepsilon_0\mathrm{I}$

$\Omega_{14,14} = -\varepsilon_1\mathrm{I}$

$\Omega_{15,15} = -\varepsilon_2\mathrm{I}$

$M = Q_5+\alpha_1h_MR_1+(\alpha_1h_M-h_m)R_2+\alpha_2\tau_MR_3+(\alpha_2\tau_M-\tau_m)^2R_4$
$\quad +\dfrac{\alpha_2^2\tau_M^2}{2}R_5+\dfrac{(\alpha_1^2h_M^2-h_m^2)}{2}R_6+\dfrac{(\alpha_2^2\tau_M^2-\tau_m^2)}{2}R_7$

Proof. Choose a new augmented Lyapunov–Krasovskii functional as:

$$V(t) = V_1(t)+V_2(t)+V_3(t)+V_4(t) \tag{9}$$

where

$V_1(t) = x^T(t)Px(t),$

$$V_2(t) = \int_{t-\alpha_1h_M}^{t} x^T(s)Q_1x(s)ds + \int_{t-h(t)}^{t} x^T(s)Q_2x(s)ds + \int_{t-h_m}^{t} x^T(s)Q_3x(s)ds$$
$$+\int_{t-\tau(t)}^{t} x^T(s)Q_4x(s)ds + \int_{t-\tau(t)}^{t} \dot{x}^T(s)Q_5\dot{x}(s)ds + \int_{t-\alpha_2\tau_M}^{t} x^T(s)Q_6x(s)ds$$
$$+\int_{t-\tau_m}^{t} x^T(s)Q_7x(s)ds,$$

$$V_3(t) = \int_{-\alpha_1h_M}^{0}\int_{t+\theta}^{t} \dot{x}^T(s)R_1\dot{x}(s)ds\,d\theta + \int_{-\alpha_1h_M}^{-h_m}\int_{t+\theta}^{t} \dot{x}^T(s)R_2\dot{x}(s)ds\,d\theta$$
$$+\int_{-\alpha_2\tau_M}^{0}\int_{t+\theta}^{t} \dot{x}^T(s)R_3\dot{x}(s)ds\,d\theta + (\alpha_2\tau_M-\tau_m)\int_{-\alpha_2\tau_M}^{-\tau_m}\int_{t+\theta}^{t} \dot{x}^T(s)R_4\dot{x}(s)ds\,d\theta,$$

$$V_4(t) = \int_{-\alpha_2\tau_M}^{0}\int_{\theta}^{0}\int_{t+\lambda}^{t} \dot{x}^T(s)R_5\dot{x}(s)ds\,d\lambda\,d\theta + \int_{-\alpha_1h_M}^{-h_m}\int_{\theta}^{0}\int_{t+\lambda}^{t} \dot{x}^T(s)R_6\dot{x}(s)ds\,d\lambda\,d\theta$$
$$+\int_{-\alpha_2\tau_M}^{-\tau_m}\int_{\theta}^{0}\int_{t+\lambda}^{t} \dot{x}^T(s)R_7\dot{x}(s)ds\,d\lambda\,d\theta.$$

with

$$\xi^T(t)=\begin{bmatrix} x^T(t) & x^T(t-h_m) & x^T(t-h(t)) & x^T(t-\alpha_1 h_M) & x^T(t-\tau_m) & x^T(t-\tau(t)) & x^T(t-\alpha_2\tau_M) \\ \dot{x}^T(t) & \dot{x}^T(t-\tau(t)) & \left(\int\limits_{t-\alpha_2\tau_M}^{t} x(s)\,ds\right)^T & \left(\int\limits_{t-\alpha_1 h_M}^{t-h_m} x(s)\,ds\right)^T & \left(\int\limits_{t-\alpha_2\tau_M}^{t-\tau_m} x(s)\,ds\right)^T & f_0^T \;\; f_1^T \;\; f_2^T & \end{bmatrix}$$

Then, the time derivative of $V(t)$ along the trajectory of system (1) is given by:

$$\dot{V}(t)=\dot{V}_1(t)+\dot{V}_2(t)+\dot{V}_3(t)+\dot{V}_4(t) \qquad (10)$$

where

$$\dot{V}_1(t)=\dot{x}^T(t)P\,x(t)+x^T(t)P\,\dot{x}(t),$$

$$\begin{aligned}\dot{V}_2(t)&=x^T(t)\left(Q_1+Q_2+Q_3+Q_4+Q_6+Q_7\right)x(t)\\ &\quad -x^T(t-\alpha_1 h_M)Q_1\,x(t-\alpha_1 h_M)-(1-\dot{h}(t))\,x^T(t-h(t))Q_2\,x(t-h(t))\\ &\quad -x^T(t-h_m)Q_3\,x(t-h_m)-(1-\dot{\tau}(t))\,x^T(t-\tau(t))Q_4\,x(t-\tau(t))\\ &\quad -(1-\dot{\tau}(t))\,\dot{x}^T(t-\tau(t))Q_5\,\dot{x}(t-\tau(t))+\dot{x}^T(t)Q_5\dot{x}(t)\\ &\quad -x^T(t-\alpha_2\tau_M)Q_6\,x(t-\alpha_2\tau_M)-x^T(t-\tau_m)Q_7\,x(t-\tau_m)\end{aligned}$$

$$\begin{aligned}\dot{V}_3(t)&=\dot{x}^T(t)((\alpha_1 h_M)R_1+(\alpha_1 h_M-h_m)R_2+(\alpha_2\tau_M)R_3\\ &\quad +(\alpha_2\tau_M-\tau_m)^2R_4)\dot{x}(t)-\int\limits_{t-\alpha_1 h_M}^{t}\dot{x}^T(s)R_1\,\dot{x}(s)\,ds\\ &\quad -\int\limits_{t-\alpha_1 h_M}^{t-h_m}\dot{x}^T(s)R_2\,\dot{x}(s)\,ds-\int\limits_{t-\alpha_2\tau_M}^{t}\dot{x}^T(s)R_3\,\dot{x}(s)\,ds\\ &\quad -(\alpha_2\tau_M-\tau_m)\int\limits_{t-\alpha_2\tau_M}^{t-\tau_m}\dot{x}^T(s)R_4\,\dot{x}(s)\,ds,\end{aligned}$$

and:

$$\begin{aligned}\dot{V}_4(t)&=\dot{x}^T(t)\Big(\frac{\alpha_2^2\tau_M^2}{2}R_5+\frac{(\alpha_1^2h_M^2-h_m^2)}{2}R_6\\ &\quad +\frac{(\alpha_2^2\tau_M^2-\tau_m^2)}{2}R_7\Big)\dot{x}(t)-\int\limits_{-\alpha_2\tau_M}^{0}\int\limits_{t+\theta}^{t}\dot{x}^T(s)R_5\,\dot{x}(s)\,ds\,d\theta\\ &\quad -\int\limits_{-\alpha_1 h_M}^{-h_m}\int\limits_{t+\theta}^{t}\dot{x}^T(s)R_6\,\dot{x}(s)\,ds\,d\theta-\int\limits_{-\alpha_2\tau_M}^{-\tau_m}\int\limits_{t+\theta}^{t}\dot{x}^T(s)R_7\,\dot{x}(s)\,ds\,d\theta.\end{aligned}$$

The upper bound of the integral terms in inequality $\dot{V}_3(t)$ is estimated as:

$$\begin{aligned}&-\int\limits_{t-\alpha_1 h_M}^{t}\dot{x}^T(s)R_1\dot{x}(s)ds-\int\limits_{t-\alpha_1 h_M}^{t-h_m}\dot{x}^T(s)R_2\dot{x}(s)ds-\int\limits_{t-\alpha_2\tau_M}^{t}\dot{x}^T(s)R_3\dot{x}(s)ds\\ &-\left(\alpha_2\tau_M-\tau_m\right)\int\limits_{t-\alpha_2\tau_M}^{t-\tau_m}\dot{x}^T(s)R_4\dot{x}(s)ds\\ &=-\int\limits_{t-\alpha_1 h_M}^{t-h(t)}\dot{x}^T(s)R_1\dot{x}(s)ds-\int\limits_{t-h(t)}^{t}\dot{x}^T(s)R_1\dot{x}(s)ds-\int\limits_{t-\alpha_1 h_M}^{t-h(t)}\dot{x}^T(s)R_2\dot{x}(s)ds\\ &-\int\limits_{t-h(t)}^{t-h_m}\dot{x}^T(s)R_2\dot{x}(s)ds-\int\limits_{t-\alpha_2\tau_M}^{t-\tau(t)}\dot{x}^T(s)R_3\dot{x}(s)ds-\int\limits_{t-\tau(t)}^{t}\dot{x}^T(s)R_3\dot{x}(s)ds \qquad (11)\\ &-\left(\alpha_2\tau_M-\tau_m\right)\int\limits_{t-\alpha_2\tau_M}^{t-\tau_m}\dot{x}^T(s)R_4\dot{x}(s)ds\end{aligned}$$

Using Jensen's inequality, such that:

$$-\left(\alpha_2\tau_M-\tau_m\right)\int\limits_{t-\alpha_2\tau_M}^{t-\tau_m}\dot{x}^T(s)R_4\,\dot{x}(s)\,ds\le\begin{bmatrix}x(t-\tau_m)\\x(t-\alpha_2\tau_M)\end{bmatrix}^T\begin{bmatrix}-R_4 & R_4\\R_4 & -R_4\end{bmatrix}\times\begin{bmatrix}x(t-\tau_m)\\x(t-\alpha_2\tau_M)\end{bmatrix} \qquad (12)$$

$$-\int\limits_{-\alpha_2\tau_M}^{0}\int\limits_{t+\theta}^{t}\dot{x}^T(s)R_5\,\dot{x}(s)\,ds\,d\theta\le\frac{2}{\alpha_2^2\tau_M^2}\begin{bmatrix}\alpha_2\tau_M x(t)\\ \int\limits_{t-\alpha_2\tau_M}^{t}x(s)\,ds\end{bmatrix}^T\begin{bmatrix}-R_5 & R_5\\R_5 & -R_5\end{bmatrix}\begin{bmatrix}\alpha_2\tau_M x(t)\\ \int\limits_{t-\alpha_2\tau_M}^{t}x(s)\,ds\end{bmatrix} \qquad (13)$$

$$-\int\limits_{-\alpha_1 h_M}^{-h_m}\int\limits_{t+\theta}^{t}\dot{x}^T(s)R_6\,\dot{x}(s)\,ds\,d\theta\le\frac{2}{(\alpha_1^2h_M^2-h_m^2)}\begin{bmatrix}(\alpha_1 h_M-h_m)x(t)\\ \int\limits_{t-\alpha_1 h_M}^{t-h_m}x(s)\,ds\end{bmatrix}^T\begin{bmatrix}-R_6 & R_6\\R_6 & -R_6\end{bmatrix}\times\begin{bmatrix}(\alpha_1 h_M-h_m)x(t)\\ \int\limits_{t-\alpha_1 h_M}^{t-h_m}x(s)\,ds\end{bmatrix} \qquad (14)$$

$$-\int\limits_{-\alpha_2\tau_M}^{-\tau_m}\int\limits_{t+\theta}^{t}\dot{x}^T(s)R_7\,\dot{x}(s)\,ds\,d\theta\le\frac{2}{(\alpha_2^2\tau_M^2-\tau_m^2)}\begin{bmatrix}(\alpha_2\tau_M-\tau_m)x(t)\\ \int\limits_{t-\alpha_2\tau_M}^{t-\tau_m}x(s)\,ds\end{bmatrix}^T\begin{bmatrix}-R_7 & R_7\\R_7 & -R_7\end{bmatrix}\times\begin{bmatrix}(\alpha_2\tau_M-\tau_m)x(t)\\ \int\limits_{t-\alpha_2\tau_M}^{t-\tau_m}x(s)\,ds\end{bmatrix} \qquad (15)$$

By using Lemma 3, an upper bound of integral term of $\dot{V}(t)$ can be obtained as:

$$\begin{aligned}-\int\limits_{t-\alpha_1 h_M}^{t-h(t)}\dot{x}^T(s)(R_1+R_2)\,\dot{x}(s)\,ds&\le(\alpha_1 h_M-h(t))\,\xi^T(t)T\,(R_1+R_2)^{-1}T^T\xi(t)\\ &\quad +2\xi^T(t)T\left[x(t-h(t))-x(t-\alpha_1 h_M)\right],\end{aligned} \qquad (16)$$

$$\begin{aligned}-\int\limits_{t-h(t)}^{t}\dot{x}^T(s)\,R_1\,\dot{x}(s)\,ds&\le h(t)\,\xi^T(t)\,Y\,R_1^{-1}\,Y^T\xi(t)\\ &\quad +2\,\xi^T(t)\,Y\left[x(t)-x(t-h(t))\right],\end{aligned} \qquad (17)$$

$$\begin{aligned}-\int\limits_{t-h(t)}^{t-h_m}\dot{x}^T(s)\,R_2\,\dot{x}(s)\,ds&\le(h(t)-h_m)\,\xi^T(t)W\,R_2^{-1}W^T\xi(t)\\ &\quad +2\,\xi^T(t)W\left[x(t-h_m)-x(t-h(t))\right]\end{aligned} \qquad (18)$$

$$\begin{aligned}-\int\limits_{t-\alpha_2\tau_M}^{t-\tau(t)}\dot{x}^T(s)\,R_3\,\dot{x}(s)\,ds&\le(\alpha_2\tau_M-\tau(t))\,\xi^T(t)\,N\,R_3^{-1}\,N^T\xi(t)\\ &\quad +2\,\xi^T(t)\,N\left[x(t-\tau(t))-x(t-\alpha_2\tau_M)\right]\end{aligned} \qquad (19)$$

$$\begin{aligned}-\int\limits_{t-\tau(t)}^{t}\dot{x}^T(s)\,R_3\,\dot{x}(s)\,ds&\le\tau(t)\,\xi^T(t)\,X\,R_3^{-1}\,X^T\xi(t)\\ &\quad +2\,\xi^T(t)\,X\left[x(t)-x(t-\tau(t))\right]\end{aligned} \qquad (20)$$

For any matrices F_1, F_2 with appropriate dimensions, the following equation, from the system (1), verifies:

$$2\left[x^T(t)F_1 + \dot{x}^T(t)F_2\right]\left[Ax(t) + A_1x(t-h(t)) + A_2\dot{x}(t-\tau(t)) - \dot{x}(t) + f_0 + f_1 + f_2\right] = 0$$
(21)

Therefore, combining Equations (10) and (21) yields:

$$\dot{V}(t) \le \xi^T(t)\Omega\ \xi(t) \qquad (22)$$

with

$$\xi^T(t) = \begin{bmatrix} x^T(t)\ x^T(t-h_m)\ x^T(t-h(t))\ x^T(t-\alpha_1 h_M)\ x^T(t-\tau_m)\ x^T(t-\tau(t))\ x^T(t-\alpha_2\tau_M) \\ \dot{x}^T(t)\ \dot{x}^T(t-\tau(t)) \left(\int_{t-\alpha_2\tau_M}^{t} x(s)ds\right)^T \left(\int_{t-\alpha_1 h_M}^{t-h_m} x(s)ds\right)^T \left(\int_{t-\alpha_2\tau_M}^{t-\tau_m} x(s)ds\right)^T f_0^T\ f_1^T\ f_2^T \end{bmatrix}$$

and Ω is given in Equation (8).

By using the Schur Complement, it is clear to see that the results $\dot{V}(t) < 0$ holds if $\Omega < 0$, $h_m \le h(t) \le \alpha_1 h_M$ and $\tau_m \le \tau(t) \le \alpha_2\tau_M$ .Thus, the system (1) is asymptotically stable according to the Lyapunov-Krasovskii theory.

Remark 1: Inspired by the previous works [2-10], some Lyapunov-Krasovskii functional including triple integral terms involving lower and upper bounds of each interval time varying delays have been improved an important role in reduction of conservatism to estimate the maximum allowable delay bound.

Theorem 2. In Case I, if $\alpha_1 h_M \le h(t) \le h_M$ $(0 < \alpha_1 < 1)$ and $\alpha_2\tau_M \le \tau(t) \le \tau_M$ $(0 < \alpha_2 < 1)$, for given positive scalars $h_m,\ h_M,\ \tau_m, \tau_M, \eta, \mu, \beta_0, \beta_1$ and β_2, the system (1) with uncertainty (5) and mixed time-varying delays satisfying Equations (2) and (3) is asymptotically stable if there exist symmetric positive definite $n \times n$ matrices $P,\ Q_i\,(i = 1,..,7)$,

$R_j\,(j = 1,..,7)$, for any free matrix variables T_a, Y_a, W_a, X_a,

$F_a\ (a = 1,2)$ and scalars $\varepsilon_i \ge 0\ (i = 0,\ 1,\ 2)$ such that the following symmetric LMI holds:

$$\begin{bmatrix} \Pi & \sqrt{(1-\alpha_1)h_M}\,T & \sqrt{h_M}\,Y & \sqrt{(1-\alpha_1)h_M}\,W & \sqrt{\tau_M}\,X \\ * & -R_2 & 0 & 0 & 0 \\ * & * & -R_1 & 0 & 0 \\ * & * & * & -R_2 & 0 \\ * & * & * & * & -R_3 \end{bmatrix} < 0, \qquad (23)$$

where $\Pi = \left(\Pi_{i,j}\right)_{15\times15}$

with

$$\Pi_{1,1} = Q_1 + Q_2 + Q_3 + Q_4 + Q_6 + Q_7 + Y_1 + Y_1^T + X_1 + X_1^T - 2R_5 - 2R_6 - 2R_7 + F_1A + A^TF_1^T + \varepsilon_0\beta_0^2\mathrm{I}$$

$$\Pi_{1,3} = -Y_1 + Y_2^T + F_1A_1$$

$$\Pi_{1,6} = -X_1 + X_2^T$$

$$\Pi_{1,8} = P - F_1 + A^TF_2^T$$

$$\Pi_{1,9} = F_1A_2$$

$$\Pi_{1,10} = \frac{2}{\tau_M}R_5$$

$$\Pi_{1,11} = \frac{2}{(1+\alpha_1)h_M}R_6$$

$$\Pi_{1,12} = \frac{2}{(1+\alpha_2)\tau_M}R_7$$

$$\Pi_{1,13} = \Omega_{1,14} = \Omega_{1,15} = F_1$$

$$\Pi_{2,2} = -Q_1 + W_1 + W_1^T$$

$$\Pi_{2,3} = -W_1 + W_2^T$$

$$\Pi_{3,3} = -(1-\mu)Q_2 + T_1 + T_1^T - Y_2 - Y_2^T - W_2 - W_2^T + \varepsilon_1\beta_1^2\mathrm{I}$$

$$\Pi_{3,4} = -T_1 + T_2^T$$

$$\Pi_{3,8} = A_1^TF_2^T$$

$$\Pi_{4,4} = -Q_3 - T_2 - T_2^T$$

$$\Pi_{5,5} = -R_4 - Q_6$$

$$\Pi_{5,7} = R_4$$

$$\Pi_{6,6} = -(1-\eta)Q_4 - X_2 - X_2^T$$

$$\Pi_{7,7} = -R_4 - Q_7$$

$$\Pi_{8,8} = Z - F_2 - F_2^T$$

$$\Pi_{8,9} = F_2A_2$$

$$\Pi_{8,13} = \Omega_{8,14} = \Omega_{8,15} = F_2$$

$$\Pi_{9,9} = -(1-\eta)Q_5 + \varepsilon_2\beta_2^2\mathrm{I}$$

$$\Pi_{10,10} = -\frac{2}{\tau_M^2}R_5$$

$$\Pi_{11,11} = -\frac{2}{(1-\alpha_1^2)h_M^2}R_6$$

$$\Pi_{12,12} = -\frac{2}{(1-\alpha_2^2)\tau_M^2}R_7$$

$$\Pi_{13,13} = -\varepsilon_0\mathrm{I}$$

$$\Pi_{14,14} = -\varepsilon_1\mathrm{I}$$

$$\Pi_{15,15} = -\varepsilon_2\mathrm{I}$$

$$Z = Q_5 + \alpha_1 h_M R_1 + (1-\alpha_1) h_M R_2 + \alpha_2 \tau_M R_3 + (1-\alpha_2)^2 \tau_M^2 R_4 + \frac{\tau_M^2}{2} R_5 + \frac{(1-\alpha_1^2) h_M^2}{2} R_6 + \frac{(1-\alpha_2^2)\tau_M^2}{2} R_7$$

Proof. Choose a new augmented Lyapunov–Krasovskii functional as:

$$V(t) = V_1(t) + V_2(t) + V_3(t) + V_4(t) \tag{24}$$

where

$$V_1(t) = x^T(t) P x(t),$$

$$V_2(t) = \int_{t-\alpha_1 h_M}^{t} x^T(s) Q_1 x(s) ds + \int_{t-h(t)}^{t} x^T(s) Q_2 x(s) ds + \int_{t-h_M}^{t} x^T(s) Q_3 x(s) ds + \int_{t-\tau(t)}^{t} x^T(s) Q_4 x(s) ds + \int_{t-\tau(t)}^{t} \dot{x}^T(s) Q_5 \dot{x}(s) ds + \int_{t-\alpha_2\tau_M}^{t} x^T(s) Q_6 x(s) ds + \int_{t-\tau_M}^{t} x^T(s) Q_7 x(s) ds,$$

$$V_3(t) = \int_{-\alpha_1 h_M}^{0}\int_{t+\theta}^{t} \dot{x}^T(s) R_1 \dot{x}(s) ds\, d\theta + \int_{-h_M}^{-\alpha_1 h_M}\int_{t+\theta}^{t} \dot{x}^T(s) R_2 \dot{x}(s) ds\, d\theta + \int_{-\alpha_2\tau_M}^{0}\int_{t+\theta}^{t} \dot{x}^T(s) R_3 \dot{x}(s) ds\, d\theta + (1-\alpha_2)\tau_M \int_{-\tau_M}^{-\alpha_2\tau_M}\int_{t+\theta}^{t} \dot{x}^T(s) R_4 \dot{x}(s) ds\, d\theta,$$

$$V_4(t) = \int_{-\tau_M}^{0}\int_{\theta}^{0}\int_{t+\lambda}^{t} \dot{x}^T(s) R_5 \dot{x}(s) ds\, d\lambda\, d\theta + \int_{-h_M}^{-\alpha_1 h_M}\int_{\theta}^{0}\int_{t+\lambda}^{t} \dot{x}^T(s) R_6 \dot{x}(s) ds\, d\lambda\, d\theta + \int_{-\tau_M}^{-\alpha_2\tau_M}\int_{\theta}^{0}\int_{t+\lambda}^{t} \dot{x}^T(s) R_7 \dot{x}(s) ds\, d\lambda\, d\theta.$$

with

$$\zeta^T(t) = \left[\begin{array}{l} x^T(t)\ \ x^T(t-\alpha_1 h_M)\ \ x^T(t-h(t))\ \ x^T(t-h_M)\ \ x^T(t-\alpha_2\tau_M)\ \ x^T(t-\tau(t))\ \ x^T(t-\tau_M) \\ \dot{x}^T(t)\ \ \dot{x}^T(t-\tau(t)) \left(\int_{t-\tau_M}^{t} x(s)ds\right)^T \left(\int_{t-h_M}^{t-\alpha_1 h_M} x(s)ds\right)^T \left(\int_{t-\tau_M}^{t-\alpha_2\tau_M} x(s)ds\right)^T\ f_0^T\ \ f_1^T\ \ f_2^T \end{array}\right]$$

Then, the time derivative of $V(t)$ along the trajectory of system (1) is given by:

$$\dot{V}(t) = \dot{V}_1(t) + \dot{V}_2(t) + \dot{V}_3(t) + \dot{V}_4(t) \tag{25}$$

where

$$\dot{V}_1(t) = \dot{x}^T(t) P x(t) + x^T(t) P \dot{x}(t),$$

$$\begin{aligned}\dot{V}_2(t) &= x^T(t)\left(Q_1 + Q_2 + Q_3 + Q_4 + Q_6 + Q_7\right) x(t) \\ &\quad - x^T(t-\alpha_1 h_M) Q_1 x(t-\alpha_1 h_M) - (1-\dot{h}(t)) x^T(t-h(t)) \\ &\quad Q_2 x(t-h(t)) - x^T(t-h_M) Q_3 x(t-h_M) - (1-\dot{\tau}(t)) x^T(t-\tau(t)) \\ &\quad Q_4 x(t-\tau(t)) - (1-\dot{\tau}(t)) \dot{x}^T(t-\tau(t)) Q_5 \dot{x}(t-\tau(t)) + \dot{x}^T(t) Q_5 \dot{x}(t) \\ &\quad - x^T(t-\alpha_2\tau_M) Q_6 x(t-\alpha_2\tau_M) - x^T(t-\tau_M) Q_7 x(t-\tau_M)\end{aligned}$$

$$\begin{aligned}\dot{V}_3(t) &= \dot{x}^T(t)((\alpha_1 h_M) R_1 + (h_M - \alpha_1 h_M) R_2 + (\alpha_2\tau_M) R_3 \\ &\quad + (\tau_M - \alpha_2\tau_M)^2 R_4)\dot{x}(t) - \int_{t-\alpha_1 h_M}^{t} \dot{x}^T(s) R_1 \dot{x}(s) ds \\ &\quad - \int_{t-h_M}^{t-\alpha_1 h_M} \dot{x}^T(s) R_2 \dot{x}(s) ds - \int_{t-\alpha_2\tau_M}^{t} \dot{x}^T(s) R_3 \dot{x}(s) ds \\ &\quad - (1-\alpha_2)\tau_M \int_{t-\tau_M}^{t-\alpha_2\tau_M} \dot{x}^T(s) R_4 \dot{x}(s) ds,\end{aligned}$$

$$\begin{aligned}\dot{V}_3(t) &\le \dot{x}^T(t)((\alpha_1 h_M) R_1 + (h_M - \alpha_1 h_M) R_2 + (\alpha_2\tau_M) R_3 \\ &\quad + (\tau_M - \alpha_2\tau_M)^2 R_4)\dot{x}(t) - \int_{t-h(t)}^{t} \dot{x}^T(s) R_1 \dot{x}(s) ds \\ &\quad - \int_{t-h_M}^{t-\alpha_1 h_M} \dot{x}^T(s) R_2 \dot{x}(s) ds - \int_{t-\tau(t)}^{t} \dot{x}^T(s) R_3 \dot{x}(s) ds \\ &\quad - (1-\alpha_2)\tau_M \int_{t-\tau_M}^{t-\alpha_2\tau_M} \dot{x}^T(s) R_4 \dot{x}(s) ds\end{aligned}$$

and

$$\begin{aligned}\dot{V}_4(t) &= \dot{x}^T(t)\left(\frac{\tau_M^2}{2} R_5 + \frac{(h_M^2 - \alpha_1^2 h_M^2)}{2} R_6 + \frac{(\tau_M^2 - \alpha_2^2\tau_M^2)}{2} R_7\right)\dot{x}(t) \\ &\quad - \int_{-\tau_M}^{0}\int_{t+\theta}^{t} \dot{x}^T(s) R_5 \dot{x}(s) ds\, d\theta - \int_{-h_M}^{-\alpha_1 h_M}\int_{t+\theta}^{t} \dot{x}^T(s) R_6 \dot{x}(s) ds\, d\theta \\ &\quad - \int_{-\tau_M}^{-\alpha_2\tau_M}\int_{t+\theta}^{t} \dot{x}^T(s) R_7 \dot{x}(s) ds\, d\theta.\end{aligned}$$

The upper bound of the integral terms in inequality $\dot{V}_3(t)$ is estimated as:

$$\begin{aligned}&- \int_{t-h(t)}^{t} \dot{x}^T(s) R_1 \dot{x}(s) ds - \int_{t-h_M}^{t-\alpha_1 h_M} \dot{x}^T(s) R_2 \dot{x}(s) ds - \int_{t-\tau(t)}^{t} \dot{x}^T(s) R_3 \dot{x}(s) ds \\ &- (1-\alpha_2)\tau_M \int_{t-\tau_M}^{t-\alpha_2\tau_M} \dot{x}^T(s) R_4 \dot{x}(s) ds \\ &= - \int_{t-h(t)}^{t} \dot{x}^T(s) R_1 \dot{x}(s) ds - \int_{t-h_M}^{t-h(t)} \dot{x}^T(s) R_2 \dot{x}(s) ds - \int_{t-h(t)}^{t-\alpha_1 h_M} \dot{x}^T(s) R_2 \dot{x}(s) ds \\ &- \int_{t-\tau(t)}^{t} \dot{x}^T(s) R_3 \dot{x}(s) ds - (1-\alpha_2)\tau_M \int_{t-\tau_M}^{t-\alpha_2\tau_M} \dot{x}^T(s) R_4 \dot{x}(s) ds\end{aligned} \tag{26}$$

Using Jensen's inequality, such that

$$-(1-\alpha_2)\tau_M \int_{t-\tau_M}^{t-\alpha_2\tau_M} \dot{x}^T(s) R_4 \dot{x}(s) ds \le \begin{bmatrix} x(t-\alpha_2\tau_M) \\ x(t-\tau_M) \end{bmatrix}^T \begin{bmatrix} -R_4 & R_4 \\ R_4 & -R_4 \end{bmatrix} \begin{bmatrix} x(t-\alpha_2\tau_M) \\ x(t-\tau_M) \end{bmatrix} \tag{27}$$

$$-\int_{-\tau_M}^{0}\int_{t+\theta}^{t} \dot{x}^T(s) R_5 \dot{x}(s) ds\, d\theta \le \frac{2}{\tau_M^2} \begin{bmatrix} \tau_M x(t) \\ \int_{t-\tau_M}^{t} x(s) ds \end{bmatrix}^T \begin{bmatrix} -R_5 & R_5 \\ R_5 & -R_5 \end{bmatrix} \begin{bmatrix} \tau_M x(t) \\ \int_{t-\tau_M}^{t} x(s) ds \end{bmatrix} \tag{28}$$

$$-\int_{-h_M}^{-\alpha_1 h_M}\int_{t+\theta}^{t}\dot{x}^T(s)R_6\dot{x}(s)ds\,d\theta \le \frac{2}{(h_M^2-\alpha_1^2h_M^2)}\begin{bmatrix}(1-\alpha_1)h_M\,x(t)\\ \int_{t-h_M}^{t-\alpha_1 h_M}x(s)ds\end{bmatrix}^T\begin{bmatrix}-R_6 & R_6\\ R_6 & -R_6\end{bmatrix}\begin{bmatrix}(1-\alpha_1)h_M\,x(t)\\ \int_{t-h_M}^{t-\alpha_1 h_M}x(s)ds\end{bmatrix}$$

(29)

$$-\int_{-\tau_M}^{-\alpha_2 \tau_M}\int_{t+\theta}^{t}\dot{x}^T(s)R_7\dot{x}(s)ds\,d\theta \le \frac{2}{(\tau_M^2-\alpha_2^2\tau_M^2)}\begin{bmatrix}(1-\alpha_2)\tau_M\,x(t)\\ \int_{t-\tau_M}^{t-\alpha_2 \tau_M}x(s)ds\end{bmatrix}^T\begin{bmatrix}-R_7 & R_7\\ R_7 & -R_7\end{bmatrix}\begin{bmatrix}(1-\alpha_2)\tau_M\,x(t)\\ \int_{t-\tau_M}^{t-\alpha_2 \tau_M}x(s)ds\end{bmatrix}$$

(30)

By using Lemma 3, an upper bound of integral term of $\dot{V}(t)$ can be obtained as:

$$\begin{aligned}-\int_{t-h_M}^{t-h(t)}\dot{x}^T(s)R_2\,\dot{x}(s)\,ds &\le (h_M-h(t))\,\zeta^T(t)T\,R_2^{-1}T^T\zeta(t)\\ &\quad+2\,\zeta^T(t)T\left[x(t-h(t))-x(t-h_M)\right]\end{aligned} \tag{31}$$

$$\begin{aligned}-\int_{t-h(t)}^{t}\dot{x}^T(s)R_1\,\dot{x}(s)\,ds &\le h(t)\,\zeta^T(t)Y\,R_1^{-1}Y^T\zeta(t)\\ &\quad+2\,\zeta^T(t)Y\left[x(t)-x(t-h(t))\right]\end{aligned} \tag{32}$$

$$\begin{aligned}-\int_{t-h(t)}^{t-\alpha_1 h_M}\dot{x}^T(s)R_2\,\dot{x}(s)\,ds &\le (h(t)-\alpha_1 h_M)\,\zeta^T(t)W\,R_2^{-1}W^T\zeta(t)\\ &\quad+2\,\zeta^T(t)W\left[x(t-\alpha_1 h_M)-x(t-h(t))\right]\end{aligned} \tag{33}$$

$$\begin{aligned}-\int_{t-\tau(t)}^{t}\dot{x}^T(s)R_3\,\dot{x}(s)\,ds &\le \tau(t)\,\zeta^T(t)X\,R_3^{-1}X^T\zeta(t)\\ &\quad+2\,\zeta^T(t)X\left[x(t)-x(t-\tau(t))\right]\end{aligned} \tag{34}$$

From Equation (5), the following inequalities hold:

$$\begin{cases}\beta_0^2x^T(t)\;x(t)-f_0^T f_0\ge 0\\ \beta_1^2x^T(t-\tau_1(t))\;x(t-\tau_1(t))-f_1^T f_1\ge 0\\ \beta_2^2\dot{x}^T(t-\tau_2(t))\;\dot{x}(t-\tau_2(t))-f_2^T\;f_2\ge 0\end{cases} \tag{35}$$

Further, for any scalars $\varepsilon_i > 0$ (i = 0, 1, 2), it follows from Equation (35), that

$$\begin{cases}\varepsilon_0\,[\beta_0^2x^T(t)\;x(t)-f_0^T\;f_0]\ge 0,\\ \varepsilon_1\;[\beta_1^2x^T(t-\tau_1(t))\;x(t-\tau_1(t))-f_1^T f_1]\ge 0,\\ \varepsilon_2\,[\beta_2^2\dot{x}^T(t-\tau_2(t))\;\dot{x}(t-\tau_2(t))-f_2^T f_2]\ge 0,\end{cases} \tag{36}$$

Therefore, combining Equations (25) and (36) yields:

$$\dot{V}(t)\le\zeta^T(t)\Pi\;\zeta(t), \tag{37}$$

with

$$\zeta^T(t)=\begin{bmatrix}x^T(t)\;\;x^T(t-\alpha_1h_M)\;\;x^T(t-h(t))\;\;x^T(t-h_M)\\ x^T(t-\alpha_2\tau_M)\;\;x^T(t-\tau(t))\;\;x^T(t-\tau_M)\\ \dot{x}^T(t)\;\;\dot{x}^T(t-\tau(t))\left(\int_{t-\tau_M}^{t}x(s)ds\right)^T\left(\int_{t-h_M}^{t-\alpha_1h_M}x(s)ds\right)^T\\ \left(\int_{t-\tau_M}^{t-\alpha_2\tau_M}x(s)ds\right)^T\;f_0^T\;f_1^T\;f_2^T\end{bmatrix}$$

and Π is given in Equation (23). By using the Schur Complement, it is clear to see that the results $\dot{V}(t)<0$ holds if $\Pi<0$, $\alpha_1h_M\le h(t)\le h_M$, and $\alpha_2\tau_M\le\tau(t)\le\tau_M$.

Thus, the system (1) is asymptotically stable according to the Lyapunov-Krasovskii method.

Theorem 3. In Case II, for given positive scalars h_m,h_M, τ_m, $\tau_M,\eta,\mu,\beta_0,\beta_1$ and β_2, the system (1) with uncertainty (5) and mixed time-varying delays satisfying Equations (2) and (3) is asymptotically stable if there exist symmetric positive definite $n\times n$ matrices $P, Q_i\,(i=1,3,..7)$, $R_j\,(j=1,..,7)$, for any free matrix variables $T_a,Y_a,W_a,N_a,X_a,F_a\,(a=1,2)$ and scalars $\varepsilon_i\ge 0\;(i=0,1,2)$ such that the following symmetric LMI holds:

$$\begin{bmatrix}\Sigma & \sqrt{\alpha_1h_M-h_m}\,T & \sqrt{\alpha_1h_M}\,Y & \sqrt{\alpha_1h_M-h_m}\,W & \sqrt{\alpha_2\tau_M-\tau_m}\,N & \sqrt{\alpha_2\tau_M}\,X\\ * & -(R_1+R_2) & 0 & 0 & 0 & 0\\ * & * & -R_1 & 0 & 0 & 0\\ * & * & * & -R_2 & 0 & 0\\ * & * & * & * & -R_3 & 0\\ * & * & * & * & * & -R_3\end{bmatrix}<0, \tag{38}$$

where $\Sigma=\left(\Sigma_{i,j}\right)_{15\times15}$

with

$$\Sigma_{1,1}=Q_1+Q_3+Q_4+Q_6+Q_7+Y_1+Y_1^T+X_1+X_1^T$$
$$-2R_5-2R_6-2R_7+F_1A+A^TF_1^T+\varepsilon_0\beta_0^2\mathrm{I}$$
$$\Sigma_{1,3}=-Y_1+Y_2^T+F_1A_1$$
$$\Sigma_{1,6}=-X_1+X_2^T$$
$$\Sigma_{1,8}=P-F_1+A^TF_2^T$$
$$\Sigma_{1,9}=F_1A_2$$
$$\Sigma_{1,10}=\frac{2}{\alpha_2\tau_M}R_5$$
$$\Sigma_{1,11}=\frac{2}{\left(\alpha_1h_M+h_m\right)}R_6$$
$$\Sigma_{1,12}=\frac{2}{\left(\alpha_2\tau_M+\tau_m\right)}R_7$$
$$\Sigma_{1,13}=\Omega_{1,14}=\Omega_{1,15}=F_1$$
$$\Sigma_{2,2}=-Q_3+W_1+W_1^T$$
$$\Sigma_{2,3}=-W_1+W_2^T$$
$$\Sigma_{3,3}=T_1+T_1^T-Y_2-Y_2^T-W_2-W_2^T+\varepsilon_1\beta_1^2\mathrm{I}$$
$$\Sigma_{3,4}=-T_1+T_2^T$$
$$\Sigma_{3,8}=A_1^TF_2^T$$
$$\Sigma_{4,4}=-Q_1-T_2-T_2^T$$
$$\Sigma_{5,5}=-R_4-Q_7$$

$\Sigma_{5,7} = R_4$

$\Sigma_{6,6} = -(1-\eta)Q_4 + N_1 + N_1^T - X_2 - X_2^T$

$\Sigma_{7,7} = -R_4 - Q_6 - N_2 - N_2^T$

$\Sigma_{8,8} = M - F_2 - F_2^T$

$\Sigma_{8,9} = F_2 A_2$

$\Sigma_{8,13} = \Omega_{8,14} = \Omega_{8,15} = F_2$

$\Sigma_{9,9} = -(1-\eta)Q_5 + \varepsilon_2 \beta_2^2 \mathrm{I}$

$\Sigma_{10,10} = -\dfrac{2}{\alpha_2^2 \tau_M^2} R_5$

$\Sigma_{11,11} = -\dfrac{2}{\left(\alpha_1^2 h_M^2 - h_m^2\right)} R_6$

$\Sigma_{12,12} = -\dfrac{2}{\left(\alpha_2^2 \tau_M^2 - \tau_m^2\right)} R_7$

$\Sigma_{13,13} = -\varepsilon_0 \mathrm{I}$

$\Sigma_{14,14} = -\varepsilon_1 \mathrm{I}$

$\Sigma_{15,15} = -\varepsilon_2 \mathrm{I}$

$$\xi^T(t) = \begin{bmatrix} x^T(t)\ x^T(t-h_m)\ x^T(t-h(t))\ x^T(t-\alpha_1 h_M)\ x^T(t-\tau_m)\ x^T(t-\tau(t))\ x^T(t-\alpha_2\tau_M) \\ \dot{x}^T(t)\ \dot{x}^T(t-\tau(t)) \left(\int_{t-\alpha_2\tau_M}^{t} x(s)\,ds\right)^T \left(\int_{t-\alpha_1 h_M}^{t-h_m} x(s)\,ds\right)^T \left(\int_{t-\alpha_2\tau_M}^{t-\tau_m} x(s)\,ds\right)^T f_0^T\ f_1^T\ f_2^T \end{bmatrix}$$

In Case II, a Lyapunov-Krasovskii functional can be chosen as (8) with $Q_2 = 0$. Similar to the above analysis, one can get that the results $\dot{V}(t) < 0$ holds if $\Sigma < 0$, $h_m \le h(t) \le \alpha_1 h_M$, and $\tau_m \le \tau(t) \le \alpha_2 \tau_M$. Thus, the proof is completed.

Remark 2: By introducing a new class of augmented Lyapunov-Krasovskii functional approach, new delay-decomposition stability criteria for nonlinear neutral systems with mixed time-varying delays are obtained in Theorems 1-3. The proposed augmented Lyapunov functional using the novel triple integral inequality is more robust than existing results in literature. It gives the upper bounds of time-varying delays $h(t), \tau(t)$ for the asymptotic stability of system (1) which can be provided larger stability domain. In addition, by applying free-weighting matrices and Jensen integral inequality, our decomposition approach, developed in Theorems 1-3, yields a much less conservative delay bounds and extends the feasible region of stability method for system (1).

Remark 3: In order to derive a fewer restrictive stability criteria for system (1), many free-weighting matrix variables are employed in Theorems 1-3. In fact, this technique of decision variables reduces the computational complexity of the obtained stability approach which is less than the previous methods.

Remark 4: In this work and from the practical point of view, several problems related to this studied field are still open such as singular descriptor systems with multiple mixed time-varying delays, chaotic systems with varying delays and neural networks systems.

IV. Illustrative Examples

In this section, two examples are presented in order to show the less conservatism of the elaborated stability condition and to demonstrate the effectiveness of the proposed approach.

Example 1.

Consider the following nonlinear neutral system with mixed time-varying delays, as given in [10]:

$$A = \begin{bmatrix} -1.2 & 0.1 \\ -0.1 & -1 \end{bmatrix},\ A_1 = \begin{bmatrix} -0.6 & 0.7 \\ -1 & -0.8 \end{bmatrix},\ A_2 = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}, \qquad (39)$$

where $0 \le |c| < 1$, $\beta_0 \ge 0, \beta_1 \ge 0,$ and $\beta_2 \ge 0.$

Case I. For $c = 0.1$, $\beta_1 = 0.1,\ \tau_M = 1,\ \mu = 0.5,\ \eta = 0,$ $\alpha_2 = 0.2$ and different values of β_2, the maximal allowable delay of h_M estimated by Theorems 1 and 2 are illustrated in Table 1. This table shows the numerical results for different values of β_2, $\beta_0 = 0$ and $\beta_0 = 0.1$. As β_2 increases, h_M decreases. In addition, the proposed stability technique gives a much less conservative result than other recent ones.

Case II. For $\beta_0 = 0.1,\ \beta_1 = 0.2,\ \beta_2 = 0.1,\ \alpha_2 = 0.2,\ \mu = 1$ and different values of *c*, the maximum admissible upper bound on the allowable time delay of $h_M = \tau_M$ obtained from Theorem 1 are listed in Table 2. As *c* increases, h_M decreases. It is clear that the proposed stability method in this paper provides larger upper bounds of delay system than the previous results for different values of *c*.

Case III. For c = 0.1, $\alpha_2 = 0.2$, $\beta_1 = 0.1,\ \beta_2 = 0,\ \beta_0 = 0$ and $\beta_0 = 0.1$, and different values of $\mu = \eta$, the maximum upper bounds on the allowable delay of $h_M = \tau_M$ obtained from Theorems 1 and 2 are illustrated in Table 3. As μ increases, h_M decreases. The presented stability criterion is less conservative than existing results.

TABLE I. Maximum Allowable Delay Bound of h_M with $\mu = 0.5, \eta = 0$ and Different Values of β_2

$\beta_0 = 0$				
β_2	0	0.1	0.2	0.3
Rakkiyappan et al.[18]	1.4886	1.2437	0.9921	0.7367
Lakshmanan et al.[13]	1.6325	1.3386	1.0816	0.8563
Cheng et al. [23]	1.6865	1.3721	1.0923	0.8613

Qiu and Zhang [21]	2.2937	1.8505	1.4565	1.1105
Theorem1($\alpha_1 = 0.25$)	6.0782	5.1772	3.4872	1.9325
Theorem2 ($\alpha_1 = 0.1$)	4.7856	4.0752	2.7424	1.5156
$\beta_0 = 0.1$				
β_2	0	0.1	0.2	0.3
Rakkiyappan et al.[18]	1.3244	1.0901	0.8475	0.6300
Lakshmanan et al.[13]	1.4440	1.1950	0.9734	0.7760
Cheng et al. [23]	1.4721	1.2466	0.9996	0.7804
Qiu and Zhang [21]	2.0417	1.6541	1.3062	0.9982
Theorem1($\alpha_1 = 0.25$)	5.6888	4.9444	3.4125	1.9128
Theorem2 ($\alpha_1 = 0.1$)	4.4785	3.8915	2.6834	1.5000

TABLE II. MAXIMUM UPPER BOUND OF $h_M = \tau_M$ WITH DIFFERENT VALUES OF C

c	0.1	0.2	0.3
Zhang and Yu [17]	0.4911	0.4125	0.3382
Qiu et al. [15]	1.8567	1.6242	1.3917
Qiu and Zhang [21]	2.1916	1.6632	1.4743
Theorem1($\alpha_1 = 0.1$)	6.4209	5.5817	4.7240
c	0.4	0.5	0.6
Zhang and Yu [17]	0.2671	0.1975	0.1294
Qiu et al. [15]	1.1592	0.9270	0.6945
Qiu and Zhang [21]	1.2396	0.9288	0.7446
Theorem1($\alpha_1 = 0.1$)	3.8303	2.8351	1.5622

TABLE III. MAXIMUM UPPER BOUND OF $h_M = \tau_M$ WITH DIFFERENT VALUES OF $\mu = \eta$

	$\beta_0 = 0$	$\beta_1 = 0.1$	$\beta_0 = 0.1$	$\beta_1 = 0.1$
$\mu = \eta$	0	0.5	0	0.5
Chen et al. [23]	2.7423	1.1425	1.8753	1.0097
Liu [22]	2.7429	1.4462	1.8895	1.1485
Qiu and Zhang [21]	3.8066	1.6402	2.6039	1.4534
Theorem1($\alpha_1 = 0.6$)	8.1497	2.3438	5.6008	2.1964
Theorem2 ($\alpha_1 = 0.7$)	5.8444	1.6738	4.0165	1.5683

Example 2.

Consider the mixed time-varying delay systems as depicted in Equation (40):

$$A = \begin{bmatrix} -2 & 0.5 \\ 0 & -1 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0.4 \\ 0.4 & -1 \end{bmatrix}, A_2 = \begin{bmatrix} 0.2 & 1 \\ 0 & 0.2 \end{bmatrix}, \quad (40)$$

with

$f_0^T(t, x(t))\ f_0(t, x(t)) \le \beta_0^2\, x^T(t)\ x(t)$ and

$f_1^T(t, x(t-h(t)))\, f_1(t, x(t-h(t))) \le \beta_1^2\, x^T(t-h(t))\ x(t-h(t))$.

While using the parameters $\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $\eta = \mu = 0$, $\beta_0 = 0.2$, $\beta_1 = 0.1$ and $\beta_2 = 0$, the upper bound of Time Delay $h_M = \tau_M$ obtained from Theorem 3 is feasible for any delay $0 < h_M \le 3.8561$.

It is remarkable that this proposed criterion is much less conservative than the results shown in [14, 16].

V. CONCLUSION

This paper studied the problem of asymptotic stability for nonlinear neutral mixed time-varying delays systems. By using the Lyapunov–Krasovskii functional with triple integral terms and free weighting matrices approach, new delay-dependent stability criteria are derived by developing a delay decomposition technique. The elaborated approach is then expressed in terms of LMIs. Finally, numerical simulations have been investigated in order to show the robustness and the flexibility of the proposed stability method.

REFERENCES

[1] M.N. Alpaslan Parlakçı, Delay-Dependent Robust Stability criteria for uncertain neutral systems with mixed time-varying discrete and neutral delays, Asian Journal of Control, vol. 9, no. 4, pp. 411-421, December 2007.

[2] I. Amri, D. Soudani and M. Benrejeb, A Delay Decomposition Approach for Exponential Stability of Perturbed Neutral Systems with Mixed Delay, 7th International Multi-Conference on Systems, Signals and Devices (SSD'10), Amman, Jordan, 2010.

[3] P. L. Liu, A delay decomposition approach to stability analysis of neutral systems with time varying delay, Appl. Math. Modell, vol. 37, pp. 5013-5026, 2013.

[4] F. Qiu, B.Cui, and Y. Ji, A delay-dividing approach to stability of neutral system with mixed delays and nonlinear perturbations, Applied Mathematical Modelling, vol. 34, pp.3701-3707, 2010.

[5] Ali, MS, On exponential stability of neutral delay differential system with nonlinear uncertainties, Nonlinear Sci. Numer. Simul, vol. 17, pp. 2595-2601, 2012.

[6] J. Cheng, H. Zhu, S. M. Zhong, G. H. Li, Novel delay-dependent robust stability criteria for neutral systems with mixed time-varying delays and nonlinear perturbations, Appl. Math. Comput, vol. 219, pp. 7741-7753, 2013.

[7] Junjun HUI, Hexin ZHANG, Xiangyu KONG, and Fei MENG, A Less Conservative Robust Stability Criteria for Neutral System with Mixed Time-varying Delays and Nonlinear Perturbations, Journal of Computational Information Systems, vol. 4, no. 10, pp. 1543-1553, 2014.

[8] X.M. Zhang, Q.L. Han, A delay decomposition approach to delay-dependent stability for linear systems with time-varying delays, International Journal of Robust and Nonlinear Control, vol. 19, no. 17, pp. 1922-1930, 2009.

[9] S. Boyd, L. EL Ghaoui, E. Feron, V. Balakrishnan, Linear matrix inequalities in system and control theory, Studies in Applied Mathematics, SIAM, Philadelphia, USA, 1994.

[10] O.M. Kwon, J.H. Park, S.M. Lee, An improved delay-dependent criterion for asymptotic stability of uncertain dynamic systems with time-varying delays, J. Optim. Theory Appl., vol. 145, pp. 343-353, 2010.

[11] F. Gouaisbaut and D. Peaucelle, Delay-dependent stability of time delay systems, 5th IFAC Symposium on Robust Control (ROCOND'06), Toulouse, France, 5–7 July 2006.

[12] Q. L. Han, Robust stability for a class of linear systems with time-varying delay and nonlinear perturbations, Computer and Mathematics with Applications, vol.47, pp.1201-1209, 2004.

[13] S. Lakshmanan, T. Senthilkumar, and P. Balasubramaniam, Improved results on robust stability of neutral systems with mixed time-varying delays and nonlinear perturbations, Applied Mathematical Modelling, vol. 35, pp. 5355-5368, 2011.

[14] J.H. Park, Novel robust stability criterion for a class of neutral systems with mixed delays and nonlinear perturbations, Applied Mathematics and Computation, vol. 161, no. 2, pp. 413-421, 2005.

[15] F. Qiu, B. Cui, and Y. Ji, Further results on robust stability of neutral system with mixed time-varying delays and nonlinear perturbations, Nonlinear Analysis: Real World Applications, vol. 11, pp. 895-906, 2010.

[16] L. L. Xiong, S. M. Zhong, and D. Y. Li, Novel delay-dependent asymptotical stability of neutral systems with nonlinear perturbations, Journal of Computational and Applied Mathematics, vol. 232, pp. 505-513, 2009.

[17] [17] W. A. Zhang and L.Yu, Delay-dependent robust stability of neutral systems with mixed delays and nonlinear perturbations, Acta Automatica Sinica, vol. 33, pp. 863-866, 2007.

[18] [18] R. Rakkiyappan, P. Balasubramaniam, R. Krishnasamy, Delay dependent stability analysis of neutral systems with mixed time-varying delays and nonlinear perturbations, J. Comput. Appl. Math., vol. 235, pp. 2147-2156, 2011.

[19] I. Amri and D. Soudani, Robust Exponential Stability of Uncertain Perturbed Systems with TimeVarying Delays, IFAC-12th LSS symposium: Theory and Applications, Lille, France, 2010.

[20] I. Amri, D. Soudani and M. Benrejeb, On Robust α-Stability Analysis of Uncertain Neutral Systems with Time Varying Delays: A Novel Augmented Lyapunov Functional Approach, 8th International Multi-Conference on Systems, Signals and Devices (SSD'11), Sousse, Tunisia, 2011.

[21] F. Qiu and Q. Zhang, Robust Stability of Neutral System with Mixed Time-Varying Delays and nonlinear perturbations Using Delay Decomposition Approach, Abstract and Applied Analysis, vol. 2014, Article ID 825715, 11 pages, 2014.

[22] P. L. Liu, Robust stability for neutral time-varying delay systems with non-linear perturbations, International Journal of Innovative Computing, Information and Control, vol.7, no. 10, pp.5749-5760, 2011.

[23] Y. Chen, A. K. Xue, R. Q. Lu, and S. S. Zhou, On robust exponential stability of uncertain neutral systems with time-varying delays and nonlinear perturbations, Nonlinear Analysis, vol. 68, pp.2464-2470, 2008.

Analyzing Virtual Machine Live Migration in Application Data Context

Mutiullah Shaikh
Faculty of Electrical, Electronic and Computer Engineering,
Mehran University of Engineering and Technology,
Jamshoro, Pakistan

Muhammad Ali Memon
School of Information Technology,
Shaheed Benazir Bhutto University,
Shaheed Benazirabad, Pakistan

Asadullah Shaikh
College of Computer Science and Information Systems,
Najran University,
Najran, Saudi Arabia

Farah Deeba
Faculty of Engineering Science and Technology,
Hamdarad Institute of Engineering and Technology,
Karachi, Pakistan

***Abstract*—Virtualization plays a very vital role in the big cloud federation. Live and Real-time virtual machine migration is always a challenging task in virtualized environment, different approaches, techniques and models have already been presented and implemented by many re- searchers. The aim of this work is to investigate various parameters of Real-time and live data migration of virtual machines in stateful and data context at the application level. The migration of one virtual machine to another requires some time depending on the network bandwidth, guest availability, hardware limitation overcomes, resource allocation, server reallocation, hypervisor compatibility and many more. To enhance and ensure the performance and optimization of the time this work presents the some analysis in the form of different time stacks in multiple piece of data stored in the virtual machines. To optimize the migration time virtual machine checkpoints are used in order to achieve the better results by using the xen hypervisor memory technique which dynamically allows the migration of the configured memory while the allocated memory could be discarded for a while. By this the bad memory remains un-migrated only the good memory consisting the used data would be migrated by means of Real-time.**

Keywords—component; Cloud Computing; Virtualization; Virtual Machine Monitor VMM; Xen; VMResume; Xen Save and Restore; DC Data Centers Copy on Write CoW

I. INTRODUCTION

As the demand of cloud computing is increasing, storage and communication resources within data centers (DC) are developing new ways for the distributed resources of computing and sharing infrastructure by using virtualization. Virtualization actually was deployed for the cost saving. But very soon organizations realized that it is also effective in terms of speed, flexibility and robustness. In general,” virtualization” refers to the process of turning a hardware-based entity into a software-embedded component and this is encapsulated in an entity called Virtual Machine (VM). By using Virtual Machines technique the resources are utilized in much more effective manner [2]. Virtualization has attracted considerable interest in recent years, particularly from the data centers and cluster computing communities. Since clusters are costly to own, therefore transferring and sharing access to a single general cluster is an optimal solution when demands vary time by time [3]. In other words, sharing of access or clusters is known as migration of virtual machines, means moving a VM from one source host to another sink host. If one VM has lot of load to carry, it can move and share some of load to another VM for better performance and results. Migration is also useful in maintenance of VMs. Additionally, if one VM fails, then through live migration the VM host failure recovery could be achieved. Live migration makes these invisible and seamless to users and end users [4]. Hence, this research typically focuses on the problems and different approaches to analyze the performance of the parameters for the Real-time on live data. Additionally, virtual machine migration between the single/multiple virtual machines on the basis of data availability, state maintenance in terms of time using multiple scenarios of data context in virtualized environment.

Virtual machine live migration in the cloud federation virtualized environment is always a much spirited task. Live migration of Virtual Machines plays vital role by providing virtual machine robustness. The main objective behind this research is to investigate and analyze different parameters achieved after implementation of live virtual machine migration. Specifically, this work aims to conduct the analysis for the optimization of migration time and live migration down time in the multiple scenarios such as:

– Time required for the Data Migration

– Time required for the State Maintenance

– Time Required for the Network Migration

In order to achieve the optimization and results in terms of time required by migration this research aims to perform the Migration between the virtual machines with the different amount of data and memory.

To implement the Virtual Machine Migration in the Cloud environment ini- tially tools required are:

– Ubuntu Cloud server 14.04 LTS TrustY

– Virtual Machine Monitor/Hypervisor Xen

Xen is the enhanced and updated type1 hypervisor which helps in creation of guest VMs and provide full access to the created guest VMs. It has the capability to save a Virtual Machine in a running state. After taking the snapshot of the saved Virtual Machine xen migrates that Virtual Machine to the another Virtual Machine in Real-time. This technology of xen called the xen Virtual Machine Xen save and resume.

II. Related Work

This section aims to describe some related work about the virtualization in the big cloud federation as well as the virtual machine live migration in the virtualized cloud environment and cluster based system. The idea behind this chapter is to present the earlier work for the sharing of resources between the virtual machines in terms of real-time and some discussion on enhancement in live migration. For the shared resources and to avoid the congestion of huge work load from a virtual machine various models and techniques have been specified by numerous researchers to study the workload and to present the overhead solutions.

A. Background & overview: live migration

As far as migration is concerned, it also took place in terms of offline migration, typically non-real-time. Web suspends/resume highlights on saving/resumption the current computing position and state on unspecified hardware [5]. Sapuntzakis et. al. attends to end user mobility and system management by encapsulating the computing atmosphere into capsules that could be shared between distinctive guests [6]. Schmidt et. al. by deploy capsules, in the form of interrelated processes with their network addresses i.e., IP and the entire network states, as the shape of migration decades [7]. In the same way, Zap uses process groups (pods) as well as their state at the kernel level in the shape of migration decades [8]. In all these proposed preceding, the running execution suspended for a while and the processes which is in use in the form of applications within the VM remained un-progressed.

To get the always availability of the data and other computing resources, presently many techniques of live migration exits in the virtualization-based environment [9, 10]. From all two of the illustrations are live migration in Xen [11] i.e., xen motion and migration of VMwares i.e., Vmotion [12], which migrates in the same manner as pre-copy strategy. For the duration of migration, pages of physical memory transfers from the one primary sink host to the another new host (backup), while the state of the VM is running on the sink host (pri-mary).Memory pages which are replicated during the migration of VMs should ensure their consistency and integrity. After that iterative procedure of VM shar- ing phase, stop-and-copy phase will be initiated and executed for a while and that caused VM suspended, the remaining pages of the configured memory are transferred , due to this the machine monitor (Hypervisor) of the destination (backup) VM generate a signal for the resumption of the executed VM. Though, the pre-copy takes minimal downtime of VM being migrated in comparison of the others [13].

In addition with pre-copy procedure according to Kemari, there are some previous related techniques which proposed the better solution for optimization during the migration [14–20]. Such as post-copy migration technique has pointed out the cons of pre-copy migration [21, 22]. Some experimental results shows that the downtime taken by post-copy migration for a VM being migrated is less than the time spend by the pre-copy migration [21]. However, pre-copy implementation supports the (PV) para-virtualized users as the catching memory method is involved which accesses and manages a memory based pseudo-paging system within a guest. Since, as the upgraded/patched version of OS needed by the post- copy procedure so due to this it could not be commonly used as the pre-copy. Hines et. al. introduces the design by combining the post-copy and pre-copy mutually [21]. By this combination an adaptive pre-paging method proposed by them, which maintains the access patterns of the user applications.

B. Improved live vm migration

Now apart from that Remus is of the official Xen warehouse [23]. It gains the huge and always availability by keeping an copy of the updated VM along with its running state on the secondary host computer i.e., (backup), which is alert and activated whenever the primary host get failed and its state being destroyed. LLM initially updates the memory after that copied dirty data and then uses the pages of the memory excluding copying [24]. Although, the tracking of the bad data is not so much efficient, the onwards goal is to further enhance and save the memory and processing time as well as power by analyzing the performance in the different decades of data context within the VMs would be migrated from sink to backup.

Lu et al. also achieved the always availability by using three state memory synchronization [28], like in systems Remus: bad memory tracing, active VM backup along with tentative transfer of the state. This method describes main idea of the proposed work in the paper and tells us about the actual pros of the migration, however, it implies with the associated memory migration overhead. For instance, in the shared environment the swapping of workload and its estimation, the memory overhead is more than the 50%. Since the main and configured memory is the essential resource, the ratio in high percentage overhead is a trouble. To overcome the memory issues in the systems which are Xen-based, several ways are available to inflexibility memory redundancy in guest VMs, like as patching and sharing of memory pages. Some previous efforts have shown the potential memory sharing in virtualized systems. Some changes in working sets were inspected and their results demonstrated that changes were essential for the host to host VMs migration [11, 26]. For the guest virtual machine with 512MB allocated

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

memory, roughly changed low load with 20MB, roughly changed medium load with 80MB, roughly changed high load with 200MB. Therefore, the workloads normally take places between these boundaries. The previous evaluation also makes known the amount in memory changes with different workload running in the VMs (within minutes) [11, 26]. Within two minutes none of the VM make changes in the memory more than 4MB. The Content-Based Page Sharing (CBPS) technique also revealed the memory sharing potentially [21]. CBPS typically based on the technique known as Compare-by-hash introduced in [23, 24]. As claimed, the CBPS was capable to recognize form all pages 42.9% as much as sharable, and reclaimed from the all pages 32.9% doing real-world workload in the ten instances of Windows NT. Nine guests VMs were capable to illustrate the sharable pages as much as 29.2% and reclaimed when decreased from nine to five guest VMs as 18.7%, and the resultant numbers were 10.0% and 7.2%, correspondingly.

For Sharing the memory pages in the efficient manner, nowadays, the technique known as CoW (Copy-on-Write) was broadly adopted in Xen Hypervisor [27]. Unlike the OS that uses CoW method for the sharing of memory pages in a conventional way, in virtualized environment, pages are shared between the multiple guest VMs. as an alternative of using CoW to migrate the memory pages between the VMs, here we use the same idea but in the more efficient manner like by sharing the pages between smaller blocks. The Difference Engine illustrated the potential saving in memory obtainable from the leveraging a mixture of page patching, sharing and in-core level memory compression [13]. It also reveals the vast potential of exploiting memory redundancy in guest VMs. On the other hand, Difference Engine also faces difficulty problems when using the patching technique because some additional modification required by Xen.

III. Contributions

In this section of paper we aim to describe the work done during the conduct of this research. Specifically, few proceeding overviewed and detailed described in this section in order to achieve our aims and objectives. To overcome the issues faces during the Live migration we discussed an adaptive way of Live migration to improve the Load balancing, optimize the Downtime, VM disk/data migration. Furthermore, this section also defined resumption of a VM into another VM in the Real-time. The key idea behind this research work i.e., Live migration in data context is discussed in this section as well.

A. Virtual machine resumption

While during the resumption of saved VM using a stored checkpoint file from slow-access storage, the saved states in a checkpoint file should be retrieved. Those saved states are virtual/shared CPU states, the states for emulated de- vices as well as contents of memory disk of VM. Usually, most of the data saved in the checkpoint file retrieves from contents of VM memory. Therefore, a straightforward method for the resumption of VM is to restore first all the saved memory data from the saved check-point file of VM, and then retrieve the rest of data which includes device and CPU data. The VM cannot start without the required device and CPU states, it cannot be initialized until all data have been retrieved from memory along with its all previously stored memory pages have been set up. Presently, Xen hypervisor uses this method for VM resumption from a check-point file.

we can summarize that the same issue of the check-pointing VM mechanism also takes place in VM resumption: as amount of contents of VM memory dominate the stored and saved data in check-point file, when assigned memory to VM increases, the time consumed on restoring its saved data rapidly develop into bottleneck. As illustrated in figure 1, with the increase of size of VM memory, the time took by command xm restore would also increase linearly. For small amount of memory (i.e. with a size of VM memory 128MB), but while retrieving in gigabytes from a saved check-point file (i.e. 10s in the 1GB), there is a significant increase in time to resumption.

However, in the first solution, which typically restores data from memory before the device and CPU data, which results in-effective. Also, it is very difficult to consider data before the CPU and device while restoring the data memory in reverse order.

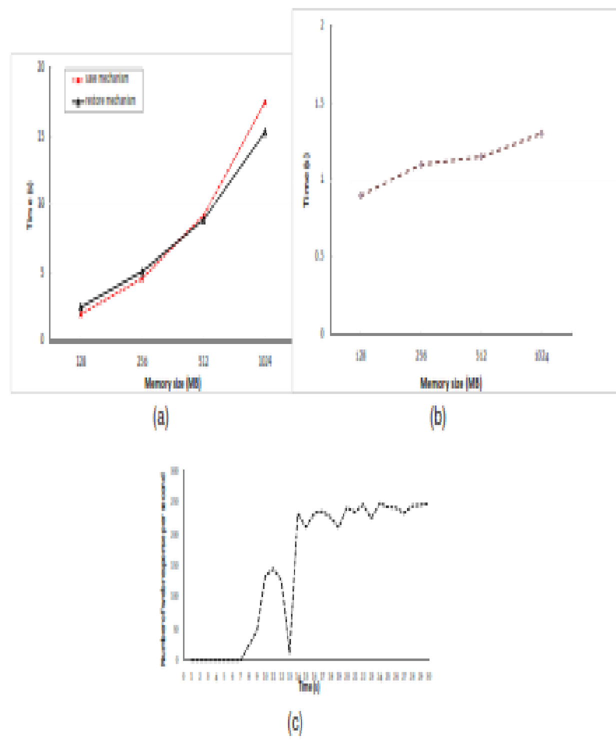


Fig. 1. Comparison of mechanisms for VM resumption

That means, letting the VM to boot initially to load required necessary devices and CPU states, after this loading, now restore the data of memory pre saved in check-pointed file after VM starts. In a certain case, when the VM requires to access a page from memory which has not still been loaded, then this corresponding data is retrievable from on-disk check-point file and get sets up the pages. This solution provide the benefit as it starts of VM very quickly, and always keeps it in running state while restoring the data memory. Moreover, since in this manner, the VM only require the restoration of small and less amount of necessary device and CPU states to start VM, its performance would not be affected by the size of VM memory.

However, in contrast with the first solution, the second one has some demerits. In the first solution, after the startup of VM (although it takes 10s or even few minutes), the works well and as better as before checkpointing. In comparison, with the second solution, the VM appears and supposed to be in running state after restoring the necessary device and CPU states. However, when it wants to accesses a page from memory which has not still be stored, a page fault immediately occurs. Then the in process execution must be paused by hypervisor, then the checkpoint file restores the memory pages, and then it will be resumed. Since at the beginning significant number of page faults occur due memory data is not restored by VM at first, which degrades somehow VM performance. Our performed experiments shows, for a VM consist of 1GB RAM, the VM runs so much slow to be useful in duration of first 10 seconds. Almost all of this time in seconds was consumed on restoring the necessary required memory data.

To pick the dual benefits of both solutions for VM resumption and to over- come their limitations, a mechanism for dual purpose is hybrid resumption called as VMresume.

Our aim is to start a VM in the running state as soon as possible, but to avoid degradation in performance caused by page faults when VM starts. Our basic aim and purpose is to examine the memory pages which have high possibility to access at the beginning period after the VM startup, restore those pages form check-point file, and then boot the VM by loading the necessary required device and CPU states. By reason of preloading all likely-to-be-accessed pages of memory, we ensure the after the startup of VM, there could not be as much as page fault found in the second solution. Also, we can ensure the earlier startup of VM compared with first solution, this is due to reason we do not preload all the data from memory which is saved in the checkpoint file before the restoring the device and CPU states. This hybrid mechanism for resumption provides some benefits in data context by ensuring the high availability of memory pages during the resumption in the running state.

So now, for the likely-to-be-accessed pages from memory, how it can be determined? Their determination can be achieved by principle of temporal locality, pages recently updated are likely to get updated in near future. Therefore, we can trust on the facts on the recent activities of memory access to forecast the upcomin activities for memory access. For suppose we get an upgraded checkpoint and now we wish to VM resume from the latest checkpoint. According to the principle temporal locality, pages which would have highest possibility to be access during the initial period, so those memory pages get accessed during the latest checkpoint interval. Thus, while receiving the checkpoint file from a checkpointing interval, we maintain a record of the memory pages recently likely to be accessed during that interval, and use that maintained record to forecast such memory pages which are likely to accessed after the VM resumption. This needs a predictive mechanism for checkpointing.

B. Virtual machine disk migration

In comparison with live migration in LANS (Local Area Networks), migration for VM possess additional challenges in WANS (Wide Area Networks). While migrating a VM within a LAN network, the storage in disk for both source and destination/target VMs are often shared by network-attached storage (NAS) or SAN (Storage area Network) media. Therefore, in LAN-based migration most part of the data that requires to be migrated is derived from run-time state of memory of VM. However, while migrating a VM within a WAN network, besides the state of memory, entire disk data, along with file system and I/O devices state, should also be migrated, this is due to they are not shared between both source and destination/target VMs. The disk data, particularly for I/O likewise applications, is commonly very large (e.g., in order of 100s of GBs). Therefore, LAN-based VM migration approaches that can only migrate the data from memory (usually in order of GBs) may not be go well with when applied to WAN-based VM migration.

A straightforward method to migrate a VM at the data context level is to suspend the VM on source host machine, then transfer the stored data in local disk memory (in the form of a self-contained image file) on the network, them towards the destination/target host machine, after then reload memory and file system to resume the VM [6]. Although, these stop-and-resume faces long downtime. In order to decrease the larger amount of data within a disk that is going to be transmitted, many optimization techniques have also been earlier introduced-i.e. data compression while migration and content-based comparison between disk data and memory [25, 26]. However, sometimes these optimization techniques introduces either computational or memory overheads. Therefore, it really needs to develop new scenarios for migration of VMs with their potentially larger file systems with acceptable overhead and minimal downtime.

In order to achieve quick live migration within WANs, the disk data with larger amount that is going to be transferred over the WAN should be reduced. Traditional migration techniques in LAN-based systems include in previous work uses checkpointing/resumption approaches discussed earlier to migrate data of memory [23, 24]. Moreover, to migrate the shared disk they use some incremental checkpointing mechanism to decrease the updated memory data that has to be transferred during each migration stage. These incremental checkpointing are often used for the virtual disk migration to gain the low downtime, but the problem occurs when larger data in disk and memory combined can still prevents in unacceptable total migration time.

C. Live migration in data context

In order to get effective live migration as well as high availability of computational resources along with rapid resource allocation VM resumption is the most important feature in virtualization. Also, it is a straightforward approach for the maintenance of data consistency in data sharing context and application context from source machine to the destination/target machine. However, by implementing the predictive checkpointing technique and hybrid solution for VM resumption i.e. VMresume, we get the data in run-time with application context without any unacceptable migration downtime. This is our main goal behind this work.

Since in contrast with previous work and citations, we aim to present a new technique based on incremental checkpointing mechanism to share the data pages during the resumption i.e. a predictive checkpointing for resumption of VM mechanism. In this mechanism, we assume VMresume, when the system initialized, first of all the complete image of data in VM memory as well as emulated devices/CPU states saved by the VMresume to on-disk file, which becomes the VM initial checkpoint. After that, it checkpoints the VM at constant and fixed frequency. Then all off the memory pages are set to be as read-only at the start of the checkpointing interval (i.e. which typically shows the time between the previous and leads to next checkpoint). Thus, if there is found any write mode in memory pages, it triggers an alert, that alert is coded for page fault. By leveraging shadow-paging feature in Xen, VMresume captures whether a page settle as read-only and tracks either it a dirty or otherwise. Whenever write mode found in a read-only page, alert triggers a page fault and it would be reported to VMM, then that page is set up as writable. Thereafter, VMresume adds the address of the triggered faulted page in the list of changed pages and discards the write mode protection from the page for the application proceeding in write. The list of changed pages which are modified during that particular interval updated, at the end of that interval. VMresume copies to the checkpoint state of all changed pages, and resets again all pages as read-only. This provides high availability of data resumption during the migration.

By using this incremental checkpointing approach, it helps to find the entire write mode accessed pages in memory during the latest checkpoint interval. These all write accessed pages are probably to get accessed after VM resumption. However, write accessed pages are often a small section of pages that are likely to get accessed after resumption of the VM. Besides this, there are some more pages in memory which are purely read accessed during the same checkpoint interval. The pages which are read accessed are not recorded by in our used checkpoint method, but they should also get preloaded while resuming the VM to decrease the potential paging faults on those pages.

Live migration in data context perspective from our work proposes a quick VM resumption by taking the snapshot/VM image. Here in our performed experiments as shown in f, A and B, we take 3 host VMs in the active state and enough configured. From the data context perspective, we assigned VMs with memory of 1GB, 2GB and 3 GB respectively. The assigned memory is al- located memory which is a complete memory of that particular VM. But, from the allocated memory some portion of memory either has data pages or otherwise. Portion of memory that has some data pages is configured memory. We proposed VMresume with the mechanism of allocated and configured memory. As discussed above about the read-accessed and write-accessed memory pages. The entire allocated memory is read-accessed while the configured memory is write accessed. While performing the Live migration from the source machine to the target/destination machine some our work focus on transferring the con- figured memory pages which has the actual data that has the high possibility to be accessed rather than the allocated memory. This solution overcomes and overheads unacceptable downtime during the migration, data duplication and trigger page faults.

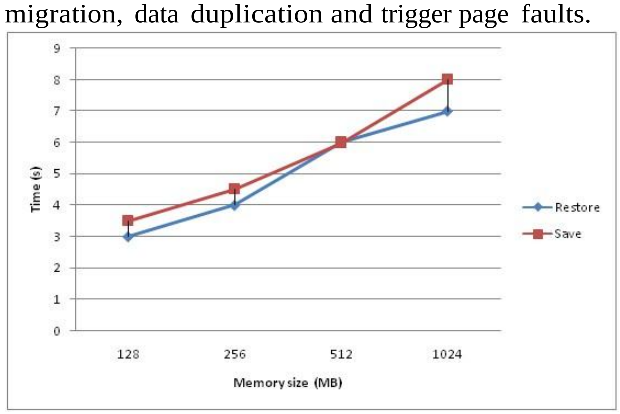


Fig. 2. Comparison of VM resumption from configured memory

Therefore, figure 2 illustrates the migration downtime needs to transfer the configured memory. The configured memory is the main focus behind our pro- posed work, which includes the facts and figures that does not involve any bad memory. All of the memory is configured means the data is also configured data (useful for the user). As shown in the figure 2, the comparison of configured data migration is meanwhile seems to be not too high, the both lines are coinciding with each other, this could happened due to the memory migration based on the configuration memory which saves the computational and migration downtime in the meanwhile which can also overcome the load balancing of VMs and also helps in performance degradation.

Additionally figure 3 illustrates comparison of VMs with allocated memory rather than the configured. The allocated memory consist all the memory storage which is assigned to a particular VM, and data within the allocated memory consist all over the memory that can be either user needed or otherwise. As shown in figure the migration downtime consumption is quite high due to big storage and raw data and the lines are too far to each other, so in the resultant it presents performance degradation due to the high storage and raw allocated migration.

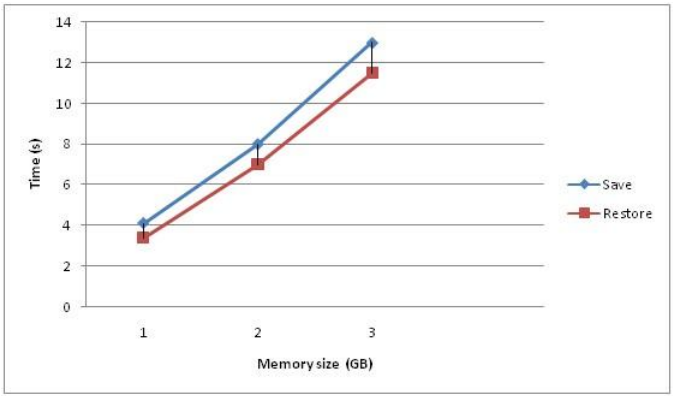


Fig. 3. Comparison of VM resumption from allocated memory

Moreover in the summary of our performed experiments aims the migration of configured memory (i.e. memory that has the actual data and write-accessed pages) rather than the allocated memory (i.e. memory that has whole VM data read-accessed and write-accessed pages), while migration of entire allocated memory results a high downtime etc. Therefore, to perform an efficient migration we propose VMresume mechanism with migration of configured memory technique. Therefore, By implementing this, at the end of checkpoint interval, for the data pages (i.e. write-accessed) in the interval, VMresume saves in checkpoint file and observes their R/W bit. For those memory pages whose A bit set to 1, then VMresume determines whether they are read-accessed or write-accessed. If pages are write-accessed, then they are already saved in checkpoint file and going to be migrated. If they are read-accessed, VMresume keeps copy of those pages read-accessed pages for future prediction purpose whenever resuming VM from corresponding checkpoint file. It is completely unnecessary to save the read-accessed memory pages contents because they are not updated and modified during migration checkpoint interval.

While resuming the VM, VMresume initially reloads all of the write-accessed pages (i.e. they are newly saved in checkpoint file), also other likely-to-accessed by tracking the record of all read-accessed pages. Then the VM will be resumed and started with required CPU/devices states along with the data resumption which is typically configured and likely-to-be-accessed with any unacceptable downtime and delay.

IV. Experiments & Observations

In this section, the performance from the proposed work and techniques (i.e. Xen Hypervisor) and estimated results are presented. We measured some overhead and migration downtime by using under FGBI, and analyzed and compared the achieved results with that under Remus and LLM.

A. Experimental setup

We have designed an experimental setup which includes two hosts. One host is primary or master and second one is used as backup. The two hosts are Intel core2 Duo processor 2.6 GHz and 4 GB RAM. The two hosts are connected through a 2 Mbps network connection. The network connection is used for migration of the Primary host to the second host.

TABLE I. Specifications of Guest VMs

Parameter / VM	Guest vm1	Guest vm2	Guest vm3
ID:	2	3	4
Name:	Testvm1	Testvm2	Testvm3
Hypervisor:	Xen	Xen	Xen
OS Type:	Hvm (Ubuntu)	Hvm (Ubuntu)	Hvm (Ubuntu)
State:	Running	Running	Running
CPU:	1	1	1
CPU Time:	11.2 s	12.5 s	14 s
Virtual Memory (RAM):	524288 KiB	524288 KiB	524288 KiB
Allocated Memory (ROM):	1048576 KiB	2097152 KiB	3145728 KiB
Disk Space:	1 GB	2 GB	3 GB
UUID:	192.168.0.60	192.168.0.30	192.168.0.20

B. Experimental results

Here VMs are migrated in two situations. The first one when there is no work load on VMs and second one when there is workload of different applications on VMs.

VM Migration with no Load Live Migration is the practice of transfer of the Virtual Machines active memory state and accurate execution state over a high- speed network, which permits the VM to shift from running on the source host to destination host. Live migration of a VM has some time parameters named as Real, User and Sys. These parameters have different values, depending on the virtual RAM and hard disk of the VM.

TABLE II. Specifications of Guest VMs

	Real	User	Sys
Testvm1	10.378s	2.068s	5.252s
Testvm2	19.789s	3.015s	9.651s
Testvm3	28.651s	3.859s	14.798s

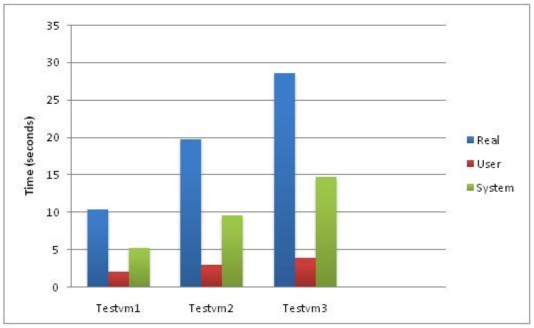


Fig. 4. Migration Time taken by XenServer in our system

Real: It is the time taken by host VM for live migration of guest VM. User: It describes the time taken by host VM for state migration of guest VM. Sys: It expresses the time which is taken by host for memory migration of guest VM.

The table 2 specifies the three migration time parameters of the entire guest VMs. The parameters specify that how much time it takes to migrate the entire VM from source to destination.

By plotting the above parameters data in pictorial form it looks like as follows:

From figure 4, one can estimate the total migration time, state migration time and memory migration time taken by the VMs in our system. The three VMs named as Testvm1, Testvm2 and Testvm3. All the VMs are of different sizes consecutively of 1GB, 2GB and 3GB. Greater the virtual hard disk is longer time is taken by Xen Hypervisor to migrate, as shown in fig 4. Here in the result, the three parameters we achieved are the Real (time taken by Data migration), user (time taken by the Network migration), system (time taken by the State migration). All the parameters and their data in accordance with time are shown meanwhile in figure 4.

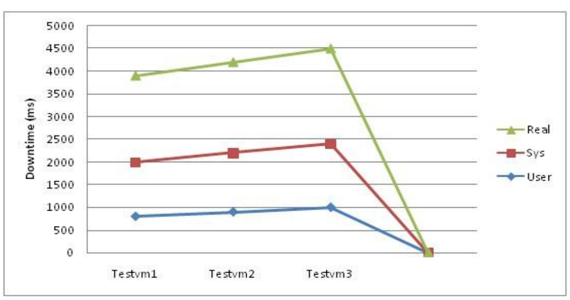


Fig. 5. Downtime of VMs with different memory size m

While the VM is migrated from source host to destination host, the both hosts will be down. In other words, the hosts will not be able to communicate or run any application. This downtime also depends on the RAM and hard disk of virtual machine (VM). The graph below (Fig 5) shows the down time of migration time parameters of the three VMs. The Testvm1 is smaller is size as compared to Testvm2 and Testvm3, therefore it takes lesser downtime of the parameters (Real, Sys and User).

Live migration consists of three parameters: migration of entire VM, state migration and memory migration. The memory migration also depends on the hard disk and RAM of VM. As mentioned earlier and our main focus is to overall migration the figure 5 states that how much time VMs take to transfer the memory state. Moreover, it is illustrated from figure 5 the individual parameters are plotted of each of all three VMs. Hence the size of each VM is different the time taken by each parameter for each depends on the size and processing speed in order to achieve the better mean time while performing the migration.

Hence the primary motivation behind this proposed work is to perform live migration the data context level. So figure 6 shows the memory and data migration of all of three VMs. Therefore, here two parameters Real and sys are plotted which typically shows the data migration and memory migration. In order to mi- grate the data between the VMs the hypervisor will check either the memory is configured or otherwise then the migration will be performed accordingly. If the hypervisor found the memory which configured and then perform migration so we can assume this strategy ensures the best results and performance.

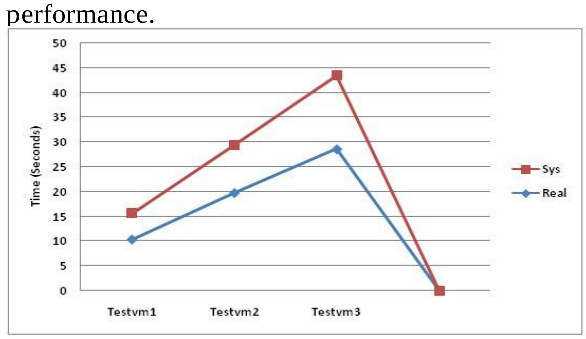


Fig. 6. Memory Migration of Xen Hypervisor

V. CONCLUSIONS AND FUTURE WORK

This work presents some set of methods and approaches for the analysis off VM live migration in data sharing perspective between the VMs in run-time. Further- more live migration in data sharing perspective without any unacceptable halt, delay and performance degradation. This happened possible with VM resumption techniques i.e. save and restore technique and VMresume hybrid solution. The design, implementation and analysis of our proposed techniques ensure the high availability with unacceptable downtime and performance degradation. This happened possible with the help of Xen hypervisors save and restore commands and VMresume mechanism. In this research we analyzed and aim to propose that, from our performed experiments and evaluations it can be analyzed that we can achieve minimal downtime with unacceptable performance degradation and high availability of data and resources by migrating the configured memory of VM (i.e. likely to be accessed memory pages or real/occupied data) rather than the allocated memory (i.e. entire assigned data may be blank space) while performing the live migration

REFERENCES

[1] What is cloud computing?? Luit Infotech

[2] Diego Perez-Botero, A Brief Tutorial on Live Virtual Machine Migration From a Security Perspective, Princeton University, Princeton, NJ, USA

[3] Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen Eric July, Christian Limpach, Ian Pratt, Andrew Warfield, Live Migration of Virtual Ma- chines, University of Cambridge Computer Laboratory 15 JJ Thomson Avenue, Cambridge, UK, Department of Computer Science University of Copenhagen, Denmark

[4] NSRC, Virtual Machine Migration

[5] Mahadev Satcyanarayanan, B. Gilbert, M. Toups, N. Tolia, D.R. OHallaron, Ajay Surie,A. Wolbach, J. Harkes, A. Perrig, D.J. Farber, M.A. Kozuch, C.J. Helfrich, P. Nath, and H.A. Lagar-Cavilla. Pervasive personal computing in an internet suspend/resume system. Internet Computing, IEEE, 11(2):1625, 2007.

[6] Constantine P. Sapuntzakis, Ramesh Chandra, Ben Pfaff, Jim Chow, Monica S.Lam, and Mendel Rosenblum. Optimizing the migration of virtual computers. SIGOPS Oper. Syst.Rev. 36(SI):377390, December 2002.

[7] Brian Keith Schmidt. Supporting ubiquitous computing with stateless consoles and computation caches. PhD thesis, Stanford, CA, USA, 2000. AAI9995216.

[8] Steven Osman, Dinesh Subhraveti, Gong Su, and Jason Nieh. The design and im- plementation of Zap: a system for migrating computing environments. SIGOPS Oper. Syst. Rev, 36:361376, December 2002.

[9] Wei Huang, Qi Gao, Jiuxing Liu, and Dhabaleswar K. Panda. High perfor- mance virtual machine migration with RDMA over modern interconnects. In CLUSTER 07:Proceedings of the 2007 IEEE International Conference on Clus- ter Computing, pages 1120, Washington, DC, USA, 2007. IEEE Computer So- ciety,

[10] Haikun Liu, Hai Jin, Xiaofei Liao, Liting Hu, and Chen Yu. Live migration of virtual machine based on full system trace and replay. In Proceedings of the 18th ACM international symposium on High performance distributed computing, HPDC 09, pages 101110, New York, NY, USA, 2009. ACM.

[11] Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen, Eric Jul, Christian Limpach, Ian Pratt, and Andrew Warfield. Live migration of virtual machines. In Proceedings of the 2nd conference on Symposium on Networked Systems Design and Implementation - Volume 2, NSDI05, pages 273286, Berke- ley, CA, USA, 2005. USENIX Association.

[12] Michael Nelson, Beng Hong Lim, and Greg Hutchins. Fast transparent migration for virtual machines. In ATEC 05: Proceedings of the annual conference on USENIX Annual Technical Conference, pages 2525, Berkeley, CA, USA, 2005. USENIX Association.

[13] Diwaker Gupta, Sangmin Lee, Michael Vrable, Stefan Savage, Alex C. Snoeren, George Varghese, Geoffrey M. Voelker, and Amin Vahdat. Difference engine: harnessing memory redundancy in virtual machines. Commun. ACM, 53:8593, October 2010.

[14] Renato J. Figueiredo, Peter A. Dinda, and Jose A. B. Fortes. A case for grid computing on virtual machines. In Proceedings of the 23rd International Con- ference on Distributed Computing Systems, ICDCS 03, pages 550,Washington, DC, USA, 2003. IEEE Computer Society.

[15] John R. Lange and Peter A. Dinda. Transparent network services via a vir- tual traffic layer for virtual machines. In Proceedings of the 16th international symposium on High performance distributed computing, HPDC 07, pages 2332, New York, NY, USA, 2007. ACM.

[16] A. Feldmann R. Bradford, E. Kotsovinos and H. Schioeberg. Live wide-area mi- gration of virtual machines including local persistent state. In VEE07: Proceed- ings of the third International Conference on Virtual Execution Environments, pages 169116, San Diego,CA, USA, 2007. ACM Press.

[17] Yoshiaki Tamura, Koji Sato, Seiji Kihara, and Satoshii Moriai. Kemari: Vir- tual machine synchronization for fault tolerance using DomT (technicalre- port).http://wiki.xen.org/xenwiki/Open Topics For Discussion?action=Attach File&do=get&target=Kemari 08.pdf, 2008.

[18] Franco Travostino, Paul Daspit, Leon Gommans, Chetan Jog, Cees de Laat, Joe Mambretti, Inder Monga, Bas van Oudenaarde, Satish Raghunath, and Phil Yonghui Wang. Seamless live migration of virtual machines over the man/wan. Future Gener. Comput. Syst.22 (8):901907, October 2006.

[19] William Voorsluys, James Broberg, Srikumar Venugopal, and Rajkumar Buyya. cost of virtual machine live migration in clouds: A performance evaluation. In Proceedings of the 1st International Conference on Cloud Computing, Cloud- Com 09, pages 254265, Berlin, Heidelberg, 2009. Springer-Verlag.

[20] Ming Zhao and Renato J. Figueiredo. Experimental study of virtual machine migration in support of reservation of cluster resources. In VTDC 07:Proceedings of the 2nd international workshop on Virtualization technology in distributed computing, pages 5:15:8, New York, NY, USA, 2007. ACM.

[21] Michael R. Hines and Kartik Gopalan. Post-copy based live virtual machine migration using adaptive pre-paging and dynamic self-ballooning. In Proceed- ings of the 2009 ACM SIGPLAN/SIGOPS international conference on Virtual execution environments, VEE 09, pages 5160, New York, NY, USA, 2009. ACM.

[22] Takahiro Hirofuchi, Hidemoto Nakada, Satoshi Itoh, and Satoshi Sekiguchi. Re- active consolidation of virtual machines enabled by postcopy live migration. In Proceedings of the 5th international workshop on Virtualization technologies in distributed computing, VTDC 11, pages 1118, New York, NY, USA, 2011. ACM.

[23] Brendan Cully, Geoffrey Lefebvre, Dutch Meyer, Mike Feeley, Norm Hutchin- son, and Andrew Warfield. Remus: high availability via asynchronous virtual machine replication. In Proceedings of the 5th USENIX Symposium on Net- worked Systems Design and Implementation, NSDI08, pages 161174, Berkeley, CA, USA, 2008. USENIX Association.

[24] Bo Jiang, Binoy Ravindran, and Changsoo Kim. Lightweight live migration for high availability cluster service. In Proceedings of the 12th international conference on Stabilization,safety, and security of distributed systems, SSS10, pages 420434, Berlin, Heidelberg,

[25] Hai Jin, Li Deng, Song Wu, Xuanhua Shi, and Xiaodong Pan. Live virtual machine migration with adaptive, memory compression. In Cluster Computing and Workshops, 2009. CLUSTER 09. IEEE International Conference on, pages 1-10,31 2009 sept 4, 2009

[26] Pierre Riteau, Christine Morin, and Thierry Priol. Shrinker: Improving live mi- gration of virtual clusters over wans with distributed data deduplication and content-based addressing. In Emmanuel Jeannot, Raymond Namyst, and Jean Roman, editors, Euro-Par 2011 Parallel Processing - 17th International Confer- ence, Euro-Par 2011, Bordeaux, France, August 29 - September 2, 2011, volume

[27] Yifeng Sun, Yingwei Luo, Xiaolin Wang, Zhenlin Wang, Binbin Zhang, aogang Chen, and Xiaoming Li. Fast live cloning of virtual machine based ceedings of the 2009 11th IEEE International Conference on High Performance Computing and Communications, pages 392399, Washington, DC, USA, 2009. IEEE Computer Society.

[28] Maohua Lu and Tzi cker Chiueh. Fast memory state synchronization for virtualization-based fault tolerance. In Dependable Systems Networks, 2009.DSN 09. IEEE/IFIP International Conference on, pages 534543, 2009

Survey of Techniques for Deep Web Source Selection and Surfacing the Hidden Web Content

Khushboo Khurana
Assistant Professor, Shri Ramdeobaba College of Engineering and Management, Nagpur, India

Dr. M. B. Chandak
Professor, Shri Ramdeobaba College of Engineering and Management, Nagpur, India

***Abstract*—Large and continuously growing dynamic web content has created new opportunities for large-scale data analysis in the recent years. There is huge amount of information that the traditional web crawlers cannot access, since they use link analysis technique by which only the surface web can be accessed. Traditional search engine crawlers require the web pages to be linked to other pages via hyperlinks causing large amount of web data to be hidden from the crawlers. Enormous data is available in deep web that can be useful to gain new insight for various domains, creating need to access the information from the deep web by developing efficient techniques. As the amount of Web content grows rapidly, the types of data sources are proliferating, which often provide heterogeneous data. So we need to select Deep Web Data sources that can be used by the integration systems. The paper discusses various techniques that can be used to surface the deep web information and techniques for Deep Web Source Selection.**

Keywords—Deep Web; Surfacing Deep Web; Source Selection; Deep Web Crawler; Schema Matching

I. Introduction

Tremendous increase in collection of web content has created new opportunities for large-scale data analysis. Working of search engine is based on index creation by crawling web pages. The web crawler retrieves the contents of web pages and parses the web pages to get the data and hyperlinks. It then continues to crawl the found hyperlinks. Parsed data is sent to the indexer and stored in database. A search is performed by referring the search engine database, consisting of web page index [1].

Most of the search engines access only the Surface Web, which is a part of web that can be discovered by following hyperlinks and downloading the snapshots of pages for including them in the search engine's index [2]. As a traditional search engine crawler requires pages to be linked to other pages via hyperlinks or the page must be static, large amount of web data is hidden. Hidden web is also referred as Deep Web.

The deep Web is qualitatively different from the surface web. The term "Deep Web" refers to web pages that are not accessible to search engines. The existing automated web crawlers cannot index these pages, thus they are hidden from the Web search engines [3].

The data in digital libraries, various government organizations, companies is available through search forms. A deep web site is a web server that provides information maintained in one or more back-end web databases, each of which is searchable through one or more HTML forms as its query interfaces [4]. Deep web consists of following types of content:

- **Dynamic Data**: Data that can only be accessed through the query interface they support. These interfaces are based on input attribute(s), and a user query involves specifying value(s) for these attributes. In response to such a query, dynamically generated HTML pages returned as the output, comprising output attributes [5].
- **Unlinked Content**: Data that is not available during link analysis done by web crawlers.
- **Non-Text Content**: Various multimedia files, PDF and non-HTML documents.

The information in the deep web is about 500 times larger than the surface web, with 7,500 Terabytes of data, across 200,000 deep web sites [6]. This wealth of information is missed since the standard search engines cannot find the information generated by dynamic sites. So, there is a need to access the data that is deep by developing efficient techniques.

II. Traditional Web Crawler

Traditional web crawlers are used to index the surface web. Fig. 1 shows the working of a traditional crawler. Initially a URL is selected as the start for the web crawler. Crawler then retrieves the web pages. From the web pages, data is extracted and resource discovery is done to extract the hyperlinks, which are further processed. Data is sent to the indexer which is used as an index during search. Hyperlinks are used as the new URL and the loop continues. Traditional crawler does not distinguish between pages with and without forms, structured and semi-structured data cannot be retrieved; hence form processing phase has to be added to the web crawler loop, to access the data present in the dynamic pages and web databases.

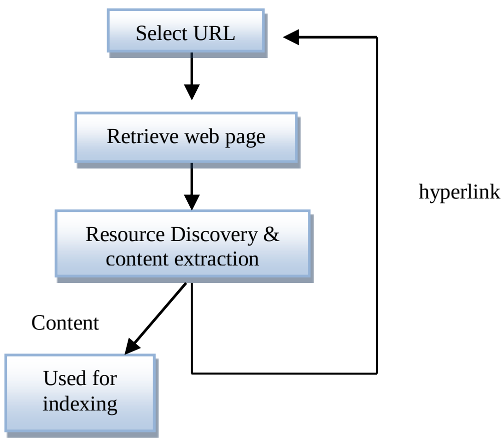


Fig. 1. Working of traditional crawler

III. Accessing the Deep Web

Deep web data can be accessed by surfacing the web data that is not accessible to the traditional search engines. Following are the major steps to access the deep web content:

Step 1: Find the Data sources.

Step 2: Data source Selection

Step 3: Send the selected data source to the Data Integration System.

Data sources for accessing the deep web may be web databases, web servers and many other dynamic web pages. Depending upon the integration system, the data sources can be added. But all sources should not be included in the Integration System. The disadvantage of including all found data sources are as follows:

- Redundant data may be added
- Irrelevant data may be added reducing the overall quality of the Data Integration System
- Low quality data can be included
- High cost of including data since, networking and processing cost are associated with including a data source in the integration system.

Various deep web source selection algorithms are discussed in section (IV). The data sources or the deep web content can be accessed by one of the following techniques as shown in fig. 2.

a) Web Surfacing by Hidden Web Crawlers by form processing and querying the Web databases:

Huge amount of data is present in the hidden web and to access this data from the deep web sites, forms need to be filled and submitted, to get the data from the web databases. Deep web crawlers are discussed in section (V). General deep web crawlers do a breadth search on the deep web, for retrieving general web data; whereas vertical deep web crawlers do a depth based search, focusing on a particular domain to extract the deep web sites based on a specific topic.

b) Schema Matching for Web source Virtual Integration system

In schema matching, instead of filling the form of the deep web site and then extracting the data to find if they are relevant to the search, a schema of the required data is prepared and only those sites which match the schema are retrieved. This technique greatly reduces the cost of extraction of web pages and then processing them. Schema matching can be done by web source virtual integration system as discussed in section (VI).

c) Data extraction by deep web search using various techniques such as Data Mining

Various techniques can be used to extract relevant information from the deep web (Refer Section VII). In Vision based approach the web page is assumed to be divided into sections that contain particular type of information. Rather than extracting the complete web page information and then parsing it, only the section that contains the relevant information is extracted using this technique.

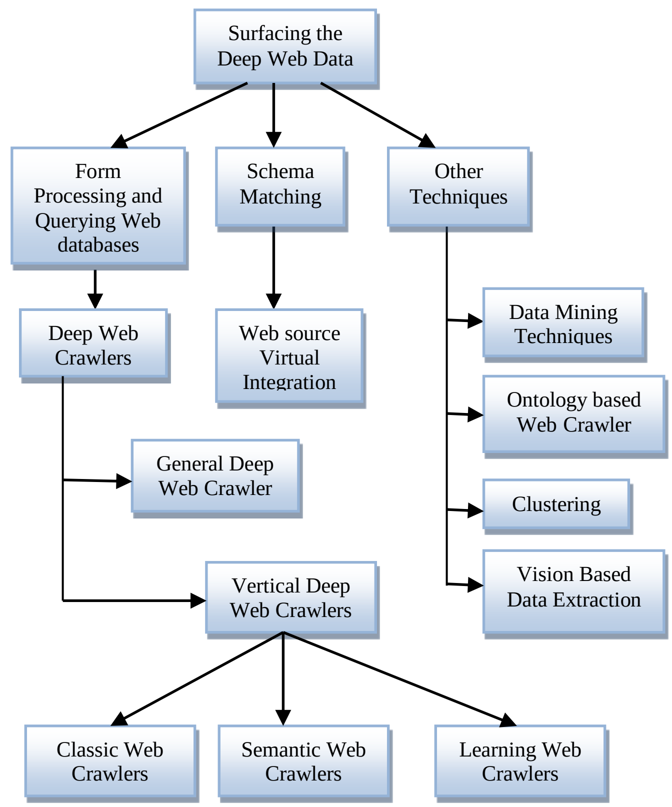


Fig. 2. Surfacing Hidden Web

IV. Data Source Selection Based on Quality Parameter

There may be hundreds and thousands of deep web data sources providing data for particular domain. The user may not want to include all the available data sources in the data integration system as there may be large number of data sources that may be of low quality. The data source selection can be broadly summarized to have following steps:

- Define quality dimensions for deep web
- Define the quality assessment model for deep web source.
- Depending on the source quality order the data sources
- Consider the highest quality set of deep web sources based on threshold.

After the web pages are extracted by the web crawler or using schema matching technique or any other technique, the web pages are checked for quality to decide whether the web page must be included or not. Fig. 3 illustrates this concept.

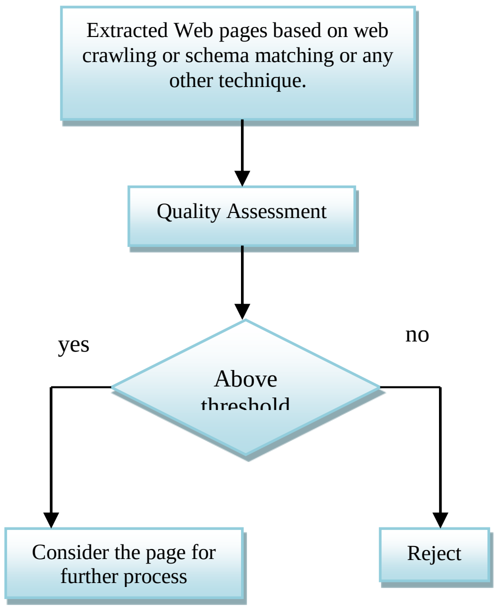


Fig. 3. Quality based Data Source Selection

In [7], an effective and scalable approach for selection of deep web source based on quality of data source is proposed for the deep web data integration system. Highest quality set of deep web sources related to particular domain are found by evaluating the quality dimensions representing the characteristics of the deep web source. Completeness, consistency, size, response time and available services are the quality dimensions considered.

The amount and type of data sources are proliferating. Data sources often provide heterogeneous or conflicting data, so we need to resolve data conflicts. There are several advanced techniques that consider accuracy of sources, freshness of sources and dependencies between sources to solve the conflicts. To improve the data fusion, a quality estimation model of Deep Web data sources (DSQ) is proposed in [8]. According to the characteristics of data fusion, the estimation model selects three dimensions of factors-data quality, interface quality and service quality as estimation criteria, and estimates the quality of data sources.

C. Hicks, et. al. [9], have proposed a new paradigm for discovery and cataloging of deep web sites. The approach divides the discovery into three phases. The first phase discovers potential deep Web sites based on a crawler configuration file for a given domain. The second phase consists of generating a set of probing queries using simple domain knowledge supplied in the query configuration file. If the submission of a probing query to a potential deep Web site is successful, the result page will be analyzed. If the result page contains the data types as specified, the site can be marked as successful identification. The third phase consists of creating a catalogue entry for that site. For this initial prototype, the catalogue entry can be as simple as the URL and the set of form parameters together with the associated values that are required for the successful query submission.

The process of data source selection can be automated by periodically analyzing different deep web sources and user can be given recommendations about a small number of data sources that will be most appropriate for their query. A data mining method to extract a high-level summary of the differences in data provided by different deep web data sources is proposed in [10]. Pattern of values are considered with respect to the same entity and a new data mining problem is formulated, referred as differential rule mining. An algorithm for mining such rules is developed. It includes a pruning method to summarize the identified differential rules. For efficiency, a hash-table is used to accelerate the pruning process.

V. DEEP WEB CRAWLER

Deep Web Crawlers are similar to traditional crawlers, but traditional crawlers do not distinguish between pages with and without forms. The results provided by the search engine are based on the copy of its local index. If additional steps are added to process pages, on which forms are detected and extraction of hidden information in databases is done then the crawler is termed as Hidden/ Deep Web Crawler [11], [4]. A user accesses the data in the Hidden Web by issuing a query through the search form provided by the web site, which in turn gives a list of links to relevant pages on the Web. The user then looks at the obtained list and follows the associated links to find interesting pages. Resource discovery and data extraction are the main task of Deep Web Crawler.

Fig. 4 shows the working of deep web crawler by addition of some extra steps. In this, the retrieved pages are checked if they have form. If form is present then it is processed to build an internal representation. Forms are filled with untried values and submitted. The returned page is then analyzed to check if a valid search result was returned. If the response page contains hypertext links, these are followed and the loop continues. Deep Web crawlers enable indexing, analysis and mining of hidden web content. The extracted content can then be used to categorize and classify the hidden databases. The Hidden Web crawler automatically process the search forms after downloading it form the hidden web site and submit the filled form so as to download the response pages which can then be used with existing index structures of the search engine.

To extract value from millions of HTML forms that span many languages and hundreds of domains, various deep web crawlers are designed. There are large numbers of forms that have text inputs and require valid input values to be submitted. In [12], an algorithm is presented for selecting input values for text search inputs that accept keywords and an algorithm for identifying inputs which accept values of specific type. HTML forms often have more than one input and hence very large number of URLs can be generated. An algorithm navigates the search space of possible input combinations to identify only those that generate URLs suitable for inclusion into the web search index.

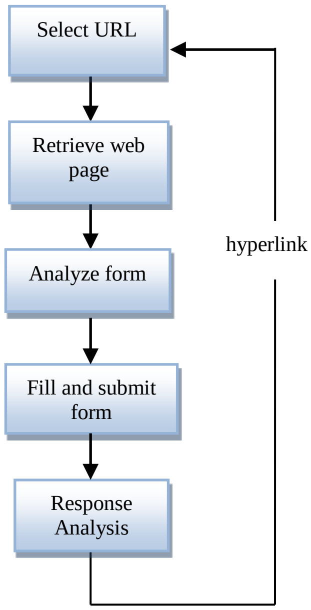


Fig. 4. Working of Deep web crawler

To meet the needs of deep web search, in [13], a new structure of crawler is designed to have innovative parts such as the mainframe extracting module and the algorithm to distinguish different websites with the same URL using improved Bayesian classification and to expand the function to AJAX form dealing. Dom Tree is also used to make easier and more visual analysis and treatment of downloaded web pages.

K.Bharati, P.Premchand, et. al., have proposed an effective design of a vertical Hidden Web Crawler that can automatically discover pages from the Hidden Web by employing multi-agent Web mining system. A framework for deep web with genetic algorithm is designed for resource discovery problem and the results show improvement in the crawling strategy and harvest rate. The focused crawler URL analysis model based on improved genetic algorithm proposed in this paper can improve accuracy rate, recall rate effectively, and avoid getting into the local optimal solution [14].

An entity extraction system, which extracts data from Deep Web automatically, is presented in [15]. A web crawler based on the characteristics of Deep Web is designed. Non- standard pages are normalized and the entity data from Deep Web are located and extracted accurately, based on the hierarchy and layout features in DOM tree, combined XPath with Regular Expression to locate entity data. Then the extracted entity attributes and attribute values are stored.

Crawling the Deep Web is requires huge amount of computing resources, but most of search engine companies hardly meet the needs. A design of the Grid-based middleware, OGSA-DWC for crawling the Deep Web is proposed in [16]. With the middleware, a Grid-based Deep Web crawling system can be implemented. It is based on two functions: Search Form Collecting and Deep Web Crawling.

A significant portion of deep web sites, including almost all online shopping sites, curate structured entities as opposed to text documents. Although crawling such entity-oriented content is clearly useful for a variety of purposes, existing crawling techniques optimized for document oriented content are not best suited for entity-oriented sites.

In [17], a prototype system is built that specializes in crawling entity-oriented deep web sites. A sub-problem is tailored to tackle important sub problems including query generation, empty page filtering and URL de-duplication in the specific context of entity oriented deep web sites. All information on the web is not in document or structured form. Multimedia data is also available is huge amount. Images can be a good source of information extraction from the deep web.

A focused crawler can miss a relevant page if there does not exist a chain of hyperlinks that connects one of the starting pages to that relevant page. Also, unless all the hyperlinks on the chain point to relevant pages, the crawler will give up searching in the relevant direction before it reaches the final target. Because of this limitation, crawlers using local search algorithms can only find relevant pages within a limited sub-graph of the Web that surrounds the starting URLs and any relevant pages outside this sub-graph will be ignored.

In [18], Tunneling technique is proposed which addresses the problems of local search. It is a heuristic based method that solves simple global optimization problem. A focused crawler using Tunneling will not give up probing a direction immediately after it encounters an irrelevant page. Instead, it continues searching in that direction for a pre set number of steps. This allows the focused crawler to travel from one relevant Web community to another when the number of irrelevant pages between them is within a limit.

Semantic Crawlers [19] are a variation of classic focused crawlers. Download priorities are assigned to pages by applying semantic similarity criteria for computing page-to-topic relevance: a page and the topic can be relevant if they share conceptually (but not necessarily lexically) similar terms. Learning Crawlers can be used to assist, visiting of web pages based on priorities. A learning crawler is supplied with a training set consisting of relevant and not relevant Web pages which is used to train the learning crawler [20] [21]. Higher visit priority is assigned to links extracted from web pages classified as relevant to the topic. Methods based on Context Graphs [22] and Hidden Markov Models (HMM) [23] can be used which consider the page content with the corresponding classification of web pages as relevant or not relevant to the topic, the link structure of the Web and the probability that a given page leading to a relevant page within a small number of steps.

Image extractor for extracting images from the result pages of deep web called AIE is proposed in [24]. Images from deep web result pages are extracted along with the images from the deep web which has no images on the result pages but has images on the detailed data record pages. The extractor can also get the images from the surface web sites which have some relations with the records on deep web. Multimedia data provide large amount of useful information. Using focused vertical crawler, image/video data can be extracted. Relevant videos can be extracted based on video annotations [25,26], based on the domain the crawler is designed for.

Some attributes in Query co-occur and some are exclusive. To generate a valid query, we have to reconcile the key attributes and their semantic relations. To address the problem, a method based on the HTML codes is presented in this paper. Different kinds of semantic containers can be got through analyzing the codes of the query interface. A Query based approach is proposed in [27].

VI. Web Source Virtual Integration

The virtual integration system, also called information mediation system [28], tries to discover relevant web sources to user's query and avoids the user to ask each web source separately using their own vocabulary. The mediation system uses a logical schema based on the source description and called the mediated schema. The source description is all the properties of the data source that the mediated schema need to know to access and to use their data.

Virtual Data Integration System creates specific virtual schema for each domain and map the fields of the search forms in that domain to the attributes of the virtual schema. This enables the user to query over all the resources in its domain of interest just by filling a single search form in the domain. Search systems using such vertical schema are called vertical search engines.

Another vision of a deep web virtual integration system uses a mediated schema built with a relational schema describing each deep web [29]. The paper proposes an approach to extract a relational schema describing a deep web source. The key idea is to analyze two structured information: the HTML Form and the HTML Table extracted from the deep web source to discover its data structure and to allow us to build a relational schema describing it. A knowledge table is also used to take profit of the learning experience on extracting relational schema from deep web source.

To automatically accomplish deep web sources discovery a method is proposed by importing focused crawler. Web sites for domain specific data sources based on focused crawling are selected. These web sites are checked if there exists deep web query interface in the former three depths. Lastly, the deep web query interface is judged to check if they are relevant to the given topic. Importing focused crawling technology makes the identification of deep web query interface locate in a specific domain and capture relative pages to a given topic instead of pursuing high overlay ratios. This method dramatically reduces the quantity of pages for the crawler to identify deep web query interfaces [30].

There are two types of virtual integration approach. The first one is the vertical search engine that integrates the same kind of information from multiple web sources like a flight ticket search engine for all flight companies. And the second one is the topical search portal that integrates all information about a particular domain. For example a topic search portal for travel will provide user data about all what concern our travel organization: flight ticketing, hotel, car rental, monuments to visit, security information etc. [31]

In [32] the authors have designed a conceptually novel approach by viewing schema matching as correlation mining, for the task of matching Web query interfaces to integrate the myriad databases on the Internet. DCM framework, which consists of data preparation, dual mining of positive and negative correlations, and finally matching selection, is proposed. The algorithm cares both positive and negative correlations, especially the subtlety of negative correlations, due to its special importance in schema matching.

Statistical Schema matching is proposed in [33]. A general statistical framework MGS for such hidden model discovery, which consists of hypothesis modeling, generation, and selection, is proposed. The algorithm targets synonym discovery and schema matching, by designing a model that specially captures synonym attributes.

Various other approaches of schema matching such as schema-only based, content based, hybrid and composite matchers are explained in [34].

Holistic Schema Matching (HSM) to find matching attributes across a set of Web database schemas of the same domain is proposed in [35]. HSM takes advantage of the term occurrence pattern within a domain and can discover both simple and complex matching efficiently with-out any domain knowledge.

VII. Data Extraction

A. Data Mining on Deep Web

Data mining on the deep web can produce important insights. For example, to show the price of electronic gadgets from different web sites and offer the customer with the site providing the requested gadget in the lowest price. Data mining on deep web must be performed based on sampling of the datasets. The samples, in turn, can only be obtained by querying the deep web databases with certain inputs. Data Mining is applied on the data that is obtained by querying the deep web database.

In [36], a stratified sampling method to support association rule mining and differential rule mining on the deep web is proposed. A pilot is selected at random from the deep web for identifying interesting rules. Then, the data distribution and relation between input attributes and output attributes are learnt from the pilot random sample. Greedy stratification approach is then applied, which processes the query space of a deep web data source recursively, and considers both the estimation error and the sampling costs. The optimized sample allocation method integrates estimation error and sampling costs.

Dasgupta et al.[6] proposed HDSampler, a random walk scheme over the query space provided by the interface, to select a simple random sample from hidden database.

A novel query-oriented, mediator based biological data querying tool, SNPMiner, is proposed in [5]. It is a domain specific search utility, which can access and collect data from the deep web. The system includes three important components, which are the web server interface, the dynamic query planner, and the web page parser. The web server interface can provide end users a unified and friendly interface. The dynamic query planner can automatically schedule an efficient query order on all available databases according to user's query request. The web page parser analyzes the layout of HTML pages and extracts desired data from those pages.

B. Clustering

Clustering can be performed on Web sources to process only domain specific content. A novel method DWSemClust, is proposed in [37] to cluster deep web databases based on the semantic relevance found among deep web forms by employing a generative probabilistic model Latent Dirichlet Allocation (LDA) for modeling content representative of deep web databases. A document comprises of multiple topics, the task of LDA is to cluster words present in the document into topics. Parameter estimation is done to discover the document's topic and tell about its proportionate distribution in documents. Deep web has a sparse topic distribution. Due to this LDA is used as a clustering solution for the sparse distribution of topics.

C. Ontology Assisted Deep Web Search

Deep web has huge amount of information, hence the number of relevant Web pages returned might be very less. It is necessary to develop a methodology that increases the number of relevant pages returned by a search. Ontologies can assist the web search, to reduce the number of irrelevant web pages returned by the search engine.

Domain ontologies can be integrated with the web search engine for efficient search. Combining of Deep Web information with ontology is suggested in [38]. The paper considers the problem of constructing domain ontologies that support users in their Web search efforts and that increase the number of relevant Web pages that are returned. A semi-automatic process to interpret information needed against the backdrop of the Deep Web is designed to utilize domain ontologies to meet Web users' needs.

In [39], a novel approach is proposed that combines Deep Web information, which consists of dynamically generated Web pages that cannot be indexed by the existing automated Web crawlers, with ontologies built from the knowledge extracted from Deep web sources. The Ontology based search is divided into different modules. The first module constructs attribute-value ontology. Second module constructs the attribute-attribute ontology. Third module formulate the user query, fills the search interface using domain ontology, extract results by looking into the index database.

G.Liu, K.Liu, et.al., [40] have put forward an automatic method for Deep Web discovery. The information from specific fields of Deep Web entry form is used to establish domain ontology, then the crawler extracts a URL from links queue as the start link, and using Bayesian Classifier to do theme classification. If the page belongs to the theme then the Form viewer module is used to check the HTML code to determine whether these have a form. If form exists in the HTML, then entry found modules are used to calculate weights between the ontology and the attributes of the form, if the value is according to the requirements, the page is downloaded.

Duplicate entity identification is done to discover the duplicate records from the integrated Web databases. However, most of existing works address this issue only between two data sources. That is, one duplicate entity matcher trained over two specific Web databases cannot be applied to other Web databases. In addition, the cost of preparing the training set for n Web databases is higher than that for two Web databases. A holistic solution to address the new challenges posed by deep Web, whose goal is to build one duplicate entity matcher over multiple Web databases is proposed in [41].

D. Visual Approach

Web information extraction based on visual approach is programming language independent. This approach utilizes the visual features of the web pages to extract data from deep web pages including data record extraction and data item extraction. There are many semi-automatic and automatic methods for visual based information extraction.

In [42], a vision-based approach is proposed to extract the structured data, including data records and data items automatically, from the deep Web pages. Given a sample deep Web page from a Web database, its visual representation is obtained and it is transformed into a Visual Block tree. Then data records are extracted from the Visual Block tree. The extracted data records are then partitioned into data items and the data items are aligned based on semantics. Finally, visual wrappers for the Web database based on sample deep is generated, such that both data record extraction and data item extraction can be preformed.

A coordinate system can be built for every Web page. The origin locates at the top left corner of the Web page. The X-axis is horizontal left-right, and the Y-axis is vertical topdown. Suppose each text/image is contained in a minimum bounding rectangle with sides parallel to the axes. Then, a text/image can have an exact coordinate (x, y) on the Web page. Here, x refers to the horizontal distance between the origin and the left side of its corresponding rectangle, while y refers to the vertical distance between the origin and the upper side of its corresponding box. The size of a text/ image is its height and width. The coordinates and sizes of texts/images on the Web page make up the Web page layout. Fig. 5a shows a popular presentation structure of deep Web pages and Fig. 5b gives its corresponding Visual Block tree. The technical details of building Visual Block trees can be found in [43]. An actual Visual Block tree of a deep Web page may contain hundreds even thousands of blocks.

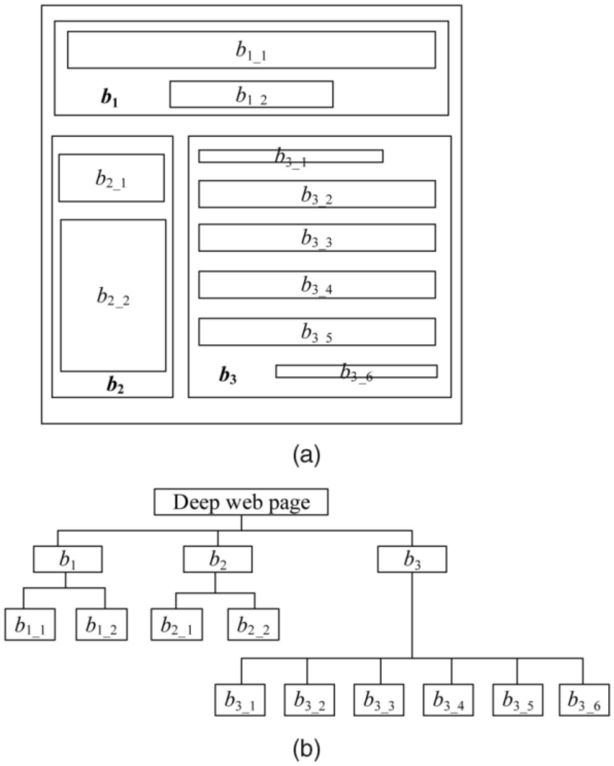


Fig. 5. (a) The presentation structure and (b) its Visual Block tree (Referred from [43])

A page division method is proposed in [44], which divides the pages into separate parts, after analyzing source codes and visual information of pages, into several segments by applying block-division algorithm. After that the parts which don't contain search interfaces are removed. At last topic-specific queries are constructed to obtain results and distinguish deep web interfaces by analyzing the results.

Some of these approaches perform only data record extraction but not data item extraction, such as Omini [45], RoadRunner [46]. These methods do not generate wrappers, i.e., they identify patterns and perform extraction for each Web page directly without using previously derived extraction.

Similar structures are recognized when comparing the differences between two Web pages in [47]. Visual structural information of Web pages are recognized. The technique is based on a classification of the set of html - tags which is guided by the visual effect of each tag in the whole structure of the page. This allows translating the web page to a normalized form where groups of html tags are mapped into a common canonical one. A metric to compute the distance between two different pages is also introduced.

VIII. COMPARATIVE ANALYSIS

Table 1 shows the comparative analysis of the two techniques widely used for surfacing the hidden web-form processing and querying the deep web by Hidden Web Crawlers and Schema Matching for Virtual Integration systems. The comparison is based on various parameters such as- technique, type, usefulness, main challenges, cost, storage requirement, advantages, and limitations with their solutions, as discussed in the previous sections.

The major requirement for any of these systems is huge computing requirement. For this Grid based middleware can be used as discussed in section-V.

TABLE I. COMPARATIVE ANALYSIS OF SURFACING TECHNIQUES

Parameter	**Form processing and querying the deep web by Hidden Web Crawlers**		**Schema Matching for Virtual Integration systems**	
Technique	1. Give the initial URL to start the process 2. Extract the pages 3. Fill the form and submit 4. Extract the data and index them.		1. Construct the schema based on the requirement. 2. Extract those sites which match the schema 3. Send the web pages to virtual integration system	
Types	**General deep web crawlers**	**Vertical / focused web crawlers**	Various techniques exist for schema matching such as logical, relational, statistical, co-relation based schema matching, element or structure based, content based, etc.	
			Vertical search engine	**Topical search engine**
	They perform a breadth search on the deep web, for retrieving general web data.	They perform a depth based search, focusing on a particular domain to extract the deep web sites based on a specific topic.	Integrates the same kind of information from multiple web sources.	Integrates all information about a particular domain.
Use	Can be used when surfacing for generalized search engine on web rather than a domain specific search engine.	Vertical crawling can be used to generate data for an individual user.	Used for specific data extraction.	
Main challenges / issues.	1. Decide which form inputs to fill when submitting queries to a form 2. To find appropriate values to fill in these inputs. 3. It needs to determine the relevance of a retrieved web page. 4. Design an algorithm that balances the trade-off between number of URLs	3. Predict and identify potential URLs that can lead to relevant pages. 4. Rank and order	1. Design an approach to extract the deep web source description needed by the mediated schema or extract a relational schema describing the deep web source. 2. Design a virtual integration system that accepts the web pages based on the schema.	

	and to achieve high coverage of the site's content.	the relevant URLs so the crawler knows exactly what to follow next.	
Cost of web page extraction	Cost of web page extraction is high because in this method the all the deep web pages are extracted following the link analysis.	Cost is less than general web crawler since only pages related to a domain are extracted. Also the probability of an unvisited page being relevant or not is calculated before actually downloading the page.	This technique greatly reduces the cost of extraction of web pages.
Processing cost	All the extracted pages are analyzed to remove the irrelevant data, hence processing cost is more.		Processing cost is less
Storage Requirement	Very high storage is required	Compared to general crawler, the storage requirement is reduced.	Very less storage. Only schema needs to be stored and the extracted pages.
Advantages	Starting from popular seed pages, leads to collecting large-Page Rank pages early in the crawl [48]	Higher density of value pages	Maximum relevant pages are retrieved based on the schema.
Limitations	1. HTML forms typically have more than one input and hence a naive strategy of enumerating the entire Cartesian product of all possible inputs can result in a very large number of URLs being generated. 2. Crawling too many URLs will drain the resources of a web crawler preventing the good URLs from getting crawled, and posing an unreasonable load on web servers hosting the HTML forms. 3. Large number of empty result pages	1. Focused crawlers have a limitation of local search because it may not follow a path that does not have relevant content. (as discussed in section V) 2. Classical focused crawler fail to associate documents with semantically similar but lexically different terms	1. Schema generation for a large domain can be time consuming 2. Dynamic changes, so greedy algorithms may fail. 3. Simple matching 4. Domain knowledge required
Probable Solutions	1. URL de-duplication 2. Use DOM trees and pruning method to reduce the number of URLs to be searched. 3. Combine techniques discussed in section (VII) to improve. Ontologies can be combined with any type of deep web crawler to get only relevant results. 4. Data mining technique, association rule mining, clustering, etc can be used. 5. Tunneling can be a solution to limitation-1 of focused web crawler. Other techniques discussed in section V. 6. Semantic focused crawlers and Learning crawlers can be used for better results. (for limitation-2 of focused crawler)		1. Faster design process for schema and updating to accommodate the dynamic nature. 2. Holistic Schema Matching (HSM), DCM and other techniques discussed in section VI.

IX. Conclusion

The paper discusses the way to extend the traditional web crawlers to surface the Deep Web. Hidden Web content can be accessed by Deep Web Crawlers that can fill and submit forms to query the online databases for information extraction. In this technique the extracted content is analyzed to check if it is relevant. Schema Matching has proved to be an efficient technique for extracting relevant content. Data from the Deep Web can be extracted by applying various techniques such as mining, building ontology to assist domain specific data retrieval. Visual approach is an efficient technique to extract only the required data. The paper also shows the comparative analysis of the two techniques widely used for surfacing the hidden web form processing and querying the deep web by Hidden Web Crawlers and Schema Matching for Virtual Integration systems. Depending upon the application area the surfacing technique can be selected and be combined with other techniques to overcome the drawbacks in the original method.

References

[1] H. T. Yani Achsan, W. C. Wibowo, "A Fast Distributed Focused-web Crawling," Procedia Engineering 69 (2014), pp. 492-499.

[2] M. Bergman, "The deep Web: surfacing hidden value", in the Journal Of Electronic Publishing 7(1) (2001).

[3] Y.J. An, J. Geller, Y.T. Wu, S. Chun, "Automatic Generation of Ontology from the Deep Web," in Database and Expert Systems Applications, 2007. DEXA'07. 18th International Workshop, IEEE, pp. 470-474.

[4] A. Ntoulas, P. Zerfos, J.Cho., "Downloading Textual Hidden Web Content Through Keyword Queries," in Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL'05, Denver, USA, Jun 2005 IEEE , pp. 100-109.

[5] F.Wang, G.Agrawal, R. Jin, and H. Piontkivska, "Snpminer: A domain-specific deep web mining tool," In Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering, 2007. BIBE 2007, IEEE, pp. 192-199.

[6] A. Dasgupta, X. Jin, B. Jewell, N. Zhang, and G. Das, "Unbiased Estimation of Size and Other Aggregates Over Hidden Web Databases," in SIGMOD '10, Proceedings of the 2010 international conference on Management of data, New York, NY, USA, 2010, ACM, pp. 855–866.

[7] X.F. Xian, P.P. Zhao, W. Fang, J. Xin, "Quality Based Data source selection for Web-scale Deep Web Data Integration," in Machine Learning and Cybernetics, 2009 International Conference,IEEE, Vol. 1, pp. 427-432.

[8] M Sun, H Dou, Q Li, Z Yan, "Quality Estimation of Deep Web Data Sources for Data Fusion," Procedia Engineering 29 (2012), pp. 2347-2354.

[9] C. Hicks, M. Scheffer, A.H. Ngu, and Q.Z. Sheng, "Discovery and cataloging of deep web sources," in Information Reuse and Integration (IRI), 2012 IEEE 13th International Conference ,pp. 224-230.

[10] T Liu, F Wang, J Zhu, G Agrawal, "Differential Analysis on Deep Web Data Sources," In Data Mining Workshops (ICDMW), 2010 IEEE International Conference, pp. 33-40.

[11] L. Barbosa, J. Freire , "Siphoning hidden-web data through keyword based interfaces", in SBBD, 2004, Brasilia, Brazil, pp. 309-321.

[12] J Madhavan, D Ko, Ł Kot,V. Ganapathy, "Google's Deep-Web Crawl," Proceedings of the VLDB Endowment,1(2),2008,pp. 1241-1252.

[13] W. Ma, X. Chen, and W. Shang. "Advanced deep web crawler based on Dom," in Fifth International Joint Conference on Computational Sciences and Optimization (CSO), 2012, IEEE, pp. 605-609.

[14] K. F. Bharati, P. Premchand, A. Govardhan, K. Anuradha, and N. Sandhya, "A Framework for Deep Web Crawler Using Genetic Algorithm," International Journal of Electronics and Computer Science Engineering, IJECSE, 2(2), pp.602-609.

[15] H Yu, JY Guo, ZT Yu, YT Xian, "A Novel Method for Extracting Entity Data from Deep Web Precisely," in Control and Decision Conference (2014 CCDC), The 26th Chinese, IEEE, pp. 5049-5053.

[16] J. Song, D.H. Choi, Y.J. Lee, "OGSA-DWC: A Middleware for Deep Web Crawling Using the Grid," in IEEE Fourth International Conference on eScience, 2008, pp. 370-371.

[17] Y. He, D. Xin, V. Ganti, S. Rajaraman, N. Shah, "Crawling Deep Web Entity Pages," in Proceedings of the sixth ACM international conference on Web search and data mining, pp. 355-364.

[18] D. Bergmark, C. Lagoze and A. Sbityakov, "Focused Crawls, Tunneling, and Digital Libraries," in Proceedings of the 6th European Conference on Digital Libraries, Rome, Italy, 2002.

[19] M. Ehrig, A. Maedche, "Ontology-Focused Crawling of Web Documents". Proc. of the Symposium on Applied Computing (SAC 2003). March 9-12, 2003.

[20] G. Pant and P. Srinivasan, "Learning to Crawl: Comparing Classification Schemes". ACM Transactions on Information Systems (TOIS). 23(4), 2005, pp.430-462.

[21] Li, Jun, K. Furuse, and K. Yamaguchi, "Focused Crawling by Exploiting Anchor Text Using Decision Tree". Proceedings of the 14th International World Wide Web Conference. 2005, pp. 1190-1191.

[22] M. Diligenti, F. Coetzee, S. Lawrence, C. Giles and M. Gori, "Focused Crawling Using Context Graphs.". Proc. 26th International Conference on Very Large Databases (VLDB 2000). 2000, pp. 527–534.

[23] H. Liu, J. Janssen, and E. Milios, "Using HMM to Learn User Browsing Patterns for Focused Web Crawling". Data & Knowledge Engineering. 59(2), 2006, pp.270–29.

[24] J. Li, D. Shen, and Y. Kou, "AIE: An Automatic Image Extractor for Deep Web and Surface Web," in Web Information Systems and Applications Conference (WISA), 7th, IEEE, 2010,pp. 137-141.

[25] K. Khurana, and M.B. Chandak, "Video annotation methodology based on ontology for transportation domain," International Journal of Advanced Research in Computer Science and Software Engineering, 3(6), 2013, pp. 540-548.

[26] K. Khurana, and M.B. Chandak, "Study of Various Video Annotation Techniques," International Journal of Advanced Research in Computer and Communication Engineering, 2(1),pp. 909-914.

[27] H Liang, J Chen, W Zuo, Y Mao, "Generating the Semantic Containers for the Query Interfaces of Deep Web," In Management and Service Science, 2009. MASS'09. International Conference, IEEE, pp. 1-4.

[28] J. Madhavan, S. Jeffery, S. Cohen, X. Dong, D. Ko, et.al., "Web-scale data integration: You can only afford to pay as you go," CIDR, January, 2007.

[29] Y. Saissi, A. Zellou, A. Idri, "Extraction of relational schema from deep web sources: a form driven approach," in Complex Systems (WCCS), 2014 Second World Conference, pp. 178-182.

[30] Y. Wang, W. Zuo, T. Peng, F. He, "Domain-Specific Deep Web Sources Discovery," in Natural Computation, 2008, Fourth International Conference, ICNC'08, IEEE, Vol. 5, pp. 202-206.

[31] A. Doan, A. Halevy, Z. Ives, Principles of Data Integration, Elsevier, 2012.

[32] B. He, K. Chen-Chang, and J. Han, "Discovering complex matchings across web query interfaces: a correlation mining approach," Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004, pp. 148–157.

[33] B. He, K. Chen-Chang, "Statistical schema matching across web query interfaces," in Proceedings of the 2003 ACM SIGMOD international conference on Management of data, 2003, pp. 217–228.

[34] E. Rahm and P. A.Bernstein, "A survey of approaches to automatic schema matching," The International Journal on Very Large Data Bases, vol. 10,2001, pp. 334–350.

[35] W. Su, J. Wang, F. Lochovsky, "Holistic Schema Matching for Web Query Interface", Advances in Database Technology-EDBT 2006. Springer Berlin Heidelberg, 2006. pp. 77-94.

[36] T. Liu, F.Wang, G. Agrawal. "Stratified Sampling for Data Mining on the Deep Web." Frontiers of Computer Science 6.2, 2012, pp. 179-196.

[37] U. Noor, et.al., "Latent Dirichlet Allocation Based Semantic Clustering of Heterogeneous Deep Web Sources," in Intelligent Networking and Collaborative Systems (INCoS), 2013 5th International Conference, IEEE, pp. 132-138.

[38] Y.J. An, S. Chun, K. Huang, J. Geller, "Assessment for Ontology-supported Deep Web Search," In E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services, 2008 10th IEEE Conference, pp. 382-388.

[39] A .K. Sharma, "Accessing the Deep Web Using Ontology," in Emerging Trends in Engineering and Technology (ICETET), 2010 3rd International Conference, pp. 565-568.

[40] G Liu, K Liu, Y Dang, "Research on discovering Deep web entries Based ontopic crawling and ontology," in Electrical and Control Engineering (ICECE), 2011 International Conference, IEEE, pp. 2488-2490.

[41] W. Liu, X. Meng, "A Holistic Solution for Duplicate Entity Identification in Deep Web Data Integration," In Semantics Knowledge and Grid (SKG), 2010 Sixth International Conference, IEEE, pp. 267-274.

[42] S.J. Pusdekar, S.P. Chhaware, "Using Visual Clues Concept for Extracting Main Data from Deep Web Pages," In Electronic Systems, Signal Processing and Computing Technologies (ICESC), 2014 International Conferenc, IEEE, pp. 190-193.

[43] W Liu, X Meng, W Meng, "ViDE: A Vision-Based Approach for Deep Web Data Extraction," Knowledge and Data Engineering, IEEE Transactions on, 22(3), pp. 447-460.

[44] X. Du, Y. Zheng, Z. Yan, "Automate Discovery of Deep Web Interfaces," in Information Science and Engineering (ICISE), 2010 2nd International Conference, IEEE, pp. 3572-3575.

[45] D. Buttler, L. Liu, and C. Pu, "A Fully Automated Object Extraction

System for the World Wide Web," Proc. Int'l Conf. Distributed Computing Systems (ICDCS), 2001, pp. 361-370.

[46] V. Crescenzi, G. Mecca, and P. Merialdo, "RoadRunner: Towards Automatic Data Extraction from Large Web Sites," Proc. Int'l Conf. Very Large Data Bases (VLDB),2001, pp. 109-118.

[47] M. Alpuente, D. Romero, "A Visual Technique for Web Pages Comparison," Electronic Notes in Theoretical Computer Science, 235, 2009, pp. 3-18.

[48] M. Najork & J. L. Wiener, "Breadth-first crawling yields high-quality pages," In Proceedings of the 10th international conference on World Wide Web, ACM, April 2001, pp. 114-118.

Optimization of Channel Coding for Transmitted Image Using Quincunx Wavelets Transforms Compression

Mustapha Khelifi
Department of Electrical Engineering, Tahri Mohammed University-Bechar / lab: TIT, Street independence bp 417
Bechar Algeria

Abdelmounaim Moulay Lakhdar
Department of Electrical Engineering, Tahri Mohammed University-Bechar/ lab: TIT, Street independence bp 417
Bechar Algeria

Iman Elawady
Department of Electrical Engineering, Tahri Mohammed University-Bechar, Street independence bp 417
Bechar Algeria

***Abstract*—Many images you see on the Internet today have undergone compression for various reasons. Image compression can benefit users by having pictures load faster and webpages use up less space on a Web host. Image compression does not reduce the physical size of an image but instead compresses the data that makes up the image into a smaller size. In case of image transmission the noise will decrease the quality of recivide image which obliges us to use channel coding techniques to protect our data against the channel noise. The Reed-Solomon code is one of the most popular channel coding techniques used to correct errors in many systems ((Wireless or mobile communications, Satellite communications, Digital television / DVB,High-speed modems such as ADSL, xDSL, etc.). Since there is lot of possibilities to select the input parameters of RS code this will make us concerned about the optimum input that can protect our data with minimum number of redundant bits. In this paper we are going to use the genetic algorithm to optimize in the selction of input parameters of RS code acording to the channel conditions wich reduce the number of bits needed to protect our data with hight quality of received image.**

Keywords—Code Rate; Optimization; Quincunx Wavelets Transforms compression; Genetic Algorithm; BSC channel; Reed-Solomon codes

I. Introduction

In information theory and coding theory with applications in computer science and telecommunication, error detection and correction or error control are techniques that enable reliable delivery of digital data over unreliable communication channels. Many communication channels are subject to channel noise, and thus errors may be introduced during transmission from the source to a receiver. Error detection techniques allow detecting such errors, while error correction enables reconstruction of the original data in many cases.

The purpose of channel coding theory is to find codes which transmit quickly, contain many valid code words and can correct or at least detect many errors. While not mutually exclusive, performance in these areas is a trade off. So, different codes are optimal for different applications. The needed properties of this code mainly depend on the probability of errors happening during transmission.

The Reed–Solomon codes are efficient error-correcting codes that used in many important applications. The input parameters can provide good results on the level of receiver however this will maximize the number of redundant bits which make us wonder about what is the efficient coding rate that have minimum number of redundant bits.

The objective here is to develop the selection of inputs parameters of RS code according to the channel noise using artificial intelligence technique (GA). The paper is organized as follow: the first part introduce the QWT and vector quantization as bases for compression, in the second part the RS coding with main input paramters, than how to creat the data base of the coding rate acording to inputs conditions (limitations), after that introduction to the main bases of GA, the formulation of solving steps, finally the simulation results and conclusion.

II. Quaternion Wavelet Transform

The foundation of the 2-D dual-tree QWT rests on the quaternion definition of the 2-D HT and analytic signal [1]. By organizing the four quadrature components of a real wavelet as a quaternion, we obtain a 2-D analytic wavelet and their associated quaternion wavelets transform (QWT).

Each quaternion wavelet consists of a standard DWT tensor wavelet plus three additional real wavelets obtained by 1-D Hilbert transforms along either or both coordinates. More specifically, if we denote the 1-D Hilbert transform operators along the x and y coordinates by Hx and Hy, respectively, then given the usual real tensor product wavelet $\psi_h(x)\,\psi_h(y)$ (for the diagonal subband in this case), we complement it with:

$$Hx\{\,\psi h(x)\,\psi h(y)\} = \psi g(x)\,\psi h(y) \qquad (1)$$

$$Hy\{\,\psi h(x)\,\psi h(y)\} = \psi\, h(x)\,\psi g(y) \qquad (2)$$

$$HyHx\{\,\psi h(x)\,\psi h(y)\} = \psi g(x)\,\psi g(y) \qquad (3)$$

Conveniently, each component can be computed as a combination of 1-D dual-tree complex wavelets. Using quaternion algebra, we can organize the four wavelet components to obtain a quaternion wavelet

$$\psi q(x,y) = \psi h(x)\,\psi h(y) - j1\,\psi g(x)\,\psi h(y) - j2\,\psi h(x)\,\psi g(y) + j3\,\psi g(x)\,\psi g(y) \quad (4)$$

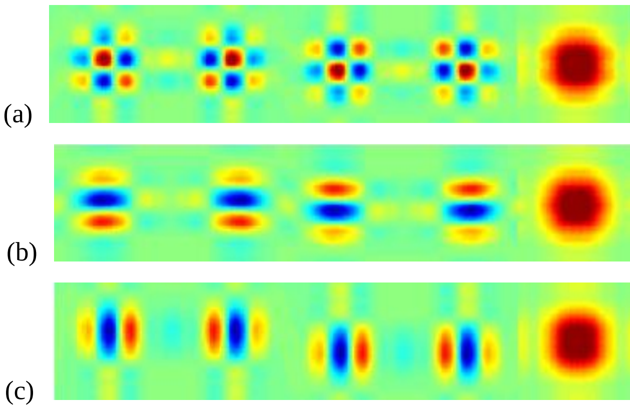


Fig. 1. Each quaternion wavelet basis contains four components that are 90° phase shifts of each other in the vertical, horizontal, and both directions. (a) The set from the diagonal subband, from left to right: ψh(x) ψh(y) (a usual, real-valued tensor wavelet), ψh(x) ψg(y), ψg(x) ψh(y), ψg(x) ψg(y). The image on the far right is the quaternion wavelet magnitude | ψq(x, y)|, a non-oscillatory function, which implies the shift-invariance of the QWT [2]. (b) The set from the horizontal subband, from left to right: Φh(x) ψh(y) (a usual, real-valued tensor wavelet), Φh(x) ψg(y), Φg(x) ψh(y), Φg(x) ψg(y). (c) The set from the vertical subband, from left to right: ψh (x) Φh(y) (a usual, real-valued tensor wavelet), ψh(x) Φg(y), ψg(x) Φh(y), ψg(x) Φg(y)

Note the 90° phase shift of the components relative to each other. The magnitude of the quaternion wavelet | $\psi^q(x, y)$| (square root of the sum-of-squares of all four components) is non-oscillatory. The construction and properties are similar for the other two subband quaternion wavelets based on $\Phi_h(x)\psi_h(y)$ and $\psi_h(x)\Phi_h(y)$ (see Figure (b) and (c)).

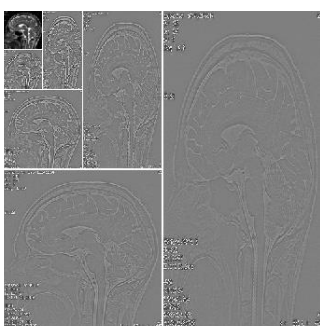


Fig. 2. Image decompostion using QWT 3 levels

III. VECTOR QUANTIZATION

A quantizer simply reduces the number of bits needed to store the transformed coefficients by reducing the precision of those values [3]. The basic principle of vector quantization based image compression techniques is to match each input vector with a code-vector in the codebook so that the distortion between the input vector and the chosen code-vector is minimum [4] by evaluating the Euclidean distance between the input vector and each codeword in the codebook [5] . Once the closest codeword is found, the index of that codeword is sent through a channel .When the encoder receives the index of the codeword, it replaces the index with the associated [6]. Quantization is an irreversible process. That is, in general, there is no way to find the original value from the quantized value [7]. The difference between the input and output signals of the quantizer becomes the quantizing error, or quantizing noise [8].

IV. RS CODE

In 1960, Irving Reed and Gus Solomon published a paper in the Journal of the Society for Industrial and Applied Mathematics [9]. Reed-Solomon codes are nonbinary cyclic codes with symbols made up of m-bit sequences, where m is any positive integer having a value greater than 2. R-S (n, k) codes on m-bit symbols exist for all n and k for which:

$$0 < k < n < 2m + 2 \quad (5)$$

Where k is the number of data symbols being encoded, and n is the total number of code symbols in the encoded block. For the most conventional R-S (n, k) code

$$(n,k) = (2m - 1, 2m - 1 - 2t) \quad (6)$$

Where t is the symbol-error correcting capability of the code, and n - k = 2t is the number of parity symbols.

With m ≤ 16 and t ≤ 16

V. CODE RATE

The rate of a block code (code rate R) is defined as the ratio between its message length and its block length.

In case of using RS code the code rate will be R=k/n it should be positive and taking values between 0 and 1

Where n = 2^m – 1(Codeword length) and k= N - 2×t (Message length, or information bytes length)

So k should be positive and n should be > 0 .The question her is how to select m and t values to respect those conditions.

We define a new function f that gather all this conditions where f= (2^m)-(2×t)-1.

In this case f should be > 0 also there are some exceptions which are m and t doesn't take the values 2 or 1 in the same time.

By taking in consideration the value of m and t in RS code (should be less than or equal 16) we get 127 possibility without repetition those code rates are representing the data base to select the optimum code rate using GA

VI. TRANSMISSION CHANNEL

In this section, it is explained the results of research and at the same time is given the comprehensive discussion. Results can be presented in figures, graphs, tables and others that make the reader understand easily [10], [11]. The discussion can be made in several sub-chapters. The binary symmetric channel (BSC) is defined by the channel diagram shown in Figure 3, and its channel matrix is given by Eq. 7:

$$[P(y/x)] = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \quad (7)$$

The channel has two inputs (x1 = 0, x2 = 1) and two outputs (y1 = 0, y2 = 1). The channel is symmetric because the probability of receiving a 1 if a 0 is sent is the same as the probability of receiving a 0 if a 1 is sent. This common transition probability is denoted by p [12]. The error events are also independent of the data bits [13]. This is the simplest model of a channel with errors, yet it captures most of the complexity of the general problem [14]. The capacity of this channel given by Eq. 8:

$$C = 1 - H(p) \text{ in } \frac{\text{bit}}{\text{channel use}} \quad (8)$$

With the binary entropy function given by Eq. 9:

$$H(p) = -p\log_2(p) - (1-p)\log_2(1-p) \quad (9)$$

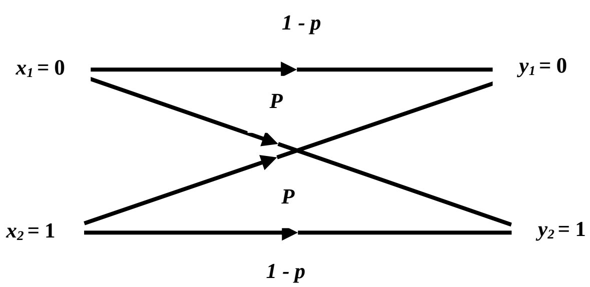


Fig. 3. Binary symmetric channel

VII. Genetic Algorithm

Genetic Algorithm (GA) is a calculus free optimization technique based on principles of natural selection for reproduction and various evolutionary operations such as crossover, and mutation. Various steps involved in carrying out optimization through GA are described. The following outline summarizes how the genetic algorithm works:

- The algorithm begins by creating a random initial population.
- The algorithm then creates a sequence of new populations. At each step, the algorithm uses the individuals in the current generation to create the next population. To create the new population, the algorithm performs the following steps:
  - Scores each member of the current population by computing its fitness value
  - Scales the raw fitness scores to convert them into a more usable range of values
  - Selects members, called parents, based on their fitness.
  - Some of the individuals in the current population that have lower fitness are chosen as elite. These elite individuals are passed to the next population.
  - e. Produces children from the parents. Children are produced either by making random changes to a single parent—mutation—or by combining the vector entries of a pair of parents—crossover.
  - Replaces the current population with the children to form the next generation.
- The algorithm stops when one of the stopping criteria is met.

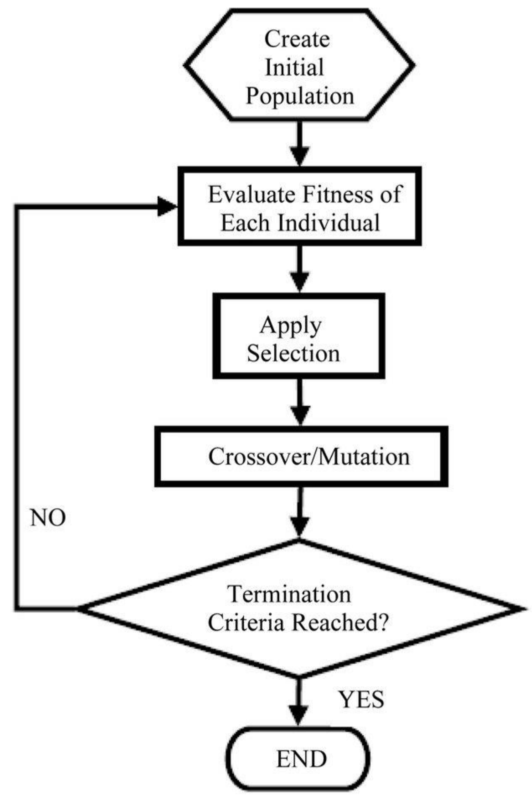


Fig. 4. The flow chart of a general Genetic Algorithm [15]

VIII. Formulation of Solving Steps

The genetic algorithm will select randomly the coding rate from the table (127 value). According to the selected value of code rate m and t will be selected. After that the selected paramters will applied the in the Reed–Solomon code, then a check of the effectiveness selected parameters will be made by calculating the MSE after transmission using BSC channel and RS decoder. The algorithm will stop if there isn't a big improvement in the optimum MSE (ε ≤ 10^-3)

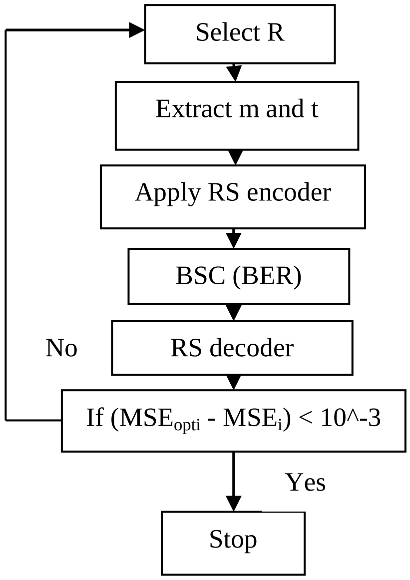


Fig. 5. Flowchart searching the optimum CODE RATE using GE

IX. Simulation Results

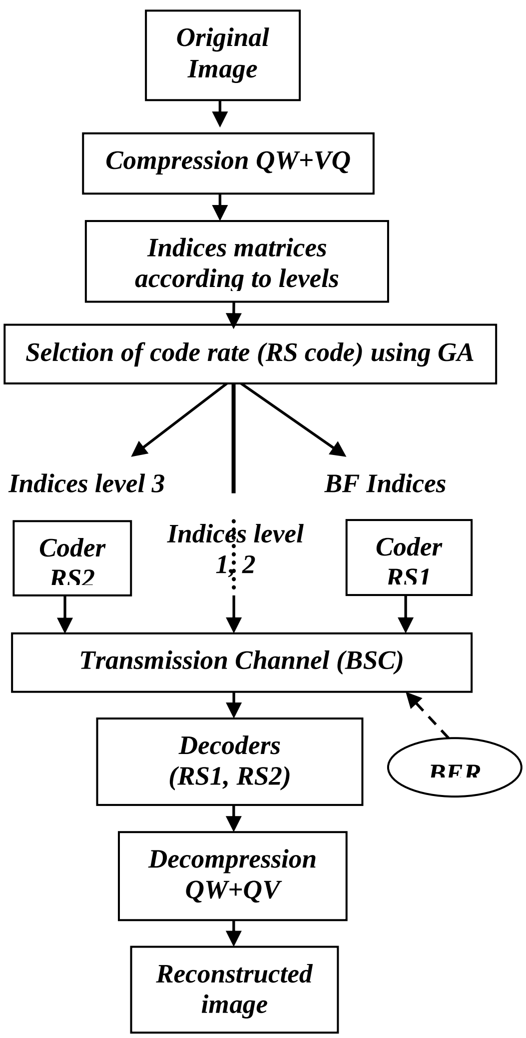


Fig. 6. Unequal error protection flowchart (UEP)

For the simulation we choose to apply the protection in base frequencies (Bf) since it represents the most important data in image and level 3 which takes the second place in terms of important. The rest of levels will be transmitted without channel coding since they are less important.

The input parameters of GA (during the simulation) are selected as follow:

PopulationSize: 20, EliteCount: 2, Crossover Fraction: 0.8000, Migration Interval: 20, Migration Fraction: 0.2000, Generations: 100, PopulationSize=20, Penalty Factor: 100, The range of the individuals in the initial population] 0, 1[.

The image is decomposed in 3 levels using QWT. For VQ we use a codebook generated by LBG size (256×16) block size 4×4. We use three gray level images (Lena, Goldhill, boat) size 512×512.

Fig. 7. MSE_r = 613.7287 PSNR_r = 20.7027 dB, BER=10^-1

Fig. 8. MSE_r = 17.0479 PSNR_r = 35.8141 dB, BER=10^-2

Fig. 9. MSE_r = 17.0479 PSNR_r = 35.8141 dB, BER=10^-3

TABLE I. INPUT PARAMETERS OF RS CODE IN DIFFERENT BER (IMAGE LENA)

BER	Input parameters of rs code					
	Optimum code rate (1)	***m (1)***	***t (1)***	***Optimum code rate (2)***	***m (2)***	***t (2)***
10^-1	0.2000	4	6	0.2000	4	6
10^-2	0.3000	4	5	0.5000	5	7
10^-3	0.3000	4	5	0.8000	4	1

Fig. 10. MSE_r = 553.3593 PSNR_r = 20.5291 dB, BER=10^-1

Fig. 11. MSE_r = 27.9659 PSNR_r = 33.6645 dB, BER=10^-2

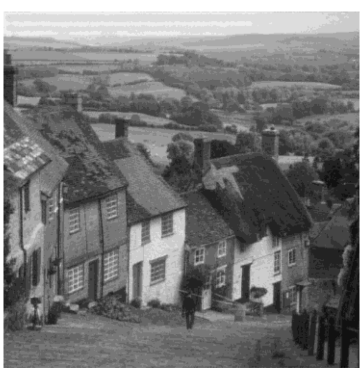


Fig. 12. MSE_r = 29.5101 PSNR_r = 33. 6645 dB, BER=10^-3

TABLE II. INPUT PARAMETERS OF RS CODE IN DIFFERENT BER (IMAGE GOLDHILL)

BER	Input parameters of rs code					
	Optimum code rate (1)	***m (1)***	***t (1)***	***Optimum code rate (2)***	***m (2)***	***t (2)***
10^-1	0.2000	4	6	0.2000	4	6
10^-2	0.3000	4	5	0.5000	5	7
10^-3	0.3000	4	5	0.6000	4	3

Fig. 13. MSE_r = 533.6945 PSNR_r = 20.8579, BER=10^-1

Fig. 14. MSE_r = 15.8678 PSNR_r = 36.1256, BER=10^-2

Fig. 15. MSE_r = 15.8678 PSNR_r = 36.1256, BER=10^-3

TABLE III. INPUT PARAMETERS OF RS CODE IN DIFFERENT BER (IMAGE BOAT)

BER	**Input parameters of rs code**					
	Optimum code rate (1)	***m (1)***	***t (1)***	***Optimum code rate (2)***	***m (2)***	***t (2)***
10^-1	0.2000	4	6	0.2000	4	6
10^-2	0.3000	4	5	0.5000	5	7
10^-3	0.3000	4	5	0.8000	4	1

The simulation results show a big optimzation in the selection of code rate for the RS code according to the channel noise. We remark that the GA gives a hight code rates for high BER levels and decreasing this value with low levels of BER. However the GA keeps a high coding rate for Bf since it represents the most important data in the image. Compared to published papers,[16],[17],[18],our proposition provides a more precision for the selection of RS inputs which optimize in the number of redundant bits and the time of transmission with efficient correction.

X. CONCLUSION

In this paper we study the protection of data using RS code according to the important of data in the original image by creating an algorithm that calculate the coding rate automatically in function of mean squared error (MSE) using genetic algorithm (GA). In general we used to select the coding rate randomly which give us satisfactory results with no precision or minimization in redundant bits needed to protect the data. However the use of GA provide an automatic selection of coding rate according to the channel conditions (BER) which minimization in redundant bits needed to protect the data with more precision and correction. We still believe that a big improvement can be achieved if we change in the fitness function with more constraints.

ACKNOWLEDGMENT

The authors would like to be grateful to the Editor and anonymous reviewers for their valuable comments and suggestions, which have improved the quality of the paper.

REFERENCES

[1] T. B¨ulow, Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images, Ph.D. thesis, Christian Albrechts University, Kiel, Germany, 1999.

[2] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," J. App. Harm. Analysis, vol. 10, no. 3, pp. 234–253, May 2001.

[3] H Malepati ,Digital Media Processing, 1st Edition ,DSP Algorithms Using C. Newnes; 1 edition (June 3, 2010) : 586

[4] Ajol Kumar Ray, Tinku Acharya. Information Technology Principles and Applications. Prentic Hall of India, October 30, 2004.

[5] K.V. Kale, S.C. Mehrotra. Computer Vision and Information Technology: Advances and Applications. Jan 7, 2010, 50

[6] Ryszard S. Choras.Image Processing and Communications Challenges 5.Springer Science & Business Media., 19 juil. 2013 pp 236-238.

[7] Yun Q. Shi, Huifang Sun. Image and Video Compression for Multimedia Engineering. CRC Press, second edition, 2008.

[8] Zhou Wang, Alan C.Bovik. Mean squared error love it or leave it. IEEE signal processing magazine, January. 2009, 99-117.

[9] Reed, I. S. and Solomon, G., "Polynomial Codes Over Certain Finite Fields," SIAM Journal of Applied Math., vol. 8, 1960, pp. 300-304

[10] Ali M. Fadhil, Haider M. AlSabbagh. Performance Analysis for Bit-Error-Rate of DS-CDMA Sensor Network Systems with Source Coding. TELKOMNIKA. March 2012;10(1): 165-170

[11] Binit Amin, Patel Amrutbhai. Vector Quantization based Lossy Image Compression using Wavelets – A Review. International Journal of Innovative Research in Science, Engineering and Technology. March 2014; 3(3):10517-10523.

[12] Hwei P. Hsu, Analog and Digital Communications (Schaum's Outlines), McGraw-Hill Education; second edition, December 10, 2002.

[13] Milan Sonka, Vaclav Hlavac , Roger Boyle, Image Processing Analysis and Machine Vision.THOMSON, CL Engineering, March 19, third edition; 2007.

[14] John Miano, Compressed Image File Formats , Addison Wesley; 1999

[15] Gilberto C. Pereira, Marilia M. F. de Oliveira, Nelson F. F. Ebecken,' Genetic Optimization of Artificial Neural Networks to Forecast Virioplankton Abundance from Cytometric Data',Journal of Intelligent Learning Systems and Applications,Vol. 5 No. 1 (2013)

[16] Lei Yao, Lei Cao,Turbo Codes-Based Image Transmission for Channels With Multiple Types of Distortion, IEEE Transactions on Image Processing ,Vol:17 , Issue: 11 ,September 2008 , pp: 2112-2121

[17] Lamia Chaari , Mohamed Fourati, Lotfi Kamoun,Image transmission quality analysis over adaptive Reed-Solomon coding ,IEEE Mediterranean Electrotechnical Conference Melecon 2010 - 2010 15th,pp: 409 – 414

[18] Phat Nguyen Huu,Vinh Tran-Quang ,Takumi Miyoshi, Multi-hop Reed-Solomon encoding scheme for image transmission on wireless sensor networks,Fourth International Conference on Communications and Electronics (ICCE), 2012 ,pp: 74 - 79

Fault-Tolerant Fusion Algorithm of Trajectory and Attitude of Spacecraft Based on Multi-Station Measurement Data

YANG Xiaoyan
State Key Laboratory of Astronautic Dynamics
Xi'an Satellite Control Center
Xi'an China

YU Hui
Xi'an Satellite Control Center
XSCC
Xi'an China

HU Shaolin
State Key Laboratory of Astronautic Dynamics
Xi'an University of Technology
Xi'an China

LI Shaomini Xi'an
Xi'an Satellite Control Center
XSCC
Xi'an China

***Abstract*—Aiming at the practical situation that the navigation processes of spacecrafts usually rely on several different kinds of tracking equipments which track the spacecraft by turns, a series of new outlier-tolerant fusion algorithms are build to determine the whole flight path as well as attitude parameters. In these new algorithms, the famous gradient descent methods are used to find out the outliers-tolerant flight paths from an integrated data-fusion function designed delicately. In this paper, these new algorithms are used to determine reliably the flight paths and attitude parameters in the situation that a spacecraft is tracked by a series of equipments working by turns and there are some outliers arising in the data series. Advantages of these new algorithms are not only plenary fusion of all of the data series from different kinds of equipments but also discriminatory usage: on the one hand, if the data are dependable, the useable information contained in these data are sufficiently used; on the other hand, if the data are outliers, the bad information from these data are efficiently eliminated from these algorithms. In this way, all of the computational flight paths and attitude parameters are insured to be consistent and reliable.**

Keywords—trajectory; fault-tolerance; data fusion

I. Introduction

This It is necessary for a spacecraft that reliable TT&C network can track, measure and determine its trajectory during the whole flight process. TT&C network usually consists of optical measuring equipments (photoelectric theodolite and laser theadolite) and radio measuring equipments (pulse radar and continuous wave interferometer). Through these equipments partially overlapped relay tracking link, TT&C network can realize the tracking and measurement of spacecrafts in the expansive universe and the orientation and navigation during long-term operation [1-3].

Assuming that there are s_a laser theadolites, s_b pulse radars , s_c ground stations distributed s_d continuous wave interferometer system under the flight trajectory of a spacecraft, through the relay tracking of the partial overlapped link, that can obtain azimuth angle A, elevation angle E and radial distance R in the axis orthogonal coordinates of the spacecraft relative to each related ground station's equipments, as well as the radial distance difference P between two stations and the spacecraft. How to scientifically and effectively use the tracking data from different types of measurement equipments and accurately calculate the flight trajectory of the spacecraft is a research project with engineering background?

Currently, this problem is solved mainly by executing subsection calculation and piecewise series connection based on data rationality check. If subsection calculation result series connection method is used, it will unavoidably cause the lost of partial measurement data and sidesteps of several connection points, which make the trajectory calculation result incoherent. If the measurement data includes outliers, the conventional method even obtains a partial abnormal trajectory, which will influence the analysis on flight state of the spacecraft. This paper proposes and designs a rapid calculation method of spacecraft trajectory and attitude parameter based on fault-tolerance fusion during the whole tracking process.

II. Point-by-Point Fusion Calculation of Spacecraft Trajectory

In order to simply describe the algorithm, this paper divides the tracking data from $s_a + s_b + s_c + s_d$ measurement equipments on the tracking link of the spacecraft into 4 data types:

Type I: the radial distance data set $S_R = \{\{R_i(t), t \in [t_i^a, T_i^a]\}, i = 1,2,\cdots,s_1\}$, which means spacecraft- station ranging data of s_1 time periods with sampling interval of h s , the i th tracking interval is $[t_i^a, T_i^a]$, and the station's coordinate is $X_{ai} = (x_{ai}, y_{ai}, z_{ai})$;

Type II: the azimuth angle data set $S_A = \{\{A_i(t), t \in [t_i^b, T_i^b]\}, i = 1,2,\cdots,s_2\}$, which means the azimuth angle of the spacecraft relative to the station with sampling interval h s in s_2 time periods, the i th tracking interval is $[t_i^b, T_i^b]$, the coordinate of the station is $X_{bi} = (x_{bi}, y_{bi}, z_{bi})$ and the transformation matrix is Φ_{bi};

Type III: the elevation angle data set $S_E = \{\{E_i(t), t \in [t_i^c, T_i^c]\}, i = 1,2,\cdots,s_3\}$, which means the elevation angle of the spacecraft relative to the station with sampling interval h s in s_3 time periods, the i th tracking interval is $[t_i^c, T_i^c]$, the station coordinate is $X_{ci} = (x_{ci}, y_{ci}, z_{ci})$ and the transformation matrix is Φ_{ci};

Type IV: the range difference data set $S_P = \{\{P_i(t), t \in [t_i^d, T_i^d]\}, i = 1,2,\cdots,s_4\}$, which means the distance difference data of main station-spacecraft-assistant station with sampling interval h s in S_4 time periods, the i th tracking time is $[t_i^d, T_i^d]$, the corresponding main station coordinate is $X_{d0} = (x_{d0}, y_{d0}, z_{d0})$ and the assistant station coordinate is $X_{di} = (x_{di}, y_{di}, z_{di})$.

Based on the above classification, a new multi-source data fusion approach is applied to establish the trajectory coordinate fusion calculation method for the spacecraft during long-term flight with multi-station's link tracking. A vector function is established as

$$\begin{pmatrix} x_\kappa(X(t)) \\ y_\kappa(X(t)) \\ z_\kappa(X(t)) \end{pmatrix} = \Phi_\kappa^\tau \begin{pmatrix} x(t) - x_\kappa \\ y(t) - y_\kappa \\ z(t) - z_\kappa \end{pmatrix}, \quad \kappa \in \{b_1,\cdots,b_{s2},c_1,\cdots,c_{s3}\} \tag{1}$$

where, Φ_κ^τ is the transformation matrix of b or c (azimuth angle or elevation angle) data source station $\tau \in \{1,2,\cdots,b_1+\cdots+b_{s2}+c_1+\cdots+c_{s3}\}$.

Denoting $D_\kappa(X) = \sqrt{(x-x_\kappa)^2 + (y-y_\kappa)^2 + (z-z_\kappa)^2}$ where $\kappa \in \{a_1,\cdots,b_{s1},\cdots,d_1,\cdots,d_{s4}\}$ and to establish an objective function as following

$$\begin{aligned} F(X(t)) = & \sum_{i=1}^{s_1} \{R_i(t) - D_{ai}(X(t))\}^2 \phi(t,[t_i^a, T_i^a]) \\ & + \sum_{i=1}^{s_2} \left\{ \sqrt{x_{bi}(X(t)^2 + z_{bi}(X(t))^2} \sin A_i(t) - z_{bi}(X(t)) \right\}^2 \phi(t,[t_i^b, T_i^b]) \\ & + \sum_{i=1}^{s_3} \{D_{ci}(X(t)) \sin E_i(t) - y_{ci}(X(t))\}^2 \phi(t,[t_i^c, T_i^c]) \\ & + \sum_{i=1}^{s_4} \{P(t)_i - (D_{d0}(X(t) - D_{di}(X(t)))\}^2 \phi(t,[t_i^d, T_i^d]) \end{aligned} \tag{2}$$

where the switch function is defined as

$$\phi(s,[t,T]) = \begin{cases} 1, & s \in [t,T] \\ 0, & s \notin [t,T] \end{cases} \tag{3}$$

Sequence the tracking time period (from the beginning time to the ending time) in an order from short to long, and consider the constraint that calculable information shall be not less than 3 measurement units. Let's arrange the beginning time and ending time for all of the tracking equipment so as to find the least and the largest time point respectively

$$\{t_1^a,\cdots,t_{s1}^a; t_1^b,\cdots,t_{s2}^b; t_1^c,\cdots,t_{s3}^c; t_1^d,\cdots,t_{s4}^d\} \Rightarrow \{t_{(1)}^0,\cdots,t_{(s1+s2+s3+s4)}^0\} \tag{4}$$

$$\{T_1^a,\cdots,T_{s1}^a; T_1^b,\cdots,T_{s2}^b; T_1^c,\cdots,T_{s3}^c; T_1^d,\cdots,T_{s4}^d\} \Rightarrow \{T_{(1)}^0,\cdots,T_{(s1+s2+s3+s4)}^0\} \tag{5}$$

then determine the beginning calculation time point and the ending calculation time point

$$\begin{cases} t^0 = t_{(3)}^0 \\ t^e = T_{(s1+s2+s3+s4-3)}^0 \end{cases} \tag{6}$$

where $t_{(3)}^0$ is the 3rd time value after sequencing the beginning time and the ending time of each tracking equipment; $T_{(s1+s2+s3+s4-3)}^0$ is the 4th time value to last.

By using the multivariable non-linear function in formula (2), the minimum point can be get and the point-by-point calculation of the trajectory coordinate can be realized at each sampling moment within $[t^0, t^e]$ in the whole flight process.

$$\hat{X}(t) = \arg\min\{F(X)\} \tag{7}$$

In order to solve the multivariable non-linear function extreme value problem (7), the steepest descent method[10-11] is used in 9 steps as follows:

Step 1: properly select the initial value $X_0 = (x_\Theta(t), y_\Theta(t), z_\Theta(t))$, $k = 0$ and the threshold.

Step 2: calculate negative gradient $S_k = -\nabla F(X_k)$ and its unit vector $\hat{S}_k = -\nabla F(X_k) / \|\nabla F(X_k)\|$;

Step 3: if $\|S^k\| \le \alpha$ or $k \ge 1000$ (Set the maximum threshold of the iteration in order to prevent program's entering the infinite loop, set according to the calculation time and the calculation velocity), execute the 8th step; otherwise execute the 4th step;

Step 4: calculate the length $\rho_k = \|F(X_k) - F(X_{k-1})\|$;

Step 5: calculate $X_{k+1} = X_k + \rho_k \hat{S}_k$ and $F(X_{k+1})$;

Step 6: if $F(X_{k+1}) - F(X_k) \le \beta$, then execute the step 8; otherwise, execute the step 7;

Step 7: given $k+1 \Rightarrow k$, then execute the step 2;

Step 8: circulate the above process, complete the spacecraft coordinate calculation of each time point from t^0 to t^e and output location coordinate $X(t_i)$ of the spacecraft in launch coordinate system at each time point $t_i \in \{t^0, t^0+h,\cdots,t^e\}$.

Step 9: output the location coordinate $X(t) = X_{k+1}$.

III. Fault-Tolerant Smoothing of the Spacecraft Trajectory

The trajectory algorithm described above can realize different type of data fusion during different time periods. However, since this algorithm is based on the least square (short as LS) theory, the calculation result will result partially abnormity or distortion from outliers which exists in the tracking data. In order words, the point-by-point LS calculation algorithm can not eliminate the fluctuation caused by measurement outliers, which make the calculation trajectory badly match with the practical one. Thus, this section will complete fault-tolerance improvement on formula (2). The squared loss function used in the least square algorithm are replaced with the attenuation function ρ described in Fig. 1, which is composed of the even function $\rho(x)$.

$$\begin{cases} \rho(x) = \begin{cases} \dfrac{c_1 c_2}{2}, & x > c_2 \\ g(x), & c_1 < x \le c_2 \\ \dfrac{x^2}{2}, & 0 \le x \le c_1 \end{cases} \\ g(x) = \dfrac{c_1(-x^2 + 2c_2 x - c_1 c_2)}{2(c_2 - c_1)} \end{cases} \qquad (8)$$

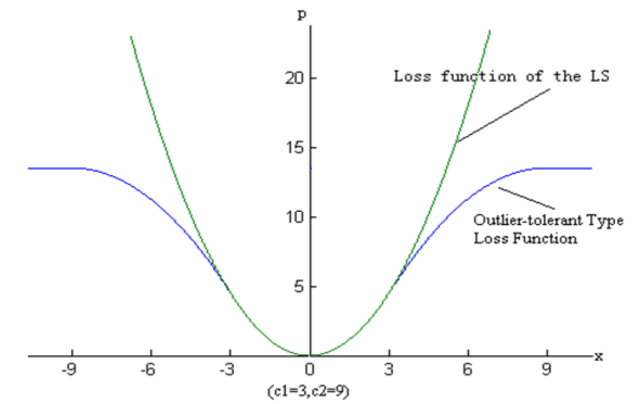


Fig. 1. Comparing between Least Square Loss Function and Fault-Tolerance Loss Function

Setting the window width $\Delta = kh$ ($k \in \{1,2,\cdots,n\}$), the sliding fault-tolerance smooth is executed on the trajectory data. The spacecraft trajectory result disturbance caused by random error is weakened and the calculation result distortion caused by outliers of measurement data is eliminated. To apply the loss function in formula (8), the fault-tolerance estimation method based on the coefficient of cubic fitting curve $\theta_\omega = (\alpha_\omega, \beta_\omega, \gamma_\omega, \tau_\omega)$ are established as follows

$$\begin{pmatrix} \hat{\alpha}_\omega \\ \hat{\beta}_\omega \\ \hat{\gamma}_\omega \\ \hat{\tau}_\omega \end{pmatrix} = \underset{\alpha,\beta,\gamma,\tau}{\arg\min izing} \left\{ \sum_{t=t_i-\Delta}^{t_i+\Delta} \rho \left(\frac{\omega(t) - (\alpha + \beta t + \gamma t^2 + \tau t^3)}{\sqrt{\dfrac{1}{2k+1} \sum_{t=t_i-\Delta}^{t_i+\Delta} \left(\omega(t) - (\alpha + \beta t + \gamma t^2 + \tau t^3) \right)^2}} \right) \right\} \qquad (9)$$

where $\omega \in \{x, y, z\}$, $\alpha_\omega, \beta_\omega, \gamma_\omega, \tau_\omega$ are respectively the minimums of $\alpha, \beta, \gamma, \tau$ .From formula (9), fault-tolerant estimation in the sliding window can be get.

$$\begin{pmatrix} \hat{x}(t) \\ \hat{y}(t) \\ \hat{z}(t) \end{pmatrix} = \begin{pmatrix} \hat{\alpha}_x + \hat{\beta}_x t + \hat{\gamma}_x t^2 + \hat{\tau}_x t^3 \\ \hat{\alpha}_y + \hat{\beta}_y t + \hat{\gamma}_y t^2 + \hat{\tau}_y t^3 \\ \hat{\alpha}_z + \hat{\beta}_z t + \hat{\gamma}_z t^2 + \hat{\tau}_z t^3 \end{pmatrix} \qquad (10)$$

It is easy to validate, due to the stability and continuity of time, the trajectory coordinate in formula (10) only relates to time, and will be hardly influenced by data outliers or random error disturbance, which reflects the operation state of the spacecraft more precisely.

IV. Fault-Tolerance Calculation of Spacecraft Attitude Parameter

The common parameters to describe the space attitude of plane or other spacecraft are pitch angle, yaw angle and rolling angle. Assuming that the direction of longitudinal axis of the spacecraft is identical with the tangent line of trajectories, the calculation result can be applied in formula (10) of trajectory algorithm and a simplification algorithm of space craft attitude parameters can be established through numerical differentiation.

The 1st step is to execute sliding polynomial fault-tolerance differential flatness on the spacecraft trajectory coordinate $\{(x(t_i), y(t_i), z(t_i)) \mid t_i = t^0, t^0 + h, \cdots, t^c\}$ .So, the spacecraft's flight velocity is determined in the measurement range of TT&C network at any time.

According to $(\hat{\alpha}_\omega, \hat{\beta}_\omega, \hat{\gamma}_\omega)$ obtained during the calculation of coordinate components $w \in \{x, y, z\}$, the fault-tolerance estimation is performed to calculate velocity $(\dot{x}(t), \dot{y}(t), \dot{z}(t))$ in the measurement range of TT&C network at any $t \in [t^0 + 5h, t^e - 4h]$ as

$$\begin{pmatrix} \hat{\dot{x}}(t) \\ \hat{\dot{y}}(t) \\ \hat{\dot{z}}(t) \end{pmatrix} = \begin{pmatrix} \hat{\alpha}_x & \hat{\beta}_x & \hat{\gamma}_x & \hat{\tau}_x \\ \hat{\alpha}_y & \hat{\beta}_y & \hat{\gamma}_y & \hat{\tau}_y \\ \hat{\alpha}_z & \hat{\beta}_z & \hat{\gamma}_z & \hat{\tau}_z \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2t \\ 3t^2 \end{pmatrix} \qquad (11)$$

When $t_i \le t^0 + 5h$, $\theta_\omega = (\alpha_\omega, \beta_\omega, \gamma_\omega, \tau_\omega)^\tau$ can be replaced in formula (11) with the calculation result of formula (12) (θ_ω and w can be referenced in formula (9) and (10)).

$$\begin{pmatrix} a_0^w \\ \vdots \\ a_3^w \end{pmatrix} = \left\{ \begin{pmatrix} (t^0)^0 & \cdots & t_{i+5}^0 \\ \vdots & & \vdots \\ (t^0)^3 & \cdots & t_{i+5}^3 \end{pmatrix} \begin{pmatrix} (t^0)^0 & \cdots & (t^0)^3 \\ \vdots & & \vdots \\ t_{i+5}^0 & \cdots & t_{i+5}^3 \end{pmatrix} \right\}^{-1} \begin{pmatrix} (t^0)^0 & \cdots & t_{i+5}^0 \\ \vdots & & \vdots \\ (t^0)^3 & \cdots & t_{i+5}^3 \end{pmatrix} \begin{pmatrix} w(t_0) \\ \vdots \\ w(t_{i+5}) \end{pmatrix} \qquad (12)$$

When $t_i \ge t^e - 4h$, $\theta_\omega = (\alpha_\omega, \beta_\omega, \gamma_\omega, \tau_\omega)^\tau$ can be replaced in formula (11) with the result form formula (13).

$$\begin{pmatrix} a_0^w \\ \vdots \\ a_3^w \end{pmatrix} = \left\{ \begin{pmatrix} t_{i-5}^0 & \cdots & (t^e)^0 \\ \vdots & & \vdots \\ t_{i-5}^3 & \cdots & (t^e)^3 \end{pmatrix} \begin{pmatrix} t_{i-5}^0 & \cdots & t_{i-5}^3 \\ \vdots & & \vdots \\ (t^e)^0 & \cdots & (t^e)^3 \end{pmatrix} \right\}^{-1} \begin{pmatrix} t_{i-5}^0 & \cdots & (t^e)^0 \\ \vdots & & \vdots \\ t_{i-5}^3 & \cdots & (t^e)^3 \end{pmatrix} \begin{pmatrix} w(t_{i-5}) \\ \vdots \\ w(t^e) \end{pmatrix} \quad (13)$$

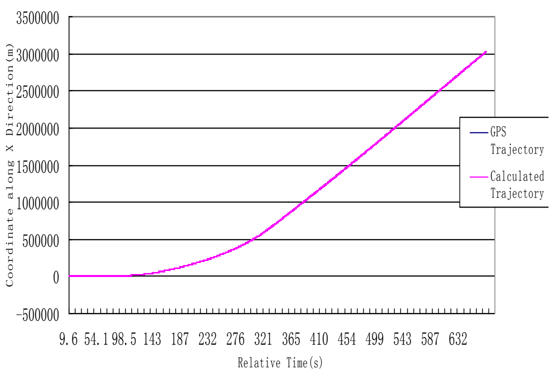
Circularly calculate the velocity parameters at each time point $t_i \in \{t^0, \cdots, t^e\}$

The 2nd step, according to formula (9)-(13), if the direction of longitudinal axis of the spacecraft is identical with the tangent line of trajectory, calculate the pitch angle and yaw angle:

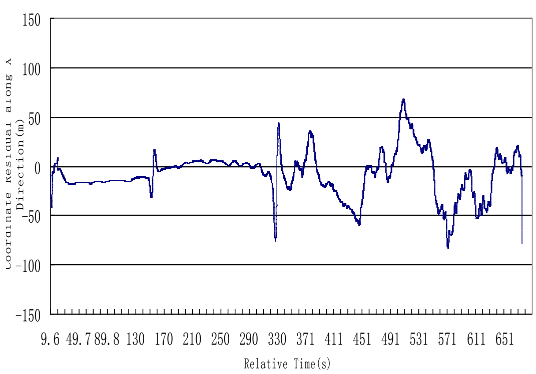
$$\theta = \arcsin\left(\frac{\hat{\dot{y}}}{\sqrt{\hat{\dot{x}}^2 + \hat{\dot{y}}^2}}\right), \ \sigma = \arcsin\left(\frac{-\hat{\dot{z}}}{\sqrt{\hat{\dot{x}}^2 + \hat{\dot{y}}^2 + \hat{\dot{z}}^2}}\right) \quad (14)$$

V. Practical Applications

This paper applies measurement data of some spacecrafts from some equipments (3 single pulse radars, 1 USB device and 4 multiple velocity measurement system) to validate and the results are shown in following figures. The blue line is theoretical trajectory, and the pink line is the calculated trajectory.

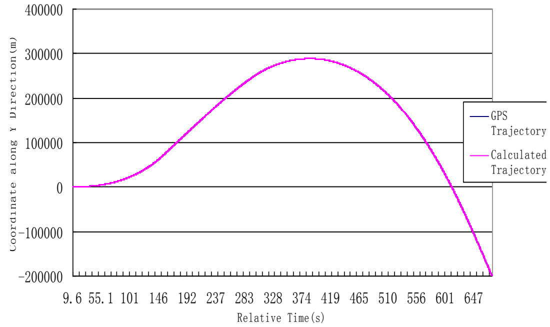


(a) coordinate along X direction

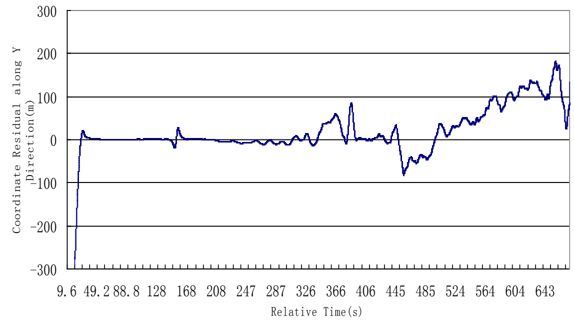


(b) coordinate residual along X direction

Fig. 2. coordinate contrastive diagram along X direction

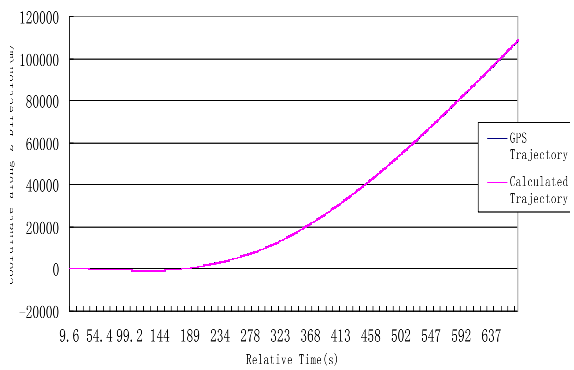


(a) coordinate along Y direction

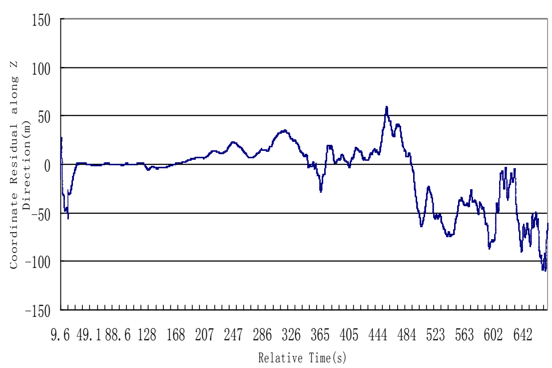


(b) coordinate residual along Y direction

Fig. 3. coordinate contrastive diagram along Y direction

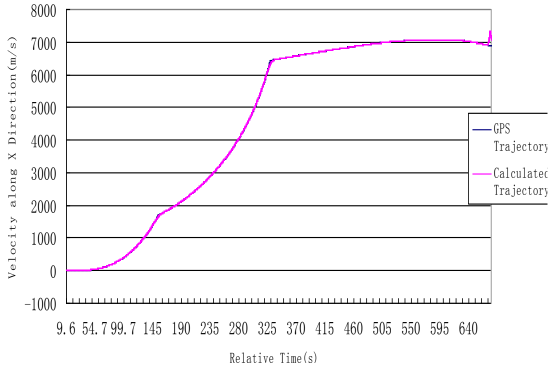


(a) coordinate along Z direction

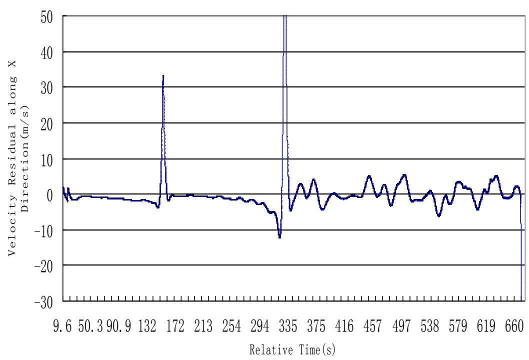


(b) coordinate residual along Z direction

Fig. 4. coordinate contrastive diagram along Z direction

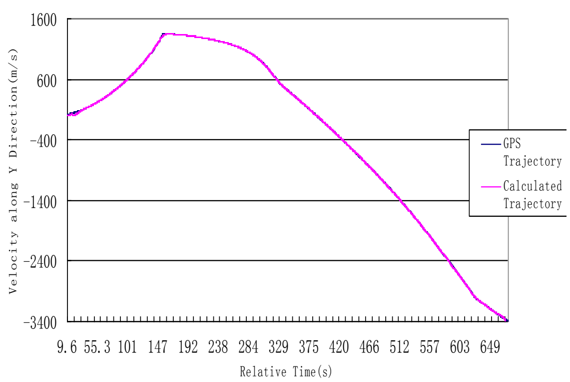


(a) velocity along X direction

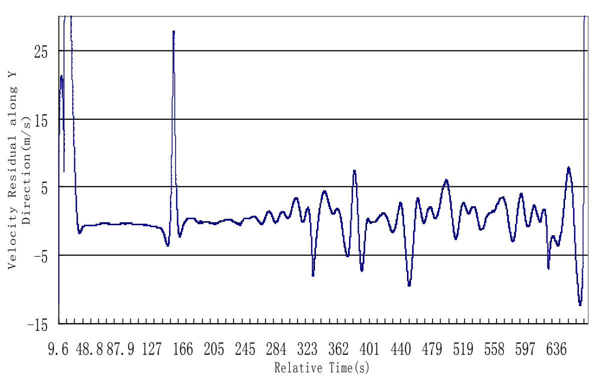


(b) velocity residual along X direction

Fig. 5. velocity contrastive diagram along Z direction

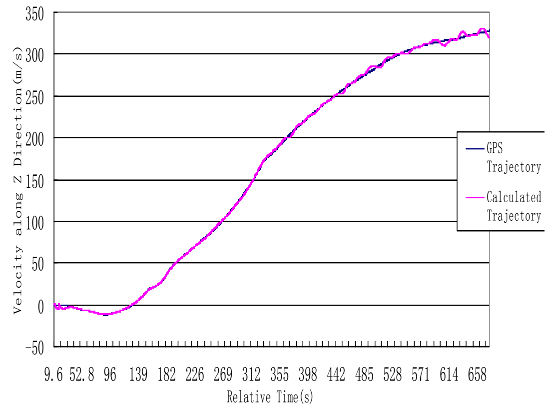


(a) velocity along Y direction

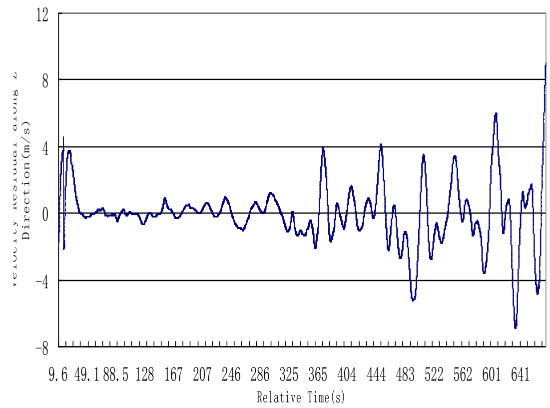


(b) velocity residual along Y direction

Fig. 6. velocity contrastive diagram along Y direction

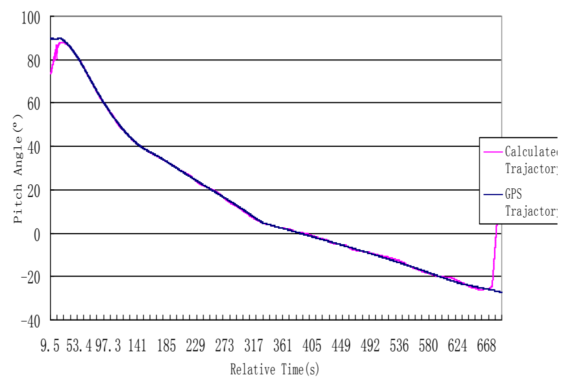


(a) velocity along Z direction

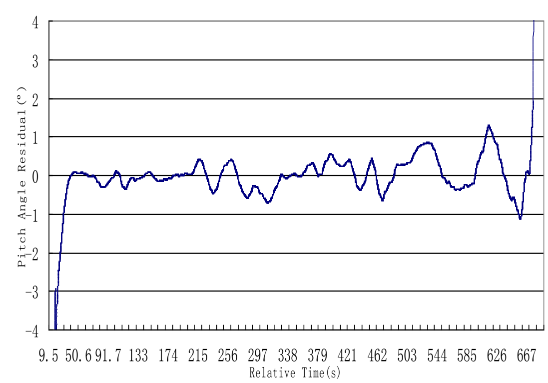


(b) velocity residual along Z direction

Fig. 7. velocity contrastive diagram along Z direction

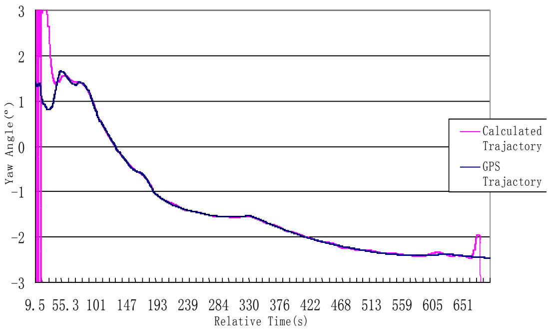


(a) pitch angle

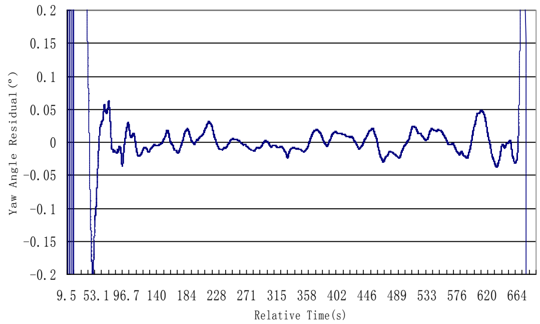


(b) pitch angle residual

Fig. 8. pitch angle contrastive diagram of spacecraft



(a) yaw angle



(b) yaw angle residual

Fig. 9. yaw angle contrast diagram of spacecraft

Fig. 2 ~9 show out the effects which the trajectory computed by this paper's method contrasts with the theoretical trajectory. It can be perceived that this paper's method can compute a whole and continuous trajectory, and this trajectory without distortion and deformation demonstrates the reliability of this paper's method.

VI. Conclusion

In both aeronautic and astronautic field, the monitoring and navigation of the spacecraft is realized by using multiple tracking measurement equipments through partial overlapped relay tracking mode. How to take full advantage of measurement data from multiple TT&C equipments to realize the accurate calculation in tracking trajectory and attitude, which is a technological subject mainly concerned in spacecraft navigation and flight performance analysis field.

This paper establishes a flight trajectory and flight attitude parameter multi-source data fault-tolerance fusion algorithm based on multi-variable non-linear function of the extreme value steepest descent method. This algorithm can rapidly and reliably calculate the trajectory and attitude of a spacecraft in multi- equipment overlapped tracking mode. It can take full advantage of effective data from different equipments, effectively avoid the bad influence of data outliers without data outlier detecting and repair, and remarkably improve the consistency and reliability of spacecraft's trajectory and attitude calculation result without piecewise calculation according to equipments. According to Fig. 1~8, the trajectory obtained by using the algorithm proposed by this paper is complete and continuous without distortion, which proves the reliability of this algorithm.

Acknowledgment

This paper is financially supported by the National Nature Science Fund of China (No.61473222).

References

[1] LU Li-sheng, ZHANG Yu-xiang, LI Jie. External Trajectory Measurement Data Processing. Beijing: National Defence Industry Press, 2002,pp:270-366

[2] Hu Shao-lin, XU Ai-hua, GUO Xiao-hong. Technique of Processing Pulse Radar Measurement Data. Beijing: National Defence Industry Press, 2007, pp:120-139

[3] CUI Shu-hua, HU Shao-lin. Technique of Optical Tracking Measurement Data Processing. Beijing: National Defence Industry Press, 2014,pp:26-107

[4] H Durrant-Whyte, T C Henderson. Multisensor Data Fusion. In Springer Handbook of Robotics,Springer Press,2008,585-610

[5] M E Liggins,D L Hall,J Llinas.Handbook of Multisensor Data Fusion:Theory and Practice. Taylor & Francis Group,CRC Press, 2009,1-44

[6] J R Raol.Multi-Sensor Data Fusion with Mathlab.Taylor & Francis Group,CRC Press, 2009,1-568.

[7] P Kmiotek. Multi-sensor Data Fusion for Representing and Tracking Dynamics Objects. Disseration of PhD,Universite de Technologe de Delfort-Montbeliard, Karkow,2009,pp:5-41 & 125-156.

[8] Shaolin Hu,Xiaofeng Wang,Karl Meinke,etal.Outlier-tolerant Fitting and Online Diagnosis of Outliers in Dynamic Process Sampling Data Series. Proc of 3rd International Conference of Artificial Intelligence and Computational Intelligence, Lecture Notes in Computer Science, vol7004,2011,pp:195-203

[9] LI Hong-yi. Idealized Steepest Descent Method and Its Approximate Example. [J] China: Journal of Shanghai Second Polytechnic University, 2011,28(1):8-13

[10] LIU Ang-ran. Iterative Solution of Linear Equation and Steepest Descent Method [J] China: Journal of Chifeng University(Natural Science Edition), 2014,30(2):10-13.

Discrete-Time Approximation for Nonlinear Continuous Systems with Time Delays

Bemri H'mida
University of Tunis El Manar
National Engineers School of Tunis
Laboratory of Research on Automatic (LARA)
BP 37, Belvedere 1002 Tunis. Tunisia

Soudani Dhaou
University of Tunis El Manar
National Engineers School of Tunis
Laboratory of Research on Automatic (LARA)
BP 37, Belvedere 1002 Tunis. Tunisia

***Abstract*—This paper is concerned with the discretization of nonlinear continuous time delay systems. Our approach is based on Taylor-Lie series. The main idea aims to minimize the effect of the delay and neglects the importance of nonlinear parameter by the linearization of the system study in an attempt to make its handling and easier programming as possible. We investigate a new method based on the development of new theoretical methods for the time discretization of nonlinear systems with time delay .The performance of these proposed discretization methods was validated by doing the numerical simulation using a nonlinear system with state delay. Some illustrative examples are given to show the effectiveness of the obtained results.**

Keywords—Discrete-time systems; Time-delay systems; Taylor-Lie series; non-linear systems; Simulation

I. Introduction

Research on discrete time delay systems has not attracted as much attention as that of continuous time delay systems. Many engineering applications need a compact and accurate description of the dynamic behavior of the considered system. This is especially true of automatic control applications. Dynamic models describing the system of interest can be constructed using the first principles of physics, chemistry, biology and so forth.

Time delay systems often appear in industrial systems and information networks. Thus, it is important to analyze time delay systems and design appropriate controllers. Control systems with time delays exhibit complex behaviors because of their infinite dimensionality. Even in the case of linear time-invariant systems that have constant time delays in their inputs or states have infinite dimensionality if expressed in the continuous time domain. It is therefore difficult to apply the controller design techniques that have been developed during the last several decades for finite dimensional systems to systems with any time delays in the variables. Thus, new control system design methods that can solve a system with time delays are necessary.

As a result, controller design techniques developed for finite dimensional systems are difficult to apply to time delay systems with some effectiveness, time delay is often encountered in various engineering systems and its existence is frequently a source of instability. Many of these models are also significantly nonlinear which motivates research in the control of nonlinear systems with time delay. For this reasons, it's difficult to analyze and design the control algorithm for the nonlinear time delay system in the continuous time domain. It is necessary to develop a method to solve the time delay problem. Most of the proposed approaches deal with linear time-delay control systems and, in particular, with the stability analysis and behavior of such systems with constant and/or uncertain time delays [19,21,11]. Quite recently and on the nonlinear front, nonlinear controllers were systematically synthesized for multivariable nonlinear systems in the presence of sensor and actuator dead time [9,5].

In practice, most of industrial controllers are currently implemented digitally. In the design of model based digital control systems two general approaches can be identified. First, a continuous time controller is designed based on a continuous-time system model, followed by a digital redesign of the controller in the discrete time domain to approximate the performance of the original continuous time controller. Second, a direct digital design approach can be followed based on a discrete time model of the system, where the controller is now directly designed in the discrete time domain. It is apparent that this alternative approach has the attractive feature of dealing directly with the issue of sampling. We can emphasize, that in both design approaches time discretization of either the controller or the system model is necessary. Furthermore, note that in controller design for time delay systems the first approach is troublesome because of the infinite dimensional nature of the underlying system dynamics. As a result, the second approach becomes more desirable and will be pursued in the present study.

In particular, the well known procedure of time discretization of linear time delay systems [7,12,4] is extended to nonlinear input driven systems with constant time delay. All these approaches require a small time step in order to be deemed accurate, and this may not be the case in control applications where large sampling periods are inevitably introduced due to physical and technical limitations [13, 8]. Due to the physical and technical limitations, slow sampling has become inevitable. A time discretization method that expands the well known time discretization of linear time delay systems [1,6,2,3] to nonlinear continuous time control systems with time delays [10,17] can solve this problem. The effect of this approach on system theoretic properties of nonlinear systems, such as equilibrium properties, relative order, stability, zero dynamics, and minimum phase characteristics has also been studied [20,16] and reveals the natural and transparent manner in which Taylor methods permeate the relevant theoretical aspects. A certainly not exhaustive sample of other approaches of notable significance, yet with certain associated practical limitations, are reported in [18], and solid theoretical results on the direct use of discrete time approximations in the control of sampled-data nonlinear systems can be found in [14,22].

In particular, the present study aims at the development of new methods for the time discretization of nonlinear input driven dynamic systems with time delay based on Taylor series. In particular, the paper is organized as follows: the next section contains some mathematical preliminaries; Section 3 discusses the discretization of system with internal point delay; Section 4 discusses the discretization of system with external point delay; Section 5 discusses the discretization of system with internal and external point delays; Section 6 discusses the linearization of nonlinear state space equation and a numerical example is given in section 7 to illustrate the proposed theoretical results and a concluding remark.

II. Preliminaries

In the present study, single-input nonlinear continuous time control systems with input output time delays is considered using a state space representation of the form:

$$\dot{x}(t) = f_1(x(t)) + f_2(x(t-\tau_0)) + g_1(x(t))u(t) + g_2(x(t-\tau_0))u(t-\tau_1) \quad (1)$$

where, τ_0 and τ_1 are the time delay and $u(t)$ is the control input.

and $$\dot{x} = f(x(t-\tau_0), u(t-\tau_1)) \quad (2)$$

where $x \in R^n$ is the vector of the states representing an open and connected set, $u \in R$ is the input variable, m and n are an integer which indicates the order of the input. τ_0 and τ_1 are the system constant time delay, that directly affects the input and the state. It is assumed that:

$f_i : R^n \to R^n$ and $g_i : R^n \to R^n$, i = 1; 2; ….

m and $f : R^n \times R \to R^n$ are smooth mappings.

An equidistant grid on the time axis with mesh $T = t_{k+1} - t_k > 0$ is considered where sampling interval is $[t_k, t_{k+1}] = [kT, (k+1)T]$ and T is the sampling period. Furthermore, we suppose the time-delay τ_0 and τ_1 mesh T are related as follows

$$\tau_0 = q_0 T \quad , (q_0 \geq 1 \text{, is an integer}) \quad (3)$$

$$\tau_1 = q_1 T \quad , (q_1 \geq 1 \text{, is an integer}) \quad (4)$$

where $q_0, q_1 \in \{0, 1, ..., m\}$. That is, the time-delay τ_0 and τ_1 are customarily represented as an integer multiple of the sampling period adding a fractional part of T [22].

It is assumed that system (1) is driven by an input that is piecewise constant over the sampling interval, i.e. the zero-order hold (ZOH) assumption holds true:

$$u(t) = u(kT) \equiv u(k) = \text{constant, for } kT < t < (k+1)T \quad (5)$$

III. Discretization of Nonlinear Systems with Internal Point Delay

The nonlinear continuous time control systems with input time delay are considered using a state space representation form:

$$\dot{x}(t) = f_1(x(t)) + g_1(x(t))u(t-\tau_1) \quad (6)$$

Based on the zero-order hold assumption and the above notation one can deduce that the delayed input variable attains the following two distinct values within the sampling interval:

$$u(t-\tau_1) = u(kT - q_1 T) \equiv u(k - q_1) \text{, for } kT < t < (k+1)T \quad (7)$$

the nonlinear system (6) can be discretized using Taylor series expansions over the subinterval $kT < t < (k+1)T$ and taking into account (7), one can obtain the state vector evaluated at $(k+1)T$ as a function of $x(k)$ and $u(k-q_1)$.

around the point $x(t_0)$, the state $x(t)$ can be expanded to Taylor series as:

$$x(t) = x(t_0) + x'(t_0)(t-t_0) + \frac{x''(t_0)}{2!}(t-t_0)^2 + \frac{x'''(t_0)}{3!}(t-t_0)^3 + ... \quad (8)$$

in the time interval $[t_k, t_{k+1}] = [kT, (k+1)T]$, equation (8) can be rewritten using equation (9):

$$x((k+1)T) = x(kT) + x'(kT)T + \frac{x''(kT)}{2!}T^2 + \frac{x'''(kT)}{3!}T^3 + ... \quad (9)$$

for simplicity and without misunderstanding, equation (9) can be rewritten as:

$$x(k+1) = x(k) + x'(k)T + \frac{x''(k)}{2!}T^2 + \frac{x'''(k)}{3!}T^3 + ... \quad (10)$$

from equation (6), we can get the differential coefficient of the state $x(t)$:

$$\dot{x}(t) = f_1(x(t)) + g_1(x(t))u(t-\tau_1) \quad (11)$$

then in the time interval $[t_k, t_{k+1}] = [kT, (k+1)T]$, equation (11) can be rewritten using equation (12):

$$\dot{x}(k) = f_1(x(k)) + g_1(x(k))u(k-q_1) \quad (12)$$

similarly, based on equation (6) we can calculate the second derivative of the state $x(t)$, shown in equation (13):

$$\begin{aligned} x''(t) &= \frac{d(x'(t))}{dt} = \frac{d(f_1(x(t)) + g_1(x(t)u(t-\tau_1)))}{dt} \\ &= \frac{df_1(x(t))}{dx}\frac{dx}{dt} + u(t-\tau_1)\frac{dg_1(x(t))dx}{dx\ dt} + g_1(x(t))\frac{du(t-\tau_1)}{dt} \\ &= \left(\frac{df_1(x(t))}{dx} + u(t-\tau_1)\frac{dg_1(x(t))}{dx}\right)\frac{dx}{dt} + g_1(x(t))\frac{du(t-\tau_1)}{dt} \end{aligned} \quad (13)$$

for the zero order hold assumption, in each sampling interval

Equation (13) is correct:

$$\frac{du(t)}{dx} = 0 \Rightarrow \frac{du(t-\tau_1)}{dx} = 0 \tag{14}$$

then in each sampling interval, equation (13) can be expressed using equation (15):

$$\begin{aligned} x''(t) &= \frac{d(x'(t))}{dt} = \left(\frac{df_1(x(t))}{dx} + u(t-\tau_1)\frac{dg_1(x(t))}{dx}\right)\frac{dx}{dt} + g_1(x(t))\frac{du(t-\tau_1)}{dt} \\ &= \left(\frac{df_1(x(t))}{dx} + u(t-\tau_1)\frac{dg_1(x(t))}{dx} + g_1(x(t))\frac{du(t-\tau_1)}{dx}\right)\frac{dx}{dt} \\ &= \frac{\partial\left(f_1(x(t)) + u(t-\tau_1)g_1(x(t))\right)}{\partial x}\frac{dx}{dt} \end{aligned} \tag{15}$$

or $\frac{dx}{dt} = \dot{x}(t) = f_1(x(t)) + g_1(x(t))u(t-\tau_1)$,

then equation (15) can be rewritten as:

$$\begin{aligned} x''(t) &= \frac{d(x'(t))}{dt} = \frac{\partial\left(f_1(x(t)) + u(t-\tau_1)g_1(x(t))\right)}{\partial x}\frac{dx}{dt} \\ &= \frac{\partial\left(f_1(x(t)) + u(t-\tau_1)g_1(x(t))\right)}{\partial x}\left(f_1(x(t)) + u(t-\tau_1)g_1(x(t))\right) \end{aligned} \tag{16}$$

in the time interval $[t_k, t_{k+1}] = [kT, (k+1)T]$, equation (16) can be rewritten using equation (17):

$$x''(k) = \frac{\partial\left(f_1(x(k)) + u(k-q_1)g_1(x(k))\right)}{\partial x}\left(f_1(x(k)) + u(k-q_1)g_1(x(k))\right) \tag{17}$$

assume that:

$$\begin{aligned} A^{[1]}(x,u) &= f_1(x(k)) + g_1(x(k))u(k-q_1) \\ A^{[2]}(x,u) &= \frac{\partial A^{[1]}(x,u)}{\partial x} f_1(x(k)) + g_1(x(k))u(k-q_1) \\ A^{[l+1]}(x,u) &= \frac{\partial A^{[l]}(x,u)}{\partial x} f_1(x(k)) + g_1(x(k))u(k-q_1) \\ l &= 1,2,3,\ldots \end{aligned} \tag{18}$$

then equation (17) can be written as:

$$\begin{aligned} x''(k) &= \frac{\partial\left(f_1(x(k)) + u(k-q_1)g_1(x(k))\right)}{\partial x}\left(f_1(x(k)) + u(k-q_1)g_1(x(k))\right) \\ &= A^{[2]}(x(k), u(k-q_1)) \end{aligned} \tag{19}$$

in the same way, we have:

$$x''' = A^{[3]}(x(k), u(k-q_1)) \tag{20}$$

then equation (10) can be written as:

$$\begin{aligned} x(k+1) &= x(k) + \sum_{l=1}^{N} \frac{T^l}{l!}\frac{d^l x}{dt^l}\bigg|_{t_k} \\ &= x(k) + \sum_{l=1}^{N} A^{[l]}\left(x(k), u(k-q_1)\right)\frac{T^l}{l!} \end{aligned} \tag{21}$$

here $x(k)$ is the value of the state $x(t)$ at the time $t = kT$, $A^{[l]}\left(x(k), u(k-q_1)\right)$ can be calculated using equation (18).

The Taylor series expansion of equation (21) can offer either an exact sampled data representation of equation (6) by remaining the full infinite series representation of the state vector. It can also provide an approximate sampled data representation of equation (6) resulting from a truncation of the Taylor series order:

$$\begin{aligned} x(k+1) &= \Phi_T^N\left(x(k), u(k-q_1)\right) \\ &= x(k) + \sum_{l=1}^{N} A^{[l]}\left(x(k), u(k-q_1)\right)\frac{T^l}{l!} \end{aligned} \tag{22}$$

where, the subscript of Φ denotes the dependence of the sampling period and the superscript N denotes the finite series truncation order of the equation (22).

IV. Discretization of Nonlinear Systems with External Point Delay

The nonlinear continuous time control systems with state delay can be represented by the following state space form:

$$\dot{x}(t) = f(x(t)) + f_1(x(t-\tau_0)) + g(x(t))u(t) + g_1(x(t-\tau_0))u(t) \tag{23}$$

where, τ_0 is the time delay and $u(t)$ is the control input.

assume that in the time interval $[t_k, t_{k+1}] = [kT, (k+1)T]$

$\tau_0 = q_0 T$, ($q_0 \geq 1$, is an integer)

In the time interval $t \in [kT, (k+1)T]$, $k = 0, 1, \ldots, n-1$,

$f_1(x(t-\tau_0)) = 0$ and $g_1(x(t-\tau_0))u(t) = 0$. Under the zero-order holds assumption and within the sampling interval, the solution described in equation (23) is expanded in a uniformly convergent Taylor series and the resulting coefficients can be easily competed by taking successive partial derivatives of the right hand side of equation (23).

An approximate sampled data representation:

$$x(k+1) = x(k) + \sum_{l=1}^{N} A^{[l]}\left(x(k-q_0), u(k)\right)\frac{T^l}{l!} \tag{24}$$

where, $A^{[l]}(x,u)$ can be calculate using equation (23)

$$\begin{aligned} A^{[1]}(x,u) &= f_1(x(k-q_0)) + g_1(x(k-q_0))u(k) \\ A^{[2]}(x,u) &= \frac{\partial A^{[1]}(x,u)}{\partial x} f_1(x(k-q_0)) + g_1(x(k-q_0))u(k) \\ A^{[l+1]}(x,u) &= \frac{\partial A^{[l]}(x,u)}{\partial x} f_1(x(k-q_0)) + g_1(x(k-q_0))u(k) \\ l &= 1,2,3,\ldots \end{aligned} \tag{25}$$

in the time interval $t \in [kT, (k+1)T]$, $k = 0, 1, \ldots, n-1$, equation (24) provides the approximates sampled data representation of equation (23):

$$x(k+1) = x(k) + \sum_{l=1}^{N}\left(A^{[l]}\left(x(k), u(k)\right)\frac{T^l}{l!} + B^{[l]}\left(x(k-q_0), u(k)\right)\frac{T^l}{l!}\right) \tag{26}$$

where, $A^{[l]}\left(x(k),u(k)\right)$ can be calculated using equation(23), and $B^{[l]}\left(x(k-q_0),u(k)\right)$ can be calculate using equation (25):

$$\begin{aligned} &B^{[1]}(x,u)=f_1(x(k-q_0))+g_1(x(k-q_0))u(k) \\ &B^{[2]}(x,u)=\frac{\partial B^{[1]}(x,u)}{\partial x}f_1(x(k-q_0))+g_1(x(k-q_0))u(k) \\ &B^{[l+1]}(x,u)=\frac{\partial B^{[1]}(x,u)}{\partial x}f_1(x(k-q_0))+g_1(x(k-q_0))u(k) \\ &\qquad l=1,2,3,... \end{aligned} \tag{27}$$

The discrete time form of the nonlinear continuous system with state delay, shown in equation (23) can be gotten by combining equation (22) and (24).

V. Discretization of Nonlinear Systems with Internal and External Point Delays

The nonlinear continuous time control systems with input output time delays are considered using a state space representation form:

$$\dot{x}(t)=f(x(t))+f_1(x(t-\tau_0))+g(x(t))u(t)+g_1(x(t-\tau_0))u(t-\tau_1) \tag{28}$$

where, τ_0 and τ_1 are the time delays and $u(t)$ is the control input.

assume that in the time interval $[t_k,t_{k+1}]=[kT,(k+1)T]$

$\tau_0=q_0T$, ($q_0\geq 1$, is an integer)

$\tau_1=q_1T$, ($q_1\geq 1$, is an integer)

based on the zero order hold assumption and the above notation one can deduce that the delayed input variable attains the following two distinct values within the sampling interval:

$$u(t-\tau_1)=u(kT-q_1T)\equiv u(k-q_1),\text{for } kT<t<(k+1)T \tag{29}$$

the nonlinear system (26) can be discretized using Taylor series expansions over the subinterval $kT<t<(k+1)T$ and taking into account (27), one can obtain the state vector evaluated at $(k+1)T$ as a function of $x(k-q_0)$ and $u(k-q_1)$.

in the time interval $t\in[kT,(k+1)T]$, $k=0,1,...,n-1$, equation (24) provides the approximates sampled data representation of equation (23):

$$x(k+1)=x(k)+\sum_{l=1}^{N}\left(A^{[l]}\left(x(k),u(k)\right)\frac{T^l}{l!}+B^{[l]}\left(x(k-q_0),u(k-q_1)\right)\frac{T^l}{l!}\right) \tag{30}$$

where, $A^{[l]}\left(x(k),u(k)\right)$ can be calculated using equation(31), and $B^{[l]}\left(x(k-q_0),u(k-q_1)\right)$ can be calculate using equation (32):

$$\begin{aligned} &A^{[1]}(x,u)=f_1(x(k))+g_1(x(k))u(k) \\ &A^{[2]}(x,u)=\frac{\partial A^{[1]}(x,u)}{\partial x}f_1(x(k))+g_1(x(k))u(k) \\ &A^{[l+1]}(x,u)=\frac{\partial A^{[1]}(x,u)}{\partial x}f_1(x(k))+g_1(x(k))u(k) \\ &\qquad l=1,2,3,... \end{aligned} \tag{31}$$

and:

$$\begin{aligned} &B^{[1]}(x,u)=f_1(x(k-q_0))+g_1(x(k-q_0))u(k-q_1) \\ &B^{[2]}(x,u)=\frac{\partial B^{[1]}(x,u)}{\partial x}f_1(x(k-q_0))+g_1(x(k-q_0))u(k-q_1) \\ &B^{[l+1]}(x,u)=\frac{\partial B^{[1]}(x,u)}{\partial x}f_1(x(k-q_0))+g_1(x(k-q_0))u(k-q_1) \\ &\qquad l=1,2,3,... \end{aligned} \tag{32}$$

The discrete time form of the nonlinear continuous system with input output time delays, shown in equation (28) can be gotten by combining equation (22) and (24).

VI. Linearization of Nonlinearstate Equation

The technique of linearization involves approximating a complicated system of equations with a simpler linear system. We hope to gain insight into the behavior of the nonlinear system through an analysis of the behavior of its linearization. We hope that the nonlinear system will behave locally like its linearization, at least in a qualitative sense.

In general, the linearization of a system of equations about an equilibrium point can be achieved by changing variables so that the equilibrium point is transformed to the origin. Points in the original system close to the equilibrium point will correspond to points close to the origin in the new system. Thus we are only concerned with values of the new variables close to zero and under certain conditions the nonlinear terms can be neglected. The equations that result are linear and are the linearization of the original system.

In order to linearize general nonlinear systems, we will use the Taylor Series expansion of functions.

Consider the nonlinear system:

$$\begin{cases} x=f(x,u) \\ y=g(x,u) \end{cases} \tag{33}$$

with the equilibrium point is (p,q). Any function which is differentiable can be written as a Taylor series expansion for $f(x,u)$ with neglect the terms of high order:

$$\dot{x}=f(x,u)=f(p,q)+\left.\frac{\partial f}{\partial x}\right|_{(p,q)}(x-p)+\left.\frac{\partial f}{\partial u}\right|_{(p,q)}(u-q)+F(x,u) \tag{34}$$

where $F(x,u)$ consists of nonlinear polynomial terms in $(x-p)$ and $(u-q)$.

since (p,q) is an equilibrium $f(p,q)=0$ and neglect high order terms, then the state space representation form (33) can be rewritten as:

$$\dot{x}=f(x,u)=\left.\frac{\partial f}{\partial x}\right|_{(p,q)}(x-p)+\left.\frac{\partial f}{\partial u}\right|_{(p,q)}(u-q)+F(x,u) \quad (35)$$

for points near to the equilibrium point $(x-p)$ and $(u-q)$ are small and the non linear terms $F(p,q)$ can be neglected.

We can write the state space model as:

$$\dot{x}=f(x,u)=\left.\frac{\partial f}{\partial x}\right|_{(p,q)}(x-p)+\left.\frac{\partial f}{\partial u}\right|_{(p,q)}(u-q) \quad (36)$$

where the elements of linearization matrices are:

$$A_{ij}=\left.\frac{\partial f_i}{\partial x_j}\right|_{(p,q)}=\begin{pmatrix}\left.\frac{\partial f_1}{\partial x_1}\right|_{(p,q)} & \left.\frac{\partial f_1}{\partial x_2}\right|_{(p,q)}\\ \left.\frac{\partial f_2}{\partial x_1}\right|_{(p,q)} & \left.\frac{\partial f_2}{\partial x_2}\right|_{(p,q)}\end{pmatrix},\ B_{ij}=\left.\frac{\partial f_i}{\partial u_j}\right|_{(p,q)}=\begin{pmatrix}\left.\frac{\partial f_1}{\partial u}\right|_{(p,q)}\\ \left.\frac{\partial f_2}{\partial u}\right|_{(p,q)}\end{pmatrix}$$

$$C_{ij}=\left.\frac{\partial g_i}{\partial x_j}\right|_{(p,q)} \text{ and } D_{ij}=\left.\frac{\partial g_i}{\partial u_j}\right|_{(p,q)}$$

where A_{ij} is called the Jacobian matrix.

VII. Result of Simulations

The performance of the proposed methods of discretization for nonlinear systems with time delays is evaluated by applying it to a nonlinear continuous system with time delays. The partial derivative terms involved in the Taylor series expansion are determined recursively. The system considered in this paper is assumed to be a nonlinear control system, as it considers the pendulum equation with friction:

$$f(x,u)=\begin{cases}\dot{x}_1=x_2\\ \dot{x}_2=-\frac{g}{l}sinx_1-\frac{k}{m}x_2+u(t-\tau)\end{cases} \quad (37)$$

The vector $u(t)$, called the input history or control input, is chosen to influence the dynamics in some desired way. The vector of functions f describes the system's dynamics and the vector of functions h provides a set of output measurements. We call any pair $(x(t),u(t))$ satisfying over some time interval including $(t=t_0)$ a solution or trajectory.

Note that any system of higher order differential equations can be written in the first order form. For example, the motion of a simple pendulum with an input torque is described by the second order nonlinear equation:

$$T=0.2s,\ \tau=0.2s,\ \frac{g}{l}=1 \text{ and } \frac{k}{m}=0,5$$

$$\begin{aligned}f(x,u)&=\begin{pmatrix}x_2\\ -sinx_1-0.5x_2+u(t-0.2)\end{pmatrix}\\ &=\begin{pmatrix}0 & 1\\ -sinx_1 & -0.5\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\begin{pmatrix}0\\ 1\end{pmatrix}u(t-0.2)\end{aligned} \quad (38)$$

with:

$$x=\begin{pmatrix}x_1\\ x_2\end{pmatrix}\ A=\begin{pmatrix}0 & 1\\ -sinx_1 & -0.5\end{pmatrix}\ B=\begin{pmatrix}0\\ 1\end{pmatrix}$$

the Jacobian matrix of the function $f(x,u)$ of the pendulum equation is given by:

$$\frac{\partial f}{\partial x}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2}\end{pmatrix}=\begin{pmatrix}0 & 1\\ -\cos(x_1) & -0.5\end{pmatrix} \quad (39)$$

$$\frac{\partial f}{\partial u}=\begin{pmatrix}\frac{\partial f_1}{\partial u}\\ \frac{\partial f_2}{\partial u}\end{pmatrix}=\begin{pmatrix}0\\ 1\end{pmatrix} \quad (40)$$

evaluating the Jacobian matrix at the equilibrium points (0, 0) and (π, 0) yields, respectively, the two matrices

$$A_1=\begin{pmatrix}0 & 1\\ 1 & -0.5\end{pmatrix} \text{ and } A_2=\begin{pmatrix}0 & 1\\ -1 & -0.5\end{pmatrix}$$

The Taylor series expansion of equation (21) can offer either an exact sampled data representation of the equation (6) by remaining the full infinite series representation of the state vector. It can also provide an approximate sampled data representation of equation (6) resulting from a truncation of the Taylor series order:

$$x(k+1)=x(k)+A(x(k),u(k-1))T \quad (41)$$

the simulation results is depicted in the figure 1

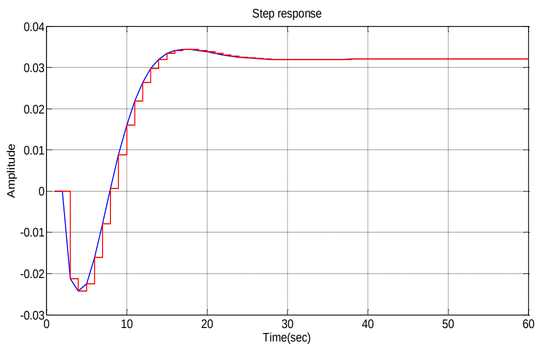


Fig. 1. Step response of nonlinear continuous and discret system with control time delay

In the second case, the nonlinear continuous time control systems with state delay can be represented by the following state space form:

$$x(k+1)=x(k)+A\left(x(k-1),u(k)\right)T \qquad (42)$$

the simulation results is depicted in the figure 2

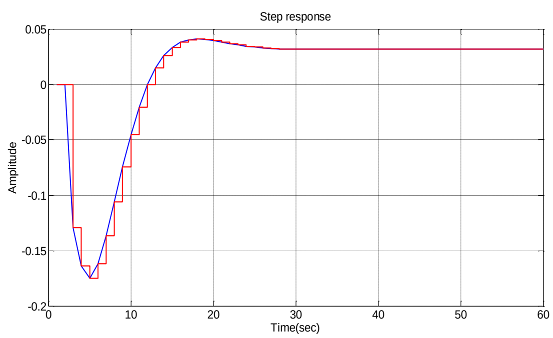


Fig. 2. Step response of nonlinear continuous and discret system with state time delay

In the third case, the nonlinear continuous time control systems with input output time delays are considered using a state space representation form:

$$x(k+1)=x(k)+\Big(A\big(x(k),u(k)\big)T+A\big(x(k-q_0),u(k-q_1)\big)T\Big) \qquad (43)$$

the simulation results is depicted in the figure 3

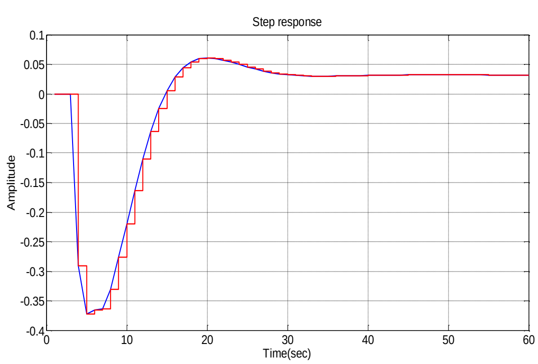


Fig. 3. Step response of nonlinear continuous and discret system with input output time delay

Eventually, the simulation using a nonlinear system with time delay is conducted to validate the proposed time discretization method.

VIII. Conclusion

This paper proposed a time discretization method for nonlinear continuous systems with internal and external point delays. This proposed discretization method is based on Taylor series. The performance of the proposed time discretization method is evaluated using a nonlinear system with time delayed. The derived time discretization method provides a finite dimensional representation for nonlinear control systems with time delay, thereby enabling the application of existing nonlinear controller design techniques to such systems.

Finally, the simulation results show that the proposed discretization method does not change the original system stability nor increase much computational burden.

References

[1] A. El moudni, b.bensassi: On a new order reduction method for discrete systems.12th world congress on scientific computation imacs,pp.95-99,1988.

[2] B.H'mida, M.Sahbi and S.Dhaou: Discrete-time approximation of multivariable continuous-time delay systems. IGI global handbook of research on advanced intelligent control engineering and automation, pp516-542, doi: 10.4018/978-1-4666-7248-2.ch019, January 2015.

[3] B.H'mida,M.Sahbi and S.Dhaou: Discretizing of linear systems with time-delay using method of euler's and tustin's approximations. International journal of engineering research and applications (ijera) ,issn: 2248-9622 , vol. 5 -issue 3, march 2015.

[4] B. H'mida, M. Sahbi and S. Dhaou: Stability of a linear discrete system with time delay via lyapunov-krasovskii functional. International journal of scientific research & engineering technology (ijset). Issn: 2356-5608, vol.3, issue 3, copyright cpco,pp .62-66, 2015.

[5] E. Fridman, U. Shaked, and V. Suplin., :Input/output delay approach to robust sampled-data h ∞ control.Systems& control letters, vol. 54, no. 3, pp. 271–282, 2005.

[6] J. Jugo.: Discretization of continuous time delay stems. 15th triennial word congress, d. Electricidad y electronica, f. De ciencias,upv/ehu,apdo.644,bilbao,barcelona, spainifac, 2002.

[7] L.s. Shieh, Yeung, C.k., and Mclnnis, B.C.: Solutions of state-space equations via block-pulse functions. Int. J. Control, 28, pp. 383-392,1978.

[8] M. Vidyasagar, Nonlinear Systems Analysis. Prentice Hall, Englewood Cliffs, NJ, USA.1978.

[9] N.k., Sinha and Lastman, G.J.: Transformation algorithm for identification of continuous-time multivariable systems from discrete data. Electron. Lett, 17, pp. 779-780,1981.

[10] N.ksinha and zouq. :Discrete time approximation of multivariable continuous time systems. Iee proceedings, vol 130,pt.d,n°3 ,pp 103-110,1983.

[11] 8S.MagdiMahmoud.:Robust filtering for time-delay systems. Marcel dekker,elsevier science inc. New york, ny, usa.an international journal,vol 176,isuue 2,pp.186-200,2000.

[12] W. Michiels, v. Van assche, and s.-i. Niculescu: stabilization of time-delay. Ieee transactions on automatic control, vol. 50, no. 4, april 2005.

[13] W. Michiels and S.-I. Niculescu . : Stability and stabilization of time-delay systems. An eigenvalue-based approach, ser. Advances in design and control. Vol. 12.,siam, 2007.

[14] Yuan-liang Zhang and Kil To Chong: Time-discretization of time delayed non-affine system via Taylor-lie series using scaling and squaring technique. International journal of control, automation, and systems, vol. 4, no. 3, pp. 293-301, june 2006.

[15] Y. L. Zhang, O. Kostyukova, K. T. Chong: A new time discretization for delay multiple-input nonlinear systems using the Taylor method and first

order hold. Discrete Applied Mathematics, vol. 159, no. 9, pp. 924-938, 2011.

[16] Y. L. Zhang: Discretization of nonlinear non-affine time delay system using first order hold assumption with scaling and squaring technique. International Review on Computers and Software, vol. 7, no. 4, pp. 1860-1865, 2012.

[17] Yuan-liang Zong : A discretization method for the nonlinear state delay system. Information technology journal, 13(6), pp.1222-1227, 2014.

[18] Yuan-Liang Zhang: Discretization of nonlinear non-affine time delay systems based on second-order hold. International journal of automation and computing, 11(3),pp 320-327 june 2014.

[19] Z. Kowalczuk.: Discrete approximation of continuous-time systems. A survey, proc. Iee-g,144, pp. 264–278, 1993.

[20] Zidong Wang, James lam, senior and Xiaohui Liu: Nonlinear filtering for state delayed systems with markovian switching. Ieee transactions on signal processing, vol. 51, no. 9, september 2003.

[21] Z. Qing-chang: Robust control of time-delay systems.Ieee transactions on automatic control ,53, pp. 636-637, 2008.

[22] Zheng Zhang and KilTo Chong: second order hold and Taylor series based discretization of siso input time-delay systems. Journal of mechanical science and technology 23 pp136-148, 2009.

New Data Clustering Algorithm (NDCA)

Abdullah Abdulkarem Mohammed Al-Matari
Information Technology Department,
Faculty of Computers and Information,
Cairo University, Cairo, Egypt

Prof. Imane Aly Saroit Ismail
Information Technology Department,
Faculty of Computers and Information,
Cairo University, Cairo, Egypt

Dr. Amira Mohammed Kotb
Information Technology Department,
Faculty of Computers and Information,
Cairo University, Cairo, Egypt

***Abstract*—Wireless sensor networks (WSNs) have sensing, data processing and communicating capabilities. The major task of the sensor node is to gather the data from the sensed field and send it to the end user via the base station (BS). To satisfy the scalability and prolong the network lifetime the sensor nodes are grouped into clusters. This paper proposes a new clustering algorithm named New Data Clustering Algorithm (NDCA). It takes optimal number of the clusters and the data packets sent from the surrounding environment to be the cluster head (CH) selection criteria.**

***Keywords*—Wireless sensor network; clustering; energy efficiency; cluster head selection**

I. INTRODUCTION

Wireless Sensor Networks (WSNs) is composed of hundreds or thousands of nodes that cooperatively sense the environment to provide interaction between computers or persons and the surrounding environment [1]. A wireless sensor network (WSN) consists of a large number of small sensors with low battery power, low cost, limited storage, a radio transceiver, a tiny microprocessor and a set of transducers [2]. The main task of a sensor node is to gather data from the sensed field and sends it to a common destination called base station (BS). Energy efficiency is the major design goal in these networks. The problem is how to develop a clustering technique that takes into consideration the energy consumption of the sensor nodes. This paper produces a New Data Clustering Algorithm (NDCA) that aims to improve the clustering algorithms in the wireless sensor networks by reducing the energy consumption and maximizing the lifetime of WSNs. The contributions of the NDCA algorithm are the fair divisions of the network area and the change in the cluster head selection criteria.

This paper is organized as follows: Section 2 illustrates the literature review of the clustering algorithms that are used in the wireless sensor networks. Section 3 describes the proposed algorithm in details. Section 4 presents the description of the simulation environment and the performance evaluation of the proposed algorithm. Finally, section 5 presents the conclusion and suggests future work.

II. LITERATURE REVIEW

There are many clustering approaches have been proposed by the research community to address the challenging of WSNs lifetime issue.

The Low Energy Adaptive Clustering Hierarchy (LEACH) protocol is the first attempts in the field of nodes clustering in WSN [3]. The essential idea behind LEACH is a CH rotation role among all the nodes to achieve load balancing. Each sensor elects itself to be a local cluster-head at any given time with a specific probability. Each node generates a random number between 0 and 1, then compares this number with the threshold T(n), if the random number is less than the pre-determined threshold T(n), the node is selected as CH.

The Hybrid Energy-Efficient Distributed clustering (HEED) protocol is a multi-hop clustering protocol for WSNs [4]. It periodically selects CHs based on a hybrid of the node remaining energy and a secondary parameter. The secondary parameters can be the node adjacency to its neighbors or the node degree.

Rotated Hybrid Energy-Efficient and Distributed clustering (R-HEED) protocol developed the HEED protocol by preventing the clustering approach at every round according to certain rules and schedule [5]. At the beginning of every round, CHs wait a period for a re-clustering message from the BS. If they do not receive a re-clustering message, they continue rotating the cluster head role within the same cluster. However, this protocol does not take into account the CH selection criteria which stay randomly rotating.

The Hybrid Energy Efficient Reactive (HEER) protocol is designed to deal with the characteristics of active homogeneous WSNs [6]. In HEER, the Cluster Head (CH) selection is based on the ratio of the residual energy of the node and the average residual energy of the network. Moreover, to conserve more energy, it introduced Hard Threshold (HT) and Soft Threshold (ST) as constraints when the data packets are transmitted over the network.

Another well-known but more efficient hierarchical-based algorithm is LEACH Inner Cluster Election algorithm (LEACH-ICE) [7]. The algorithm selects the cluster head based on the residual energy of the node using the following equation:

$$T_{(n)} = \frac{p}{1-p[r \; mod(1/p)]} * \frac{E_{n\_resident}}{E_{n\_initial}} , \qquad n \in G \qquad (1)$$

Where $T(n)$ is the threshold of nodes to be selected as CH, P is the probability of the node to be CH, r is the current round, *En_resident* is the resident energy of the node, *En_initial* is the initial energy of the node and G is the set of nodes that have not been cluster-heads in the last $\frac{1}{p}$ rounds. Also, the LEACH-ICE algorithm specifies that some cluster head should choose the nearest node inside the cluster to be the new cluster head, when the resident energy of the node is lower than *(∂ Esend*), where ∂ is a constant value equal to 31.5 and *Esend* is the node's minimum energy level equal to 0.02 Joule.

Distributed Energy-Efficient Clustering (DEEC) protocol, In this protocol, the CH selection criteria is based on the ratio of the residual energy of each node and the average energy of the network [8].

An Efficient Ad-Hoc Routing using a Hybrid Clustering Method in a Wireless Sensor Network algorithm uses a single set-up process to achieve high energy efficiency in wireless sensor networks [9]. It relies on the rotation sequence for selecting a CH instead of the random rotation. The CH node is replaced in each round according to a schedule based on an internal procedure within a sensor node without sending or receiving any additional information.

Energy-Balanced Unequal Clustering (EBUC) protocol is a centralized protocol that organizes the network in unequal clusters and the CHs relay the data of other CHs via multi-hop routing [10]. The operation of clustering is done by the BS. The BS computes the average energy level of each node and uses this information for CH selection.

III. THE PROPOSED ALGORITHM (NDCA)

This paper proposed a new data clustering algorithm named New Data Clustering Algorithm (NDCA).

The nodes are generated randomly and distributed in the network region. The BS gets the nodes locations. Then the network area is divided by the BS to zones that have the same size called clusters. After the nodes have been sensing their environment, some nodes have more events in their field than others. So the nodes have an unequal number of the data packets at a period of time, based on the place of the event. So the minimum number of the data packets is the important factor in this phase. So the node that has the minimum number of packets to send is selected as a cluster head. Also, the residual energy of a selected CH must be higher than the pre-determined threshold. The cluster head is replaced only when its level of the energy falls under a specific threshold. The threshold is computed using formula:

$$Threshold = C * E_{send} \quad (2)$$

Where C is a constant number equal to 31.5, *Esend* is the minimum energy level that enables the node to send packets, which is equal to 0.02 Joule.

Then cluster head broadcasts its unique number to all the nodes that allocated in the same cluster. Fig 1 shows the clustering setup stage of the NDCA algorithm. The NDCA algorithm depends on the LEACH-ICE algorithm with changing in the CH selection criteria. The figure shows the additional steps are used by NDCA over LEACH-ICE.

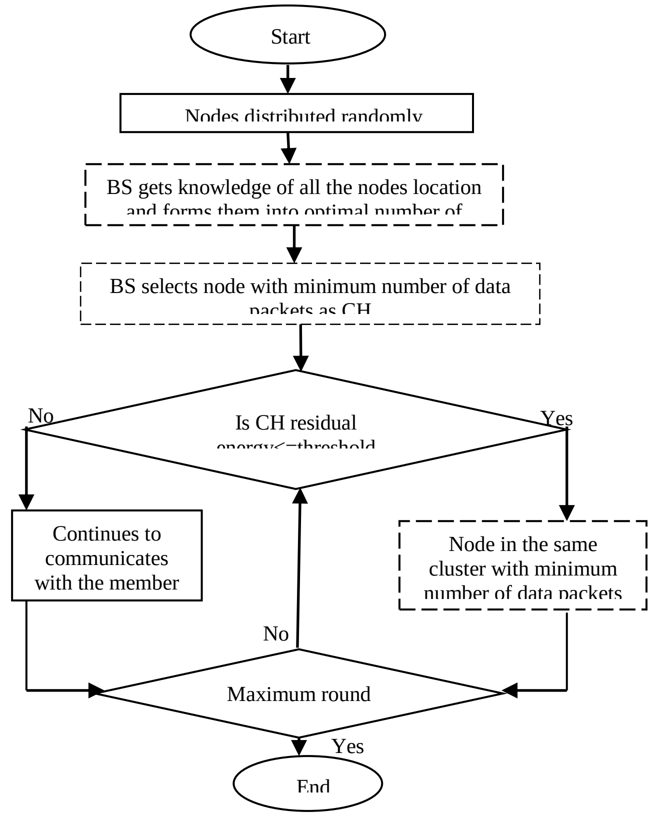


Fig. 1. Proposed NDCA Algorithm Flow Chart

— — — — Additional Steps Compared to LEACH-ICE

IV. SIMULATION AND RESULTS

A simulation program was built by MATLAB for both the NDCA algorithm and the LEACH-ICE algorithm to evaluate their performance and compares between them. The network consists of 100 and 250 nodes. Randomly, nodes deploy in the network into an area of 100m x 100m and 250m x 250m respectively. The base station is located away from the network area and have unlimited resources. The position of the base station (BS) is at location (50,175) and (125,300) respectively. Also it knows the information of the node's ID and location. All nodes have the same capabilities of initial energy, processing resources, and communication. The energy consumption for reception and transmission to all nodes are the same. The nodes states can be active or sleep. The energy consumption for the data transmission depends on the distance between (the sender and the receiver). A node dies when the residual energy is below or equal to 0.02 Joule [7].

Table l lists the used simulation parameters [8, 11].

TABLE I. SIMULATION PARAMETERS

The Parameter	The Value
Number of nodes	100, 250
Size of network	(100 x 100 m^2), (250 x 250 m^2)
Base Station position	(50 x 175 m^2), (125 x 300 m^2)
Packet size	3500 bits
Control packet	150 bits
Initial energy	2.5 Joule
Threshold of energy depleted	0.02J
Energy dissipated per bit (Eelec)	50nJ/bit
Data aggregation energy consumption (EDA)	5nJ/bit
Distance of threshold (do)	87 m
Amplifier of transmit if $D_{BS}>=$ do (Emp)	0.0013pJ/bit/ m^4
Amplifier of transmit if $D_{BS}<=$ do (Efs)	10nJ/bit/ m^2
Method of deployment	Hybrid

A. Performance Metrics

Five evaluation metrics are used to evaluate and test the performance of the proposed NDCA algorithm and compare it to the LEACH-ICE algorithm.

1) Average Energy Consumption

It is the average energy consumption for all the alive nodes of the whole network.

$$\text{Average Energy Consumption} = \sum_1^n \frac{\text{energy consumption of nodes}}{\text{Number of alive node (n)}} \quad (3)$$

It is better to keep it as a least as possible.

2) Number of Alive Nodes per Round

It is the lifetime interval of the nodes between the network operation start until the death of the last node. It is computed as follows:

$$\text{Average number of alive nodes} = \sum_1^n \frac{\text{Number of alive nodes}}{\text{Number of round (n)}} \quad (4)$$

It is better to maximize the metric.

3) Number of Various Nodes Dying per Round

a) FND (First Node Dies): It is the first node that has depleted its whole energy (specified threshold) in a round during the network life time.

b) HND (Half of the Nodes Die): It is the rounds (time) elapsed until half of the nodes have consumed their whole energy (specified threshold).

c) LND (Last Node Dies): It is the last node that has exhausted its whole energy (specified threshold) in a round during the network life time.

It is better to maximize the metric.

4) Number of Packets Delivered to Cluster Head

It is the number of packets that the CHs received over the network lifetime.

$$\text{Average Packets Delivered to CH} = \sum_1^n \frac{\text{Number of packets reached to CH}}{\text{Number of round (n)}} \quad (5)$$

It is better to maximize the metric.

5) Number of Packets Delivered to Base Station

It is the total number of packets that the base station received during the network lifetime.

$$\text{Average Packets Delivered to BS} = \sum_1^n \frac{\text{Number of packets reached to BS}}{\text{Number of round (n)}} \quad (6)$$

It is better to maximize the metric.

B. Evaluated Scenarios

A simulation program was built for two scenarios. In the first scenario the network size is 100 nodes in (100m x 100m) area and the base station location is at (50m x 175m). While in the second scenario the network size is 250 nodes in (250m x 250m) area and the base station location is at (125m x 300m).

1) First Scenario

In this scenario; the network size is 100 nodes distributed over (100m x 100m) area and the base station location is at (50m x 175m).

a) Average Energy Consumption

Fig 2 shows the energy consumption for both NDCA and LEACH-ICE algorithms. Ten rounds have been extracted randomly from along the rounds range, in order to compare the proposed NDCA with LEACH-ICE algorithms. It is obvious from the figure that the NDCA algorithm saves energy more than that of the LEACH-ICE algorithm by 27%. This energy conservation is due to the criteria of CH selection which chooses the node that has the minimum packets to send to be selected as CH. Unlike LEACH-ICE which it is based on the random distribution of CHs and selection criteria.

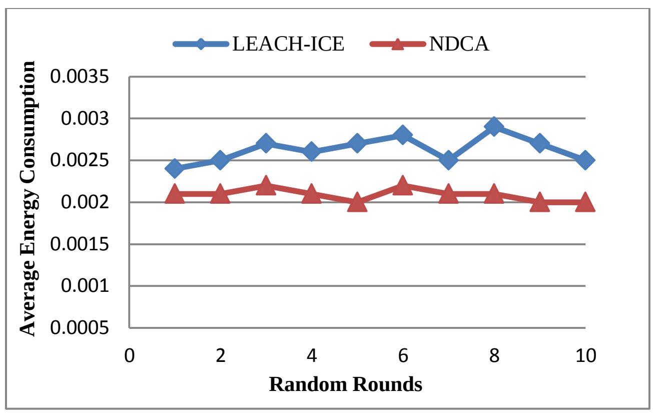


Fig. 2. Average Energy Consumption

b) Number of Alive Nodes per Round

Fig 3 shows the number of alive nodes per rounds for the NDCA and the LEACH-ICE algorithms. It is obvious from the figure that the proposed NDCA algorithm achieves better performance than that of the LEACH-ICE algorithm. The proposed algorithm extended the network life-time more than LEACH-ICE by 272 rounds, which is about 16%. This improvement is due to the energy conservation during the clusters setup phase achieved, as while LEACH-ICE treats all the nodes without discrimination according to the data load of a node, but NDCA takes data load of nodes into account; therefore, has a longer period of the network lifetime than LEACH-ICE.

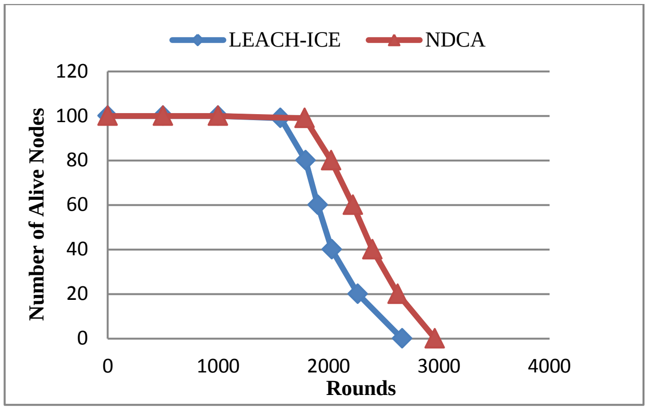


Fig. 3. Number of Alive Nodes per Round

c) Number of Various Nodes Dying per Round

Fig 4 demonstrates the performance measurement in term of the number of various of nodes that die per round for NDCA and LEACH-ICE algorithms. The graph shows that the lifetime of FND metric in the NDCA algorithm compared to the LEACH-ICE algorithm increased by 14%, HND metric by 16%, and LND metric by 10% respectively. Since NDCA optimize energy consumption of nodes communication, as it keeps the cluster head with a light load, high residual energy and with fair distribution which leads node to be close to the cluster head, this has a vital impact on the lifetime of the nodes in the network.

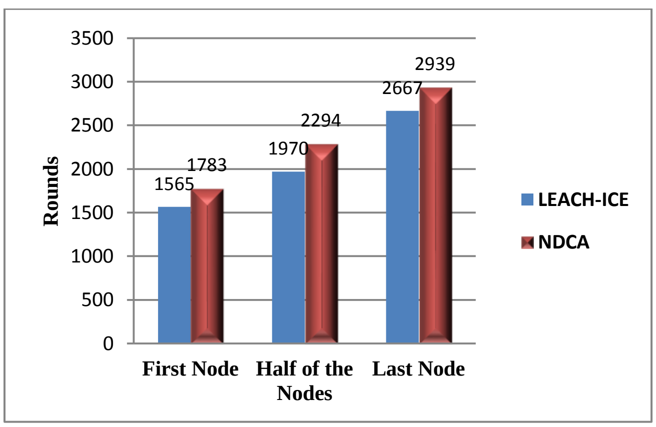


Fig. 4. Number of Various Nodes Dying per Round

d) Number of Packets Delivered to Cluster Head

Fig 5 shows the number of packets delivered to the CH for the NDCA and LEACH-ICE algorithms. The figure obviously shows that the data packets transmitted over throughout the network to CHs by NDAC are more than that transmitted by LEACH-ICE. The proposed algorithm achieved better performance than that of the LEACH-ICE algorithm by 16% in this term. This achievement is due to the increase in the network lifetime produced by the process of the cluster heads distribution and selection. The number of data packets sent to the CHs is slightly more than that of LEACH-ICE because the node with a minimum number of the packet is selected as CH, so others nodes with a large number of packets send to the CH.

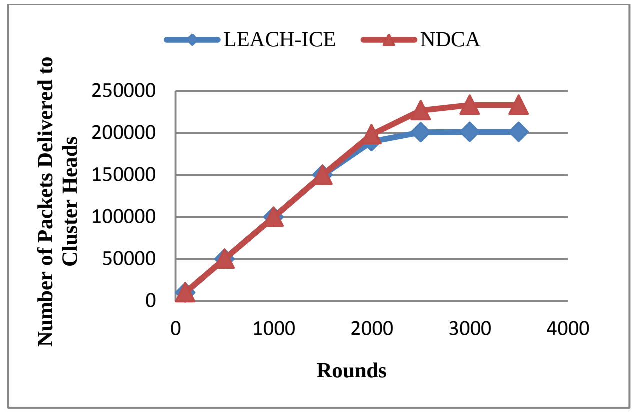


Fig. 5. Number of Packets Delivered to Cluster Heads

e) Number of Packets Delivered to Base Station

Fig 6 shows the number of packets delivered to the BS for the NDCA and LEACH-ICE algorithms. It is clear from the figure that the data packets transmitted through the network from the CHs to the BS by the NDAC algorithm are more than that transmitted by the LEACH-ICE algorithm. The proposed algorithm outperforms the LEACH-ICE algorithm by 18% in this term. Since in NDCA, CHs received much more data packets than that in LEACH-ICE as mentioned in the above subsection, cluster heads periodically aggregate this data and then send them to the BS.

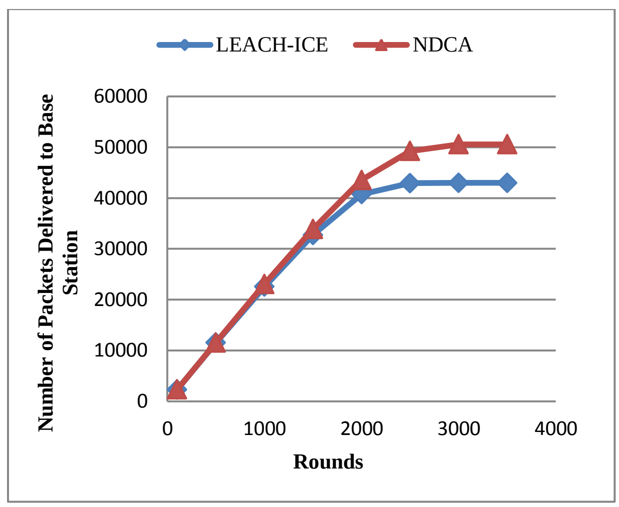


Fig. 6. Number of Packets Delivered to Base Station

2) Second Scenario

In this scenario the network size is 250 nodes distributed over (250m x 250m) area and the base station location is at (125m x 300m).

a) Avergae Energy Consumption

Fig 7 displays the energy consumption for both NDCA and LEACH-ICE algorithms. Fig 7 shows that the NDCA algorithm consumes less energy than that of the LEACH-ICE algorithm; therefore, the NDCA algorithm saves energy by 38%. This improvement is due to the fact that the LEACH-ICE algorithm consumes more energy than NDCA because of the unfair distribution of CHs that causes a large distance between the node and the CH. Also, CH selection criteria plays a significant role in prolonging the lifetime of the network.

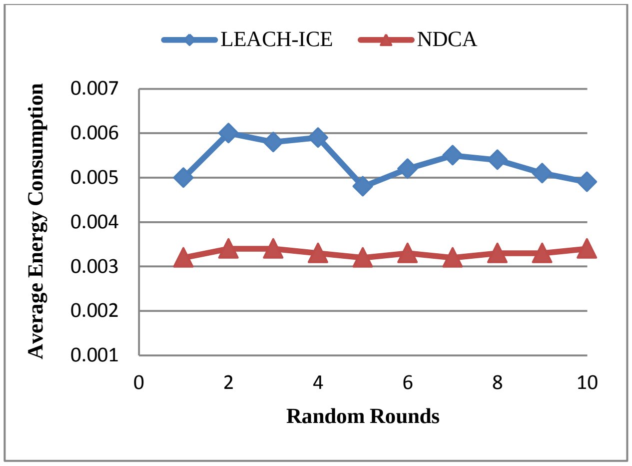


Fig. 7. Average Energy Consumption

b) Number of Alive Nodes per Round

Fig 8 shows the number of the alive nodes per rounds for the NDCA and LEACH-ICE algorithms. It is clear from the figure that the proposed algorithm saves great amount of the energy in this scenario. Therefore, the NDCA algorithm achieves better performance than that of the LEACH-ICE algorithm by prolongs the network life-time by 411 round, which is about 26%. While in the first scenario the proposed algorithm extends the network life-time by 272 round. This improvement is due to the selection of cluster head based on the minimum data packets to send. Also the residual energy of a node is taken into consideration.

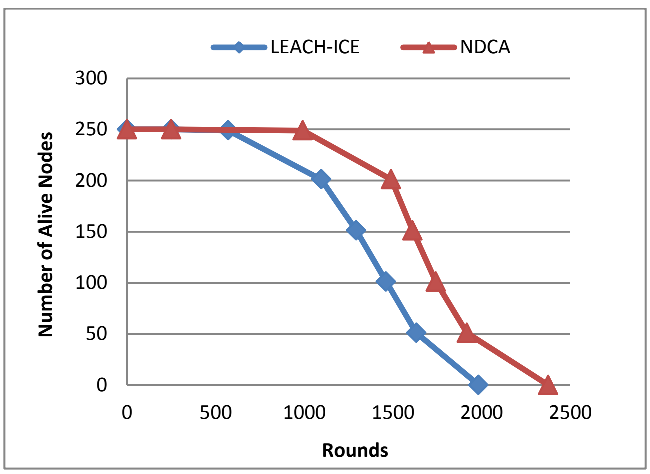


Fig. 8. Number of Alive Nodes per Round

c) Number of Various Nodes Dying per Round

The numbers of various nodes dying per round for the NDCA and LEACH-ICE algorithms are illustrated in fig 9. It is obvious from the figure that the NDCA algorithm achieves better performance than that of the LEACH-ICE algorithm by 73% for the FND, by 23% for the HND and by 20% for the LND. The increase achievement percentage compared to the first scenario is by 59%, 7% and 10% respectively. This enhancement is also due to the selection of cluster head based on the minimum data packets to send and the residual energy parameter of the nodes.

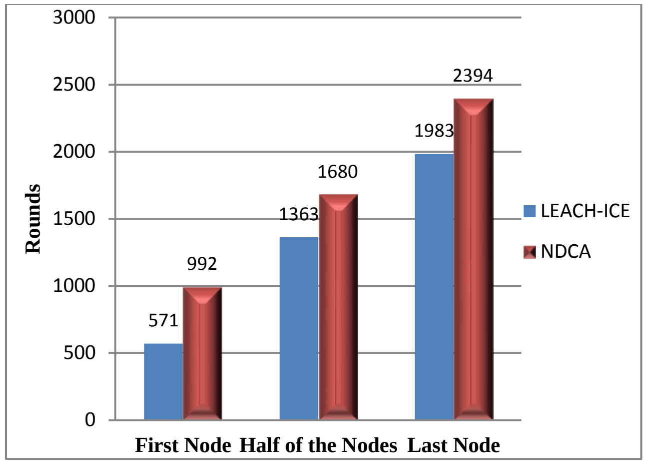


Fig. 9. Number of Various Nodes Dying per Round

d) Number of Packets Delivered to Cluster Head

Fig 10 demonstrates the number of packets delivered to CH for the NDCA and the LEACH-ICE algorithms. It is observed from the figure that the total packets sent by the proposed NDCA algorithm to the cluster heads in each cluster are much more data packets than that of the LEACH-ICE algorithm. The proposed NDCA algorithm overcomes the LEACH-ICE algorithm in this metric by 26%. The difference in the amount of data packets send to CHs between the first scenario and the second scenario is by 10%. This is because, here the nodes that forward the data packets have huge number of packets, while the CH who received the data packets are selected based on the minimum number of data packets to send, this saves energy for prolonged lifetime. On other hand; in NDCA, the nodes are alive for longer time compared to LEACH-ICE.

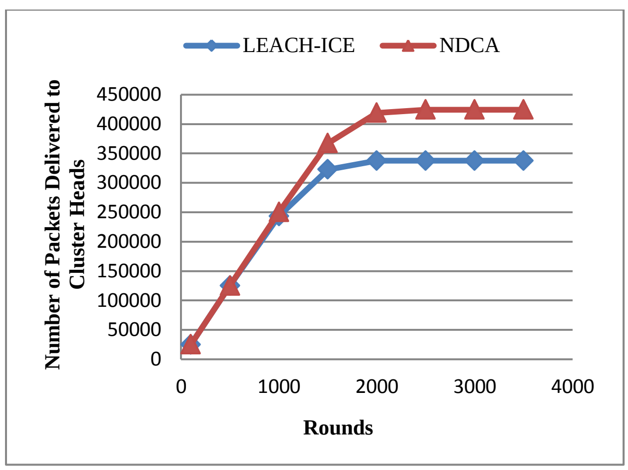


Fig. 10. Number of Packets Delivered to Cluster Heads

e) Number of Packets Delivered to Base Station

Fig 11 shows the number of packets delivered to BS for the NDCA and LEACH-ICE algorithms. Fig 11 displays that the packets received at the base station by the NDCA algorithm are much higher than that received by the LEACH-ICE algorithm.

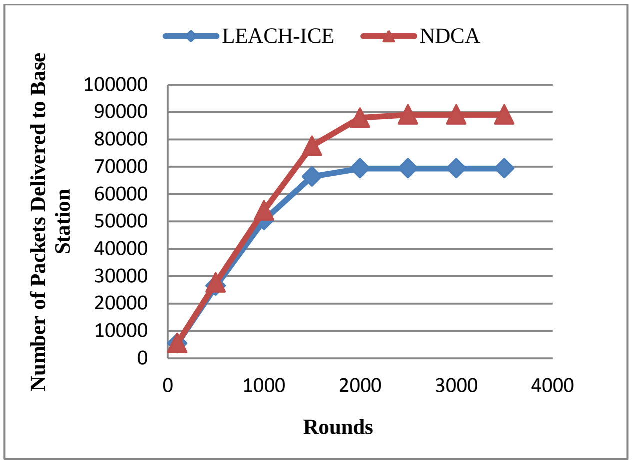


Fig. 11. Number of Packets Delivered to Base Station

In this term, the result of the proposed NDCA algorithm improves the LEACH-ICE algorithm by 28%. Better network cluster head distribution and efficient cluster head selection criteria are two main reasons of increased the data packets in the NDCA algorithm.

3) General Comparison

Table 2 contains summarization of the implementation results and the comparison of the algorithms; LEACH-ICE and the proposed NDCA. From the table, It can be observed that the average values of all the performance metrics when applying the proposed NDCA algorithm overcome those when applying the LEACH-ICE algorithm in the all cases.

TABLE II. SUMMARIZES THE COMPARISON BETWEEN NDCA AND LEACH-ICE USING TWO SCENARIOS

Performance Matric	**Average Percentage of Improvement Using NDCA**	
	First Scenario	***Second Scenario***
Average Energy Consumption	27%	38%
Average Alive Nodes	16%	26%
First Node Die	14%	73%
Half of the Nodes Die	16%	23%
Last Node Die	10%	20%
Number of Packets Delivered to Cluster Head (CH)	16%	26%
Number of Packets Delivered to Base Station (BS)	18%	28%

V. CONCLUSION AND FUTURE WORK

In this paper, a clustering algorithm, named NDCA, for WSN is proposed. This algorithm chooses the node with minimum data packets to send to be as CH and the residual energy of the new CH must be higher than that of the predefined threshold level. Also the proposed algorithm divided the network into zones (clusters). A simulation program was built for both the NDCA algorithm and the LEACH-ICE algorithm to evaluate their performance and compare between them. Two different scenarios are used to evaluate the proposed NDCA algorithm. It is concluded from the results, that the proposed NDCA algorithm outperforms the LEACH-ICE algorithm on all the performance metrics.

For future work, there are many important subjects such as: introducing the clustering technique for Wireless Sensor Network over Cloud Computing. It is important to add an intelligent method for secure communications in this filed.

REFERENCES

[1] R. Verdone, D. Dardari and G. Mazzini, "Wireless Sensor and Actuator Networks," Elsevier, London, UK, 2008.

[2] R. Badi, P. Suri and S. Gupa, "Review Paper on Various Clustering Protocols Used In Wireless Sensor Network (WSN)," Electronics, Signals, Communication and Optimization (EESCO) International Conference, Visakhapatnam, pp. 1- 4, Jan 2015.

[3] W. Heinzelman, A. Chandrakasan and H. Balakrishnan "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," Proceedings of the 33rd Hawaii International Conference on System Sciences, Maui , HI , USA, IEEE, 4-7 Jan 2000.

[4] O. Younis and S. Fahmy ," A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad hoc Sensor Networks," Mobile Computing, IEEE Transactions, vol. 3, no. 4, pp. 366-79, 2004.

[5] W. Mardini, M. B. Yassein, Y. Khamayseh and B. Ghaleb, "Rotated hybrid, energy-efficient and distributed (R-HEED) clustering protocol in WSN," WSEAS Trans. Comm, vol. 13, pp. 275–290, 2014.

[6] S. Mohammed, K. Latif and U. Qasim, "HEER: Hybrid Energy Efficient Reactive protocol for Wireless Sensor Networks," Electronics Communications and Photonics Conference (SIECPC), 2013 Saudi International, Riyadh, pp. 1 – 4, April 2013.

[7] Y. Miao, "Cluster-head Election Algorithm for Wireless Sensor Networks based on LEACH Protocol," Trans Tech Publications, Switzerland, vols. 738-739, pp. 19-22, 2015.

[8] L. Qing, Q. Zhu and M. Wang, "Design of a Distributed Energy-Efficient Clustering Algorithm for Heterogeneous Wireless Sensor Networks," ELSEVIER, Computer Communications 29, pp. 2230- 2237, 2006.

[9] D. Nam and H. Min, " An Efficient Ad-hoc Routing using a Hybrid Clustering Method in a Wireless Sensor Network," Wireless and Mobile Computing, Networking and Communications, Third IEEE International Conference, New York, USA, p. 60, Oct 2007.

[10] J. Chang, S. Wei, X. Min and T. Xian, "Energy-Balancing Unequal Clustring Protocol for Wireless Sensor Network," The Journal of Chain Universities of Posts and Telecommunications, vol.4, p. 94–99, August 2010.

[11] R. N. Enam, S. Misbahuddin and M. Imam, "Energy efficient round rotation method for a random cluster based WSN," Collaboration Technologies and Systems (CTS), International Conference, Denver, vol. 8, p. 157 – 163, May 2012.

Load Balancing in Partner-Based Scheduling Algorithm for Grid Workflow

Muhammad Roman
Institute of Information Technology
Kohat University of Science and Technology
Kohat, Pakistan

Asad Habib
Institute of Information Technology
Kohat University of Science and Technology
Kohat, Pakistan

Jawad Ashraf
Institute of Information Technology
Kohat University of Science and Technology
Kohat, Pakistan

Gohar Ali
Information Systems and Technology
Sur University College
Sur, Sultanate of Oman

***Abstract*—Automated advance reservation has the potential to ensure a good scheduling solution in computational Grids. To improve global throughput of Grid system and enhance resource utilization, workload has to be distributed among the resources of the Grid evenly. This paper discusses the problem of load distribution and resource utilization in heterogeneous Grids in advance reservation environment. We have proposed an extension of Partner Based Dynamic Critical Path for Grids algorithm named Balanced Partner Based Dynamic Critical Path for Grids (B-PDCPG) that incorporates a hybrid and threshold based mechanism to achieve load balancing to an allowed value of variation in workload among the resources in Partner Based Dynamic Critical Path for Grids algorithm. The proposed load balancing technique uses Utilization Profiles to store the reservation details and check the loads from these profiles on each of the resources and links. The load is distributed among resources based on the processing element capacity and number of processing units on resources. The simulation results, using Gridsim simulation engine, show that the proposed technique has balanced the workload very effectively and has provided better utilization of resources while decreasing the workflow makespan.**

Keywords*—Load Balancing; Advance Reservation; Resource Utilization; Workflow Scheduling; Job Distribution

I. Introduction

Large-scale distributed and parallel computing has been changed by the growth of the Internet, powerful computing and network speed. With the coordination of distributed computing power, resources and applications, Grid computing has emerged as distributed computing platform. It utilizes the power of wide range of heterogeneous distributed resources for execution of compute- and data-intensive applications. It provides consistent, inexpensive, and pervasive access to geographically widely distributed resources to solve large scale scientific, engineering, and commerce problems. The resources joining the Grid are independent of each other in terms of performance, memory speed, and bandwidths. They are owned and administered by different providers. The motivation of Grid computing is to provide users and applications a seamless access to a number of high performance resources by creating illusion of a single system image [1]. User submitted applications are distributed among various Grid resources and executed in parallel. This makes the execution of applications efficient. In order to decrease the overall execution time of the application, effective and efficient load balancing algorithms are needed to be designed for Grid computing. Advance reservation (AR) provides facility to the user to reserve CPU and bandwidth for an application before the actual execution of the application [2]. AR jobs cause other non-AR jobs to wait as the resources are reserved in advance. Also consecutive reservations may leave fragments between them which cannot be used by other jobs, leaving un-used empty time slots. This leads to the under utilization of Grid resources. We are trying to fit the jobs in such a way that it is balancing the load over resources by considering all resource's load, which results in reducing empty fragments between the consecutive reservations. Although a lot of work has been done in non-AR environment for load balancing, yet there is a little contribution by the researchers in providing a solution for load balancing in advance reservation environment [3]. The objective is to prevent the condition where some of the resources are heavily loaded by the tasks submitted by users and others are lightly loaded or not being utilized at all [4]. The advantages of implementing good load balancing policies are better utilization of resources, low rejection rate, minimized wait time, high performance, maximized throughput, and reduced cost of job execution. Some of them are good for the resource providers, in terms of, resource utilization and throughput while others are good from Grid user's perspective, in terms of, reduced cost and minimized completion time of application.

Large scale scientific applications, modeled as workflow and represented as Directed Acyclic Graph (DAG), are submitted to Grid by using Workflow Management System [5]. For efficient execution of workflow good scheduling heuristics are trivial. Partner Based Dynamic Critical Path for Grids (PDCPG) proposed in [6], is one of the proposed algorithms for advance reservation environment in Grid and is based on Dynamic Critical Path for Grids (DCPG) [7]. PDCPG tries to schedule partner jobs on same resource in

Sur University College, Sur, Sultanate of Oman (*sponsor*)

order to minimize the communication involved to transfer required files of the child jobs. This may lead the overall Grid to an un-balanced state in which some of the resources are highly utilized while others remain under-utilized. We have proposed a hybrid policy restricted load balancing technique named as B-PDCPG to solve the problem of load balancing in PDCPG.

The rest of the paper is organized as follows: Section 2 presents load balancing problem. Section 3 explains related work to solve load balancing problem. Section 4 defines keywords and explains system model. Section 5 explains proposed load balancing technique and section 6 discusses observations and results. The paper is concluded in Section 7 and future work is discussed.

II. Load Balancing

Grid architecture involves a large number of geographically distributed worker nodes connected together to achieve a level of performance. However, increasing the number of worker nodes does not always guarantee increased level of computing power. The resources involved in the system must be used such that all of the resources are utilized appropriately. The unequal demands and heterogeneity of Grid resources leads to the problem of job distribution. Algorithms try to map jobs on resources such that all the resources are utilized equally and makespan of the applications is reduced. Workload is the amount of work to be done by resource which can be heavy, light or moderate. Load balancing is sometime confused with load sharing and load leveling. Load sharing confirms that there is no idle node when there is highly loaded node in the Grid. This is the most basic level of load distribution which only checks if there is a load available on a resource or not, i.e., it is viewed as a binary set. Load balancing is the finest form of load distribution which tries to achieve strictly balanced distribution of load among all of the resources of the system. Load leveling is in the middle of the two extremes of load distribution which seeks to avoid congestion on any of the resources in the system.

For defining load balancing algorithm we need to implement certain policies. Following are some of the policies of load balancing [8]:

A. Information Policy

Defines what information is required for load balancing algorithm at what time and from where. It decides the type of information that will be collected, based on which algorithm will take decision. The information may include both static, like number of CPUs, size of memory, and dynamic information, like current load on the resources, memory being utilized. It also decides when to collect and update this information, to understand the current status of the system. Updating this information very frequently will lead to communication overhead. Normally a special information agent does this job either after a specific period of time or when an event is triggered [9]. Agents may directly or indirectly collect this information from the working nodes of the system. The effectiveness of load balancing algorithm is very much associated with how the load information is gathered. Simple load index calculates load based on one metric like CPU load, bandwidth utilization, or disk storage, while complex load index combines more than one metrics to aggregate them to single load information. The information gathered is then exchanged periodically, on demand, or when the state of a node has been changed [9].

B. Transfer type Policy

Determines when to start load balancing and transfer of load from a sender to receiver of the load. It performs the classification of resources as source or receiver based on the information gathered from information policy and the current status of the working nodes based on a predefined threshold value [10]. The resource which is heavily loaded and is going to transfer load to another resource is called sender and the resource which is going to receive load from other resource is called receiver.

C. Selection Policy

Decides which of the processes should be transferred from overloaded nodes (source) to the idle nodes (receiver). The criteria of selection may be defined at sender or receiver nodes based on the type of initiation policy. In sender-initiated approach, the sender decides which tasks should be selected for transfer. In receiver-initiated approach, the receiver defines the selection policy in terms of what type of, how big, and how many. Normally the criteria is based on; shortest remaining time, longest remaining time, First-in-First-Out (FIFO), Last-in-First-Out (LIFO), or a process is randomly selected [8].

D. Location Policy

The responsibility of location policy is to find a suitable partner node for the heavily loaded node for transferring some of its load to it. This policy works on the basis of information collected in information policy. In sender-initiated approach, heavily loaded node tries to find a lightly loaded node to share the load with it. In receiver-initiated approach lightly loaded node searches for a node which is heavily loaded. As the load information can be centralized or distributed, location selected decision can also be central or distributed. In case of distributed, each node tries to find a light resource in its neighbors and from the partial information saved at it. A node which has light load on it may not always be selected by location decision, other properties such as communication link and processing power are also consider for selecting it. Ultimately, the goal is to reschedule a job, which is already scheduled on a highly loaded resource, on a resource after which it will be executed earlier and will not delay other jobs or increase the makespan of the application.

Load balancing problem is described by different terminologies by many literatures which gives different meanings to the same problem in different situations. In [8] a detailed hierarchical model of load balancing is defined, as shown in Fig. 1.

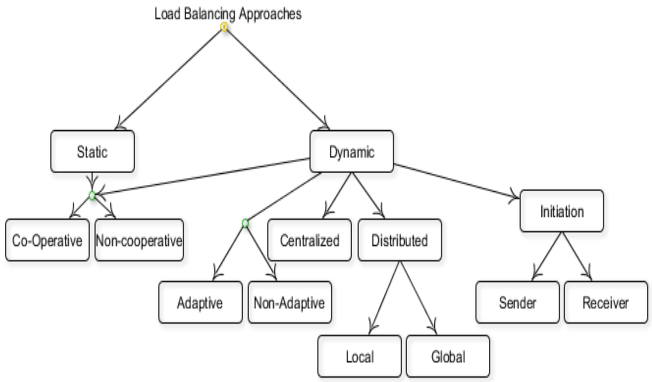


Fig. 1. Hierarchical Taxonomy for Load Balancing Approaches

A. *Static vs. Dynamic Load Balancing*

Static load balancing is based on the static information such as CPU capacity, memory size, link speed, and etc., whereas dynamic load balancing is based on the current status of the system and working nodes [11]. Static load balancing is good as it does not require information to be collected at all the time but this may lead to low utilization rate [12]. Dynamic load balancing algorithm may change the distribution of tasks among resources at runtime. This is similar to job re-scheduling, where the scheduled jobs are rescheduled after some changes are encountered in the system, like resource failure. In dynamic load balancing, if a resource is assigned large number of tasks and there are some lightly loaded resources available in the system, it will transfer some jobs to the resources which are lightly loaded or idle, selected on the bases of selection policy. A question arises, to which resource it has to be transferred and who will decide when to transfer them. There are two approaches in common: sender-initiated and receiver-initiated. In sender-initiated, a heavily loaded node requests to transfer some of its load to an idle or lightly loaded resource. In receiver-initiated, an idle resource initiates the process of transfer [13]. In some cases a combination of these two is used in which both sender and receiver can initiate the process. The main problem of the dynamic load balancing is how to distribute work among the resources as the resources are continuously changing their status, especially in case of heterogeneous system like Grid where the number of resources and bandwidth changes more frequently.

B. *Centralized vs. Distributed*

The load balancing decision may be taken at a centralized location or this may be taken at various levels of the system as in [14] and [3] where Grid is represented as a tree having four different levels. Centralized algorithm is easy to implement but has the problem of scalability, fault tolerance, and also becomes a bottleneck as all the decisions are to be made at one central location. In distributed strategy the tasks are distributed among different locations, where each location takes decision of its own region and sometime communicates with each other to take a global decision. This approach is difficult to implement but is highly scalable, fault tolerant and the load is divided which prevents them from becoming a bottleneck.

C. *Adaptive vs. Non-Adaptive*

In adaptive techniques the decision is based on the past and current system behavior and is based on the previously taken decisions and the changes in the environment. Confusion may arise as to differentiate between dynamic and adaptive techniques. The dynamic load balancing technique keeps into account the current status of the system at that time. On the other hand, adaptive technique considers the environmental stimuli in taking load balancing decision.

D. *Local vs. Global*

In local load balancing, each processing element gets information from its neighbor nodes and makes the decision based on this locally gathered information. In global load balancing, information from all or part of the system is collected to take the balancing decision. This may require a lot of information exchange.

E. *Initialization*

The task migration can by initialized by the sender, receiver or both (symmetric). In sender-initiated model, the resource which is highly overloaded transfers some of its selected processes to an idle or under loaded resource. In receiver-initiated model, the under loaded resource requests the system that it has the capacity to engage some more jobs. The probability of finding a lightly loaded resource is high in case of a lightly loaded system, therefore, sender-initiated algorithms works better in lightly loaded systems. In contrast, receiver-initiated algorithms are good in case of heavily loaded systems, as the probability of finding a highly loaded node is higher (heavily loaded nodes are more than lightly loaded nodes) [10]. A combination of these two policies is used in symmetrically initiated approach. The nodes behave intelligently by working as sender-initiated in low system load and receiver-initiated in highly loaded system.

F. *Co-operative vs. Non-cooperative*

Both in static as well as dynamic load balancing, nodes may or may not work together in taking balancing decision. In co-operative mode the nodes work together in order to make a decision that is based on the collective objectives of the overall system. In non-cooperative mode, however, individual system works autonomously in taking decision keeping their own objectives in account, irrespective of the effects on the rest of the system.

III. RELATED WORK

There are various ways to balance load over the resources connected to a Grid, mainly categorized as static and dynamic load balancing techniques. There are fewer studies on static approaches for Grid environment [15] due to the heterogeneity of the system. Dynamic load balancing is considered better for such systems, as it responds better to the changes in the system [16]. In dynamic load balancing, if a resource is assigned large number of tasks and there are some lightly loaded resources available in the system, it will transfer some jobs to the resources which are lightly loaded or idle, selected as per selection policy of the algorithm. The disadvantages of dynamic load balancing algorithms are clear; these policies are more complex and require a lot of communication to share dynamic state information. A good dynamic load balancing algorithm tries to minimize this cost in order to decrease the overhead and yet achieve the best load balancing. Some of the algorithms use combination of these two in order to make scheduling decisions, called hybrid load balancing algorithms [17]. In [18], a hybrid technique is proposed based on table of effectiveness to keep information of each Grid resource. When a resource is requested for a job, dispatcher selects it based on the information in the table of effectiveness and maps the job to the selected resource. If the resource is overloaded, dispatcher updates the table of effectiveness accordingly. In [9], the table of effectiveness is updated after a fix interval; consequently the execution of the task is not delayed. This algorithm keeps only static information of nodes, like remaining capacity of CPU and remaining memory. Reference [19] also updates the table on timely bases, the difference is that it stores dynamic load information rather than static load information.

Some other algorithms are developed based on fuzzy algorithms which are easy to implement and provide fast response time with good load balancing results [20]. Fuzzy algorithms are rule based algorithms and use knowledge of experts in creation of rules for a specific domain. Genetic based algorithms are also used for load balancing which performs the mapping of jobs to nodes by genetic operators which include three operators of reproduction, exchange and mutation [21]. Reference [22] combines the features of genetic algorithm, simulated annealing, and clonal selection algorithm and introduces genetic clonal annealing algorithm to task scheduling problem. Hierarchical load balancing algorithms try to achieve load balancing at different levels of Grid while representing the Grid as a tree model [14]. Agent based algorithms try to achieve load balancing by using agents working together to find load balancing [23]. Policy based approaches try to get an optimized scheduling under the constraint of load on a resource policy [24].

Our objective is to propose a load balancing mechanism in advance reservation environment. Partner Based Dynamic Critical Path for Grid (PDCPG), proposed in [6], is based on Dynamic Critical Path for Grid (DCPG) for AR environment. For job selection, it uses same technique as of DCPG. For resource mapping, it gets all the scheduled partner jobs of the selected most critical job in reverse order of their criticality. It calculates the completion time of the selected job on the resource on which its partner is scheduled. If the completion time is less than combined execution time and data transfer time of a partner job minus data transfer time of the selected job on that resource, then the resource is selected. Otherwise the next partner job is considered. In case, there was no partner job of a job or none of them was already scheduled on a resource, then a resource that gives minimum completion time is selected. In this way, PDCPG tries to minimize the communication overhead involved in transferring the required data files from parent jobs to the child jobs in a workflow. It encourages majority of the jobs to get scheduled on a single most efficient resource. This may lead to the problem of un-equal load on resources; some of the resources get highly loaded while the rest of them are under-loaded. We tried to balance the load in PDCPG by introducing a policy based technique.

The main contribution of this paper is to design a policy to get an optimized job scheduling solution under the constraint of resource usage policy. We adopt some ideas from DCPG and PDCPG to design our algorithm in AR environment for Grid.

IV. Proposed Load Balancing Algorithm

Before going into the detailed discussion of the problem and proposed solution, we explain the workflow models, resource model and other related terminology used in the proposed work.

DAG Workflow: Large business and scientific applications are submitted to Grid in the form of workflows, usually represented by DAG,G=(V,E), where V is the set of *v* nodes representing jobs and E is the set of *e* edges/lines connecting two nodes; parent (predecessor) and child(successor). The order in which parent and child nodes are connected shows the interdependency of the jobs. Edge(i,j) $\in$E shows that job *j* is dependent on job *i*, therefore, job *i* must be completed before scheduling job *j*. A job is said to be dependent on another job if it uses a file generated after execution of that job. The entry job is the only job which takes as input an already existing file, so having no incoming edge, while the rest of the jobs get the files as inputs which are generated by other jobs. F= $\{f_1, f_2, f_3, \ldots, f_g\}$ is the set of files which are submitted as required files to jobs and generated as result of execution of jobs in the workflow. DS_{f_i} denotes the size of a file f_i in bits, where $1 \leq i \leq g$. $PTR_i = \{ptr_1,\ ptr_2, \ldots, ptr_p\}$ is the set of partner jobs of job *i*, where jobs are said to be partners if they have at least one common direct child job. Job computation length is described in Million Instructions (MI) and resource speed is described in Million Instructions per Second (MIPs). Each workflow has a single entry job and single exit job; in case there are more than one entry jobs, a 0 MI length job is added to the start and end of the workflow. The average of the ratio of the computation cost and communication cost of the jobs in a workflow is called granularity of a job, except for the exit job *snk* having no child jobs.

$$Granularity = \frac{1}{v-snk}\left(\sum_{i=0}^{v-snk} \frac{\mathrm{expET(j_i)}}{\mathrm{ADTT_{j_i}}}\right) \qquad (1)$$

Grid Resources: All the resources in the proposed work support advance reservation and is represented as a set of $R = \{r_1,\ r_2, \ldots, r_m\}$. Resource r has W_r CPUs of same speedofPS_r. The user who submits the workflow for execution is connected to r_0, where $r_0 \notin R$. Workflow is submitted to the system from this resource and the starting files are also stored on it. All the resources are connected by a set of links L=$\{l_1,l_2,\ldots,l_n\}$, which support advance reservation. Capacity of the link l_i is measured in Mbps and denoted by C_{l_i}, whereas delay is measured in milliseconds and denoted by d_{l_i}. Our proposed work uses Optical Burst Switching (OBS) network architecture which is suitable for Grids supporting advance reservation [25]. In OBS, the files are transmitted in a single burst from source to destination. Resources and links can be requested and reserved by a user in advance. The reservation details are stored in utilization profiles [26] stored at each resource and link, called UP_r and UP_l, respectively. The reservation request can be made for a CPU called r_{cpu} or for a bandwidth calledr_{bw}. The availability of resource and link is checked against the utilization profile which contains the information of number of reserved CPUs and reserved bandwidth at time *t*. Further details of how to maintain reservation details and processing reservation request can be found in [24].

We have proposed a technique that prevents PDCPG from taking the overall system into an un-balanced state. It is a policy based technique which tries to schedule jobs on resources while implementing a policy to stop a resource from being over-loaded. The job selection technique is the same as inherited by PDCPG from DCPG. The resource selection technique of PDCPG has been modified to balance the load on the computing resources of the Grid. For implementing information policy; all the reservations and current load information of the system is kept in utilization profiles of the computing resources and communication links. At the time of resource selection this information is used to check the percent load on each of the resource, calculated as:

$$Load_r = \frac{\sum_{t=0}^{z_r} UP_r\,(t) * PS_r}{Z_r * PE_r * PS_r} * 100 \qquad (2)$$

where Z_r is the total size of UP_r. $Load_r$ is the percent load on the resource *r*. UP_r (*t*) is the number of CPUs already reserved at time *t* in the dimension Z_r. PE_r is the number of processing elements on resource r. PS_r in the numerator and denominator can be cancelled out, but they are shown in the calculation for better understanding. For selecting a resource R_i for a job *j*, it is checked if some of the partner jobs of job *j* PTR_j have already been scheduled. The resource, on which job *j*'s most critical job partner is scheduled, is selected. The selected job is mapped onto this resource if the following condition [18] is met:

$$expCT_j^{r_{ptr_i}} \le CT_{ptr_i} + ADTT_{ptr_i} - ADTT_j \qquad (3)$$

where $expCT_j^{r_{ptr_i}}$ is the expected completion time of the job *j* on the resource on which i^{th} partner job of *j* is scheduled. CT_{ptr_i} is the completion time of i^{th} partner job on that resource, $ADTT_{ptr_i}$being the communication time required to transfer the output file generated by the i^{th} partner job of *j* to the common child job and $ADTT_j$ is the communication time required to transfer the output file from the job *j* to the common child job [27].

If the condition is not satisfied, the next most critical partner job is considered. By doing so, the communication time will decrease for the most critical child job. However, this may increase the probability of scheduling more jobs on single resource, the resource on which the very first partner job was scheduled. This shall un-balance the load in Grid resources and some of the resource will get overloaded while others lightly loaded or not being used at all. To prevent this state, the resources selected after satisfying condition in (3) is checked whether its load is going beyond a predefined allowed variance from the load of other resources in the Grid. To check this, percent load on the selected resource $Load_r$ is calculated as in (2). Then standard deviation of the loads on resourcesSD_{Load_R} is determined, excluding the resource on which load is begin checked. Based on this data the resource is selected if the following condition is satisfied:

$$\left(Load_r - \frac{\sum_{i=0}^{m-r} Load_i}{m-1}\right) - SD_{Load_{R-r}} \le SD_{Load_{allowed}} \qquad (4)$$

Where $SD_{Load_{R-1}}$is the variation of loads on resources from the mean load of all the resources, excluding the load on the resource itself, calculated as standard deviation. $SD_{Load_{allowed}}$ is the allowed threshold value of standard deviation, provided by the algorithm as input parameter. The allowed threshold controls the variation in the loads in Grid system. Keeping the threshold value too small ensures that all the resources are strictly balanced and before assigning any extra load to a resource the rest of the resources also gets similar load. But this may lead to selecting a resource which is not very fast and may delay the completion time of the job. Keeping the threshold value too large relaxes the policy and allows the resource selector to make decision without considering the load balancing. Setting a threshold value helps us in tuning the overall functionality of the algorithm as per the preferences and goals (better resource utilization and throughput) of the Grid owner.

If the condition is not satisfied for the given resource, the next partner will be checked to see if it satisfies both of the conditions. If no partner job was scheduled or none of the partner jobs satisfy the conditions, then the resource that gives earliest completion time is selected. In this way, resource utilization is improved and also in most of the cases makespan of the workflow is also reduced. Makespan is the difference between submission time of the workflow and completion time of the last job of the workflow. The complete process is described in Algorithm 1.

Algorithm 1: Pseudo-Code for Balanced-PDCPG

$\textbf{function}$ SELECTRESOURCE(j)
$PTR_j \leftarrow list\ of\ partner\ jobs\ of\ j$
$\textbf{foreach}\ partner\ ptr_i\ in\ PTR_j\ do$
$r_{ptr_i} \leftarrow resource\ r\ on\ which\ partner\ ptr_i\ is\ scheduled$
$\textbf{if}\ IsNotOverLoaded(r_{ptr_i})\ then$
$expCT_j^{r_{ptr_i}} \leftarrow expected\ completion\ time\ of\ j\ on\ r_{ptr_i}$
$CT_{ptr_i} \leftarrow completion\ time\ of\ ptr_i\ on\ r_{ptr_i}$
$ADTT_{ptr_i} \leftarrow data\ transfer\ time\ of\ ptr_i$
$ADTT_j \leftarrow data\ transfer\ time\ of\ j\ when\ scheduled\ on\ r_{ptr_i}$
$\textbf{if}\ expCT_j^{r_{ptr_i}} \leq CT_{ptr_i} + ADTT_{ptr_i} - ADTT_j\ then$
$return\ r_{ptr_i}$ // selected this resource
$\textbf{end if}$
$\textbf{end if}$
$\textbf{end foreach}$
$\textbf{return}\ resource\ r\ from\ R\ which\ gives\ minimum\ expCT$
$\textbf{end function}$

$\textbf{function}$ ISNOTOVERLOADED(r)
$Load_r \leftarrow CHECKLOAD(r)$
$Load_{mean} \leftarrow \frac{\sum_{i=0}^{m-r} Load_i}{m-1}$
$Load_{allowed} \leftarrow allowed\ standard\ deviation$
$\textbf{if}\ (Load_r - Load_{mean}) - SD_{Load_{R-r}} \leq SD_{Load_{allowed}}\ then$
$return\ false$ // resources in not overloaded
$\textbf{else}$
$return\ true$ // resource is overloaded
$\textbf{end if}$
$\textbf{end function}$

V. Evaluation

The proposed algorithm is compared with PDCPG and DCPG. The algorithms are compared in different working environments of workflow and loads. The comparison is based on utilization of resources and makespan of the workflows. This section explains the experimental environment followed by the results.

A. Environment

The simulation is carried out by using Gridsim [28] simulator in advance reservation environment under OBS network architecture. The underlying machine configurations are; 2.7 GHz (8 CPUs) Core i7 with 8 GB primary memory running Windows 7 Professional64 bit Operating System. The resources are connected with 100 Mbps link. For every reservation request the requested number of CPUs is one and requested bandwidth is 10 Mbps. The users connected to resource 0 submit a workflow with the objective of earliest completion time and the Grid trying to complete the job while utilizing all of the resources equally. Multiple users are connected to the Grid submitting workflows at the same time in order to check the load balancing in the Grid on the resources. At start, the utilization profiles are empty and there are no reservations in them, to check how each of the algorithms makes reservations in the utilization profile. As we are mainly interested to balance the resource load and increase throughput, we have provided enough bandwidth at each of the links.

B. Workflow

To compare the algorithms in different situations there are various workflows selected to be submitted to the system as application. E-protien has minimum number of 15 jobs, and is selected from real world applications [29]. Job60PSPLIB, Job90PSPLIB, and Job120PSPLIB having 60, 90, and 120 jobs, respectively, are taken from Project Scheduling Problem Library (PSPLIB) [30]. LIGO, Montage, SIPHT, and Cybershake [31] have 83, 25, 30, and 21 jobs respectively. For each of these workflows, jobs are created with different granularity settings. For every setting simulation is run 100 times and mean values are taken to remove the errors in individual simulation execution. During every execution, job lengths are randomly assigned ranging from 20,000 to 50,000 MI, based on the number of the jobs in a workflow. The output file size is given randomly according to the granularity. For low granularity the output file size ranges from 420 to 450 MB and for higher granularity it ranges from 80 to 120 MB. The number of users submitting jobs in each 100 rounds is kept the same and ranges from 5 to 45 users based on the number of jobs in the workflow. This is to keep the load in the reservation profile of the resources always from 60% to 95%. For small workflows like e-protein we have submitted jobs by 45 users to populate the utilization profiles with reservations in them and check how the load is balanced among all the resources of the Grid.

VI. Results

Our main objective is to observe the effectiveness of our load balancing policy and see the effects of this balancing technique on makespan of different workflows. To understand this, we have experimented our proposed technique on low granularity workflow followed by higher granularity workflows. We have studied the improvement in resource utilization and have also studied the effects of this on makespan of individual workflows, as listed above. As we are trying to balance the load, the balancing policy will disparate the idea of selecting a resource on which a partner job has already been scheduled. Apparently, this may increase the communication cost and increase the makespan of a workflow. In contrast, the results show that makespan has also been decreased in many cases by using our proposed balancing technique, because of better resource utilization and future optimal approach. The results contain the average of 100 simulations for each of the settings. For low granularity we have taken mean of 0.25, 0.5, and 0.75 granularities, calculated as in (1). For each of the granularities we have simulated 100 times and then taken mean of them. This means that we have simulated 300 times for low granularity for each of the algorithms, and similarly for high granularity as well. For understanding the effects of our proposed balancing technique, we have embedded our load balancing policy in DCPG and PDCPG named Balanced-DCPG (B-DCPG) and Balanced-PDCPG (B-PDCPG), respectively.

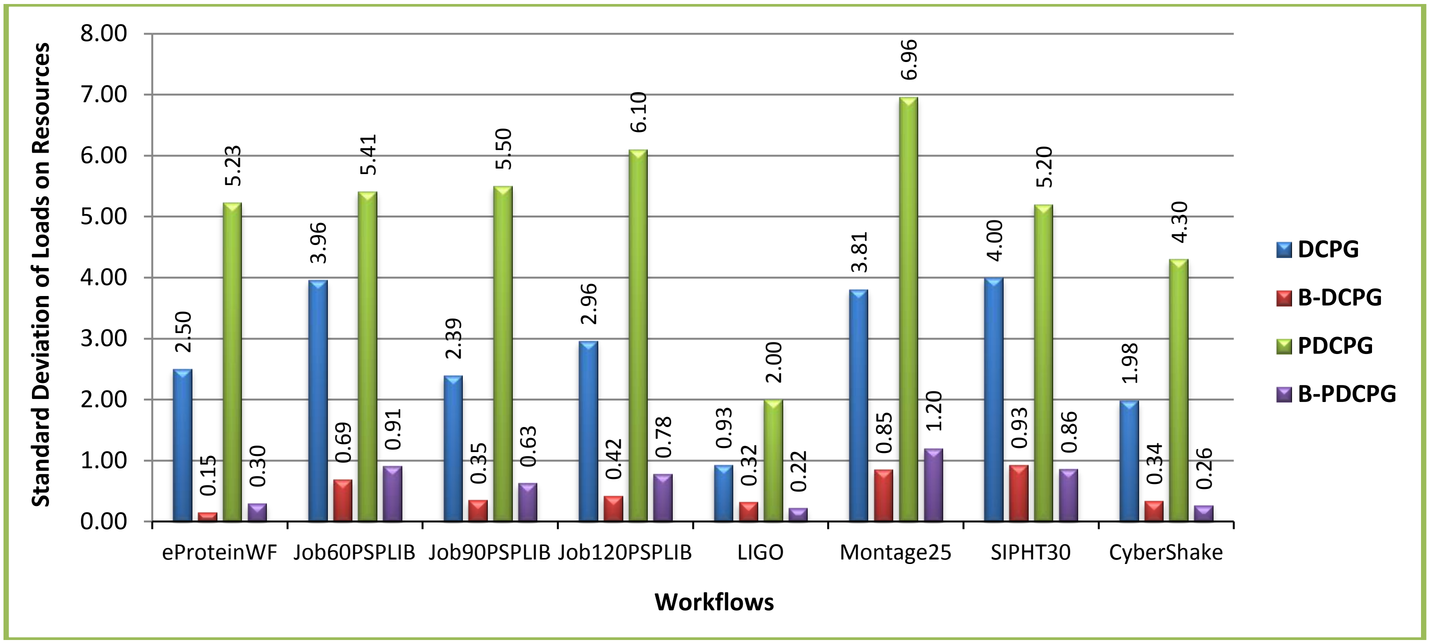


Fig. 2. Standard Deviation of Loads on Resources

Fig 2, shows a comparison of how different algorithms have distributed load on resources of the Grid. The values shown are the standard deviations of the loads. We have compared the variation of loads on resources of the Grid in DCPG to that of B-DCPG and similarly PDCPG to that of B-PDCPG. It can be observed that the load on resources has been distributed equally with negligible difference in balanced version of DCPG and PDCPG. The results are very similar in low and high granularities; therefore, we have combined them in a single chart. It can be seen that PDCPG has a very large variation in load distribution, which shows that the load is not equally distributed among the resources of the Grid system. In other words, PDCPG does not guarantee a good utilization of resources.

The balancing technique has achieved significant improvement in utilization of resources. As discussed above, the allowed variance can be controlled by setting the allowed threshold in balancing mechanism. Here we set a very small value of allowed threshold value, to check how much we can balance the loads on Grid resources.

To check the effects of load balancing on execution time required for a single application, we have also measured the effects of this load balancing technique on individual makespans of the user applications.

TABLE I. AVERAGE MAKESPAN OF THE WORKFLOWS IN LOW GRANULARITY

WF	MAKESPANS				PERCENT IMPROVEMENT			
	DCPG	*B-DCPG*	*PDCPG*	*B-PCPG*	*B-DCPG over DCPG*	*B-PDCPG over PDCPG*	*PDCPG over DCPG*	*B-PCDPG over DCPG*
eProteinWF	631.67	630.67	622.67	623.00	0.16	-0.05	1.45	1.39
Job60PSPLIB	976.67	962.00	957.33	955.33	1.52	0.21	2.02	2.23
Job90PSPLIB	1364.00	1363.33	1343.00	1345.33	0.05	-0.17	1.56	1.39
Job120PSPLIB	1114.00	1114.00	1094.67	1093.67	0	0.09	1.77	1.86
LIGO	1175.33	1177.33	1176.00	1172.67	-0.17	0.28	-0.06	0.23
Montage25	480.57	480.30	470.87	470.33	0.06	0.11	2.06	2.18
SIPHT30	424.33	422.33	420.67	419.00	0.47	0.4	0.87	1.27
CyberShake	373.67	374.67	372.67	370.67	-0.27	0.54	0.27	0.81

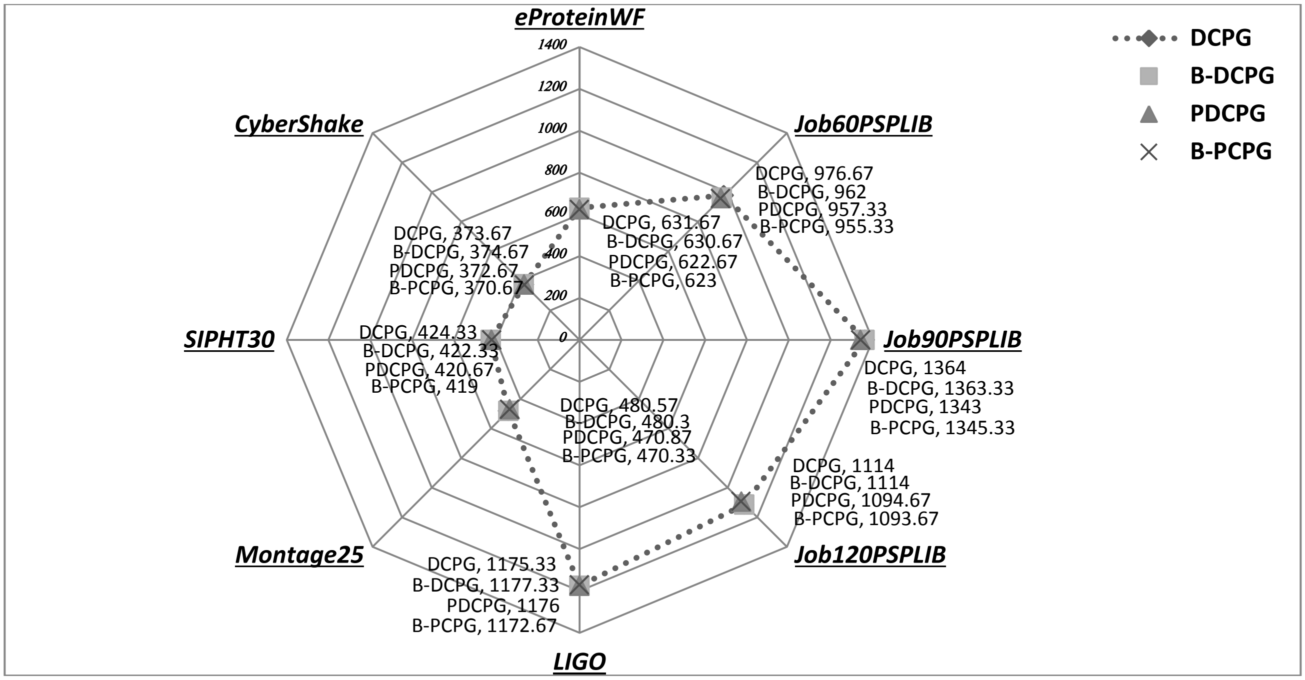


Fig. 3. Average Makespan of the Workflows in Low Granularity

Table 1 and Fig. 3, explains the effects of load balancing on makespans of workflows in low granularity jobs. It can be seen that the overall execution time required to complete an application is unchanged in most of the cases and has even been reduced in some cases, especially in the workflows which have more breadth rather than depth. Although the completion time of individual jobs may get delayed due to increased communication cost in our proposed balancing technique, it has no effect on the overall execution time of a workflow. Table 1 and Fig. 3, also shows the percent improvements in the makespans of the workflows. We can see that there are some cases in which our balanced approached has not achieved any improvement but that is not our goal at all. Our objective is not to achieve improvements in makespan reduction, but to show that our balancing technique has no effect on individual makespan of the workflows.

TABLE II. AVERAGE MAKESPAN OF THE WORKFLOWS IN HIGH GRANULARITY

WF	MAKESPAN				PERCENT IMPROVEMENT			
	DCPG	*B-DCPG*	*PDCPG*	*B-PCPG*	*B-DCPG over DCPG*	*B-PDCPG over PDCPG*	*PDCPG over DCPG*	*B-PDCPG over DCPG*
eProteinWF	637.67	637.67	636.67	636.67	0	0	0.16	0.16
Job60PSPLIB	1324.33	1324.33	1322.00	1322.00	0	-0.05	0.18	0.13
Job90PSPLIB	1340.67	1337.67	1333.33	1333.33	0.22	-0.4	0.55	0.15
Job120PSPLIB	1575.00	1575.00	1566.67	1566.67	0	0.11	0.53	0.64
LIGO	1731.33	1731.00	1735.67	1735.67	0.02	0.08	-0.25	-0.17
Montage25	626.10	626.47	622.00	622.00	-0.06	0.02	0.66	0.68
SIPHT30	556.67	553.67	556.33	556.33	0.54	0.48	0.06	0.54
CyberShake	485.00	484.67	484.33	484.33	0.07	0.07	0.14	0.21

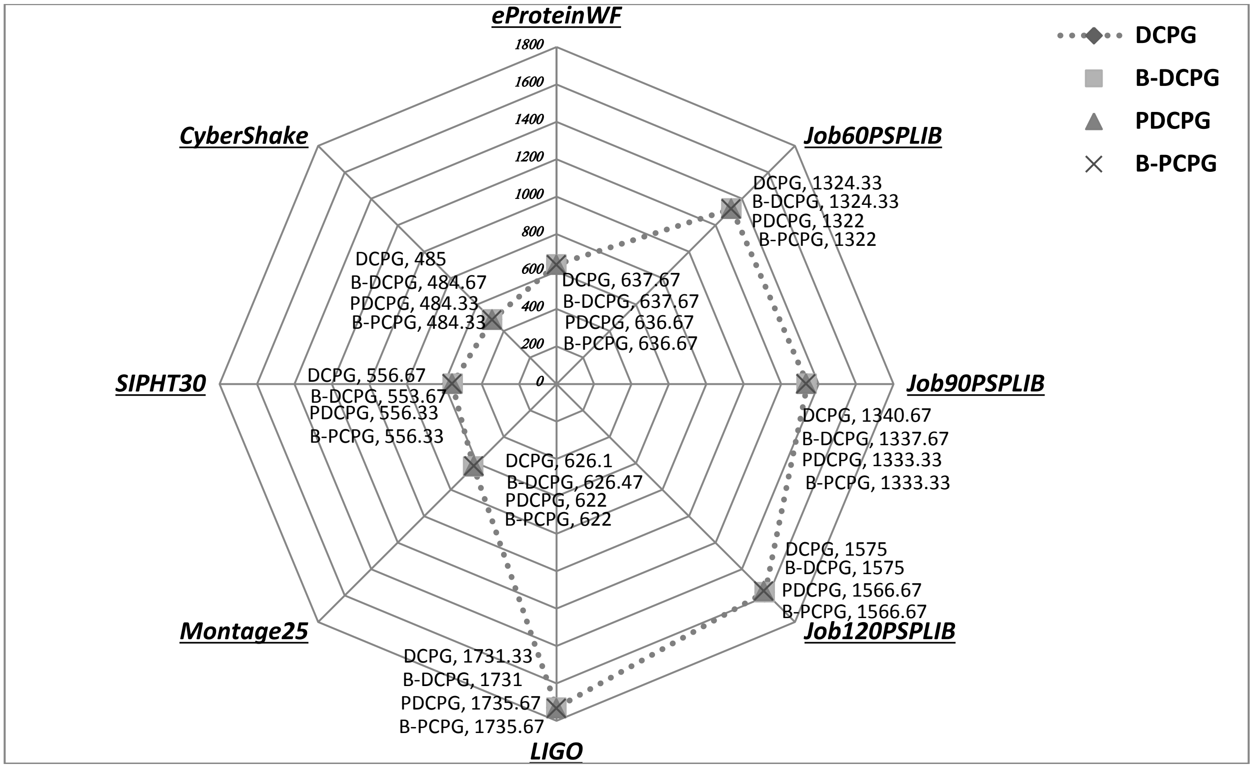


Fig. 4. Average Makespan of the Workflows in High Granularity

Table 2 and Fig. 4 below, explains the effects of load balancing on makespans of workflows in high granularity jobs. For high granularity, we have taken mean of simulations with 1.0, 1.25, 1.50, and 1.75 granularities and combined them together. The results show that, like PDCPG, B-PDCPG is also providing very small percentage improvement over DCPG. The reason is that there is not much communication involved in high granularity jobs.

It has been observed that the standard deviation of the makespans also reduces by balancing the loads on resources. What this means, is that PDCPG completes some of the workflows in a very short time by scheduling most of its jobs (especially partner jobs) on the fastest resource, while others are scheduled on relatively slower resources resulting in large completion time. By balancing the workload on resources the makespan is not that much different from each other, as the balancing policy stops a workflow from reserving all of the slots on a high speed resource.

VII. Conclusion and Future Work

In this study, we have addressed the problem of load balancing in advance reservation environment. We have proposed a policy based dynamic and centralized load balancing technique for balancing the loads on different resources of the Grid. We have observed that by trying to reduce communication, PDCPG leads the overall system to an un-balanced state where some of the resources are highly reserved in advance while the rest are rarely used or idle. This can result in resource under-utilization and increased rejection rate at some time. To solve this problem we have presented a technique based on a policy restriction while assigning job of a workflow to a resource. Simulation results shows that the proposed technique enhances the performance of PDCPG and increase utilization of resources and jobs throughput.

Future work will consider a distributed and adaptive technique that will implement the load balancing policy at different level of the Grids in a hierarchy while adapting itself to the changes. The implementation of proposed technique in non-advance reservation environment will also be considered.

Acknowledgement

The authors would like to thank Sur University College (Sur, Sultanate of Oman) for support and sponsorship of this research.

References

[1] B. Yagoubi and Y. Slimani, "Task load balancing strategy for grid computing," Journal of Computer Science, vol. 3, no. 3, pp. 186-194, 2007.

[2] M. Siddiqui, A. Villazon and T. Fahringer, "Grid capacity planning with negotiation-based advance reservation for optimized QoS," in Proceedings of the 2006 ACM/IEEE conference on Supercomputing, 2006.

[3] S. F. El-Zoghdy, "A load balancing policy for heterogeneous computational grids," International Journal of Advanced Computer Science and Applications, vol. 2, no. 5, 2011.

[4] D. S. Gawande, R. C. Dharmik and C. Panse, "A load balancing in grid environment," International Journal of Engineering Research and Applications (IJERA), vol. 2, no. 2, pp. 445-450, 2012.

[5] V. Curcin and M. Ghanem, "Scientific workflow systems - can one size fit all?," in Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International, 2008.

[6] J. Ashraf and T. Erlebach, "A new resource mapping technique for grid workflows in advance reservation environments," in High Performance Computing and Simulation (HPCS), 2010 International Conference on, 2010.

[7] M. Rahman, S. Venugopal and R. Buyya, "A dynamic critical path algorithm for scheduling scientific workflow applications on global grids," in e-Science and Grid Computing, IEEE International Conference on, 2007.

[8] B. Yagoubi and Y. Slimani, "Load balancing strategy in grid environment," Journal of Information Technology and Applications, vol. 1, no. 4, pp. 285-296, 2007.

[9] L. Mohammad Khanli, "A new hybrid load balancing algorithm in grid computing systems," International Journal of Computer Science & Emerging Technologies, vol. 2, no. 5, 2011.

[10] L. M. Khanli, S. Razzaghzadeh and S. V. Zargari, "A new step toward load balancing based on competency rank and transitional phases in grid networks," Future Generation Computer Systems, vol. 28, no. 4, pp. 682-688, April 2012.

[11] Deepika, D. Wadhwa and N. Kumar, "Performance analysis of load balancing algorithms in distributed system," Advance in Electronic and Electric Engineering, vol. 4, no. 1, pp. 59-66, November 2014.

[12] B. Yagoubi and M. Meddeber, "Towards hybrid dependent tasks assignment for grid computing," Java in Academia and Research, iConcept Press.

[13] P. Bindu, R. Venkatesan and K. Ramalakshmi, "Perspective study on resource level load balancing in grid computing environments," in Electronics Computer Technology (ICECT), 2011 3rd International Conference on, 2011.

[14] J. C. Patni, M. S. Aswal, O. P. Pal and A. Gupta, "Load balancing strategies for grid computing," in Electronics Computer Technology (ICECT), 2011 3rd International Conference on, 2011.

[15] S. Penmatsa and A. T. Chronopoulos, "Cooperative load balancing for a network of heterogeneous computers," in Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International, 2006.

[16] J. Balasangameshwara and N. Raju, "A hybrid policy for fault tolerant load balancing in grid computing environments," Journal of Network and Computer Applications, vol. 35, no. 1, pp. 412-422, 2012.

[17] Y. Li, Y. Yang, M. Ma and L. Zhou, "A hybrid load balancing strategy of sequential tasks for grid computing environments," Future Generation Computer Systems, vol. 25, no. 8, pp. 819-828, 2009.

[18] K.-Q. Yan, S.-C. Wang, C.-P. Chang and J. Lin, "A hybrid load balancing policy underlying grid computing environment," Computer Standards & Interfaces, vol. 29, no. 2, pp. 161-173, 2007.

[19] S. Zikos and H. D. Karatza, "Communication cost effective scheduling policies of nonclairvoyant jobs with load balancing in a grid," Journal of Systems and Software, vol. 82, no. 12, pp. 2103-2116, 2009.

[20] Y.-K. Kwok and L.-S. Cheung, "A new fuzzy-decision based load balancing system for distributed object computing," Journal of Parallel and Distributed Computing, vol. 64, no. 2, pp. 238-253, 2004.

[21] V. Di Martino and M. Mililotti, "Sub optimal scheduling in a grid using genetic algorithms," Parallel computing, vol. 30, no. 5, pp. 553-565, 2004.

[22] Z. Wenpeng and L. Hongzhao, "A load balancing method based on genetic clonal annealing strategy in grid environments," in Educational and Network Technology (ICENT), 2010 International Conference on, 2010.

[23] J. Cao, D. P. Spooner, S. A. Jarvis and G. R. Nudd, "Grid load balancing using intelligent agents," Future generation computer systems, vol. 21, no. 1, pp. 135-149, 2005.

[24] J. U. In, S. Lee, S. Rho and J. H. Park, "Policy-based scheduling and resource allocation for multimedia communication on grid computing environment," Systems Journal, IEEE, vol. 5, no. 4, pp. 451-459, 2011.

[25] E. M. Varvarigos, V. Sourlas and K. Christodoulopoulos, "Routing and scheduling connections in networks that support advance reservations," Computer Networks, vol. 52, no. 15, pp. 2988-3006, 2008.

[26] K. Christodoulopoulos, N. Doulamis and E. Varvarigos, "Joint communication and computation task scheduling in grids," in Cluster Computing and the Grid, 2008. CCGRID'08. 8th IEEE International Symposium on, 2008.

[27] J. Ashraf, "Partner-based scheduling and routing for grid workflows," University of Leicester, 2012.

[28] A. Sulistio, C. S. Yeo and R. Buyya, "Visual modeler for grid modeling and simulation (gridsim) toolkit," in ICCS 03 Proceedings of the 2003 International Conference on Computational, Berlin, 2003.

[29] A. O'Brien, S. Newhouse and J. Darlington, "Mapping of scientific workflow within the e-protein project to distributed resources," in UK e-Science All Hands Meeting, 2004.

[30] F. Wittemann, "The Library PSBLIB," December 2015. [Online]. Available: http://www.om-db.wi.tum.de/psplib/library.html. [Accessed 5 January 2015].

[31] January 2015. [Online]. Available: https://confluence.pegasus.isi.edu/display/pegasus/WorkflowGenerator. [Accessed 3 January 2015].

TMCC: An Optimal Mechanism for Congestion Control in Wireless Sensor Networks

Razieh Golgiri
Department of computer engineering
Science and Research Branch, Islamic Azad University
Khuzestan, Iran

Reza Javidan
Department of Computer Engineering and IT
Shiraz University of Technology
Shiraz, Iran

***Abstract*—Most proposed methods for congestion control of Wireless Sensor Networks (WSNs) have disadvantages such as central congestion control mechanism through the sink node, using only one traffic control or resource control mechanism and also having the same throughput on all nodes. For the purpose of addressing these problems, in this paper, a new congestion control protocol is presented in order to increase network lifetime and reliability of WSNs. Since the priority of generated traffic in network level is not uniform in WSNs, an architecture framework is proposed based on priority of generated traffics for service identification in network level in order to meet better service quality and efficiency. The proposed method called TMCC has been compared with Traffic-Aware Dynamic Routing (TADR) method to show the effectiveness of the proposed method in terms of end to end delay, throughput, power consumption and lifetime of network.**

Keywords—Wireless Sensor Network (WSN); traffic management; resource control; alternative path; QoS; TMCC

I. Introduction

Wireless Sensor Networks (WSNs) consist of a number of sensor nodes which are widely distributed in environment [1]. They can be operated in air, underwater, or ground [2]-[3]. Recent advances in electronics and wireless communication have led to design and build sensors with low power consumption, small size and low price. These small sensors can perform actions such as receiving environmental information based on sensor type, processing them and sending data through other sensor nodes [4]. With the capabilities of pervasive surveillance, sensor networks have attracted significant attention in many applications domains, such as habitat monitoring [5], [6], object tracking [7], [8], environment monitoring [9]–[11], military [12, 13], disaster management [14], as well as smart environment [15]. Well-designed congestion control techniques allow efficient transmission of significant volumes of data from a large number of nodes along one or more routes towards the data processing centers (usually known as 'sink') [16]. With the increasing on the application of WSNs [17], different traffic levels generated by these sensors require providing different quality of services. In such condition, congestion occurs when data capacity transmitted through the network is above the network packet handling capacity [18]. Congestion occurs when the traffic load exceeds the available capacity on node level (buffer overflow) or link level (interference or contention) [19].

WSNs are divided into two categories of event-based and data stream- based in terms of collecting data and sending data [20]. In both methods, data stream is formed from source nodes to base station that can lead to the creation of congestion in network. In fact, mismatch of sent and received data rate leads to congestion formation in network. Therefore, reliability of sending packets and network throughput will be reduced. Lost packets due to congestion should be resent resulting waste of energy and reducing overall network lifetime [21]. Thus, congestion has direct effect on energy efficiency and quality of service. Moreover, congestion in WSNs interacts with node energy limitation, low buffer capacity of nodes, sensitivity to delay and topology changing of sensors. Network congestion must be controlled in order to increase the reliability and longevity of network. To this end, the proposed protocol must be capable of controlling congestion step by step and end to end by rate setting in order to prevent packet loss and reducing traffic at each node and increasing network reliability. Most of the proposed protocols for congestion control of WSNs have disadvantages like having central congestion control mechanism (rate setting by sink), using only one of the mechanisms for traffic control or resource control as well as having the same throughput on all nodes [22].

Since the priority of generated traffic in network level is not uniform in WSNs, an architecture framework is needed based on priority of generated traffics for service identification in network level in order to meet better service quality and efficiency. According to above materials, paying attention to different priorities for different levels of network traffic is essential which is discussed in this paper.

This remainder of this paper is organized as follows: in Section 2 recent works on congestion control protocols in WSNs are reviewed. In Section 3, the proposed method explained. Simulation results and performance evaluation are explained in Section 4. Finally in Section 5, conclusion is outlined.

II. Related Works

Recently many researched are performed for congestion control in WSNs [23]-[25]; among them, some of related works are reviewed.

Wan et al. [26] proposed a complete research for detecting and preventing congestion in WSNs in which congestion is detected by sampling wireless areas and supervising queue occupation. When congestion is detected, an upward inverse message is propagated and the upward nodes reduce the traffic volume to moderate congestion. In addition, closed circuit resource adjustment is used in which end-to-end constant long-term feedback from the base station to source nodes requires adjusting the transmission rate by using additive gain and multiplicative reduction. Although, this method supports congestion reduction, it does not insure balance among resources.

Heikalabad et al. proposed a dynamic predictive congestion control which prevents congestion using precise rate control based on dynamic priority [27]. Some applications of multimedia wireless sensor networks may need urgent traffic sending to base station. The essential traffic requires low latency and high reliability so that immediate escape and defensive measures can be used when needed. The proposed approach in this protocol consists of three different types of traffic; immediate type that is important real-time traffic, fast type that is real-time traffic that is not so important and normal traffic. Any traffic with different priority is initialized, that may change in different phases of congestion. This strategy can prevent effectively the problem with extreme demands of quality and, service performance maintenance, when congestion is high in multimedia WSNs.

Zhao et al. [28] proposed an alternative path-based congestion control mechanism that uses free resources in order to reduce the congestion of multi networks. Congestion of nodes is classified into two groups and two distributed algorithms are offered that can find an alternate route through a neighborhood table. Alternative path-based congestion control uses existing paths, instead of creating new paths due to the limited energy of sensor nodes. In addition, according to tree topology of sensor networks, one distributed algorithm is used in order to select the appropriate way through preventing new congestion simultaneously. The protocol focuses on WSN congestion that occurs by sudden data. The main focus of this protocol is on reducing traffic rate to avoid congestion.

In traffic-aware dynamic routing to alleviate congestion protocol of TADR, dynamic capacity timing is used in order to reduce congestion and maintain reliability through an aware of traffic dynamic routing protocol [29]. The main intention of this protocol is that if two nodes send their packets in the shortest path and the parent node is congested, source nodes must use a suitable alternative path consisting of passive nodes or low load once. Based on this method, a hybrid scalar potential field is defined which contains a depth field and a queue length field. The depth field provides the basic routing backbone which routes the packets directly to the sink along the shortest path. The queue length field makes TADR traffic aware. When the congestion appears, the excessive packets are dynamically rerouted to multiple paths consisting of idle or under-loaded nodes.

In a separate and fair congestion control a distributed and adaptive mechanism is provided for congestion control in sensor networks that seeks to find an optimal transmission rate for nodes [30]. In this scenario, two separate modules are used for network control and fairness among flows. Each node controls the overall rates of input and output. Firstly, one node calculates the increase (if the output rate is higher) or decrease (if the input rate is greater) of needed traffic. Then, the fairness control module is applied on this assembly and divides it to independent streams in order to achieve the desired fairness. Control information is carried in data packets.

Sergiou et al. in [31] proposed a congestion control algorithm that creates alternative paths from sources to sinks to prevent congestion. Alternative path algorithm is based on a hierarchical tree that changes routs based on local information such as congestion of neighbors. In fact, a "source control" algorithm tries to reduce congestion of WSNs through simple steps to create dynamic alternative paths to sink.

Table I shows brief comparisons between different protocols including the proposed method. As mentioned so far, almost all of the protocols for congestion control in WSNs have disadvantages like having central congestion control mechanism, using only one of the mechanisms for traffic control or resource control as well as having the same throughput on all nodes.

TABLE I. COMPARISONS BETWEEN DISCUSSED METHODS

Protocol	**Congestion Detection Phase**	**Congestion Declaration Phase**	**Congestion Control Phase**
CODA [26]	buffer occupancy + channel load	explicit	rate adjustment (AIDM)
PCC [27]	buffer occupancy	implicit	rate adjustment (packet drop)
APCC	buffer occupancy	implicit	alternative path
TADR	queue depth and length	implicit	alternative path
DCCF	feedback delay	implicit	rate adjustment
HTAP	buffer occupancy	implicit	alternative path
TMCC	buffer occupancy, Packet reaplication, reciving rate > sending rate	explicit	rate adjustment, alternative path

III. THE PROPOSED METHOD

The proposed method in this paper called TMCC, tries to implement an acceptable level of reliability factor in network during the congestion avoidance and congestion control of sensor networks. The main motivation of the proposed method is that the mechanism uses both congestion control methods including source control and traffic management. The purpose of the proposed method is to obtain a congestion control mechanism that applies the highest reliability in sensor networks.

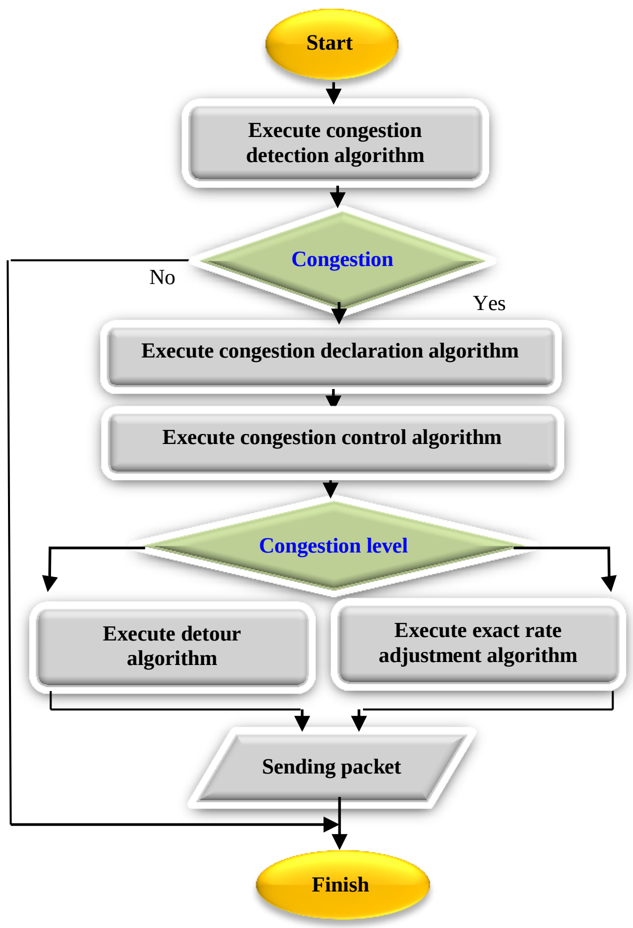


Fig. 1. The proposed congestion control plan

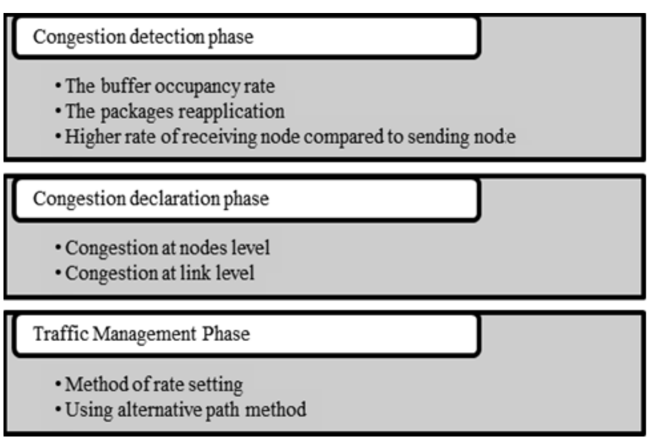


Fig. 2. The main stages of congestion control in the proposed method. The Proposed Method

Fig. 1 shows the proposed congestion control plan. If two source nodes send their packets in the shortest route and the receiving node is congested, the receiving node after detecting congestion and notifying the source nodes, will begin to implement an congestion control algorithm. Under normal circumstances, if a packet is received by node but receiving node cannot send packet to the sink node, it must prioritize the data and place it in its buffer based on some preferences. However, if congestion conditions prevail, an appropriate congestion control algorithm must be implemented while each of data will be manipulated in node buffer based on their priorities. According to output phase, the congestion detection algorithm will send data to the sink or base station node by precise rate setting or by creating a suitable alternative path with low load or having passive neighbor nodes. Therefore high reliability and more energy saving scheme is achieved.

In the proposed method, we need to design congestion detection mechanism, congestion declaration mechanism and congestion control mechanisms in order to provide a general method of congestion control. The proposed congestion control method as shown in Fig. 2, includes three main phases of congestion detection, congestion declaration and traffic management that traffic management phase itself is divided into two traffic control (rate setting) and resource control (alternative path). The details of the proposed method are explained in the following subsections.

A. Congestion Detection Phase

The main task of this phase is to detect congestion type. Two congestion types are distinguished: node level congestion and link level congestion. Node level congestion can be detected through every node buffer overflow while link level congestion is created because of multiple accesses to a common media or link between two nodes or node shutdown because of ending energy.

In the proposed method, such information will be used: crowded queue of a receiving node, demanding rate for packet retransmission in buffer, higher receiving rate to sending rate and etc. Node congestion in the buffer can be performed through implementing the algorithm on receiving packet. That is, if the buffer at each node reaches the congestion threshold right after receiving a new packet, it will begin to implement a congestion detection algorithm after calculating the absolute priority of packet and putting it in queue. Buffer threshold is usually set 0.9 to queue length. Necessary conditions for occurrence or non-occurrence of congestion can be seen in flowchart of Fig. 3.

```
If Buffer_used >=Buffer_Treshold  OR  Double_Req = true OR
    Rate_in >  Rate_out  Then
                Congestion_Detection = True
                Call Congestion_Type Procedure (Buffer_used,
                         Double_Req, Rate_in, Rate_in)
Else
        Congestion_Happen = False
End if
```

Fig. 3. Boolean congestion detection procedure of the proposed method

The congestion detection algorithm will be implemented according to the above parameters after occurrence of one or more conditions mentioned so far in each node and output of this algorithm will determine reasonably the presence or absence of congestion in the vicinity of node. In fact, in congestion detection phase, in addition to the detection of congestion, the type of congestion must also be determined. Finally congestion opposite method of the third phase will be selected according to conditions of these three criteria: queue occupancy, the ratio of buffer packets that are reapplied to total number of packets in buffer, and the rate difference between receiving node and sending node that can be rate setting or path alternative method.

B. Congestion Declaration Phase

Usually congestion declaration can be performed in two implicit and explicit ways. In the proposed method, since nodes must know the type of mechanism that is used to relieve congestion and they should be aware of congestion occurrence in their neighbors, the explicit congestion declaration is used. Therefore, they can choose an appropriate congestion control mechanism.

C. Congestion Control Phase

In the proposed method, instead of using just a single scheme, the main objective is to use both congestion control mechanisms: "traffic management" and "resource control". According to congestion type that has been occurred, one of these two methods will be adopted. Based on explicit congestion method, necessary information will be sent to each neighboring nodes. Afterward, resource control, path alternative or a new alternative path creation can be used for link type congestion control. All these are explained in the following subsections.

❖ *Traffic Management*

In general, the incoming traffic at a node is in two main types: source traffic sensed by sensor and transit traffic which is the traffic transmitted from child nodes. Assume that these traffic rates denoted by R_s and R_t. As a result, the total amount of traffic at each node called R_{in} is defined as $R_{in} = R_t + R_s$. R_{in} is the total rate of all data flow in node i. Driving rate of node at time t is considered $A_{r,t}$ calculated according to Eq. (1).

$$A_{r,t} = Max_{rate} - R_{in,t} \tag{1}$$

Where $R_{in,t}$ shows total rate of all data flow in node i at time t. For node i, it can be concluded that the total rate of child nodes must be less that $A_{r,t}$ in order to keep away node i from congestion.

If the rate of packet transfer R_{trans} is greater than the rate of incoming traffic R_{in}, output rate will be equal to the rate of incoming traffic $R_{out} = R_{in}$. Otherwise, the packet transmission rate reaches to the output rate of packet that is R_{out} is close to R_{trans}. In the proposed method, Eq. (2) is used to calculate the rate of packet output.

$$R_{out} = Min\,(R_{trans}, R_{in}) \tag{2}$$

This suggests that R_{out} can be partially reduced through reducing R_{in} and increase through increasing it. It can be effective in achieving the main objective of this research that is congestion control be reliable.

❖ *Control of Resources*

With increasing the number of nodes in a sensor network, "resource control" method can enhance the performance in comparison with "traffic management" methods. It can deliver a lot of packets to sinks through using different paths to distribute the sudden traffic in order to reduce congestion. In the proposed approach this feature of resource control method is used for solving the problem of link level congestion. When the desired node receives the control packet of congestion declaration, it will find the congestion control method and will implement the resource control mechanism.

❖ *Alternative path Method*

If the congestion is the result of incompetency of a node, the damaged node regulates its node available transmission rate $A_r(i)$ to null value and informs it to the other adjacent nodes. The child nodes update their adjacency table by regulating $A_r(i)$ and start to find an alternative. When a node damaged, all of its child nodes start to find other ways to forward the packets. The suggested method is presented using pseudo code in Fig. 4.

1. Sort neighbor table (except the father node) according to ascending of H
2. For each node k
   If $H_i > H_k$ **AND** $N_k <> N_i$ **AND** $A_r(k) > 0$ **Then** send request
   Else if $A_r(k)>0$ **Then** send request for node with minimum H_k
3. Receive response, start forwarding packets using new path

Fig. 4. Alternative path search pseudo code of the proposed method

Since sink node is the root of routing tree, the nodes with the lowest height is closer to the sink node, therefore, they spend less cost for transportation. For nodes on the same level, algorithm must guarantee that the node is not a child node. When node i found a path alternative candidate for node k, requested packet will be sent to the node k by node i. when node k replied to request, another alternative path will be established and node i can forward data to new path. If node i does not receive an answer, it will research other neighboring nodes.

If a node fails to find the candidate node in step 3, it will send a path alternative request to node with minimum H_k and nonzero Ar. Request includes request node ID, faulty node ID, and congestion height and the node that receives the request will find an alternative path in neighborhood table. In case of ring request, when a node starts to find the requested path, firstly, it must check the flag, if the flag is not set, the algorithm will be implemented. It is easy to find an alternative path to send packets from congestion due to the large number of idle nodes. Also, if more than one neighboring node is found for changing the path with equal conditions, the other two criteria can be used, that is, the remaining energy of node and occupancy rate of the node buffer to select one from several candidate nodes. Therefore, traffic management process is performed after declaring congestion and traffic can be controlled through replacing other path.

IV. Evaluation and Experimental Results

In order to demonstrate the capability of the proposed method, it has been compared with Traffic-Aware Dynamic Routing (TADR) method [16] that is a well-known method to reduce traffic in WSNs. The evaluation is performed using NS-2 simulator. Important routing parameters such as energy consumption, end to end delay, network lifetime and throughput are measured to measure the performance.

Simulations with different parameters such as simulation time are performed and the results of evaluations are outlined.

In this paper, several scenarios with different number of network nodes equal to 50, 100 and 250 nodes are carried out. The dimension of simulation environment for these scenarios is equal to 700 m×700 m. Radio propagation range for each node is considered 250 and its MAC layer protocol is selected as IEEE 802.11. In addition, there are two traffic flows simulates on the network that send fixed rate of packets to the network. Simulations have been carried out at different times: at 100, 250 and 500 seconds. Buffer size is selected as 150 packets. Finally, position of nodes is considered to be random. All the simulations are performed once using the proposed method of TMCC and again on the method of TADR to compare the results as demonstrated in the following subsections.

A. End-to-end delay

End to end delay of wireless sensor networks refers to the time when information packets are transferred across the network from source node to destination node. Fig. 5 shows a diagram where end to end delay for TMCC is significantly lower compared to TADR. It shows that our proposed method, TMCC has better performance at different times and this argument reflects the better performance of the proposed method.

B. Throughput

The measure is achieved through dividing received amount of data received at destination by data delivery time. In throughput, criteria such as packet delivery rate and end to end delay are also involved. The better these criteria, the network throughput will be higher. Throughput is expressed as:

$$X = C / T \qquad (3)$$

In Eq. (3) X represents the throughput and C shows the number of requests that have been completed. T represents the entire time. Fig. 6 shows the results.

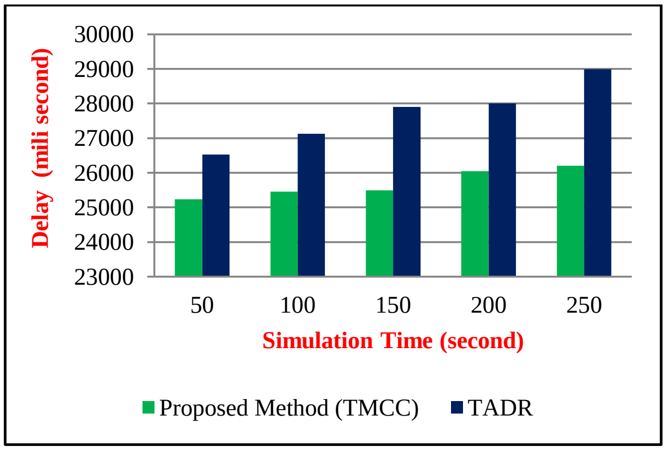


Fig. 5. The delay time of 250 seconds and 50 nodes

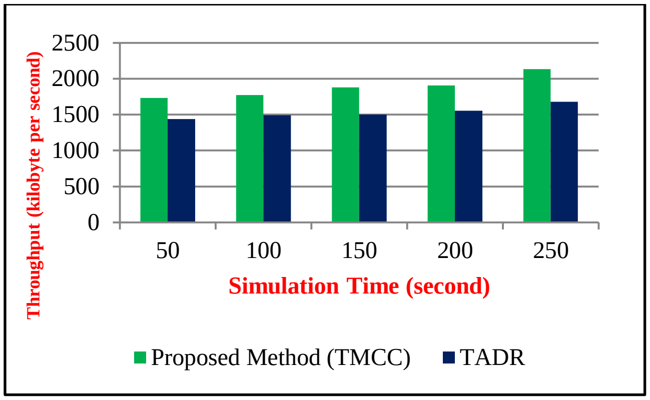


Fig. 6. The throughput with 250 seconds time and 50 nodes

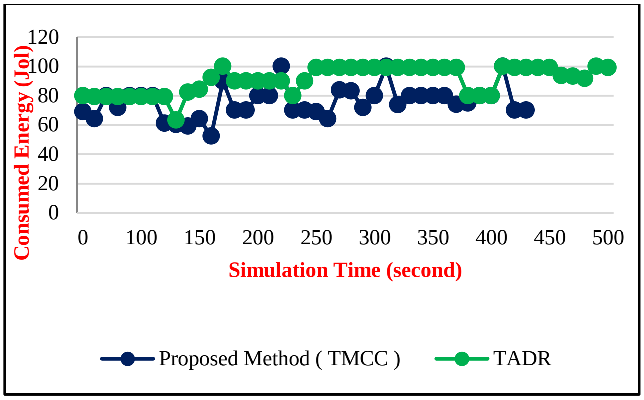


Fig. 7. Energy consumption by changing simulation time

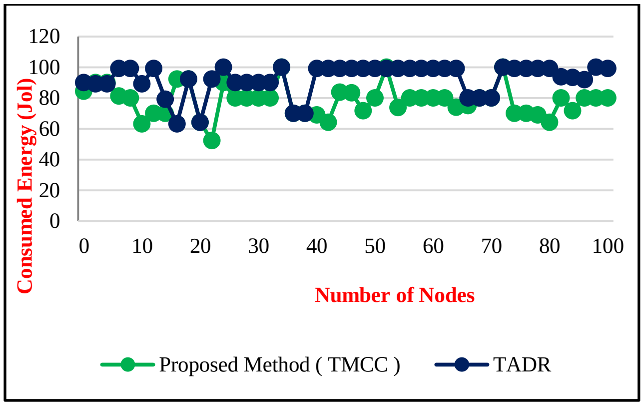


Fig. 8. Energy consumption by varying the number of nodes

C. Consumed energy

This measure represents the amount of energy that is consumed in routing. Fig. 7 shows the results for energy consumption for different simulation time while Fig. 8 shows energy consumption by varying of the number of nodes in the network. As shown in Fig. 7 and 8, consumed energy in the proposed method of TMCC is lower compared to the TADR method. This is due to the rapid selection of alternatives which makes traffic not be on a specific route and total energy consumption on the network will be reduced through quick notification of congestion and selecting alternative appropriate routes. So in this case, the proposed method has better performance.

D. Network Lifetime

Network overall lifetime is the time when the network is stable and can send data packets and beyond this time the network is not under operation. As it can be seen in Fig. 9, the proposed TMCC method has acted better than TADR in terms of network lifetime. This shows the superiority of the proposed method both on energy consumption and increasing lifetime of the network.

A brief summarization of advantages and disadvantages or limitations regarding the proposed method is shown in Table II.

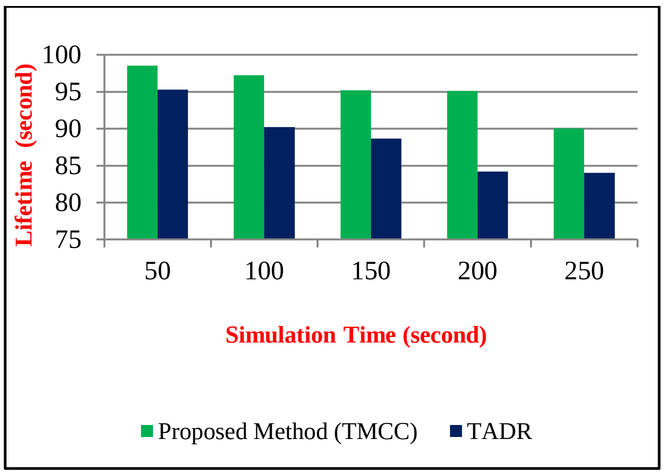


Fig. 9. network lifetime with 250 nodes and 100 seconds

V. Conclusion

In order to resolve the congestion problem in Wireless Sensor Networks (WSNs), in this paper a solution for congestion control based on safe traffic management is proposed. Most proposed protocols for congestion control of WSNs suffer from some problems including central sink congestion control, using only one traffic control or resource control mechanisms and also having the same throughput for all nodes. However, the proposed method used both of "resource control" and "traffic management" approaches to congestion control in WSNs. In "resource control" part, the proposed method uses the available resources to reduce congestion and in "traffic management" part; it uses the exact rate setting. The main objective of the proposed solution is to avoid congestion problem and after that in second step, managing traffic with minimal energy consumption. Moreover, this method uses resource control strategy by selecting an alternative path that will control the traffics

The simulation results show that the proposed method improves the throughput, end to end delay, energy consumption and lifetime as compared to TADR base method.

TABLE II. Summarize the Advantages and Disadvantages/Limitations of the Proposed Method

Advantage	Since the proposed congestion control method uses both rate adjustment and alternative path method, it realized congestion occurrence immediately, then controled it by an optimal alternative path or rate adjustment.
Disadvantage or limitation	Since in explicit congestion declaration method, sensor nodes for exchanging congestion information, thy use special control packets, that led to the imposition of additional overhead to the network.

References

[1] R. Mohammadi, , Reza Javidan, and A. Jalili, "Fuzzy Depth Based Routing Protocol for Underwater Acoustic Wireless Sensor Networks", Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 7, no. 1, pp. 81-86, 2015.

[2] M. Keshtgary, Reza Javidan, and R. Mohammadi, "Comparative Performance Evaluation of MAC Layer Protocols for Underwater Wireless Sensor Networks", Modern Applied Science, Vol. 6, No. 3, 2012.

[3] R Mohammadi, S.Y. Nabavi, and Reza Javidan, "MAC Protocols in Underwater Wireless Sensor Networks: Issues and Simulations", Journal of Advances in Computer Research, Vol. 5, No. 2, pp. 97-109, 2014.

[4] K. Sohraby, D. Minoli, and T. Znati, Wireless Sensor Networks Technology, Protocols, and Applications, Hoboken, John Wiley & Sons, Inc, 2007.

[5] A. Rozyyev, H. Hasbullah, and F. Subhan, "Indoor child tracking in wireless sensor network using fuzzy logic technique," Research Journal of Information Technology, vol. 3, no. 2, pp. 81–92, 2011.

[6] R. Szewczyk, E. Osterweil, J. Polastre, M. Hamilton, A. Mainwaring, and D. Estrin, "Habitat monitoring with sensor networks," Communications of the ACM, vol. 47, no. 6, pp. 34–40, 2004.

[7] S. H. Chauhdary, A. K. Bashir, S. C. Shah, and M. S. Park, "EOATR: energy efficient object tracking by auto adjusting transmission range in wireless sensor network," Journal of Applied Sciences, vol. 9, no. 24, pp. 4247–4252, 2009.

[8] P. K. Biswas and S. Phoha, "Self-organizing sensor networks for integrated target surveillance," IEEE Transactions on Computers, vol. 55, no. 8, pp. 1033–1047, 2006.

[9] L. T. Lee and C. W. Chen, "Synchronizing sensor networks with pulse coupled and cluster based approaches," Information Technology Journal, vol. 7, no. 5, pp. 737–745, 2008.

[10] N. Sabri, S. A. Aljunid, B. Ahmad, A. Yahya, R. Kamaruddin, and M. S. Salim, "Wireless sensor actor network based on fuzzy inference system for greenhouse climate control," Journal of Applied Sciences, vol. 11, no. 17, pp. 3104–3116, 2011.

[11] D. Kumar, "Monitoring forest cover changes using remote sensing and GIS: a global prospective,"Research Journal of Environmental Sciences, vol. 5, pp. 105–123, 2011.

[12] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," Computer Networks, vol. 52, no. 12, pp. 2292–2330, 2008.

[13] T. Arampatzis, J. Lygeros, and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks," in Proceedings of the 20th IEEE International Symposium on Intelligent Control (ISIC '05), pp. 719–724, June 2005.

[14] Y.-C. Tseng, M.-S. Pan, and Y.-Y. Tsai, "Wireless sensor networks for emergency navigation," Computer, vol. 39, no. 7, pp. 55–62, 2006.

[15] A. Mahmood, K. Shi, Sh. Khatoon, and M. Xiao, "Data Mining Techniques for Wireless Sensor Networks: A Survey," International Journal of Distributed Sensor Networks Volume 2013, 24 pages, 2013.

[16] Brahma S, Chatterjee M, Kwiat K. "Congestion Control and Fairness in wireless sensor networks." Int.Conf. on 8th IEEE Pervasive Computing and Communications Workshops (PERCOM Workshops), IEEE; Mannheim; pp.413-18, 2010.

[17] M.C. Vuran, and I.F.Akyildiz. "XLP: A cross layer protocol for efficient communication in wireless sensor networks", IEEE Trans. Mobile Computing, 9(11), pp. 1578-1591, Nov. 2010.

[18] J. Sayyad, and N.K. Choudhari, "Congestion Control Techniques in WSN and Their Performance Comparisons", International Journal of Multidisciplinary and Current Research, 3, 108-113, 2015.

[19] M.A. Kafi, D. Djenouri, J. Ben-Othman, and N Badache, "Congestion Control Protocols in Wireless Sensor Networks: A Survey", IEEE Communications Surveys & Tutorial, 16(3), 2014.

[20] H. Karl, and A. Willig, Protocols and architectures for wireless sensor networks. John Wiley & Sons, New York, New York, USA; 2005.

[21] Y.M. Lu, and V.W.S.Wong, "An energy efficient multipath routing protocol for wireless sensor networks", International journal of communication systems, 14; 20(7):747-766, 2007.

[22] R. Hashemzehi, R. Nourmandipour, and F. koroupi, "Congestion in Wireless Sensor Networks and Mechanisms for Controling Congestion", Indian Journal of Computer Science and Engineering (IJCSE), 4(3), 204-207, 2013.

[23] R.A. Ostwal, M.B. Kalkumbe, and S.A. Bhosale, "An explore to congestion control in wireless sensor network", International Journal of Engineering and Innovative Technology (IJEIT), 4(7), 118-120, 2015.

[24] S. Pansare, and C.V. Kulkarni, "Design of Congestion Control Protocol for WMSN", International Journal of Innovative Technology and Exploring Engineering (IJITEE), 5(1), 58-62, 2015.

[25] P. Gowthaman, and R. Chakravarthi, "Survey on Various Congestion Detection and Control Protocols in Wireless Sensor Networks", International Journal of Advanced Computer Engineering and Communication Technology (IJACECT), 2(4), 15-19, 2013.

[26] C.Y. Wan, S.B. Eisenman, and A.T. Campbell, "Energy-Efficient Congestion Detection and Avoidance in Sensor Networks", ACM Transactions on Sensor Networks, 7(4), 32.1-32.31.

[27] S.R. Heikalabad, A. Ghaffari, M.A. Hadian, and H. Rasouli, "DPCC: Dynamic predictive congestion control in wireless sensor networks",. International Journal of Computer Science Issues, 8(1):472-477, 2011.

[28] K. Zhao, W. Liu, M.S. Wong, and J. Song, "Alternative path-based congestion control in many-to-one sensor networks", In 5th IEEE International ICST Conference on Communications and Networking in China (CHINACOM); Aug 25-27; Beijing, China, 2010.

[29] F. Ren, S.K. Das, and C.H. Lin, "Traffic-aware dynamic routing to alleviate congestion in wireless sensor networks" IEEE Transactions on Parallel and Distributed Systems, 21; 22(9):1585-1599, 2011.

[30] S. Brahma, M. Chatterjee, K. Kwiat, and P.K.Varshney, "Traffic management in wireless sensor networks: Decoupling congestion control and fairness. Computer Communications", 35(6):670-681, 2012.

[31] C. Sergiou, V. Vassiliou, and A. Paphitis, "Hierarchical Tree Alternative Path (HTAP) Algorithm for congestion control in wireless sensor networks",. Ad Hoc Networks, 11(1):257-272, 2013.

Static Filtered Sky Color Constancy

Ali Alkhalifah
Information Technology Department
College of Computer, Qassim University
Buraidah, Saudi Arabia

***Abstract*—In Computer Vision, the sky color is used for lighting correction, image color enhancement, horizon alignment, image indexing, and outdoor image classification and in many other applications. In this article, for robust color based sky segmentation and detection, usage of lighting correction for sky color detection is investigated. As such, the impact of color constancy on sky color detection algorithms is evaluated and investigated. The color correction (constancy) algorithms used includes Gray-Edge (G_E), Gray-World (G_W), Max-RGB (M_{RGB}) and Shades-of-Gray (S_G). The algorithms G_E, G_W, M_{RGB}, and S_G, are tested on the static filtered sky modeling. The static filter is developed in the LAB color space. This evaluation and analysis is essential for detection scenarios, especially, color based object detection in outdoor scenes. From the results, it is concluded that the color constancy before sky color detection using LAB static filters has the potential of improving sky color detection performance. However, the application of the color constancy can impart adverse effects on the detection results. For images, the color constancy algorithms depict a compact and stable representative of the sky chroma loci, however, the sky color locus might have a shifting and deviation in a particular color representation. Since the sky static filters are using the static chromatic values, different results can be obtained by applying color constancy algorithms on various datasets.**

***Keywords*—Static Filter; Color Constancy; LAB color space; Sky Color Detection; Horizon detection**

I. Introduction

In Computer Vision, the color based horizon/sky detection is used for lighting enhancement, image color enhancement, horizon alignment, image indexing, outdoor image classification and in countless many other applications [1]. From applications perspectives, the sky color is the most important feature that can be visualized in the outdoor scenes and image [2]. In a chromatic image, a set of pixels or region depicts a sky color if it corresponds to a set of color tones (blue to white) in the particular scene. In general, a pixel or region depicts sky region, if it corresponds to an image of the horizon [2]. The successful detection of sky region in an image facilitates a variety of imaging tasks, especially, the image based visual quality enhancement and object based manipulation tasks.

From the color perspective, the sky color falls in the area of ranging from blue tone to half white. However, this is the ideal scenario. Due to the atmospheric conditions, somewhere, other components are also present in sky patch. This naturally unbounded domain of the sky color ranges makes it a difficult proposition for the algorithms to optimally and correctly detect sky color in unconstrained environments. The challenges for optimally detecting the color ranges of the sky include mostly the clouds color, planes, sun alignment in the sky, forests, water and seas and other confused sky-like backgrounds in outdoor images.

The color of the sky in any digital image is possible to be detected from the image pixel values alone. This unique property benefits the processing speed of the algorithm. The algorithm thus based on color pixel alone can be executed extremely fast compare to the region based and complex features extraction based detection techniques. This also helps in the creation of a real-time sky color based detection and image enhancement algorithm. Moreover, regarding the color based sky detection; the detection is thus independent of image dimensions and pixels or scene orientation [3]. From the inherent nature of the pixel based sky detection, it is significantly dependent on the lightning variations in the image. Therefore, the strength of the algorithms is limited in uncontrolled (natural scenes and different lighting conditions) environments. This research work thus tackles the problem of lighting corrections for sky detection.

Zafarifar et al. [4] describe horizon detection approach using two features. There method takes benefit of the adapting positioning and color based analysis for sky segmentation and the extraction of the horizon part in an image/video. Authors advocate that the proposed algorithm produces superior performance compared to the state-of-the-art approaches especially, in natural outdoor image and scenes. Schmitt et al. in [5] use color, shape and position vector as an input feature for the position of the sky detection in an image. The performance their method is measured and tested on a number of out-door images and in different weather and lighting conditions. The authors in their work describe the superior performance of their algorithm compared to the other similar algorithms in separating sky regions from non-sky regions under different weather conditions. The authors in [6] develop sky color based solar exposure system using the image processing techniques, segmenting the outdoor images recorded under different lighting conditions. Authors in [7] explain through physics phenomenon that due to the properties of light scattered by small particles in the atmosphere, the clear color of the sky often appears deep, saturated blue shade at the top of the image and slowly unsaturates to typically white toward an infinite horizon line in an image. Since the detection is based on robust representation, the detector developed in [7] is unlikely to be fooled by similar colored materials and objects, for example, bodies of water, frontal walls, trees, and clothes. The classification is based on an acceptable gradient signal. The method achieves an excellent performance; however, it fails to correctly identify minor sky regions, for example, a small sky patch/region visible between buildings or trees. A random forest [8], Artificial neural networks [9], Bayesian networks [10], Radial Basis Function (RBF) [11,12] are most of the time used for feature based object detection in images.

In this article, for robust color based sky (horizon) segmentation and detection, usage of lighting correction for sky color detection is investigated. As such, the impact of color constancy on sky color detection algorithms is evaluated and investigated. The color constancy algorithms used includes G_E, G_W, M_{RGB} and S_G. The algorithms G_E, G_W, M_{RGB} and S_G, are tested on the static filtered sky modeling. The static filter is developed in the LAB color space. This evaluation and analysis are essential for detection scenarios, especially, color based object detection in outdoor scenes. From the results, it is concluded that the color constancy before sky color detection using LAB static filters has the potential of improving sky color detection performance. However, the application of the color constancy can impart negative effects on the detection results. For digital images, the color correction methods represent a stable representative of the horizon in chromaticity space loci. Also, the sky color representation can have a shifting in the corresponding space. Since the sky static filters are using the fixed color space values, different results are obtained by applying color constancy algorithms on different datasets.

II. Color Constancy

Color constancy is the inherent ability of the vision for resolving an object color in a particular scene that is basically independent of the illuminant source. In general, the color constancy is also defined as the ability to estimating the undefined/unknown light source of an active scene from a given image. From mathematical perspectives, assume that an image f is composed of [14]:

$$f(x) = \int_{w} e(\lambda)c(\lambda)s(x,\lambda)d\lambda$$

The equation shows that the light source of a color spectrum is e (λ), the apparent surface reflectance is s(x, λ) and the image capturing source sensitivity is c(λ). In the color constancy, the visible spectrum is represented by w and spatial coordinates of the image are represented by x. From implementation point of view, the color constancy is an estimation of the light source e(λ) from given image. We use four color correction algorithms. Color correction approaches provide an estimation of e depending on the paradigm they are using. In the following, we discuss Gray Edge (G_E), Gray World (G_W), Max RGB (M_{RGB}), and Shades of Gray (S_G).

A. Gray Edge (G_E)

The G_E is presented in [14]. It proposes the G_E hypothesis as: “the mean of the difference of the reflectance in a particular scene is achromatic." Mathematically, [14]:

$$\left(\frac{\int |f_x^\sigma(x)|^p dx}{\int dx}\right)^{\frac{1}{p}} = ke$$

B. Gray World (G_W)

In [15], Buchsbaum proposes the G_W hypothesis. Contrary to the G_E Algorithm, the G_w assumes that the average reflectance in the particular scene is achromatic. The G_E, however, accounts for the differences of the reflectance in a scene. In [15], the author derives: That the average reflectance for the short, middle and longer oscillations are identical [16].

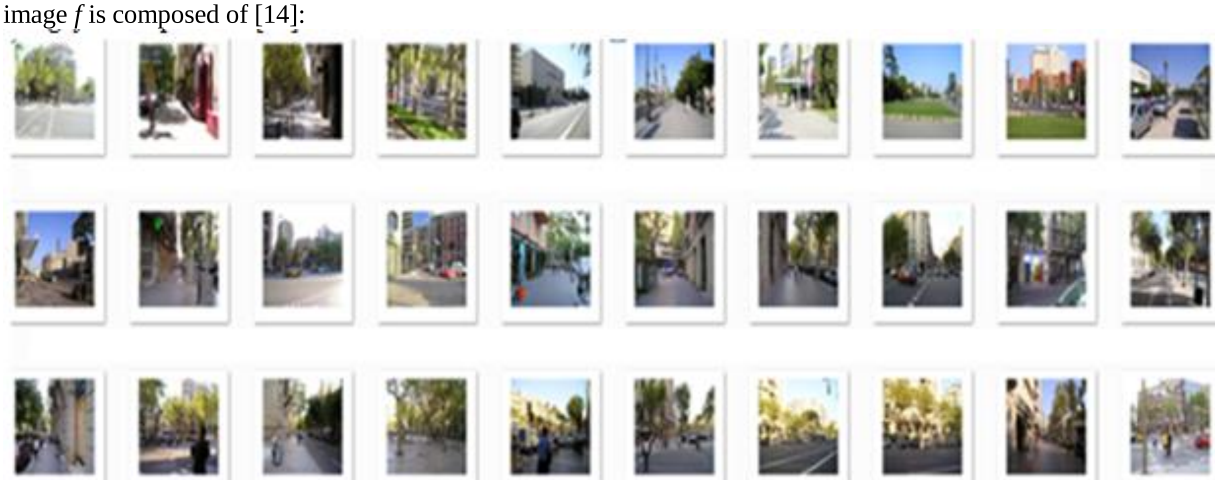


Fig. 1. Original images containing sky patch

C. Max RGB (M_{RGB})

M_{RGB} relies on the assumption that the reflectance value of each of the three Red, Green and Blue channels (from image perspectives) is equal [14].

D. Shades of Gray (SG)

The G_W and the M_{RGB} algorithms are the bi instances of the color constancy paradigm based on the Minkowski norm [14]. The authors in [17] analyzed the performance of the lighting source e as a function of the Minkowski norm. It was deduced that the optimal performance is obtained with a Minkowski while setting p to 6 is termed as Shades of Gray [17].

III. Methodology

In this section, the method to filtering and color constancy application is explained. The LAB static filter is created in the LAB color space on a set of training sky pixels. We choose the LAB as the color space and its static filter because of its non-linear transformation, proper separation of lighting and color components and its generally good performance for sky color detection.

The LAB static filter is a rule based filter represented as follows:

$$F_{sf}(p)_L = min\_L+C < L < max\_L-C \text{ ------(1)}$$
$$F_{sf}(p)_a = min\_a+C < a < max\_a-C \text{ ------(2)}$$
$$F_{sf}(p)_{b=} \ min\_b < b < max\_b \text{ ------(3)}$$

Whereas:

p is an image pixel, C=0.1, min_L=58.42, max_L=100, min_a=-19.96, max_a=4.76, min_b=-29.62 and max_b=9.65.

Given an image, I, static filter created in Equations 1-3 are used to process and obtain the sky blobs from the image.

$$I_{\_sky\_blobs}=F_{sf}(I)\text{------(4)}$$
$$[R_e\ G_e\ B_e]=CC(I_{\_sky\_blobs})\text{------(5)}$$
$$I_{\_e}=I/[R_e\ G_e\ B_e]\text{------(6)}$$

Where $I_{\_sky\_blobs}$ is the detection from the static filter. From an Equation 5, the R_e, G_e, B_e are the lighting parameters from the color constancy algorithm (CC). The final image ($I_{\_e}$) is obtained by applying parameters from the color constancy on the original image.

IV. Experimental Evaluation

The algorithms G_E, G_W, M_{RGB} and S_G are used with the static filtered sky modeling. The static filter is developed in the LAB color space, represented as S_{LAB}. For experimental analysis, the Sky Data Set (S_{DS}) is used which is extracted from [13].

Fig. 2. Mask images of original images used in the experiments

A. Sky Data Set (S_{DS})

The S_{DS} used for evaluation is extracted from [13]. The S_{DS} contains natural scenes images and the manually annotated pixel-level ground truth of the original images. The original images are shown in Figure 1. The S_{DS} has 1000 images. The images used for testing (annotated) images are called the mask images of original images as show in Figure 2. The mask images are used for performance evaluation of the static filtered color constancy algorithms.

B. Evaluation

We show the effect of color constancy algorithms on LAB static sky filter using G_E, G_W, M_{RGB} and S_G.

Figure 3 shows the output of the color constancy algorithms used.

Fig. 4 is showing two examples of the results of sky color detection after applying G_E color constancy. Figure 4 (Column 1) shows actual images, Figure 4 (Column 2) depicts sky detection using the fixed filter in the Lab color space and Figure 4 (Column 3) shows improvement in sky space detection by using the same fixed static filter (after applying the G_E algorithm). This result shows that some of the pixels that were not blocked by static filter are now blocked by static filter after applying the G_E. Figure 4 is also showing that color correction can benefit the sky detection.

Color constancy may not improve results in all lighting conditions and all images across dataset. Also, the application of color constancy can negatively affect the results. As shown in Figure 5, applying G_e (first row, third column) has the same result as the static filter alone (first row, second column). In Figure 5, (second row and third column), applying color constancy has blocked sky pixels and thus increasing the false positives.

Fig. 3. Color constancy algorithms and the resultant images

Fig. 4. Horizon detection can be improved by applying color correction algorithm. Column 1 depicts actual sky images. Column 2 shows the result of the filter without applying color correction. Column 3 shows the output of the static filter after applying color constancy. As can be seen in Column 3, improved results are obtained

Fig. 5. Color correction may reduce performance in some scenarios. Column 1 is showing original sky images. Column 2 is showing the result without applying color correction. Column 3 shows the output of the algorithm after applying color constancy

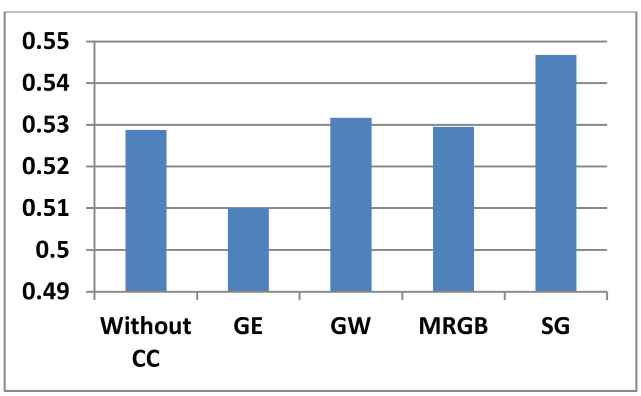


Fig. 6. shows the evaluation results (accuracy) for static filtered based color constancy on dataset S_{DS}

Figure 6 shows the evaluation of the static filter using color constancy algorithms. From the evaluation perspectives, we use the Accuracy as an evaluation measure. For S_{DS}, we find that the Accuracy of 0.528 achieved by not using color correction is reduced to 0.51 when using the static sky filter with the G_E. The color constancy algorithm G_w reports an increased performance with Accuracy of 0.531. In the case of M_{RGB}, the sky color performance is slightly enhanced to 0.529. S_G reports increased performance of 0.546.

From the results, it is concluded that the color constancy before sky color detection using LAB static filters has the potential of improving sky color detection performance. However, the application of the color constancy can impart negative effects on the detection results. This negative behavior can be explained as follows: For digital images, the color correction methods are depicting a stable and compact structure of the sky chromaticity space; however, the sky color part in color space can have a shifting and deviation in the color space. As, the sky static filters are using the static chromatic parameters, therefore, varying results might be obtained when applying color constancy algorithms on different datasets.

V. Conclusion

In this article, for robust color based sky (horizon) segmentation and detection, usage of lighting correction for sky color detection is investigated. As such, the impact of color constancy on sky color detection algorithms is evaluated and investigated. The color constancy algorithms used includes: Gray Edge (G_E), Gray World (G_W), Max RGB (M_{RGB}) and Shades of Gray (S_G). , it is concluded that the color constancy before sky color detection using LAB static filters has the potential of improving sky color detection performance. However, the application of the color constancy can impart negative effects on the detection results. This negative behavior can be explained as follows: For digital images, the color correction methods are depicting a stable and compact structure of the sky chromaticity space; however, the sky color part in color space can have a shifting and deviation in the color space. As, the sky static filters are using the static chromatic parameters, therefore, varying results might be obtained when applying color constancy algorithms on different datasets.

References

[1] I. Khan, K. Haider, Q. Sattar, S. U. Rehman, A. Ali "An Efficient Approach for Sky Detection" IJCSI, 2013, Pages: 222-226
[2] A. C. Gallagher, Eastman; J. Luo and W. Hao "Improved Blue Sky Detection Using Polynomial Model Fit", ICIP, 2004, pages: 2367-2370
[3] R. Khan, A. Hanbury, J. Stöttinger, A. Bais "Color based sky color classification. Pattern Recogn Lett 33(2): Pages: 157–163.
[4] B. Zafarifar and P. Henry , "Blue Sky Detection for Content based Television Picture Quality Enhancement," in Proc. IEEE International Conference on Consumer Electronics, 2007, pp. 437–438.
[5] F. Schmitt, L. Priese., "Sky detection in CSC segmented color images". In Proceedings. Computer Vision Theory and Applications (VISAPP), (2009), vol. 2, pp. 101–106.
[6] N. Laungrunthip, A. Mckinnon, C. Churcher, K. Unsworth, "Sky Detection in Images for Solar Exposure Pediction", (2008),
[7] J. Luo and S. Etz, "A Physical Model-based Approach to Sky Detection in Photographic Images:' IEEE Pans. On Image Processing, vol. II, no. 3, pp. 201-212, 2002.
[8] T. Ho "Random Decision Forests." IEEE ICIP, 1995, 1: 278–82.
[9] M. Christopher, Bishop "Neural Networks for Pattern Recognition" Oxford University Press, 1 edition, January 1996.
[10] N. Friedman, D. Geiger, and M. Goldszmidt "Bayesian network classifiers" Mach. Learning., 29:131–163, November 1997.

[11] P. Yee and S. Haykin. "A dynamic regularized radial basis function network for nonlinear, nonstationary time series prediction" IEEE Transactions on Signal Processing, 47:2503–2521, 1998.

[12] M. Carlin. "Radial basis function networks and nonlinear data modelling" In Neural Networks and their Applications, volume 1, pages 623–633, 1992.

[13] Russell, C. Bryan, A. Torralba, K. P. Murphy, and W. Freeman "LabelMe: A Database and Web-Based Tool for Image Annotation." International Journal of Computer Vision 77: 157–73.

[14] J. Weijer, T. Gevers, and A. Gijsenij "Edge-based color constancy" IEEE Transactions on Image Processing, 16(9):2207–2214, 2007.

[15] G. Buchsbaum. "A spatial processor model for object color perception" Journal of the Franklin Institute, 310:1 – 26, 1980.

[16] K. Barnard, V. Cardei, and B. Funt. "A comparison of computational color constancy algorithms. Methodology and experiments with synthesized data". Image Processing, IEEE Transactions on, 11(9):972–984, 2002.

[17] D. Graham, Finlayson and E. Trezzi. "Shades of gray and color constancy" In Color Imaging Conference, pages 37–41, 2004.

An Approach to Finding Similarity Between Two Community Graphs Using Graph Mining Techniques

Bapuji Rao
PhD Scholar (CSE)
Biju Patnaik University of Technology (BPUT)
Rourkela, Odisha, India

Saroja Nanda Mishra
CSE & A
Indira Gandhi Institute of Technology (IGIT)
Sarang, Odisha, India

***Abstract*—Graph similarity has studied in the fields of shape retrieval, object recognition, face recognition and many more areas. Sometimes it is important to compare two community graphs for similarity which makes easier for mining the reliable knowledge from a large community graph. Once the similarity is done then, the necessary mining of knowledge can be extracted from only one community graph rather than both which leads saving of time. This paper proposes an algorithm for similarity check of two community graphs using graph mining techniques. Since a large community graph is difficult to visualize, so compression is essential. This proposed method seems to be easier and faster while checking for similarity between two community graphs since the comparison is between the two compressed community graphs rather than the actual large community graphs.**

Keywords—community graph; compressed community graph; dissimilar edges; self-loop; similar edges; weighted adjacency matrix

I. Introduction

A graph arises in many situations like web graph of documents, a social network graph of friends, a road-map graph of cities. Graph mining has grown rapidly for the last two decades due to the number, and the size of graphs has been growing exponentially (with billions of nodes and edges), and from it, the authors want to extract much more complicated information. Graph similarity has numerous applications in social networks, image processing, biological networks, chemical compounds, and computer vision, and therefore it has suggested many algorithms and similarity measures. Graph similarity is that "a node in one graph is similar to a node in another graph if their neighborhoods are similar" [1].

II. Literature Survey

Graphs are general object model; graph similarity has studied in many fields. Similarity measures for graphs have used in systems for shape retrieval [2], object recognition [3] or face recognition [4]. For all those measures, graph features specific to the graphs in the application, are exploited to define graph similarity. Examples of such features are given one to one mapping between the vertices of different graphs or the requirement that all graphs are of the same order.

A very common similarity measure for graphs is the edit distance. It uses the same principle as the well-known edit distance for strings [5, 6]. The idea is to determine the minimal number of insertions and deletions of vertices and edges to make the compared graphs isomorphic. In [7], Sanfeliu and Fu extended this principle to attributed graphs, by introducing vertex relabeling as a third basic operation beside insertions and deletions. In [8], the measure is used for data mining in a graph.

The main idea behind the feature extraction method is that similar graphs probably share certain properties, such as degree distribution, diameter, and Eigen values [9]. After extracting these features, a similarity measure [10] is applied to assess the similarity between the aggregated statistics and, equivalently, the similarity between the graphs. In the iterative method "two nodes are similar if their neighborhoods are also similar".

In each iteration, the nodes exchange similarity scores, and this process ends when convergence has achieved. A successful algorithm belongs to this category is the similarity flooding algorithm by Melnik et al. [11] applies in database schema matching. It solves the "matching" problem, and attempts to find the correspondence between the nodes of two given graphs. Another successful algorithm is SimRank [12], which measures the self-similarity of a graph, i.e., it assesses the similarities between all pairs of nodes in one graph. Furthermore, another successful recursive method related to graph similarity and matching is the algorithm proposed by Zager and Verghese [13]. This method introduces the idea of coupling the similarity scores of nodes and edges to compute the similarity between two graphs.

A new method to measure the similarity of attributed graphs proposed in [14]. This technique solves the problems mentioned in similarity measures for attributed graphs and is useful in the context of large databases of structured objects. First, BP-based algorithm implemented for graph similarity [1] uses the original BP algorithm as it is proposed by Yedidia [15]. This algorithm is naive and runs in O (n^2) time.

III. Proposed Method

In the literature survey the authors have studied thoroughly the existing methods which checks for similarity of two graphs. In this paper, the authors have proposed graph mining techniques for checking of similarity between two community graphs. Further, the authors have proposed a community graph which is depicted in "Fig. 1". For similarity measure of two community graphs, the authors have first compressed both the community graphs. Then the compressed community graphs are used for comparison for similarity. The authors have adopted the compression of large community graph to smaller one technique from [16].

The authors have proposed a village community graph having ten communities namely C_1 to C_{10}, and the total number of community members is 118. The black color edge represents the edge among the community members of similar communities. Whereas the blue color edge represents the edge among the community members of dissimilar communities.

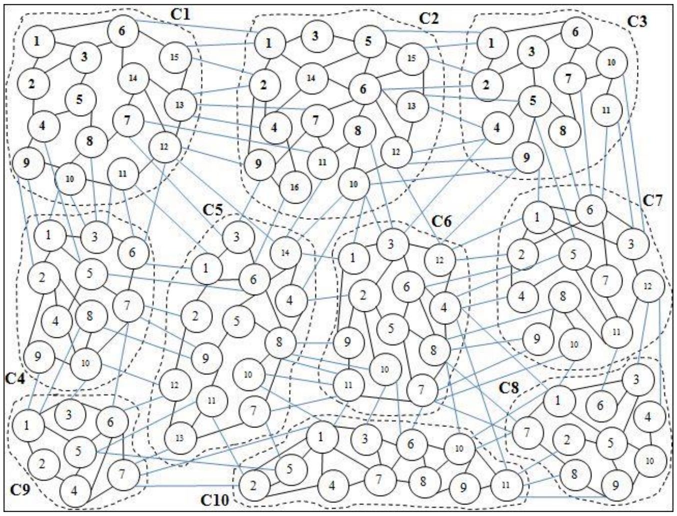


Fig. 1. Community graph with communities {C1, C2, C3, C4, C5, C6, C7, C8, C9, C10}

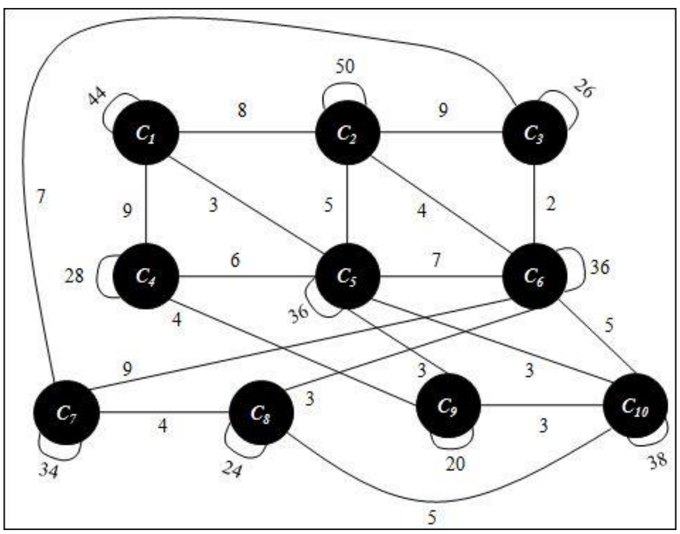


Fig. 2. Compressed community graph of Fig.1

To compress the community graph to a smaller one depicted in "Fig. 1", the authors have adopted the logic from [16]. The compressed community graph is depicted in "Fig. 2". Then its corresponding adjacency matrix is represented in the memory and depicted in "Fig. 3". In this weighted adjacency matrix, the self-loop of the community has some weight and considered as a total number of edges among the community members of that particular community. Similarly, the edge between the pair of communities as the total number of edges between the community members of dissimilar communities. For this proposed approach, the authors have considered "Fig. 1" community graph as the principle community graph for comparison with six more community graphs namely CG_2 to CG_7. Before comparison, these six community graphs, i.e., CG_2 to CG_7's adjacency matrices are compressed and represented in the memory. Finally, the principle community graph CG_1's compressed adjacency matrix has compared with all the six community graphs, i.e., CG_2 to CG_7's compressed adjacency matrices for similarity check. The details of all the seven community graphs, CG_1 to CG_7 has considered as datasets for the proposed algorithm is listed in "Table I".

Edges of similar community members of community

Community Codes	Community Codes									
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
C_1	44	8	0	9	3	0	0	0	0	0
C_2	8	50	9	0	5	4	0	0	0	0
C_3	0	9	26	0	0	2	7	0	0	0
C_4	9	0	0	28	6	0	0	0	4	0
C_5	3	5	0	6	36	7	0	0	3	3
C_6	0	4	2	0	7	36	9	3	0	5
C_7	0	0	7	0	0	9	34	4	0	0
C_8	0	0	0	0	0	3	4	24	0	5
C_9	0	0	0	4	3	0	0	0	20	3
C_{10}	0	0	0	0	3	5	0	5	3	38

Fig. 3. Adjacency matrix of Fig.2

IV. Proposed Algorithm

The proposed algorithm has three phases. Phase-1 is to open for reading four dataset files. The dataset files *commun1.txt* and *commun2.txt* for reading number of communities, and community code and their total number of community members of two community graphs CG_1 and CG_2, and assign to the matrices NCM1[][] and NCM2[][] respectively. Similarly two more dataset files *data1.txt* and *data2.txt* for reading edge details of two community graphs CG_1 and CG_2, and assign to the matrices CMM1[][] and CMM2[][] respectively. So Phase-1 is about read data and creation of community member matrices, and creation of initial form of compressed community matrices.

Pahse-2 for counting edges of community members of same communities by calling procedure SCED( ) and counting edges of community members of dissimilar communities by calling procedure DCED( ). Using procedures SCED( ) and DCED( ), the compressed community adjacency matrices CCM1[][] and CCM2[][] are assigned with the edge values and self loop values.

Finally, Phase-3 for comparison of both the compressed community matrices by calling procedure CS( ). Further, it returns a numerical value i.e., from 0 to 3. So based on the numerical value, the similarities of both of community graphs are judged. The numerical value 1 for similarity; whereas values 0, 2, and 3 for no similarity between communities graph CG_1 and CG_2.

A. Algorithm for Community Graph Similarity

Algorithm Community_Graph_Similarity ( )

Algorithm Convention [17]

//CG_1, CG_2: Given two community graphs with 'n1' and 'n2' //number of community members.

//tcm1, tcm2: To assign total community members of //community graphs CG_1 and CG_2.

//NCM1[n1][2], NCM2[n2][2]: Matrices to hold community //member's community number and number of community //members of CG_1 and CG_2.

//CMM1[tcm1+1][tcm1+1], CMM2[tcm2+1][tcm2+1]: //Adjacency matrices of CG_1 and CG_2.

//CCM1[n1][n1], CCM2[n2][n2]: Adjacency matrices of //compressed community graphs of CG_1 and CG_2.

//*commun1.txt*, *commun2.txt*: Text file contains number of //communities, and community code and their total number of //community members of CG_1 and CG_2.

//*data1.txt*, *data2.txt*: Text file contains edge details of CG_1 //and CG_2.

//flag: To assign the similarity check value from 0 to 3.

```
{
  n1:=RCD (NCM1, "commun1.txt");  // CG1 details
  n2:=RCD (NCM2, "commun2.txt");  // CG2 details
  tcm1:=ACMC (NCM1, n1, CMM1, CCM1);
  tcm2:=ACMC (NCM2, n2, CMM2, CCM2);
  CMMatrix (CMM1, tcm1, "data1.txt");
  CMMatrix (CMM2, tcm2, "data2.txt");
  SCED (NCM1, n1, CMM1, CCM1);
  SCED (NCM2, n2, CMM2, CCM2);
  DCED (NCM1, n1, CMM1, tcm1, CCM1);
  DCED (NCM2, n2, CMM2, tcm2, CCM2);
  flag:=CS (CCM1, n1, CCM2, n2);
if(flag=0) then write("Both the Community Graphs are not Similar");
if(flag=1) then write("Both the Community Graphs are Similar");
if(flag=2) then write("Both the Community Graphs are Similar on Similar Edges");
if(flag=3) then write("Both the Community Graphs are Similar on Dissimilar Edges");
}
```

B. Procedure for Community Data Read

Procedure RCD (NCM, FileName)

```
// n: To assign number of communities.
// cc: To assign community code.
// tcm: To assign total community members
{
    open(FileName);
    read(n);
    for i:=2 to (n+1) do
    {
      read(cc, tcm);
      NCM[i-1][1]:=cc;
      NCM[i-1][2]:=tcm;
    }
    close(FileName);
    return(n);
}
```

C. Procedure for Assignment of Community Member Codes

```
Procedure ACMC (NCM, n, CMM, CCM)
// k: index variable, tcm: to count total community members
{
  k:=2;
  tcm:=0;
 for i:=1 to n do
 {
    tcm:=tcm+NCM[i][2];
  for j:=1 to NCM[i][2] do // assignment of community codes
                            // in community member matrix
  {
     CMM[1][k]:=CMM[k][1]:=j;
     k:=k+1;
   }
 }
//assignment of community codes in compressed community
//matrix CCM[][]
  for i:=2 to (n+1) do
  {
    CCM[i][1]:= CCM[1][i]:= NCM[i-1][1];
   }
   return(tcm);
}
```

D. Procedure for Community Member Matrix Creation

```
Procedure CMMatrix (CMM, tcm, FileName)
{
  open(FileName);
  i:=2;
  j:=2;
  while (i ≠ (tcm+1)) do
  {
    read(data);
    if (j=(tcm+1)) then { i:=i+1;   j:=2; }
    CMM[i][j]:=data;
    j:=j+1;
   }
    close(FileName);
}
```

E. Procedure for Same Community Edge Detection

```
Procedure SCED (NCM, n, CMM, CCM)
{
  d:=1;
  s:=0;
 for i:=1 to n do
 {
    s:= s + NCM[i][2];
      for j:=d to s do
        for k:=d to s do
```

```
      if (CMM[j+1][k+1]=1) then
        CCM[i+1][i+1]:=CCM[i+1][i+1]+1;  //check for edge
                                          // at CMM[j+1][k+1]
      d:=s;
   }
}
```

F. Procedure for Counting Dissimilar Edges

```
Procedure DCED (NCM, n, CMM, tcm, CCM)
{
   a:=1;
   b:=NCM[1][2];
   c:=b;
   d:=b;
   for i:=2 to (n+1) do
   {
      d:= d + NCM[i][2];
      // to count dissimilar communities edges
      Count_Edge (i-1, a, b, c, d, NCM, CMM, tcm, CCM);
      a:=b;
      b:=b + NCM[i][2];
      c:=d;
    }
}
```

G. Procedure to Count Dissimilar Communities Edges

```
Procedure Count_Edge (p, a, b, c, d, NCM, CMM, tcm, CCM)
// a, b: Initial and final index of row.
// c, d: Initial and final index of column.
// p: Initial index of CCM[][].
{
   x:= c;
   y:=d;
   k:=p+1;
   for i:=a to b do
   {
      k:=p+1;
   Bapu:
      for j:=c to d do
  if(CMM[i+1][j+1]=1) then
  {
CCM[p+1][k+1]:=CCM[p+1][k+1]+1;    // row-side dissimilar
                                   // community edges counting
CCM[k+1][p+1]:=CCM[k+1][p+1]+1;             // column-side
                          // dissimilar community edges counting
  }
      k:=k+1;
      if(d<tcm) then
     {
        c:=d;
        d:=d+NCM[k][2];
        goto Bapu;
      }
      c:=x;
      d:=y;
   }
}
```

H. Procedure for Similarity Check Between Community Matrices

```
Procedure CS (CCM1, n1, CCM2, n2)
{
   flag:=flag1:=flag2:=count:=0;
   if(n1 ≠ n2) then return(0); // both the community graphs are
                               // dissimilar
   else
   {
     // arrange both matrices in ascending order
     Arrange(CCM1, n1);
     Arrange(CCM2, n2);

     // check for dissimilar communities
    for i:=2 to (n1+1) do
     for j:=2 to (n2+1) do
        if(CCM1[1][i]=CCM2[1][j]) then count:=count+1;

     if(count=n1) then flag:=1; else flag:=0;
     // check for same communities
     if(flag=1) then
     {
       // check for same number of edges of each communities
       for i:=2 to (n1+1) or (n2+1) do
          if(CCM1[i][i] ≠ CCM2[i][i]) then
          {
            flag1:=1;
            break;
          }
    // check for different number of edges among communities
       for i:=2 to (n1+1) do
         for j:=2 to (n2+1) do
            if(j>i) then
               if(CCM1[i][j] ≠ CCM2[i][j])  then
               {
                 flag2:=1;
                 break;
               }
   if(flag1=1 and flag2=1) then return(0);     // same number
          // communities but different number of similar and
          // dissimilar edges
   else if(flag1=0 and flag2=1) then return(2);   // similarity on
                                                  // similar edges
   else if(flag1=1 and flag2=0) then return(3);   // similarity on
                                                  // dissimilar edges
   else return(1);      // same number of similar and dissimilar
                                                  // edges
   }
   else
       return(0);      // number of communities same but not its
                                  // community codes (numbers)
  }
}
```

I. Procedure for Sorting of Compressed Community Matrix

```
Procedure Arrange (mat, n)
// t[]: temporary array for swap.
```

```
{
  // row-side community code arrangement
   for i:=2 to n do
    for j:=i+1 to (n+1) do
      if(mat[i][1]>mat[j][1]) then
         for k:=1 to (n+1) do
         {
           t[k]:=mat[i][k];
           mat[i][k]:=mat[j][k];
           mat[j][k]:=t[k];
          }
  // column-side community code arrangement
    for i:=2 to n do
     for j:=i+1 to (n+1) do
       if(mat[1][i]>mat[1][j]) then
          for k:=1 to (n+1) do
          {
            t[k]:=mat[k][i];
            mat[k][i]:=mat[k][j];
            mat[k][j]:=t[k];
           }
}
```

V. Evaluation of Algorithm and Results

To evaluate the performance of the proposed algorithm, the authors have considered seven community graphs namely CG_1 to CG_7, where 1st community graph CG_1 is considered as principle community graph for comparison with the remaining six community graphs for finding similarities.

For the seven examples of community graphs, two sets of dataset files were created for each example of community graphs. The 1st dataset file contains community graph details such as number of communities, community number, and number of community members. So for the seven community graphs, these dataset files were from *datacom1.txt* to *datacom7.txt*. Similarly the 2nd dataset file contains community graphs edge details i.e., edge between community members which only consist of 1s and 0s. So for the seven community graphs, these dataset files were from *dataedg1.txt* to *dataedg7.txt*. These fourteen dataset file details are depicted in "Table I".

The algorithm was written in C++ and compiled with TurboC++ and run on Intel Core I5-3230M CPU +2.60 GHz Laptop with 4GB memory running MS-Windows 7. The comparison results of CG_1 with CG_2 to CG_7 are depicted from "Fig. 6" to "Fig. 17".

The datasets for community graphs CG_1 to CG_7 are in text files from *datacom1.txt* to *datacom7.txt* and for *datacom1.txt* is depicted in "Fig. 4", which contains the total number of communities, community numbers, and a total number of community members. Similarly, the datasets for community graphs CG_1 to CG_7 are in text files from *dataedg1.txt* to *dataedg7.txt* and for *dataedg1.txt* is depicted in "Fig. 5", which contains the edge details, i.e., 0s (no edge) and 1s (edge) between the community members of similar communities as well as dissimilar communities of the community graphs.

TABLE I. Dataset Table

Sl. No.	Community Graphs	Communities	Total Communities	Total Community Members	Number of Edges	Community Data File Name	Community Member Edge File Name
1.	CG_1	C_1 to C_{10}	10	118	435	datacom1.txt	dataedg1.txt
2.	CG_2	C_1 to C_{10}	10	118	466	datacom2.txt	dataedg2.txt
3.	CG_3	C_1 to C_{10}	10	118	439	datacom3.txt	dataedg3.txt
4.	CG_4	C_1 to C_{10}	10	118	471	datacom4.txt	dataedg4.txt
5.	CG_5	C_3 to C_8 C_{10} to C_{13}	10	118	434	datacom5.txt	dataedg5.txt
6.	CG_6	C_1 to C_{12}	12	118	515	datacom6.txt	dataedg6.txt
7.	CG_7	C_1 to C_{10}	10	118	435	datacom7.txt	dataedg7.txt

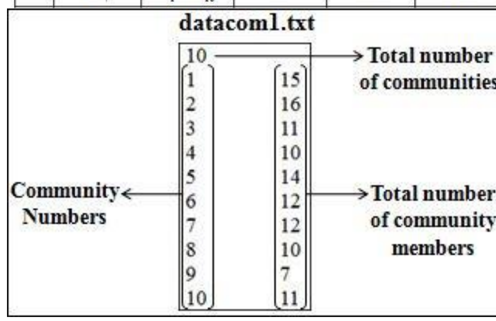


Fig. 4. Dataset file of CG_1

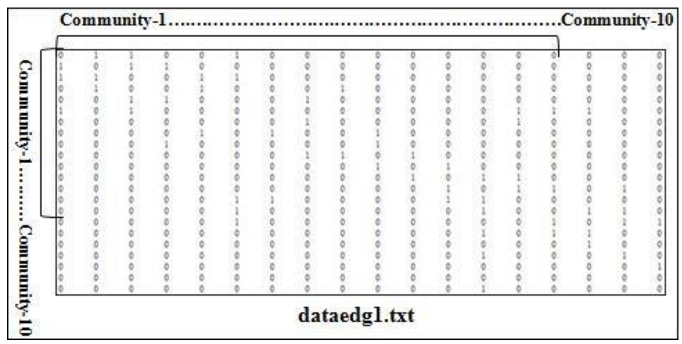


Fig. 5. Dataset of CG_1 contains edge details of community members C_1 to C_{10}

The authors have studied the existing techniques of Danai Koutra et al. method [1], Sergey Melnik et al. method [2], Glen Jeh et al. method [12], L Zager et al. method [13], and Hans-Peter Kriegel et al. method [14] for graph similarity.

In Danai Koutra et al. method [1], two graphs $G_1(N_1, E_1)$ and $G_2(N_2, E_2)$, with possibly different number of nodes and edges for similarity check, then adopting belief propagation (BP) into the proposed method for finding similarity between two graphs which finally returns a similarity value i.e., a real number between 0 and 1.

In Sergey Melnik et al. method [2], the matching of two graphs based on a fixed point computation. It takes two graphs as input, which is preferably a schema or catalog or other data structures for similarity check. Finally, it produces the result as mapping between the corresponding nodes of the graphs. Depending on the matching goal, a sub-set of the mapping is chosen using some filtering methods. Moreover, it allows the user to adjust the results if it is necessary.

In Glen Jeh et al. method [12], to find similarity between two objects based on their relationships. Two objects are said to be similar, if they are related to similar objects. This similarity measure is called SimRank. This method is based on the simple graph-theoretic model.

In L Zager et al. method [13], it is a node-edge coupling, i.e., two graph elements is similar if their neighborhoods are similar. So edge score is constructed "when an edge in G_1 is like an edge in G_2 if their respective source and terminal nodes are similar". This is called edge similarity.

In Hans-Peter Kriegel et al. method [14], attributed graphs are considered as a natural model for the structured data. The authors proposed a new similarity measure between two attributed graphs, called "matching distance". The matching distance is calculated by sum of the cost for each edge matching.

The proposed method in this paper is different from the above existing methods. In the proposed method two community graphs with possibly equal number of nodes (communities) and different number of edges for similarity check. Each node (community) is labeled with a unique community number. Based on the community number of node, the similarity measure takes place by considering the weight of self-loop of community as well as the weight of edge between the communities. After similarity between two community graphs, it finally returns a similarity value i.e., a number from 0 to 3. Based on this number, the similarity of two community graphs can be judged. The proposed algorithm has capable of showing similarity and five different ways of dissimilarity. The five different dissimilarities are "similar on dissimilar edges", "similar on similar edges", "communities same but different edges", "communities not same", and "number of communities are different". Moreover, the proposed method is completely based on labeled community graphs and simple graph-theoretic model. So the authors conclude that the proposed community graph similarity is simply different from the above existing methods and fast since the time complexity is $O(n^3)$.

A. *Comparison of CG_1 and CG_2*

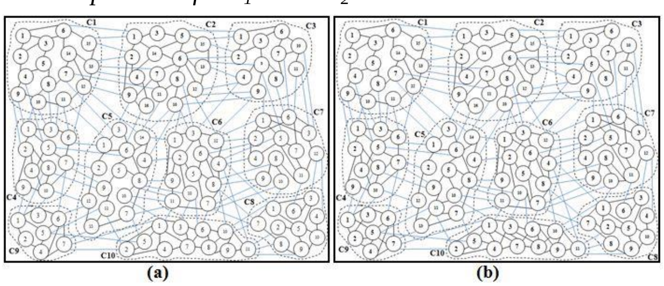


Fig. 6. (a) Community Graph CG_1 (b) Community Graph CG_2

In community graph CG_1, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:9, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

In community graph CG_2, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 46, 24, 32, 42, 42, 38, 28, 26, 46}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:9, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

The comparison takes place on community graph CG_1 and CG_2's community codes and the number of edges belonging to dissimilar community codes member since these two are same. So finally the algorithm shows as "*Both the Community Graphs are Similar on Dissimilar Edges*".

```
< Enter 1st Community Graph's Details >

Enter the Community Data File Name :  datacom1.txt

Enter the Edge Data File Name :  dataedg1.txt

< Enter 2nd Community Graph's Details >

Enter the Community Data File Name :  datacom2.txt

Enter the Edge Data File Name :  dataedg2.txt

Press Any Key to see the 1st Compressed Adjacency Matrix

 1st Compressed Community Graph's Adjacency Matrix

 C    1    2    3    4    5    6    7    8    9   10
 1   44    8    0    9    3    0    0    0    0    0
 2    8   50    9    0    5    4    0    0    0    0
 3    0    9   26    0    0    2    7    0    0    0
 4    9    0    0   28    6    0    0    0    4    0
 5    3    5    0    6   36    7    0    0    3    3
 6    0    4    2    0    7   36    9    3    0    5
 7    0    0    7    0    0    9   34    4    0    0
 8    0    0    0    0    0    3    4   24    0    5
 9    0    0    0    4    3    0    0    0   20    3
10    0    0    0    0    3    5    0    5    3   38

 Press Any Key to see the 2nd Compressed Adjacency Matrix

 2nd Compressed Community Graph's Adjacency Matrix

 C    1    2    3    4    5    6    7    8    9   10
 1   44    8    0    9    3    0    0    0    0    0
 2    8   46    9    0    5    4    0    0    0    0
 3    0    9   24    0    0    2    7    0    0    0
 4    9    0    0   32    6    0    0    0    4    0
 5    3    5    0    6   42    7    0    0    3    3
 6    0    4    2    0    7   42    9    3    0    5
 7    0    0    7    0    0    9   38    4    0    0
 8    0    0    0    0    0    3    4   28    0    5
 9    0    0    0    4    3    0    0    0   26    3
10    0    0    0    0    3    5    0    5    3   46

 Press Any Key to Compare Both Compressed Adjacency Matrices

 Both the Community Graphs are Similar on Dissimilar Edges
```

Fig. 7. Comparison result of CG_1 and CG_2

B. Comparison of CG_1 and CG_3

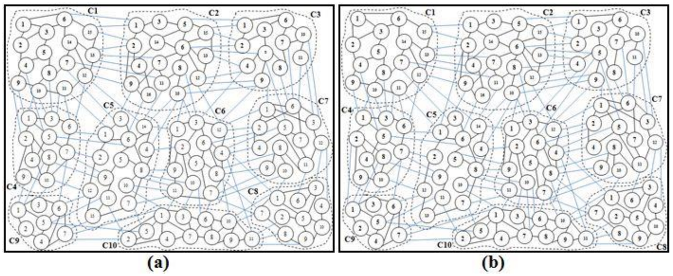
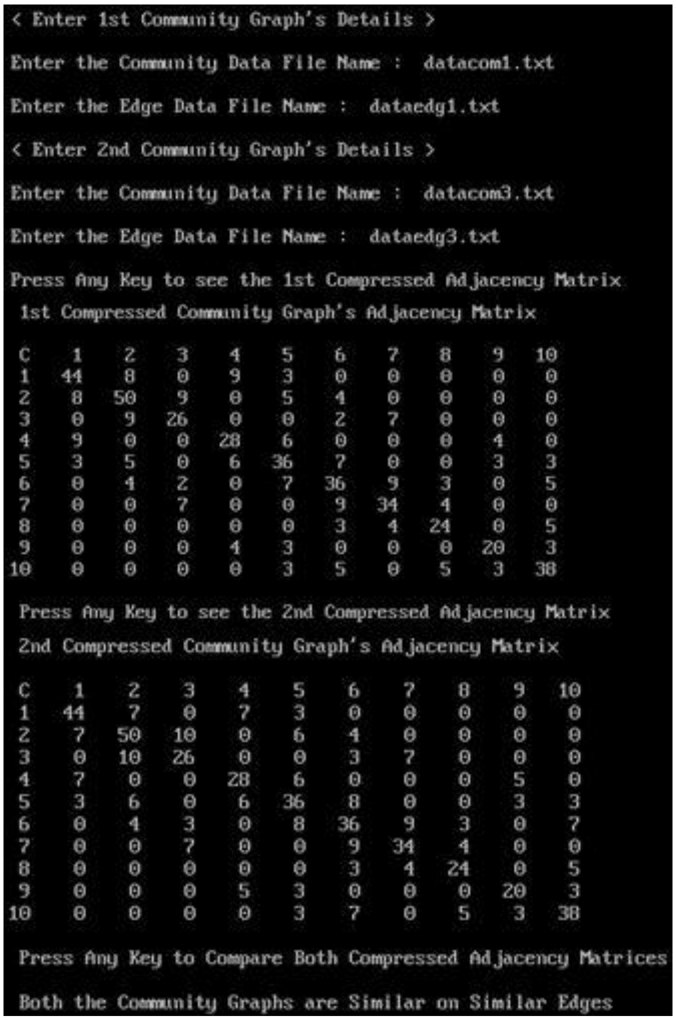


Fig. 8. (a) Community Graph CG_1 (b) Community Graph CG_3



```
< Enter 1st Community Graph's Details >

Enter the Community Data File Name :  datacom1.txt

Enter the Edge Data File Name :  dataedg1.txt

< Enter 2nd Community Graph's Details >

Enter the Community Data File Name :  datacom3.txt

Enter the Edge Data File Name :  dataedg3.txt

Press Any Key to see the 1st Compressed Adjacency Matrix
 1st Compressed Community Graph's Adjacency Matrix

 C    1    2    3    4    5    6    7    8    9   10
 1   44    8    0    9    3    0    0    0    0    0
 2    8   50    9    0    5    4    0    0    0    0
 3    0    9   26    0    0    2    7    0    0    0
 4    9    0    0   28    6    0    0    0    4    0
 5    3    5    0    6   36    7    0    0    3    3
 6    0    4    2    0    7   36    9    3    0    5
 7    0    0    7    0    0    9   34    4    0    0
 8    0    0    0    0    0    3    4   24    0    5
 9    0    0    0    4    3    0    0    0   20    3
10    0    0    0    0    3    5    0    5    3   38

 Press Any Key to see the 2nd Compressed Adjacency Matrix

 2nd Compressed Community Graph's Adjacency Matrix

 C    1    2    3    4    5    6    7    8    9   10
 1   44    7    0    7    3    0    0    0    0    0
 2    7   50   10    0    6    4    0    0    0    0
 3    0   10   26    0    0    3    7    0    0    0
 4    7    0    0   28    6    0    0    0    5    0
 5    3    6    0    6   36    8    0    0    3    3
 6    0    4    3    0    8   36    9    3    0    7
 7    0    0    7    0    0    9   34    4    0    0
 8    0    0    0    0    0    3    4   24    0    5
 9    0    0    0    5    3    0    0    0   20    3
10    0    0    0    0    3    7    0    5    3   38

 Press Any Key to Compare Both Compressed Adjacency Matrices

 Both the Community Graphs are Similar on Similar Edges
```

Fig. 9. Comparison result of CG_1 and CG_3

In community graph CG_1, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:9, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

In community graph CG_3, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:7, C_1-C_4:7, C_1-C_5:3, C_2-C_3:10, C_2-C_5:6, C_2-C_6:4, C_3-C_6:3, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:8, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:7, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

The comparison takes place on community graph CG_1 and CG_3's community codes and the number of edges belonging to similar community codes member since these two are same. So finally the algorithm shows as "*Both the Community Graphs are Similar on Similar Edges*".

C. Comparison of CG_1 and CG_4

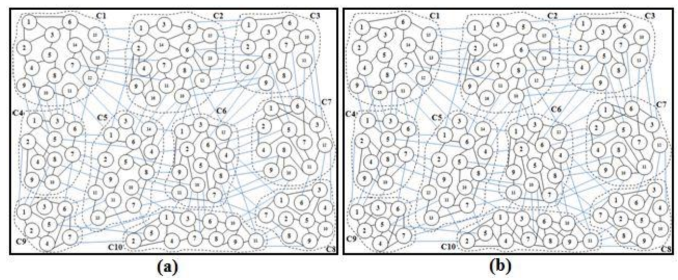
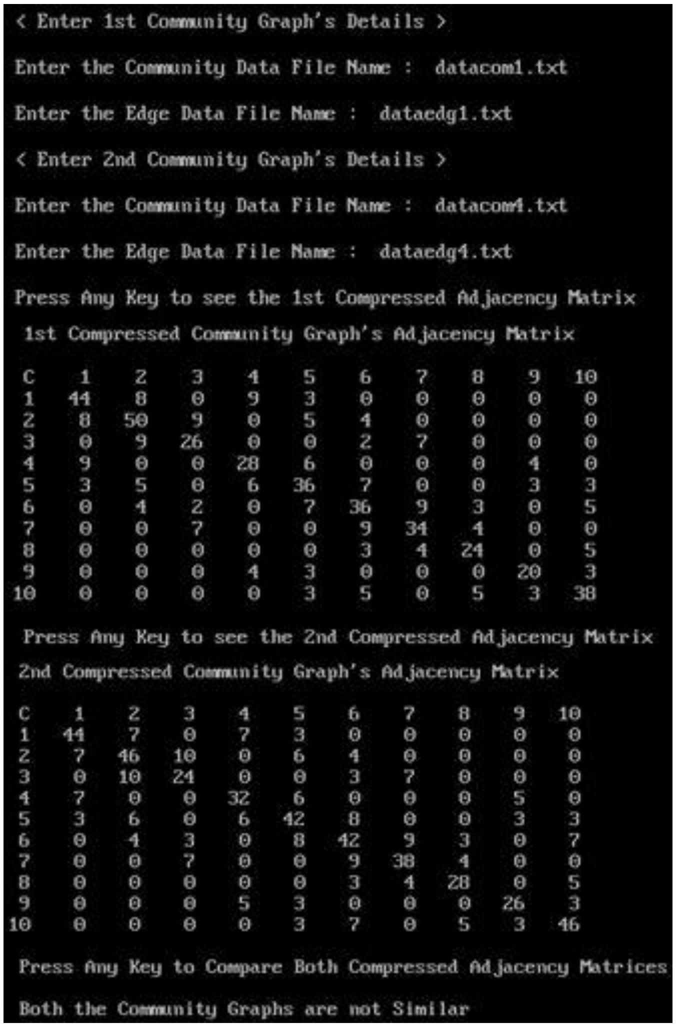


Fig. 10. (a) Community Graph CG_1 (b) Community Graph CG_4

In community graph CG_1, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:9, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

In community graph CG_4, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 46, 24, 32, 42, 42, 38, 28, 26, 46}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:7, C_1-C_4:7, C_1-C_5:3, C_2-C_3:10, C_2-C_5:6, C_2-C_6:4, C_3-C_6:3, C_3-C_7:7, C_4-C_5:6, C_4-C_9:5, C_5-C_6:8, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:7, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

The comparison takes place on community graph CG_1 and CG_4's number of edges belonging to similar community codes member and the number of edges belonging to dissimilar community codes member since these two are not same. So finally the algorithm shows as "*Both the Community Graphs are not Similar*".



< Enter 1st Community Graph's Details >

Enter the Community Data File Name : datacom1.txt

Enter the Edge Data File Name : dataedg1.txt

< Enter 2nd Community Graph's Details >

Enter the Community Data File Name : datacom4.txt

Enter the Edge Data File Name : dataedg4.txt

Press Any Key to see the 1st Compressed Adjacency Matrix

1st Compressed Community Graph's Adjacency Matrix

C	1	2	3	4	5	6	7	8	9	10
1	44	8	0	9	3	0	0	0	0	0
2	8	50	9	0	5	4	0	0	0	0
3	0	9	26	0	0	2	7	0	0	0
4	9	0	0	28	6	0	0	0	4	0
5	3	5	0	6	36	7	0	0	3	3
6	0	4	2	0	7	36	9	3	0	5
7	0	0	7	0	0	9	34	4	0	0
8	0	0	0	0	0	3	4	24	0	5
9	0	0	0	4	3	0	0	0	20	3
10	0	0	0	0	3	5	0	5	3	38

Press Any Key to see the 2nd Compressed Adjacency Matrix

2nd Compressed Community Graph's Adjacency Matrix

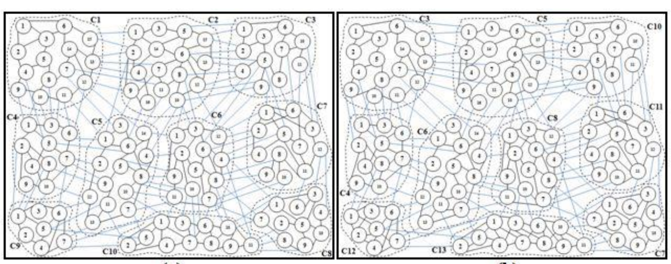
C	1	2	3	4	5	6	7	8	9	10
1	44	7	0	7	3	0	0	0	0	0
2	7	46	10	0	6	4	0	0	0	0
3	0	10	24	0	0	3	7	0	0	0
4	7	0	0	32	6	0	0	0	5	0
5	3	6	0	6	42	8	0	0	3	3
6	0	4	3	0	8	42	9	3	0	7
7	0	0	7	0	0	9	38	4	0	0
8	0	0	0	0	0	3	4	28	0	5
9	0	0	0	5	3	0	0	0	26	3
10	0	0	0	0	3	7	0	5	3	46

Press Any Key to Compare Both Compressed Adjacency Matrices

Both the Community Graphs are not Similar

Fig. 11. Comparison result of CG_1 and CG_4

D. Comparison of CG_1 and CG_5



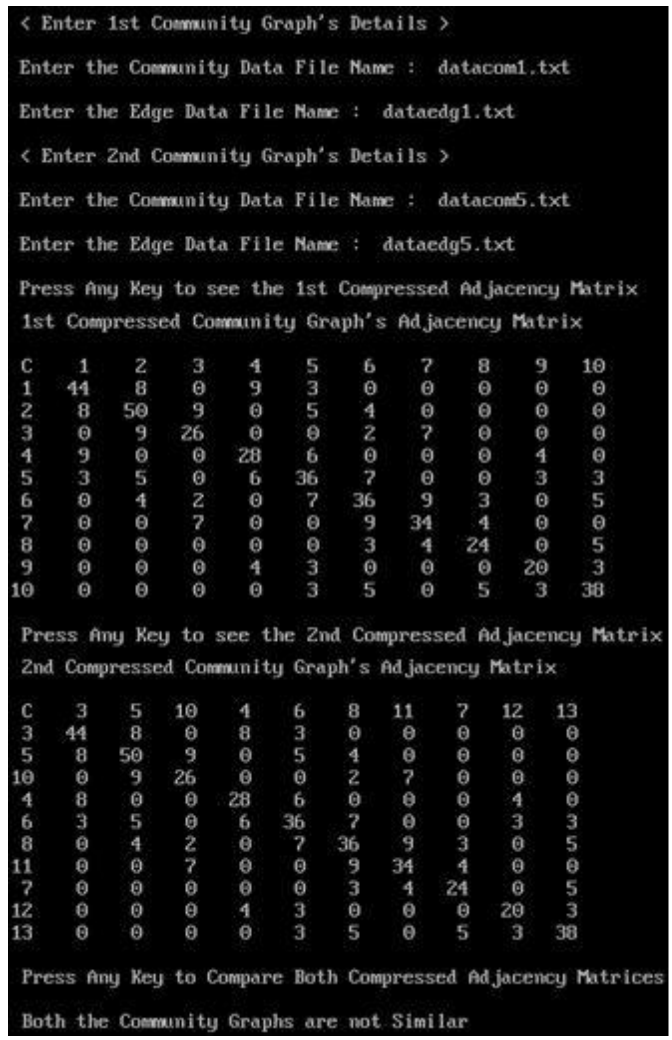
(a) (b)

Fig. 12. (a) Community Graph CG_1 (b) Community Graph CG_5

In community graph CG1, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:9, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

In community graph CG_5, the community codes (numbers) are {C_3, C_5, C_{10}, C_4, C_6, C_8, C_{11}, C_7, C_{12}, C_{13}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_3-C_5:8, C_3-C_4:8, C_3-C_6:3, C_5-C_{10}:9, C_5-C_6:5, C_5-C_8:4, C_{10}-C_8:2, C_{10}-C_{11}:2, C_4-C_6:6, C_4-C_{12}:4, C_6-C_8:7, C_4-C_{12}:3, C_6-C_{13}:3, C_8-C_{11}:9, C_8-C_7:3, C_8-C_{13}:5, C_{11}-C_7:4, C_7-C_{13}:5, and C_{12}-C_{13}:3.

The comparison takes place on community graph CG_1 and CG_5's community codes. Since the community codes of community graphs CG_1 and CG_5 are not same. So the algorithm shows as "*Both the Community Graphs are not Similar*".



< Enter 1st Community Graph's Details >

Enter the Community Data File Name : datacom1.txt

Enter the Edge Data File Name : dataedg1.txt

< Enter 2nd Community Graph's Details >

Enter the Community Data File Name : datacom5.txt

Enter the Edge Data File Name : dataedg5.txt

Press Any Key to see the 1st Compressed Adjacency Matrix

1st Compressed Community Graph's Adjacency Matrix

C	1	2	3	4	5	6	7	8	9	10
1	44	8	0	9	3	0	0	0	0	0
2	8	50	9	0	5	4	0	0	0	0
3	0	9	26	0	0	2	7	0	0	0
4	9	0	0	28	6	0	0	0	4	0
5	3	5	0	6	36	7	0	0	3	3
6	0	4	2	0	7	36	9	3	0	5
7	0	0	7	0	0	9	34	4	0	0
8	0	0	0	0	0	3	4	24	0	5
9	0	0	0	4	3	0	0	0	20	3
10	0	0	0	0	3	5	0	5	3	38

Press Any Key to see the 2nd Compressed Adjacency Matrix

2nd Compressed Community Graph's Adjacency Matrix

C	3	5	10	4	6	8	11	7	12	13
3	44	8	0	8	3	0	0	0	0	0
5	8	50	9	0	5	4	0	0	0	0
10	0	9	26	0	0	2	7	0	0	0
4	8	0	0	28	6	0	0	0	4	0
6	3	5	0	6	36	7	0	0	3	3
8	0	4	2	0	7	36	9	3	0	5
11	0	0	7	0	0	9	34	4	0	0
7	0	0	0	0	0	3	4	24	0	5
12	0	0	0	4	3	0	0	0	20	3
13	0	0	0	0	3	5	0	5	3	38

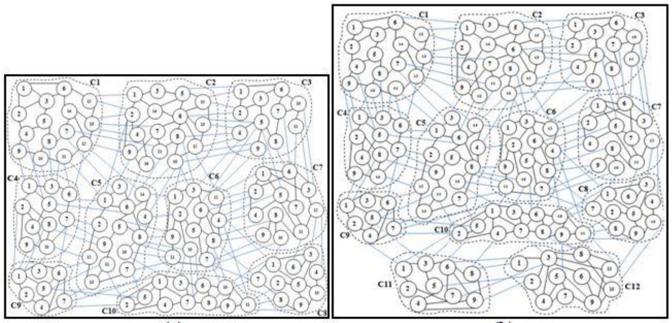
Press Any Key to Compare Both Compressed Adjacency Matrices

Both the Community Graphs are not Similar

Fig. 13. Comparison result of CG_1 and CG_5

E. Comparison of CG_1 and CG_6

In community graph CG_1, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:9, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.



(a) (b)

Fig. 14. (a) Community Graph CG_1 (b) Community Graph CG_6

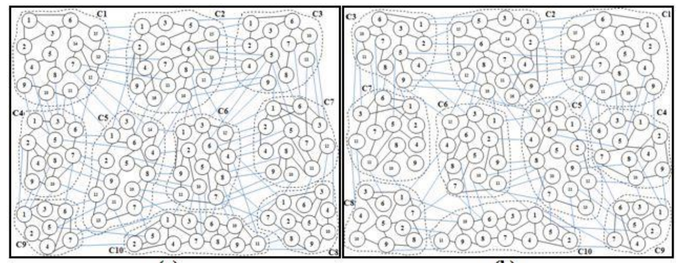
```
< Enter 1st Community Graph's Details >

Enter the Community Data File Name :  datacom1.txt

Enter the Edge Data File Name :  dataedg1.txt

< Enter 2nd Community Graph's Details >

Enter the Community Data File Name :  datacom6.txt

Enter the Edge Data File Name :  dataedg6.txt

Press Any Key to see the 1st Compressed Adjacency Matrix
 1st Compressed Community Graph's Adjacency Matrix

C    1    2    3    4    5    6    7    8    9   10
1   44    8    0    9    3    0    0    0    0    0
2    8   50    9    0    5    4    0    0    0    0
3    0    9   26    0    0    2    7    0    0    0
4    9    0    0   28    6    0    0    0    4    0
5    3    5    0    6   36    7    0    0    3    3
6    0    4    2    0    7   36    9    3    0    5
7    0    0    7    0    0    9   34    4    0    0
8    0    0    0    0    0    3    4   24    0    5
9    0    0    0    4    3    0    0    0   20    3
10   0    0    0    0    3    5    0    5    3   38

 Press Any Key to see the 2nd Compressed Adjacency Matrix
 2nd Compressed Community Graph's Adjacency Matrix

C    1    2    3    4    5    6    7    8    9   10   11   12
1   44    8    0    8    3    0    0    0    0    0    0    0
2    8   50    9    0    5    4    0    0    0    0    0    0
3    0    9   26    0    0    2    7    0    0    0    0    0
4    8    0    0   28    6    0    0    0    4    0    0    0
5    3    5    0    6   36    7    0    0    3    3    0    0
6    0    4    2    0    7   36    9    3    0    5    0    0
7    0    0    7    0    0    9   34    4    0    0    0    0
8    0    0    0    0    0    3    4   24    0    5    0    2
9    0    0    0    4    3    0    0    0   20    3    2    0
10   0    0    0    0    3    5    0    5    3   38    4    4
11   0    0    0    0    0    0    0    0    2    4   26    5
12   0    0    0    0    0    0    0    2    0    4    5   38

 Press Any Key to Compare Both Compressed Adjacency Matrices

 Both the Community Graphs are not Similar
```

Fig. 15. Comparison result of CG_1 and CG_6

In community graph CG_6, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}, C_{12}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11, 9, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38, 26, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:8, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, C_8-C_{12}:2, C_9-C_{10}:3, C_9-C_{11}:2, C_{10}-C_{11}:4, C_{10}-C_{12}:4, and C_{11}-C_{12}:5.

The comparison takes place on community graph CG_1 and CG_6's community codes. Since the number of community codes of community graphs CG_1 and CG_6 are not same. So the algorithm shows as "*Both the Community Graphs are not Similar*".

F. Comparison of CG_1 and CG_7



(a) (b)

Fig. 16. (a) Community Graph CG_1 (b) Community Graph CG_7

```
< Enter 1st Community Graph's Details >

Enter the Community Data File Name :  datacom1.txt

Enter the Edge Data File Name :  dataedg1.txt

< Enter 2nd Community Graph's Details >

Enter the Community Data File Name :  datacom7.txt

Enter the Edge Data File Name :  dataedg7.txt

Press Any Key to see the 1st Compressed Adjacency Matrix
 1st Compressed Community Graph's Adjacency Matrix

C    1    2    3    4    5    6    7    8    9   10
1   44    8    0    9    3    0    0    0    0    0
2    8   50    9    0    5    4    0    0    0    0
3    0    9   26    0    0    2    7    0    0    0
4    9    0    0   28    6    0    0    0    4    0
5    3    5    0    6   36    7    0    0    3    3
6    0    4    2    0    7   36    9    3    0    5
7    0    0    7    0    0    9   34    4    0    0
8    0    0    0    0    0    3    4   24    0    5
9    0    0    0    4    3    0    0    0   20    3
10   0    0    0    0    3    5    0    5    3   38

 Press Any Key to see the 2nd Compressed Adjacency Matrix
 2nd Compressed Community Graph's Adjacency Matrix

C    3    2    1    7    6    5    4    8   10    9
3   26    9    0    7    2    0    0    0    0    0
2    9   50    8    0    4    5    0    0    0    0
1    0    8   44    0    0    3    9    0    0    0
7    7    0    0   34    9    0    0    4    0    0
6    2    4    0    9   36    7    0    3    5    0
5    0    5    3    0    7   36    6    0    3    3
4    0    0    9    0    0    6   28    0    0    4
8    0    0    0    4    3    0    0   24    5    0
10   0    0    0    0    5    3    0    5   38    3
9    0    0    0    0    0    3    4    0    3   20

 Press Any Key to Compare Both Compressed Adjacency Matrices

 Both the Community Graphs are Similar
```

Fig. 17. Comparison result of CG_1 and CG_7

In community graph CG_1, the community codes (numbers) are {C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}} with total community members are {15, 16, 11, 10, 14, 12, 12, 10, 7, 11}. The total number of edges belonging to same community codes member are {44, 50, 26, 28, 36, 36, 34, 24, 20, 38}. Similarly, the total number of edges belonging to dissimilar community codes member are C_1-C_2:8, C_1-C_4:9, C_1-C_5:3, C_2-C_3:9, C_2-C_5:5, C_2-C_6:4, C_3-C_6:2, C_3-C_7:7, C_4-C_5:6, C_4-C_9:4, C_5-C_6:7, C_5-C_9:3, C_5-C_{10}:3, C_6-C_7:9, C_6-C_8:3, C_6-C_{10}:5, C_7-C_8:4, C_8-C_{10}:5, and C_9-C_{10}:3.

In community graph CG_7, the community codes (numbers) are {C_3, C_2, C_1, C_7, C_6, C_5, C_4, C_8, C_{10}, C_9} with total community members are {11, 16, 15, 12, 12, 14, 10, 10, 11, 7}. The total number of edges belonging to same community codes member are {26, 50, 44, 34, 36, 36, 28, 24, 38, 20}. Similarly, the total number of edges belonging to dissimilar community codes member are C_3-C_2:9, C_3-C_7:7, C_3-C_6:2, C_2-C_1:8, C_2-C_6:4, C_2-C_5:5, C_1-C_5:3, C_1-C_4:9, C_7-C_6:9, C_7-C_8:4, C_6-C_5:7, C_6-C_8:3, C_6-C_{10}:5, C_5-C_4:6, C_5-C_{10}:3, C_5-C_9:3, C_4-C_9:4, C_8-C_{10}:5, and C_{10}-C_9:3.

The comparison takes place on community graph CG_1 and CG_7's community codes. Since the community codes of community graphs CG_1 and CG_7 are same. Then the comparison takes place on a number of edges belonging to similar community codes member and number of edges belonging to dissimilar community codes member. So finally the algorithm shows as "*Both the Community Graphs are Similar*".

VI. Conclusions

Graph similarity technique is helpful in the fields of shape retrieval, object recognition, face recognition and many more areas. This paper starts with literature survey related to various techniques implemented for graph similarity. So it is important to compare two community graphs for similarity check to extract the reliable knowledge from a large community graph. This paper proposes an algorithm for similarity check of two community graphs using graph mining techniques. The authors have implemented the proposed algorithm using C++ programming language and obtained satisfactory results.

References

[1] Danai Koutra, Ankur Parikh, Aditya Ramdas, and Jing Xiang, "Algorithms for Graph Similarity and Subgraph Matching," Dec 4, 2011. https://www.cs.cmu.edu/~jingx/docs/DBreport.pdf

[2] B. Huet, A. Cross, and E. Hancock, "Shape retrieval by inexact graph matching." in Proc. IEEE Int. Conf. on Multimedia Computing Systems. Volume 2., IEEE Computer Society Press, 1999, pages 40–44.

[3] E. Kubicka, G. Kubicki, and I.Vakalis, "Using graph distance in object recognition." in Proc. ACM Computer Science Conference, 1990, pages 43–48.

[4] L. Wiskott, J. M. Fellous, N. Kr¨uger, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching." in IEEE PAMI 19, 1997, pages 775–779.

[5] V. Levenshtein. "Binary codes capable of correcting deletions, insertions and reversals," in Soviet Physics Doklady (10), 1966, 707–710.

[6] R. A. Wagner and M. J. Fisher, "The string-to-string correction problem," in Journal of the ACM (21), 1974, 168–173.

[7] A. Sanfeliu and K. S. Fu, "A distance measure between attributed relational graphs for pattern recognition," in IEEE Transactions on Systems, Man and Cybernetics (13), 1983, 353–362.

[8] D. J. Cook and L. B. Holder, "Graph-based data mining," in IEEE Intelligent Systems (15), 2000, 32–41.

[9] Duncan J. Watts, "Small worlds: the dynamics of networks between order and randomness", Princeton University Press, 1999.

[10] Sung-Hyuk Cha, "Comprehensive survey on distance/similarity measures between probability density functions," in International Journal of Mathematical Models and Methods in Applied Sciences, 1(4), 2007, 300–307.

[11] Sergey Melnik, Hector Garcia-Molina, and Erhard Rahm, "Similarity flooding: A versatile graph matching algorithm and its application to schema matching," in 18th International Conference on Data Engineering (ICDE 2002), 2002.

[12] Glen Jeh and Jennifer Widom, "SimRank: A measure of structural-context similarity," in Proceedings of the eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '02, New York, USA, 2002, pages 538–543.

[13] L. Zager and G. Verghese, "Graph similarity scoring and matching," in Applied Mathematics Letters, 21(1), 2008, 86–94.

[14] Hans-Peter Kriegel and Stefan Schonauer, "Similarity Search in Structured Data," in Proceedings 5th International Conference on Data Warehousing and Knowledge Discovery (DaWaK' 03), Prague, Czech Republic, 2003, pages 224-233.

[15] Jonathan S. Yedidia, William T. Freeman, and Yair Weiss, "Understanding belief propagation and its generalizations," in Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003, pages 239–269.

[16] Bapuji Rao, Anirban Mitra and D. P. Acharjya, "A New Approach of Compression of Large Community Graph Using Graph Mining Techniques," in Proceedings of 3rd ERCICA 2015, Volume 1, Springer Verlag, NMIT, Bangalore, India, Pp. 127 – 136, July 31–Aug 1, 2015. DOI: 10.1007/978-81-322-2550-8_13

[17] Horowitz, Sahani, and Rajasekaran, "Fundamentals of Computer Algorithms", Galgotia Publications Pvt. Ltd., 5, Ansari Road, Darya Ganj, New Delhi-110 002 © 1998 by W. H. Freeman and Company.

A New Approach for Enhancing the Quality of Medical Computerized Tomography Images

Mu'taz Al-Frejat
Department of Computer Science
Al-Balqa' Applied University
Al Salt, Jordan

Mohammad Hjouj Btoush
Department of Computer Science
Al-Balqa' Applied University
Al Salt, Jordan

Abstract—**Computerized tomography (CT) images contribute immensely to medical research and diagnosis. However, due to degradative factors such as noise, low contrast, and blurring, CT images tend to be a degraded representation of the actual body or part under investigation. To reduce the risk of imprecise diagnosis associated with poor-quality CT images, this paper presents a new technique designed to enhance the quality of medical CT images. The main objective is to improve the appearance of CT images in order to obtain better visual interpretation and analysis, which is expected to ease the diagnosis process. The proposed technique involves applying a median filter to remove noise from the CT images and then using a Laplacian filter to enhance the edges and the contrast in the images. Also, as CT images suffer from low contrast, a Contrast Limited Adaptive Histogram Equalization transform is also applied to solve this problem. The main strength of this transform is its modest computational requirements, ease of application, and excellent results for most images. According to a subjective assessment by a group of radiologists, the proposed technique resulted in excellent enhancement, including that of the contrast and the edges of medical CT images. From a medical perspective, the proposed technique was able to clarify the arteries, tissues, and lung nodules in the CT images. In addition, blurred nodules in chest CT images were enhanced effectively. Therefore the proposed technique can help radiologists to better detect lung nodules and can also assist in diagnosing the presence of tumours and in the detection of abnormal growths.**

Keywords—Spatial domain; CT image; Laplacian filter; MedPix database; Lung nodules

I. Introduction

Image processing is a method that is used to perform certain operations on an image to obtain an enhanced image and useful information from the image. It can be described as signal processing, where the input is an image and the output may be an image or characteristic/feature from that image [3]. Image enhancement, also known as filtering, is the most simple and beneficial step in image processing. Sometimes images taken by a satellite or a conventional/digital camera are not clear because of the limitations of the system taking the pictures or the illumination conditions when the picture is taken. In image enhancement, certain image features are accentuated for image analysis or display. However, image enhancement itself does not increase the information present in the image; rather, it enhances the required image characteristics [8]. Enhancement involves modifying the attributes of an image to make it more suitable than the original for a certain observer and a specific activity. Image enhancement mainly includes contrast and intensity manipulation, background removal, noise reduction, filtering, and edge enhancement [15]. In the medical domain, a computerized tomography (CT) scan image is a detailed digital representation or picture of cross-sections of the body created through the use of X-rays. The images are created by using a CT scanner that produces a beam of X-rays while rotating around the body as a computer generates separate images, known as slices, of the body area under examination. The images are then processed by the computer for storage, printing on a film or viewing on a monitor. Computerized tomography reveals a lot of information about the human body through the cross-sectional images it produces and this technique has contributed immensely to medical research and diagnosis. However, due to degradative factors such as noise, low contrast, and blurring, a CT image tends to be a degraded representation of the original body or part under investigation. These factors are caused by the use of a low radiation dose and inappropriate or poor enhancement or restoration algorithms. Also, noise leads to low-contrast details being present in CT images, which causes poor visibility and lack of clarity in the images [2].

To reduce the risk of imprecise diagnosis associated with poor-quality CT images, there is a need to improve the visual detail of CT images and thus aid medical practitioners such as radiologists in making more informed diagnoses by addressing the issues of contrast enhancement and feature enhancement. Image enhancement makes it easier to interpret the results through improving visualization. Several techniques are available to reduce the noise, enhance the edges, and improve the contrast in images [6]. The main objective of this paper is to improve the appearance of medical CT-scan images in order to obtain better visual interpretation and analysis, which is expected to ease the diagnostic process.

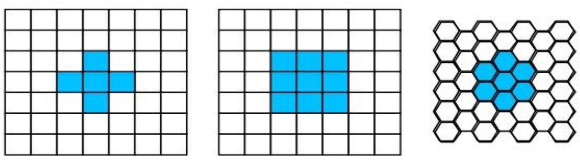
Many approaches can be employed to enhance a digital image without destroying it and they can be categorized into two main types: spatial domain and frequency domain methods. Spatial domain methods directly deal with the pixels of the image to obtain the desired enhancement, whereas frequency domain methods involve initially transferring the image into a frequency domain so that the image's Fourier transform is computed first. All the enhancement operations are done on the Fourier transform of the image, after which an Inverse Fourier transform is performed to obtain the intended final image [11].

The main operations that can be applied to process digital images are [20].

1) The point operation, where the result of processing a pixel depends only on its value.

2) The local operation, where the result of processing a pixel depends on the values of its neighbourhood. The neighbourhood is a finite set of pixels around the pixel being processing. Defining the neighbourhood is an important part of modern digital image processing and the most common neighbourhoods are 4-connected and 8-connected for rectangular sampling and 6-connected for hexagonal sampling (see Fig. 1).

3) The global operation, where the result of processing a specific coordinate depends on the value of all the coordinates in the input image.



(a) **(b)** **(c)**

Fig. 1. Representation of commonly used neighbourhoods: (a): 4-connected neighbourhood; (b): 8-connected neighbourhood; (c): 6-connected neighbourhood [20]

Image quality is the characteristic of an image that is based on the visible image degradation. Imaging systems introduce a quantity of artifacts or distortion to the images produced. The factors that affect image quality are [1]:

a) ***Blur:*** Sometimes the boundary of the organs or the lesion is sharp but the image is not sharp. This can occur due to geometric blur or lack of sharpness of the object or receptor.

b) ***Noise:*** Uncertainty or lack of precision in the recording of a signal can occur due to the presence of fewer photons, of grains in the radiographic film, intensifying screens with large grains, and the electronic noise of the detector or amplifier.

c) ***Contrast****:* Image contrast refers to the minute differences in the brightness between two regions in an image, which can be affected by tissue thickness and density, electron density in the tissue, X-ray energy (kV), X-ray spectrum, scatter rejection, windowing level of CT and digital subtraction angiography, film characteristics, and screen characteristics.

d) ***Distortion and artifacts:*** The inaccurate recording of the real size, shape, and relative positions of the items in an image may be due to the visualization of grid on the film, dust on the screen, bad contact of the film with the screen, and bad positioning of the patient. The motion artifact appears when there is patient motion during the acquisition of the CT image, where even a small movement will cause blurring of the image.

II. LITERATURE REVIEW

Image enhancement is an important aspect of medical image processing because it can assist in correct diagnosis and treatment planning. Numerous articles have been published on this topic and some of the works most relevant to the technique proposed in this paper are discussed here.

Bhardwaj and Singh [5] address the problem of the low contrast in and poor quality of medical images by proposing a process to enhance medical images based on the Haar transform, soft thresholding, and a nonlinear approach for contrast enhancement. The problem with using the Haar wavelet transform lies in finding a suitable way to extract high-frequency information and to decompose high-frequency sub-images of wavelets. However, the approach proposed by the authors helps in the effective extraction of high-frequency information because the experimental results showed that the proposed algorithm not only improves image contrast, but also preserves the original image's edge property. The final result is good visibility, which is essential for accurate medical diagnosis.

Rajput et al [17] propose the use of nonlinear enhancement techniques such as the discrete wavelet transform and histogram equalization to improve the contrast in medical images. These techniques together improve the lower and the higher contrast areas of an image in both the spatial and frequency domain. Different parameters such as the mean square error, peak signal, and noise ratio are used to evaluate the effectiveness of the proposed method. The results indicated that the peak signal is a better parameter for the pre processed histogram technique than the mean square error or noise ratio. The results also proved that the wavelet transform is better than the simple histogram technique. The discrete wavelet transform enhances the image edges, which implies that it can mainly be used for de-noising images. Based on the experiments performed, the algorithm proposed by the authors can be used to enhance contrast in the image while preserving the quality of the image.

Ritika [16] proposes a novel approach to enhance the contrast of medical images with the help of morphology. The method uses a multiscale structuring element, where white and black (bright and dark) regions are extracted for various scales of the image that correspond to different scales of the structuring element. Then these features are combined with the original image to produce a final image with enhanced contrast. The author's proposed algorithm was implemented and executed on a set of greyscale medical images including a synthetic image obtained through the Monte Carlo simulation method. A comparison of the results of the proposed approach with those obtained from existing methods for contrast enhancement showed that the proposed approach gives better results and makes the features of the images clearer, but there is a slight increase in noise. However, this drawback could be reduced through further research and the algorithm could be further extended through the use of non-flat structuring elements.

Hirani and Totsuka [7] developed a hybrid technique based on the concept of Projection on Convex Sets (POCS) to reduce noise in images caused by a mixture of types of noise. The approach works by searching for noisy pixels so as to replace them with pixels from the neighbourhood. This is because the algorithm can be able to handle images that feature varying intensities. The method was shown to be quite useful in the reconstruction of images that show repeating patters. However, this method showed challenges in that the proposed technique is that the contents of repair images and the sample have to be approximately translated versions of each other.

III. Materials and Methods

The technique proposed in this paper consists of three main steps:

1) A median filter for noise reduction:

2) A Laplacian filter under an eight neighbour (-8 in centre) mask for edge enhancement; and

3) A Contrast Limited Adaptive Histogram Equalization (CLAHE) transform for contrast enhancement.

The proposed technique was implemented using Matlab language. See Appendix A

A. Median Filter

The median represents a partitioning value between the higher weights and lower weights. It is categorized as an order-statistic filter. The filter uses the median value for each pixel by using a sliding kernel around it. If the kernel is square, then there is an odd number of cells in the kernel and picking the median value is easy. If the kernel is not square, then the mean value from the two contesting cells is used [10].

In image processing, the median filter is used extensively to replace unwanted pixel values with a more suitable median value from the surroundings. It is a nonlinear filtering technique that is often used to reduce noise. This noise reduction technique is often used as a preprocessing step to improve the results of image processing. The median filter often gives better results than a mean or average filter. Moreover, it can be used to remove Gaussian noise as well as pepper and salt noise [9].

Like the mean filter, an n x n filter is used to replace the pixel value with its median according to the values of its neighbouring pixels. Instead of replacing the pixel value with its mean, it replaces the central value with the median value. Filtering is applied to each pixel of the image separately with help of an n × n filter kernel. To calculate the median value, pixel values are sorted in ascending order to find the number that has an equal number of values before and after it. The central pixel value is then replaced by the calculated median value [13]. Figure 2 provides an example of the application of the median filter on a pixel using a 3 x 3 kernel size. From the figure,

The pixel values in ascending order are 6,9,11,15,19,20,24,32,34 and the median= 19.

11	15	9
32	6	34
19	24	20

Fig. 2. Part of image source [13]

The median filter is comparatively more robust than the mean filter. In addition, the median filter is more accurate because only one neighbour participates in the final assignment of the central pixel value. As the median value is only one value from the neighbours, it is not affected by the other neighbours at all. This filter preserves edges while removing noise [13].

B. Laplacian Filter (using second-order derivatives in image enhancement)

The use of second derivatives for enhancement involves defining a discrete formulation of the second-order derivative and then developing a filter mask using that formulation [21] [19]. In this work, the interest is isotropic filters whose response does not depend on the direction of the discontinuities within the image to which the filter is applied. Isotropic filters are described as rotational invariant. According to [6], the simplest isotropic derivative operator is the Laplacian, which is defined (for a function f(x, y) of two variables) as [6].

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad \textbf{(1)}$$

The Laplacian is a linear operator because the derivatives of any order are linear operations. Equation **(1)** has to be expressed in discrete form for use in digital image processing. There are several ways to define the digital Laplacian employing neighbourhoods. The definition must satisfy the attributes of the second derivative. As there are two variables, the notation for the partial second-order derivative in the x-direction is employed [6], [21] [19] as follows:

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y) \quad \textbf{(2)}$$

And likewise in the y-direction as:

$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y) \quad \textbf{(3)}$$

The digital adoption of the two-dimensional Laplacian in Eq. (**1**) is obtained by summing the two parts [21]:

$$\nabla^2 f = \left[f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) \right] - 4f(x, y) \quad \textbf{(4)}$$

Equation (**4**) can be adopted at all points (x, y) in an image through convolving the image with the following mask, which provides an isotropic outcome for rotations in increments of 90 degrees [21].

0	1	0
1	−4	1
0	1	0

Fig. 3. Filter mask used to implement Laplacian [6]

In the definition of the digital Laplacian, the diagonal directions can be integrated by adding two terms to Eq. (**4**); one term for each of the two diagonal directions. The form of each new term is identical to either Eq. (**3**) or (**2**). However, the coordinates are along the diagonals and because each diagonal term also comprises a -2f(x, y) term, the total deducted from the difference term would be -8f(x, y). Figure4 demonstrates the mask employed to adopt this new definition [19].

1	1	1
1	−8	1
1	1	1

Fig. 4. Diagonal Laplacian mask [6]

The Laplacian filter is a linear spatial filter that is used to enhance the quality of grey-scale images. This filter is very suitable for reducing noise in an image. In addition, it enhances the linear features of images [21]. The Laplacian is used for image enhancement as follows [22] [19]:

$$g(x,y)=\begin{cases} f(x,y)-\nabla^2 f(x,y) & \text{If the centre coefficient of the Laplacian mask is negative} \\ f(x,y)+\nabla^2 f(x,y) & \text{If the centre coefficient of the Laplacian mask is positive.} \end{cases} \quad \textbf{(5)}$$

Figure5 (a) shows a non-contrast axial CT-scan image of a lung with multiple parenchymal nodules. Figures5(b)and 5(c) depict the effects of filtering with a Laplacian mask [0 1 0; 1 -4 1; 0 1] and a diagonal Laplacian mask [1 1 1; 1 -8 1; 1 1 1;],respectively.

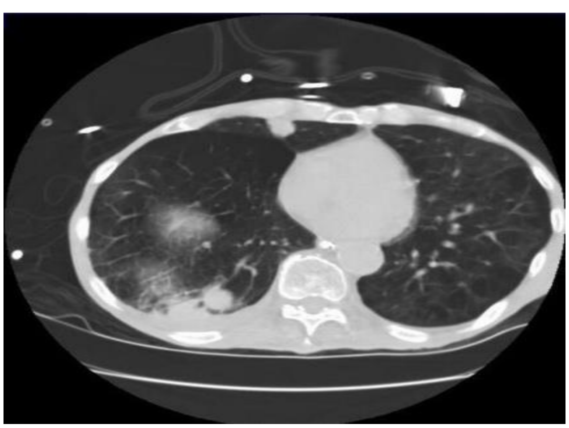


Fig. 5. (a): Lung CT-scan image [12]

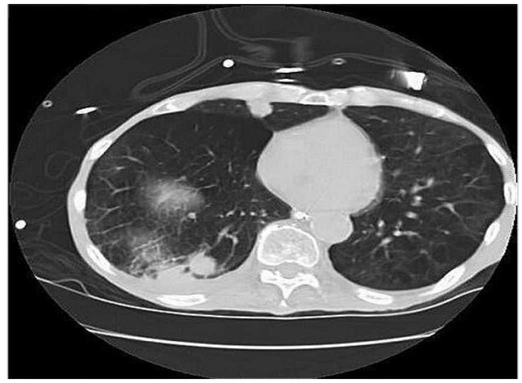


Fig. 5. (b): Enhancement obtained by four-neighbour Laplacian mask (-4 in the centre).

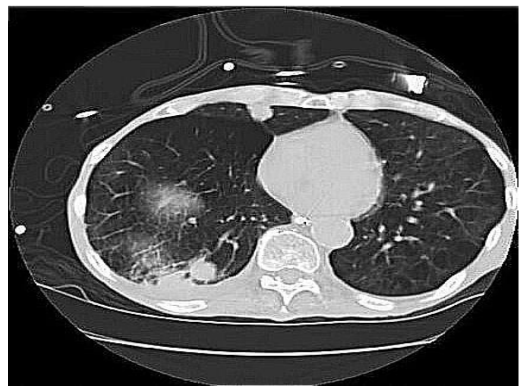


Fig. 5. (c): Enhancement obtained by diagonal Laplacian mask (-8 in the centre)

C. Contrast Limited Adaptive Histogram Equalization

Unlike the Adaptive Histogram Equalization (AHE) method, the CLAHE method can be applied to small regions of data known as tiles rather than to the whole image in order to enhance the local contrast in an image. In CLAHE, the eventual neighbouring tiles are then combined by using bilinear interpolation. Also, to avoid noise amplification, the contrast in the homogeneous part can be limited. These specific parts can be characterized by a high peak in the histogram linked with the contextual parts because many pixels fall inside the same grey range. By adopting CLAHE, the slope that comes along with the grey-level assignment scheme is highly limited. This can be possible by only allowing a maximum number of pixels in each particular bin associated with the local histogram [23].

The clip limit is defined as a multiple of the average histogram contents. With a low clip limit, the maximum slope of the local histogram will be low and therefore lead to limited contrast enhancement. A factor of one hinders contrast enhancement; redistribution of histogram bin values can be avoided by using a very high clip limit, which is equivalent to the AHE method. The main strength of the CLAHE transform is its modest computational requirements, ease of application, and excellent results for most images [23].

Figure 6(a) shows an original CT-scan image. Figures 6(b)-(d) depict the effect of CLAHE on the image under different clip limits.

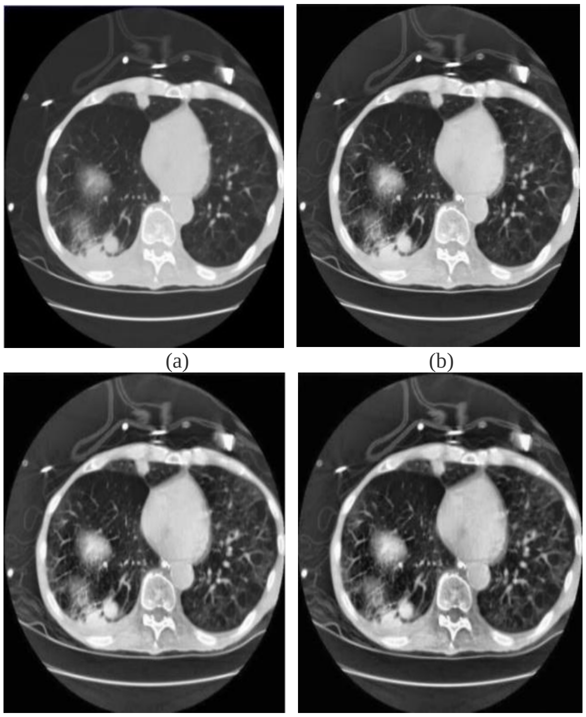


Fig. 6. The original CT scan image and the results of applying CLAHE Transform **(a):** The original CT scan image; **(b)**: image at clip limit =0.002; **(c):** image at 0.005; **(d):** image at 0.008

IV. RESLUTS

This work is focused on enhancing the appearance of the nodules and tumours present in medical CT-scan images of the chest and lungs of human patients. Some results of applying the proposed technique are illustrated in Figures 7-14 below.

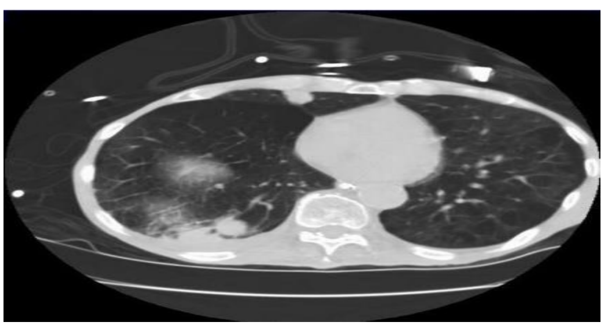


Fig. 7. Lung CT-scan image [12] (Test image #1)

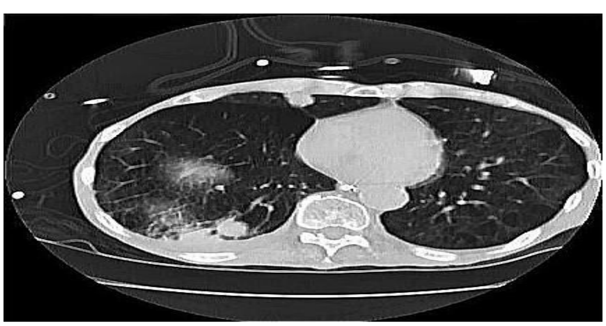


Fig. 8. Enhancement obtained by applying proposed technique to Fig. 7.

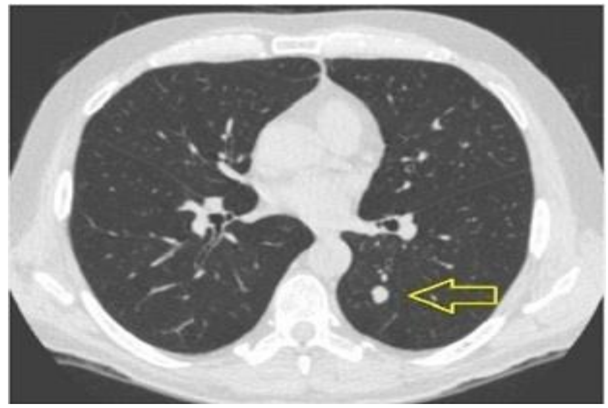


Fig. 9. Lung CT-scan image [12] (Test image #2)

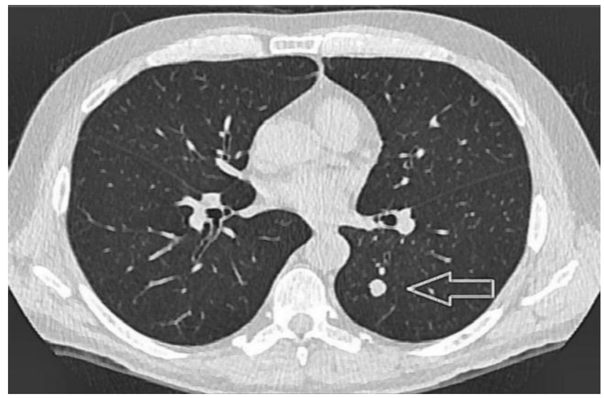


Fig. 10. Enhancement obtained by applying proposed technique to Fig. 9

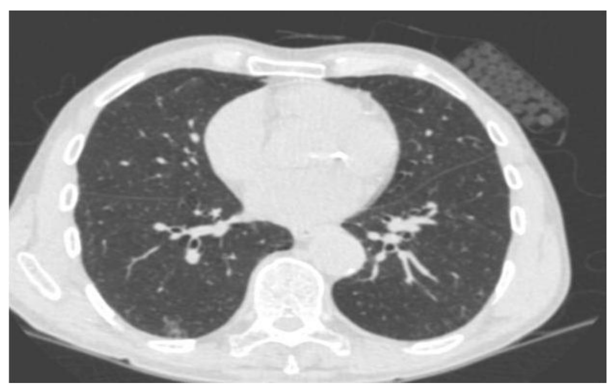


Fig. 11. Lung CT-scan image [12] (Test image #3)

Fig. 12. Enhancement obtained by applying proposed technique to Fig. 11.

Fig. 13. Lung CT-scan image [12] (Test image #4)

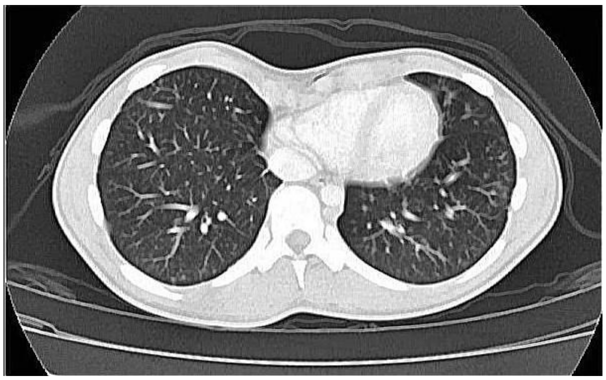


Fig. 14. Enhancement obtained by applying proposed technique to Fig. 13.

V. Discussion

Image enhancement is indisputably crucial in medical imaging. It increases contrast, gets rid of noise and blurring, and improves the quality of photographs for human viewers. To enhance medical CT-scan images, the first step of the proposed algorithm addresses the issue of noise reduction by applying a median filter to remove high-frequency components such as salt and pepper noise, Gaussian noise, and impulsive noise. As shown by the figures 8, 10, 12, and 14 in the Results section, the median filter was able to effectively remove above mentioned types of noise medical CT-scan images to improve image clarity.

The second step of the proposed algorithm is the application of a Laplacian filter under a diagonal (-8 in the centre) mask, which helps in enhancing the edges and the contrast in the CT-scan images. This step is necessary to clarify the details of the image. Edge enhancement by Laplacian filter is useful in the detection and extraction of objects in images. Moreover, edge enhancement with this filter helps in detecting nodules and tumours in CT-scan images. Figures 8, 10, 12, and 14 above show that filtering by a Laplacian filter under an 8-neighbour (-8 in centre) mask was able to enhance the edges and contrast effectively. In addition, the details in these images are clearer and sharper. The lung nodules and their surrounding tissue have been enhanced and sharpened effectively, which makes their detection through human visual perception easier.

The third and final step of the proposed algorithm is the use of the CLAHE technique in order to enhance the contrast and to clarify the details in the CT-scan images. In medical applications, contrast enhancement is very important as it improves the perceptibility of the cross-sectional images of body. Moreover, this process increases the brightness variance between the body and its related background. This allows for an accurate analysis of low-contrast CT images. From Figures 8, 10, and 12 above, it can be seen that application of the CLAHE transform was able to enhance the appearance of the tissues and the blood vessels. In addition, new details have appeared.

Thus, it is apparent from the figures in the Results section that the proposed algorithm enhances medical CT-scan images. However, it is also important to seek the opinion of experts regarding the effectiveness of the proposed technique from the medical perspective. While image restoration methods are assessed objectively through the use of mathematical models to test image degradation, when comparing and evaluating the performance of image enhancement techniques it is necessary to judge the visual quality of the enhanced images [18] as perceived by the human eye.

In the context of this work, doctors in hospitals are the best category of human viewer to check if the proposed technique gives a better result to aid them in making the correct diagnosis by comparing the enhanced images with the original images [4]. As this work is targeted at assisting radiologists, they were chosen as image quality assessors. All the enhanced images together with the original images were presented to a group of five specialist radiologists to measure the enhancement percentage. The experts presented percentage ratio for enhancement of images which is 85% from the original images.

VI. Conclusion and Fututre Work

One of the biggest challenges in digital image processing, particularly in computerized tomography, is the occurrence of degradation due to several factors such as noise, blurring, and low contrast, which are associated with various real-world limitations and which corrupt the quality of images. It is important therefore to subject CT images to an image enhancement process to improve their clarity before they are used in the diagnostic process.

The principal goal of image enhancement involves modifying the attributes of an image to make it more suitable than the original image for a certain observer and a specific activity. Image enhancement encompasses manipulation of contrast and intensity, background removal, reduction of noise, filtering, and sharpening of edges to improve quality.

The results presented in this paper show that the proposed technique can enhance medical CT-scan images effectively, and this finding is supported by the results of a subjective assessment by a group of medical experts. Enhancements were made to the contrast and the edges of a range of medical CT-scan images. From the medical perspective, the proposed technique clarified the arteries, tissues, and nodules. In addition, blurred nodules were enhanced effectively. Thus the proposed technique can help radiologists in the detection of lung nodules as well as assist in diagnosing the presence of tumours and in the detection of abnormal growths. Future work will focus on design a new technique to enhance the medical image quality.

ACKNOWLEDGMENT

This paper would not have been produced without Allah's aid. The authors would like to thank the radiologists and the staff members at Prince Hamza and Al-Isra hospitals for facilitating part of the assessment. Special thanks also to Doctor Kifah Adib Al-Hafez for his efforts and help.

REFERENCES

[1] Accra, 2011. Image quality and patient dose. Available at: http://documents.mx/documents/lecture-05-image quality-and-patient-dose.html. [Accessed 2015-10-5].

[2] A. Zohair, A.Shamil, and S. Ghazali, "Latest Methods of Image Enhancement and Restoration for Computed Tomography: A Concise Review," *Applied Medical Informatics*, Vol. 36, Issue 1, pp. 1-12, (2015).

[3] I. A. Ansari and R. Y. Borse, "Image Processing and Analysis ", International Journal of Engineering Research and Applications, Vol.3, issue.4, pages.1655-1658,ISSN.2248-9622,(2013).

[4] P. Aparma, I. Anand, and B. Sanjay, "Contrast Limited Adaptive Histogram with edge enhancement for medical X-ray images," *International Journal of Humanities, Arts, Medicine and Sciences*, Vol. 3, Issue 6, pp. 9-16, (2015).

[5] A. Bhardwaj and M. Singh, "A novel approach of medical image enhancement based on wavelet transform", *International journal of engineering research and application*,Vol.2, issue.3, pages.2356-2360,ISSN.2248-9622,(2012).

[6] R. C. Gonzalez and R. E. Woods, "Digital Image Processing,", New York: Prentice Hall, (2008).

[7] A. N. Hirani and T. Totsuka, "Combining Frequency and Spatial Domain Information for Fast Interactive Image Noise Removal", *in Proceedings of the International Conference on Computer Graphics and Interactive Techniques*, pages.269-276,(1996).

[8] S. Jayaraman, S.Esakkiranjan and T. Veerakumar, T, "Digital image processing", New Delhi: Tata McGraw Hill,2015.

[9] R. C. Kenneth, "Digital image processing",USA: Prentice Hall.1996

[10] V. Kumar, P. Gupta, "Importance of statistical measures in digital image processing," *International Journal of Emerging Technology and Advanced Engineering*, Vol. 2, Issue 8, pp. 56-62, (2012).

[11] R. Maini and H. Aggarwal, "A Comprehensive Review of Image Enhancement Techniques ", *JOURNAL OF COMPUTING*, Vol.2, issue.3, pages.8-13, ISSN.2151-9617, (2010).

[12] MedPix, *Medical Image Database*, Radiology. At: https://medpix.nlm.nih.gov/home. [Accessed 2015-7-22].

[13] S. Mitra and J.Sicuranza, "Nonlinear image processing,", San Diego: Academia Press, 2001.

[14] S. Mundhada and V. Shandilya, "Spatial and transformation domain techniques for image enhancement", *International journal of engineering science and innovative technology*, Vol.1, issue.2, pages.213-216, ISSN.2319-5967(2012).

[15] M. Ritika And S. Kaur, "Contrast Enhancement Techniques for Images– A Visual Analysis," *International Journal of Computer Applications*, Vol. 64, issue.17, pages.20-25, ISSN.0975 – 8887, (2013).

[16] M. Ritika, "A novel approach of local contrast enhancement of medical images using mathematical morphology", *International journal of computer science and information technology and security*, Vol.2, issue.2, pages.392-397, ISSN.2249-9555, (2012).

[17] Y. Rajput, V. Rajput, A. Thakur and G. Vyas, "Advanced image enhancement based on wavelet and histogram equalization for medical images", *IOSR Journal of electronics and communication engineering*, Vol. 2, issue.6, pages.12-16, ISSN.2278-2834, (2012).

[18] H.E. Awad, M. Hjouj Btoush and E. Tadros, "A Robust Watermarking Algorithm Based on Two Dimensional Cellular Automata," *International Journal of engineering Research and Industry Applications*, Vol. 5, Issue.4, pp. 249-264, (2012).

[19] I. Shakhrah, "Digital High-Pass Filters with Milder High-Pass Effect on Digital Images", *American Journal of Engineering and Applied Sciences*, Vol.8, issue.3, pages.360-370, ISSN.10.3844, (2015).

[20] I. Suneetha and T. Venkateswarlu, "Enhancement Techniques for True Color Images in Spatial Domain ", *International Journal of Computer Science and Technology*, Vol.3, issue.2, pages.814-820, ISSN.2229-433, (2015).

[21] K. Thangadurai. And K. Padmavathi, "Noise Reduction and Image Sharpening Using Linear Spatial Filtering in Plant Leaves Disease Detection ", *Journal Computer Technology & Applications*, Vol.5, issue.4, pages.1561-1565, ISSN.2229.6093, (2014).

[22] A. M. Trifas, Medical image enhancement. PhD Dissertation, Louisiana State University, East Lansing, MI, USA.2005.

[23] K. Zuiderveld, "Contrast Limited Adaptive Histograph Equalization," In Graphics gems IV, Paul S. Heckbert (Ed.). Academic Press Professional, Inc., San Diego, CA, USA, 1994, pp. 474–485.

APPENDIX A

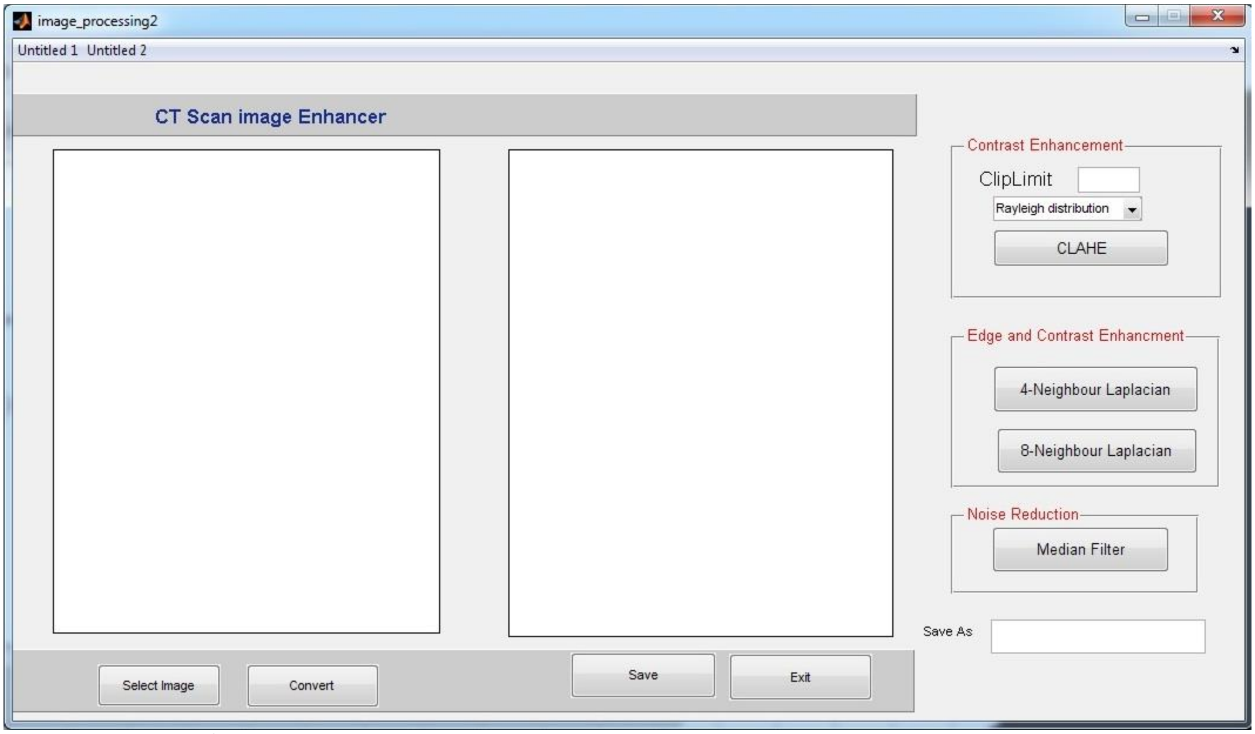


Fig A1: Main window of proposed system

An Enhencment Medical Image Compression Algorithm Based on Neural Network

Manel Dridi
Electronics and Micro-Electronics Laboratory
(E. μ. E. L), University of Monastir – Tunisia
National Engineering School of Monastir, University of Monastir – Tunisia

Mohamed Ali Hajjaji
Electronics and Micro-Electronics Laboratory
(E. μ. E. L), University of Monastir – Tunisia
Higher Institute of Applied Sciences and Technology of Kasserine, University of Kairouan

Belgacem Bouallegue
Electronics and Micro-Electronics Laboratory
(E. μ. E. L), University of Monastir – Tunisia

Abdellatif Mtibaa
Electronics and Micro-Electronics Laboratory
(E. μ. E. L), University of Monastir – Tunisia
National Engineering School of Monastir, University of Monastir – Tunisia

***Abstract*—The main objective of medical image compression is to attain the best possible fidelity for an available communication and storage [6], in order to preserve the information contained in the image and does not have an error when they are processing it. In this work, we propose a medical image compression algorithm based on Artificial Neural Network (ANN). It is a simple algorithm which preserves all the image data. Experimental results performed at 8 bits/pixels and 12bits/pixels medical images show the performances and the efficiency of the proposed method. To determine the 'acceptability' of image compression we have used different criteria such as maximum absolute error (MAE), universal image quality (UIQ), correlation and peak signal to noise ratio (PSNR).**

Keywords—Artificial Neural Network; medical image; compression; DICOM; PSNR; CR

I. Introduction

DICOM (Digital Imaging and Communications in Medicine) image is composed of two files [15]: header files and image pixel data. Therefore, the treatment of digital medical images as textual data, such as computed tomography (CT) and the magnetic resonance images (MRIs) required an extensive amount of hard disk space, also a high transmission time needed to store and to transmit because of the large file size [9]. One of the solutions is to compress medical images before its transmission and storage, then decompress it at the receiver for use. At present, the main steps of image compression are: pixels transform, quantization and entropy coding [11]. In the last years, compression methods have attracted the interest of many researchers all around the world; many compression methods have been proposed. Among the already proposed methods such as JPEG2000, and JPEG cannot be satisfied due to the long execution time, complexity of algorithms or/and lower compression ratio.

As flown we present some researchers work. In [5] M.ElZorkany has proposed a new image compression approach which combines NN and DCT. In [9] W.K.Yeo et al. have presented their medical image compressed algorithm, which is based on Hebbian process and quantization of the extracted components. In [10] S.manimurugan et al. proposed a crypto-compression medical image using the block pixel sort scheme. In [18] A.Younus et al. have proposed a hybrid medical image compression technique based on Discrete Cosines Transform (DCT) and Lapped Biorthogonal Transform (LBT). In [7] S.kuamo et al. present an experimental study of some image compression methods and propose a new hybrid method based on a neural network.

This paper is organized as follows, in the second section, we define neural network and give its different features. In the third section we present the proposed scheme. In the fourth section, we present some treated examples and tests followed by some comparison with respect to other works. Finally, we conclude this paper.

II. Artificial Neural Network (ANN)

In this section, we present a brief introduction of Artificial Neural Network (Fig.1), which is inspired from human brains. It's a data modeling tool able to capture and represent a complex input/output relationship. Because of the specified characteristic for ANN such as high degree of interconnection, nonlinear mapping, massive parallel structure and self-organization, we can estimate that ANN can resolve some compression problems [12].

The ANN parameters are : the number of training iterations and hidden neurons , the input layer units for each hidden layer unit (N) and the range of synaptic weight values (MINMAX) {-L...+L}. The main steps for ANN are:

A. Initiation

Put all weights and activation thresholds of the network to random values uniformly distributed in a small range. This initialization is done for one neuron. Set the value of the learning rate to a small positive value.

B. Activation

The activation function in ANNs algorithm is a weighted sum: the sum of the inputs X multiplied by their respective weights w_{ji}):

$$Xj = f(\sum_{i=1}^{N} w_{ij} \times x_i) \text{With } 1 \leq j \leq M \quad (1)$$

Activate the neural network using input signals and desired output.

Calculate the output signals of the neurons of the successive layers and the first hidden layer to the output layer (see eq 1).

$$\text{Where} \quad f(x) = \frac{1}{1+e^{-x}} \quad (2)$$

C. Training weights:

Update the weights of the network by propagating the error in the opposite direction

$$w_i(p+1) = w_i + \Delta w_i \quad (3)$$

Calculate the error gradient for the hidden layers (reverse order) and for the output layer.

Calculate the weight corrections for each layer and update the weights.

D. Iteration

Increment of p, return to step 2 and repeat the process while the error is reached

1) Set a random weight were used to start. The weights were usually between 0.01 and .99.

2) Calculate each output layer

3) Calculate errors: the error of the neuron k is the difference between the value of the desired output and the actual value of the output neuron k

$$e_k(p) = y_{d,k} - y_k \quad (4)$$

4) Compute delta errors

$$\Delta w_{jk}(p) = \alpha \times y_j(p) \times \delta_k(p) \quad (5)$$

Where $\delta_k(p)$ is the neuron error gradient k at p iteration.

a) The error Gradient for output layer neurons

$$\delta_k(p) = f'[X_k(p)] \times e_k(p)$$

$$= y_k(p) \times [1 - y_k(p)] \times e_k(p) \quad (6)$$

The second part is determined by the derivative of the sigmoid function applied to the input net of the neuron k at p iteration

b) The error Gradient for hidden layer neurons

$$\delta_j(p) = y_j(p) \times [1 - y_j(p)] \sum_{k=1}^{l} \delta_k(p) w_{j,k}(p) \quad (7)$$

Where l is the number of neurons in the next layer

III. Compression Algorithm: The Proposed Algorithm

In this section, we present the proposed method in detail (Fig 2). There are many different types from ANN [5] in our work; we have used Multilayer Perceptron (MLP) with three layer input layer, one hidden layer and an output layer as shown in Fig.1 above.

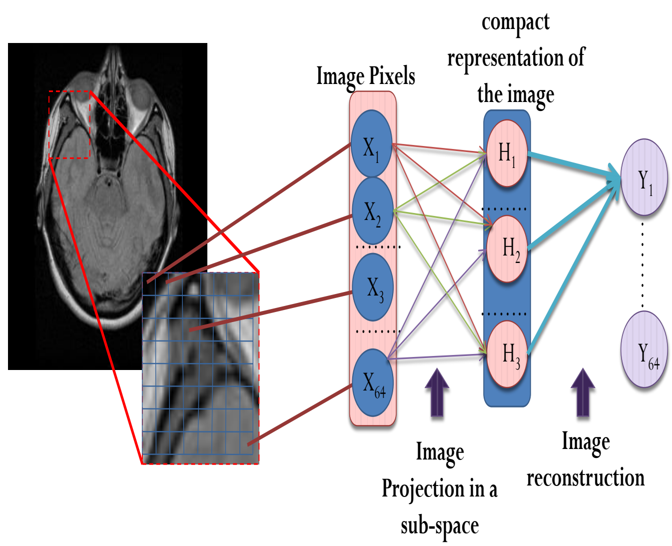


Fig. 1. Architecture of neural network

First off, the entire pretreatment step: the image, I, pixels is converted to double data type then normalized (I=I/4000). Secondly, the image is divided into square blocks of n by n size in order to reduce execution time. Then each of these blocks are transformed into a vector of N-dimensional. The input vectors of attributes, X_N, to the input layer can be written as $X_K=\{x_1 \ldots\ldots xi \ldots\ldots x_N\}$ where K = (1,2,...,K), K is the total vector number. Thirdly, for the training step in which the data of every vector are entered in the network as an input data iteratively: In each iteration, the network weights are updated according to equation (2), and the residues error obtained as the difference between the actual and the predicted calculated output. The error is used to adjust the weights. The training continues until the predictive error is reached. The allowable error is equal $1e^{-10}$. Finally, the corresponding encoding from the obtained hidden layer is saved with the output weight. To get the decompressed image, we use the appropriate compressed image data and extracted weights to train the neural network. The appropriate compressed image is entered in the output layer block by block and the output data are computed in order to recreate the original image.

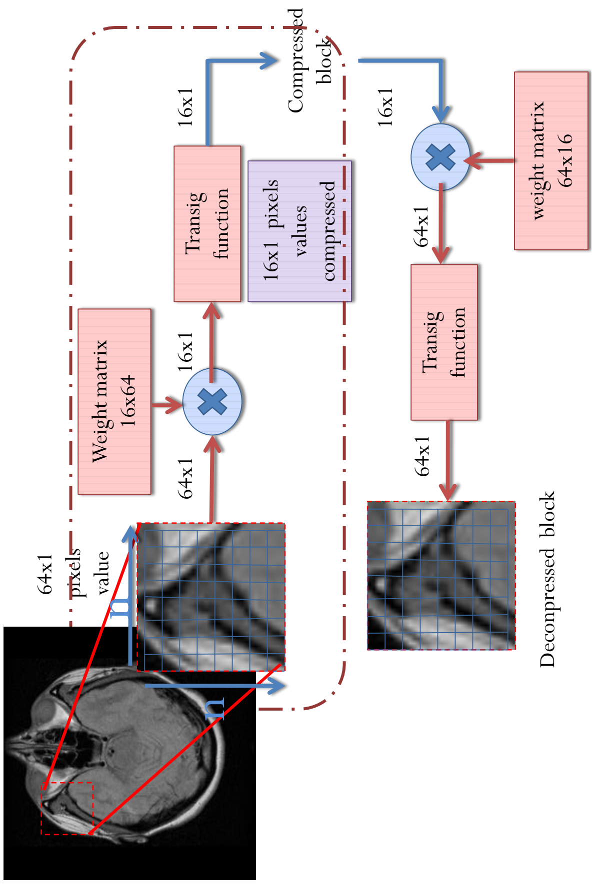


Fig. 2. The proposed process

The algorithm

Step 1: [Start] Generate a random weight and divided the image in sub blocks

$(w_{ij}$=[0.5 - rand(nbc2,nbc1)]

Step 2: calculate hidden and output layer

$$s1j=\sum_{i=0}^{n} w_j(i) \times imp(i)$$

$$s2k=\sum_{j=0}^{n} w_k(j) \times H(j)$$

H=[2./(1+exp(-s1))]-1;

out=[2./(1+exp(-s2))]-1;

Step 3: Create a new weight by repeating the following steps until we get a perfect quality of image.

e(k) = d(k)-out(k);

```
E=E+(1/2*((e(k))^2));
derfs2=(1-out).*(1+out);
delta1 = e.*derfs2;
gradES(j,k)=mu*delta1(k)*H(j);
w1=w1+ gradES;
gradEC(i,j)=mu*delta2(j)*imp(i);
w=w+ gradEC;
```

Step 4: Use new generated weight for a further algorithm run.

Step 5: [Test] if the end condition is satisfied, stop, and return the best solution

Step 6: Go to step 2.

According to the Neural network the process of compression is presented by eq (8) for encoding:

$$h_j= f(\sum_{i=1}^{N} w_{ij} \times x_i) \quad \text{with } 1\leq j\leq M \tag{8}$$

Where f(x) is a sigmoid activation function:

$$f(x) = \frac{2}{1+e^{-x}} \tag{9}$$

and by eq (9) for decoding

$$yi= f\left(\sum_{j=1}^{M} w_{ij} \times h_j\right) - 1 \tag{10}$$

with 1≤i≤N

Where $y_i \epsilon$ [0, 1] and presents the normalized pixel values. We used normalized pixel values because the neural network can perform more efficiently when their inputs and outputs belong to the interval [0, 1].

IV. EXPERIMENTAL RESULTS

The proposed approach for the compressed medical image is implemented in Matlab using a personal computer with an intel dual core 2.2 GHz processor, windows7.

In order to evaluate the performances of our approach, we were considered some indicators which are the most used in the research area:

A. Definition

- **CR**

CR is the ratio between the uncompressed and compressed image size

$$\text{Compressionratio} = \frac{\text{Original image size}}{\text{Compressed image size}} \tag{11}$$

- **MSE**

Mean Squared Error is the cumulative squared error between the compressed and the original image. The large value of MSE means that image is poor quality. It is defined as follows:

$$\text{MSE} = \frac{1}{W\times L}\sum_{i=1}^{W}\sum_{j=1}^{H}[P(i,j) - C(i,j) \tag{12}$$

- **MAE** (absolute error)

The MAE between two images is defined according to two images, P(i,j) and C(i,j), which correspond to the plain-image and the decompressed image. It's defined by equation (12)

$$\text{MAE} = \frac{1}{W\times L}\sum_{i=1}^{W}\sum_{j=1}^{H}|P(i,j)\_C(i,j)| \times 100\% \tag{13}$$

- **PSNR** [9]

It means the Peak Signal to Noise Ratio which can be used to evaluate a compression scheme. It's defined as follow by equation (14) in which b is the number of bit coding.

$$\text{PSNR} = 10 \times \log_{10}\left[\frac{L\times W \times \max(P)^2}{\sum_{i=1}^{L}\sum_{j=1}^{W}(P(i,j)-C(i,j))^2}\right] \tag{14}$$

Typically, the values of PSNR, for lossy compression of an image, are between 30 and 50 dB and when the PSNR is greater than 40 dB, then the two images are indistinguishable.

- Correlation analysis

In an ordinary image, each pixel is highly correlated with its adjacent pixel. However, the test of the adjacent pixels correlations in the decompressed image determines the superior confusion and diffusion characteristics [8]. The correlation of the adjacent pixels is calculated with the use of Eq. (15), (16), (17) and (18).

$$E(x) = \frac{1}{N}\sum_{i=1}^{N} x_i \; ; \quad (15)$$

$$D(x) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(x_i))^2 \quad (16)$$

$$cov(x,y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(x))\big(y_i - E(y)\big) \quad (17)$$

$$r_{xy} = \frac{cov(x,y)}{\sqrt{D(x)}\sqrt{D(y)}}, \quad (18)$$

where x and y are grayscale values of two adjacent pixels in the image. E(x)is the expectation of x, D(x) is the estimation of the variance of x and cov(x,y) is the estimation of the covariance between x and y.

B. *Results and interpretation*

1) Results

In this part, we have given different performance analysis of the proposed process. By using the formula presented in the previous part, the evaluation of the reconstructed image was performed.

Table 1, Table 2 and Table 3 represent the PSNR results, Universal Image Quality Index, SNR and correlation of some medicals images. These results illustrate that, if the input and the hidden layer numbers decrease the PSNR and Universal Image Quality Index will also increase. As shown in a tab, it was found that the PSNR was between 34dB and 50dB.

We remark from correlation values results that there is an important relation between the pixels of the original image and the decompressed one.

2) The influence of the size of sub-block

In this section, we show the influence of the number of the input layer and the number of the hidden layer on the quality of the obtained image after decompression. Presence of hidden neurons enhances the flexibility of the system and increases power processing, but if the number of hidden neurons is taken too small, then robustness of the system can reduce due to improper fitting of input data. Therefore, in order to have a good result, we don't use few neurons in the hidden layers.

If we observe the graph (Fig 3, Fig 4 and Fig5) and the Fig 6 presented above, we can conclude that the use of 64 inputs for the input layer provides a more reliable quality image than the use of 36 inputs for the input layer.

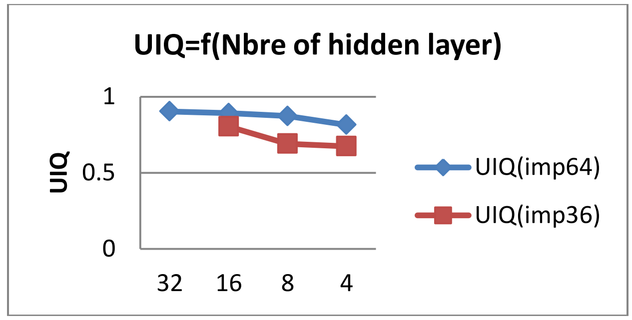


Fig. 3. universal image quality

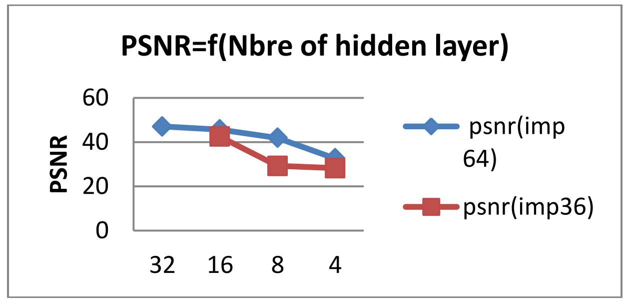


Fig. 4. PSNR (db)

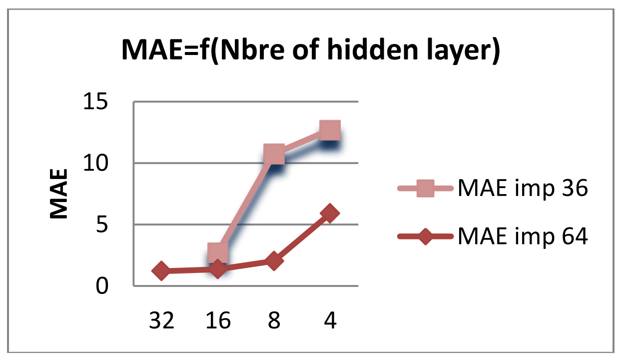
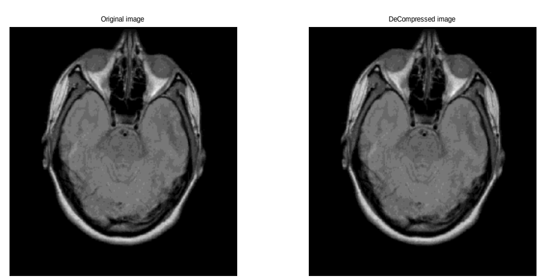
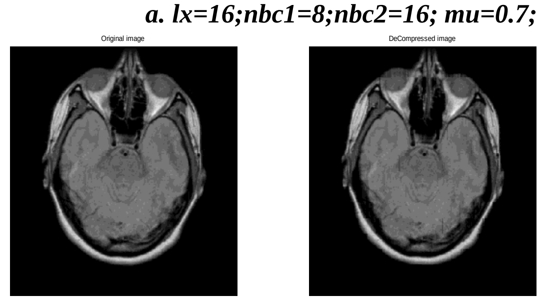


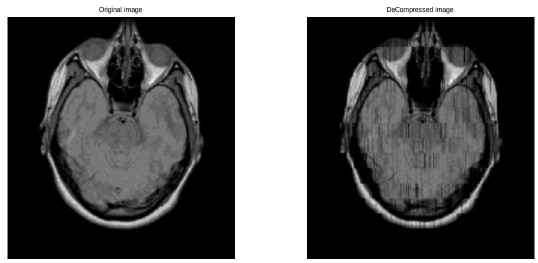
Fig. 5. MAE



a. lx=16;nbc1=8;nbc2=16; mu=0.7;



b. Lx=64;nbc1=32;nbc2=64, mu=0.7



c. lx=16;nbc1=8;nbc2=16;mu=0.7;

Fig. 6. sub-blocks size influence

We have performed several simulations to find the best parameters in order to have a better decompressed image quality. As seen in previous figure5, we note that for mu = 0.7 the resulting image is acceptable. But this result is just for some images, but for another one we should find the parameter mu which gives a good result.

TABLE I. NUMBER OF INPUT =64, NUMBER OF HIDDEN LAYER= 32, NUMBER OF OUTPUT =64; MU=0.6 EXPECT FOR 2 MU=0.2

	PSNR(db)	UIQ	SNR	Correlation of decompressed image
1	47.2245	0.81371	35.7104	0.99954
2	45.0224	0.68115	33.0231	0.9997
3	44.3494	0.77337	33.8874	0.99951
4	45.5447	0.78747	34.537	0.99952
5	40.7745	0.78747	28.7745	0.99952
6	42.9413	0.38702	26.2487	0.99747
7	48.5782	0.69393	34.0252	0.99937
8	50.2165	0.67070	35.402	0.99965

TABLE II. NBRE OF INPUT =64, NBRE OF HIDEN LAYER= 16, NBRE OF OUTPUT =64;MU=0.6

	PSNR(dB)	UIQ	SNR	Correlation of decompressed image
1	45.7308	0.89591	34.2158	0.99974
2	46.0224	0.77833	34.0232	0.99982
3	43.311	0.84740	32.8491	0.99966
4	44.1974	0.86226	33.1964	0.99970
5	37.363	0.99857	25.4508	0.99857
6	39.5329	0.42105	23.8405	0.99816
7	44.6791	0.83524	30.1262	0.99961
8	48.1061	0.76283	33 .2916	0.99975

TABLE III. NBRE OF INPUT =36, NBRE OF HIDEN LAYER= 18, NBRE OF OUTPUT =36; MU=0.6

	PSNR(dB)	UIQ	SNR	Correlation of decompressed image
1	42.9718	0.80863	31.4568	0.99949
2	43.9466	0.64007	31.9474	0.99974
3	41.5185	0.76174	31.9474	0.99931
4	41.8933	0.78607	30.8923	0.99949
5	34.1588	0.50070	22.2459	0.99699
6	38.05	0.37824	22.3635	0.99739
7	42.6312	0.70103	28.0783	0.99938
8	46.4447	0.66797	31.6303	0.99965

V. COMPARISON BETWEEN OTHER COMPRESSION METHODS

In order to show the efficiency of the proposed process, we have compared it firstly with JPEG compression and with wavelet compression secondly. This comparison is carried out by calculating the compression ratio (CR) and peak signal to noise ratio (PSNR) for medical image compressed by JPEG, then by comparing the image resulting using the proposed algorithm with the image resulting using other proposed method.

As seen in Table 4, we can conclude that our proposed process gives a good result. In fact, our result is better than the result in 13 and 14 for CR value.

TABLE IV. COMPARISON WITH OTHER WORKS

	CR	PSNR(db)
our	7,44	32,08
13	Max =3,05	23.9
14	5.387	42.499

and as seen as the obtained image in [17] (Fig. 8.) we remark that the original image and the compressed one haven't the same psycho-visual characteristic. Contrariwise, we obtained the image (Fig.7) has the same psycho-visual characteristic. Since, we are working with medical images that may cause some problems.

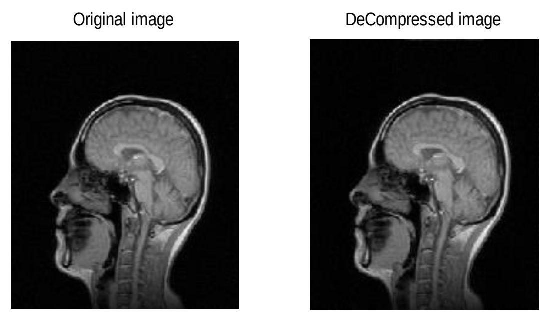
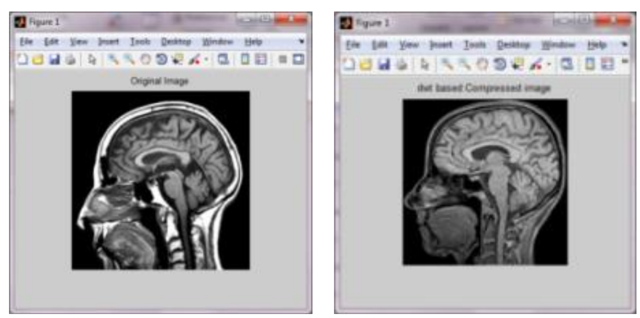


Fig. 7. The result image obtained by our proposed algorithm



a.Origanel image b.Compressed image using WBCT algorithm

c. Compressed image using WDT

Fig. 8. The result image obtained in the related work [16]

Compared to the other works, our proposed scheme presents a compromise between the quality of the obtained image and the compression ratio. In addition, the proposed algorithm is more easily adapted than the other algorithm such as JPEG and JPEG2000.

VI. CONCLUSION

In this work, we have proposed a compression method for medical image which preserve the quality of the original image. This method is based on neuronal network with some modification in the back-propagation usual algorithm. We have shown the efficiency and the performance of the proposed scheme by carrying out many measurements for differential analysis such, the correlation, the PSNR and the Universal Image Quality Index. The simplicity of the algorithm makes real-time implementation easy therefore in the future work we will implement this algorithm on FPGA.

REFERENCES

[1] Y. Shantikumar Singh, B. Pushpa Devi, Kh.Manglem Singh, Image Compression using Multilayer Feed Forward Artificial Network with conjugate Gradient ,World Congress on Information and Communication Technologies2012 IEEE.

[2] P.V Rao, S.Madhusudana, N.SS and KusumaKeerthi, Image Compression using Artificial Neural Networks, Second International Conference on Machine Learning and Computing, 2010.

[3] D.Bouslimi, G. Coatrieux, Christian Rouxa, A joint encryption/watermarking algorithm for verifying the reliability of medical images: Application to echographic images, computer methods and programs in (2012) pp47–54.

[4] U.Seiffert, ANNIE-Artificial Neural Network-based Image encoder, Neurocomputing 125(2014), pp:229-235.

[5] M.ElZorkany, A hybrid image compression technique using neural network and vector Quantization with DCT, Image Processing and Communications Challenges 5, Advance in Intelligent Systems and Computing 233,Springer International Publishing Switzerland, 2014.

[6] A.Saffor, A.R.Ramli and Kwan-Hoong Ng, A Comparative Study of Image Compression Between JPEG and WAVELET, Vol. 14 No. 1, June 2001, pp. 39-45.

[7] S.Kouano and C.Tangha, Image Compression with Artificial Neural Networks, Springer-Verlag Berlin Heidelberg 2013. Int. joint conf. CISIS'12-SOCO'12.AISC 189, 2012, pp: 515-524.

[8] Chong Fu a,n, Wei-hongMeng b, Yong-fengZhan b, Zhi-liangZhu c, FrancisC.M.Lau d, Chi K.Tse d, Hong-fengMae ,An efficient and secure medical image protection scheme based on chaotic maps, Science Direct journal ,Computers inBiology and Medicine,2013.

[9] W.K Yeo, D.F.W.Yap,D.P.Andito, MK.Suaidi, Grayscale MRI Image Compression Using feedforward Neural Networks, Proceedings of the 6th International conference on Broadband Communications & Biomedical Application , November 21-24,2011.

[10] S.Manimurugan, porkumaran, Secure Medical Image Compression Using Block pixel Sort Algorithm, European Journal of scientific Research, Vol.56 No2 (2011), pp: 129-138.

[11] Jiang.J, Image compression with neural network, A survey. Signal Processing: image communication, 1998.

[12] N.Sriraam, Ahigh-Performance Lossless Compression Scheme for EEG Signals Using Wavelet Transform and Neural network Predictors, International journal of telemedicine and application, Volume 2012.

[13] K Gopi, Dr. T. Rama Shri, Medical Image Compression Using Wavelets, IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 2, Issue 4 (May. – Jun. 2013), pp:1-6.

[14] G.Vijayvargiya, S.Silakari,R.Pandey,A Novel Medical Image Compression Technique based on Structure Reference Selection using Integer Wavelet Transform Function and PSO Algorithm, International Journal of Computer Applications (0975 – 8887) Volume 91 – No.11, April 2014, pp:9-13.

[15] D.Ravi Varma, "Managing DICOM images: Tips and tricks for the radiologist", Indian J Radiol Imaging. 2012 Jan-Mar; 22(1), pp: 4–1

[16] Tjokorda A.B.W, Adiwijaya,Febri Puguh Permana , Medical Image Watermarking with Tamper Detection and Recovery Using Reversible Watermarking with LSB Modification and Run Length Encoding (RLE) Compression

[17] J. Hyma1, P.V.G.D. Prasad Reddy2, A. Damodaram, Comparative Analysis of Compression Techniques and Feature Extraction for Implementing Medical Image Privacy Using Searchable Encryption, ,Advances in Intelligent Systems and Computing, Springer International Publishing Switzerland 2015, pp:63-69.

[18] A.Younus,G.Raja, A.K.Khan, Hybrid Compression of Medical Images Based on Lapped Biothogonal Transform & DISCRETE COSINE TRANSFORM

A QoS Solution for NDN in the Presence of Congestion Control Mechanism

Abdullah Alshahrani
The Catholic University of America

Izzat Alsmadi
University of New Haven

Abstract—**Both congestion control and Quality of Service (QoS) are important quality attributes in computer networks. Specifically, for the future Internet architecture known as Named Data Networking (NDN), solutions using hop-by-hop interest shaping have shown to cope with the traffic congestion issue. Ad-hoc techniques for implementing QoS in NDN have been proposed. In this paper, we propose a new QoS mechanism that can work on top of an existing congestion control based on interest shaping. Our solution provides four priority levels, which are assigned to packets and lead to different QoS. Simulations show that high priority applications are consistently served first, while at the same time low priority applications never starve. Results in ndnSIM simulator also demonstrate that we avoid congestion while operating at optimal throughputs.**

Keywords—Named Data Networking; Quality of Service; Congestion Control

I. Introduction

Throughout the years, observation of todays Internet architecture has shown that its usage has shifted from its original route sough in early design. In order to have a more adequate design for future Internet, the U.S. National Science Foundation [21] is funding the NDN project [19], which focuses on developing a Content-Centric architecture. In contrast to the host-centric architecture of IP networks, NDN allows users to request data without the need of referencing the server IP address. Therefore, the routers are responsible to find the data provider, relieving the user of complex configurations and making the data-hosting more flexible and network independent.

In NDN architecture, requests are called *interests* while data itself is named *content*. End-points that send interests are called *consumers* and end-points that provide contents are called *producers*. In this paradigm, an interest packet, which contains the name of the resource it requires, is replied by a single content packet, which also contains the resource name. Content packets always traverse the inverse path of the interests that triggered them. This characteristic can be translated as Property 1.
Property1: In NDN the sum of all interest packets that traverse a link in one direction is equal to the sum of all content packets that traverse the same link on the reverse direction. This property implies that hop-by-hop techniques are particularly interesting for NDN, because the flow of content packets is predictable (as shown in Fig. 1)

The technique proposed in [11] proposes a hop-by-hop congestion control mechanism for NDN that also relies on Property 1. This solution is based on the idea that controlling the interest packet rate in one direction can regulate the traffic of contents in the reverse direction. By shaping the outgoing interest rate to an optimal value, the incoming content rate is indirectly shaped to maximize the use of available bandwidth while minimizing the loss of content packets.

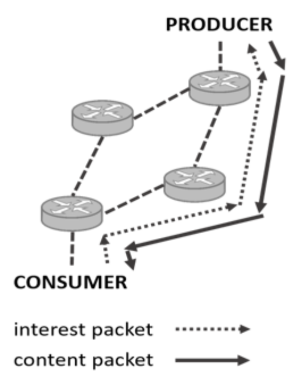


Fig. 1: Content packets always travel the inverse path of interest packets in NDN

Besides congestion control concerns, NDN should offer the possibility of discerning among different service priorities. For instance, video streaming and gaming applications must be treated with more priority by routers than a regular download. Thus when co-existing, high-priority applications should be served faster than low-priority ones. In existing networks, this service is among one of the many features provided by Quality of Service (QoS) mechanisms. In other words, QoS techniques should offer to higher priority services higher throughput, shorter transmission delays, more availability and less jitter.

In this paper, we propose a practical hop-by-hop solution for QoS in NDN that can be integrated with the existing congestion control solution proposed in [11]. Our mechanism offers four priority levels, each one designed for a particular class of applications. *Real-time traffic (higher priority apps)* is always served first, guaranteeing high QoS; while low-priority services have a minimum QoS with no starvation. Results at ns-3 ndnSIM [13] simulator show that the proposed mechanism reaches optimal bandwidth usage with a negligible congestion. In addition, the solution offers a flexible QoS configuration, making it possible to change at will the way that routers treat different priorities.

The rest of the paper is organized as the following: Section II introduced relevant related papers. Next section introduced interest shaping congestion control. Main paper contribution is in next sections experiments and analysis. Paper is then concluded in section VI.

II. Related Work

Named Data Networking (NDN) [22] extended original project:(Content-Centric Networking (CCN),[23] as an alternative architecture to classical IP-based Internet networking architecture. Unlike the former IP-based networking, NDN aims at uniquely identifying Internet content rather than users or hosts. This information-centric architecture is proposed to allow the Internet to accommodate usage scenarios that it is believed that IP-based architecture was not natively built to handle or accommodate such scenarios. For example, a popular page in the Internet can be redundantly saved and stored in different locations. The large volume of users' traffic with such popular pages can be significantly reduced if the Internet architecture handles all those instances of similar content synonymously. Consumers or data requesters express their interests and routers exchange this traffic based on content name. Receivers or owners of this content send back requested content in the same-but reversed paths. The focus or scope of our paper; a user- or content-driven resources'-allocation network is an example of those usage scenarios. If the underlying architecture is a host-driven, it will be hard for such architecture to accommodate dynamically different bandwidth and resources for different applications coming from the same users or hosts. Their are two key elements in NDN, related to our paper scope. Those are stateful forwarding which enables transmitted packets to traverse the same, reversed, path of packet requests and second adaptive forwarding which enables routers ti change their future routing decisions based on current traffic and also consumers' requests.

In the context of NDN and user-driven congestion controls [24] proposed a receiver-driven congestion algorithm to predict the location of desired content. For each packet, timeout is defined in their algorithm based on originating nodes. For each packet, the originating nodes are expected to list next data packets. As a result, each router can judge if they have those next packets or their destinations where ultimately the receiver can have a map for future packets and hence better manage or control traffic congestion. The idea of the involvement of receivers or consumers in smart or adaptive routing decisions is discussed also in similar other contributions (e.g. [28]). In this paper, authors proposed an Explicit Congestion Notification (ECN) based routing algorithm in which consumers can utilize network-wide range information to make smart routing decisions. [25] proposed an architecture for an adaptive routers' multi-path forwarding algorithm based on NDN to accommodate in addition to traffic congestion, security problems and traffic failures. In order to minimize congestion, routers should try efficient routing algorithms to avoid the last resort of sending interests to all interfaces. Their algorithm is based on ranking router interfaces and sending interests to interfaces with highest ranks. Routing information and forwarding performance are the two main factors to consider in this ranking scheme. Their algorithm extended original algorithm in NDN prototype software implementation:CCNx [26]. Upon a link failure, next-ranked interface will take the rank precedence.

In the same scope of NDN content adaptive forwarding, [27] proposed nCDN to enhance Content Delivery Network (CDN) based on NDN and allow NDN over UDP/TCP to utilize best features of NDN such as: Caching, multi-casting and stateful forwarding to reduce network congestion. Identical interests can be aggregated to optimize resources' utilization.

Other QoS solutions have been proposed for NDN. In [14], the authors propose to include a data lifetime information on the NDN packet header. The lifetime tag at the packets are translated into the storage lifetime at Content Stores. This is implied into improvement of service availability for applications that emit larger lifetime packets.

In [6], IP differentiated services are used as a guideline to implement QoS in NDN. The authors propose the use of traffic classification and packet tagging at edge routers. The core routers offer a special treatment to packets with certain tags.

None of the previous solutions considered the traffic congestion issue. Therefore, we believe that this is the first mechanism that make use of the control congestion structure to provide QoS. Regarding the traffic congestion issue, there are several proposals of solutions. In [15], [16], the authors proposed a credit-based interest shaping scheme to control data traffic volume. Each flow is giving certain credits at a rate that is equal to the available downlink capacity divided by the number of bottlenecked flows. If the outgoing Interest rate for a flow exceeds the credit, the flows Interests will be dropped or delayed. Note that a content flow is defined as the packets bearing the same content name or prefix identified on the fly by parsing the packet headers. In [17], the authors proposed a rate-based hop-by-hop congestion control mechanism to shape the Interest rate that follows the same principles as the rate control mechanism for Available Bit Rate (ABR) in ATM networks [18] and the interest sending rate is computed as a function of the target flow queue occupancy.

In [11], a per-interface based Interest limit mechanism is used in which a router limits the rate in forwarding interest packets for each interface to prevent congestion and an interface becomes unavailable when it reaches its interest limit. The interest rate limit is based on link capacity and estimated requested packet size. This work is elaborated in Section III, since it is used as a basis for our QoS solution.

In [19], a fair share interest shaping (FISP) is proposed. In this solution, a content router decides whether to forward an interest immediately or delay it temporarily based upon the data queue sizes and the flow demands. Then it realizes fair bandwidth sharing among flows by having per-flow interest queues at an interface. Next, it uses a modified round robin scheduling mechanism for these per-flow interest queues (the deficit counter of the queue is decreased by the size of corresponding data after sending an interest).

These existing Interest shaping mechanisms assume that the link capacity is equally divided among the bottlenecked flows during congestion. They do not consider different priorities of various flows. An unloaded network can meet the

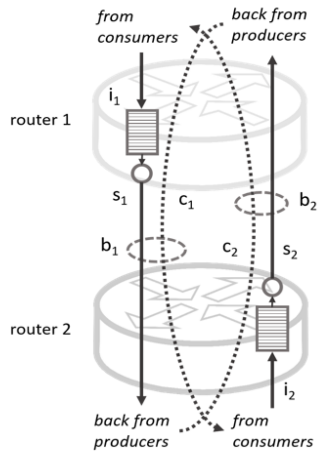


Fig. 2: Control congestion solution presented in [3]

needs of all applications and subscribers. However, the importance of quality of service (QoS) increases during periods of congestion. In an overloaded network, it is vital to ensure the QoS of the important, delay-sensitive traffic (e.g., continuous uninterrupted multicast TV flows viewed by thousands), while less important and delay-tolerant traffic can be buffered or discarded.

The rest of our paper is structured as follows: Section II contains a description of related work. In Section III, we briefly explain the control congestion technique presented in [11], since we reused the same principles. Section IV describes our QoS solution for NDN. In Section V we present our results and make a qualitative analysis of the ideal configurations. Finally, we conclude our paper in Section VI.

III. Interest Shaping Congestion Control

In [11], the authors proposed a hop-by-hop interest shaping mechanism to cope with network congestion. As shown in Fig. 2, each router has an uplink and downlink bandwidth related to that hop, named b_1 and b_2 respectively. The interfaces at each side of the link receive interests that need to be sent out through that link (at rates i_1 for router 1 and i_2 for router 2). In addition, returning contents arrive at rates c_1 and c_2. At router 1, interests are placed at the shaping queue at i_1 rate. Then the shaper decides whether to limit or not the outgoing interest rate (named hereafter shaping rate or s_1 for router 1). The same occurs at router 2. In the NDN architecture, each interest packet sent through a hop must be replied with exactly one content packet (cf. Property 1). Therefore, any device interface will send the same number of interests o contents it receives (except if there is packet loss). So if we measure s_1 and c_2 in number of packets per second, s1 should be equal to c_2, meaning that controlling only the interest rate (s_1) implies in indirectly controlling the content rate (c_2). In that way, wisely choosing the shaping rate leads to an optimal use of the bandwidth with minimal congestion. For instance, suppose that at router 1 we measure the s_2 rate, and it occupies 10% of the b_2 bandwidth. Then s_1 can be chosen so that c_2 occupies the 80% left of b_2, reaching a full usage with no congestion. That same logic applies to router 2. The adaptation of s_1 and

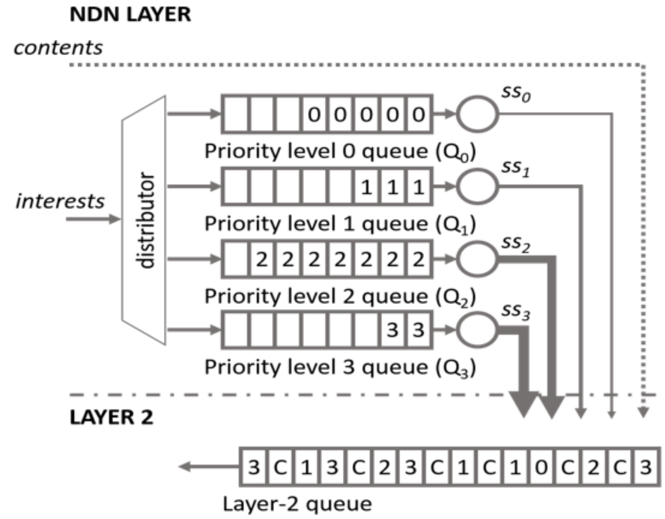


Fig. 3: Implementing priorities

s_2 occurs instantly. Actually, s_1 and s_2 may assume less than optimal values to have some slack for burstiness. In [11], the shaping rate is chosen according to a theoretical analysis of the congestion problem at a single hop. First, the authors define a range for shaping rates, with maximum and minimum shaping rates. This range depends on the instant values of b_1, b_2, i_1, i_2, and the ratios between interests and content sizes. In real-world scenarios, all these data can be measured at any instant. Then, actual shaping rate s_1 is chosen inside this range by confronting the observed instant shaping rate s_2 (called here obs s_2) with the minimum expected shaping rate s_2 (called here expmin s_2). Eq. 1 gives the actual shaping rate s_1.

$$s_1 = \min s_1 + (\max s_1 - \min s_1)(1 - \frac{obs\ s_2}{expmin\ s_2})^2 \quad (1)$$

$$\text{where } \frac{obs\ s_2}{expmin\ s_2} \leq 1$$

As shown in Eq. 1, the greater the ratio between obs s_2 and expmin s_2, more s_1 approaches the minimum value ($\min s_1$). In other words, higher s_2 rates occupy more downlink bandwidth (b_2 in Fig. 2), so c_2 rate should be reduced to avoid congestion. The shaper alone does not guarantee congestion avoidance. Consumers must be noticed when their interest rates are too fast, otherwise there are too much packet losses. Therefore, each dropped packet in the shaper queue generates a negative acknowledgment (NACK) message to the consumer. Thus, it reduces the interest rate to match the instant shaping rate. Results [11] show that this solution achieves high bandwidth use for both downlink and uplink while avoiding congestion conflicts, thus, minimizing packet loss. In Section III, this solution is extended to provide QoS, maintaining the advantages of congestion control while offering the user different QoS options.

IV. Implementation of QOS in NDN

In the control congestion solution [11], only interest packets are shape-able: content packets skip the shaper queue. Shaping only interest packets is effective against traffic congestion because controlling the shaping rate means an indirect control of the returning content rate. Notice that this is only valid in NDN due to Property 1. Our solution also assumes that Property 1 holds true.

A. Priority Levels

In order to offer QoS, our solution offers 4 *priority* levels. Applications can pro-grammatically choose the priority of the packets it sends, as needed:

- Priority 0: This priority is assigned to background services that have no need of shorter transmission delays neither higher data throughput.
- Priority 1: Best effort. Should have a better quality of service than priority 0.
- Priority 2: Excellent effort. Assigned to critical applications that need better quality but that cannot interfere with priority 3 services.
- Priority 3: assigned to real-time services (e.g. streaming) or network control services.

Having only 4 priority levels makes the classification meaningful. In other words, that the difference of treatment (difference of QoS) between one priority level and the next one is significant. In Section IV, we show that the difference of treatment given to different priorities is configurable. Therefore, if in practice one observes that priority 2 services should have more throughput than it had in a previous time, it is just a matter of updating the router configurations. The priority levels must be included in the header of interest packets. NDN headers are currently defined as a sequence of Type-Length-Value definitions. Interest packets have special types, such as Nonce, InterestLifeTime and Selectors. We propose to include the priority levels as a new special type (using one of many available type values). Remember that contents are not shapeable, thus only half of the packets need this extra parameter. Additionally, packets without this extra field might be considered as having priority 0, which ultimately reduce the overhead.

B. Implementation

Regarding the implementation of the priority levels, the concept of congestion control with interest shaping is reused. Instead of implementing a single shaper (and a single queue) for each interface, 4 sets of shapers and queues must be present. Each set is assigned to a different priority level, as shown in Fig. 3. This solution concerns only outgoing packets. Incoming packets are not shapeable as in [original]. Also, content packets arriving from other interfaces are directly placed at the layer-2 queue (The below layer, e.g. Ethernet). On the other hand, interest packets are given to a distributor. The distributor places the interest packets on the queue corresponding to the priority level set on the packet header. Queue sizes are set to a maximum value. The queue management policy used in our implementations is drop-tail, meaning that if the queue size is below the maximum value, it pushes the packet into it, otherwise it drops it. When a drop occurs, a NACK is sent back to the source, in order to reduce the consumer throughput. Interest packets placed on the queue wait to be served according to their sub-shaping rate. Each queue has a different sub-shaper rate, named hereafter ss_0, ss_1, ss_2 and ss_3 (ss_x is the shaping rate of priority level x). In order to choose the optimal sub-shaper rates, first the main shaping rate s is calculated from Eq. 1, exactly as described in [11]. Once s is defined, the sub-shaping rates are calculated using Eq. 2:

$$s = ss_0 + ss_1 + ss_2 + ss_3 \tag{2}$$

As it can be seen from Eq. 2, the sum of sub-shaping rates is equal to the main shaping rate. It means that, even with multiple priority queues, the main shaping rate targets optimal value for congestion control.

In order to differentiate the priority levels, giving more throughput to higher levels, ss_0 , ss_1 , ss_2 and ss_3 must be chosen accordingly to Eq. 3:

$$ss_3 > ss_2 > ss_1 > ss_0 \tag{3}$$

With only Eq. 2 and Eq. 3, the sub-shaping rates cannot be calculated. Thus, we correlate all the sub-shaping rates using Eqs. 4-7:

$$ss_3 = w_3 \cdot s \tag{4}$$

$$ss_2 = w_2 \cdot s \tag{5}$$

$$ss_1 = w_1 \cdot s \tag{6}$$

$$ss_0 = w_0 \cdot s \tag{7}$$

Where w_x is called weight. From Eqs. 4-7, one can deduce that the sum of weights result in 1, as shown in Eq. 8:

$$1 = w_0 + w_1 + w_2 + w_3 \tag{8}$$

It means that each weight is the percentage of the main shaping rate allocated for that sub-shaping rate. Therefore, choosing the weights allows the network engineer to define how much throughput will be given to each priority level. It must be noticed that w_0 has actually fixed value, once the other weights are chosen (due to Eq. 7).

It is not the purpose of this paper to discuss the better configurations of weights. However, Eq. 3 forces higher priority levels to have higher sub-shaping rates, guarantying higher QoS to these services. In Section IV, results show that our solution works with different choices of weights. Additionally, it is easy to use this methodology to define any number of priorities, if required. In this paper, we only consider 4 priority levels.

C. Dynamic weights

The sum of weights shown in Eq. 8 expects that all sub-shaper queues have packets, i.e. packets tagged with all priorities are being sent through that interface. Therefore, all the weights will have values different than 0 and will thus have a share on the main shaping rate. However, there are cases where not all priority levels are going to be present.

For instance, assume that only services with priority level 0 are using a specific router. Assume also that the network engineer has set w_1 to 2, w_2 to 4 and w_3 to 8. In this case, $N+1$ level services would have twice throughput as N services.

In that case, if we keep the sharing of the main shaping rate as shown in Eqs. 4-7, those priority level 0 services would use only 6.66% of the main shaping rate. Since other priority level services are not present, there would be a waste of bandwidth. In order to always make optimal bandwidth usage, Eqs. 4-7 should be slightly changed to Eqs. 9-12:

$$\alpha_0 = w_0 \text{ if } Q_0 \text{ is not empty, otherwise } 0 \tag{9}$$

$$\alpha_1 = w_1 \text{ if } Q_1 \text{ is not empty, otherwise } 0 \tag{10}$$

$$\alpha_2 = w_2 \text{ if } Q_2 \text{ is not empty, otherwise } 0 \tag{11}$$

$$\alpha_3 = w_3 \text{ if } Q_3 \text{ is not empty, otherwise } 0 \tag{12}$$

Which can be applied to Eqs 13-17:

$$1 = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 \tag{13}$$

$$ss_0 = \alpha_0 \cdot s \tag{14}$$

$$ss_1 = \alpha_1 \cdot s \tag{15}$$

$$ss_2 = \alpha_2 \cdot ss_0 \tag{16}$$

$$ss_3 = \alpha_3 \cdot ss_0 \tag{17}$$

According to Eqs 9-17, the main shaping rate is now shared only between priority level services that are alive. For instance, if only priority level 0 service is alive, Q_1, Q_2 and Q_3 will be empty, thus a_1, a_2 and a_3 should be equal to 0, while a_0 should be equal to 1. Therefore, w_0 is equal to 1 (100%), meaning that ss_0 is equal to s itself. In other words, priority level 0 services take the whole throughput.

It must be noticed that with four queues, and Q_0, Q_1, Q_2 and Q_3 assuming two possible states (empty or not empty), there are 16 possible cases where our solution theoretically makes optimal use of bandwidth. For those our solution theoretically makes optimal use of bandwidth.

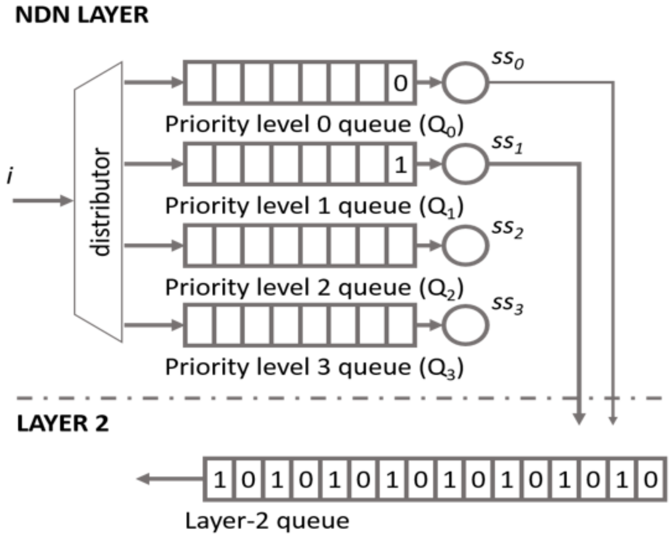


Fig. 4: Some scenarios may cause instability

D. Stabilization

As described in the previous subsection, the sub-shaping rates can assume different values dynamically. However with rates changing instantly, the system may reach undesirable states due to instability.

For instance, suppose that the weights are set to w_1 to 2, w_2 to 4 and w_3 to 8. Suppose also that consumers are sending only packets with priority level 0 and 1. Thus, sub-shaping rate ss_0 should be 33% of the main shaping rate and ss_1 should be the double: 66% (from Eqs. 9-17). ss_2 and ss_3 would be equal to zero. Now suppose that the interest rates are equal to the shaping rates and that each queue has only one packet, as shown in Fig. 4.

When the Q_0 is fed at the same rate as ss_0 rate, the queue size oscillates from 0 to 1, never reaching more than 1. In other words, the shaper consumes the same number of interest packets sent by the consumer. The same behavior occurs to Q_1. Simulations showed that this scenario may quickly become one of the below scenarios:

1) Q_0 size equal to 1 and Q_1 size equal to 1. In this case, the sub-shaping rates assume the expected values.
2) Q_0 size equal to 1 and Q_1 size equal to 0. It happens since ss_1 is faster than ss_0, Q_1 may be served before Q_0. In this case, Q_0 will have 100% of the main shaping rate.
3) Q_0 size equal to 0 and Q_1 size equal to 1. In this case, Q_1 will have 100% of the main shaping rate.

It is clear that scenarios 2 and 3 lead to incorrect shaping rates, violating Eqs.4-7. Additionally, Scenario 2 violates Eq. 3. If these scenarios occur frequently, in average, the sub-shaping rates ss_0 and ss_1 will be 50% of main shaping rate each, leading to an undesired QoS.

In order to add stability to this system, we consider that a queue is empty only when it has been empty after a while. It

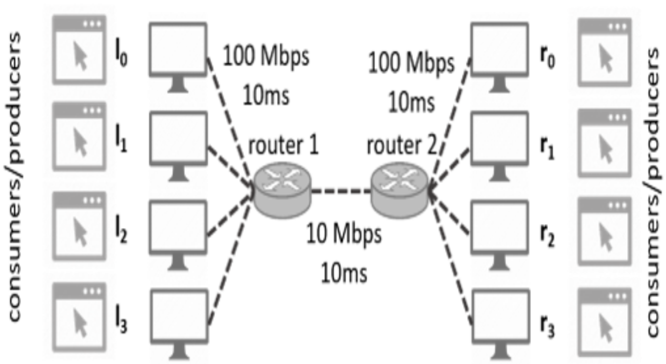


Fig. 5: Dumbbell topology

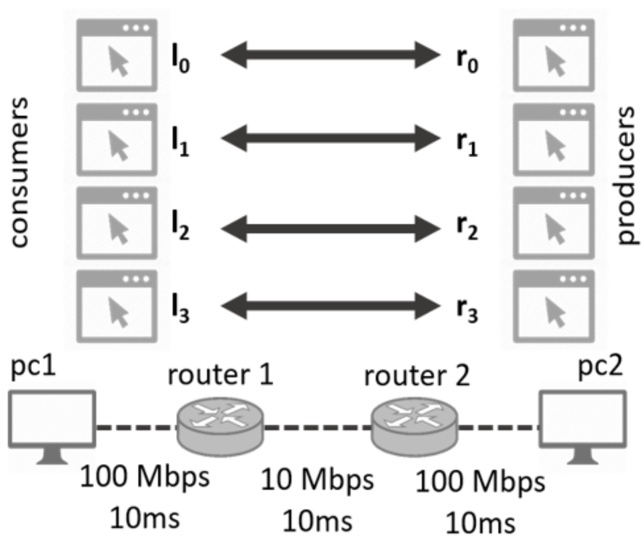


Fig. 6: Baseline topology

implies that the sub-shaping rates resist changing to transitional states, improving stability. It must be noticed that, this delay does not apply to the inverse logic: empty queues that are filled with packets are instantly served accordingly with its sub-shaping rate. For instance, in the example of Fig. 5, if a service with priority level 3 arrives, the system automatically shares the main rate among ss_0, ss_1 and ss_3 (which assume the values of 14.28%, 28.57% and 57.14%).

E. Software implementation

This solution was implemented using the ndn-extension of ns-3 simulator. It is an event-driven simulator where events such as packet sending can be scheduled programmatically at the desired simulation time. Each event triggers a function that treats it when in occurs.

In order to code a shaper, we reuse the code presented at [20]. Then, each time an outgoing packet arrives at a NDN device interface, the distributor places it on the right queue. If that queue is empty, a send event for that queue is scheduled.

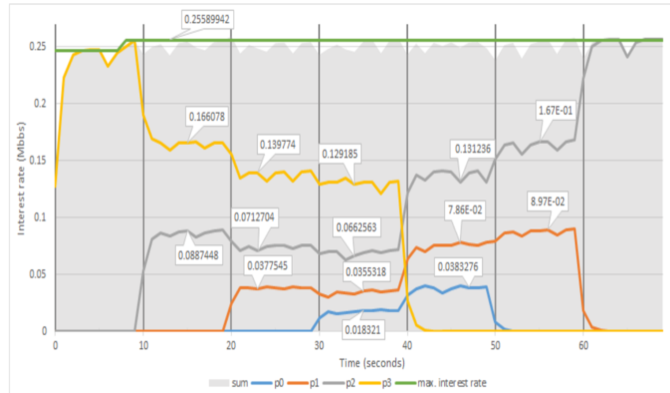


Fig. 7: Interest rate vs time (scenario 1)

Each time a send event triggered, the main shaping rate is calculated as well as all sub-shaping rates. Therefore, update frequency of s and ss_x is equal to the shaping rates.

For the incoming packets arriving at the interface, the values of obs s_2 and $\mathbb{E}(\min s_2)$ are updated as well. It means that at simulation all values are updated instantly.

V. Simulation Results

Two main topologies were defined in order to extract the results: a baseline topology (see Fig. 6) and a dumbbell topology (see Fig. 7). The baseline topology consists of two routers (router 1 and router 2) and two computers (pc1 and pc2). The link between routers has a 10 Mbps bandwidth while the other links have 100 Mbps bandwidths. All links have a 10ms delay. Up in the application layer, pc1 has four consumer applications (named p0, p1, p2 and p3) with priority levels 0, 1, 2 and 3. All consumers use the AIMD (Additive-Increase, Multiplicative-Decrease) algorithm, due to its congestion avoidance nature. Each consumer searches for a content prefix defined at a different producer (the prefixes are /p0, /p1, /p2 and /p3).

The second topology also contains two routers, but it contains four computers at each side. Each computer has a consumer application running at a unique priority level. l_i is the application running on the left side while r_i is the application running on the right side, where i also represents the priority level. l_i consumes data from r_i, and r_i consumes data from l_i (i goes from 0 to 3) For all scenarios (except explicitly said otherwise), the weights are set as: w_3 is equal to 8, w_2 is equal to 4 and w_1 is equal to 2. In this case, the sub-shaping rate for packets with priority level $N+1$ is twice as fast as the sub-shaping rate for packets with priority level N. If services from all priority levels are alive, the sub-shaping rates are distributed as show in Eq. 18-21:

$$ss_3 = s \times 0.5333 \tag{18}$$

$$ss_2 = s \times 0.2666 \tag{19}$$

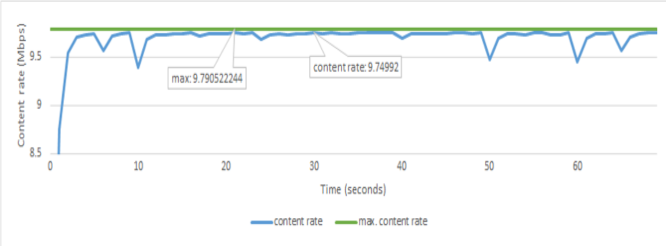


Fig. 8: Content rate vs time (scenario 1)

$$ss_1 = s \times 0.1333 \tag{20}$$

$$ss_0 = s \times 0.0666 \tag{21}$$

To extract the results, an ndnSIM tracer is used. It shows periodically the exponentially weighted moving average (EWMA) of all interests and contents flowing through each interface. This module is modified so it also shows QoS information. For all simulations, the main shaping rate was multiplied by 0.98 (named headroom) to make 2% room for unattended bursts. Additionally, all simulations were repeated at least 10 times to obtain deviation information. All shaper queues and layer-2 queue were set to have at most 60 packets. In the next subsections, we present simulation results for several scenarios. The scenario described in Subsections A, B, C and D use the baseline topology while the others use the dumbbell topology. The communication for scenario 1 is unidirectional, meaning the left side applications are only consumers. All the other scenarios use bidirectional communication, meaning that all apps are both consumer and producers.

A. Scenario 1: Priorities dynamics

The first scenario aims at simulating a baseline topology, exactly as shown at Fig. 6. All producers payloads are set to 1000B. Contents have a 32B header (leading to 1032-sized packets). Consumers send interest packets with an average size of 26B. Applications p0, p1, p2 and p3 start and stop operating at different times, in order to show the shapers dynamics. From time 0s to 10s only application p3 is alive. P_3 is later shut down at time 40s. p2 is turned on from time 10s until 70s. P_1 is turned on from time 20s until 60s. P_0 is alive from time 30s until 50s. Fig. 7 shows the sub-shaping rates for all four applications during a 70-second simulation. The green lines represent the maximum achievable rates for interest and contents. All throughputs are traced at the NDN layer, meaning that the bandwidth at this layer is not the full 10 Mbps (due to the overhead of link layer). The maximum throughput at NDN layer can be calculated from Eq. 22:

$$\text{Link layer throughput efficiency} \times \text{bandwidth} \times \text{headroom} \tag{22}$$

Where the link layer adds only 2 extra header bytes. Therefore, for content packets (1032 bytes long) the maximum content rate is 9.7810 $Mbps$. Since the prefixes increase with time (e.g. from /p0/1, /p0/2, , to /p0/10) the size of the header of both interest and content packets get larger. For 1033-byte long content packets, the content rate is 9.7905 $Mbps$. In order to generate these content rates, the maximum interest rates are 0.2464 $Mbps$ (for 26 B) and 0.2558 $Mbps$ (for 27 B), obtained from Eq. 23:

$$\frac{\text{max.content rate} \times \text{interest packet size}}{\text{content packet size}} \tag{23}$$

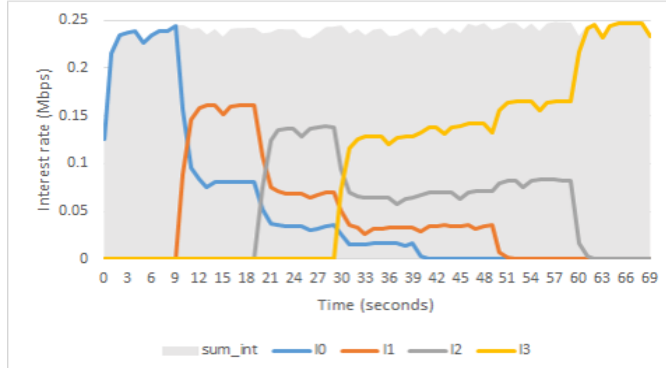


Fig. 9: Interest rate vs time for left side (scenario 5)

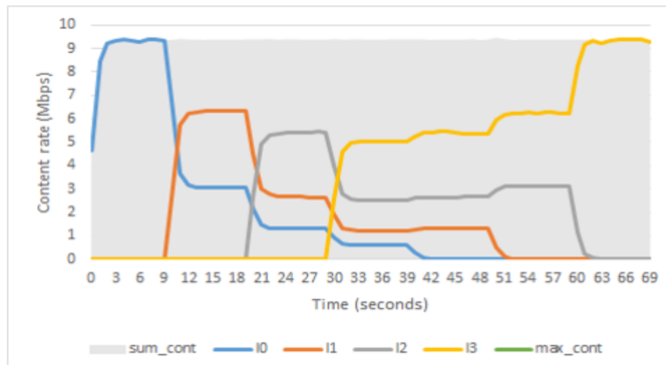


Fig. 10: Content rate vs time for left side (scenario 5)

The sum of sub-shaping rates is shown as the gray area. As it can be seen from Fig. 7, the sum of rates approaches the maximum interest rates described earlier (green line). Fig. 8 shows the returning content rate. The incoming content rate (blue line) also approaches the maximum content rate (green line), showing that our solution achieves optimal bandwidth usage.

As shown in Fig. 7 and Fig. 8, the saw behavior of rates is due to the AIMD consumers. Each time that a shaper queue drops a packet, a NACK is sent to the consumer, which causes a multiplicative decrease of its window size. Later, the consumer increases additively the window size. This decrease-increase behavior generates saw waves.

The proportions of sub-shaping rates in relation to the sum of interest rates (in Fig, 7) respect the configurations specified

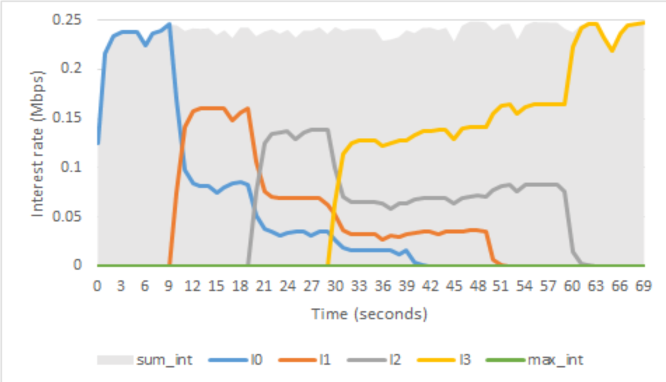


Fig. 11: Interest rate vs time for left side (scenario 6)

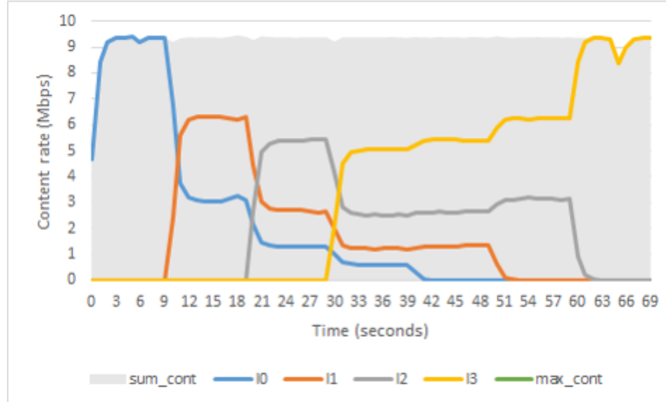


Fig. 12: Content rate vs time for left side (scenario 6)

at Section IV. For instance, at time 35s (when all applications are alive), p3 obtains 0.1291 $Mbps$ out of 0.25 $Mbps$, which gives 53.25% as expected from Eq.18. P_2, P_1 and P_0 have rates equal to 0.066 $Mbps$, 0.035 $Mbps$ and 0.018 $Mbps$, whose percentages are 26.66%, 13.33%, 6.25%, also matching Eqs. 19-21.

It must be noticed that at transition times (i.e. in every seconds), the content rates may have valleys (Fig. 8). It may happen because when one rate decreases and another one increases, the decrease is multiplicative while the increase is additive. Thus, the sum of rates tends to decrease faster than increase until it reaches a stable state.

B. Scenario 2: Randomized packet size

In this scenario, the baseline topology is used. However, the consumers reply with random payloads, instead of using a fixed-size (1000 bytes). The payloads vary randomly from 600 bytes to 1400 bytes. In addition, the four applications are alive during the whole simulation time (20 seconds). In this and for all the subsequent scenarios the transmission is bidirectional. In other words, the computer on the right side has also four consumer applications (r_0, r_1, r_2 and r_3) that are sending interests towards the left side (similarly to the dumbbell topology).

Table 1 shows the results for scenario 2, 3 and 4. As shown, the left side applications (l_0 to l_3) respect the sharing of the main shaping rate (represented by the sum row) specified at Eqs. 18-21. The same occurs for right side applications.

As shown in Table 1, the actual content rate for this scenario (for both left and right sides) matches the maximum rate. Since the maximum content rate depends on the protocol efficiency, which depends on the payload size, this maximum value is slightly different from the scenario 1. It must be calculated from the average payload is 1000B, we calculate those maximum values from Eqs 22-23.

It must be noticed that even when both sides are sending interests the system achieves optimal bandwidth usage respecting the priority weights.

C. Scenario 3: Asymmetric Size Ratio

In this scenario, the payload is also modified: producers on the right side send 500B-payload packets while producers on the left side still send 1000B-payload packets. Table 1 shows the interest and content rates for both sides.

The results show that the proportions of sub-shaping rates in comparison with the main shaping rate (sum) respect Eqs. 18-21. The content rate also reaches optimal use (almost equal to maximum value). The 1000B maximum is the same as the one calculated in Scenario 2, where the average payload has 1000 bytes. However, the maximum rate for the 500-bytes payload size is slightly different, because the protocol overhead charges more on smaller payloads.

In addition to that, the interest rate for left side has increased. This behavior is expected because reducing the payload to 500B (in average, half of the previous scenario) allows the double of packets to flow in that link. It implies thus in almost doubling the shaping rate (as shown in Table 1). Additionally, the maximum values here are the ones that appear in [11].

D. Scenario 4: Asymmetric Link Bandwidth

The purpose of this scenario is to validate the proposed technique on routers with asymmetric bandwidth, meaning that downlink and uplink have different throughputs. In this case, baseline topology with bidirectional communication is used (4 apps on each side). The link from router 1 to router 2 is set to 1Mbps while the reverse link is set to 10Mbps. Table 1 also present the results for this case. As we can see:

The 10Mbps link present similar results to the random packet size scenario. However, the content rate sum is greater than previous cases because the interest rates for that link occupy less bandwidth (only 0.0173 Mbps) since their contents are on the reverse 1Mbps link. In addition, the content rate sum matches with the value presented in [11].

The reverse link is under 1Mbps and its sum corresponds to the value presented at [11]. It is important to notice that even for the smaller bandwidth the priority levels are respected.

TABLE I: Interest and Content rates (in Mbps) for scenarios 2, 3 and 4

	Senarios					
	Random packet size		Asymmetric size ratio		Asymmetric link bandwidth	
	Interest rate	Content rate	Interest rate	Content rate	Interest rate	Content rate
l_0	0.0170 ± 0.0002	0.6820 ± 0.0171	0.0331 ± 0.0003	0.6717 ± 0.0087	0.0006 ± 0.001	0.0158 ± 0.023
l_1	0.0340 ± 0.0003	1.3341 ± 0.0242	0.0661 ± 0.0007	1.3382 ± 0.016	0.002 ± 0.0011	0.0904 ± 0.0248
l_2	0.0679 ± 0.0007	2.6688 ± 0.0348	0.1313 ± 0.0021	2.6429 ± 0.0277	0.0047 ± 0.0003	0.1935 ± 0.0095
l_3	0.1213 ± 0.0024	4.7553 ± 0.0696	0.1752 ± 0.0084	3.5001 ± 0.1381	0.0098 ± 0.0001	0.3958 ± 0.0044
Sum	0.2404 ± 0.0031	9.4404 ± 0.0785	0.4058 ± 0.0112	8.153 ± 0.1617	0.0173 ± 0.001	0.6956 ± 0.0295
Max	0.2479 (from Eq.23)	9.52 (from Eq. 22)	0.39 (from Eq. 23)	8.34 (from Eq. 22)	0.017 (from eq. 23)	0.7195 (from [3])
r_0	0.0171 ± 0.0001	0.6673 ± 0.0149	0.0166 ± 0.0002	0.6619 ± 0.0089	0.0174 ± 0.0002	0.6911 ± 0.0059
r_1	0.0341 ± 0.0003	1.3465 ± 0.0247	0.0332 ± 0.0002	1.3064 ± 0.0144	0.0347 ± 0.0003	1.3669 ± 0.019
r_2	0.0681 ± 0.0006	2.6728 ± 0.0252	0.066 ± 0.0005	2.5949 ± 0.0251	0.0631 ± 0.0021	2.4813 ± 0.084
r_3	0.1224 ± 0.0011	4.7852 ± 0.0461	0.1214 ± 0.0009	4.7506 ± 0.0328	0.1329 ± 0.0033	5.1872 ± 0.074
Sum	0.2416 ± 0.0020	9.4718 ± 0.0254	0.2374 ± 0.0015	9.3139 ± 0.0122	0.2483 ± 0.0021	9.7267 ± 0.0034
Max	0.2479 (from Eq.23)	9.52 (from Eq. 22)	0.236 (from Eq. 23)	9.3736 (from [3])	0.246 (from Eq. 23)	9.7744 (from [3])

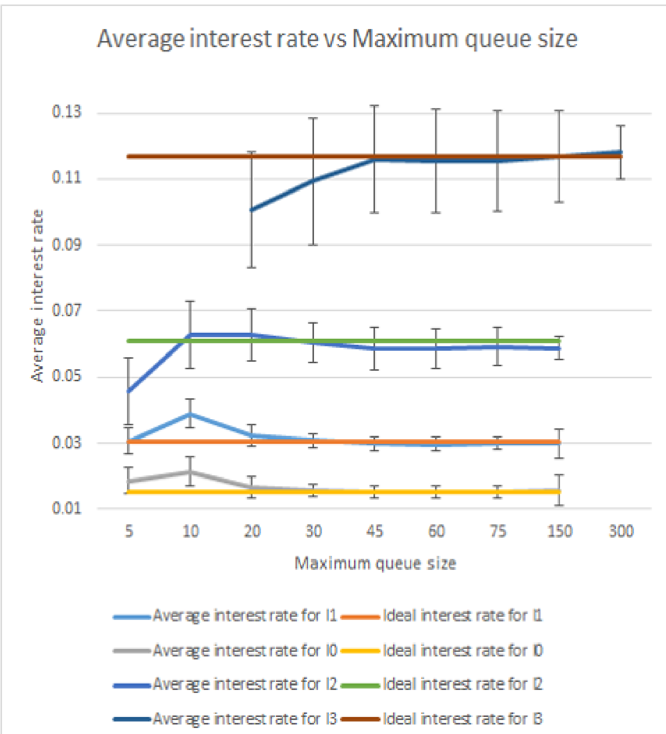


Fig. 13: Average interest rate vs Maximum queue size (scenario 6)

E. Scenario 5: Homogeneous RTT

From this scenario on, we use the dumbbell topology. Applications have different lifetimes and we change start and stop times using a 10 seconds step. Fig. 11 and Fig 12 show the interest and content rate for the applications on the right side (left side applications are omitted due to similarity of results).

As shown in Fig. 9, l_3 starts at 0s and stop at 50s. l_2 starts at 20s and live through the rest of simulation time. l_1 lives from 10s to 30s while l_0 lives from 10s to 40s. Since interest rates are smaller than content rates, oscillations are more visible at Fig. 12. As expected, even in the dumbbell topology, the sub-shaping rates have the right proportions.

TABLE II: Interest and content rates for scenario 7

	Interest Rate	Content Rate
l_0	0.0253 ± 0.0003	0.9986 ± 0.0205
l_1	0.0506 ± 0.0003	1.9892 ± 0.0127
l_2	0.0761 ± 0.0004	2.9915 ± 0.0189
l_3	0.0892 ± 0.0008	3.5085 ± 0.0326
Sum	0.2413 ± 0.0003	9.4879 ± 0.0038
Max	0.2479 (from Table 1)	9.5231 (from Table 1)
r_0	0.0253 ± 0.0003	0.9986 ± 0.0205
r_1	0.0506 ± 0.0003	1.9892 ± 0.0127
r_2	0.0761 ± 0.0004	2.9915 ± 0.0189
r_3	0.0892 ± 0.0008	3.5085 ± 0.0326
Sum	0.2413 ± 0.0003	9.4879 ± 0.0038
Max	0.2479 (from Table 1)	9.5231 (from Table 1)

F. Scenario 6: Heterogeneous RTT

This scenario aims at testing our solution in heterogeneous round-trip times (RTT). For that, links connected to computers that host l_1 and l_3 are now set 20ms instead of 10ms. Fig. 13. and Fig. 14 show the interest and content rates for the right side (remember that right side applications are communicating with those in left side).

As shown in both figures, the behavior is similar to the ones in scenario 5, demonstrating that heterogeneous RTT have no impact on this solution. It must be noticed that without the stabilization technique described in Section III, results would show an unstable system.

G. Scenario 7: Custom priority

In scenario 7, we apply other values for w_1, w_2 and w_3 in order to test the flexibility of our solution. In the results presented Table 2, w_1 is equal to 2, w_2 is equal to 3 and w_3 is equal to 4, reducing the ratios used in previous scenarios.

For both sides, the interest and content rates respect the proportion 1 to 2 to 3 to 4. It must be noticed that even with other weights, the achieved rates correspond to the ones show in Table 1.

H. Scenario 8: Queue sizes

The last scenario aims at evaluating the ideal queue size for all priority levels. Fig. 23 shows the average interest rate vs maximum queue size for priority levels 0, 1 and 2; while Fig. 24 show it for priority level 3. The deviations are the bars at some points in the line. The ideal interest rate for each priority level is also shown (the ideal rate is the rate predicted from the Table 1, considering any functional scenario, for instance scenario 2).

Shorter queues (below size 20) do not reach the ideal rate. This is especially visible for l_2 and l_3, since they usually have higher rates. This effect happens due to the fact that shorter queues are more prone to send NACKs (they can easily drop packets).

On the other hand, the deviation bars show that for lower priorities (l_0 and l_1), larger queues (150) tend to have more deviation for the interest rates. It happens mainly when the sub-shaping rates are transitioning (as Fig. 19): longer queues will take more time to send NACKs and therefore the AIMD consumer will synchronize later.

Higher priority services always get faster shaping rates, which produce NACKs more easily. Thus, the higher deviations appear for p3.

VI. Conclusion

In this paper, we presented a mechanism that offers QoS on top of a congestion control solution. The hop-by-hop interest shaper is improved to use 4 sub-shapers in order to provide differentiated QoS to 4 priority levels. Results showed that higher priority applications have higher throughputs, while lower priority applications never starve. Our solution reacts instantly and dynamically to any configuration of requests, always using optimal bandwidth (even in cases where all not all priorities exist). Several simulations with complex scenarios have emphasized the validity of the proposed method.

References

[1] Abdel-Hadi, A., and Clancy, C. (2014, February). A utility proportional fairness approach for resource allocation in 4G-LTE. In Computing, Networking and Communications (ICNC), 2014 International Conference on (pp. 1034-1040). IEEE.

[2] Abdel-Hadi, A., and Clancy, C. (2013, September). A robust optimal rate allocation algorithm and pricing policy for hybrid traffic in 4G-LTE. In Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on (pp. 2185-2190). IEEE.

[3] author = Ingo Ltkebohle Itu, R. (2008). ITU-R M. 2135: Guidelines for evaluation of radio interface technologies for IMT-Advanced.

[4] Oh, J., and Han, Y. (2012, September). Cell selection for range expansion with almost blank subframe in heterogeneous networks. In Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on (pp. 653-657). IEEE.

[5] Richter, F., Fehske, A. J., and Fettweis, G. P. (2009, September). Energy efficiency aspects of base station deployment strategies for cellular networks. In Vehicular Technology Conference Fall (VTC 2009-Fall), 2009 IEEE 70th (pp. 1-5). IEEE.

[6] Wireless Systems Research Lab, Hitachi America Ltd, Cellular Network Topology Toolbox Implementation Documentation, http://www.ltehetnet.com/sites/default/files/NetworkToolBox_Documentation.pdf,2014.

[7] Chinipardaz, M., Rasti, M., and Nourhosseini, M. (2014, September). An overview of cell association in heterogeneous network: Load balancing and interference management perspective. In Telecommunications (IST), 2014 7th International Symposium on (pp. 1250-1256). IEEE.

[8] Ghosh, A., Mangalvedhe, N., Ratasuk, R., Mondal, B., Cudak, M., Visotsky, E., and Dhillon, H. S. (2012). Heterogeneous cellular networks: From theory to practice. Communications Magazine, IEEE, 50(6), 54-64.

[9] Fehske, A., Fettweis, G., Malmodin, J., and Biczok, G. (2011). The global footprint of mobile communications: The ecological and economic perspective. Communications Magazine, IEEE, 49(8), 55-62.

[10] Blume, O., Eckhardt, H., Klein, S., Kuehn, E., and Wajda, W. M. (2010). Energy savings in mobile networks based on adaptation to traffic statistics. Bell Labs Technical Journal, 15(2), 77-94.

[11] Wang, Y., Rozhnova, N., Narayanan, A., Oran, D., and Rhee, I. (2013, August). An improved hop-by-hop interest shaper for congestion control in named data networking. In ACM SIGCOMM Computer Communication Review (Vol. 43, No. 4, pp. 55-60). ACM.

[12] Zhang, L., Afanasyev, A., Burke, J., Jacobson, V., Crowley, P., Papadopoulos, C., and Zhang, B. (2014). Named data networking. ACM SIGCOMM Computer Communication Review, 44(3), 66-73. Chicago

[13] Mastorakis, S., Afanasyev, A., Moiseenko, I., and Zhang, L. (2015). ndnSIM 2.0: A new version of the NDN simulator for NS-3. Technical Report NDN-0028, NDN. Chicago

[14] Wu, T. Y., Lee, W. T., Duan, C. Y., and Wu, Y. W. (2014). Data Lifetime Enhancement for Improving QoS in NDN. Procedia Computer Science, 32, 69-76.

[15] Oueslati, S., Roberts, J., and Sbihi, N. (2012, March). Flow-aware traffic control for a content-centric network. In INFOCOM, 2012 Proceedings IEEE (pp. 2417-2425). IEEE.

[16] Carofiglio, G., Gallo, M., and Muscariello, L. (2012, August). Joint hop-by-hop and receiver-driven interest control protocol for content-centric networks. In Proceedings of the second edition of the ICN workshop on Information-centric networking (pp. 37-42). ACM. Chicago

[17] Floyd, S. (1994). TCP and explicit congestion notification. ACM SIGCOMM Computer Communication Review, 24(5), 8-23.

[18] Moret, Y., and Fdida, S. (1997, November). ERAQLES an efficient explicit rate algorithm for ABR. In Global Telecommunications Conference, 1997. GLOBECOM'97., IEEE (Vol. 2, pp. 801-805). IEEE. Chicago

[19] Zhang, F., Zhang, Y., Reznik, A., Liu, H., Qian, C., and Xu, C. (2014, August). A transport protocol for content-centric networking with explicit congestion control. In Computer Communication and Networks (ICCCN), 2014 23rd International Conference on (pp. 1-8). IEEE. Chicago

[20] https://github.com/wygivan/ndnSIM

[21] https://github.com/wygivan/ndnSIM

[22] Zhang, L., Estrin, D., Burke, J., Jacobson, V., Thornton, J. D., Smetters, D. K., and Abdelzaher, T. (2010). Named data networking (ndn) project. Relatrio Tcnico NDN-0001, Xerox Palo Alto Research Center-PARC. Chicago

[23] Jacobson, V., Smetters, D. K., Thornton, J. D., Plass, M. F., Briggs, N. H., and Braynard, R. L. (2009, December). Networking named content. In Proceedings of the 5th international conference on Emerging networking experiments and technologies (pp. 1-12). ACM. Chicago

[24] Saino, L., Cocora, C., and Pavlou, G. (2013, June). Cctcp: A scalable receiver-driven congestion control protocol for content centric networking. In Communications (ICC), 2013 IEEE International Conference on (pp. 3775-3780). IEEE.

[25] Yi, C., Afanasyev, A., Wang, L., Zhang, B., and Zhang, L. (2012). Adaptive forwarding in named data networking. ACM SIGCOMM computer communication review, 42(3), 62-67.

[26] http://blogs.parc.com/ccnx/

[27] Jiang, X., and Bi, J. (2014, April). nCDN: CDN enhanced with NDN. In Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on (pp. 440-445). IEEE. Chicago

[28] Zhou, J., Wu, Q., Li, Z., Kaafar, M. A., and Xie, G. (2015, June). A proactive transport mechanism with Explicit Congestion Notification for NDN. In Communications (ICC), 2015 IEEE International Conference on (pp. 5242-5247). IEEE.

An Algorithmic approach for abstracting transient states in timed systems

Mohammed Achkari Begdouri, Houda Bel Mokadem and Mohamed El Haddad
Labtic – Ecole Nationale des Sciences Appliqu´ees de Tanger,
BP 1818 Tanger Principal, Tangier, Morrocco

***Abstract*—In previous works, the timed logic TCTL was extended with importants modalities, in order to abstract *transient* states that last for less than k time units. For all modalities of this extension, called TCTL^{Δ}, the decidability of the model-checking problem has been proved with an appropriate extension of Alur and Dill's region graph. But this theoretical result does not support a natural implementation due to its state-space explosion problem. This is not surprising since, even for TCTL timed logics, the model checking algorithm that is implemented in tools like UPPAAL or KRONOS is based on a so-called zone algorithm and data structures like DBMs, rather than on explicit sets of regions.**

In this paper, we propose a symbolic model-checking algorithm which computes the characteristic sets of some TCTL^{Δ} formulae and checks their truth values. This algorithm generalizes the zone algorithm for TCTL timed logics. We also present a complete correctness proof of this algorithm, and we describe its implementation using the DBM data structure.

Keywords: *Timed automata, symbolic model checking, back-ward analysis algorithm, correctness, data structures.*

I. Introduction

Timed verification. Temporal logic is a convenient formalism for specifying systems and reasoning about them. Furthermore, model-checking techniques lead to the automatic verification that a model of a system satisfies some temporal logic specification. These methods have been extended to real-time verification: systems are modeled with timed automata [6], [7] and timed logics like TCTL [3] are used to express timed specification like “any problem is followed by an alarm within 3 seconds”. Analysis tools have been developped [22], [25], [30] and successfully applied to numerous case studies.

Timed temporal logics and duration properties. Along with the study of timed automata, various timed logics have been defined to extend the classical temporal logics with quantitative modalities. For example, this was done with MTL [29], [8], [31], an extension of LTL, and TCTL [9], [3], [26], where CTL modalities are augmented with time comparisons of the form $\sim c$, where $\sim$ is a comparison operator. Another related logic is the Parametrized TCTL [18] where TCTL and the timed model are in turn extended with parameters.

In another direction, since the introduction of the *duration calculus* [19] in order to express duration properties, numerous works have been devoted to the algorithmic computation of such properties for timed systems. Since *clocks*, which evolve at the rate of time (as in timed automata), are sometimes not expressive enough, hybrid variables (with multiple slopes) have been considered. The resulting model of hybrid automata has been largely studied in the subsequent years [27]. However, while some decidability results could be obtained [5], [28], using stopwatches (*i.e.* variables with slopes 0 and 1) already leads to undecidability for the reachability problem [4].

Further research has thus been devoted to weaker models where hybrid variables are only used as *observers*, *i.e.* are not tested in the automaton and thus play no role during a computation. These variables, sometimes called costs or prices in this context can be used in an optimization criterium [5], [10], [11], [16] or as constraints in temporal logic formulas. For instance, the logic WCTL [17], [15], interpreted over timed automata extended with costs, adds cost contraints on modalities: it is possible to express that a given state is reachable within a fixed cost bound.

Abstracting transient states. When practical examples are considered, the need for abstracting transient states often happens. This is the case for systems which handle variables, subject to instantaneous changes of value. This motivated the work in [12], [13], where events that do not last continuously for at least k time units could be abstracted by introducing an extension of TCTL called TCTL^{Δ}. The theoretical decidability result of TCTL^{Δ} model-checking problem rely on an extension of the region graph proposed in [13]. However, the region graph is not used for implementation, but tools like UPPAAL or KRONOS use a so-called “zone algorithm”. This algorithm computes on-the-fly the set of reachable symbolic states, that is pairs (q, Z) where q is a control state and Z a zone. One of the major advantage of zones is that they can be easily implemented using data structures like DBMs [24].

Contribution. The aim of this paper is to provide an implementable algorithm for TCTL^{Δ} model-checking. The algorithm we propose is an extension of the zone algorithm used for TCTL timed logics in tools like UPPAAL and KRONOS. We also provide a possible implementation of this algorithm using the DBM data structure. The main result of this paper is the proof of correctness of our algorithm. This proof uses several techniques, from properties of zones and symbolic model-checking to properties of fixed point theory.

Outline. The structure of the paper is the following: we first recall the main features of timed automata model and give definitions for the syntax and semantics of TCTL^{Δ} timed logic (Section 2); we present after some known decidability results of the TCTL^{Δ} model-checking (Section 3); we then describe the classical zone algorithm for TCTL timed logics (Section 4); we present thereafter our algorithm, we give a complete proof of its correctness (Section 5) and the following section is devoted to explain how to implement it using the DBMs (Section 6); we end this paper with some concluding remarks (Section 7).

II. Basic Notions

Let $\mathbb{N}$ and $\mathbb{R}$ denote the sets of natural and non-negative real numbers, respectively. Let X be a set of real valued clocks. We write $\mathcal{C}(X)$ for the set of boolean expressions over atomic formulae of the form $x \sim k$ with $x \in X$, $k \in \mathbb{N}$, and $\sim \in \{<, \leq, =, \geq, >\}$. Constraints of $\mathcal{C}(X)$ are interpreted over *valuations* for clocks, i.e. mappings from X to $\mathbb{R}$. The set of valuations is denoted by $\mathbb{R}^X$. For every $v \in \mathbb{R}^X$ and $d \in \mathbb{R}$, we use $v + d$ to denote the time assignment which maps each clock $x \in X$ to the value $v(x) + d$. For every $r \subseteq X$, we write $v[r \leftarrow 0]$ for the valuation which maps each clock in r to the value 0 and agrees with v over $X \setminus r$. Let AP be a set of atomic propositions.

A. Timed Automata

Definition 1. *A* timed automaton *(TA) is a tuple* $A = \langle X, Q_A, q_{\text{init}}, \rightarrow_A, \mathsf{Inv}_A,\ l_A\rangle$ *where X is a finite set of clocks, Q_A is a finite set of* locations or control states *and $q_{\text{init}} \in Q_A$ is the* initial location. *The set $\rightarrow_A \subseteq Q_A \times \mathcal{C}(X) \times 2^X \times Q_A$ is a finite set of* action transitions*: for $(q, g, r, q') \in \rightarrow_A$, g is the enabling condition and r is a set of clocks to be reset with the transition (we write $q \xrightarrow{g,r}_A q'$). $\mathsf{Inv}_A \colon Q_A \rightarrow \mathcal{C}(X)$ assigns an* invariant *to each control state. Finally $l_A \colon Q_A \rightarrow 2^{\mathsf{AP}}$ labels every location with a subset of* AP.

A *state* (or *configuration*) of a TA A is a pair (q, v), where $q \in Q_A$ is the current location and $v \in \mathbb{R}^X$ is the current clock valuation. The initial state of A is (q_{init}, v_0) with $v_0(x) = 0$ for any x in X. There are two kinds of transition. From (q, v), it is possible to perform the *action transition* $q \xrightarrow{g,r}_A q'$ if $v \models g$ and $v[r \leftarrow 0] \models \mathsf{Inv}_A(q')$ and then the new configuration is $(q', v[r \leftarrow 0])$. It is also possible to let time elapse, and reach $(q, v + d)$ for some $d \in \mathbb{R}$ whenever the invariant is satisfied along the delay. Formally the semantics of a TA A is given by a Timed Transition System (TTS) $\mathcal{T}_A = (S, s_{\text{init}}, \rightarrow_{\mathcal{T}_A}, l)$ where:

- $S = \{(q, v) \mid q \in Q_A \text{ and } v \in \mathbb{R}^X \text{ s.t. } v \models \mathsf{Inv}_A(q)\}$ and $s_{\text{init}} = (q_{\text{init}}, v_0)$.
- $\rightarrow_{\mathcal{T}_A} \subseteq S \times S$ and we have $(q, v) \rightarrow_{\mathcal{T}_A} (q', v')$ iff
  - either $q' = q$, $v' = v + d$ and $v + d' \models \mathsf{Inv}_A(q)$ for any $d' \leq d$. This is a delay transition — we write $(q, v) \xrightarrow{d} (q, v + d)$ —,
  - or $\exists q \xrightarrow{g,r}_A q'$ and $v \models g$, $v' = v[r \leftarrow 0]$ and $v' \models \mathsf{Inv}_A(q')$. This is an action transition — we write $(q, v) \rightarrow_a (q', v')$.
- $l \colon S \rightarrow 2^{\mathsf{AP}}$ labels every state (q, v) with the subset $l_A(q)$ of AP .

An execution (or run) of A is an infinite path $s_0 \rightarrow_{\mathcal{T}_A} s_1 \rightarrow_{\mathcal{T}_A} s_2 \ldots$ in $\mathcal{T}_A$ such that (1) time diverges and (2) there are infinitely many action transitions. Note that an execution can be described as an alternating infinite sequence $s_0 \xrightarrow{d_1}\rightarrow_a s_1 \xrightarrow{d_2}\rightarrow_a \cdots$ for some $d_i \in \mathbb{R}$. Such an execution ρ goes through any configuration s' reachable from some s_i by a delay transition of duration $d \in [0, d_i]$. Let $\mathrm{Exec}(s)$ be the set of all executions from s. With a run $\rho : (q_0, v_0) \xrightarrow{d_1}\rightarrow_a (q_1, v_1) \xrightarrow{d_2}\rightarrow_a \ldots$ of A, we associate the sequence of absolute dates defined by $t_0 = 0$ and $t_i = \sum_{j \leq i} d_j$ for $i \geq 1$, and in the sequel, we often write ρ as the sequence $((q_i, v_i, t_i))_{i \geq 0}$.

Example 1. An example of timed automaton is given below (Fig. 1), where P is an atomic proposition and x, y are clocks.

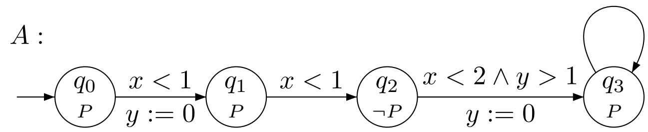


Figure 1: Example of timed automaton.

An example of run is depicted below,

$$\rho : (q_0, (0, 0)) \xrightarrow{0.1}\rightarrow_a (q_1, (0.1, 0)) \xrightarrow{0.8}\rightarrow_a (q_2, (0.9, 0.8)) \ldots$$

□

A state (q, v) can occur several times along a run ρ, the notion of *position* [1] allows us to distinguish them: every occurrence of a state is associated with a unique position. Given a position p, the corresponding state is denoted by s_p. The standard notions of prefix, suffix and subrun apply to paths in TTS: given a position $p \in \rho$, $\rho^{\leq p}$ is the prefix leading to p, $\rho^{\geq p}$ is the suffix issued from p. Finally a subrun σ from p to p' is denoted by $p \overset{\sigma}{\mapsto} p'$.

Note that the set of positions along ρ is totally ordered by $<_\rho$. Given two positions p and p', we say that p *precedes strictly* p' along ρ (written $p <_\rho p'$) iff there exists a finite subrun σ of ρ s.t. $p \overset{\sigma}{\mapsto} p'$ and σ contains at least one non null delay transition **or** one action transition (*i.e.* σ is not reduced to $\xrightarrow{0}$). We write $\sigma <_\rho p$ when for any position p' in the subrun σ, we have $p' <_\rho p$.

Given a position $p \in \rho$, the prefix $\rho^{\leq p}$ has a *duration*, $\mathsf{Time}(\rho^{\leq p})$, defined as the sum of all delays along $\rho^{\leq p}$. Since time diverges along an execution, we have: for any $t \in \mathbb{R}$, there exists $p \in \rho$ such that $\mathsf{Time}(\rho^{\leq p}) > t$.

For a subset $P \subseteq \rho$ of positions in ρ, we define a natural measure $\hat{\mu}(P) = \mu\{\mathsf{Time}(\rho^{\leq p}) \mid p \in P\}$, where μ is Lebesgue measure on the set of real numbers. In the sequel, we only use this measure when P is a subrun of ρ: in this case, for a subrun σ such that $p \overset{\sigma}{\mapsto} p'$, we simply have $\hat{\mu}(\sigma) = \mathsf{Time}(\rho^{\leq p'}) - \mathsf{Time}(\rho^{\leq p})$.

B. Definition of TCTL^Δ.

The syntax of TCTL was extended in [13] to express that a formula holds everywhere except on subruns with duration a parameter $k \in \mathbb{N}$: TCTL^Δ is obtained by adding to TCTL the modalities $\mathsf{E\_U}^k_{\sim c}\mathsf{\_}$ and $\mathsf{A\_U}^k_{\sim c}\mathsf{\_}$, where $k \in \mathbb{N}$.

Definition 2 (Syntax of TCTL^Δ)**.** TCTL^Δ *formulae are given by the following grammar:*

$$\varphi, \psi ::= P_1 \mid P_2 \mid \ldots \mid \neg\varphi \mid \varphi \wedge \psi \mid \mathsf{E}\varphi\mathsf{U}_{\sim c}\psi \mid \mathsf{A}\varphi\mathsf{U}_{\sim c}\psi \mid \mathsf{E}\varphi\mathsf{U}^k{}_{\sim c}\psi \mid \mathsf{A}\varphi\mathsf{U}^k_{\sim c}\psi$$

where $P_i \in \mathsf{AP}$, $\sim$ belongs to the set $\{<, >, \leq, \geq, =\}$ and $c, k \in \mathbb{N}$.

[1] Note that as it is possible to perform a sequence of action transitions in 0 t.u., we cannot replace the notion of positions by a function from f_ρ from $\mathbb{R}$ to S.

Standard abbreviations include $\top, \varphi \vee \psi, \varphi \Rightarrow \psi, \ldots$ as well as :

$$\begin{array}{ll} \mathsf{EF}^k_{\sim c}\,\varphi \overset{\text{def}}{=} \mathsf{E}(\top\,\mathsf{U}^k_{\sim c}\,\varphi) & \mathsf{AF}^k_{\sim c}\,\varphi \overset{\text{def}}{=} \mathsf{A}(\top\,\mathsf{U}^k_{\sim c}\,\varphi) \\ \mathsf{EG}^k_{\sim c}\,\varphi \overset{\text{def}}{=} \neg\mathsf{AF}^k_{\sim c}\neg\varphi & \mathsf{AG}^k_{\sim c}\,\varphi \overset{\text{def}}{=} \neg\mathsf{EF}^k_{\sim c}\neg\varphi \end{array}$$

Moreover U^k stands for $\mathsf{U}^k_{\geq 0}$.

Definition 3 (Semantics of TCTL^Δ). *The following clauses define when a state s of some TTS $\mathcal{T} = \langle S, s_{init}, \rightarrow, l\rangle$ satisfies a TCTL^Δ formula φ, written $s \models \varphi$, by induction over the structure of φ.*

$$\begin{array}{ll} s \models \neg\varphi & \textit{iff } s \not\models \varphi \\ s \models \varphi \wedge \psi & \textit{iff } s \models \varphi \textit{ and } s \models \psi \\ s \models \mathsf{E}\varphi\mathsf{U}_{\sim c}\psi & \textit{iff } \exists\rho \in \textit{Exec}(s) \textit{ s.t. } \rho \models \varphi\mathsf{U}_{\sim c}\psi \\ s \models \mathsf{A}\varphi\mathsf{U}_{\sim c}\psi & \textit{iff } \forall\rho \in \textit{Exec}(s) \textit{ we have } \rho \models \varphi\mathsf{U}_{\sim c}\psi \\ s \models \mathsf{E}\varphi\mathsf{U}^k_{\sim c}\psi & \textit{iff } \exists\rho \in \textit{Exec}(s) \textit{ s.t. } \rho \models \varphi\mathsf{U}^k_{\sim c}\psi \\ s \models \mathsf{A}\varphi\mathsf{U}^k_{\sim c}\psi & \textit{iff } \forall\rho \in \textit{Exec}(s) \textit{ we have } \rho \models \varphi\mathsf{U}^k_{\sim c}\psi \end{array}$$

$$\begin{array}{ll} \rho \models \varphi\mathsf{U}_{\sim c}\psi & \textit{iff } \exists p \in \rho \textit{ s.t. } \mathsf{Time}(\rho^{\leq p}) \sim c \\ & \wedge\, s_p \models \psi \,\wedge\, \forall p' <_\rho p,\ s_{p'} \models \varphi \\ \rho \models \varphi\mathsf{U}^k_{\sim c}\psi & \textit{iff there exists a subrun } \sigma \textit{ along } \rho, \\ & \textit{a position } p \in \sigma \textit{ s.t. } \mathsf{Time}(\rho^{\leq p}) \sim c \,\wedge \\ & \hat{\mu}(\sigma) > k \,\wedge\, \forall p' \in \sigma,\ s_{p'} \models \psi \textit{ and for all} \\ & \textit{subrun } \sigma' \textit{ s.t. } \sigma' <_\rho p \,\wedge\, \forall p' \in \sigma',\ s_{p'} \models \neg\varphi \\ & \textit{we have } \hat{\mu}(\sigma') \leq k \end{array}$$

Modality $\mathsf{E}\varphi\mathsf{U}^k_{\sim c}\psi$ means that it is possible to reach a sufficiently long interval $(> k)$ where ψ is true, around a position at a distance $\sim c$ and, before this position, φ is everywhere true except along negligible duration subpaths $(\leq k)$. Whereas modality $\mathsf{A}\varphi\mathsf{U}^k_{\sim c}\psi$ means that along any path, ψ lasts long enough $(> k)$ around a position at a distance $\sim c$ and, before this position, φ is everywhere true except along negligible duration subpaths $(\leq k)$.

III. Decidability result for TCTL^Δ

In this section we recall the decidability result for the TCTL^Δ model checking [13]. First, we remind that the classical notion of region proposed by Alur, Courcoubetis and Dill [3] for TCTL is also correct for TCTL^Δ. Nevertheless it needs a stronger notion of equivalence for the runs in order to preserve the truth value of TCTL^Δ formulae [13]. Then we recall that adding the modalities U^k does not increase the complexity of the verification.

A. Region graph

Given a set X of clocks and $M \in \mathbb{N}$, two valuations $v, v' \in \mathbb{R}^X$ are *M-equivalent* [3] (written $v \cong_M v'$) if:

1) for any $x \in X$ $\lfloor v(x)\rfloor = \lfloor v'(x)\rfloor$ or $(v(x) > M \wedge v'(x) > M)$,
2) for any $x, y \in X$ s.t. $v(x) \leq M$ and $v(y) \leq M$, we have: $\mathsf{frac}(v(x)) \leq \mathsf{frac}(v(y)) \Leftrightarrow \mathsf{frac}(v'(x)) \leq \mathsf{frac}(v'(y))$ and $\mathsf{frac}(v(x)) = 0 \Leftrightarrow \mathsf{frac}(v'(x)) = 0$.

An equivalence class of $\cong$ is called a *region*; and a region is called a *boundary region* if it contains valuations v s.t. the fractional part of $v(x)$ is 0, for some clock x. Given a TA A, we use M_A to denote the maximal constant occurring in A (in its guards or invariants). We write simply $\cong$ instead of $\cong_M$ when M is clear from the context.

Example 2. Consider a automaton with two clocks x and y and the constant M equal to 2. The set of regions associated with this automaton can be described by the figure beside (Fig. 2). The region drawn in gray corresponds to the valuations satisfying the following constraints:

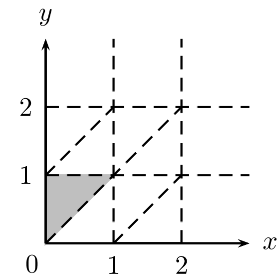


Figure 2: Example of Region.

$$0 < x < 1 \,\wedge\, 0 < y < 1 \,\wedge\, frac(y) < frac(x).$$

□

Moreover, the equivalence $\cong_{M_A}$ is consistent w.r.t. TCTL^Δ formulae [13], *i.e.* for all $\Phi \in \mathsf{TCTL}^\Delta$ and $v, v' \in \mathbb{R}^X$ s.t. $v \cong_{M_A} v'$, we have: $(q, v) \models \Phi \;\Leftrightarrow\; (q, v') \models \Phi$.

To illustrate this result, consider the formula $\Phi = \mathsf{E}\varphi\mathsf{U}^k_{\sim c}\psi$ and assume that $(q, v) \models \Phi$, *i.e.* there exists a run $\rho = ((q_i, v_i, t_i))_{i\geq 0}$ from (q, v) satisfying $\varphi\mathsf{U}^k_{\sim c}\psi$. The consistency of $\cong$ for TCTL^Δ timed logics, means that there exists an *equivalent* run ρ' from (q, v') which also satisfies $\varphi\mathsf{U}^k_{\sim c}\psi$, with v, v' are in the same region.

For this, the equivalence over runs is defined as follows [13]: Given a TA A, two runs $\rho = ((q_i, v_i, t_i))_{i\geq 0}$ and $\rho' = ((q'_i, v'_i, t'_i))_{i\geq 0}$ are equivalent (written $\rho \cong^* \rho'$) if

1) for all $i \geq 0$, $q_i = q'_i$,
2) for all $i \geq 0$, $(v_i, t_i) \cong_{M_A} (v'_i, t'_i)$,
3) for all $0 \leq j < i$, (i) $\mathsf{frac}(t_j) < \mathsf{frac}(t_i)$ iff $\mathsf{frac}(t'_j) < \mathsf{frac}(t'_i)$
and (ii) $\mathsf{frac}(t_j) = \mathsf{frac}(t_i)$ iff $\mathsf{frac}(t'_j) = \mathsf{frac}(t'_i)$.

Such that the equivalence $\cong$ is extended to pairs (v_i, t_i) as follows: $(v_i, t_i) \cong (v'_i, t'_i)$ iff (1) $v_i \cong v'_i$, (2) $\lfloor t_i\rfloor = \lfloor t'_i\rfloor$ and $\mathsf{frac}(t_i) = 0$ iff $\mathsf{frac}(t'_i) = 0$ and (3) for each clock $x \in X$, (i) $\mathsf{frac}(v_i(x)) < \mathsf{frac}(t_i)$ iff $\mathsf{frac}(v'_i(x)) < \mathsf{frac}(t'_i)$ and (ii) $\mathsf{frac}(v_i(x)) = \mathsf{frac}(t_i)$ iff $\mathsf{frac}(v'_i(x)) = \mathsf{frac}(t'_i)$.

The equivalence on runs used in [3] to prove that regions are compatible with TCTL formulae only requires conditions *(ER 1)* and *(ER 2)*. This is however not sufficient for proving the compatibility of regions with TCTL^Δ formulae. Indeed, back to the *Example 1* and consider the two following runs (Fig. 3), which are equivalent in [3]:

ρ : $(q_0, (0, 0)) \xrightarrow{0.1}\!\!\rightarrow_a (q_1, (0.1, 0)) \xrightarrow{0.8}\!\!\rightarrow_a (q_2, (0.9, 0.8)) \xrightarrow{0.3}\!\!\rightarrow_a (q_3, (1.2, 0)) \ldots$

ρ' : $(q_0, (0, 0)) \xrightarrow{0.8}\!\!\rightarrow_a (q_1, (0.8, 0)) \xrightarrow{0.1}\!\!\rightarrow_a (q_2, (0.9, 0.1)) \xrightarrow{1.05}\!\!\rightarrow_a (q_3, (1.95, 0)) \ldots$

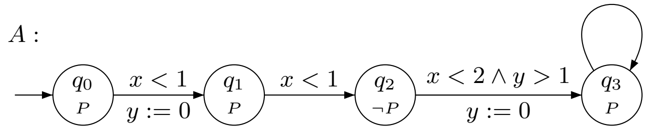


Figure 3: Example of equivalence over runs.

□

The runs ρ and ρ' satisfy conditions *(ER 1)* and *(ER 2)* but the delays spent in state q_2 where P does not hold are respectively 0.3 and 1.05, so that $\rho \models G^1 P$ whereas $\rho' \not\models G^1 P$.

This is why we need the stronger equivalence above which also requires condition *(ER 3)*. Note that the proof of the equivalence

$\cong_{M_A}$ consistency for TCTL^{Δ} timed logics is given in [13].

B. Labeling algorithm

The main result of the labeling algorithm is reducing the model-checking problem $A \models \Phi$ with a TA $A = \langle X, Q_A, q_{\text{init}}, \rightarrow_A, \mathsf{Inv}_A, l_A \rangle$ and $\Phi \in \mathsf{TCTL}^{\Delta}$, to a model-checking problem $A' \models \Phi'$ where A' is a *region graph* (*i.e.* a finite Kripke structure) and Φ' is a CTL-like formula [13].

Let X^* be the set of clocks $X \cup \{z, z_r, z_{\bar{l}}\}$. The three extra clocks are used to verify timing constraints in the formula: z is used to handle subscripts $\sim c$ in U modalities (as in TCTL model checking) and the clock $z_{\bar{l}}$ (resp z_r) is used to measure time elapsing when the left (resp. right) part in U^k modalities is false (resp true).

Let M_Φ be the maximal constant occurring in the timing constraints in Φ and k_m be the maximal k occurring in a modality U^k in Φ. Let M be $\max(M_A, M_\Phi + k_m)$.

The region graph $\mathcal{R}_{A,\Phi} = (V, \rightarrow, l, F)$ for A and Φ is defined as usual over X^* and M [3]: its set of states V is $\{(q,\gamma) | q \in Q_A \text{and} \gamma \in \mathbb{R}^{X^*}/{\cong_M}\}$, the transitions correspond to action transitions ($\rightarrow_a$) in A or delay transitions ($\rightarrow_t$, leading to the *successor region* denoted by $succ(\gamma)$). The states are labeled with atomic propositions AP and we also use additional propositions for the extra clocks: a state (q,γ) is labeled with the proposition $(\!|y \sim a|\!)$ with $y \in \{z, z_{\bar{l}}, z_r\}$ and $0 \leq a \leq M$, when $\gamma \models y \sim a$ (see [3], [12] for the detailed construction of $\mathcal{R}_{A,\Phi}$).

Labeling algorithm.: The algorithm consists in labeling the vertices of $\mathcal{R}_{A,\Phi}$ with the subformulae of Φ they satisfy, starting from the subformulae of length 1 and length 2 and so on.

Consider a formula Ψ of the form $\mathsf{E}\varphi_l \mathsf{U}^k_{\sim c} \varphi_r$ or $\mathsf{A}\varphi_l \mathsf{U}^k_{\sim c}\varphi_r$. At this step we know for every state (q,γ) of $\mathcal{R}_{A,\Phi}$ whether it satisfies (or not) φ_l and φ_r (*i.e.* whether any (q,v) with $v \in \gamma$ satisfies φ_l or/and φ_r). First we define a variant of $\mathcal{R}_{A,\Phi}$, called $\mathcal{R}^{\varphi_l,\varphi_r}_{A,\Phi}$, where some transitions are modified according to the truth value of φ_l and φ_r:

1) we replace the transitions $(q,\gamma) \rightarrow_t (q, succ(\gamma))$ by $(q,\gamma) \rightarrow_a (q, \gamma[z_{\bar{l}} \leftarrow 0])$ when $(q,\gamma) \models \varphi_l$, $(q, succ(\gamma)) \models \neg\varphi_l$ and $\gamma \not\models z_{\bar{l}} = 0$.
2) we replace the transitions $(q,\gamma) \rightarrow_a (q',\gamma')$ by $(q,\gamma) \rightarrow_a (q', \gamma'[z_{\bar{l}} \leftarrow 0])$ when $(q,\gamma) \models \varphi_l$, $(q',\gamma') \models \neg\varphi_l$.
3) we replace the transitions $(q,\gamma) \rightarrow_t (q, succ(\gamma))$ by $(q,\gamma) \rightarrow_a (q, \gamma[z_r \leftarrow 0])$ when $(q,\gamma) \models \neg\varphi_r$, $(q, succ(\gamma)) \models \varphi_r$ and $\gamma \not\models z_r = 0$.
4) we replace the transitions $(q,\gamma) \rightarrow_a (q',\gamma')$ by $(q,\gamma) \rightarrow_a (q', \gamma'[z_r \leftarrow 0])$ when $(q,\gamma) \models \neg\varphi_r$, $(q,\gamma') \models \varphi_r$.

Due to these changes, in $\mathcal{R}^{\varphi_l,\varphi_r}_{A,\Phi}$, the clock $z_{\bar{l}}$ (resp. z_r) measures the time elapsed since $\neg\varphi_l$ (resp. φ_r) is true : they are reset when the truth value of the corresponding formula changes. In the following we will use two abbreviations:

$$\overleftarrow{\varphi_l} \stackrel{\text{def}}{=} \varphi_l \vee (\!|z_{\bar{l}} \leq k|\!) \qquad\qquad \overleftarrow{\varphi_r} \stackrel{\text{def}}{=} \varphi_r \wedge (\!|z_r > k|\!)$$

The first one states that φ_l holds or did hold less than k t.u. ago. And the second one states that φ_r lasts for more than k t.u. We will also use the abbreviation $\overleftarrow{\neg\varphi_l}$ to denote $\neg\varphi_l \wedge (\!|z_{\bar{l}} > k|\!)$: the formula $\neg\varphi_l$ has held for more than k t.u. And we use $\overleftarrow{\neg\varphi_r}$ for $\neg\varphi_r \vee (\!|z_r \leq k|\!)$. Therefore, the construction of the region graph $\mathcal{R}^{\varphi_l,\varphi_r}_{A,\Phi}$ allows us to decide the values of $\overleftarrow{\varphi_l}$ and $\overleftarrow{(\neg\varphi_l)}$, for any formula Ψ of the form $\mathsf{E}\varphi_l\mathsf{U}^k_{\sim c}\varphi_r$ or $\mathsf{A}\varphi_l\mathsf{U}^k_{\sim c}\varphi_r$. Furthermore, for all TA A and TCTL^{Δ} formula Ψ the labeling algorithm labels (q,γ) with Ψ in $\mathcal{R}_{A,\Phi}$ iff $(q,v) \models \Psi$ for any $v \in \gamma$ [13].

The proof of this decidability result is based on a generalization of the construction of the region graph for TCTL timed logics (as presented in [6], [7]). Instead of it, and for reasons of efficiency to avoid the state-space explosion problem, model-checkers like UPPAAL or KRONOS use a symbolic analysis algorithm to explore finitely the reachable state-space (this algorithm is called the "zone algorithm"). The implementation of this algorithm uses a data structure initially proposed by [24], the Difference Bounded Matrices, DBMs for short.

The aim of this paper is precisely to propose such an algorithm for decidable TCTL^{Δ} model-checking. The algorithm we propose is an extension of the algorithm used in UPPAAL and KRONOS. Hence, we will first recall the zone algorithm for TCTL timed logics. After this brief presentation of a so much used algorithm, we will come back to TCTL^{Δ} timed logic and present our algorithm for its symbolic model-checking. The remainder of the paper is devoted to present a complete correctness proof of our algorithm and we describe its implementation using the DBM data structure.

IV. Classical zone algorithm, State of the Art

In this section, we describe the on-the-fly analysis algorithm, which is implemented in some tools for the verification of classical timed logics [14], [21], [33], [23], [2].

A. Zones

For timed automata, the set of configurations is infinite. To check this model, it is therefore necessary to manipulate sets of configurations, and therefore to provide a symbolic representation, called zone. A zone is a set of valuations defined by a conjunction of simple constraints $x \sim c$ or $x - y \sim c$ where x and y are clocks, $\sim$ is a comparison sign, and c is a integer constant. In forward and backward analysis, the objects that will be handled are pairs (q, Z) where q is a control state of the automaton and Z a zone.

$$(q, Z) = \{(q, v) \mid v \in Z\}$$

On zones, multiple operations can be performed:

- **Future** of Z, defined by $\overrightarrow{Z} = \{v + t \mid v \in Z \wedge t \in \mathbb{R}_{\geq 0}\}$;
- **Past** of Z, defined by $\overleftarrow{Z} = \{v - t \mid v \in Z \wedge t \in \mathbb{R}_{\geq 0}\}$;
- **Intersection** of two zones, defined by $Z \cap Z' = \{v | v \in Z \wedge v \in Z'\}$;
- **Clock reset** in a zone, defined by $[Y \leftarrow 0]Z = \{[Y \leftarrow 0]v | v \in Z\}$;
- **Inverse clock reset** of a zone, defined by $[Y \leftarrow 0]^{-1}Z = \{v \mid [Y \leftarrow 0]v \in Z\}$.

These operations, defined through the first order formulas on the zones, preserve zones [32].

Example 3. Consider the zone Z drawn in (dark) gray on the figure beside (Fig. 4): Z is defined by the clock constraint

$$1 < x < 4 \wedge 2 < y < 4 \wedge x - y < 1 \wedge y - x < 2.$$

Taking the operation Future of Z, $\overrightarrow{Z}$ is drawn in light gray and in dark gray; it is defined by the clock constraint

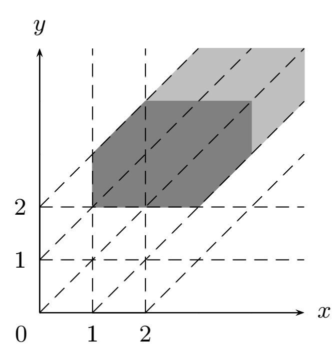


Figure 4: Example of Zone.

$$x > 1 \wedge y > 2 \wedge x - y < 1 \wedge y - x < 2.$$

□

B. The Algorithm

We give now an idea about how it is possible to check the TCTL properties [33]. The construction to be described avoids building region graph, because such an approach would not be very effective, and there's no data structures really adapted to the regions in terms of complexity. The idea of the algorithm [33] is to calculate for each formula, its characteristic set defined as set of pairs (q, Z) where q is a control state of the automaton and Z a zone, *i.e.*

$$[\![\psi]\!] = \bigcup_{(q,Z)\models\psi} (q, Z) \ = \{((q, v), w) \mid (q, v) \models \psi\}$$

Where w is a valuation on clocks corresponding to the *Until* operators in the TCTL formula.

The construction is by induction on the structure of the formula:

$$\begin{aligned}
&[\![p]\!] = \{((q, v), w) \mid q \text{ labeled by } p\}\\
&[\![True]\!] = \{((q, v), w) \mid q \text{ is a state }\}\\
&[\![\neg\varphi]\!] = [\![True]\!] \setminus [\![\varphi]\!]\\
&[\![\varphi_1 \vee \varphi_2]\!] = [\![\varphi_1]\!] \cup [\![\varphi_2]\!]\\
&[\![\varphi_1 \wedge \varphi_2]\!] = [\![\varphi_1]\!] \cap [\![\varphi_2]\!]
\end{aligned}$$

It remains to describe the characteristic sets of formulas that have the *Until* operator. For the formula $\mathsf{E}\varphi_1\mathsf{U}_{\sim c}\varphi_2$, the characteristic set is given by the following recurrent sequence [33]:

$$[\![\mathsf{E}\varphi_1\mathsf{U}_{\sim c}\varphi_2]\!] = EU([z \leftarrow 0][\![\varphi_1]\!], [\![\varphi_2]\!] \cap [\![z \sim c]\!])$$

Where z is the clock corresponding to the operator U and $EU(R_1, R_2) = \bigcup_{i\geq 0} E_i$ with

$$\begin{cases} E_0 = R_2\\ E_{i+1} = Pre[R_1](E_i) \cup Pre(E_i)\end{cases}$$

$Pre[R_1](E_i)$ represents the set of configurations that allow to reach E_i by letting time pass while staying in R1, while $Pre(E_i)$ represents the configurations that allow to reach E_i by taking an action transition.

A clock is attached to each U operator in the formula. it's used to handle subscripts $\sim c$ in *Until* modalities. We note that the above analysis is in fact a backward analysis. We do not describe the algorithm of $\mathsf{A}\varphi_1\mathsf{U}_{\sim c}\varphi_2$ which also uses a backward analysis, but slightly more complicated, it is described for example in [26].

V. Back to TCTL$^\Delta$ timed logic: Symbolic Model-Checking Algorithm

In this section, we propose a symbolic model-checking algorithm which computes the characteristic sets of some TCTL$^\Delta$ formulae and checks their truth values using a backward analysis. This algorithm extends the zone algorithm for TCTL timed logics. We also present a complete correctness proof of this algorithm, and we describe its implementation using the DBM data structure in the next section.

A. *Modality* $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$

For this modality, the approach we have opted is to split the semantics of formula $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$ in two parts, the right and the left part (as depicted in Fig. 5). The left part represents the subrun where φ_1 is true everywhere except along negligible duration subpaths ($\leq k$), until reaching the right part which represents the subrun where φ_2 lasts long enough around a position ($z \sim c$), and before this position φ_1 is true except along negligible duration subpaths.

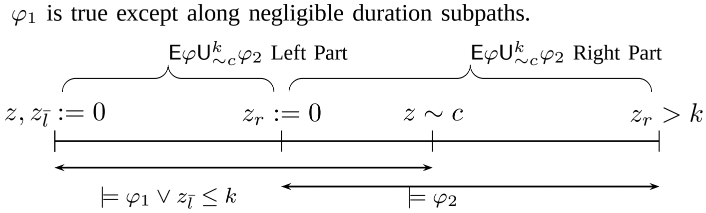


Figure 5: Illustration of $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$ Modality.

1) $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$ *Right part:*

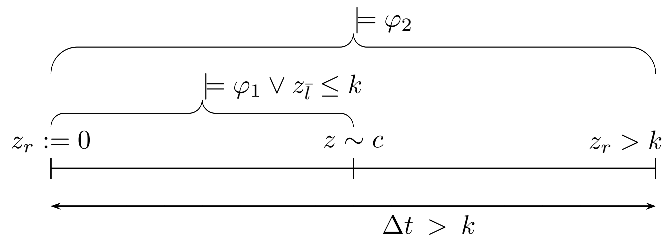


Figure 6: Illustration of $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$ Right part.

First case: $\sim \in \{<, \leq\}$

In this case, it is necessary and sufficient that constraint $z \sim c$ be verified at the beginning of the subrun where φ_2 lasts long enough ($> k$). Thus all the right part of $\mathsf{E}\varphi\mathsf{U}^k_{\sim c}\varphi_2$, as depicted in the figure above (Fig. 6), can be reduced andx expressed using the following TCTL formula :

$$(z \sim c) \ \wedge \ [z_r \leftarrow 0](\mathsf{E}\ \varphi_2 \ \mathsf{U}\ (\varphi_2 \ \wedge \ z_r > k))$$

Second case: $\sim \in \{>, \geq, =\}$

In this case, we will split the subrun where φ_2 is true for more than k t.u into two parts, one satisfying $(\varphi_1 \vee z_{\bar{l}} \leq k)$, followed by the other which satisfying $z \sim c$ at its first position as depicted in the figure above (Fig. 6). Thus, the semantic of the $\mathsf{E}\varphi\mathsf{U}^k_{\sim c}\varphi_2$ Right part is deduced from *Definition 3* as follows:

Let s be a state of some TTS $\mathcal{T} = \langle S, s_{\text{init}}, \rightarrow, l\rangle$ which satisfies the $\mathsf{E}\varphi\mathsf{U}^k_{\sim c}\varphi_2$ Right part, written $s \models \mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)$, we have :

$$\begin{aligned}
s \models \mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2) \quad \textit{iff} \quad & \exists\,\sigma \in \mathrm{Exec}(s) \text{ s.t. } \hat{\mu}(\sigma) > k\\
& \wedge \sigma \ \models \varphi_2 \wedge \ (\exists p \in \sigma \mid s_p \models z \sim c)\\
& \wedge \forall p' <_\sigma p,\ s_{p'} \models \varphi_1 \vee z_{\bar{l}} \leq k
\end{aligned}$$

Now, we propose and prove that the following sequence is increasing by inclusion, stationary, and its least upper bound represents the set of all symbolic states (*i.e.*, characteristic sets defined in section 4.2, as set of pairs (q, Z) where q is a control state of the automaton and Z a zone) that satisfying $\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)$:

$$\begin{cases}
Y_0 & = [\![(z \sim c) \ \wedge \ (\mathsf{E}\ \varphi_2 \ \mathsf{U}\ (\varphi_2 \ \wedge \ z_r > k))]\!]\\
Y_{n+1} & = Y_n \vee \Bigg(\Big([\![\varphi_2 \ \wedge \ \varphi_1]\!] \rhd [z_{\bar{l}} \leftarrow 0](Y_n \ \wedge \ \neg\varphi_1)\Big)\\
& \qquad \vee \ \Big([\![\varphi_2 \ \wedge \ \varphi_1]\!] \rhd (Y_n \ \wedge \ \varphi_1)\Big)\\
& \qquad \vee \ \Big([\![\varphi_2 \ \wedge \ (\neg\varphi_1 \ \wedge \ z_{\bar{l}} \leq k)]\!] \rhd Y_n\Big)\Bigg)
\end{cases}$$

Note that z_r is reset when the stationary value of the sequence Y_n is reached, *i.e.* after that the set of symbolic states satisfying $\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)$ is computed.

The recurrent sequence Y_n computes for each iteration the predecessors of current states represented by Y_n. As we said in section 3.2, the clock $z_{\bar{l}}$ measures time elapsing when φ_1 is false, so it will be reset at each transition from set of states satisfying φ_1 to another satisfying not $\neg\varphi_1$. Without losing information about clock $z_{\bar{l}}$, and in order to further optimize our sequence, $z_{\bar{l}}$ can also be reset when transition from set of states satisfying φ_1 to set of states satisfying φ_1, and therefore the sequence Y_n becomes as follows:

$$\begin{cases}
Y_0 & = [\![(z \sim c) \ \wedge \ (\mathsf{E}\ \varphi_2 \ \mathsf{U}\ (\varphi_2 \ \wedge \ z_r > k))]\!]\\
Y_{n+1} & = Y_n \vee \Bigg(\Big([\![\varphi_2 \ \wedge \ \varphi_1]\!] \rhd [z_{\bar{l}} \leftarrow 0]Y_n\Big)\\
& \qquad \vee \ \Big([\![\varphi_2 \ \wedge \ (\neg\varphi_1 \ \wedge \ z_{\bar{l}} \leq k)]\!] \rhd Y_n\Big)\Bigg)
\end{cases}$$

Now we define the operator $\triangleright$ as follows:

Definition 4 (Predecessor operator $\triangleright$). *Given a TA A, a TTS $\mathcal{T} = \langle S, s_{init}, \rightarrow, l\rangle$, an alphabet Σ which denotes a finite set of actions and two characteristic sets Q_1 and Q_2. Calculate $Q_1 \triangleright Q_2$ is to determine:*

- $Q_1 \triangleright Q_2$:
  - All the instantaneous predecessors of Q_2 states that verify Q_1, *i.e.* the states satisfying Q_1 and can reach Q_2 by an action transition denoted $Q_1 \triangleright^a Q_2$.
  - Union, all temporal predecessors of Q_2 that verify Q_1, *i.e.* all states that can reach a state of Q_2 by a delay transition, such that all intermediates states are in Q_1.
    $q \in Q_1 \triangleright^t Q_2 \;\Leftrightarrow\; q \in Q_1 \wedge \exists t > 0$ s.t. $q + t \in Q_2$ and $\forall t' < t \;\; q + t' \in Q_1$

Back to our sequence Y_n, it can be written as follows :

$$\begin{cases} Y_0 & = & [\![(z \sim c) \wedge (\mathsf{E}\,\varphi_2\,\mathsf{U}\,(\varphi_2 \wedge z_r > k))]\!] \\ Y_{n+1} & = & g(Y_n) \end{cases}$$

Such that: $g(Y) = Y \vee \Big(\big([\![\varphi_2 \wedge \varphi_1]\!] \triangleright [z_{\bar{l}} \leftarrow 0]Y \big) \vee \big([\![\varphi_2 \wedge (\neg\varphi_1 \wedge z_{\bar{l}} \le k)]\!] \triangleright Y \big) \Big)$

The particularity of the backward analysis is that the iterative calculating described by the sequence Y_n terminates, the reason is quite simple; it is fairly easy to show that if Z' is a zone and that this zone is an union of regions, then the zone $Z = g(Z')$ is not only a zone, but also is an union of regions [1]. As there's a finite number of regions, the number of pairs (q, Z) that can be computed in a backward analysis is finite.

Thus we show that the sequence Y_n is increasing by inclusion, stationary, and its least upper bound represents the characteristic set of $\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)$:

$$[\![\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)]\!] = [z_r \leftarrow 0] Sup\ Y_n$$

Proof: (sketch.). In order to prove this result we show at first that the least upper bound of the sequence Y_n is the least fixpoint of g.

Let be E the set of symbolic states defined as :

$$E = \{(q, Z) \mid q \in Q_A \text{ and Z is a Zone}\}$$

We know that E is a finite set, hence the power set $P(E)$ is also finite.

Furthermore, the first term of Y_n is given as the characteristic set of a TCTL formula, then we have $Y_0 \in P(E)$. As all operations in the function g preserve zones [32], so $\forall n \in \mathbb{N}\ Y_n = g^n(Y_0) \in P(E)$.

The sequence Y_n is monotonic by inclusion, because $Y_n \subseteq Y_{n+1}$ $\forall n \in \mathbb{N}$. Thus Y_n is monotonic in the finite set $P(E)$, so Y_n is stationary, *i.e.* $\exists\ r \in \mathbb{N}$ such that $\forall\ n \ge r$, $Y_n = Y_r$.

Moreover, since $(P(E), \subseteq)$ is a complete partially ordered set, then its finite subset $(W, \subseteq)$ defined as $W = \{Y_0, Y_1, ..., Y_n, ...\}$ is also a complete totally ordered set.

Also, $g : W \longmapsto W$ is a monotonic (order-preserving) function, because $\forall\ Y, Y' \in W$, if $Y \subseteq Y'$ we have: $g(Y) \subseteq g(Y')$ (immediate using the definition of the operator $\triangleright$).

Since in finite sets, monotonic function is always Scott-continuous, so using *Kleene's fixed-point theorem*, the least fixpoint of g is the least upper bound of the sequece $g^n(Y_0) = Y_n$, such that Y_0 is the least element of W (intersection of its elements).

$$Sup\ Y_n = \mu.Y.g(Y)$$

On another side, if we prove that $[\![\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)]\!]$ $= [z_r \leftarrow 0]\mu.Y.g(Y)$, we have then the result we're looking for.
Let be $Q = [\![\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)]\!]$.
1. First of all we show that Q is a fixpoint of $g : Y \to g(Y)$. We prove that:

$$\begin{cases} Q & = & g(Q),\ \textit{i.e.:} \\ Q & = & Q \vee \big([\![\varphi_2 \wedge \varphi_1]\!] \triangleright [z_{\bar{l}} \leftarrow 0]Q \big) \\ & & \vee \big([\![\varphi_2 \wedge (\neg\varphi_1 \wedge z_{\bar{l}} \le k)]\!] \triangleright Q \big) \end{cases}$$

$\subseteq$ / We have obviously $Q \subseteq g(Q)$

$\supseteq$ / Let $q \in Q \vee \big([\![\varphi_2 \wedge \varphi_1]\!] \triangleright [z_{\bar{l}} \leftarrow 0]Q \big) \vee \big([\![\varphi_2 \wedge (\neg\varphi_1 \wedge z_{\bar{l}} \le k)]\!] \triangleright Q \big)$, we prove that $q \in Q = [\![\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)]\!]$

☐ Suppose that $q \in \big([\![\varphi_2 \wedge \varphi_1]\!] \triangleright [z_{\bar{l}} \leftarrow 0]Q \big)$, then $q \in [\![\varphi_2 \wedge \varphi_1]\!]$ and $\exists q' \in [z_{\bar{l}} \leftarrow 0]Q$, $\exists \alpha \in \mathbb{R}^*_+ \cup \Sigma$ s.t. :
$q \triangleright^{\alpha} q'$, and if $\alpha = t > 0$, we have $\forall t' < t \quad q + t' \in [\![\varphi_2 \wedge \varphi_1]\!]$

Moreover, $q' \in [z_{\bar{l}} \leftarrow 0]Q = [z_{\bar{l}} \leftarrow 0][\![\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)]\!]$, *i.e.:*
$\exists\, \sigma \in \text{Exec}(q')$ s.t. $\hat{\mu}(\sigma) > k \wedge \sigma \models \varphi_2$
$\wedge (\exists p \in \sigma \mid s_p \models z \sim c) \wedge \forall p' <_{\sigma} p,\ s_{p'} \models \varphi_1 \vee z_{\bar{l}} \le k$

So whatever the type of the transition α, action or delay, we can verify that the subrun σ' defined as: $\sigma' = (q \triangleright^{\alpha} q').\sigma$ also verify :
$\hat{\mu}(\sigma') > k \wedge \sigma' \models \varphi_2$
$\wedge (\exists p \in \sigma' \mid s_p \models z \sim c) \wedge \forall p' <_{\sigma'} p,\ s_{p'} \models \varphi_1 \vee z_{\bar{l}} \le k$

Given that $\sigma' \in Exec(q)$, so we have $q \in [\![\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)]\!]$, *i.e.:* $q \in Q$.

☐ Suppose that $q \in \big([\![\varphi_2 \wedge (\neg\varphi_1 \wedge z_{\bar{l}} \le k)]\!] \triangleright Q \big)$, in the same manner as the previous proof, we show taht $q \in Q$.

Consequently it follows that: $\forall q \in g(Q)$ we have $q \in Q$, *i.e.:* $g(Q) \subseteq Q$. Hence the result $Q = g(Q)$, which means that $Q = [\![\mathsf{RP}(\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2)]\!]$ is a fixpoint of $g : Y \to g(Y)$.

2. Now we prove that $Q = [z_r \leftarrow 0]\mu.Y.g(Y)$:

☐ Let $q \in Q$, so: $\quad \exists\, \sigma \in \text{Exec}(q)$ s.t. $\hat{\mu}(\sigma) > k \wedge \sigma \models \varphi_2$
$\wedge (\exists p \in \sigma \mid s_p \models z \sim c) \wedge \forall p' <_{\sigma} p,\ s_{p'} \models \varphi_1 \vee z_{\bar{l}} \le k$

In other words, $\exists\, \sigma \in \text{Exec}(q) : \sigma = q \triangleright^{\alpha_0} q_1 \triangleright^{\alpha_1} ... \triangleright^{\alpha_{i-1}} q_i \triangleright ...$, with $\alpha_i \in \mathbb{R}^*_+ \cup \Sigma$, and z_r is reset at the beginning of σ, such that $q_i \in Y_0 = [\![(z \sim c) \wedge (\mathsf{E}\,\varphi_2\,\mathsf{U}\,(\varphi_2 \wedge z_r > k))]\!]$, and $\forall j < i$, we have:

$$\begin{cases} \mathrm{q}_j & \models & \varphi_2 \wedge (\varphi_1 \vee z_{\bar{l}} \le k\,), & & if\ \alpha_j \in \Sigma \\ \mathrm{q}_j + t' & \models & \varphi_2 \wedge (\varphi_1 \vee z_{\bar{l}} \le k\,), & \forall t' < t & if\ \alpha_j = t \in \mathbb{R}^*_+ \end{cases}$$

Let be q_{i-1} from the subrun σ, we have $q_{i-1} \triangleright^{\alpha_{i-1}} q_i$. So q_{i-1} is a predecessor of $q_i \in Y_0$, that verifies $\varphi_2 \wedge (\varphi_1 \vee z_{\bar{l}} \le k\,)$. Then, according to the definition of the function g, $q_{i-1} \in g(Y_0) = Y_1$.

By the same reasoning we deduce that $q_{i-2} \in g^2(Y_0) = Y_2$. This is repeated until reaching $q \in [z_r \leftarrow 0]g^i(Y_0) = [z_r \leftarrow 0]Y_i$, *i.e.* $Q \subseteq [z_r \leftarrow 0]Y_i$. As $Y_i \subseteq Sup\ Y_n = \mu.Y.g(Y)$, then $Q \subseteq [z_r \leftarrow 0]\mu.Y.g(Y)$.

Moreover Q is already a fixpoint of g, *i.e.* $\mu.Y.g(Y) \subseteq Q$, which means that $[z_r \leftarrow 0]\mu.Y.g(Y) \subseteq [z_r \leftarrow 0]Q = Q$.
Hence, $Q = [\![\mathsf{RP}(\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2)]\!]$ is the least fixpoint of $g : Y \rightarrow g(Y)$:

$$[\![\mathsf{RP}(\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2)]\!] = [z_r \leftarrow 0]\mu.Y.g(Y)$$

Since we proved that the least fixpoint of g is the least upper bound of the sequence Y_n, we have finally:

$$[\![\mathsf{RP}(\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2)]\!] = [z_r \leftarrow 0]Sup\ Y_n$$

■

2) $\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2$ Left part:

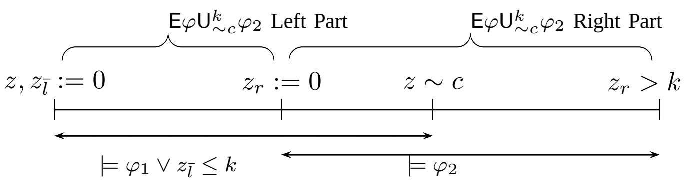


Figure 7: Illustration of $\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2$ Left part.

Now we propose and prove that the characteristic set of $\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2$ (deduced from the left part modality Fig. 7) is given by least upper bound of the following stationary and increasing (by inclusion) sequence:

$$\begin{cases} X_0 & = [\![\mathsf{RP}(\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2)]\!] \\ X_{n+1} & = X_n \vee \Big(\big([\![\varphi_1]\!] \rhd [z_{\bar{l}} \leftarrow 0]X_n \big) \\ & \qquad \vee \big([\![(\neg\varphi_1 \ \wedge\ z_{\bar{l}} \leq k)]\!] \rhd X_n \big) \Big) \end{cases}$$

Where $z,\ z_{\bar{l}}$ are reset when the stationary value of the sequence X_n is reached, *i.e.* after that the set of symbolic states satisfying $\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2$ is computed.

Proof: (sketch.). We show in the same way as the previous proof that: $[\![\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2]\!] = \mu.X.f(X)$, s.t. :

$$\begin{cases} X_0 & = & [\![\mathsf{RP}(\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2)]\!] \\ X_{n+1} & = & f(X_n) \end{cases}$$

Such that:

$$f(X) = X \vee \Big(\big([\![\varphi_1]\!] \rhd [z_{\bar{l}} \leftarrow 0]X \big) \vee \big([\![(\neg\varphi_1 \wedge z_{\bar{l}} \leq k)]\!] \rhd X \big) \Big).$$

Therefore we have the following result:

$$[\![\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2]\!] = [z \leftarrow 0][z_{\bar{l}} \leftarrow 0](Sup\ X_n)$$

Note that when computing iterations of the sequence X_n (resp Y_n), the stop condition is given by convergence to the fixed point of f (resp g), *i.e.* $X_{n+1} = X_n$ (resp $Y_{n+1} = Y_n$).

■

VI. Implementation of the Algorithm using DBMs

To prove that the DBMs are appropriate to implement algorithms proposed in the previous section, we will show how to compute using the DBMs the new operations on zones appearing in the TCTL^{Δ} model-checking algorithm. Indeed, we first recall the main features of the DBM data structure, then we give an effective method for computing the operation $Q_1 \rhd\ Q_2$. We present after pseudocode for $\mathsf{E}\varphi_1 \mathsf{U}^k_{\sim c}\varphi_2$ Model-Checking algorithm.

A. The Implementation: the DBM Data Structure

In order to implement the TCTL model-checking algorithm, we need a data structure to represent the zones and this data structure must allow to test for inclusion of zones and to compute easily the different operations used in the algorithm, that is the intersection of two zones, the past of a zone, the image of a zone by a reset and the normalization of a zone. Tools like UPPAAL or KRONOS use the data structure proposed by Dill in [24], the DBM data structure. A detailed presentation of this data structure can be found in [20].

A *difference bounded matrice* (say DBM for short) for n clocks is an $(n+1)$-square matrice of pairs:

$$(m; \prec) \in \mathbb{V} = (\mathbb{Z} \times \{<, \leq\}) \cup \{(\infty; <)\}.$$

A DBM $M = (m_{i,j}; \prec_{i,j})_{i,j=1...n}$ defines the following subset of $\mathbb{R}^n$ (the clock x_0 is supposed to be always equal to zero, i.e. for each valuation v, $v(x_0) = 0$):

$$\{v : \{x_1, ..., x_n\} \longrightarrow \mathbb{R} \,|\, \forall\ 0 \leq i, j \leq n,\ v(x_i) - v(x_j) \prec_{i,j} m_{i,j}\}$$

where $\gamma < \infty$ means that γ is some real (there is no bound on it). This subset of $\mathbb{R}^n$ is a zone and will be denoted, in what follows, by $[\![M]\!]$. Each DBM on n clocks represents a zone of $\mathbb{R}^n$. Note that several DBMs can define the same zone.

Example 4. The zone defined by the equations $x_1 > 3\ \wedge\ x_2 \leq 5\ \wedge\ x_1 - x_2 < 4$ can be represented by the two DBMs

$$\begin{pmatrix} (0, \leq) & (-1, \leq) & (0, \leq) \\ (4, \leq) & (0, \leq) & (2, \leq) \\ (5, \leq) & (0, \leq) & (0, \leq) \end{pmatrix} \text{ and } \begin{pmatrix} (0, \leq) & (-1, \leq) & (0, \leq) \\ (4, \leq) & (0, \leq) & (2, \leq) \\ (5, \leq) & (0, \leq) & (0, \leq) \end{pmatrix}$$

Thus the DBMs are not a canonical representation of zones. Moreover, it isn't possible to test syntactically whether $[\![M]\!] = \emptyset$ or $[\![M_1]\!] = [\![M_2]\!]$. A normal form has thus been defined for representing zones. Its computation uses the Floyd-Warshall algorithm and some syntactic rewritings (see [24], [20] for a description of this procedure). In what follows, we denote by $\Phi(M)$ the normal form of M. Before stating some very important properties of the normal form, we define a total order on V in the following way: if $(m; \prec), (m'; \prec') \in \mathbb{V}$, then $(m; \prec) \leq (m'; \prec') \iff$

$$\begin{cases} m < m' \\ \text{or} \\ m = m' \text{ and either } \prec = \prec' \text{ or } \prec' = \leq \end{cases}$$

Of course, for each $m \in \mathbb{Z}$, it holds that $m < \infty$. We define $>, \geq$ and $<$ in a natural way. These orders are extended to the DBMs in the following way: let $M = (m_{i,j}; \prec_{i,j})_{i,j=0...n}$ and $M' = (m'_{i,j}; \prec'_{i,j})_{i,j=0...n}$ be two DBMs, then

$M \leq M' \iff$ for evry $i, j\ = 0...n$, $(m_{i,j}; \prec_{i,j}) \leq (m'_{i,j}; \prec'_{i,j})$.

We can now state some (very useful) properties of normal forms. If M and M' are DBMs, then:

(i) $[\![M]\!] = [\![\phi(M)]\!]$ and $\phi(M) \leq M$,
(ii) $[\![M]\!] \subseteq [\![\phi(M)]\!] \iff \phi(M) \leq M' \iff \phi(M) \leq \phi(M')$

The last point expresses the fact that the test for inclusion of zones can be checked syntactically on the normal forms of the DBMs (representing the zones).

Normal forms of DBMs can be characterized in a natural way. if $M = (m_{i,j}; \prec_{i,j})_{i,j=0...n}$ is a DBM such that $[\![M]\!] \neq \emptyset$, then the two following properties are equivalent:

1) M is in normal form,
2) for every $i, j = 0...n$, for every real $-m_{j,i} \prec_{j,i} r \prec_{i,j} m_{i,j}$, there exists a valuation $v \in [\![M]\!]$ such that $v(x_j) - v(x_i) = r$ (still assuming that $v(x_0) = 0$).

This property expresses the fact that if a DBM is in normal form, then no constraint of this DBM can be tightened using the Floyd-Warshall algorithm.

Computation of Some Operations on DBMs. As we argued at the beginning of the section, the data structure used to represent zones must also be appropriate to compute all the operations on zones that are used by the TCTL model-chacking algorithm, namely future, past, intersection, image by resets and normalization. These operations on DBMs are described nicely in [20], here we recall them quickly.

Intersection. Assume that $M = (m_{i,j}; \prec_{i,j})_{i,j=1...n}$ and $M' = (m'_{i,j}; \prec'_{i,j})_{i,j=1...n}$ are two DBMs in normal form. Then, defining $M'' = (m''_{i,j}; \prec''_{i,j})_{i,j=1...n}$ by

$(m''_{i,j}; \prec''_{i,j}) = min\left((m_{i,j}; \prec_{i,j}), (m'_{i,j}; \prec'_{i,j})\right)$ for every indexes $i, j = 0...n$.

We get that $[\![M'']\!] = [\![M]\!] \cap [\![M']\!]$. Note that it can be the case that M'' is not in normal form.

Future. Assume that $M = (m_{i,j}; \prec_{i,j})_{i,j=1...n}$ is a DBM in normal form. We define the DBM $\overrightarrow{M} = (m'_{i,j}; \prec'_{i,j})_{i,j=1...n}$ by:

$$\begin{cases} (m'_{i,j}; \prec'_{i,j}) = (m_{i,j}; \prec_{i,j}) \text{ if } j \neq 0 \\ (m'_{i,0}; \prec'_{i,0}) = (\infty; <) \end{cases}$$

We get that $[\![\overrightarrow{M}]\!] = \overrightarrow{[\![M]\!]}$ and that the DBM $\overrightarrow{M}$ is in normal form.

Past. Assume that $M = (m_{i,j}; \prec_{i,j})_{i,j=1...n}$ is a DBM in normal form. We define the DBM $\overleftarrow{M} = (m'_{i,j}; \prec'_{i,j})_{i,j=1...n}$ by:

$$\begin{cases} (m'_{i,j}; \prec'_{i,j}) = (m_{i,j}; \prec_{i,j}) \text{ if } i \neq 0 \\ (m'_{0,j}; \prec'_{0,j}) = (0; \leq) \end{cases}$$

We get that $[\![\overleftarrow{M}]\!] = \overleftarrow{[\![M]\!]}$ and note that it can be the case that $\overleftarrow{M}$ is not in normal form.

Image by resets. Assume that $M = (m_{i,j}; \prec_{i,j})_{i,j=1...n}$ is a DBM in normal form. We define the DBM $M_{x_k:=0} = (m'_{i,j}; \prec'_{i,j})_{i,j=1...n}$ by:

$$\begin{cases} (m'_{i,j}; \prec'_{i,j}) = (m_{i,j}; \prec_{i,j}) & \text{if } i, j \neq k \\ (m'_{k,k}; \prec'_{k,k}) = (m'_{0,k}; \prec'_{0,k}) = (m'_{k,0}; \prec'_{k,0}) = (0; \leq) & \\ (m'_{i,k}; \prec'_{i,k}) = (m_{i,0}; \prec_{i,0}) & \text{if } i \neq k \\ (m'_{k,i}; \prec'_{k,i}) = (m_{0,i}; \prec_{0,i}) & \text{if } i \neq k \end{cases}$$

We get that $[\![M_{x_k:=0}]\!] = [x_k \leftarrow 0][\![M]\!]$ and that the DBM $M_{x_k:=0}$ is in normal form.

DBM Normal form (Zone Normalization). The DBM normal form can be computed using a shortest path algorithm. Floyd-Warshall algorithm is often used to transform DBM to canonical form. We define the DBM $\Phi(M) = (m'_{i,j}; \prec'_{i,j})_{i,j=1...n}$ by:

$(m''_{i,j}; \prec''_{i,j}) = min\left((m_{i,j}; \prec_{i,j}), (m'_{i,k}; \prec'_{i,k}) + (m'_{k,j}; \prec'_{k,j})\right)$ for every index $k = 0...n$.

We get that $\Phi(M)$ the normal form of M.
Note that to manipulate DBMs efficiently we need two operations on bounds: comparison and addition. We define that $(m_1; \prec_1) < (m_2; \prec_2)$ if $m_1 < m_2$ and $(m; <) < (m; \leq)$. Further we define addition as $(m_1; \leq) + (m_2; \leq) = (m_1 + m_2; \leq)$ and $(m_1; <) + (m_2; \prec_2) = (m_1 + m_2; <)$.

B. Computing the operator $\triangleright$

We present now an effective method for computing the operation $Q_1 \triangleright Q_2$, where Q_1 and Q_2 are characteristic sets represented by sets of symbolic states of the form $\bigcup_q (q, {Z_q}^i)$, for $i = 1, 2$. The method consists in determining all instantaneous and temporal predecessors as follows:

1) Instantaneous predecessors:

Let $q_1, q_2 \in Q_A$ be two control states, $e = q_1 \xrightarrow{g,r}_A q_2$ an edge from q_1 to q_2. Let (q_2, Z_2) be a symbolic state. We have [33] :

$$pre_e(q_2,\ Z_2) = \Big(q_1,\ g \wedge [r \leftarrow 0](Z_2)\Big)$$

i.e., $pre_e(q_2,\ Z_2)$ is the symbolic state representing predecessors, through instantaneous transition via the edge e, of states characterized by $(q_2,\ Z_2)$.

Example 5. Let A be the timed automaton depicted below (Fig. 8), with two clocks x and y.

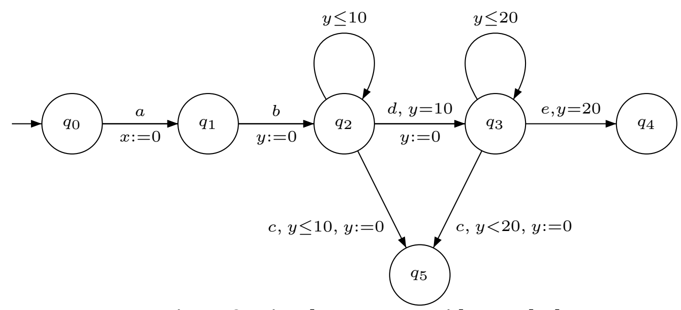


Figure 8: Timed automaton with two clocks.

The symbolic state that characterizes instantaneous predecessor of the symbolic state $(q, Z) = (q_5,\ y = 0 \wedge z > 35)$ through the edge from q_2 to q_5, is calculated as follows:

$$\begin{aligned} \text{pre}_{q_2,q_5}(q, Z) &= (q_2,\ y \leq 10\ \wedge\ [y \leftarrow 0](y = 0 \wedge z > 35)) \\ &= (q_2,\ y \leq 10\ \wedge\ (0 = 0 \wedge z > 35)) \\ &= (q_2,\ y \leq 10\ \wedge\ z > 35) \end{aligned}$$

□

Therefore, for a symbolic state (q, Z), we have :

$$pre(q, Z) = \bigcup_{e \in E} pre_e(q, Z)$$

i.e. $pre(q, Z)$ is the set of symbolic states that characterizes all instantaneous predecessors of the symbolic state (q, Z).

2) Temporal predecessors:

Let (q, Z_1) and (q, Z_2) be two symbolic states. We define:

$$(q, Z_1) \triangleright_t (q, Z_2) = (q, \exists t > 0 \text{ s.t. } (Z_2 + t\ \wedge\ \forall t' < t\ \ Z_1 + t'))$$

i.e. $(q, Z_1) \triangleright_t (q, Z_2)$ represents temporal predecessors of states characterized by (q, Z_2), such that all intermediates valuations satisfy Z_1.
Furthermore, Yovine proved in [33] that $(q, Z_1) \triangleright_t (q, Z_2)$ is also a symbolic state. Indeed he demonstrated that quantifiers can be eliminated and then the result is a timing constraint.

Example 6. Let Z_1 and Z_2 be zones defined as:

$$\begin{aligned} \text{Z}_1 &= (0 < x\ \vee\ 0 < y) \\ \text{Z}_2 &= (y \leq 10\ \wedge\ 35 < z) \end{aligned}$$

Back to the timed automaton depicted in *Example 5*, the symbolic state $(q_2, Z_1) \triangleright_t (q_2, Z_2)$ is calculated as follows :

$$\exists t > 0 \text{ s.t. } \Big((y + t \leq 10 \ \wedge \ 35 < z + t) \ \wedge \ \forall t' < t \ \ (0 < x + t' \ \vee \ 0 < y + t')\Big)$$

The universal quantifier can be written in existential quantifier. We obtain:

$$\exists t > 0 \text{ s.t. } \Big((y + t \leq 10 \ \wedge \ 35 < z + t) \ \wedge \ \neg\exists t' \geq 0 \ \ (t' < t \ \wedge \ \neg(0 < x + t' \ \vee \ 0 < y + t'))\Big)$$

Which is equivalent to

$$\exists t > 0 \text{ s.t. } \Big((y + t \leq 10 \ \wedge \ 35 < z + t) \ \wedge \ \neg\exists t' \geq 0 \ \ (t' < t \ \wedge \ x + t' \leq 0 \ \wedge \ y + t' \leq 0)\Big)$$

It is possible to eliminate the quantifier [33]

$$\exists t > 0 \text{ s.t. } \Big((y + t \leq 10 \ \wedge \ 35 < z + t) \ \wedge \ \neg(0 \leq t' < t \ \wedge \ x + t' \leq 0 \ \wedge \ y + t' \leq 0)\Big)$$

From which we deduce

$$\exists t > 0 \text{ s.t. } \Big((y + t \leq 10 \ \wedge \ 35 < z + t) \ \wedge \neg(x \leq 0 \ \wedge \ y \leq 0)\Big)$$

Whose negation is

$$\exists t > 0 \text{ s.t. } \Big((y + t \leq 10 \ \wedge \ 35 < z + t) \ \wedge (0 < x \ \vee \ 0 < y)\Big)$$

The quantifier is removed by the same procedure. The result is then:

$$(q_2, Z_1) \triangleright_t (q_2, Z_2) = \Big(q_2, \ (y \leq 10 \ \wedge \ y - z < -25)\Big) \qquad \square$$

Finally, using operators pre and $\triangleright_t$ it is possible to compute the operation $Q_1 \triangleright Q_2$. Therefore, we reduce all operations appearing in the TCTL^{Δ} model checking algorithm to known operations on zones, which are obviously implemented through the DBM data structure.

C. Pseudo-Code for $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$ Model-Checking algorithm

We now give the pseudo-code for Model-Checking algorithm of $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$ modality.

Algorithm 1 Model-Checking of $\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$ Modality

function RIGHT PART($\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2 \ : \ \mathsf{TCTL}^{\Delta}$)
    // TCTL formula
    TargetSet := $[\![(z \sim c) \wedge (\mathsf{E}\varphi_2\mathsf{U}(\varphi_2 \wedge z_r > k))]\!]$;

    repeat
        CurrentSet := TargetSet;
        TargetSet := TargetSet $\cup$ CurrentSet;
TargetSet := TargetSet $\cup\big([\![\varphi_2 \wedge \varphi_1]\!] \triangleright [z_{\bar{l}} \leftarrow 0]$ CurrentSet $\big)$;
TargetSet := TargetSet $\cup\big([\![\varphi_2 \wedge (\neg\varphi_1 \wedge z_{\bar{l}} \leq k)]\!] \triangleright$ CurrentSet $\big)$;
    until TargetSet = CurrentSet
    return $[z_r \leftarrow 0]$TargetSet;
end function

function CHARACTERISTIC SET($\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2 \ : \ \mathsf{TCTL}^{\Delta}$)
    TargetSet := RIGHT PART($\mathsf{E}\varphi_1\mathsf{U}^k_{\sim c}\varphi_2$);

    repeat
        CurrentSet := TargetSet;
        TargetSet := TargetSet $\cup$ CurrentSet;
        TargetSet := TargetSet $\cup\big([\![\varphi_1]\!] \triangleright [z_{\bar{l}} \leftarrow 0]$ CurrentSet $\big)$;
        TargetSet := TargetSet $\cup\big([\![\neg\varphi_1 \wedge z_{\bar{l}} \leq k]\!] \triangleright$ CurrentSet $\big)$;
    until TargetSet = CurrentSet
    return $[z \leftarrow 0][z_{\bar{l}} \leftarrow 0]$TargetSet;
end function

VII. CONCLUSION

In this paper, we proposed a symbolic model-checking algorithm that computes the characteristic sets of some TCTL^{Δ} formulae and checks their truth values. Moreover, we gave an accurate description of an implementation of our algorithm using zones and DBMs, the same approach as the one used in model-checkers like UPPAAL or KRONOS, in order to avoid the state-space explosion problem caused by the explicit construction of region graphs. Indeed, to get a tool from this algorithm, no much work is now necessary : the computation of each step of the algorithm is precisely described in this paper. Moreover, our algorithm appears really as an extension of the zone algorithm for TCTL timed logic, and its complexity is not more important.

Thus, this work is the link that was missing between the theoretical work did by (Houda Bel Mokadem et al.) to abstract transient events in [13] (namely decidability and expressiveness) and a tool that would deal with TCTL^{Δ} timed logic.

REFERENCES

[1] L. Aceto, P. Bouyer, A. Burgueño, and K. G. Larsen. The power of reachability testing for timed automata. *Theoretical Computer Science*, 300(1–3):411–475, 2003.

[2] R. ALUR. Timed automata. In *In Proc. 11th Int. Conf. Computer Aided Verification (CAVâ99), Trento, Italy, July 1999, vol. 1633 of Lecture Notes in Computer Science, pp. 8â22.* Springer-Verlag, 1999.

[3] R. Alur, C. Courcoubetis, and D. Dill. Model-checking in dense real-time. *Information and Computation*, 104(1):2–34, 1993.

[4] R. Alur, C. Courcoubetis, N. Halbwachs, T. A. Henzinger, P.-H. Ho, X. Nicollin, A. Olivero, J. Sifakis, and S. Yovine. The algorithmic analysis of hybrid systems. *Theoretical Computer science*, 138(1):3–34, 1995.

[5] R. Alur, C. Courcoubetis, and T. A. Henzinger. Computing accumulated delays in real-time systems. *Formal Methods in System Design*, 11(2):137–156, 1997.

[6] R. Alur and D. Dill. Automata for modeling real-time systems. In *Proc. 17th International Colloquium on Automata, Languages and Programming (ICALP'90)*, volume 443 of *Lecture Notes in Computer Science*, pages 322–335. Springer, 1990.

[7] R. Alur and D. Dill. A theory of timed automata. *Theoretical Computer Science (TCS)*, 126(2):183–235, 1994.

[8] R. Alur, T. Feder, and T. A. Henzinger. The benefits of relaxing punctuality. *Journal of the Association for Computing Machinery (JACM)*, 43(1):116–146, 1996.

[9] R. Alur and T. A. Henzinger. Logics and models of real-time: a survey. In *Real-Time: Theory in Practice, Proc. REX Workshop 1991*, volume 600 of *Lecture Notes in Computer Science*, pages 74–106. Springer, 1992.

[10] R. Alur, S. La Torre, and G. J. Pappas. Optimal paths in weighted timed automata. In *Proc. 4th International Workshop on Hybrid Systems: Computation and Control (HSCC'01)*, volume 2034 of *Lecture Notes in Computer Science*, pages 49–62. Springer, 2001.

[11] G. Behrmann, A. Fehnker, Th. Hune, K. G. Larsen, P. Pettersson, J. Romijn, and F. Vaandrager. Minimum-cost reachability for priced timed automata. In *Proc. 4th International Workshop on Hybrid Systems: Computation and Control (HSCC'01)*, volume 2034 of *Lecture Notes in Computer Science*, pages 147–161. Springer, 2001.

[12] H. Belmokadem, B. Bérard, P. Bouyer, and F. Laroussinie. A new modality for almost everywhere propeties in timed automata. In *Proc. 16th International Conference on Concurrency Theory (CONCUR05)*, volume LNCS 3653, pages 110–124. Springer, 2005.

[13] H. Belmokadem, B. Bérard, P. Bouyer, and F. Laroussinie. Timed temporal logics for abstracting transient states. In *Proc. 4th International Symposium on Automated Technology for Verification and Analysis.* Springer., 2006.

[14] J. BENGTSSON and F. LARSSON. Uppaal, a tool for automatic verification of real-time systems. In *Masterˆas thesis, Department of Computer Science, Uppsala University (Sweden).* ISSN 0283-0574, 1996

[15] P. Bouyer, Th. Brihaye, and N. Markey. Improved undecidability results on weighted timed automata. *Information Processing Letters*, 2006. To appear.

[16] P. Bouyer, E. Brinksma, and K. G. Larsen. Staying alive as cheaply as possible. In *Proc. 7th International Workshop on Hybrid Systems: Computation and Control (HSCC'04)*, volume 2993 of *Lecture Notes in Computer Science*, pages 203–218. Springer, 2004.

[17] Th. Brihaye, V. Bruyère, and J.-F. Raskin. Model-checking for weighted timed automata. In *Proc. Joint Conference on Formal Modelling and Analysis of Timed Systems and Formal Techniques in Real-Time and Fault Tolerant System (FORMATS+FTRTFT'04)*, volume 3253 of *Lecture Notes in Computer Science*, pages 277–292. Springer, 2004.

[18] V. Bruyère, E. Dall'Olio, and J.-F. Raskin. Durations, parametric model-checking in timed automata with presburger arithmetic. In *Proc. 20th Annual Symposium on Theoretical Aspects of Computer Science (STACS'03)*, volume 2607 of *Lecture Notes in Computer Science*, pages 687–698. Springer, 2003.

[19] Z. Chaochen, C. Hoare, and A. Ravn. A calculus of duration. *Information Processing Letters*, 40(5):269–276, 1991.

[20] E. CLARKE, O. GRUMBERG, and D. PELED. Model-checking. In *The MIT Press, Cambridge, Massachusetts.* Springer-Verlag, 1999.

[21] C. DAWS. Analyse par simulation symbolique des systÃ¨mes temporisÃ©s avec kronos. In *Research report*. Verimag, 1997.

[22] C. Daws, A. Olivero, S. Tripakis, and S. Yovine. The tool KRONOS. In *Proc. Hybrid Systems III: Verification and Control (1995)*, volume 1066 of *Lecture Notes in Computer Science*, pages 208–219. Springer, 1996.

[23] C. Daws and S. Tripakis. Model-checking of real-time reachability properties using abstractions. In *In Proc. 4th Int. Conf. Tools and Algorithms for the Construction and Analysis of Systems, vol. 1384 of Lecture Notes in Computer Science, pp. 313â329.* Springer-Verlag., 1998.

[24] D. DILL. Timing assumptions and verification of finite-state concurrent systems. In *In Proc. the Workshop Automatic Verification Methods for Finite State Systems, vol. 407 of Lecture Notes in Computer Science, pp. 197â212*. Springer-Verlag., 1989.

[25] T. A. Henzinger, P.-H. Ho, and H. Wong-Toi. HYTECH: the next generation. In *Proc. 16th IEEE Real-Time Systems Symposium (RTSS'95)*, pages 56–65. IEEE Computer Society Press, 1995.

[26] T. A. Henzinger, X. Nicollin, J. Sifakis, and S. Yovine. Symbolic model-checking for real-time systems. *Information and Computation*, 111(2):193–244, 1994.

[27] Th. A. Henzinger. The theory of hybrid automata. In *Proc. 11th Annual Symposim on Logic in Computer Science (LICS'96)*, pages 278–292. IEEE Computer Society Press, 1996.

[28] Y. Kesten, A. Pnueli, J. Sifakis, and S. Yovine. Decidable integration graphs. *Information and Computation*, 150(2):209–243, 1999.

[29] R. Koymans. Specifying real-time properties with metric temporal logic. *Real-Time Systems*, 2(4):255–299, 1990.

[30] K. G. Larsen, P. Pettersson, and W. Yi. UPPAAL in a nutshell. *Journal of Software Tools for Technology Transfer (STTT)*, 1(1–2):134–152, 1997.

[31] J. Ouaknine and J. Worrell. On the decidability of Metric Temporal Logic. In *Proc. 20th IEEE Symposium on Logic in Computer Science (LICS'05)*, 2005.

[32] A. SCHRIJVER. Theory of linear and integer programming. In *Interscience Series in Discrete Mathematics and Optimization*. Wiley, 1998.

[33] S. YOVINE. Model checking timed automata. In *In School on Embedded Systems, vol. 1494 of Lecture Notes in Computer Science*. Springer-Verlag, 1998.

Automatic Diagnosing of Suspicious Lesions in Digital Mammograms

Abdelali ELMOUFIDI[1], Khalid El Fahssi[1], Said Jai-andaloussi[1], Abderrahim Sekkaki[1], Gwenole Quellec[2] Mathieu Lamard[2,3], Guy Cazuguel[2,4]

[1]Deprtment of Mathematics & Computer Sciences, Faculty of sciences, Hassan II University of Casablanca, Casablanca, Morocco.
[2]Inserm UMR 1101, Brest, F-29200 France.
[3]Univ Bretagne Occidentale, Brest, F-29200 France.
[4]Institut Mines-Telecom, Telecom Bretagne, UEB Dpt ITI, Brest, F-29200 France.

***Abstract*—Breast cancer is the most common cancer and the leading cause of morbidity and mortality among women's age between 50 and 74 years across the worldwide. In this paper we've proposed a method to detect the suspicious lesions in mammograms, extracting their features and classify them as Normal or Abnormal and Benign or Malignant for diagnosing of breast cancer. This method consists of two major parts: The first one is detection of regions of interest (ROIs). The second one is diagnosing of detected ROIs. This method was tested by Mini Mammography Image Analysis Society (Mini-MIAS) database. To check method's performance, we've used FROC (Free-Receiver Operating Characteristics) curve in the detection part and ROC (Receiver Operating Characteristics) curve in the diagnosis part. Obtained results show that the performance of detection part has sensitivity of 94.27% at 0.67 false positive per image. The performance of diagnosis part has 94.29% accuracy, with 94.11% sensitivity, 94.44% specificity in the classification as normal or abnormal mammogram, and has achieved 94.4% accuracy, with 96.15% sensitivity and 94.54% specificity in the classification as Benign or Malignant mammogram.**

***Index Terms*—Breast cancer, Mammogram, Computer-aided diagnosis, Segmentation, Regions of interest, Support Vector Machine, FROC analysis, ROC analysis.**

I. Introduction

Breast cancer is the most common cancer and the leading cause of morbidity and mortality among women's age between 50 and 74 years across the worldwide. Recent statistics have shown that one in 8 women in the United States and one in 10 women in Europe develop breast cancer during their lifetime [1],[2]. So, breast cancer is a major problem of public health, and the best strategy for the fight against breast cancer is early detection. For that reason, the mammography remains the best and most accurate tool for early detection of breast cancer [2],[3]. Reading and interpretation of mammogram is a crucial step. From where, Breast Imaging-Reporting and Data System (BI-RADS) of the American College of Radiology (ACR)[4], aims at providing a standardized classification system for reporting mammographic breast densities. Faced with the increase in the number of mammograms in recent decades, and the difficulty of reading and interpretation of mammograms, different research make the effort. Either, to automatically detect breast lesions through Computer Aided detection systems (commonly referred CADe). Either, To automatically interpret mammograms through Computer Aided Diagnostic Systems (commonly referred CADx). These systems are employed as a supplement to the radiologists' assessment.

Generally, the procedure to develop a Computer-Aided-Diagnosis (CAD) system, for diagnosing of suspicious regions in mammograms takes place in four steps: 1) Preprocessing step: this step is to prepare the mammograms for the next steps of operations (segmentaton, classification); 2)Detection of regions of interest :This step is to analyze the mammogram and extract the necessary information, for example, segmentation which divides the mammogram into multiple segments, edge detection which finds the edges of objects and helps us to find regions of interest; 3) Features extraction and selection of ROIs detected: In this step , we can identify specific patterns, shapes, density and texture; 4) Classification of ROIs: The purpose of this step is to classify the mammograms as Normal or Abnormal and malignant or benign [5][6].

In this paper, we've proposed an automatic method to detect and diagnosing of suspicious lesions in mammogram. The proposed method is a very accurate technique for detecting and diagnosing breast cancer by using mammogram.

Obtained results show the efficiency of projected method and make sure chance of its use in rising breast cancer detection and the diagnosing.

Paper organization : The rest of paper organized as follows: Section I: An introduction ; Section II: Related work; Section III: Materials and method ; Section IV : Features generation and extraction; Section V : Our proposed research; Section VI : Results and performance of proposed method ; Section VII : Conclusion; and references are given at the end.

II. Related Work

For detection and diagnosing of abnormalities in mammograms, A number of methods have been proposed, generally regrouped as: Statistical methods [7]; methods based wavelets [8][9]; Methods based Markov models [10]; Methods using machine learning[11], etc. Several researches have been published about computer breast cancer detection and diagnosis. For example, K. Ganesan et al. [12] presented an overview describe recent developments and advances in the field of computer-aided breast cancer diagnosis using mammograms. M. Veta et al.[13] presented an overview of methods that have proposed for the analysis of breast cancer histopathology images. Detection of ROIs is a capital step in development a computer-aided breast cancer diagnosis system. Many researchers have published on segmentation of breast tissue regions according to differences in density and texture, for detecting ROIs. For example, Adel et al. [14] used a method to segment mammograms into three distinct regions are : pectoral muscle, fatty regions, and fibroglandular regions using Bayesian techniques with Markov random field. Elmoufidi et al. [15][16] developed a method to Detect of ROIs in Mammograms using LBP algorithm, K-Means algorithm and GLCM algorithm. K. Hu et al. [2] published an approach to detect of suspicious lesions in mammograms by adaptive thresholding based on multiresolution. In other word, many methods have been used to feature extraction and classification. For example, Veena et al. [17] proposed a CAD System for Automatic Detection and Classification of Suspicious Lesions in Mammograms. Nasseer et al. [18] developed an algorithm for Classification of Breast Masses in Mammograms using SVM. L.Jelen et al. [19] developed a method for Classification of breast cancer malignancy using cytological images of fine needle aspiration biopsies. J. Malek et al. [20] proposed a system to Automatic Breast Cancer Diagnosis Based on GVF-Snake Segmentation, Wavelet Features Extraction and Fuzzy Classification. Nra Szkely et al. [21] used A Hybrid System for Detecting Masses in Mammographic Images. [22] Used an approach for Mammogram Segmentation by Contour Searching and Massive Lesion Classification with Neural Network. S. Timp et al. [23] developed a Computer-aided diagnosis with temporal analysis to improve radiologists.

The CAD systems are powerful tools that could aid radiologists to lead better results in diagnosing a patient.

III. Materials and Method

The proposed method checked by mini Mammography Image Analysis Society (mini-MIAS) database[24] and implemented using Seed Region Growing (SRG) algorithm, Local Binary Pattern (LBP) algorithm and support vector machine (SVM) classifier. The SRG to remove the pectoral muscle, the LBP to detect the regions of interest, and SVM to classify the mammograms as normal or abnormal and benign or malignant. SRG and LBP are two simples algorithms of segmentation and better choice for easy implementation. Using SVM as classifier because provide an effective and flexible framework from which to base CAD techniques for breast mammogram [25].

A. *Mammogram Database*

To checked the proposed method we've used the mini-Mammography Image Analysis Society (mini-MIAS) database[24]. The mammograms are in gray scale file format (Portable Grey Map - PGM), the size of every image is 1024×1024 pixels, and resolution of 200 micron. This database composed of 322 mammograms of right and left breast, from 161 patients, where 207 mamograms diagnosed as normal and 115 mammograms as abnormal (22 images of CIRC - Well-defined/circumscribed masses, 19 images of SPIC - Spiculated masses, 19 images of ARCH - Architectural distortion, 15 images of ASYM - Asymmetry, 26 images of CALC - Calcification and 14 images of MISC - Other, ill-defined masses) 52 mammograms malignant and 63 benign. Fig.1 shows the different objects in the mammograms.

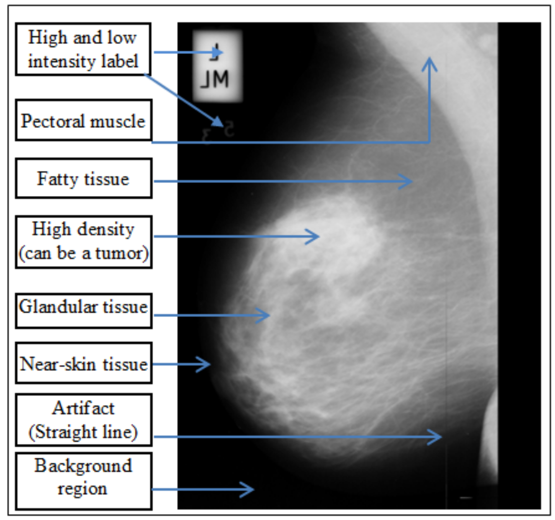


Fig. 1: The different elements in mammogram.

B. *Seed Region Growing (SRG)*

SRG algorithm for segmentation introduced by R. Adams et al. [25] is a simple method of segmentation which is free of tuning parameters and rapid. It's one of the better choice for easy implementation and applying it on a larger dataset. Seed region growing approach for image segmentation is to segment an image into regions with respect to a set of N seeds as presented in [12],[14] is discussed here.

C. *Local Binary Pattern (LBP)*

Local Binary Pattern (LBP) operator combines the characteristics of statistical and structural texture analysis. The LBP operator is used to perform gray scale invariant two-dimensional texture analysis. The LPB operator labels the pixel of an image by Thresholding the neighborhood (i.e. 3×3) of each pixel with the center value and considering the result of this Thresholding as a binary number [7],[26].When all the pixels have been labeled with the corresponding LBP codes, histogram of the labels are computed and used as a texture descriptor. Formally, given a pixel at (x_c, y_c), the resulting LBP can be expressed in decimal form as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P=1} S(i_p - i_c)2^P \qquad (1)$$

where : i_c and i_p are, respectively, gray-level values of the central pixel and P surrounding pixels in the circle neighborhood with a radius R, and function s(x) is defined as:

$$S(x) = \{^{1, x \succeq 0}_{0, x \prec 0} \tag{2}$$

D. Support Vector Machine (SVM)

SVM classifier algorithm, developed from the machine learning community is a discriminative classifier formally defined by a separating hyperplane. The hyperplane is determined in such a way that the distance from this hyperplane to the nearest data points on each side, called support vectors, is maximal [27]. SVM classifiers can be extended to nonlinearly separable data with the help of kernel function application on the data to make them linearly separable [28]. An approach with wavelet SVM was discussed in [29]. Details about SVM, its application to diagnose of breast cancer was discussed in [26][30].

IV. Feature generation and extraction

Below a list of eighteen features selected to use as input parameters of SVM classifier for training and testing our proposed method.

*a) **Mean Value**:* μ represents the average of pixels in the segmented ROI.

$$\mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} I(i,j) \tag{3}$$

Where: I(i,j) is the pixel value at point (i,j) in ROI of size MxN.

*b) **Standard Deviation**:* σ describes the dispersion within a local region.

$$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (I(i,j) - \mu)^2} \tag{4}$$

*c) **Entropy**:* H used to describe the distribution variation within ROI.

$$H = -\sum_{k=1}^{L-1} P_k * log_2(P_k) \tag{5}$$

Where: P_k is the probability of the k^{th} grey level, L is the total number of grey levels.

*d) **Skewness**:* S is a number characterizes the shape of the distribution.

$$S = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [\frac{I(i,j) - \mu}{\sigma}]^3 \tag{6}$$

Where: I(i,j) the pixel value at point (i,j), μ the mean and σ the standard deviation.

*e) **Kurtosis**:* **K** measures the flatness of a distribution relative to a normal distribution.

$$K = \{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [\frac{I(i,j) - \mu}{\sigma}]^4\} - 3 \tag{7}$$

*f) **Uniformity**:* **U** is a texture measure based on histogram :

$$U = \sum_{k=0}^{L-1} P_k^2 \tag{8}$$

Where: P_k the probability of the k^{th} grey level.

*g) **Sum Entropy**:* **SE** is a logarithmic function of the ROI in consideration.

$$SE = -\sum_{i=2}^{2N_g} p_{x+y}(i) log\{p_{x+y}(i)\}. \tag{9}$$

*h) **Sum Average**:* SA is found from the ROI in consideration and the size of the gray scale

$$SA = \sum_{i=2}^{2N_g} i p_{x+y}(i) \tag{10}$$

*i) **Difference variance**:* **DV** is a variance measure between the ROI intensities calculated as a function of the SE calculated previously

$$DV = \sum_{i=2}^{2N_g} (i - SE)^2 p_{x-y}(i) \tag{11}$$

*j) **Difference entropy**:* **DE** is an entropy measure which provides a measure of no uniformity while taking into consideration a different measure obtained from the original image

$$DE = -\sum_{i=2}^{2N_g} p_{x-y}(i) log\{p_{x-y}(i)\}. \tag{12}$$

*k) **Inverse Difference Moments**:* IDM is a measure of the local homogeneity.

$$IDM = \sum_i \sum_j \frac{1}{1 + (i-j)^2} p(i,j). \tag{13}$$

*l) **Area**:* **A** is the sum of the number of all pixels (x) within segmented ROI.

$$A = \sum_{x \in ROI} 1. \tag{14}$$

*m) **Perimeter**:* P is the length of a polygonal approximation of the boundary (B) of ROI:

$$P = \sum_{x \in B} 1. \tag{15}$$

*n) **Convexity**:* **C(S)** is the ratio of the ROI area and its convex hull, the convex hull is the minimal area of the convex polygon that can contain the ROI:

$$C(S) = \frac{A}{Area(CH(S))}. \tag{16}$$

Where: S is a ROI, CH(S) is its convex hull and A is the ROI's area.

*o) **Compactness**:* **C** is a measure of ROI's shape, which indicates how much the ROI is compact :

$$C = \frac{P^2}{4\pi A}. \tag{17}$$

Where : P the ROI's perimeter, A ROI's area.

*p) **Aspect Ratio**:* **AR** corresponds to the aspect ratio of the smallest window fully enclosing the ROI in both directions (see Fig.2):

$$AR = \frac{D_y}{D_x}. \tag{18}$$

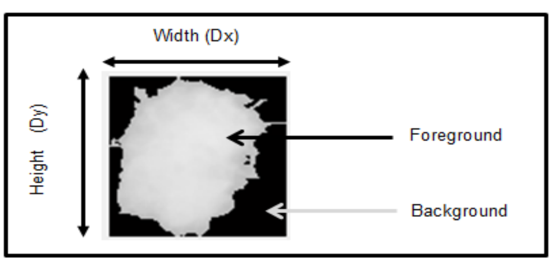


Fig. 2: Example of ROI window from which some features will be extracted.

Where: Dy the height, Dx the width of window in Fig.2

*q) **Area Ratio**:* The Area Ratio (R_Area) is defined by dividing the area of the segmented ROI in pixels by the area of the same window given in Fig.2 :

$$R\_Area = \frac{Area\_ROI(in\ pixels)}{Area\_window(in\ pixels)}. \tag{19}$$

Where: $Area\_window =$ Dx*Dy , Dx is the width's ROI and Dy is the height's ROI. The value of R_Area will range from 0 to 1. So, It takes small values for ROI with appendices and branches emitted from it, and larger values for more compacted and rounded objects.

*r) **Perimeter Ratio**:* R_Perim presents the ratio between the perimeter of the segmented ROI to the perimeter of the same rectangular window of fig.2, this can be written as:

$$R\_Perim = \frac{Perimeter\_ROI(in\ pixels)}{Perimeter\_window(in\ pixels)}. \tag{20}$$

V. Our proposed research

In this paper, we've proposed a method for automatic detecting and diagnosis of suspicious lesions in mammograms. The proposed method consists four major blocks, namely: (1) Mammogram preprocessing performs three steps are: Remove the labels and additional objects in mammograms; Suppressed the background and the pectoral muscle; Eliminate the artifact, digital noise and contrast enhancement. (2) Segmentation and detection of ROIs, In this block, we've segmented and detected the ROIs are done using Local Binary Pattern . The details about LBP are discussed above. (3) Extraction and selection of features for each ROIs, in this case, we've combined the different types of features (size, density, shape, contrast and texture). (4) Classification of mammograms to normal or abnormal in the first time and benign or malignant in the second time using Support Vector Machine (SVM).

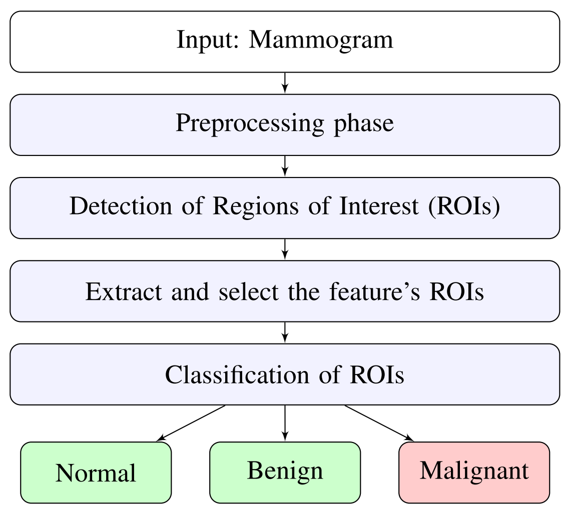


Fig. 3: Steps of the proposed method.

A. *Preprocessing*

The mammography can cause some additional objects at the resulting mammograms, like: artifact, noises, labels, etc. According to [31], the mammograms contain several sorts of noise and imaging artifacts. So, preprocessing step will be applied to get rid of the extra objects and enhance the standard of mammograms. Generally, the preprocessing step is to prepare the mammograms for the next steps, such as: segmentation of ROIs , selection and extraction of features of ROIs, and classification of ROIs. In this step, the aim is to extract only the breast profile region without additional objects, and without background. First, a threshold value is used to get rid of the labels and also the further objects within the mammograms. Second, we've used an automatic technique to take away further background, and detected mammogram orientation. From where, the pectoral muscle is within the top corner in right or left, the seed point of SRG is J[5,5] or J[5, y-5], (were J: is the mammogram when the background has removed, [x,y]=size(J)) and we've used a minimal threshold value for giving a good result with all type of mammogram (Fatty , Fatty-glandular, Dense-glandular) . Third, we used 2D median filter in a 3-by-3 neighborhood connection to remove additional objects (artifact and noise). In addition, the mammogram is basically low contrast [1], so, we've applied a step of enhancement of contrast(see Fig.4).

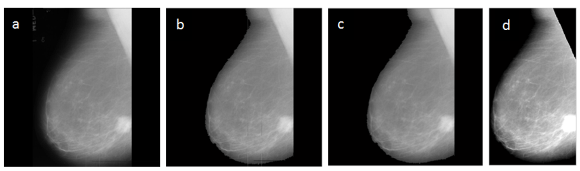


Fig. 4: Mammogram preprocessing: (a) Original mammogram, (b) Additional objects and Labels suppressed, (c) Noise and artifact removal, (d) Remove additional background and pectoral muscle. In addition, contrast enhancement.

B. *Detection of Regions of Interest (ROIs)*

Detection ROIs is a capital step in developing a CAD system, detecting several false positive result a weak system. To perform this task, we've implemented the Local Binary Pattern (LBP) algorithm for detecting the ROIs.

*1) **Experimental results of detection part**:*

*a) **Example 1**:* The first example deals with of normals mammograms. Fig.5.

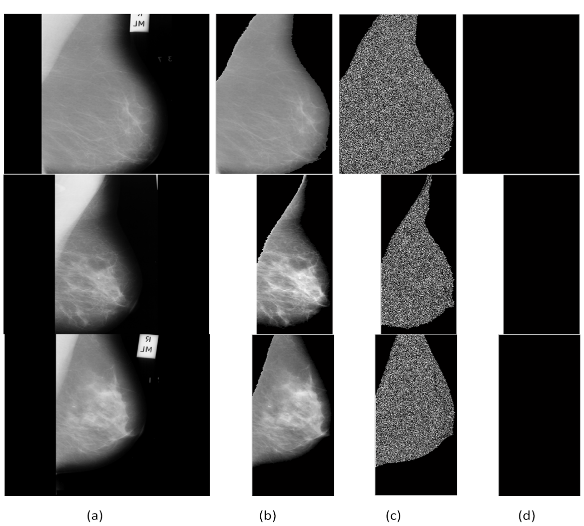


Fig. 5: Detection of ROIs :(a) Original mammograms, (b) Mammograms after preprocessing, (c) LBP algorithm is applied,(d) Any regions of interest have detected .

*b) **Example 2**:* Mammogram contains a single lesion has been correctly detected without any false positive. Fig.5.

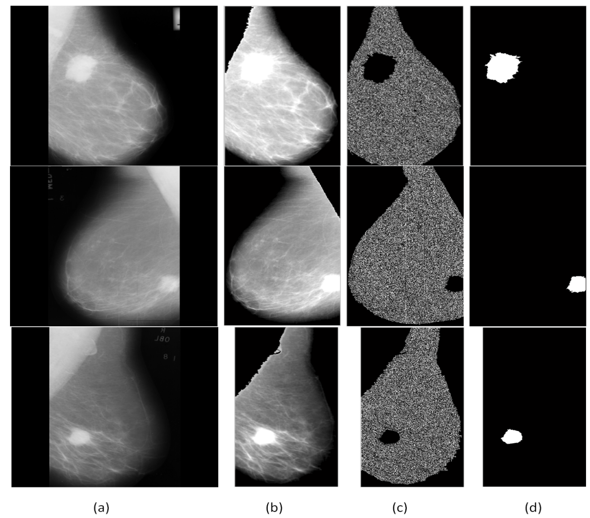


Fig. 6: ROIs detected contain the lesion without any false positive :(a) Original mammograms, (b) Mammograms after preprocessing, (c) LBP algorithm is applied,(d) regions of interest have detected .

*c) **Example 3**:* Mammograms which contains a single lesions has been correctly detected with other ROIs detected as false positive. Fig.6.

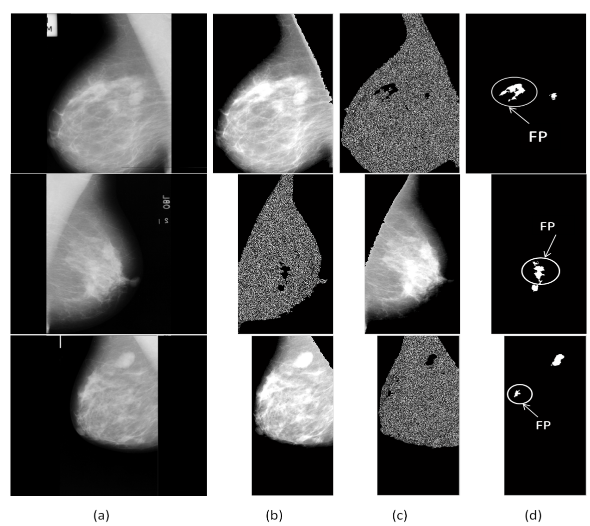


Fig. 7: ROIs detected contain the lesion and false positives :(a) Original mammograms, (b) Mammograms after preprocessing, (c) LBP algorithm is applied,(d)Regions of interst have detected.

*2) **Example of ROIs automatically detected**:* In the figure below, some ROIs automatically detected.

C. *Diagnosis of Regions of Interest (ROIs)*

After detection of regions of interest and extraction their features, the next step is to classify them as normal or

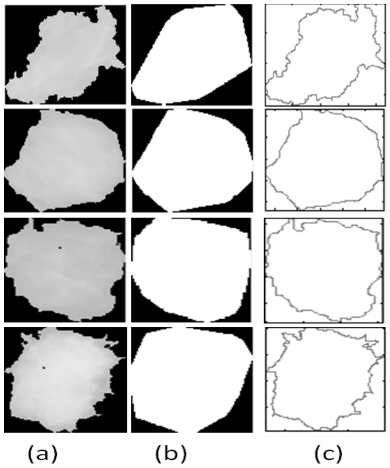


Fig. 8: Examples of ROIs detected: (a) ROIs detected, (b) convex hull of ROIs, (c) The contour of ROIs.

abnormal in the first time and as benign or malignant in the second time. One among the novelties of proposed method that a new technique to detect all suspected areas in mammogram (not just the detection of lesions) and consider them as regions of interest (ROIs). If no regions of interest detected, the mammogram is normal. In the case of detection of multiple ROIs, we are going to separate them one by one and extracted their features separately (one by one), then diagnosing them. first, in the case all ROIs belong in the same mammogram are normal, then the mammogram is normal. Otherwise, the mammogram is abnormal. Second, in the case all ROIs belong in the same mammogram are benign, then the mammogram is benign. Otherwise, the mammogram is malignant. In addition, this algorithm is able to detect the masses and the calcification.

*1) **Experimental results of diagnosis part**:* Next three figures show details of the mentioned method. 1) Button "download" for downloading a new mammogram. 2) Button "Pre-processing" is to apply a preprocessing step on original mammogram (remove label, noise, pectoral muscle and additional background). 3) Button "Apply LBP" is to apply local binary pattern algorithm on the result mammogram after preprocessing step . 4) Button "Extract ROIs" is to extract all detected objects as ROIs. If we get just one ROI, only the button "ROI1" is going to enable. If we get two ROIs, the two buttons "ROI1" and "ROI2" are going to enable, and so on. 5) Button "ROI1" is to select the first ROI, the button "ROI2" is to select the second ROI, and so on. Button "Clac-features" is to extracte ROI's features selected in the previous step. 6) Button "add-feature" is to add the features in our database. 7) Button "Classify" is to classify the ROI selected to normal, benign or malignant. if the ROI selected is normal a white button appears on the screen containing the text normal, if the ROI selected is benign a green button appears on the screen contains the text benign, if the ROI selected is malignant a red button appears on the screen containing the text malignant. In addition, if we get many ROIs, we are going to classify them one by one, if all the ROIs are normals, the mammogram is normal. if there are at least one ROI benign and no ROI malignant, the mammogram is benign. If there are at least one ROI malignant, the mammogram is malignant.

*a) **Example 1**:* A normal mammogram correctly detection and diagnosis .

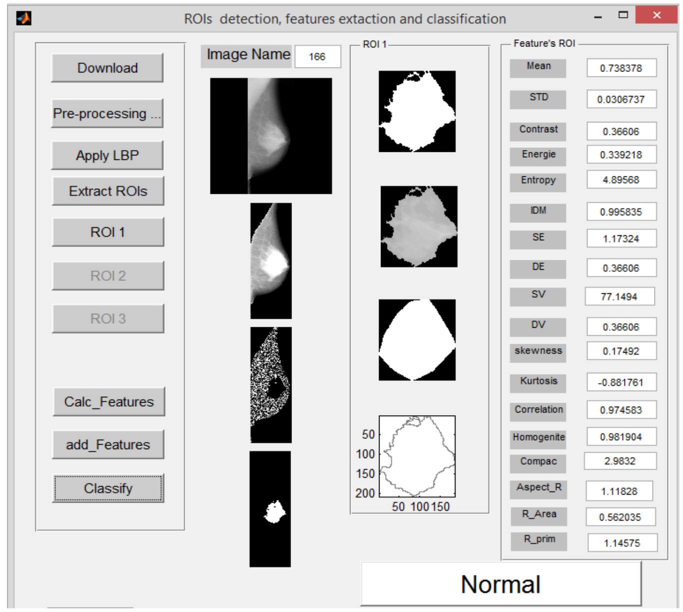


Fig. 9: Example 1: A normal mammogram Exactly diagnosis.

*b) **Example 2**:* A benign lesion correctly detection and diagnosis without false positive

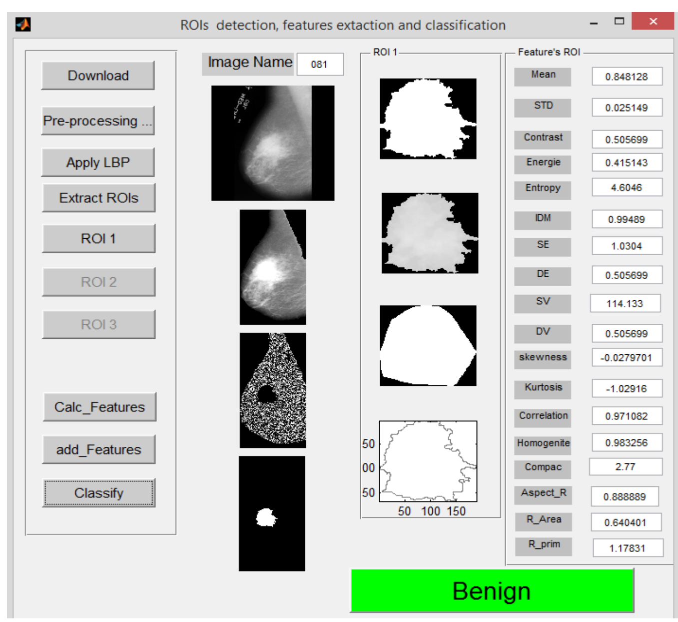


Fig. 10: Example 2: A Benign mammogram correctly diagnosing.

*c) **Example 3**:* A lesion malignant correctly detection/diagnosis without false positive

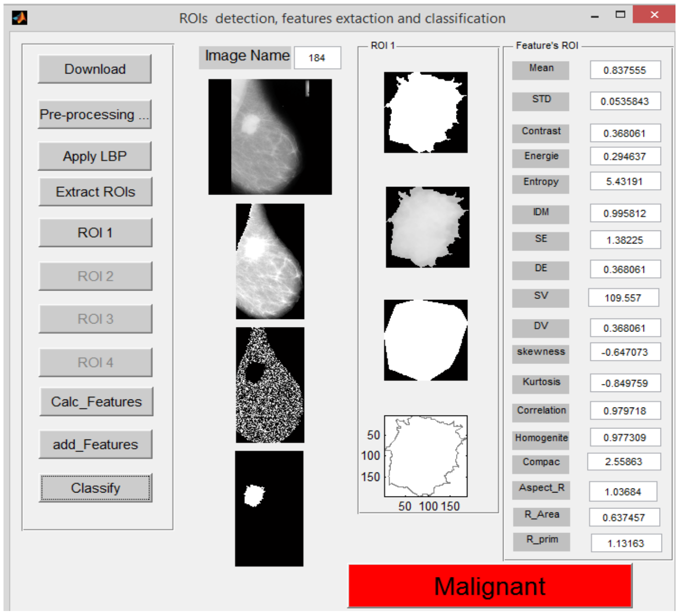


Fig. 11: Example 3: A Malignant mammogram correctly diagnosing..

VI. RESULTS AND PERFORMANCE

Our global method were checked on 322 mammograms from mini-MIAS database. The detail about mini-MIAS database is given above.

A. *Performance of detection part*

Each segmentation and classification result needs evaluation of its performance. Generally, there are three types of performance evaluations of algorithms and approaches proposed for medical imaging processing (detection of regions of interest, segmentation and classification): The first type involves qualitative assessment, the second is quantitative assessment involving the ground truth evaluation and the third is a statistical evaluation [31].

Detected and selected the suspicious regions in mammogram is a crucial step in developing a CAD system, and detecting more regions d'interet as false positive, result a weak system. For that, we've considered a ROI correctly detected if its area is overlapped by at least of 75% from ground truth. We have obtained a good detection result, i.e., 100%, for MISC and 95.45%, for CIRC. The detection result of SPIC (89.47%) is relatively reliable, because the overlapping of some SPIC is least of 75%, hence, we considered as false negative. Generally, we've obtained a sensitivity of 94.27% at 0.67 False Positive per Image in the detection stage.

FROC curve, representing the True Positive Fraction (TPF) according False Positive per Image (FP/I) see the detail below:

$$\text{False Positive per Image} = \frac{\text{Number of False Positive}}{\text{Number of image}} \qquad (21)$$

TABLE I: The obtained results grouped by anomaly classes .

Class of abnormality present	Number of images	Sensitivity (%)
Normal	207	94.21%
CIRC	22	95.45%
SPIC	19	89.47%
ARCH	19	94.73%
ASYM	15	93.3%
CALC	26	92.31%
MISC	14	100%
Total	322	94.21%

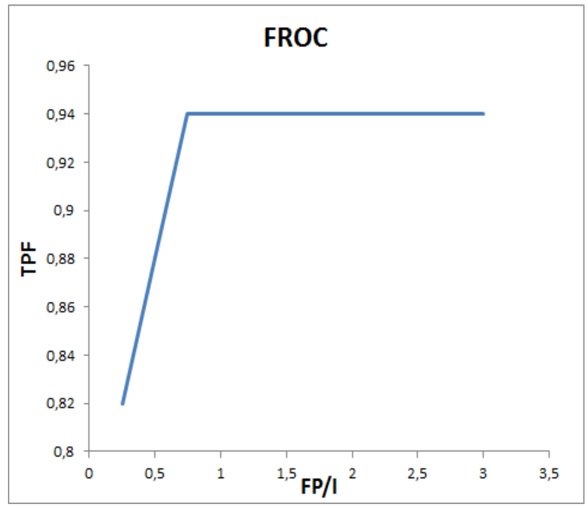


Fig. 12: FROC curve.

The evaluation procedure is as following: the database is divided into two parts: the first one for training contains the half of database (161 mammograms from 322 mammograms) selected aleatory, the second one for testing contains the rest of database(161 mammograms) the detail of the database distribution between training and testinig is given below:

TABLE II: Images's number used to train SVM Classifier.

Image	Normal	Abnormal	Benign	Malignant
Training	104	57	31	26
Testing	103	58	32	26
Total	207	115	63	52

B. *Performance of diagnosing part*

We have evaluated the performance of proposed method by calculating of accuracy, sensitivity and specificity for normal or abnormal case and benign or malignant case.

TABLE III: Diagnosing accuracy of normal or abnormal cases

	Training		Testinig	
	Normal	Abnormal	Normal	Abnormal
Normal	(102)TP	(2)FP	(100)TP	(4)FP
Abnormal	(2)FN	(53)TN	(3)FN	(52) TN

Diagnosing part of our method has achieved 94.29% accuracy, with 94.11% sensitivity and 94.44% specificity. Fig.12 shows the ROC curve of the proposed method.

TABLE IV: Diagnosing accuracy of benign or malignant cases

	Training		Testing	
	Benign	Malignant	Benign	Malignant
Benign	(30)TP	(1)FP	(29)TP	(2)FP
Malignant	(2)FN	(24)TN	(3)FN	(23) TN

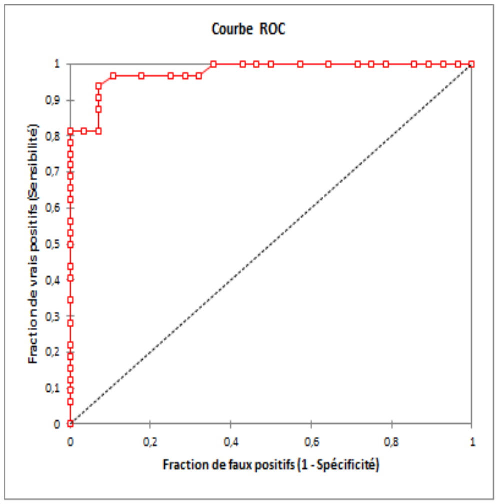


Fig. 13: ROC curve.

C. Comparison our method with existing papers.

TABLE V: Comparison our method's Performance with articles recently published .

Authors	Method proposed	Accuracy
K. Hu et al.[2]	Detection of suspicious lesions in mammograms	91.3%
Veena et al.[17]	Detection & Classification of Suspicious Lesions in Mammograms	92.13%
Nasseer et al.[18]	Classification of Breast Masses in Digital Mammograms	93.069%
K. Ganesan et al.[27]	Classification of Mammograms Using Trace Transform Functionals	92.48 %
Our method	Automatic Diagnosing of Suspicious Lesions in Digital Mammograms	94.29%

VII. Conclusion

In this paper, an automatic algorithm to breast cancer detection and diagnosing is implemented using the MATLAB environment. our algorithm's performance has evaluated using FROC curve in detection part and ROC curve in diagnosis part. Obtained results show the efficiency of this method and comparable to different solutions. The proposed algorithm will contribute to determination of main drawback in diagnostic procedure mammogram, such as: detection and diagnosing of the masses and also the calcification. The efficiency of the planned method confirms possibility of its use in up the Computer-Aided Diagnosis system.

References

[1] A. ELMOUFIDI Member IEEE et al., "Automatically Density Based Breast Segmentation for Mammograms by using Dynamic K-means Algorithm and Seed Based Region Growing," I2MTC 2015 - International Instrumentation and Measurement Technology Conference , PISA, ITALY, MAY 11-14, 2015.

[2] K. Hu et al., "Detection of suspicious lesions by adaptive thresholding based on multiresolution analysis in mammograms," IEEE Trans on Instrumentation and Measurement, vol. 60, no. 2, pp. 462-472, 2010.

[3] A. Ferrero Fellow IEEE et al., "Uncertainty evaluation in a fuzzy classifier for microcalcifications in digital mammography," I2MTC 2010 - International Instrumentation and Measurement Technology Conference Austin, TX, 3-6 May 2010.

[4] American College of Radiology. American College of Radiology Breast Imaging Reporting and Data System (BIRADS). 4^{th} ed., American College of Radiology, Reston, VA 2003.

[5] A.Jalalian et al., "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound," Clinical Imaging, 37 ,2013 420426.

[6] S. Shirmohammadi and A. Ferrero, "Camera as the Instrument: The Rising Trend of Vision Based Measurement," IEEE Instrumentation and Measurement Magazine, Vol. 17, No. 3, June 2014, pp. 41-47. DOI: 10.1109/MIM.2014.6825388

[7] H. Chan et al., "Computerized analysis of mammographic micro calcifications in morphological and feature spaces," Medical Physics, vol.25, no.10, pp.2007-2019, 1998.

[8] A. Mencattini et al., "Mammographies images enhancement and denoising for breast cancer detection using dyadic wavelet processing," IEEE Trans. Instrumentation Measure., 57: 1422-1430. DOI: 10.1109/TIM. 2007. 915470.

[9] T. Wang and N. Karayiannis, "Detection of microcalcification in digital mammograms using wavelets," IEEE Trans. Medical Imaging, vol.17, no.4, pp.498-509, 1998.

[10] H. Li et al., "Marcov random field for tumor detection in digital mammography," IEEE Trans. Medical Imaging, vol.14, no.3, pp.565-576, 1995.

[11] Drew P.J.Monson (J.R.T), "Artificial neural networks", Surgery volume 127, 2000, pp. 3-11.

[12] Karthikeyan Ganesan et al., "Computer-Aided Breast Cancer Detection Using Mammograms," IEEE Reviews in biomedical engineering, vol. 6, 2013.

[13] M. Veta, et al. "Breast cancer histopathology image analysis: a review.", IEEE transactions on bio-medical engineering, vol. 61, no. 5, pp. 140011, May 2014.

[14] Adel et al. , "Statistical Segmentation of Regions of Interest on a Mammographic Image," EURASIP Journal on Advances in Signal Processing 2007, Article ID 49482, 1-8 2007.

[15] Abdelali Elmoufidi et al.,"Detection of Regions of Interest in Mammograms by Using Local Binary Pattern, Dynamic K-Means Algorithm and Gray Level Co-occurrence Matrix," 2014 Fifth International Conference on Next Generation Networks and Services (NGNS'14) 28-30 May 2014, Casablanca, Morocco.

[16] A. Elmoufidi et al, "Detection of Regions of Interest in Mammograms by Using Local Binary Pattern and Dynamic K-Means Algorithm," International Journal of Image and Video Processing: Theory and Application Vol. 1, No. 1, 30 April 2014. ISSN: 2336-0992.

[17] Veena, et al., "CAD Based System for Automatic Detection et Classification of Suspicious Lesions in Mammograms," International Journal of Emerging Trends et Technology in Computer Science (IJETTCS) ISSN 2278-6856 , Volume 3, Issue 4 July-August 2014.

[18] Nasseer M. Basheer et al., "Classification of Breast Masses in Digital Mammograms Using Support Vector Machines," International Journal of Advanced Research in Computer Science and Software Engineering ISSN: 2277 128X, Volume 3, Issue 10, October 2013.

[19] L. Jelen et al., "Classification of breast cancer malignancy using cytological images of fine needle aspiration biopsies," int. j. appl. math. comput. sci., 2008, vol. 18, no. 1, 7583 doi: 10.2478/v10006-008-0007-x.

[20] Jihene Malek et al., "Automated Breast Cancer Diagnosis Based on GVF-Snake Segmentation, Wavelet Features Extraction and Fuzzy Classification," J Sign Process Syst. DOI: 10.1007/s11265-008-0198-2

[21] N.Szkely et al., "A Hybrid System for Detecting Masses in Mammographic Images," IEEE Trans. Instrum. Meas., vol. 55, no. 3 June 2006.

[22] Cascio D. et al., "Mammogram Segmentation by Contour Searching and Massive Lesion Classification with Neural Network," Institute of Electrical and Electronic Engineering (IEEE), 2006.

[23] S. Timp et al., "Computer-aided diagnosis with temporal analysis to improve radiologists' interpretation of mammographicmass lesions,"

IEEE Trans. Inform. Technol. Biomedicine, vol. 14, no. 3, pp. 803808, May 2010.

[24] J. Suckling et al., ”The Mammographic Image Analysis Society digital mammogram database,” Exerpta Medica, International Congress Series 1069 pp. 375-378., 1994.

[25] Jacob Levman, ”Classification of Dynamic Contrast-Enhanced Magnetic Resonance Breast Lesions by Support Vector Machines”, IEEE Transactions On Medical Imaging, Vol. 27, No. 5, May 2008.

[26] H. X. Liu, et al., ”Diagnosing Breast Cancer Based on Support Vector Machines”, J. Chem. Inf. Comput. Sci. 2003, 43, 900-907.

[27] K. Ganesan et al., ”One-Class Classification of Mammograms Using Trace Transform Functionals”, IEEE Transactions on Instrumentation and Measurement, Vol. 63, No. 2, February 2014.

[28] K. R. Muller, et al. ”An introduction to kernel based learning algorithms”, IEEE Trans.Neural Netw., vol. 12, no. 2, pp. 181201, Mar. 2001.

[29] M. Shen et al., ”A prediction approach for multichannel EEG signals modeling using local wavelet SVM”, IEEE Trans. Instrum. Meas., vol. 59, no. 5, pp. 14851492, May 2010.

[30] L. Wei, et al., ”A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications,” IEEE Trans. Med. Imag., vol. 24, no. 3, pp. 371380, Mar. 2005.

[31] Stylianos.D et al., ”A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry,” computer methods and programs in biomedicine 2011, 47-63.

Detection and Counting of On-Tree Citrus Fruit for Crop Yield Estimation

Zeeshan Malik, Sheikh Ziauddin, Ahmad R. Shahid, and Asad Safi
Department of Computer Science,
COMSATS Institute of Information Technology,
Park Road, Islamabad

***Abstract*—In this paper, we present a technique to estimate citrus fruit yield from the tree images. Manually counting the fruit for yield estimation for marketing and other managerial tasks is time consuming and requires human resources, which do not always come cheap. Different approaches have been used for the said purpose, yet separation of fruit from its background poses challenges, and renders the exercise inaccurate. In this paper, we use k-means segmentation for recognition of fruit, which segments the image accurately thus enabling more accurate yield estimation. We created a dataset containing 83 tree images with 4001 citrus fruits from three different fields. We are able to detect the on-tree fruits with an accuracy of 91.3%. In addition, we find a strong correlation between the manual and the automated fruit count by getting coefficients of determination R^2 up to 0.99.**

Keywords*—*Precision agriculture; yield estimation; k-means segmentation; leaf occlusion;, illumination; morphology

I. Introduction

In citrus groves, yield estimation is typically carried out a few weeks earlier to fruitage to estimate the resource requirement. The manual process of crop assessment is occasionally done by hand pawns. Preferably, yield would be estimated at numerous periods during crop development but it requires enormous labor cost and time. Precise, low-cost yield estimation is important for cultivators, especially if it can be done timely in the growing season. As orange juice needs to be processed within 48 hours of harvesting, orange juice manufacturers need suppliers to provide accurate yield estimates to guarantee that their juice plants can run at full capacity given the time constraints. Additionally, more accurate yield estimates will allow farmers to plan more precisely for their harvesting labor and other logistical needs. Image processing can help in improving the decision making process for irrigation, fruit sorting and yield estimation [1]. Detection of fruit is important as the subsequent fruit counting depends on accurate detection.

Citrus fruit have different properties that can be used for detection purposes. The most natural property to be used for such purpose is the color. In the past, color has been extensively used [2], [3], [4], [5], [6], [7], [8], [9]. Color on is own may not provide enough information as it may change depending upon how ripe the fruit is. For instance, unripe oranges may be greenish, while over-ripe oranges may be brownish. That poses a challenge when detection is based on filtering of orange color only.

Lighting may pose another challenge as the oranges may appear differently, under varying lighting conditions, such as bright sunlight, cloudy, and evening. Figure 1 shows a tree image under two varying light conditions. The image on the left is captured on a cloudy day and the one on the right is captured under direct sunlight. Under cloudy conditions, the images are more consistent in terms of brightness and intensity changes. While in the latter case, things may look different if they are exposed to direct sunlight as opposed to when they are under a shadow.

Another problem is that of occlusion. Occlusion may be caused either by leaves or by the neighboring oranges, as can be seen from Figure 2. In the first case, a single orange may be counted as more than one, as leaves may affect the shape of the orange. In the later case, may fruits may appear as one larger fruit with irregular shape. In the first case, a need arises for a way to count two or more smaller citrus blobs as one, while in the later case we need to break down a larger blob into smaller units, where each unit is counted as a separate fruit.

II. Related Work

Despite the above-mentioned challenges involved, image processing has been used with good results for automated fruit counting and yield estimation. In the past, a variety of techniques have been used across diverse fruits. Stajnko et al. [10] used a combination of normalized chromaticity coordinates (RGB) and thermal images to estimate on-tree count of apple fruit. In order to separate pixels more precisely, they also used the normalized difference index (NDI). Using their technique, they were able to achieve correlation coefficient R^2 of 0.83 to 0.88 at different times during the vegetation period. They observed that the values of R^2 were improved during the fruit ripening period. In a subsequent work [11], they used surface texture and color features to segment apples from the background. HSI color space was used for segmentation using a threshold which was empirically selected. Red color intensity ($3*R-G-B$), contrast, uniformity and entropy were the main features used. They achieved a detection accuracy of 89% with a false detection of 2.2%.

Zhou et al. [12] also used color features for on-tree apple counting. Fifty tree images were captured at two different times during the season. The images were captured under normal daylight conditions. For the images captured earlier in the season, they used color RGB difference ($R-B$ and $G-R$) while two different color models were used for the images captured during ripening season. They achieved R^2 scores of 0.8 and 0.85 for the two seasons, respectively. Regunathan and Lee [13] presented a technique for citrus detection and

Fig. 1: Citrus tree images captured under two different sunlight conditions.

Fig. 2: Orange fruit occluded by leaves (left). Orange fruit overlapping each other (right).

size finding. Their technique relies on an untrasonic sensor to determine the distance between the camera and the fruit. They used hue and saturation features in combination with three different classifiers namely Beyesian, neural network and Fischer's linear discriminant to segment and differentiate fruit from the background.

Hannan et al. [14] used red chromaticity coefficient for image enhancement which helped in better segmentation in variable light orange tree images. They used shape analysis technique to detect overlapped oranges achieving a detection rate of 90% with 4% false detection. Yet another citrus fruit counting scheme was presented by Chinchuluun et al. [15]. They used Bayesian classification in addition to morphological operations and watershed segmentation for citrus segmentation and detection. They reported a correlation coefficient value of 0.89 between the manual and the automated fruit count. Billingsley [16] presented a machine vision system for counting Macadamia nuts on a harvester roller. RGB color features were used to separate the nuts and the leaves from the roller. To segment the nuts from the leaves, they used a modified version of circular Hough transform. Cakir et al. [17] used color histograms and shape analysis to detect and count orange fruit on tree images. They were able to achieve a detection accuracy of 76.5%.

Annamalai and Lee [18] developed an algorithm for citrus counting. They used hue and saturation color planes and utilized histogram information to segment the fruit from the background and the leaves. Erosion and dilation operations were used to remove noise. They found R^2 of 0.76 between the manual count and the count obtained using their algorithm.

In this paper, we present a technique for automated yield estimation using k-means segmentation. The proposed technique is able to detect and count citrus fruit by catering for changes in color both due to lighting conditions as well as the state of ripeness of the fruit. Our system also takes care of the cases where the fruit is occluded either by the leaves or by other fruit. The experimental evaluation shows the accuracy and robustness of the proposed scheme as described in the subsequent sections.

The rest of this paper is organized as follows: The proposed scheme is presented in Section 2. The results of our experiments are described in Section 3 while Section 4 concludes this paper.

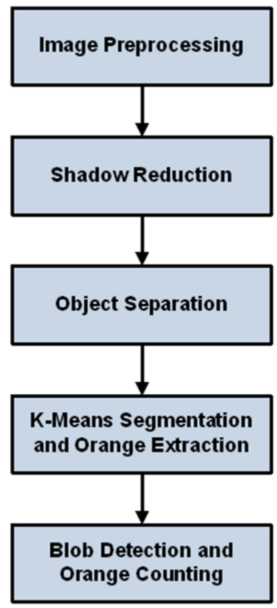


Fig. 3: An overview of the proposed technique.

III. Proposed Technique

The technique we proposed here uses k-means segmentation algorithm on orange tree images. First, a few preprocessing steps are performed including noise reduction and image enhancement. Next, we minimize shadow effects from the image. Then, we extract oranges using blob detection and size calculation which is then followed by the final yield estimation. Figure 3 gives an overview of the proposed technique by showing its different steps in a sequential order. The details of these steps are given in the following sub-sections.

1) Image Preprocessing: In general, images contain a lot of redundant information not needed for the application in hand. The image may contain noise which makes the edge detection and the segmentation tasks prone to errors. Therefore, it is often necessary to perform certain type of noise reduction and image enhancement before any meaningful processing of the images. In this paper, we use the Perona-Malik model for image enhancement and noise reduction [19]. It smoothes the image without effecting significant features of the image such as lines, edges or other important details that are necessary for analysis and interpretation of the images. In this model, image is smoothed using the following mathematical relation.

Fig. 4: Shadow reduction step. The image on left is before the shadow reduction while that on right is the same image after shadow reduction step is applied.

$$\frac{\partial I}{\partial t} = \text{div}(c(x, y, t)\nabla I) = c(x, y, t)\Delta I + \nabla c.\Delta I \quad (1)$$

where div denotes the divergence operator, while ∇ and Δ represents gradient and Laplacian, respectively. The diffusivity function in Perona-Malik model is given by the following equation.

$$c(x, y, t) = g(||\nabla I(x, y, t)||) \quad (2)$$

where g is a nonnegative, monotonically decreasing function with $g(0) = 1$.

A. Shadow Reduction

Lighting conditions play a very important role in the performance of many computer vision algorithms. In particular, shadows in an image can cause problems in recognition, segmentation and tracking algorithms to produce desirable outcomes. Distinct objects can be combined through shadows and they can obscure object recognition systems. When dealing with outdoor images, shadow minimization is an important task. In order to minimize shadow, we adjust the luminosity of the image. First, we convert the RGB image to $L*a*b$ image, where L is the luminosity layer, while a and b represents color-opponent dimensions. Next, we increase the luminosity of the image which results in reduced shadow effect in the image. Finally, the image is converted to RGB color space after replacing the luminosity layer with the processed data. Figure 4 shows an image after shadow reduction step is applied.

B. Object Separation

One of the main challenges in orange counting is that of overlapped oranges. Due to overlapping, multiple fruits may be counted as a single fruit which negatively impacts the fruit counts and the yield estimates. In order to overcome this challenge and to separate the overlapping fruit, we convolve the image with a variance mask. After the convolution, each pixel in the output RGB image contains the neighbor variance of the R, G and B channels, respectively [20]. The image is then transformed to gray scale by taking the mean of the three color channels. Finally, a threshold is applied on the image. These steps not only separate overlapping fruit, they also help in reducing undesired edges such as those within the foliage or the grass. Figure 5 shows an image after object separation step is applied.

Fig. 5: Object separation step. The image on left is before the object separation while that on right is the same image after object separation step is applied.

C. K-Means Segmentation and Orange Extraction

Image Segmentation is the most important part of the whole process of yield estimation. We use k-means clustering algorithm for orange segmentation. K-means clustering is an unsupervised classification technique which deals with finding a pattern in a collection of unlabelled data [21]. The k-means algorithm aims at minimizing a squared error function by iteratively reorganizing the clusters. The iterations continue until the cluster means do not move further than a certain cut-off value.

K-means algorithm is popular due to its simplicity and relatively low computational complexity. It is suitable in our scenario as the number of clusters (K) is easy to choose. An image of an orange tree generally consists of regions representing the oranges, leaves, branches and sky. Therefore, we select K to be 4 corresponding to these 4 regions. After the clustering, we apply thresholding to extract the oranges from the tree images. Each object in the image is segmented using a particular RGB value.

In many cases, the fruit is visually fragmented because of the obstructions caused by the leaves. This makes the counting error prone as one fruit may be counted as two or more fruits. In order to remove the smaller fragments of these fruits, we converted the image into binary and applied the erosion operation. The erosion of a binary image A by the structuring element B is defined by [22]:

$$A \ominus B = \{z|(B)_z \subseteq A\} \quad (3)$$

i.e., the output of erosion functions contains all pixels z in A such that B is in A when origin of $B = z$.

Figure 6 shows two images. On the left, there is input image while the output of k-means segmentation is shown on the right.

D. Blob Detection and Orange Counting

The final step of the proposed orange counting technique is to detect monolithic fruit regions, which are also known as blobs or objects. After the erosion operation, we find the connected components in the binary image using 8-connectivity. Each connected component corresponds to one orange. We count the number of connected components which gives us the number of oranges in a particular tree image.

Fig. 6: K-means segmenation step. The image on left is the input image while that on right is the output after k-means segmentation is applied.

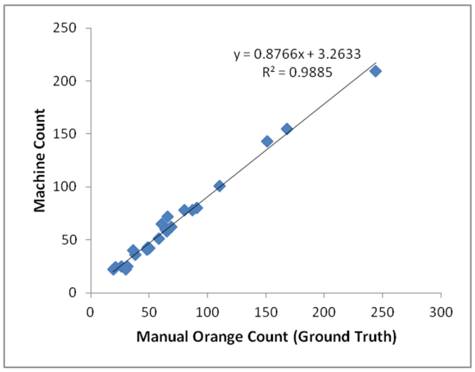


Fig. 7: Comparison of citrus fruit count per tree counted manually against that calculated using the proposed technique for Khanpur dataset.

IV. Experimental Evaluation

A. Dataset

We have collected a total of 83 tree images taken from Khapur and National Agricultural Research Council (NARC) fields. The images were taken at three times with an average gap of around one month. In the rest of this paper, we will call these three datasets Khanpur, NARC 1 and NARC 2, respectively. Khanpur dataset contains 23 tree images while NARC 1 and NARC 2 contain 16 and 44 tree images, respectively. The images contain variable illumination including shadows, bright sunlight and dusk. We prepared a ground truth by manually counting and marking all the orange fruit in all the input images.

B. Results

We performed experiments by getting automated citrus counting using the proposed scheme and comparing the results with the ground truth. The experimental results show a correct detection rate of 91.3%, as shown in Table I. In addition to finding detection rate, we used linear regression to model the relationship between the automated results and the ground truth. Figures 7, 8,and 9, show the result of plotting manual and automated orange counts for Khanpur, NARC 1 and NARC 2 datasets, respectively. The figures also show the regression equations as well as the coefficient of determination R^2. As you can see, neither the accuracy in Table I, nor the coefficients of determination in Figures 7, 8, 9 vary significantly among different datasets. As the datasets were created at different times with varying lighting conditions, the consistency in results shows the robustness of the proposed technique.

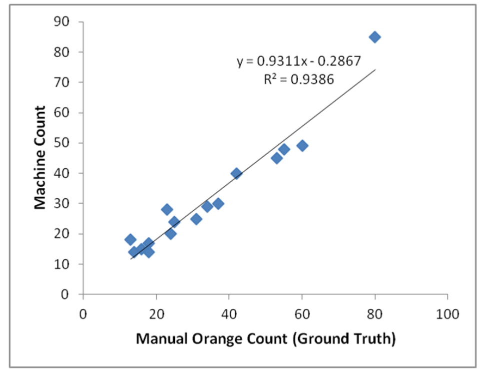


Fig. 8: Comparison of citrus fruit count per tree counted manually against that calculated using the proposed technique for NARC 1 dataset.

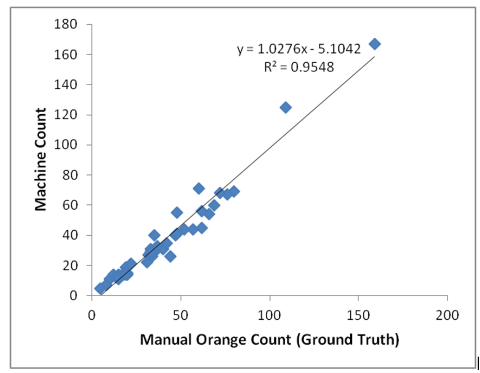


Fig. 9: Comparison of citrus fruit count per tree counted manually against that calculated using the proposed technique for NARC 2 dataset.

In Table II, we compare our results with the past work. Some of the past work has presented results in terms of detection accuracy while some has presented it in terms of coefficient of determination (R^2). Table II shows that our technique is capable of detecting and counting the fruits with more accuracy as compared to others.

V. Conclusion

In this paper, we have presented a technique for the segmentation, detection and yield measurement of citrus fruit. The proposed approach gives very good results under varying lighting conditions, occlusion of leaves and overlapping of

TABLE I: Orange detection and counting results of the proposed scheme.

Dataset	No of Images	No of Fruits	Machine Count	Accuracy (%)
Khanpur	23	1662	1532	92.2
NARC 1	16	543	501	92.3
NARC 2	44	1796	1621	90.3
Overall	**83**	**4001**	**3654**	**91.30%**

TABLE II: Comparison with the previous work.

Technique	Accuracy (%)	R^2 Value
Cakir et al. [17]	76.5	-
Hannan et al. [14]	90	-
Chinchuluunet al. [15]	-	0.89
Annamalai and Lee [18]	-	0.76
Proposed Technique	91.3	0.94 to 0.99

fruits on the images taken from varying distances from the orange trees. Our experiments on three different datasets showed an accuracy of 91.3% with R^2 values of up to 0.99. In the future, we aim to collect larger datasets for further experiments. In addition, instead of taking images manually from the citrus trees, we plan to use a camera-mounted robot for image acquisition.

REFERENCES

[1] C. Pohl and J. Van Genderen, "Review article multisensor image fusion in remote sensing: concepts, methods and applications," *International journal of remote sensing*, vol. 19, no. 5, pp. 823–854, 1998.

[2] D. Bulanon, T. Kataoka, S. Zhang, Y. Ota, and T. Hiroma, "Optimal thresholding for the automatic recognition of apple fruits," *ASAE*, no. 01-3033, 2001.

[3] T. Burks, "A real-time machine vision algorithm for robotic citrus harvesting," 2007.

[4] M. W. Hannan and T. F. Burks, "Current developments in automated citrus harvesting," in *ASABE Annual International Meeting*, 2004.

[5] R. Harrell, D. Slaughter, and P. Adsit, "A fruit-tracking system for robotic harvesting," *Machine Vision and Applications*, vol. 2, no. 2, pp. 69–80, 1989.

[6] R. Harrell, P. Adsit, T. Pool, R. Hoffman *et al.*, "The florida robotic grove-lab." *Transactions of the ASAE*, vol. 33, no. 2, pp. 391–399, 1990.

[7] A. Jimenez, R. Ceres, J. Pons *et al.*, "A survey of computer vision methods for locating fruit on trees," *Transactions of the ASAE-American Society of Agricultural Engineers*, vol. 43, no. 6, pp. 1911–1920, 2000.

[8] F. Pla, F. Juste, and F. Ferri, "Feature extraction of spherical objects in image analysis: an application to robotic citrus harvesting," *Computers and Electronics in Agriculture*, vol. 8, no. 1, pp. 57–72, 1993.

[9] E. Tutle, "Image controlled robotics in agricultural environments," 1984.

[10] D. Stajnko, M. Lakota, and M. Hočevar, "Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging," *Computers and Electronics in Agriculture*, vol. 42, no. 1, pp. 31–42, 2004.

[11] D. Stajnko, J. Rakun, M. Blanke *et al.*, "Modelling apple fruit yield using image analysis for fruit colour, shape and texture." *European journal of horticultural science*, vol. 74, no. 6, pp. 260–267, 2009.

[12] R. Zhou, L. Damerow, Y. Sun, and M. M. Blanke, "Using colour features of cv. 'gala' apple fruits in an orchard in image processing to predict yield," *Precision Agriculture*, vol. 13, no. 5, pp. 568–580, 2012.

[13] M. Regunathan and W. S. Lee, "Citrus fruit identification and size determination using machine vision and ultrasonic sensors," in *ASABE Annual International Meeting*, 2005.

[14] M. Hannan, T. Burks, and D. M. Bulanon, "A machine vision algorithm combining adaptive segmentation and shape analysis for orange fruit detection," *Agricultural Engineering International: CIGR Journal*, 2010.

[15] R. Chinchuluun, W. Lee, R. Ehsani *et al.*, "Machine vision system for determining citrus count and size on a canopy shake and catch harvester," *Applied Engineering in Agriculture*, vol. 25, no. 4, pp. 451–458, 2009.

[16] J. Billingsley, "More machine vision applications in the NCEA," in *Mechatronics and Machine Vision in Practice*. Springer, 2008, pp. 333–343.

[17] Y. Cakir, M. Kirci, E. O. Gunes, and B. B. Ustundag, "Detection of oranges in outdoor conditions," in *Agro-Geoinformatics (Agro-Geoinformatics), 2013 Second International Conference on*. IEEE, 2013, pp. 500–503.

[18] P. Annamalai and W. Lee, "Citrus yield mapping system using machine vision," in *ASAE Annual International Meeting*, 2003.

[19] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 7, pp. 629–639, 1990.

[20] A. B. Payne, K. B. Walsh, P. Subedi, and D. Jarvis, "Estimation of mango crop yield using image analysis–segmentation method," *Computers and Electronics in Agriculture*, vol. 91, pp. 57–64, 2013.

[21] D. Fradkin and I. Muchnik, "A study of k-means clustering for improving classification accuracy of multi-class SVM," Technical Report. Rutgers University, New Brunswick, New Jersey 08854, Tech. Rep., 2004.

[22] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Prentice Hall, 2007.

Diversity-Based Boosting Algorithm

Jafar A. Alzubi
School of Engineering
Al-Balqa Applied University
Al-Salt, Jordan

***Abstract*—Boosting is a well known and efficient technique for constructing a classifier ensemble. An ensemble is built incrementally by altering the distribution of training data set and forcing learners to focus on misclassification errors. In this paper, an improvement to Boosting algorithm called DivBoosting algorithm is proposed and studied. Experiments on several data sets are conducted on both Boosting and DivBoosting. The experimental results show that DivBoosting is a promising method for ensemble pruning. We believe that it has many advantages over traditional boosting method because its mechanism is not solely based on selecting the most accurate base classifiers but also based on selecting the most diverse set of classifiers.**

Keywords*—*Artificial Intelligence; Classification; Boosting; Diversity; Game Theory.

I. Introduction

Boosting is a powerful mechanism for improving the performance of classifiers that has been proven to be theoretically and empirically sound. The ADAptive BOOSTing (AdaBoost) algorithm, developed by Freund and Schapire [1] in 1995, has shown remarkable performance on solving benchmark real world problems, and it is been recognized as the best "off-the-shelf" learning algorithm.

The idea of the algorithm is to use a weak learner (e.g., decision tree) as the base classifier to be boosted. However, just like any other ensemble learning design, AdaBoost builds a composite hypothesis by combining many individual hypotheses through a weighted voting mechanism. Unfortunately, in many tasks and for the sake of reaching a reasonable accuracy, the number of base classifiers must be increased. It is obvious that enlargement in the design requires a huge amount of memory to store these hypotheses [2]. In fact, this requirement makes such ensemble method impractical to be deployed in many real applications. This drawback came to the attention of researchers in the machine learning field, prompting many solutions to be proposed [3][4][5]. One of the early suggestions called for an empirical pruning technique called Kappa pruning method to perform pruning on the boosting ensemble constructed of decision trees [4][6][7]. Their objective was to accomplish the task while maintaining the accuracy rate.

This work proposes a potential improvement to the AdaBoost by applying Coalition Based Ensemble Design algorithm (CED) [8][9] to be an intermediate phase in AdaBoost. Although the problem of pruning the boosting algorithm is intractable and hard to approximate, This work suggests a *margin*-based heuristic approach for solving this problem.

II. Related Work

This section presents a review of the methods that have been introduced to prune boosting algorithms. Pruning techniques described in the literature can be classified into two main categories: first, techniques that combine sub-ensembles based on the error rate estimated on the training set. The second, techniques which use some of the diversity measures, in particular the pair-wise measures, to build the subset of classifiers [4][10][11].

However, the first category is not very effective in producing a better sub-ensemble than the whole ensemble. As in the case of boosting, the generated classifiers are typically driven to zero training error very quickly [12]. Therefore, sub-ensembles based on this approach are similar and it is not easy to distinguish between them.

In 1997 Margineantu and Dietterich [4] were the first who studied the problem of pruning boosting algorithm and in particular AdaBoost. They presented five pruning methods: Early stopping, K-L Divergence pruning, Kappa pruning, Kappa-error convex hull pruning, and Reduce-error pruning with back fitting.

Later, Tamon and Xiang [5] suggested a modification to the Kappa Pruning method proposed by Margineantu and Dietterich [4]. They introduce what is called "weight shifting" strategy as an alternative heuristic approach to Kappa pruning. They further explained that while the voting weight of pruned hypothesis in kappa pruning assigned zero, in their proposed method it transfers that voting weight to the unpruned hypothesis.

The process by which this weight is transferred is based on measuring the similarity between the pruned hypothesis and the rest of the unpruned hypotheses, where each of them will receive a fraction of the weight proportional to its distance from the pruned hypothesis. The closer an unpruned hypothesis to prune one the higher its share of the distributed weight will be. This weight allocation mechanism has been called soft assignment according to [5][13], who claimed that it yields more faithful final ensemble, especially when a high pruning rate is required.

Hernandez-Lobato et al. [11] in 2006 presented a completely different heuristic approach for pruning AdaBoost which is based on application of a genetic algorithm. They defined the base classifiers that are returned by AdaBoost as their population. The fitness function is the created ensemble accuracy, and the optimization problem is to find the best subset of a given set of classifiers.

In [11] they conclude that the results of experiments which carried over a variety of domains support their claim that the genetic algorithm outperforms or is at least as good as the heuristic methods that have been used such as Kappa pruning and Reduce-error pruning which they compared their work with.

To avoid the drawbacks of the methods used in literature, we introduced our algorithm CED (Figure 2) which is based on calculating the contribution of diversity for each one of the classifiers in the ensemble and create a coalition based on these calculations which later will construct the sub-ensemble.

III. Design of DivBoosting Algorithm

AdaBoost is one of the most powerful and successful ensemble methods. It shows an outstanding performance in many classification tasks and it outperforms bagging in many situations. The drawback of AdaBoost is that it suffers and seriously deteriorates if there is a noise in the class labels. This disadvantage occurs because of the weight adaptation nature of the algorithm that it applies it on the training data set.

Here we shall present our improved version of AdaBoost algorithm which we called Diverse Boosting ($DivBoosting$). Figure 1 shows the flow chart of $DivBoosting$ and full details of the algorithm functionality are presented as a pseudo code in algorithm 1. It is worth mentioning here, that the implementation of AdaBoost we are considering here is the $resampling$ version.

$DivBoosting$ is an iterative algorithm. It starts by initializing the set of candidate classifiers to an empty set, then starts its training process by assigning uniform weights, w^0 on D_{trn}. Then it proceed with the following loop: generates a bootstrap sample S_k using the weight w^k. Create the classifier e_k using the sample S_k for training.

The next major step is to calculate ϵ_k which represents the weighted error for e_k on D_{trn} using the set of weights w^k. In contrast to other versions of AdaBoost, the algorithm does not stop if either of the two conditions is met; first, if the error ϵ_k is equal to zero and second if the error ϵ_k is equal to or greater than 0.5. Instead the weights w^{k+1} are reset to uniform values and process repeated. In case of the error ϵ_k is greater than zero and less than 0.5 then a new weights are calculated. This loop stops when the desire number of base classifiers is generated.

The previous iterative process produces a set of candidate base classifiers that form the input for the next phase. The subroutine $Excecuting_{C}ED$ execute the CED algorithm that is explained in details in [8]. $DivBoosting$ uses weighted majority vote as a combining method.

The final output of $DivBoosting$ is an optimal ensemble composed of base classifiers that are complementary (diverse) which means their errors are uncorrelated. The objective of $DivBoosting$ is to produce an ensemble that outperforms the original ensemble in term of both accuracy and ensemble size.

IV. Experimental Data

To verify our theoretical assertion that the DivBoosting algorithm will have an improvement in performance over the conventional AdaBoost algorithm, and further illustrate how DivBoosting works, several experiments conducted on nine real data sets from the UCI repository [14]. In addition to one experiment performed on the blog spam data set - a data set we built- in order to see the effect of DivBoosting on large data set with a large number of features. Table I summarizes

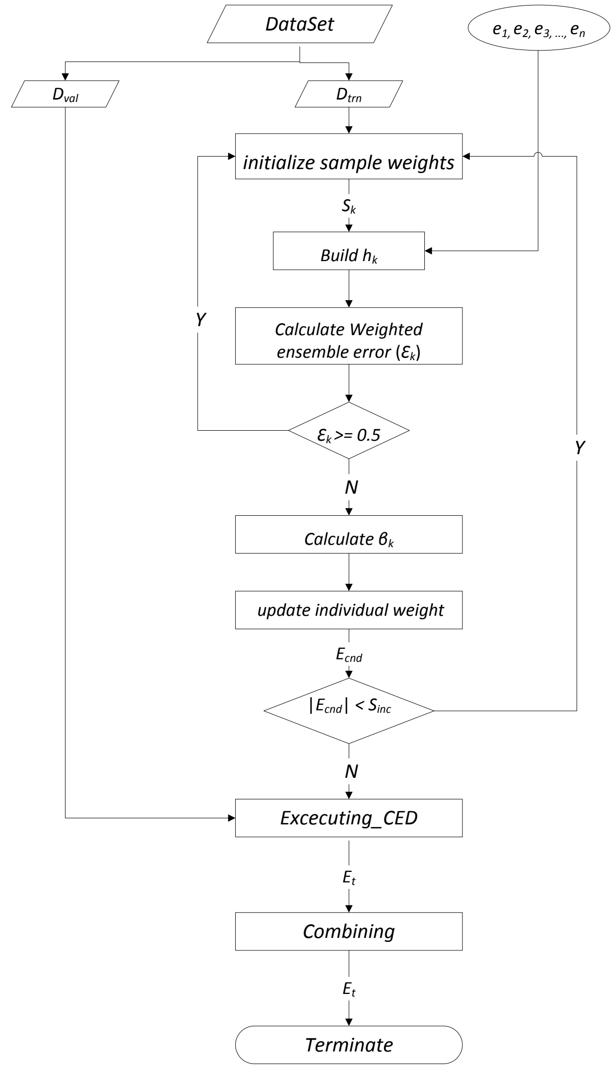


Fig. 1: Flowchart of Diverse Boosting algorithm (DivBoosting)

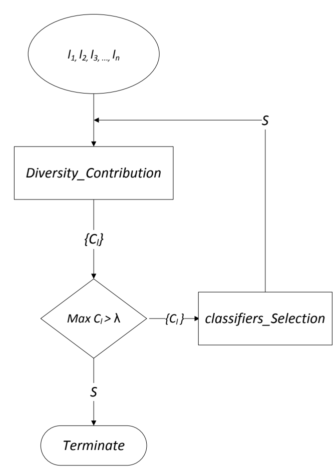


Fig. 2: Flow chart of Coalition Based Ensemble Design algorithm (CED)

Algorithm 1 pseudo code of Diverse Boosting algorithm (DivBoosting)

Input:
D_{trn} : $training\ data$
D_{val} : $validation\ data$
ζ : $a\ training\ algorithm$
E_{cnd} : $set\ of\ incremental\ cadidate\ classifiers$
S_{inc} : $size\ of\ the\ incremental\ candidate\ classifiers\ set$
η : $a\ combining\ method$
Output:
E_t : $target\ ensemble$
$Set\ weights\ w_j^1 = \frac{1}{N}$ // weights $w_1 = [w_1, ..., w_N], w_j^1 \in [0, 1], \sum_{j=1}^{N} w_j^1 = 1$
$E_{cnd} := \phi$ // Initialize the candidate classifiers set
for $k = 1, ..., S_{inc}$ **do**
  $S_k = sample\ from\ D_{trn}\ using\ distribution\ w^k$
  $e_k = Training(\zeta, S_k)$; // Create a learning base classifier
  $Calculate\ weighted\ ensemble\ error\ at\ step\ k$:

$$\epsilon_k = \sum_{j=1}^{N} w_j^k l_k^j \tag{1}$$

  $(l_j^k = 1\ if\ e_k(x_j) \neq y_j\ and\ l_j^k = 0\ if\ e_k(x_j) = y_j)$
  if $\epsilon_k = 0 \mid \epsilon_k \geq 0.5$ **then**
    $ignore\ D_k$
    $w_j^k = \frac{1}{N}$
  else
    $Calculate$

$$\beta_k = \frac{\epsilon_k}{1 - \epsilon_k} \quad , where\ \epsilon_k \in (0, 0.5) \tag{2}$$

    $Update\ individual\ weights$

$$w_j^{k+1} = \frac{w_j^k \beta_k^{(1-l_k^j)}}{\sum_{i=1}^{N} w_i^k \beta_k^{(1-l_k^j)}} \quad , j = 1, ..., N \tag{3}$$

  end if
  $E_{cnd} = E_{cnd} \bigcup e_k$
end for
$E = Excecute\_CED(E_{cnd}, D_{val})$; // Execute CED algorithm 2
$E_t = Combining(E, \eta)$;
$return\ E_t$

the used data sets in terms of number of examples, features, and classes.

The ensembles that used in the experiments were homogeneous ensembles which means the base classifiers were all the same (100 C4.5 decision trees). The performance of each decision tree was evaluated using five complete runs of five fold cross validation. In each five-fold cross-validation, each data set is randomly split into five equal size partitions and the results are averaged over five trails. In this case, one partition is set aside for testing, while the remaining data is available for training.
To test the performance on varying ensemble sizes, learning curves were generated by the system after forming sub-ensembles with different sizes (20%, 30%, 40%, 50%, 60%, 80%, and 100%). The sub-ensemble sizes of the generated ensemble represented as points on the learning curve.

TABLE I: Summary of Data Sets

Name	Examples	Classes	Attributes
Blog Spam	56000	2	547
Breast Cancer	699	2	9
Letter Recognition	20000	26	16
Iris	150	3	4
Segment	2310	7	19
ionosphere	351	2	34
Statlog (Vehicle Silhouettes)	946	4	18
Haberman's Survival	946	2	3
Contraceptive Method Choice	1473	2	3
Isolet	1559	26	617
glass	214	6	9
colic	368	2	22
heart-c	303	2	13
splice	3190	3	62
Anneal	898	6	38

For the purpose of comparing DivBoosting with AdaBoost across all domains we implemented statistics used in [15][16], specifically the win/draw/loss record and the geometric mean error ratio. The simple win/draw/loss record computed by calculating the number of data sets for which DivBoosting obtained better, equal, or worse performance than Boosting with respect to the ensemble classification accuracy. In addition to that, we computed another record representing the *statistically significant* win/draw/loss, according to this record win/loss is only computed if the difference between two values is greater than 0.05 level which was determined to be significant by computing the student paired *t*-test.

V. Experimental Results

Our results are summarized in Table II. Each cell in this table presents the accuracy of DivBoosting versus AdaBoost algorithm. We varied the sub-ensembles sizes from 20% to 100% of the generated ensemble, with more points lower on the learning curve because this is where we expect the difference to be the most between the two algorithms. A summary of the statistics is presented at the bottom of the table for each point on the learning curve.

For a better visualization of the results presented in the

TABLE II: Accuracy Rate of DivBoosting VS. AdaBoost using homogeneous ensembles

Classifiers	20%	30%	40%	50%	60%	80%	100%
Blog Spam	97.84/95.33	95.70/95.69	97.69/96.32	97.40/97.06	97.40/97.21	97.40/96.96	96.82/96.82
Breast Cancer	95.38/93.52	96.49/93.51	95.67/93.95	95.67/94.18	95.23/94.18	95.67/94.83	95.59/95.59
Ionosphere	92.20/87.14	93.36/87.89	93.36/87.90	92.71/88.97	91.21/89.56	91.21/89.66	89.64/89.64
Iris	87.69/81.73	94.38/88.06	94.61/88.11	94.01/91.97	91.50/90.85	91.20/91.00	91.66/91.66
splice	66.12/56.74	73.28/59.10	77.32/60.11	72.55/60.25	68.04/61.14	60.82/61.36	59.91/59.91
Colic	76.16/73.14	81.17/75.32	82.67/77.29	81.86/78.46	80.74/78.84	79.96/80.23	80.76/80.76
Haberman's	71.43/65.01	71.94/66.21	70.49/68.30	71.16/68.44	69.80/68.99	66.50/67.10	65.89/65.89
Auto (Statlog)	66.87/58.77	73.15/61.23	74.54/64.33	72.45/67.37	73.19/68.46	71.08/67.94	69.78/69.78
Heart-c	71.85/68.74	76.27/72.15	76.55/71.40	75.63/73.28	78.21/74.34	75.81/73.90	76.29/76.29
Letter Recognition	75.41/70.21	79.12/73.50	83.10/78.15	82.76/79.81	80.51/78.60	81.31/77.09	78.21/78.21
Win/Draw/Loss	15/0/0	15/0/0	15/0/0	15/0/0	15/0/	15/0/0	0/15/0
Sig. W/D/L	15/0/0	15/0/0	15/0/0	15/0/0	15/0/0	15/0/0	0/15/0
GM error ratio	0.7345	0.6963	0.6777	0.8043	0.8824	0.9180	1.0000

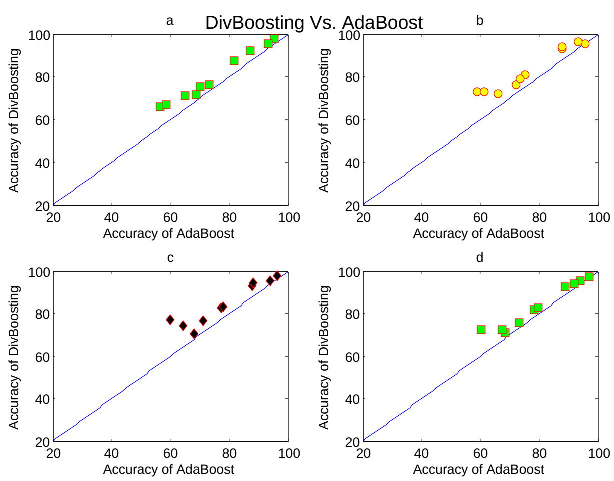


Fig. 3: A comparison showing the DivBoosting versus AdaBoost algorithm on all data sets given various homogeneous ensembles: a) 20 classifiers b) 30 classifiers c) 40 classifiers d) 50 classifiers

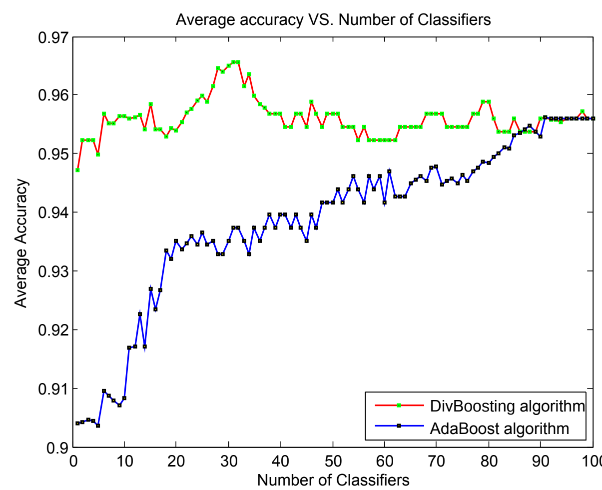


Fig. 4: Learning curve showing the average accuracy versus the number of classifiers produced by both DivBoosting and AdaBoost algorithms on breast cancer data set using 30% of data for training and a homogeneous ensemble.

table, we present plots in figures 4 to 6. Each plot present a comparison of DivBoosting versus AdaBoost for one data set over all sub-ensemble sizes from 1% to 100% of the generated ensemble.

The results in Table II confirm our assumption that combining the predictions of DivBoosting ensembles will, on average, have accuracy improvement over the AdaBoost. According to this table, we have the following general observations: 1) DivBoosting algorithm can generally improve the classification performance across all domains. 2) the best gain in performance is achieved when the ensemble accuracy of the data set is low.

For the results in Table II which represents the homogeneous ensembles, DivBoosting has more significant wins to

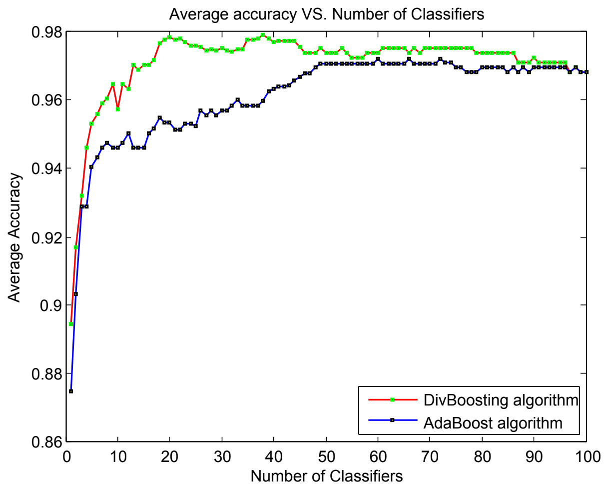


Fig. 5: Learning curve showing the average accuracy versus the number of classifiers produced by both DivBoosting and AdaBoost algorithms on Blog spam data set using 30% of data for training and a homogeneous ensemble.

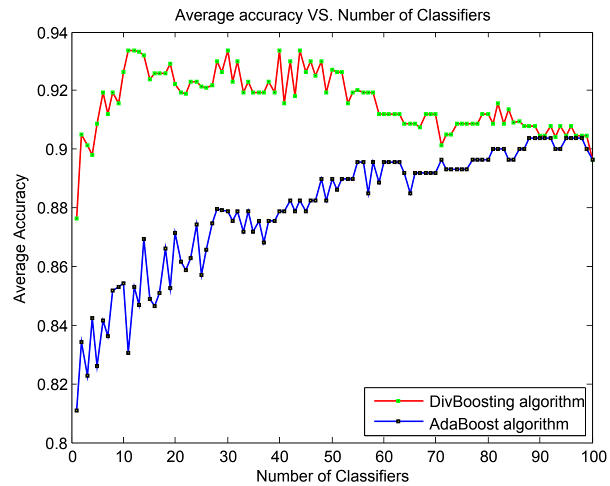


Fig. 6: Learning curve showing the average accuracy versus the number of classifiers produced by both DivBoosting and AdaBoost algorithms on Ionosphere data set using 40% of data for training and a homogeneous ensemble.

losses over AdaBoost for all data points along the learning curve. DivBoosting also outperforms AdaBoost on the geometric error ratio. This suggests that even in cases where gain is not achieved no loss occurred at any point.

We produce a scatter plots in figures 3 for various sub-ensembles sizes from Tables II. Figure 3 shows the homogeneous sub-ensembles case where DivBagging outperforms AdaBoost algorithm, in particular in sub graph c case.

DivBoosting outperforms AdaBoost early on the learning curves both on significant wins/draw/loss and geometric mean ratio. However, the trend becomes less obvious when the ensemble size increases and getting closer to the maximum size (consisting of all base classifiers). Note that even with large ensemble size, DivBoosting performance is quite competitive with AdaBoost, given ensemble sizes of 80% to 95% base classifiers, DivBoosting produces higher accuracies on all data sets with all training data set sizes.

On all data sets, DivBoosting achieves a higher accuracy rate than AdaBoost with less ensemble size. Figures 4 to 6 show learning curves that clearly demonstrate this point. To determine the influence of DivBoosting algorithm on the ensemble size, we chose to present a comparison of accuracy versus ensemble size for DivBoosting and AdaBoost on three data sets (see figures 4 to 6). The performance on other data sets is similar. We note, in general, that the accuracy of AdaBoost increases with ensemble size while the accuracy of DivBoosting increases when the diversity of the ensemble increases. So on most data sets, the performance reach its highest level when the ensemble size is between 20% and 30% of the generated ensemble size.

Figure 4 shows the performance of both algorithms on breast cancer data set for homogeneous ensembles. DivBoosting achieves an accuracy rate of 96.55% with ensemble size of 31 where AdaBoost's highest accuracy of 94.62% occurred at ensemble size of 91. These results yield a reduction of 65.93% in the ensemble size and a gain of 3% in the accuracy at the same size level.
The curve of Ionosphere data set in figure 6 illustrates that DivBoosting reaches an accuracy rate of 93.37% with ensemble size of 12 comparing to AdaBoost which it achieved an accuracy of 90.36% at ensemble size of 88. So the reduction in size here is 86.4% and at the same time the accuracy increased by 9.59%. A similar pattern observed on the Contraceptive Method Choice data set where a reduction of 64.56% in the ensemble size and 11.59% increases in the accuracy obtained. The learning curve of Blog Spam data set in figure 5 demonstrates the performance of DivBoosting and AdaBoost on a large data set in terms of both number of examples and number of features. Apart from a 2.63% improvement in the accuracy which is not small when taken in relation to the high performance of AdaBoost on this data set, the trend of ensemble reduction is the same as with other data sets which is 66.66%.

VI. Conclusions

DivBoosting is a very powerfull and effective algorithm to increase the classification accuracy and reduce the ensemble size. Throughout this paper, we introduced the algorithm and evaluated its performance through extensive experiments in comparison with conventional AdaBoost algorithm. We conducted a set of experiments using homogenous ensembles where the base learners are decision trees. DivBoosting shows the ability to increase the classification accuracy and achieves a lower ensemble size than AdaBoost.
The experimental results show that DivBoosting achieves significant improvements over AdaBoost in all domains, and yet reduces the ensemble size with more than 40% compared to the one produced by AdaBoost. Generally speaking, DivBoosting is a promising ensemble algorithm that inherits the efficiency of AdaBoost and the size reduction of CED algorithm.

References

[1] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, EuroCOLT '95

Proceedings of the Second European Conference on Computational Learning Theory, London, UK, (1995) 23-37.

[2] O. A. Alzubi, J. A. Alzubi, S. Tedmori, H. Rashaideh, and O. Almomani, *Consensus-Based Combining Method for Classifier Ensembles*, International Arab Journal of Information Technology, 15(2), (2018).

[3] G. Martinez-Munoz and A. Suarez, *Using Boosting to Prune Bagging Ensembles*, Pattern Recognition Letters, 28, (2007) 156-165.

[4] D. Margineantu and G. T. Dietterich, *Pruning Adaptive Boosting*, ICML '97 Proceedings of the Fourteenth International Conference on Machine Learning, San Francisco, CA, USA, (1997) 211-218.

[5] C. Tamon and J. Xiang, *On the Boosting Pruning Problem*, ECML '00 Proceedings of the 11th European Conference on Machine Learning, London, UK, (2000) 404-412.

[6] T. Dietterich and D. Fisher, *An experimental comparison of three methods for constructing ensembles of decision trees, in Bagging, boosting, and randomization*, Machine Learning, (2000) 139-157.

[7] S. Shylaja, K. Balasubramanya, S. Natarajan, R. Muthuraj, and S. Ajay, *Feed Forward Neural Network Based Eye Localization and Recognition Using Hough Transform*, International Journal of Advanced Computer Science and Applications, 2(3), (2011).

[8] J. A. Alzubi, *Optimal Classifier Ensemble Design Based on Cooperative Game Theory*, Research Journal of Applied Sciences, Engineering and Technology, 11(12), (2015) 1336-1343.

[9] J. A. Alzubi, *Diversity Based Improved Bagging Algorithm*, ICEMIS '15 Proceedings of the The International Conference on Engineering and MIS, Istanbul, Turkey, (2015).

[10] Y. Zhang, S. Burer, W. Street, K. Bennett, and E. Parrado-hern, *Ensemble Pruning Via Semi-definite Programming*, Journal of Machine Learning Research, 7, (2006) 1315-1338.

[11] D. Hernández-Lobato, J. Hernández-Lobato, R. Ruiz-Torrubiano, and A. Valle*Pruning adaptive boosting ensembles by means of a genetic algorithm*, IDEAL'06 Proceedings of the 7th international conference on Intelligent Data Engineering and Automated Learning, Burgos, Spain, (2006) 322-329.

[12] R. Schapire, and Y. Freund, *Boosting the margin: a new explanation for the effectiveness of voting methods*, The Annals of Statistics, 26, (1998) 322-330.

[13] D. Opitz and R. Maclin, *Popular ensemble methods: An empirical study*, Journal of Artifficial Intelligence Research, 11, (1999) 169-198.

[14] D. Newman, S. Hettich, C. Blake, and C. Merz, *UCI Repository of machine learning databases*, University of California, Irvine, Dept. of Information and Computer Sciences, (1998), http://www.ics.uci.edu/ mlearn/MLRepository.html.

[15] G. I. Webb, *MultiBoosting: A Technique for Combining Boosting and Wagging*, Machine Learning, 40(2), (2000) 159-196.

[16] E. Bauer and R. Kohavi, *An empirical comparison of voting classification algorithms: Bagging, boosting, and variants*, Machine Learning, (1998) 105-139.

Efficient Verification-Driven Slicing of UML/OCL Class Diagrams

Asadullah Shaikh
College of Computer Science and Information Systems,
Najran University,
Najran, Saudi Arabia
and
The Maersk-McKinney Moller Institute,
University of Southern Denmark,
Odense, Denmark

Uffe Kock Wiil
The Maersk-McKinney Moller Institute,
University of Southern Denmark,
Odense, Denmark

***Abstract*—Model defects are a significant concern in the Model-Driven Development (MDD) paradigm, as model transformations and code generation may propagate errors present in the model to other notations where they are harder to detect and trace. Formal verification techniques can check the correctness of a model, but their high computational complexity can limit their scalability.**

Current approaches to this problem have an exponential worst-case run time. In this paper, we propose a slicing technique which breaks a model into several independent submodels from which irrelevant information can be abstracted to improve the scalability of the verification process. We consider a specific static model (UML class diagrams annotated with unrestricted OCL constraints) and a specific property to verify (satisfiability, i.e., whether it is possible to create objects without violating any constraints). The definition of the slicing procedure ensures that the property under verification is preserved after partitioning. Furthermore, the paper provides an evaluation of experimental results from a real-world case study.

***Keywords*—MDD; UML; OCL; Model Slicing; Efficient Verification**

I. Introduction

Model-Driven Development (MDD) is a methodology widely used in the process of software development. The focus of MDD is on the use of models which can be transformed into code to save software developers time and effort. Transformation and code generation from models may spread errors in the code if the models are not verified, however.

There are formal verification tools for automatically checking correctness properties of models, but the lack of scalability of such tools is a serious problem. Addressing this problem is the goal of this paper.

At present, we face efficiency problems when verifying Object Constraint Language (OCL) constraints of complex Unified Modeling Language (UML) class diagrams. As the complexity of a model can be exponential in terms of model size (i.e., the number of classes, associations, and inheritance hierarchies), reducing the size of a model can cause a drastic speed-up in the verification process. We focus on analysis of static elements of a software specification, modelled as a UML class diagram. Complex integrity constraints will be expressed in OCL. In this context, the fundamental correctness property of a model is *satisfiability* [9], [3], [37] and whether it is possible to instantiate the model without violating any integrity constraints. Constraints can be either textual OCL invariants or graphical restrictions like multiplicities of association ends.

This property is important because it can identify inconsistent models, but also it can be used to check other interesting properties like the *redundancy* of an integrity constraint. For example, a pair of constraints C_1 and C_2 are not redundant if the following is satisfiable: $(C_1 \wedge \neg C_2) \vee (\neg C_1 \wedge C_2)$, i.e., it is possible to satisfy C_1 but not C_2 and vice versa. For instance, the *redundancy* of an integrity constraint C can be expressed as a satisfiability test: if we change the integrity constraint to $\neg C$ and the model is still satisfiable, this means that C is not redundant as it effectively avoids at least one undesired instance.

Furthermore, the addition of unrestricted[1]OCL constraints makes the problem undecidable. For example, reasoning on UML class diagrams is EXPTIME-complete [5] and, when general OCL constraints are allowed, it becomes undecidable.

In order to provide practical and workable solutions, tools for formal verification of UML/OCL class diagram must consider several aspects of the verification problem pragmatically: the desired degree of automation (fully automatic or user-guided?), the desired degree of completeness (conclusive answer for any input model?), and the degree of expressiveness allowed in OCL constraints.

Current solutions for checking satisfiability employ formalisms such as description logics [3], higher-order logics [8], database deduction systems [4], linear programming [35], SAT [21][2], or constraint satisfaction problems (CSP) [5], [9]. Each method provides different trade-offs in terms of decidability, completeness, expressiveness, and efficiency, which depend on the underlying formalism and tool support. All the approaches which support general OCL constraints share a common drawback, however: high worst-case computational complexity. Their execution time may depend exponentially on the size of the model, understanding size as the number of classes/attributes/associations in the model and/or the number

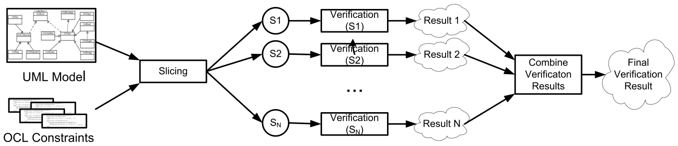


Fig. 1. High-level description of the slicing process.

of OCL constraints. This complexity is a serious limitation for the scalability of these techniques and their application in large-scale class diagrams.

A review of sample UML/OCL models highlights two observations which are relevant to this problem. First, models typically contain elements which are either unconstrained or constrained in a trivially satisfiable way. For instance, attributes acting as identifiers should have unique values, but often there are no other constraints on these attributes. Similarly, some integrity constraints regarding the multiplicity of association ends may be abstracted as well, e.g. an association end with a multiplicity of * does not constrain the model in any way which affects its satisfiability. A second observation is that some constraints refer to independent entities. For example, constraints about the password of a user and the price of a product are likely to be unrelated.

These observations can be used to improve the scalability of verification methods for satisfiability. Our proposal is based on *model slicing*: given an input UML/OCL model, the diagram and its constraints will be automatically partitioned into submodels while unnecessary model elements are abstracted. The structure of the class diagram (associations and class hierarchies) and the OCL invariants (abstract syntax tree) guide the partitioning process. Intuitively, the underlying idea is that all constraints restricting the same model element should be verified together and therefore belong to the same slice. Then, satisfiability of each slice is checked independently and the results are combined to assess the satisfiability of the entire model. Figure 1 illustrates the overall flow. To ensure soundness, slicing should not alter the outcome of the verification.

In contrast, there are a few verification and validation tools and techniques that verify the model properties and finds valid objects of the class diagrams [1], [20], [8], [17]. These tools and techniques are, however, inefficient (high Central Processing Unit and memory consumption) and unable to verify large UML/OCL class diagrams. The efficiency analysis of a few UML/OCL tools can be found in [41]. Therefore, the general question addressed in this paper is how we can improve the efficiency of the verification process for complex UML/OCL class diagrams.

A few hypotheses can be derived from the above discussion:

1) H1. Model slicing can be implemented in existing verification tools independent of their formalism.
2) H2. Model slicing will reduce the verification time.
3) H3. Model slicing enables verification of certain types of UML/OCL class diagrams that cannot be verified with current tools.

II. Contributions

This paper proposes a slicing technique for complex UML/OCL class diagrams which have a high worst-case computational complexity. This slicing technique is called the UML/OCL Slicing Technique (UOST). High worst-case computational complexity represents the amount of time in which the information in the model will be processed. The slicing procedure is based on breaking a complex UML/OCL model into several independent submodels where all irrelevant components of the model are abstracted from the complex hierarchy. The defined slicing procedure ensures that if all submodels are satisfiable then the entire model is satisfiable. If the model is unsatisfiable then some submodel is also unsatisfiable. The contributions of this paper are as follows:

1) A slicing procedure for a disjoint set of submodels
2) A procedure for analysis of OCL constraints
3) A procedure for detection of trivially satisfiable constraints
4) A procedure for analysis of UML class diagram
5) Application of the slicing technique in a real-world case study Digital Bibliography and Library Project (DBLP) conceptual schema
6) Experimental results in UMLtoCSP (UOST) [39] and in Alloy [20]

The major contribution of this paper, however, is improvement of the scalability of verification of UML class diagrams with OCL constraints. The presented slicing technique slices the model before it passes to any satisfiability analyser/engine, rather than passing a complete or complex model to any verification engine. We slice the model in the memory before checking satisfiability and this is the major reason for speed-ups in verification. In this paper, the experiments are conducted in Alloy [20] and UMLtoCSP [9]. The underlying satisfiability analyser used in Alloy is the Kodkod model finder and a variety of SAT solvers [47] whereas UMLtoCSP uses CSP [10]. With the help of the slicing technique these satisfiability analysers receive the original model as several independent submodels and check the satisfiability independently, which causes drastic speed-up in verification time. Verification time is totally dependent, however, on how fast these satisfiability analysers can find the valid objects as per conditions given in OCL constraints.

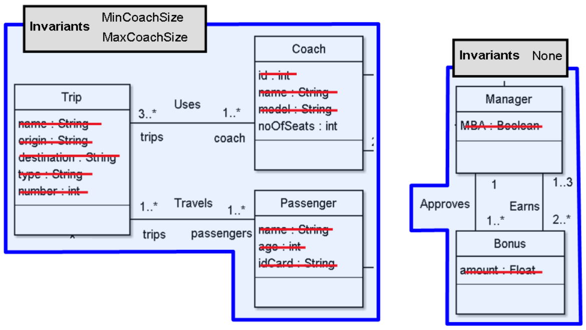
Previously, we have proposed a slicing technique for a non-disjoint set of submodels [41], [40], however, this paper extends our work published in ASE 2010 [38] with the contributions given in sections "Contributions", VI, and VII. We have implemented slicing techniques for disjoint and non-disjoint sets of submodels in UMLtoCSP (UOST) [39]. Furthermore, comparison with other tools and techniques has been discussed in [41].

III. Overview of Proposed Slicing Technique

The input for the slicing procedure is a UML class diagram annotated with OCL invariants. Figure 2 presents a

[1]Some approaches restrict the set of supported OCL constructs, e.g., to make the verification decidable. In this paper, we consider general OCL constraints with no limitations on their expressivity.

[2]"SAT stand for 'satisfiablity': a solution to a boolean formula is an assignment of values to the formula's boolean variables that 'satisfies' the forumula"[21].



(a) Slice 1 (b) Slice 2

Fig. 3. Slices for the verification of strong satisfiability in the running example.

class diagram that will be used as an example modelling the information system of a bus company. Several integrity constraints are defined as OCL invariants.

Our goal is to determine whether the input class diagram has legal instances, that is, instances that satisfy all integrity constraints. An instance of a UML class diagram is a collection of objects (according to the class definitions) and a collection of links between them (according to the associations). The output of the verification process will be either : 'satisfiable' or 'unsatisfiable'. In the case of satisfiability, a sample instance proving the satisfiability will be computed as well.

Two different notions of satisfiability will be considered for verification: *strong* and *weak* satisfiability [9], [3], [37]. A class diagram is weakly satisfiable if it is possible to create a legal instance which is non-empty, i.e., it contains at least one object from *some* class. On the other hand, strong satisfiability is a more restrictive condition requiring that the legal instance has at least one object from *each* class and a link from *each* association. Some parts of the slicing algorithm will work differently depending on the satisfiability notion to be verified.

The algorithm works by partitioning the UML/OCL class diagram into a set of disjoint *slices*. A slice S of a UML/OCL class diagram D is a subset of the original model: another valid UML/OCL class diagram where any element (class, association, inheritance, aggregation, invariant, ...) appearing in S also appears in D, but the reverse does not necessarily hold. Figure 3 represents the slices for strong satisfiability for the example system. Each slice is verified independently and the verification result of the whole model is obtained by combining the results of all slices. If we are checking strong satisfiability, it is necessary to check also whether *all* slices are strongly satisfiable. On the other hand, if we are checking weak satisfiability, it is sufficient to ensure that *at least one* slice is weakly satisfiable.

A. Preserving Satisfiability

The fundamental requirement of the slicing algorithm is that it should preserve the outcome of the verification: the answer provided by the verification with slicing should be the same as the one given by a verification tool without slicing.

Each slice is a disjoint subset of the integrity constraints and a disjoint fragment of the original class diagram. Disjoint fragment contains different classes for each slice. As each slice is less constrained than the original model, it is clear that if the original model was satisfiable, the slices will also be satisfiable. Therefore, it is only necessary to ensure that if the original model was unsatisfiable, the answer will also be 'unsatisfiable': if we are checking strong satisfiability, at least one slice will be strongly unsatisfiable, and if we are checking weak satisfiability, all slice[s] will be weakly unsatisfiable.

A class diagram can be unsatisfiable for several reasons. First, it is possible that the model provides inconsistent conditions for the number of objects of a given type. Inheritance hierarchies, multiplicities of association/aggregation ends, and textual integrity constraints (e.g., Type::allInstances() $->$size() = 7) can restrict the possible number of objects of a class. Second, it is possible that there are no valid values for one or more attributes of an object in the diagram. Within a model, textual constraints provide the only source of restrictions on the values of an attribute, e.g., self.x = 7. Finally, it is possible that unsatisfiability arises from a combination of both factors, e.g., the values of some attributes require a certain number of objects to be created which contradict other restrictions.

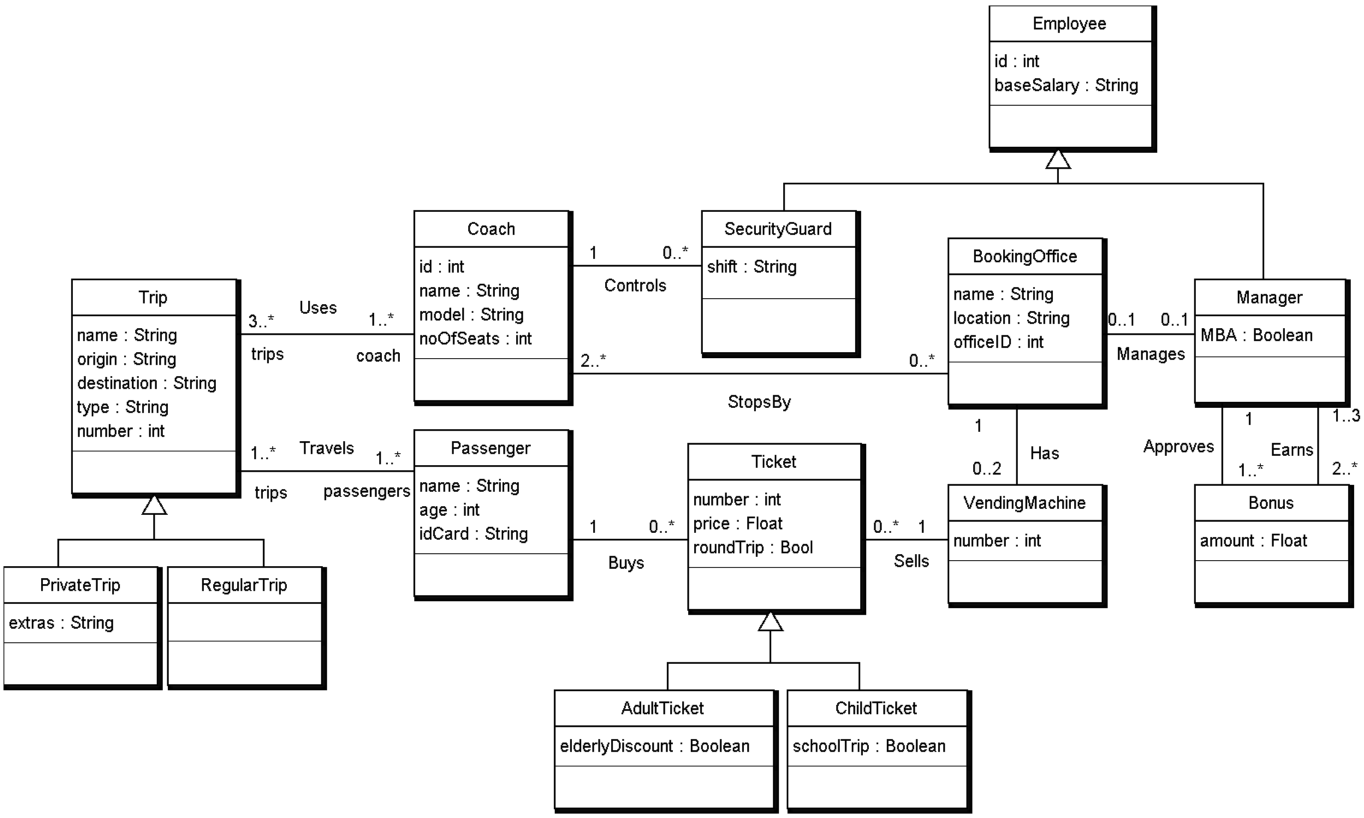
To sum up, an unsatisfiable model contains an unsatisfiable textual or graphical constraint or an unsatisfiable interaction between one or more textual or graphical constraints. To ensure that unsatisfiability is propagated in the slices, three conditions should be guaranteed:

1) No potentially unsatisfiable constraint should be removed from the problem.
2) If there are two or more constraints whose interaction could be unsatisfiable, none of them should be removed from the problem.
3) All constraints referring to the same model element should appear together in the same slice, i.e., their interaction should not be split into different slices.

The procedure presented in this paper guarantees all conditions (1 to 3). Before slicing, the class diagram and integrity constraints are analysed to detect unconstrained elements, constraints which do not affect satisfiability and constraints which cannot interact adversely with other constraints. In order to provide this assurance, it is necessary to analyse the UML class diagram before slicing in order to know what can be partitioned and abstracted and what should be kept together. The analysis is performed at two levels, a syntactic analysis of the OCL constraints and a structural analysis of the UML class diagram:

- A traversal of the syntax tree of each OCL constraint identifies which classes, attributes, and navigations are used. Additional analysis identifies trivial constraints and constraints that can be checked independently.
- The analysis of the UML class diagram reveals dependencies among the number of objects in each class, like inheritance hierarchies or multiplicity constraints of association/aggregation ends.

The following sections describe the analysis of OCL invariants and the UML class diagram. The combination of the verification results from each slice is straightforward (either all slices have to be satisfiable or at least one has to be for strong



context Coach **inv** MinCoachSize:
self.noOfSeats $\geq$ 10

context Coach **inv** MaxCoachSize:
self.trips $->$forAll(t | t.passengers $->$size() $\leq$ noOfSeats)

context Trip **inv** CorrectTripDestination:
not self.origin = self.destination

context Ticket **inv** UniqueTicketNumber:
Ticket::allInstances() $->$isUnique (t | t.number)

context Ticket **inv** MachineNumber:
self.name=self.vendingMachine.bookingOffice.location.concat(self.number.toString())

context Passenger **inv** NonNegativeAge:
self.age $\geq$ 0

Fig. 2. UML/OCL class diagram used as running example (model Coach).

and weak satisfiability, respectively) and will not be detailed further.

IV. Analysis of OCL Constraints

OCL allows the definition of *expressions* on UML class diagrams. An expression which evaluates to 'true' or 'false', e.g., a class invariant, will be called a *constraint*. OCL can also be used to define the result of *query operations*, which can then be invoked inside other expressions.

Any OCL expression is defined within the *context* of a type. Typically, an OCL expression involves several objects from one or more classes of the model. To get a starting object, we can use the keyword *self*, which denotes an object of the context type, or the method allInstances(), which can be used to access all objects of a given type, e.g., Trip::allInstances() and returns a set of all objects of class Trip. Given an object, OCL provides operators to read the values of its attributes (*attribute access*) and access the objects connected to it through associations (*navigation*). Combining these operators with arithmetic, logic, and relational operators, iterators and user-defined query operations, it is possible to write complex constraints about class diagrams.

This section describes how to analyse OCL invariants in order to extract information relevant to its satisfiability. We are interested in identifying which model elements are constrained by an invariant, as interactions between constraints appear when two or more constraints restrict the same model elements.

A. Constraint Support

The *support* of an OCL expression is the subset of classes of the class diagram referenced by the expression. For invari-

TABLE I. SUPPORT, ATTRIBUTES, AND NAVIGATIONS IN THE RUNNING EXAMPLE.

Invariant	Support	Attributes	Navigations
MinCoachSize	Coach	Coach.noOfSeats	None
MaxCoachSize	Coach, Trip, Passenger	Coach.noOfSeats	Travels, Uses
CorrectTripDestination	Trip	Trip.(origin,destination)	None
MachineNumber	VendingMachine, BookingOffice, Ticket	Ticket(name,number) BO.location	Sells, Has
UniqueTicketNumber	Ticket	Ticket.number	None
NonNegativeAge	Passenger	Passenger.age	None

ants, the support describes the set of classes restricted by the constraint. This information will be used to identify classes that appear together in the same constraint and therefore must be analysed within the same slice. Formally, the support of an expression E and the supertypes of E contains the following types:

1) The context type where E is defined and all its supertypes, as long as the 'self' variable appears within E.
2) The type of each association end navigated within E.
3) Each type referenced explicitly in E by the operation Type::allInstances() or by a type check or conversion operation, e.g., oclIsKindOf, oclIsTypeOf, or oclAsType.
4) The union of the supports of all query operations invoked from E.

Another piece of information required by the remaining steps of the analysis is the set of attributes and navigations used in each invariant. This information can be gathered with a straightforward traversal of the OCL syntax tree. Table I summarises all these data for the invariants of the running example.

The support information can be used to partition a set of OCL invariants into a set of independent *clusters* of constraints, where each cluster can be verified separately. The following procedure can be used to compute the clusters:

- Compute the support of each invariant.
- Initially, each constraint is located in a different cluster.
- Select two constraints x and y with non-disjoint supports (i.e., support(x) $\cap$ support(y) $\neq \emptyset$) and located in different clusters, and merge those clusters.
- Repeat the previous step until all pairs of constraints with non-disjoint support belong to the same cluster.

Using this procedure and the information from Table I, we can identify three clusters in our model: invariants MinCoachSize, MaxCoachSize, CorrectTripDestination and NonNegativeAge (support: Coach, Trip, Passenger); invariant MachineNumber (support: VendingMachine, BookingOffice) and invariant UniqueTicketNumber (support: Ticket). In the following sections, however, we describe additional analysis that can abstract constraints before this clustering, simplifying the problem that has to be verified.

B. Local and Global Constraints

Some parts of a verification problem can be checked in isolation within the boundaries of a class and without affecting the overall solution. Intuitively, if there is a constraint on an attribute which is not used anywhere else in the model, we can split the verification problem into two separate subproblems: checking that the constraint on the attribute is feasible and verifying the rest of the system. This section will present the techniques which identify such local constraints.

TABLE II. EXAMPLES OF LOCAL AND GLOBAL INVARIANTS.

Type	Expression (context Trip)	Description
Local	self.origin $\neq$ self.destination	Attribute access
Global	**not** self.passengers−>isEmpty()	Navigation
Global	Ticket::allInstances()−>isUnique(t|t.number)	allInstances()
Global	self.oclIsTypeOf("PrivateTrip")	oclIsTypeOf()

An expression is called *local to a class C* if it can be evaluated by examining *only* the values of the attributes in ***one*** object of class C. Expressions that do not fit into this category, because they need to examine multiple objects of the same class or some objects from another class, are called *global*.

In other words, a local expression can be defined as follows: (1) it does not use navigations through associations, (2) it does not call allInstances(), (3) it does not use attributes defined in a superclass, (4) it does not call any global query operation, and (5) it does not perform any type check or type conversion operation. Table II shows some examples of local and global expressions written in the context of class *Trip*.

Attributes may appear in local constraints, global constraints, or both. We are interested in detecting those attributes that can be studied locally, like those that do not appear in global constraints and are not related to attributes that appear there. In this sense, the set of *global* attributes will be iteratively defined as follows: (1) the attributes used in global expressions plus (2) the attributes used in local expressions where there is at least one global attribute. All other attributes of the model will be called *local*. A local expression which uses only local attributes will be called *strongly local*.

It should be noted that according to our definition the result of a strongly local invariant does not depend on (1) attributes outside those mentioned in the expression or (2) the number of objects in any class. The only chance of potential interaction with other invariants is with other strongly local invariants of the same class, if they have any attribute in common. Therefore, strongly local invariants of a class can be analysed separately from the rest of the model. The division into subproblems is as follows:

- A problem defined by the class, its local attributes, and its strongly local invariants (which can be further partitioned if these invariants restrict disjoint sets of attributes).
- Another problem defined by the original model, removing the attributes and constraints that appear in the first subproblem.

In our running example, invariants MinCoachSize, NonNegativeAge, and CorrectTripDestination are all local invariants. Of these, invariant MinCoachSize is not strongly local as the attribute 'noOfSeats' is also used in the global invariant MaxCoachSize. The remaining invariants, NonNegativeAge and CorrectTripDestination, can be abstracted from the model together with the attributes they reference and their satisfiability can be checked independently.

C. Trivially Satisfiable Constraints

A final analysis that can improve the efficiency of satisfiability verification is the detection and removal of trivially satisfiable invariants from the UML/OCL class diagram. Detecting satisfiable constraints is as hard as satisfiability itself, so we restrict ourselves to considering typical patterns which may arise in different applications.

The first trivially satisfiable pattern which can be safely removed is the *key constraint* stating that a given attribute's value must be unique, e.g. Type::allInstances() −>isUnique(obj | obj.attr). If the attribute is of Integer, Float, or String type and it is not referenced by any other constraint, it can be trivially satisfied: a different value can be assigned to each potential instance, e.g., 1, 2, 3, …The verification engine does not need to spend time computing the value of the attribute in each object and enforcing uniqueness among different objects. Therefore, the attribute and the constraint can be safely removed from the problem without affecting its satisfiability.

Another trivially satisfiable pattern which can also be removed is the *derived value constraint*, where the value of one attribute depends on the values of other attributes. The pattern is *self.attrib op expression* where *attrib* is an attribute of a basic type (Boolean, Integer, Float, String) not constrained by any other constraint, *op* is a relational operator ($=, \neq, <, >, \leq, \geq$) and *expression* is a 'safe' OCL expression which does not include any reference to *attrib*. By 'safe' we mean a side-effect-free expression which cannot evaluate the undefined value in OCL (OclUndefined). This means that we do not allow divisions that can cause a division-by-zero or collection operations which are undefined on empty collections like first().

Intuitively, this constraint cannot make the model unsatisfiable: if an instance for the rest of the model can be created, it is simply a matter of evaluating *expression* to find the right value of *attrib*. The conditions for *expression* (no self-references, no undefined values) guarantee that the evaluation always computes a feasible value for *attrib*. In some cases, derived value constraints involving recursive query operations may render/make a model to become unsatisfiable. In that case, if there is any repetition in the same constraint (i.e., using recursive query operations) with the same pattern *self.attrib op expression* will not be counted as a derived value constraint. The recurrence of values of same objects make a model unsatisfiable. Therefore, if there is any repetition in OCL constraint will not be removed from the problem. Table III briefly summarises the patterns and conditions where the column *Pattern* shows the possible expressions and the column *Condition* illustrates the criteria for including the corresponding pattern.

With regard to the running example, invariant MinCoachSize is a derived value constraint where the expression is the constant 0. This invariant is not trivially satisfiable, however, and therefore cannot be abstracted, because the attribute 'noOfSeats' is also constrained by the invariant MaxCoachSize. On the other hand, the constraints NonNegativeAge, CorrectTripDestination, and MachineNumber are derived value constraints which can be abstracted. Finally, the invariant UniqueTicketNumber is a key constraint which can also be abstracted.

TABLE III. Patterns with conditions.

Pattern	Condition
Type::allInstances() −>isUnique(at)	Key constraint if attribute is not constrained anywhere else.
self.at op exp	Derived value constraint if attribute is not used anywhere else and expression does not involve attribute.
A or B	Trivially satisfiable if either A or B are satisfiable.
A and B	Trivially satisfiable if both A and B are satisfiable and have no interdependencies.
A implies B =¬A∨B	Trivially satisfiable if either $\neg A$ or B are satisfiable.
Not A	Trivially satisfiable if A is trivially satisfiable and it is not a key constraint.
self.navigation−>isUnique(at)	Trivially satisfiable if attribute is not used anywhere else.

V. Analysis of UML Class Diagrams

In this section, we will consider a UML class diagram composed of binary associations and inheritance relations. The features of class diagrams like associative classes or n-ary associations can be expressed in terms of binary associations (and potentially additional OCL constraints) [18].

In this phase, we will compute a graph-based representation (*dependency graph*) that captures the dependencies of the elements within the UML/OCL class diagram. Then, the computation of slices will simply consist of computing the *connected components* of the graph, i.e., the maximal subgraphs where there is a path among each pair of vertices. Intuitively, each connected component represents a set of interdependent constraints which have to be analysed as a whole.

A dependency graph is an undirected graph where each vertex is a class of the model. The core challenge is the definition of the conditions under which two vertices will be connected: they should be as aggressive as possible (removing irrelevant dependencies) but also conservative (related vertices will not be separated under any circumstances).

In order to define these relationships, we will use an auxiliary graph-based representation called a *flow graph*. A flow graph is a labelled directed pseudograph, i.e., there can be arcs from a vertex to itself and multiple arcs between two vertices. The vertices of the flow graph are the classes of the class diagram and the labels in the arcs are non-negative integers. An arc $X \overset{n}{\to} Y$ has means 'if there is an object in class X, at least n objects of class Y must exist'. Using this definition, there is an arrow $X \overset{n}{\to} Y$ if:

- X is a subclass of Y ($n = 1$): each object of a subclass is also an object of the superclass.
- There is an association between X and Y and the lower bound of the multiplicity of the association end at Y is n.

Arcs with a label of zero can be removed because they are not imposing any constraint. Multiple arcs between two vertices can be replaced by a single arc labelled with maximum label. For example, Figure 4 illustrates the flow graph for the running example after these simplifications.

Intuitively, a path in the flow graph among vertices X and Y establishes a dependency from X to Y. A cycle defines

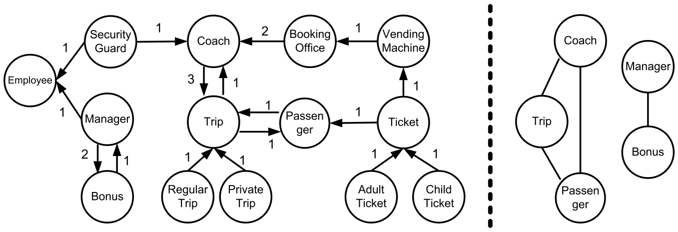


Fig. 4. Flow graph (left) and dependency graph (right) for the running example.

a cyclic dependency and it is therefore a possible source of unsatisfiability. Any cycle where the maximum label is one is inherently satisfiable, and it will be called *safe*, but cycles where (1) the maximum label ≥ 2 and (2) there are two or more participating associations/inheritance relations which also form a cycle in the class diagram *can* be unsatisfiable. Such cycles will be called *unsafe*. In our running example (Figure 4), there are three cycles: Trip-Coach, Trip-Passenger and Manager-Bonus. The first two are safe (they only involve one association so there is no cycle in the class diagram) whereas the third one is unsafe (two associations participate in the cycle and there is a multiplicity with lower bound 2).

Using this information, the dependency graph will be created in two steps. In the first step, we identify classes which are *potentially unsatisfiable*, i.e., classes constrained by OCL invariants and classes belonging to an unsafe cycle:

1) Create a vertex for each class that appears in the constraint support of an OCL constraint.
2) Add an edge $X - Y$ if both X and Y belong to the constraint support of the same constraint.
3) Create a vertex (if it does not previously exist) for each class that appears in an unsafe cycle in the flow graph.
4) Add an edge $X - Y$ among all vertices participating in the same unsafe cycle.

In the second step, we iteratively add classes that constrain vertices already in the dependency graph. Let X and Y be a pair of vertices in the dependency graph, where X and Y can be the same vertex, and Z a class that does not appear in the dependency graph. Then, if there is a path from X to Z and from Z to Y *in the flow graph*, vertex Z must be added to the dependency graph together with edges $X - Z$ and $Y - Z$. This process propagates dependencies between potentially unsatisfiable classes that cross through other classes. In our running example, the resulting flow graph is shown in Figure 4, with two connected components: one coming from the unsafe cycle in the flow graph *Manager-Bonus* and another coming from the constraints *Min/Max-CoachSize*, formed by classes *Coach*, *Trip*, and *Passenger*.

It is possible to extract its connected components from the dependency graph. Each component defines a slice of the class diagram that can be analysed independently: the set of classes from the class diagram, the set of associations and the inheritance hierarchies among them, the invariants that have some of these classes in their support and the attributes referenced by any of those invariants. For example, Figure 3 highlights the final slices passed to the verification tool for strong satisfiability. Strikethrough text indicates attributes from the original model which have been abstracted in the slice. Notice how, thanks to the detection of trivially satisfiable invariants described in the previous section, some attributes like *origin* which were originally constrained by an invariant can be simply abstracted.

With this approach, the slices of the class diagram correspond to those fragments that could be unsatisfiable. The implication is that if the slices can be populated, then the remaining classes can be populated as well. But what happens if these slices cannot be populated? This does not matter for strong satisfiability, as *all* classes must be populated so any failure means the whole model is unsatisfiable. As regards weak satisfiability, however, it could be the case that all slices are unsatisfiable but some of the remaining classes can be satisfied independently. Considering our running example, let us consider class *Employee*: creating an employee does not impose any obligation on any other class of the model. Thus, it is clear that this class can be populated and the model is weakly satisfiable. Formally, if there is any class X such that (1) X does not appear in the dependency graph and (2) the flow graph has no path from X to a class in the dependency graph, the model is weakly satisfiable. In this case X and any classes which depend on X can be populated even if no class of the dependency graph can be populated. In our running example, class *Employee* is the only class which exhibits this trait.

VI. DBLP Conceptual Schema

This section demonstrates the application of the slicing algorithm in a real-world case study: the conceptual schema of the DBLP system, modelled as a UML class diagram. It is a computer science bibliographical website, dating from the 1980s [15]. The DBLP structural schema deals with people and their publications, which can be edited books and authored publications. The class diagram has 17 classes and 26 integrity constraints. This case study is interesting for our problem since it has complex invariants and is a real-world case study. Therefore, we applied our slicing approach to this DBLP case study in order to show that our method works for external case studies and can improve the efficiency of the verification process.

Figure 5 introduces the DBLP class diagram that will be used as an example to demonstrate slicing. Several integrity constraints are defined as OCL invariants which we classify in three types of categories: *key constraints, derived value constraints,* and *indispensable integrity constraints*. Table IV describes the list of constraint names, constraint supports, and the category of constraints (key, derived value or indispensable). The key constraints and derived value constraints are considered as trivially satisfiable patterns and can safely be removed from the problem in order to improve the efficiency of the verification process without considering its satisfiability. Another category of integrity constraint is the indispensable integrity constraints which are neither key constraints nor derived value constraints and therefore cannot be abstracted. These types of OCL invariants cannot be removed because their attributes are constrained by more than one invariant which affects their satisfiability. For example, *self.journalVolume −>isUnique(volume)* could be considered

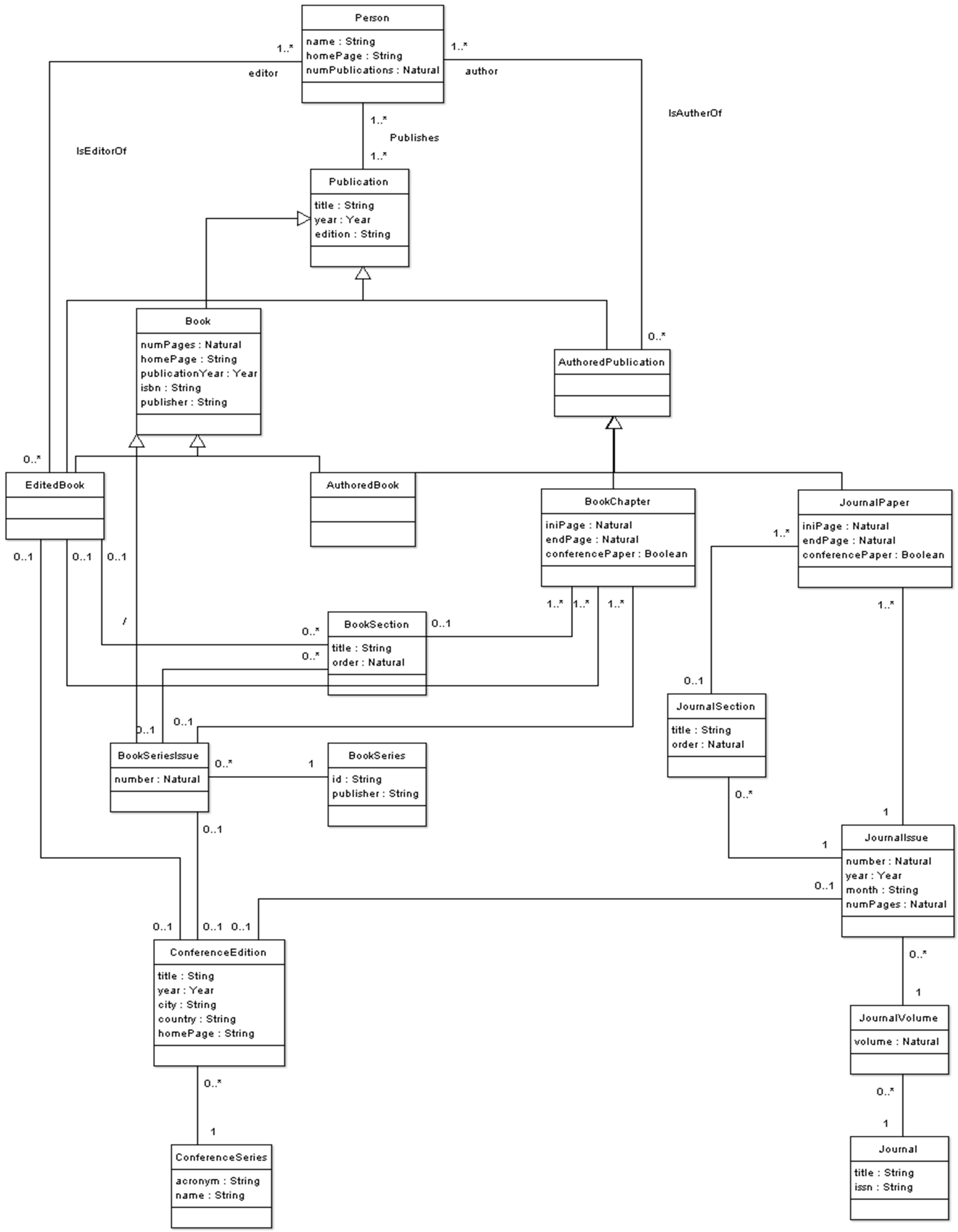


Fig. 5. DBLP class diagram.

a key constraint, but it is not a key constraint because there is another constraint which affects the same attribute such as *self.journalVolume −>sortedBy(volume).volume = sequence { 1..self.journalVolume −>size() }*.

A. Slicing DBLP

The input model is a DBLP case study annotated with 26 OCL invariants. After the elimination of key constraints and derived value constraints, we have 10 other integrity constraints whose satisfiability needs to be checked. Out of these 10 constraints, there are two *local* and eight *global* constraints. In order to identify the slices, we need, first of all, to compute a flow graph of the DBLP case study that captures the dependencies of the model elements, then the connected components of the graph will be computed respectively. In a flow graph, each vertex is a class and the arcs are non-negative integers showing the association between one vertex and another. Arcs with a label 0 are removed from the DBLP class diagram because they are not restricting any OCL invariant except ConferenceEdition-EditedBook, ConferenceEdition-BookSeriesIssue, and ConferenceEdition-JournalIssue. Figure 6 illustrates the flow graph of the DBLP class diagram after the elimination of the unnecessary arcs. The next step involves the detection of cycles among the vertices. Because of cyclic dependency, it is possible that the model may become unsatisfiable. In the case of DBLP case study, all cycles have the maximum label 1 and therefore all will be deemed safe. There are two cycles: Person-Publication and JournalPaper-JournalIssue which exists in a DBLP conceptual schema; however, no unsafe cycle exists. Applying all this information, we create the dependency graph. Initially, the classes constrained by the OCL invariants will be identified. Second, the vertices corresponding to these classes will be added. For example, there is an arc between vertex JournalPaper and JournalIssue in the graph; however, the arc between JournalVolume and JournalIssue is not included in the graph. Using the path from JournalIssue to JournalVolume, an arc between JournalVolume and Journal can now be added to the flow graph. Edges with a multiplicity 0 which are not navigated through any constraint are not depicted in the diagram.

Finally, the connected components will be extracted from the flow graph. Each of these single components is a slice and is composed of a set of classes, associations, and invariants. Figure 7 describes two resulting slices, whose satisfiability must be checked. These slices are made on the results from indispensable integrity constraints which are not trivially satisfiable. Table V summarises the constraint name, support, required classes, and submodel for the indispensable integrity constraints. As the DBLP case study involves complexity, there is only the possibility of slicing the conceptual schema between JournalIssue and JournalVolume. Consequently, five classes are eliminated from the hierarchy, i.e., Person, BookSection, BookSeries, JournalSection, and ConferenceSeries.

The detection of trivially satisfiable invariants means that some attributes like title, name, or city which were constrained before by an invariant can now be simply abstracted as can be seen in Table VI. The description of the columns is as follows: column *name* describes the names of the classes; column *attribute* outlines the list of attributes for the entire DBLP class diagram; column *restricted attributes* describes those attributes

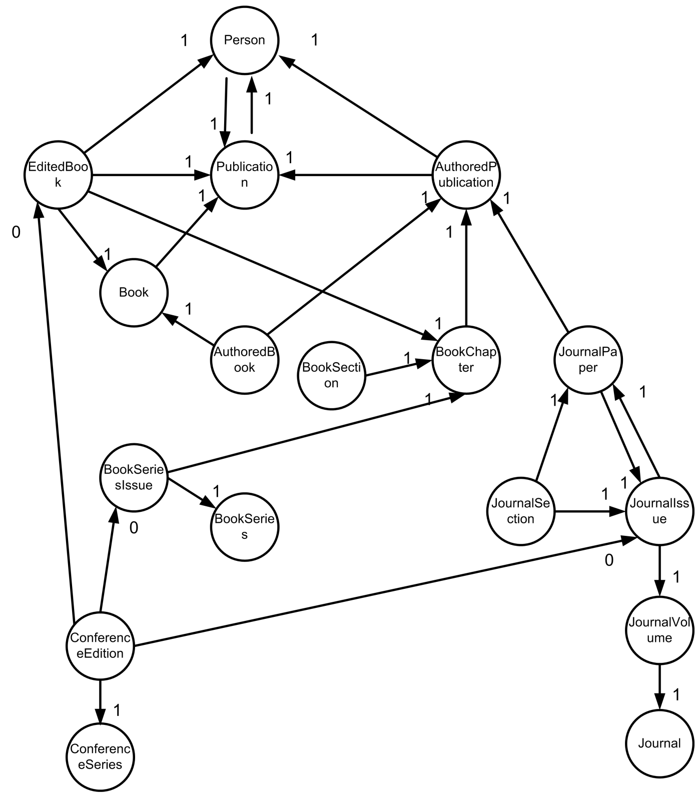


Fig. 6. DBLP flow graph.

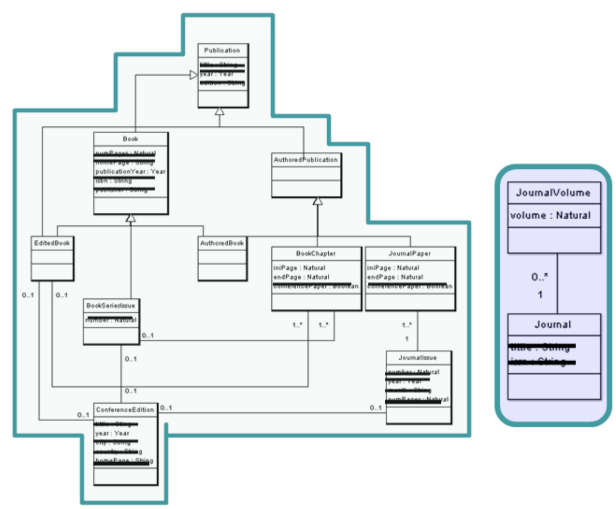


Fig. 7. Submodels 1 and 2 of DBLP.

which are constrained by indispensable integrity constraints; column *unrestricted attributes* refers to those attributes which are not constrained at all and, finally, column *unrestricted after removal* shows those constraints which were constrained by key and derived value constraints. In the end, the only attributes which need to be considered are those from column *restricted attributes*.

VII. Experimental Results

In this section, we measure the speed-up achieved by implementing the slicing technique in two different tools.

TABLE IV. NAME, SUPPORT, AND CATEGORY IN THE DBLP CLASS DIAGRAM.

Constraint Name	Constraint Support	Is Key or Is Derived Value or Indispensable Constraints
nameIsKey	Person (name)	Is Key
isbnIsKey	Book (isbn)	Is Key
idIsKey	BookSeries (id)	Is Key
BookSeriesAndNumber IdentifyBookSeriesIssue	BookSeries (number)	Is Key
issnIsKey	Journal (issn)	Is Key
titleIsKey	Journal (title)	Is Key
editedBookWithout Repetitions	EditedBook, BookSection (title)	Is Key
bookSeriesIssueWithout Repetitions	BookSeriesIssue, BookSection (title)	Is Key
journalSectionWithout Repetitions	JournalSection, JournalPaper (title)	Is Key
bookSectionWithout Repetitions	BookSection, BookChapter (title)	Is Key
journalVolumeAndNumber IdentifyJournalIssue	JournalVolume, JournalIssue (number)	Is Key
journalIssueAndTitle IdentifyJournalSection	JournalIssue, JournalSection (title)	Is Key
nameIsKey	ConferenceSeries (name)	Is Key
titleIsKey	ConferenceEdition (title)	Is Key
journalAndVolume IdentifyJournalVolume	Journal, JournalVolume (volume)	Indispensable
correctPagination	BookChapter (iniPage, endPage)	Indispensable
correctPagination	JournalPaper (iniPage, endPage)	Indispensable
correctPagination	JournalIssue, JournalPaper (iniPage, endPage)	Indispensable
correctPagination	EditedBook, BookChapter (iniPage, endPage)	Indispensable
correctPagination	BookSeriesIssue, BookChapter (iniPage, endPage)	Indispensable
consecutiveVolumes	Journal, JournalVolume (volume)	Indispensable
compatibleYear	EditedBook, ConferenceEdition (year), Book (publicationYear)	Indispensable
compatibleYear	BookSeriesIssue, ConferenceEdition (year), Book (publicationYear)	Indispensable
conferenceIsPublished	ConferenceEdition, EditedBook, BookSeriesIssue, JournalIssue	Indispensable
theSamePublisher	Book, BookSeriesIssue, BookSeries (publisher)	Is Derived Value
compatibleYear	JournalIssue (year), ConferenceEdition (year)	Is Derived Value

TABLE V. CONSTRAINT NAME, SUPPORT, TIGHTLY COUPLED CLASSES AND SUBMODEL FOR INTEGRITY CONSTRAINT.

Constraint Name	Constraint Support	Tightly Couped Classes	Submodel
correctPagination	BookChapter (iniPage, endPage)	BookChapter, AuthoredPublication Publication	1
correctPagination	JournalPaper (iniPage, endPage)	JournalPaper,JournalIssue AuthoredPublication, Publication, JournalVolume,Journal	1
correctPagination	JournalIssue, JournalPaper (iniPage, endPage)	JournalPaper,JournalIssue AuthoredPublication, Publication, JournalVolume,Journal	1 1
correctPagination	EditedBook, BookChapter (iniPage, endPage)	EditedBook,BookChapter, AuthoredPublication	1
correctPagination	BookSeriesIssue, BookChapter (iniPage, endPage)	BookSeriesIssue,BookSeries, AuthoredPublication Publication, Book,BookChapter	1
compatibleYear	BookSeriesIssue, Book (publicationYear), ConferenceEdition (year)	BookSeriesIssue,Book BookChapter, AuthoredPublication, Publication, ConferenceEdition	1
compatibleYear	EditedBook, Book (publicationYear), ConferenceEdition (year)	EditedBook,Book Publication, ConferenceEdition BookChapter, AuthoredPublication	1
conferenceIsPublished	ConferenceEdition, EditedBook, BookSeriesIssue, JournalIssue	ConferenceEdition, EditedBook, Book BookSeriesIssue, journalIssue, Publication, JournalPaper, AuthoredPublication	1
journalAndVolume IdentifyJournalVolume	Journal, JournalVolume (volume)	Journal,JournalVolume	2
consecutiveVolumes	Journal, JournalVolume (volume)	Journal,JournalVolume	2

Initially, we had developed a prototype implementation of the slicing procedure on top of the tool UMLtoCSP [9]. UMLtoCSP transforms verification problems involving UML/OCL class diagrams into *constraint satisfaction problems* (CSP) which can be solved by a constraint solver. Solutions to the CSP are instances of the diagram which prove or disprove the property to be verified. Figure 8 presents the method of applying slicing technique in UMLtoCSP. Further discussion on translation of UML/OCL class diagrams, transformation of classes and, definition of correctness of properties can be found in [10]. Second, we slice the conceptual schema of the DBLP programmed in Alloy [20] in order to show the drastic speed-up. Alloy is a widely-used structural modelling language based on first-order logic (FOL) and can generate instances of invariants and simulate the execution of operations. The purpose behind showing the results in the Alloy specification is to demonstrate that the developed slicing technique is neither tool dependent nor formalism dependent. It can be implemented in any verification-driven UML/OCL tool. These two cases support hypothesis 1 (H1).

A. Slicing in UMLtoCSP

In the first case, we compare the verification time of several UML/OCL class diagrams using (1) the original tool

TABLE VI. RESTRICTED AND UNRESTRICTED ATTRIBUTES FOR DBLP CASE STUDY.

Class Name	Attributes	Restricted Attributes	Unrestricted Attributes	Unrestricted Attributes After Removal
Person	name: String homePage: String numPublications: Natural	None	homePage: String numPublications:Natural	name: String
Publication	title: String year: Year edition: String	None	title:String year:Year edition:String	None
Book	numPages: Natural homePage: String publicationYear: Year isbn: String publisher: String	publicationYear:Year	numPages:Natural homePage:String	isbn:String publisher:String
EditedBook	None	None	None	None
AuthoredBook	None	None	None	None
AuthoredPublication	None	None	None	None
BookChapter	iniPage: Natural endPage: Natural conferencePaper:Boolean	iniPage: Natural endPage: Natural	conferencePaper:Boolean	None
BookSection	title: String order: Natural	None	order:Natural	title:String
BookSeriesIssue	number: Natural	None	None	number:Natural
BookSeries	id: String publisher: String	None	None	id:String publisher:String
ConferenceEdition	title: String year: Year city: String country: String homePage: String	year: Year	city: String country: String homePage: String	title: Sting
ConferenceSeries	acronym: String name: String	None	acronym:String	name: String
JournalPaper	iniPage: Natural endPage: Natural conferencePaper:Boolean	iniPage: Natural endPage: Natural	conferencePaper:Boolean	None
JournalSection	title: String order:Natural	None	order:Natural	title: String
JournalIssue	number: Natural year: Year month: String numPages: Natural	None	number: Natural month: String numPages: Natural	year: Year
JournalVolume	volume:Natural	volume: Natural	None	None
Journal	title: String isbn:String	None	None	title: String isbn:String

TABLE VII. DESCRIPTION OF THE UML/OCL BENCHMARKS.

Example	Classes	Associations	Attributes	Invariants	Strongly Satisfiable?
Atom-Molecule	2	1	6	1	Yes
Paper-Researcher	2	2	5	4	No
Coach	13	13	27	6	Yes
Production System	50	30	72	5	Yes
Company	100	100	100	100	Yes
DBLP Conceptual Schema	17	19	38	26	Yes
Script 1	100	53	122	2	Yes
Script 2	500	227	522	5	Yes
Script 3	1000	505	1022	5	Yes
Cycle 1	10	10	10	10	No
Cycle 2	100	100	100	100	No

UMLtoCSP and (2) the tool UMLtoCSP (UOST) with slicing [39]. UMLtoCSP (UOST) is developed in JAVA and "the basic approach behind the UMLtoCSP (UOST) tool is a model slicing technique that enables efficient verification of UML/OCL class diagrams. The tool can verify different sets of properties for UML/OCL class diagrams with disjoint and non-disjoint sets of slicing. The features strong satisfiability and weak satisfiability are same as in UMLtoCSP [9]. However, other new features in UMLtoCSP (UOST) are:

Strong satisfiability: the class diagram should have a legal instance for at least one object from each class and a link from each association.

Weak satisfiability: the class diagram should have a legal instance/object which is non-empty, i.e., it contains at least one object from some class.

Remove attributes: for weak or strong satisfiability, unrestricted attributes can be removed from the class diagram.

Non-disjoint slicing: slicing of a class diagram with non-disjoint sets of submodels.

Disjoint slicing: slicing of a class diagram with disjoint sets of submodels.

Show specific invariants: detection of failing submodel(s) in disjoint slicing and a specific unsatisfiable invariant(s) in non disjoint slicing"[39].

In each example verified in UMLtoCSP (UOST), the property to be verified has strong satisfiability. Table VII describes

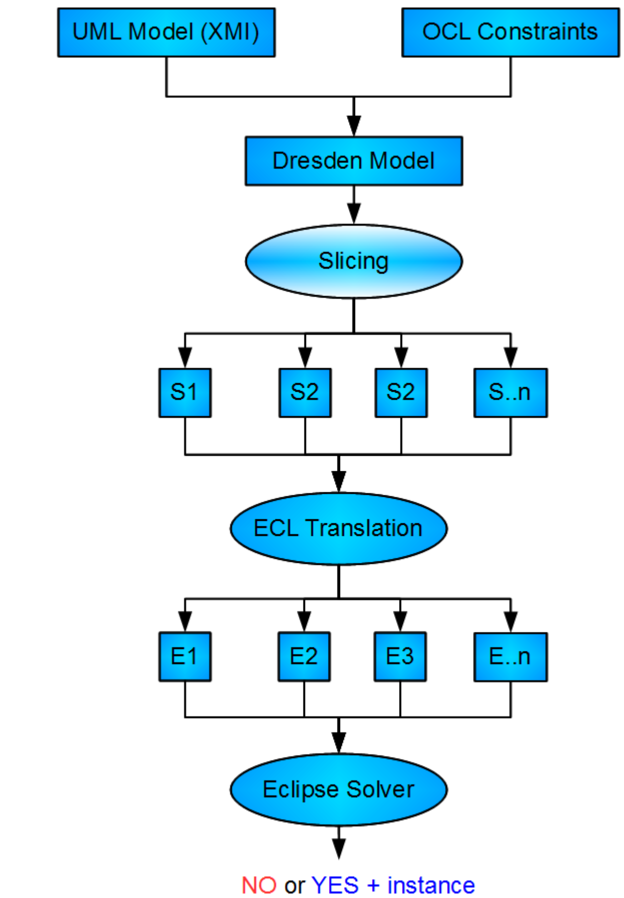
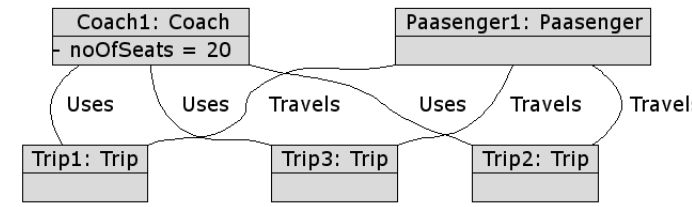


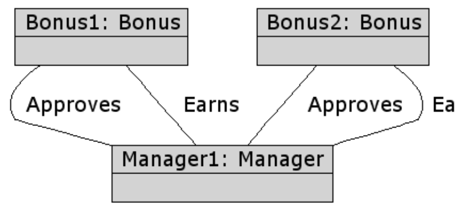
Fig. 8. Slicing procedure in UMLtoCSP.

the set of benchmarks used for our comparison: the number of classes, associations, invariants, and attributes. For each class diagram, we also indicate whether it is strongly satisfiable or not. The benchmarks ‘Company’, ‘Script’, and ‘Cycle’ were programmatically generated, in order to test large input models. Of these models, we consider the ‘Script’ models to be the best possible scenario for slicing (large models with many attributes and very few constraints). The models ‘Paper-Researcher’, ‘Atom-Molecule’, ‘Company’, and ‘Cycle’ serve as worst-case scenarios (models with many interdependent constraints, designed so they cannot be sliced).

UMLtoCSP has a set of parameters that can have a strong influence on its runtime. These parameters set an upper bound on the size of the instance (number of objects per class, number of links per association) and the domain of attributes (set of feasible values for each attribute). In UMLtoCSP, verification is not *complete* in the sense that it will only explore potential instances within these bounds. Nevertheless, the size of the solution space to be explored by UMLtoCSP is exponential in terms of these parameters. Therefore, parameters of large value will make the comparison more favourable in terms of slicing, as abstracting attributes and classes will cause a larger reduction of the solution space. In our analysis, we considered small but reasonable values for parameters: at most four objects will be created for each class, at most 10 links for each association and each attribute will have at most 10 distinct values.



(a) Submodel 1 of ‘Model Coach’



(b) Submodel 2 of ‘Model Coach’

Fig. 9. UMLtoCSP output of model coach.

Table VIII shows the experimental results computed using a Intel Core 2 Duo Processor 2.1Ghz with 2Gb of RAM. All times are measured in seconds and a time-out limit was set at two hours (7200 seconds). For each model, we describe the original verification time (OVT), the number of slices in which the model is divided, the number of attributes that we manage to abstract, the time required to perform all the UML/OCL slicing analysis (ST), and the verification time after the slicing (SVT). Figure 9 shows the UMLtoCSP output, i.e., object diagram of ‘Model Coach’.

The first conclusion is that slicing is a very fast procedure even in diagrams with hundreds of classes and it is a formalism independent technique. As expected, the effectiveness of the technique depends on the specific model analysed: small models and models where UMLtoCSP has already been performed will gain little from slicing. This also happens with models where there are no unconstrained attributes and all classes and constraints are interdependent. In the worst case, the verification time with slicing is the same as that without slicing. In models where slicing manages to partition the model and abstract attributes, however, the speed-up reaches several orders of magnitude. Therefore, its success will depend on the type of model where it is applied. Small models which have been manually preprocessed for verification will gain little from slicing. Models created for other purposes or models generated through automatic transformation can, however, benefit greatly from the application of slicing. Therefore, hypothesis 2 (H2) is supported.

The tiny overhead introduced by slicing and the tool-independent nature of this approach are additional reasons in favour of adding slicing to existing formal verification toolkits. A larger real-world case study where further benefits of slicing are illustrated in different formalism (i.e., SAT) is presented in Section “Slicing Alloy Specification (DBLP)”.

B. *Slicing in Alloy*

In the second case, we have applied the slicing technique on a few examples programmed in the Alloy specification in order

TABLE VIII. DESCRIPTION OF EXPERIMENTAL RESULTS (CASE 1).

Example	OVT	Slices	Attr	ST	SVT	Times
Atom-Molecule	0.03s	1	3	0.00s	0.03s	0x
Paper-Researcher	0.04s	1	0	0.00s	0.04s	0x
Coach	5008.76s	2	26	0.00s	0.15s	33392x
Production System	3605.35s	4	59	0.02s	0.03s	72107x
Company	0.08s	1	0	0.00s	0.08s	0x
DBLP Conceptual Schema	Time-out	2	18	0.19s	0.37s	Not verifiable without slicing
Script 1	Time-out	2	117	0.02s	0.03s	Not verifiable without slicing
Script 2	Time-out	4	509	0.09s	0.02s	Not verifiable without slicing
Script 3	Time-out	4	1009	0.29s	0.34s	Not verifiable without slicing
Cycle 1	Time-out	1	10	0.00s	Time-out	Not Available
Cycle 2	Time-out	1	100	0.00s	Time-out	Not Available

OVT Original Verification Time **Attr** # of abstracted attributes
SVT Total verification time for all slices **ST** Slicing Time

to prove that our developed slicing technique is neither tool-dependent nor formalism-dependent. To translate the models into the Alloy language, we have used the model finder to generate possible model instances, which proves satisfiability of the model.

We compare the verification time of UML/OCL class diagrams using the Alloy analyser with and without the slicing technique. Table IX describes the set of benchmarks used for our comparison: the number of classes, associations, invariants, and attributes.

Tables X, XI, and XII summarise the experimental results obtained with the Alloy analyser before and after slicing, running on an Intel Core 2 Duo Processor 2.1Ghz with 2Gb of RAM. Each table represents the results as described in the benchmark (Table IX). The execution time is largely dependent on the defined scope. Therefore, in order to analyse the efficiency of verification, the scope is limited to four. The Alloy analyser will examine all the examples with up to four objects, and try to find one that violates the property. For example, specifying scope four means that the Alloy analyser will check models whose top level signatures have up to four instances.

All times are measured in milliseconds (ms). For each scope (before slicing), the translation time (TT), solving time (ST), and the summation of the TT and ST, which is the total execution time, are described. Similarly, for each scope (after slicing) we measure the sliced translation time (STT), sliced solving time (SST), and the summation of STT and SST. Similarly, the column speed-up shows the efficiency obtained after the implementation of the slicing approach.

Previously, with no slicing, it took 200 ms (scope 4) for the execution of 'University' and 254 ms (scope 4) for 'ATM Machine'. With the UOST approach, it takes only 72 ms (scope 4) for 'University' and 24 ms (scope 4) for 'ATM Machine'. It is an improvement of 64% and 90%, respectively. In addition, the improvement can also be achieved for larger scopes as well.

C. Slicing Alloy Specification (DBLP)

We have also applied our slicing technique to the DBLP structural schema programmed in the Alloy specification. The schema that we slice with our approach defines 26 integrity constraints. The approach is manually implemented in the

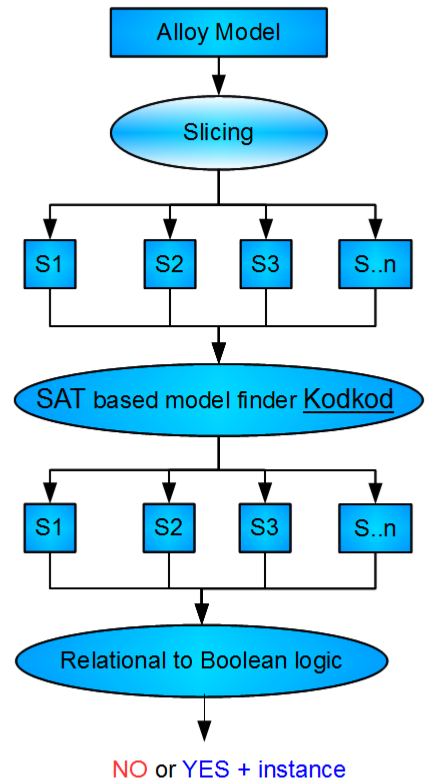


Fig. 10. Slicing procedure in Alloy analyser.

DBLP in order to show how fast it generates satisfying instances of the example before and after the slicing is applied. The same model is used for slicing in Alloy to check the advantages of slicing. The execution time is largely dependent on the defined finite scope and therefore, in order to analyse the efficiency of verification, we limit the scope to a minimum of two and a maximum of 22. Figure 10 shows the general procedure of implementation of the slicing technique in the Alloy analyser.

After application of the technique, two submodels are obtained: submodel 1 consists of 10 classes annotated with eight OCL constraints and submodel 2 comprises two classes annotated with two OCL constraints. Table XIII summarises the experimental results obtained with the Alloy analyser before and after slicing, running on an Intel Core 2 Duo

TABLE IX. DESCRIPTION OF THE EXAMPLES.

Example	Classes	Associations	Attributes	Invariants
Atom-Molecule	3	3	4	4
University	5	4	9	7
ATM Machine	50	51	51	7

TABLE X. SLICING RESULTS IN ALLOY FOR ATOM-MOLECULE EXAMPLE.

Before Slicing				After Slicing			
Scope	TT	ST	TT+ST	STT	SST	STT+SST	Speedup %
2	115ms	70ms	185ms	109ms	56ms	165ms	10%
3	138ms	76ms	214ms	117ms	65ms	182ms	15%
4	153ms	100ms	253ms	119ms	70ms	189ms	25%

TT Translation Time **ST** Solving Time
STT Sliced Translation Time **SST** Sliced Solving Time

TABLE XI. SLICING RESULTS IN ALLOY FOR UNIVERSITY EXAMPLE.

Before Slicing				After Slicing			
Scope	TT	ST	TT+ST	STT	SST	STT+SST	Speedup %
2	67ms	50ms	117ms	24ms	10ms	34ms	29%
3	92ms	56ms	148ms	35ms	30ms	65ms	56%
4	134ms	66ms	200ms	39ms	33ms	72ms	64%

TT Translation Time **ST** Solving Time
STT Sliced Translation Time **SST** Sliced Solving Time

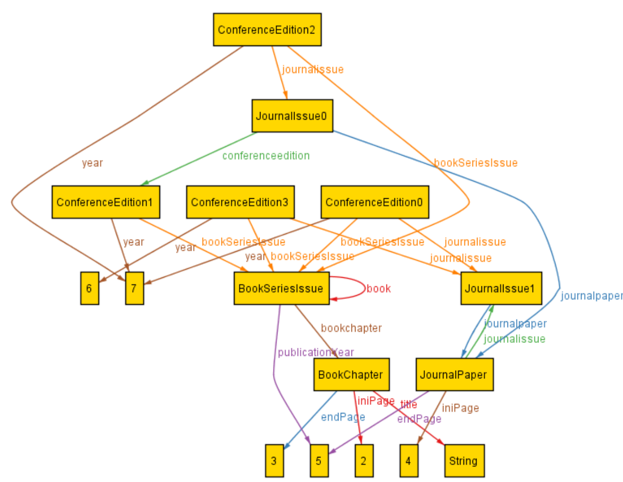
TABLE XII. SLICING RESULTS IN ALLOY FOR ATM MACHINE.

Before Slicing				After Slicing			
Scope	TT	ST	TT+ST	STT	SST	STT+SST	Speedup %
2	20ms	46ms	66ms	5ms	8ms	13ms	81%
3	83ms	91ms	174ms	9ms	11ms	20ms	89%
4	96ms	185ms	254ms	13ms	11ms	24ms	90%

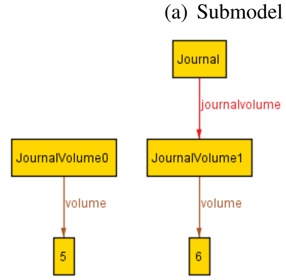
TT Translation Time **ST** Solving Time
STT Sliced Translation Time **SST** Sliced Solving Time

Processor 2.1Ghz with 2GB of RAM . All times are measured in milliseconds (ms). For each scope (before slicing), the translation time (TT), solving time (ST), and the summation of the TT and ST, which is the total execution time, are described. Similarly, each scope after slicing time is also measured: i.e., the sliced translation time (STT), sliced solving time (SST), and the summation of STT and SST, which is equivalent to the summation of TT and ST. The only difference is that the total execution time varies before and after slicing. Similarly, the column *speed-up* shows the efficiency obtained after the implementation of the slicing approach.

It took 1453 ms (scope 7) for the execution of the DBLP. Using the approach for the slice computed by this method, it takes only 828 ms (scope 7) to generate a satisfying instance for the slice. It is an improvement of 43% and we have also achieved minimum 18% (scope 2) and maximum 80% (scope 22) speed-up which is marked progress in terms of total execution time. In addition, the improvement can also be achieved for larger scopes as well. For instance, we have conducted experiments with a maximum scope of 22, and therefore, at certain scopes, it is possible to reach up to 99% improvement. Without slicing, however, we could only run the analysis only up to a limited scope. Figure 11 shows the object diagram of DBLP in the form of submodel 1 and submodel 2 as output in the Alloy analyser. This experiment supports hypothesis 3 (H3) - that model slicing enables verification of certain types of UML/OCL class diagrams that cannot be verified with current tools.



(a) Submodel 1



(b) Submodel 2

Fig. 11. Alloy output of DBLP conceptual schema.

TABLE XIII. DESCRIPTION OF EXPERIMENTAL RESULTS (CASE 2).

Before Slicing				After Slicing			
Scope	TT	ST	TT+ST	STT	SST	STT+SST	Speedup %
2	125ms	47ms.	172ms	110ms	31ms	141ms	18%
3	187ms	78ms	265ms	125ms	62ms	187ms	29%
4	281ms	172ms	453ms	219ms	78ms	297ms	34%
5	473ms	190ms	663ms	299ms	110ms	409ms	38%
6	671ms	344ms	1015ms	438ms	156ms	594ms	41%
7	969ms	484ms	1453ms	672ms	156ms	828ms	43%
8	1132ms	567ms	1699ms	602ms	240ms	842ms	50%
9	1694ms	906ms	2600ms	787ms	302ms	1089ms	58%
10	2049ms	1149ms	3198ms	854ms	208ms	1062ms	66%
11	2751ms	1297ms	4048ms	1054ms	179ms	1233ms	70%
12	3934ms	1935ms	5869ms	1283ms	351ms	1634ms	72%
13	6361ms	2838ms	9199ms	1902ms	435ms	2337ms	75%
..	....	....	....	....	....	....	...
..	....	....	....	....	....	....	...
..	....	....	....	....	....	....	...
22	26034ms	11049	37083ms	6736ms	584ms	7320ms	80%

TT Translation Time **ST** Solving Time
STT Sliced Translation Time **SST** Sliced solving Time

VIII. RELATED WORK

Slicing techniques can be classified according to two criteria: the *entity* to be sliced (e.g., a program, a UML model, an ontology, etc.) and the *goal* of the slicing process (e.g., synthesis, analysis, optimisation, visualisation, comprehension, etc.). Intuitively, all slicing techniques proceed in two steps: first, the subset of elements of interest that should appear in the slice is identified; second, elements which depend on elements of the slice are iteratively appended to the slice. The notions of 'element', 'element of interest' and 'dependency between elements' are completely determined by *what* is being sliced and *why*.

Program-slicing [56], [46] techniques work at the level of source code. Given a set of variables of interest and a program location which are provided as input, program-slicing computes the set of statements of the program that can affect (backward) or be affected (forward) by those variables. The applications of program-slicing include program analysis, optimisation, verification and comprehension. Slicing has also been used in the analysis of the *architectural specifications* of a software system [43], [26], [57]. In this context, extracting the set of components related to a component of interest can facilitate component reuse and provide a high-level view of the architecture that helps in its comprehension.

Another type of program slicing is used for *Declarative Specifications* [52], [51]. This work proposes a tool known as *Kato* which relies on heuristics to identify 'core' (slices) and it is targeted towards the relational logic underlying Alloy. Few details are provided on the set of heuristics used. The declarative modelling language Alloy plays a vital role in model verification and the complexity of declarative models is equal to that of the UML model with OCL constraints; therefore, the slicing in Alloy is needed. A novel optimisation technique based on program slicing for declarative models in order to perform efficient analysis is proposed. The algorithm works by partitioning slices for Alloy models into a case and derived slices. Afterwards, a satisfying instance of a base slice is generated to find the solution for the entire model. If the base slice is unsatisfiable then the entire model is unsatisfiable [53].

A challenging area is the testing of software product lines using SAT-based analysis. An interesting technique is proposed which incrementally generates tests for product lines. In this approach, the features of the program are the basis of incremental test cases. Afterwards, these features are converted into incremental test suites via transformation [50]. For model-based abstractions, a novel approach towards general automation for Model Driven Architecture (MDA) is introduced which is known as the FORMULA framework. FORMULA is a specification language and analysis tool that can be used to construct general MDA abstractions [22].

A further interesting angle of related work is systematic constraint-based test generation for programs and slicing for Alloy models. An interesting method for constraint-based test generation for programs is proposed which is based on organised generation of structurally complex test data from declarative constraints. The approach takes structurally complex data as an input, generates high-quality test cases and finds bugs in non-trivial programs [25]. Moreover, a novel approach which incrementally generates tests for product lines is proposed. In this work, test generation derives from the functions of the program [54]. A software system grows in size and complexity and therefore testing becomes a challenging task. In order to address the complexity issues, a novel approach to synthesising declarative specifications is presented. Partitioning is applied to enable efficient incremental analysis which defines a suite of optimisation in order to improve the analysis [49].

The most recent work on program slicing focuses on the reduction of the source code by program slicing before test generation [11]. This method is implemented in a tool called SNATE (Static Analysis and Testing). Furthermore, dynamic backward slicing for programs and the model transformation explores the idea of tracing the model transformations [48]. This work is based on program slicing for model transformations where the primary aim is to assess data and control dependencies.

Ontologies provide a formal description of a set of concepts and their relationships. General-purpose ontologies may represent a large number of concepts and their size makes them impractical for many applications. Several approaches [14], [36], [45] focus on pruning large ontologies to produce smaller ontologies which are more manageable.

Slicing methods have also been proposed in the management of different types of UML models. *Context-free slicing* [24] provides a framework for defining model slices in UML diagrams, e.g., class diagrams. This work proposes a general theory of model slicing which has to be adapted to each specific goal by defining a slicing criterion suitable for the goal. There is no discussion on the definition of suitable slicing criteria for verification. A different approach focusing on class diagram comprehension uses coupling metrics [27] to slice large models for visualisation. This type of approach would not be suitable for verification purposes, as metrics do not provide guarantees about the properties satisfied by the partitions. Finally, the slicing of models consisting of both UML class diagrams and UML sequence diagrams is considered in [29], [30]. A common representation, Model Dependency Graphs, is used to encode both types of diagrams. Again, the slicing criterion must be provided as an input to the algorithm.

Other similar work to our approach is the slicing of statecharts. *Slicing hierarchical automata for model checking* [23] presents an approach for slicing statecharts for the verification of properties. This research highlights the concept of slicing criteria for states and transitions. The algorithm is based on the removal of irrelevant hierarchies and concurrent states whereas our goal is the verification of specific properties using OCL constraints. We remove the OCL invariants to reduce the complexity of the model and then slice. The slicing criteria described in this paper are based on the scope of OCL constraints. Another piece of work related to our own is *System Verification through logic* (SV_tL) [55]. SV_tL provides a verification environment based on slicing for UML statecharts. It also removes irrelevant hierarchies in order to reduce the complexity of the verification for statecharts. The slicing criteria are based on dependency relations among states and transitions of the system.

Slicing of UML state machines is another type of related work [31]. This research provides a framework for creating smaller models for UML state machines given the fact that the behavior of the model should be the same. The method of slicing simplifies the model with the help of features. It is based on path predicates. Furthermore, *slicing of state-based models* [28], [12], [19], [2] classifies the segments of the model based on the element of interest. This approach presents a slicing technique that reduces the complexity of state-based models. The main use of slicing is for extended finite state machine (EFSM) models; however, the technique can also be applied to Specification and Description Language (SDL) models and statecharts.

Recent work on the slicing of UML models using model transformation has been presented by Kevin Lano [32], [33]. The purpose of slicing is to break the model into several submodels for better analysis and understanding. The slicing technique is applied to UML class diagrams and state machines. The main goal of slicing is model transformation. Similar practical approach for model slicing is also proposed which is based on extraction of sub-models from original model to ease software visualization. The proposed methodology uses the idea of model based slicing of sequence diagrams to extract desired sub-models [42]. Further recent work is introduced by Wuliang Sun et al. [44] where authors invented model slicing for invariant checking, and applied a slicing technique to reduce the size inputs to improve efficiency of verification process in current tools. It is further proven that model slicing can drastically reduce the verification time. A general model-based slicing framework is also proposed that can be used to define both program and model slicing. The purpose of the framework is to construct slices written in a UML-like language [13].

The Kompren language to generate model slicers for Domain Specific Modeling Languages (DSMLs) is recommended for different purposes: for example, examining and model understanding. This work presents the model properties of the slices which can be extracted from different forms of the slicer [6], [7].

In contrast to these previous works, our paper describes a slicing criterion oriented towards the verification of satisfiability of UML/OCL class diagrams. We slice UML/OCL class diagrams before transformation. Previous works either do not target UML class diagrams or do not consider OCL (other than as a notation to express slicing criteria) and none propose a slicing criterion for verification.

Another source of relevant work appears in the underlying theorem provers and solvers used to check satisfiability in UML/OCL models. At this level, similar concepts for partitioning, symmetry-breaking and other optimisations have been considered extensively, for instance [16], [34], [52]. We suggest that slicing *before* the translation into a formalism like SAT or CSP is worthwhile for several reasons. First, slicing analysis is independent of the underlying formalism, so it can benefit a variety of tools. Furthermore, at this level of abstraction the problem is smaller, so it is feasible to perform more complex analysis. Finally, we can take advantage of our knowledge of the semantics of UML/OCL and the property to be verified, information which can be lost in the translation into the formalism. For instance, the removal of derived value constraints proposed in Section “Trivially Satisfiable Constraints” would not be possible without precise information about the property to be checked.

IX. Conclusions and Future Work

This paper presents a novel slicing technique for UML/OCL class diagrams aimed at making the verification of satisfiability more efficient. The approach receives as input a UML class diagram annotated with OCL constraints and automatically breaks it into submodels whose satisfiability can be analysed independently. Then, the satisfiability of the original model can be established by checking if at least one submodel (weakly satisfiable) or all submodels (strongly satisfiable) are satisfiable. A benefit of this approach is that it is independent of the underlying formalism used to check satisfiability and can therefore be applied in many existing tools.

A prototype implementation of the slicing procedure has been developed on top of the tool UMLtoCSP. Experimental results show that slicing can produce a significant speed-up in verification time. The amount of speed-up achieved by this method depends on the specific model, from none to several orders of magnitude. As the overhead introduced by slicing analysis is negligible, we believe that slicing is a useful addition to any UML/OCL satisfiability-checking toolkit. Furthermore, we have demonstrated the slicing technique on a real-world case study (DBLP conceptual schema) to analyse the benefits. This real-world case study is programmed in Alloy, which is a popular tool and widely used for verification of models. We applied the slicing technique and achieved drastic speed-up in this tool as well.

As regards our future work, we plan to create a verification engine with slicing techniques that can break unverifiable models into several independent submodels and verify them. With the help of slicing procedures, current verification methods will be efficient enough to verify models with an additional level of complexity that no current tool can handle.

Acknowledgements

This paper is a revised and extended version of the one presented at Automated Software Engineering (ASE 2010), Antwerp, Belgium had have received 38 citations.

References

[1] K. Anastasakis, B. Bordbar, G. Georg, and I. Ray. UML2Alloy: A challenging model transformation. In *ACM/IEEE 10th Int. Conf. on Model Driven Engineering Languages and Systems (MODELS 2007)*, volume 4735 of *LNCS*, pages 436–450, 2007.

[2] K. Androutsopoulos, D. Binkley, D. Clark, N. Gold, M. Harman, K. Lano, and Z. Li. Model projection: simplifying models in response to restricting the environment. In *ICSE*, pages 291–300, 2011.

[3] M. Balaban and A. Maraee. A UML-based method for deciding finite satisfiability in Description Logics. In *DL'2008*, volume 353 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008.

[4] P. Baumgartner, U. Furbach, M. Gross-Hardt, and T. Kleemann. Model based deduction for database schema reasoning. In *KI 2004: Advances in Artificial Intelligence*, pages 168–182. Springer, 2004.

[5] D. Berardi, D. Calvanese, and G. D. Giacomo. Reasoning on UML class diagrams. *AIntelligence*, 168:70–118, 2005.

[6] A. Blouin, B. Combemale, B. Baudry, and O. Beaudoux. Modeling model slicers. In *MoDELS*, pages 62–76, 2011.

[7] A. Blouin, B. Combemale, B. Baudry, and O. Beaudoux. Kompren: Modeling and generating model slicers. *Software and Systems Modeling (SoSyM)*, 2012.

[8] A. D. Brucker and B. Wolff. The HOL-OCL book. Technical Report 525, ETH Zurich, 2006.

[9] J. Cabot, R. Clarisó, and D. Riera. UMLtoCSP: a tool for the formal verification of UML/OCL models using constraint programming. In *ASE'2007*, pages 547–548. ACM, 2007.

[10] J. Cabot, R. Clarisó, and D. Riera. Verification of uml/ocl class diagrams using constraint programming. In *ICSTW '08: Proceedings of the 2008 IEEE International Conference on Software Testing Verification and Validation Workshop*, pages 73–80, Washington, DC, USA, 2008. IEEE Computer Society.

[11] O. Chebaro, N. Kosmatov, A. Giorgetti, and J. Julliand. Program slicing enhances a verification technique combining static and dynamic analysis. In *SAC*, pages 1284–1291, 2012.

[12] D. Clark. Amorphous slicing for EFSMs. In *PLID' 07*, 2007.

[13] T. Clark. A general model-based slicing framework. In *Proc. of the Workshop on Composition and Evolution of Model Transformations*, 2011.

[14] J. Conesa and A. Olivé. Pruning ontologies in the development of conceptual schemas of information systems. In *ER'2004*, volume 3288 of *LNCS*, pages 122–135. Springer, 2004.

[15] DBLP. Digital bibliography andy library project, 2012. http://guifre.lsi.upc.edu/DBLP.pdf.

[16] V. Durairaj and P. Kalla. Guiding CNF-SAT search via efficient constraint partitioning. In *ICCAD'04*, pages 498–501. IEEE Computer Society, 2004.

[17] M. Gogolla, J. Bohling, and M. Richters. Validating UML and OCL models in USE by automatic snapshot generation. *Journal on Software and System Modeling*, 4(4):386–398, 2005.

[18] M. Gogolla and M. Richters. Expressing UML Class Diagrams Properties with OCL. In *AOM with the OCL*, volume 2263 of *LNCS*, pages 86–115. Springer, 2001.

[19] M. P. E. Heimdahl, J. M. Thompson, and M. W. Whalen. On the effectiveness of slicing hierarchical state machines: A case study. In *EUROMICRO*, pages 10435–10444, 1998.

[20] D. Jackson. Alloy: a lightweight object modelling notation. *ACM Transactions on Software Engineering and Methodology*, 11(2):256–290, 2002.

[21] D. Jackson. *Software Abstractions: Logic, Language and Analysis*. MIT Press, 2006.

[22] E. K. Jackson, E. Kang, M. Dahlweid, D. Seifert, and T. Santen. Components, platforms and possibilities: towards generic automation for mda. In *EMSOFT*, pages 39–48, 2010.

[23] W. Ji, D. Wei, and Q. Zhi-Chang. Slicing hierarchical automata for model checking uml statecharts. In *Formal Methods and Software Engineering*, volume 2495 of *Lecture Notes in Computer Science*, pages 435–446. Springer Berlin / Heidelberg, 2002.

[24] H. H. Kagdi, J. I. Maletic, and A. Sutton. Context-free slicing of UML class models. In *ICSM'05*, pages 635–638. IEEE Computer Society, 2005.

[25] S. Khurshid. *Generating structurally complex tests from declarative constraints*. PhD thesis, Massachusetts Institute of Technology, 2003.

[26] T. H. Kim, Y. T. Song, L. Chung, and D. Huynh. Software architecture analysis: A dynamic slicing approach. *International Journal of Computer & Information Science*, 1(2):91–103, 2000.

[27] R. Kollmann and M. Gogolla. Metric-based selective representation of uml diagrams. In *CSMR'02*, pages 89–98. IEEE Computer Society, 2002.

[28] B. Korel, I. Singh, L. Tahat, and B. Vaysburg. Slicing of state-based models. In *Proceedings of the International Conference on Software Maintenance*, ICSM '03, pages 34–. IEEE Computer Society, 2003.

[29] J. T. Lallchandani and R. Mall. Slicing UML architectural models. In *ACM / SIGSOFT SEN*, volume 33, pages 1–9, 2008.

[30] J. T. Lallchandani and R. Mall. A dynamic slicing technique for uml architectural models. *IEEE Trans. Software Eng.*, 37(6):737–771, 2011.

[31] K. Lano. Slicing of uml state machines. In *Proceedings of the 9^{th} WSEAS International Conference on Applied Informatics and Communications*, pages 63–69. World Scientific and Engineering Academy and Society (WSEAS), 2009.

[32] K. Lano and S. K. Rahimi. Slicing of uml models using model transformations. In *MoDELS (2)*, pages 228–242, 2010.

[33] K. Lano and S. K. Rahimi. Slicing techniques for uml models. *Journal of Object Technology*, 10:11: 1–49, 2011.

[34] Y. C. Law and J. H. Lee. Symmetry breaking constraints for value symmetries in constraint satisfaction. *Constraints*, 11(2-3):221–267, 2006.

[35] A. Maraee and M. Balaban. Efficient reasoning about finite satisfiability of UML class diagrams with constrained generalization sets. In *ECMDA-FA'2007*, volume 4530 of *LNCS*, pages 17–31. Springer, 2007.

[36] B. J. Peterson, W. A. Andersen, and J. Engel. Knowledge bus: Generating application-focused databases from large ontologies. In *KRDB '98*, volume 10 of *CEUR Workshop Proceedings*, pages 2.1–2.10. CEUR-WS.org, 1998.

[37] A. Queralt and E. Teniente. Reasoning on UML class diagrams with OCL constraints. In *ER'2006*, volume 4215 of *LNCS*, pages 497–512. Springer-Verlag, 2006.

[38] A. Shaikh, R. Clarisó, U. K. Wiil, and N. Memon. Verification-driven slicing of uml/ocl models. In *ASE*, pages 185–194, 2010.

[39] A. Shaikh and U. K. Wiil. UMLtoCSP (UOST): a tool for efficient verification of UML/OCL class diagrams through model slicing. In *20th ACM SIGSOFT Symposium on the Foundations of Software Engineering*

(FSE-20), SIGSOFT/FSE'12, Cary, NC, USA - November 11 - 16, 2012, page 37, 2012.

[40] A. Shaikh, U. K. Wiil, and N. Memon. UOST: UML/OCL aggressive slicing technique for efficient verification of models. In *System Analysis and Modeling: About Models - 6th International Workshop, SAM 2010, Oslo, Norway, October 4-5, 2010, Revised Selected Papers*, pages 173–192, 2010.

[41] A. Shaikh, U. K. Wiil, and N. Memon. Evaluation of tools and slicing techniques for efficient verification of uml/ocl class diagrams. *Adv. Software Engineering*, 2011, 2011.

[42] R. Singh and V. Arora. A practical approach for model based slicing. *IOSR Journal of Computer Engineering*, 12(4):18–26, 2013.

[43] J. A. Stafford, D. J. Richardson, and A. L. Wolf. Architecture-level dependence analysis in support of software maintenance. In *ISAW'98*, pages 129–132, 1998.

[44] W. Sun, B. Combemale, and R. B. France. Towards the use of slicing techniques for an efficient invariant checking. In *Companion Proceedings of the 14th International Conference on Modularity*, pages 23–24. ACM, 2015.

[45] B. Swartout, P. Ramesh, K. Knight, and T. Russ. Toward distributed use of large-scale ontologies. *AAAI Symp. on Ontological Engineering*, pages 138–148, 1997.

[46] F. Tip. A survey of program slicing techniques. *Journal of Programming Languages*, 3:121–189, 1995.

[47] E. Torlak and D. Jackson. Kodkod: A relational model finder. In *Tools and Algorithms for the Construction and Analysis of Systems*, pages 632–647. Springer, 2007.

[48] Z. Ujhelyi, Á. Horváth, and D. Varró. Dynamic backward slicing of model transformations. In *ICST*, pages 1–10, 2012.

[49] E. Uzuncaova. *Efficient specification-based testing using incremental techniques*. ProQuest, 2008.

[50] E. Uzuncaova, D. Garcia, S. Khurshid, and D. Batory. Testing software product lines using incremental test generation. In *Software Reliability Engineering, 2008. ISSRE 2008. 19th International Symposium on*, pages 249–258. IEEE, 2008.

[51] E. Uzuncaova and S. Khurshid. Program slicing for declarative models. *ACM SIGSOFT Software Engineering Notes*, 31(6):1–2, 2006.

[52] E. Uzuncaova and S. Khurshid. Kato: A program slicing tool for declarative specifications. In *ICSE '07*, pages 767–770, 2007.

[53] E. Uzuncaova and S. Khurshid. Constraint prioritization for efficient analysis of declarative models. In *FM 2008: Formal Methods*, pages 310–325. Springer, 2008.

[54] E. Uzuncaova, S. Khurshid, and D. Batory. Incremental test generation for software product lines. *Software Engineering, IEEE Transactions on*, 36(3):309–322, 2010.

[55] S. Van Langenhove and A. Hoogewijs. sv_tl: System verification through logic tool support for verifying sliced hierarchical statecharts. In *Recent Trends in Algebraic Development Techniques*, volume 4409 of *Lecture Notes in Computer Science*, pages 142–155. Springer Berlin / Heidelberg, 2007.

[56] M. Weiser. Program slicing. *IEEE Trans. Software Eng.*, 10(4):352–357, 1984.

[57] J. Zhao. Applying slicing technique to software architectures. In *ICECCS*, pages 87–99, 1998.

NEB in Analysis of Natural Image 8×8 and 9×9 High-contrast Patches

Shengxiang Xia
School of Science,
Shandong Jianzhu University,
Jinan 250101 P.R.China.

Wen Wang
School of Science,
Shandong Jianzhu University,
Jinan 250101 P.R.China.

***Abstract*—In this paper we use the nudged elastic band technique from computational chemistry to investigate sampled high-dimensional data from a natural image database. We randomly sample 8×8 and 9×9 high-contrast patches of natural images and create a density estimator believed as a Morse function. By the Morse function we build one-dimensional cell complexes from the sampled data. Using one-dimensional cell complexes, we identify topological properties of 8×8 and 9×9 high-contrast natural image patches, we show that there exist two kinds of subsets of high-contrast 8×8 and 9×9 patches modeled as a circle, by the new method we confirm some results obtained through the method of computational topology.**

Keywords*—*nudged elastic band; natural image high-contrast patch; cell complex; density function

I. Introduction

Computational topology becomes a very important and efficiently method to analyse high-dimensional dada in the recent years [1], [2], [3], [4]. To analyse high-dimensional dada, we usually construct a sequence of simplicial complexes from the finite sampled data set to produce a simple combinatorial presentations of the data, the most commonly used complexes include Čech complexes, Rips complexes and lazy witness complexes. As the dimensional problem, the constructed simplicial complexes usually compose of thousands (even tens of thousands) of simplices , they are sometimes too large to compute. Adams, Atanasov, and Carlsson [5] used the nudged elastic band method to construct cell complexes through density functions of sampling data, they built more low-priced reasonable models for some nonlinear data sets (such as, sets generated from social networks, from range image analysis, and from microarray analysis) by a few of cell complexes, and effectively detect the homology of the nonlinear data sets, it initially shows that cell complex models are efficient ways for analysing high-dimensional nonlinear data. Adams etc [5] obtained a circle model and the three circle model for the data set of 3×3 optical image patches from [6]. In the paper [7], Xia shown that there exist some core subsets of 8×8 and 9×9 natural image patches that are topologically equivalent to a circle and the three circle model respectively.

In this paper, we utilize the methods of the paper [5] to identify topological features of spaces of 8×8 and 9×9 natural image patches, we discover a circle model for 8×8 and 9×9 patches. The data sets used here are drawn from INRIA Holidays dataset [8], that are different from the data set in the paper [5].

II. Basic Concepts

A. Nudged elastic band

The nudged elastic band (NEB) is an effective way for finding a minimum energy path between two initial stable states. The minimum energy path have the property of any point on the path being at an energy minimum in all directions perpendicular to the path [9].

An elastic band with $N + 1$ images can be defined by $[\mathbf{U}_0, \mathbf{U}_1, ..., \mathbf{U}_N]$, $\mathbf{U}_0$ and $\mathbf{U}_N$ are initial and final states. The $N - 1$ middle images are modified by an optimization algorithm [11].

The total force acting on each image is defined as following:

$$\begin{aligned}\mathbf{F}_i = \mathbf{F}_i^S|_{\|} - \bigtriangledown E(\mathbf{U}_i)|_{\perp} = (||\mathbf{U}_{i+1} - \mathbf{U}_i|| \\ -||\mathbf{U}_i - \mathbf{U}_{i-1}||)\tau_i - (\bigtriangledown E(\mathbf{U}_i) - \bigtriangledown E(\mathbf{U}_i) \cdot \tau_i),\end{aligned} \quad (1)$$

the first part $\mathbf{F}_i^S|_{\|}$ is called the spring force, the second part $\bigtriangledown E(\mathbf{U}_i)|_{\perp}$ is true force, and $\tau_i = \frac{(\mathbf{U}_{i+1}-\mathbf{U}_{i-1})}{||\mathbf{U}_{i+1}-\mathbf{U}_{i-1}||}$ local tangent at image i. where E is the energy of the system.

The nudged elastic band method apply an optimization algorithm to shift the images depending to the force in (1) for finding the minimum energy path. For more details of NEB, please refer to papers [8], [10], [11]

B. CW complexes

A k dimensional closed ball $\{x \in \mathcal{R}^k | \, ||x|| \leq 1\}$ is called a k-cell. A CW complex is a topological space X defined by the follow inductive steps. The 0-skeleton $X^{(0)}$ of X is a set of 0-cells. The 1-skeleton $X^{(1)}$ is created by gluing the endpoints of 1-cells to the 0-skeleton. Inductively, the k-skeleton $X^{(k)}$ are built by gluing the boundaries of k-cells to the $(k-1)$-skeleton $X^{(k-1)}$.

C. Morse theory

Suppose M be a compact manifold and a smooth Morse function $f : M \longrightarrow \mathcal{R}$ has non-degenerate critical points $t_1, ..., t_k \in M$ such that

$$p_0 < f(t_1) < p_1 < f(t_2) < \cdots < p_{k-1} < f(t_k) < p_k.$$

Suppose $M_p = f^{-1}((-\infty, p])$ is the sublevel set corresponding to $p \in \mathcal{R}$. It follows from Morse theory that M_{p_i} is homotopy equivalent to a CW complex with a λ_i-cell for each critical point t_i.

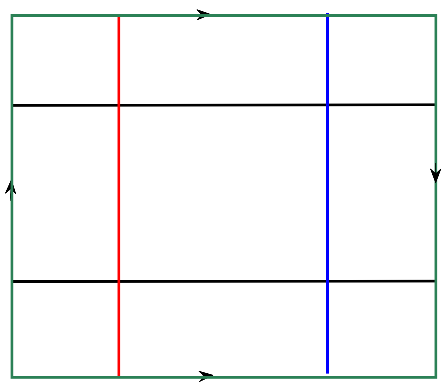


Fig.1. Denotation of Klein bottle as an identification space

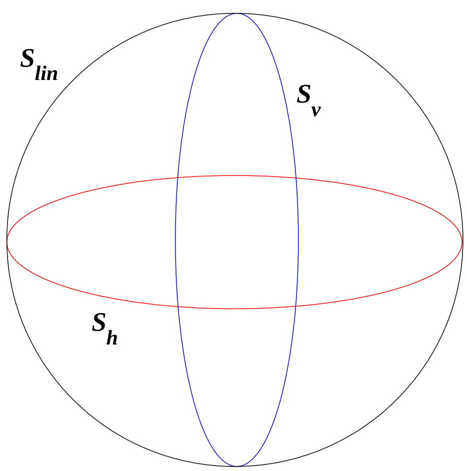


Fig.2. The three circle model.

D. The three circles model

The Klein bottle can be represented by pasting a square as Fig.1. While pasting a square, three circles are created, one is the main circle (S_{lin}) informed by horizontal segments (black lines), the other two circles (S_v and S_h) are informed from the vertical segments (red line and blue line) respectively, that is called the three circle model (Fig.2), represented by C_3. In the three circle space, the circles S_v and S_h intersect the main circle S_{lin} in exactly two points, but they themselves do not intersect.

III. The Data Sets of Natural Image Patches

We select data sets of 8×8 and 9×9 high-contrast patches from natural images of INRIA Holidays dataset [8]. Each data set consists of 5.5×10^5 high-contrast log patches. INRIA Holidays dataset is available at http://lear.inrialpes.fr/ jegou/data.php. Fig.3 is a sample.

Our spaces X_8 and X_9 are sets of 8×8 and 9×9 patches of high contrast created by the following steps.

Step 1. Select 550 images from INRIA Holidays dataset.

Step 2. Using MATLAB function rgb2gray to calculate the intensity at each pixel for each image.

Fig.3. A Sample from INRIA Holidays dataset

Step 3. We randomly choose 5000 8×8 and 9×9 patches from each image.

Step 4. We consider each patch as a n^2-dimensional vector, and take the logarithm of each coordinate.

Step 5. For any vector $\mathbf{x}$=$(x_1, x_2, ..., x_n)$, we calculate the D-norm: $\parallel \mathbf{x} \parallel_D$. Two coordinates of $\mathbf{x}$ are neighbors, expressed by $i \sim j$, if the corresponding pixels in the $n \times n$ patch are adjacent. The formula of D-norm is: $\parallel \mathbf{x} \parallel_D$= $\sqrt{\sum_{i \sim j}(x_i - x_j)^2}$.

Step 6. We pick the patches that have a D-norm in the top $t = 20\%$ percent in each image.

Step 7. Subtract an average of all coordinates from each coordinate.

Step 8. We map X_8 (X_9) into the unit sphere S^{63} (S^{80}) by dividing each vector with its Euclidean norm.

Step 9. We randomly sample 50,000 points from X_8 and X_9 for computational convenience, the subspaces of X_8 and X_9 are represented by $\bar{X}_8$ and $\bar{X}_9$ respectively.

In this paper, we use set symbols similar as in the papers [7], [12], $\bar{X}_n(15000)$ is a random subset of X_n with size 15000 $(n = 8, 9)$. We do not make the discrete cosine transform for these sets.

IV. Computing Method

we give main steps of calculating method used in this section, for more details of the method, please refer to the paper [5].

Given a data set $X \subset \mathcal{R}^n$ from unknown probability density function $f : \mathcal{R}^n \longrightarrow [0, \infty)$. We take superlevel sets

$$X^\alpha = f^{-1}([\alpha, \infty)) = \{x \in \mathcal{R}^n | f(x) \geq \alpha\}$$

, the high dense regions of data set X may give important topological information of X. We will construct CW complex models Z^α to approximate the superlevel sets X^α.

We construct only the one-dimensional skeleton of the cell complex by following three steps. First step, we create a differentiable density estimator to approximate the unknown probability density function. Second step, we acquire local maxima of the density estimate to give 0-cells. Third step, we randomly produce initial bands, then find the convergent bands by NEB, thus we obtain 1-cells.

A. Density estimator

For a data set $X \subset \mathcal{R}^n$, let $\Phi_{x,\sigma} : \mathcal{R}^n \longrightarrow [0,\infty)$ be the probability density of a normal distribution centered at $x \in X$, we apply a differentiable density estimator $g(y) = |X|^{-1}\sum_{x\in X}\Phi_{x,\sigma}(y)$ to approach the unknown density.

B. 0-cells

To find 0-cells, we randomly select an initial point $y_0 \in X$, and iteratively define a sequence $\{y_0, y_1, ...\}$ with $y_{n+1} = m(y_n)$, where $m(y) : \mathcal{R}^n \longrightarrow \mathcal{R}^n$ is the mean shift function given by the formula

$$m(y) = \frac{\sum_{x\in X}\Phi_{x,\sigma}(y)x}{\sum_{x\in X}\Phi_{x,\sigma}(y)}.$$

The sequence $\{y_n\}$ converges to a local maxima of g [13]. In order to identify different 0-cells, we use single-linkage clustering to cluster the convergent points, and choose the densest member from each cluster as a 0-cell.

C. 1-cells

For two 0-cells, there is a 1-cell between them if we find a convergent band between them by using NEB. For an initial band $[\mathbf{U}_0, \mathbf{U}_1, ..., \mathbf{U}_N]$, where $\mathbf{U}_0$ and $\mathbf{U}_N$ are 0-cells. Using a similar formula as (1), the total force on each midterm node $\mathbf{U}_i$ is computed by

$$\mathbf{F}_i = (||\mathbf{U}_{i+1} - \mathbf{U}_i|| - ||\mathbf{U}_i - \mathbf{U}_{i-1}||)\tau_i + c\bigtriangledown g(\mathbf{U}_i)|_\perp + \mathbf{F}_{sm}. \quad (2)$$

Where $c = (\sigma\sqrt{2\pi})^n\sqrt{e}$ is the gradient constant. Let $h_{\theta,\phi}(x) : [0,\pi] \longrightarrow [0,1]$ be a function defined by

$$h_{\theta,\phi}(x) = \begin{cases} 0, & x \le \theta \\ (1-\cos(\frac{x-\theta}{x-\phi}\pi))/2, & \theta < x < \phi \\ 1, & \phi \le x \end{cases}$$

Let α_i be the angle between $(\mathbf{U}_{i+1} - \mathbf{U}_i)$ and $(\mathbf{U}_i - \mathbf{U}_{i-1})$, the smoothing force $\mathbf{F}_{sm}$ is computed by $h_{\theta,\phi}(\alpha_i)(\mathbf{U}_{i+1} - 2\mathbf{U}_i + \mathbf{U}_{i-1})$, here $\theta = \pi/6$ and $\phi = \pi/2$.

V. Experimental results

The author of the paper [7] used persistent homology to detect the topological structure of spaces X_n of $n \times n$ natural image patches ($n = 8, 9$), and shown that the topologies of the core sets vary from a circle to a 3-circle model as decreasing of density estimator. Especially, there are core subsets $\bar{X}_n(300, 20)$ in X_n ($n = 8, 9$) , whose homology is that of a circle. X_8 and X_9 have core subsets $\bar{X}_8(15, 20)$ and $\bar{X}_9(15, 20)$ respectively possessing the homology of the three circle model C_3. By using the method in [7], we can check that the core subsets $\bar{X}_8(20, 25)$ and $\bar{X}_9(20, 20)$ of X_8 and X_9 have the homology of the three circle model C_3 respectively.

Now we use NEB to analyse some subsets of X_8 and X_9, here we utilize two types subsets of X_n: (1) random subsets $\bar{X}_n(15000)$ of X_n with size 15000; (2) core subsets $\bar{X}_n(k, p)$.

TABLE I. Data set information

	$\bar{X}_8(15000)$	$\bar{X}_8(300,20)$	$\bar{X}_8(15,20)$	$\bar{X}_8(20,25)$
size of data set	15000	10000	10000	12500
dimension n	64	64	64	64
standard deviation σ	0.38	0.35	0.35	0.35

TABLE II. Data set information

	$\bar{X}_9(15000)$	$\bar{X}_9(300,20)$	$\bar{X}_9(15,20)$	$\bar{X}_9(20,20)$
size of data set	15000	10000	10000	10000
dimension n	81	81	81	81
standard deviation σ	0.30	0.38	0.35	0.38

A. 8×8 patches

The data sets used for X_8 are shown in Table 1. For the set $\bar{X}_8(15000)$, we let standard deviation $\sigma = 0.38$, we get four 0-cells with densities 0.7664, 0.7826, 0.877, and 0.8818 respectively, and four 1-cells having densities 0.7597, 0.7664, 0.7672 , and 0.7772 respectively, these cells produce a loop. Therefore for $\alpha = 0.7597$, the Z^α is a circle (Fig.4). When we choose standard deviation $\sigma = 0.30$, we have four 0-cells having densities in [1.02×10^6, 1.356×10^6] and four 1-cells with densities in [8.001×10^5, 8.566×10^5], these cells also produce a circle.

For $\bar{X}_8(300, 20)$, we take the value of standard deviation $\sigma = 0.35$, we get four 0-cells having densities of 277.2, 294.5, 396.5 and 401.7 and four 1-cells having densities of 216.1, 224.1, 227.5 and 228.9, all these cells compose a circle (Fig.5).

For $\bar{X}_8(15, 20)$, let $\sigma = 0.35$, we find four 0-cells and four 1-cells which compose a circle (Fig.6). The four 0-cells have densities of 263.6, 279.4, 366, and 378.9, the four 1-cells have densities of 207.1, 217.3, 218.7, and 218.9. When we take σ=0.30, and 0.38, we obtain similar results as for $\sigma = 0.35$. For example, for $\sigma = 0.38$, the four 0-cells have densities in [1.649, 2.332] and the four 1-cells have densities in [1.42, 1.497].

For $\bar{X}_8(20, 25)$, we have the similar result as $\bar{X}_8(15, 20)$. Taking $\sigma = 0.25$, 0.30, 0.35 and 0.38, we get four 0-cells and four 1-cells for each case which form a circle. For example, when $\sigma = 0.35$, the four 0-cells have densities 240.7, 253.2, 330.1, 335.8 and the four 1-cells have densities 198.4, 205.5, 205.7, 206.1 respectively, these cells form a circle (Fig.7).

As shown in [7], the subspace $\bar{X}_8(15, 20)$ ($\bar{X}_8(20, 25)$) of X_8 has same homology as the three circle model C_3, but we can not detect that $\bar{X}_8(15, 20)$ ($\bar{X}_8(20, 25)$) has the topology of C_3 by using the current method.

B. 9×9 patches

In this section we utilize the computing method for two kind of data sets in Table 2. We consider the set $\bar{X}_9(15000)$ for standard deviation $\sigma = 0.30$, we find four 0-cells, whose densities are 1.138×10^8, 1.15×10^8, 1.587×10^8, 1.627×10^8 respectively, and four 1-cells with densities of 9.364×10^7, 9.405×10^7, 9.408×10^7, 9.884×10^7, all these cells form a circle. Thus $\alpha = 9.364 \times 10^7$, the Z^α is a circle (Fig.8). If we take $\sigma = 0.35$, we get four 0-cells with densities in [864, 1091], and four 1-cells having densities in [822.7, 850.3], which also compose a circle.

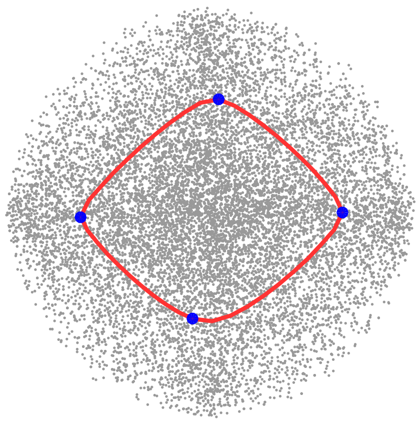


Fig.4. $\bar{X}_8(15000)$ and the circle $Z^{0.7597}$, projected to a plane.

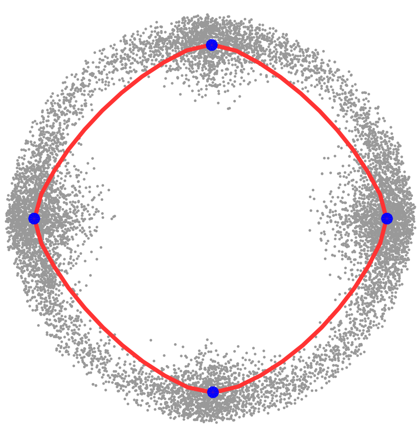


Fig.5. $\bar{X}_8(300, 20)$ and the circle $Z^{216.1}$, projected to a plane.

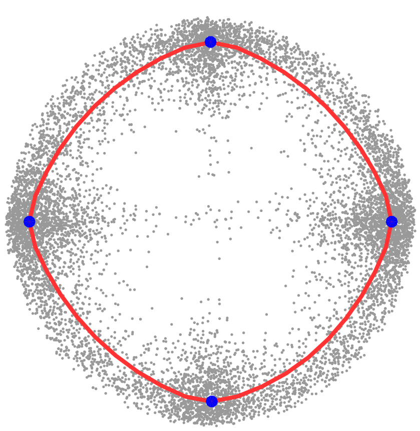


Fig.6. $\bar{X}_8(15, 20)$ and the circle $Z^{207.1}$, projected to a plane.

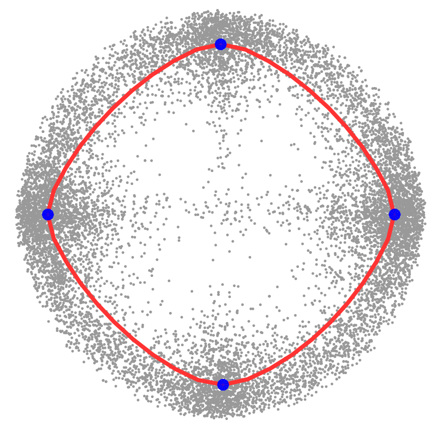


Fig.7. $\bar{X}_8(20, 25)$ and the circle $Z^{198.4}$, projected to a plane.

For $\bar{X}_9(300, 20)$, when we take $\sigma = 0.35$, we have four 0-cells with densities in [23800, 35840] and four 1-cells having densities in [1910, 2031], all these cells form a circle (Fig.9). If we let σ=0.25, 0.30, and 0.38, we get similar results as for $\sigma = 0.35$. However the densities of cells decrease as increasing of standard deviation, for example, when $\sigma = 0.38$, we get four 0-cells having densities in [3.704, 5.504] and four 1-cells with densities in [3.263, 3.436].

We take standard deviation σ=0.20, 0.25, 0.30, 0.35, and 0.38 for $\bar{X}_9(15, 20)$, for each case we obtain four 0-cells and four 1-cells, these cells compose a circle. For example, if $\sigma = 0.35$, four 0-cells have densities 2254, 2308, 3240, and 3333, four 1-cells have densities 1817, 1863, 1908, and 1935 respectively, these cells form a circle (Fig.10).

For $\bar{X}_9(20, 20)$, we get the similar result as $\bar{X}_9(15, 20)$. For example, when $\sigma = 0.38$, we find four 0-cells with densities in [3.548, 5.138] and four 1-cells with densities in [3.143, 3.292], these cells form a circle (Fig.11).

We take various values of standard deviation σ and do experiments for them, but we can not find that $\bar{X}_9(15, 20)$ ($\bar{X}_9(20, 20)$) has the homology of C_3 by the current method.

VI. Conclusions

In this paper we utilize the nudged elastic band technique to analyse spaces of 8×8 and 9×9 natural image patches, and we get some similar results as the papers [2], [7], which show that the results got in this paper and [2], [7] are native properties of natural image patches, they do not rest on the methods and databases. We experimentally show that the spaces of high-contrast 8×8 and 9×9 patches have different subsets modeled as a circle. By matching the results (method) of this paper with the results (method) of the papers [2], [7], we discover that the most advantage of the method is its simplicity. For example, to model $\bar{X}^8(300, 20)$ as a circle using cell complexes, we only use four 1-cells, if we model $\bar{X}^8(300, 20)$ as a circle using witness complexes, we may need several tens of thousands witness complexes. The disadvantages of the method are that to create higher dimensional cells is more difficult [5] and it may only detect coarse topology of a data-set.

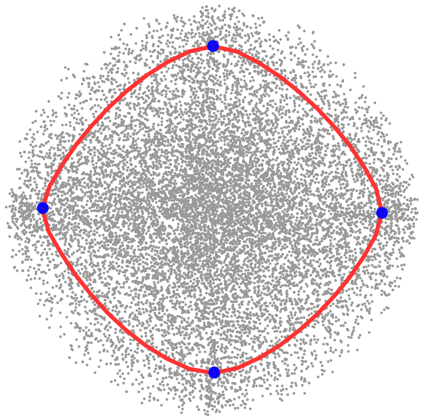


Fig.8. $\bar{X}_9(15000)$ and the circle $Z^{9.364\times10^7}$, projected to a plane.

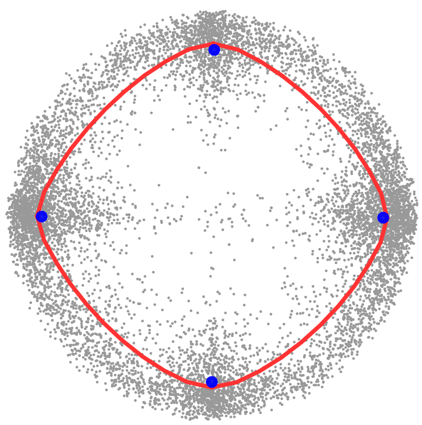


Fig.9. $\bar{X}_9(300, 20)$ and the circle Z^{1910}, projected to a plane.

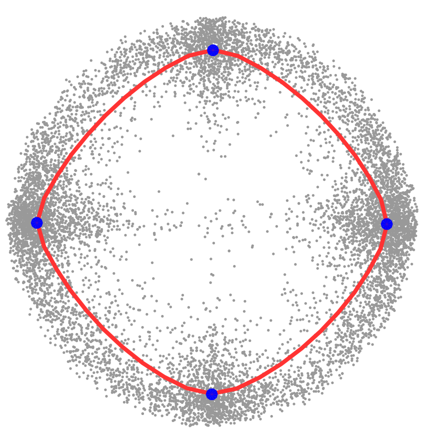


Fig.10. $\bar{X}_9(15, 20)$ and the circle Z^{1817}, projected to a plane.

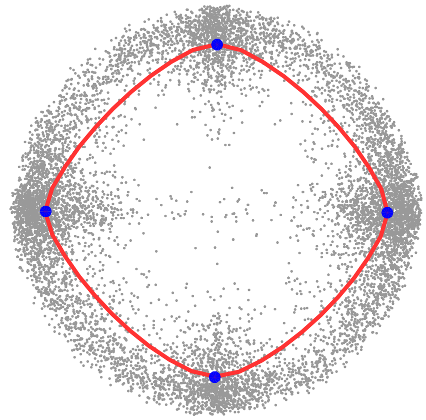


Fig.11. $\bar{X}_9(20, 20)$ and the circle $Z^{3.143}$, projected to a plane.

ACKNOWLEDGEMENTS

The authors are very grateful to the reviewers for valuable comments.

Project supported by the National Natural Science Foundation of China (Grant No.61471409).

REFERENCES

[1] H. Adams and G. Carlsson, *On the nonlinear statistics of range image patches*, SIAM J. Imag. Sci., **2** (2009), 110–117.

[2] G. Carlsson, T. Ishkhanov, V. de Silva, A. Zomorodian, *On the local behavior of spaces of natural images*, Internat. J. Computer Vision, 76 (2008), pp. 1–12.

[3] G. Carlsson, *Topology and data*, Bulletin (New Series) of the American Mathematical Society, 46 (2009), pp. 255–308.

[4] A. Zomorodian, *Topological data analysis, Advances in Applied and Computational Topology* (Proceedings of Symposia in Applied Mathematics), A. Zomorodian, eds., American Mathematical Society, 70 (2012), pp. 1–39.

[5] H. Adams, A. Atanasov, and G. Carlsson, *Nudged elastic band in toplogical data analysis*, Topological Methods in Nonlinear Analysis, 45 (2015), pp. 247–272.

[6] A. B. Lee, K. S. Pedersen, and D. Mumford, *The non-linear statistics of high-contrast patches in natural images*, Internat. J. Computer Vision, 54 (2003), pp. 83–103.

[7] S. Xia, *An Analysis on Natural Image* 8×8 *and* 9×9*-Patches*, submitted.

[8] H. Jegou, M. Douze, and C. Schmid, *Hamming Embedding and Weak geometry consistency for large scale image search*, Proceedings of the 10th European conference on Computer vision, October (2008), pp. 304–317.

[9] D. Sheppard, R. Terrell, and G. Henkelman, *Optimization methods for finding minimum energy paths*, Journal Chemical Physics, 128 (2008), pp. 134106-1–10.

[10] G. Henkelmana and H. Jónsson, *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points*, Journal Chemical Physics, 113 (2000), pp. 9978–9985.

[11] G. Henkelman, B. Uberuaga, H. Jónsson, *A climbing image nudged elastic band method for finding saddle points and minimum energy paths*, Journal Chemical Physics, 112 (2000), pp. 9901–9904.

[12] S. Xia, *A topological analysis of high-contrast patches in natural images*, J. Nonlinear Sci. Appl., 9 (2016), pp. 126–138.

[13] Y. Cheng, *Mean shift, mode seeking, and clustering*, IEEE Trans. Pattern Anal., 17 (1995), pp. 790–799.

Object Conveyance Algorithm for Multiple Mobile Robots based on Object Shape and Size

Purnomo Sejati, Hiroshi Suzuki, Takahiro Kitajima, Akinobu Kuwahara and Takashi Yasuno
Graduate School of Tokushima University
2-1 Minamijosanjima, Tokushima, 770-8506, JAPAN

***Abstract*—This paper describes a determination method of a number of a team for multiple mobile robot object conveyance. The number of robot on multiple mobile robot systems is the factor of complexity on robots formation and motion control. In our previous research, we verified the use of the complex-valued neural network for controlling multiple mobile robots in object conveyance problem. Though it is a significant issue to develop effective determination team member for multiple mobile robot object conveyance, few studies have been done on it. Therefore, we propose an algorithm for determining the number of the team member on multiple mobile robot object conveyance with grasping push. The team member is determined based on object weight to obtain appropriate formation. First, the object shape and size measurement is carried out by a surveyor robot that approaches and surrounds the object. During surrounding the object, the surveyor robot measures its distance to the object and records for estimating the object shape and size. Since the object shape and size are estimated, the surveyor robot makes initial push position on the estimated push point and calls additional robots for cooperative push. The algorithm is validated in several computer simulations with varying object shape and size. As a result, the proposed algorithm is promising for minimizing the number of the robot on multiple mobile robot object conveyance.**

***Keywords*—multiple mobile robots, object conveyance, team member determination;**

I. INTRODUCTION

In the mobile robotic research area, multiple mobile robot systems have grown significantly in size and importance in Recent years. The concern is given by the robotic researchers to this study is motivated by the variety of application which can be carried out by a group of the robot, such as exploration, agricultural foraging, military, warehouse, search, and rescue, and cleaning [1] [2]. Object conveyance is a well-studied topic in multiple mobile robot systems where robots should work Together to transport a box to a destination. In this task, the robots are required to achieve the major goal of delivering an object to a goal point as well as the maintenance of minor requirement including approaching the object with a formation, maintaining contact with the object and maintaining forward motion.

Multiple mobile robot systems offer several advantages that single robot can not achieve such as fault tolerant, efficiency, low cost, and flexibility. However, the complexity of formation and motion control may become high because of the increase in sensor information and actuators since the number of the robots is increased. Therefore, the team with minimum robots offers lower system load and can be addressed by dynamic team strategy. The dynamic team is an approach for defining robot number in a multiple mobile robot system where the mobile robots are a temporary and fluid team whose association properties are allowed to vary over time [3] [4] [5]. That is, teams dynamically and automatically grow and shrink, and the member may be substituted. Therefore, it enables multiple mobile robot systems becomes more efficient in the use of the robot.

In the object conveyance with a given object, the number of robots can be determined in advance [1] [6] [7]. However, in practical, multiple robot systems might be required to move any object in size and shape [8]. Therefore, determination of robot number before the push the object is the concern of this study. The clue of determining the number of the robots are the shape and size of the object which is moved. The object is assumed to be same in shape and size from bottom to the top like a box or a pipe. Thus, the object shape and size in two-dimensional represent its overall shape.

In this paper, a simple and efficient algorithm for determining push position is proposed. In the beginning, a robot approaches the object and stops at a certain distance. Then, the robot surrounds the object and measures the distance to the object. Recorded distance data are processed using Graham Scan algorithm to make a convex hull which is used to estimate object shape [9]. After that, object corner positions are obtained and between them, push points are determined. Also, we employ the dynamic team strategy to manage team member and use predefined formation to make the push. Moreover, each robot motion is controlled by a simple method that uses object position and subgoal as control input parameter. A simulation is attempted, and it shows satisfactory results.

The rest of this paper is organized as follows. Section II in the context of the relevant study on this problem and Section III describes the problem definition. In Section IV, the algorithm is explained. Results and discussion are presented in Section VI then the conclusion is in Section VII.

II. RELATED WORKS

The problem of object conveyance touches on many established Research areas. In this section, we discuss the relevant work in coordination, communication, decision-making system, and determination of the number of the robot.

A. Static vs. dynamic coordination

Coordination is the core task in multiple mobile robot systems and can affect overall performance system. There is static and dynamic coordination. Static coordination (also known as off line coordination [10]) is the coordination method with predefined rules for engaging a task. The rules in traffic control problem such as "keep right", "stop in intersection" And "keep sufficient distance to other robot" are the example of static coordination [11]. Dynamic coordination (also known as online coordination [10]) is carried out during execution of a mission and it based on the analysis and synthesis the information that can be obtained through the communication. The static method can handle complex tasks. However, its real time controlling might be poor. The dynamic method can well meet the capability of real-time. However, it has difficulty in dealing with more complex tasks.

B. Explicit vs. implicit communication

Communication enables each robot to interact and know the information of the position, sensor data and the state of an Environment with others in the system. Communication can be explicit and implicit.

Explicit communication refers to the means for the direct exchange of information between the robots, which can be made in the form of unicast or broadcast intentional messages. This often requires a dedicated on-board communication module. Existing coordination methods are mainly based on the use of explicit communication.

On the other hand, implicit communication refers to the way in which the robot gets information about other robots in the system through the environment. This should be achieved by embedding different kinds of sensors in the robot. Implicit communication can also be divided into two categories: active implicit communication (e.g., interaction via the environment) and passive implicit communication (e.g., interaction via sensing). Active implicit communication refers to the fact that the robots communicate by collecting the remaining information of others in the environment. The use of this form is generally related to the field of biometrics, and is usually inspired by the collective behavior of bees and ants. Passive implicit communication refers to the fact that the robots communicate by perceiving a change of environment through the use of sensors. For example, a robot needs to compute the context information (e.g., position and attitude) of others by modeling and reasoning based on the perceived data in order to cooperate with them.

C. Centralized vs. decentralized decision making

The decision-making guided by planning can be centralized or decentralized by the group architecture of The robots. There is a central control agent in centralized architectures that has the global information about the environment as well as all information about the robots, and which can Communicate with all the robots to share them. The central Control agent could be a computer or a robot. The advantage of the centralized architecture is that the central control agent has a global view of the world, whereby the globally optimal Plans can be produced. Nevertheless, this architecture: 1) is typical for a small number of robots and ineffectual for large teams with more robots; 2) is not robust about dynamic environments or failures in communications and other uncertainties; 3) produces a highly vulnerable system, and if the central control agent malfunctions, a new agent, must be available or else the entire team is disabled.

Decentralized architectures can be further divided into two categories: distributed architectures and hierarchical architectures. There is no central control agent in distributed architectures, such that all the robots are equal on control and are completely autonomous in the decision-making process. In contrast to a centralized architecture, a decentralized architecture can better respond to unknown or changing environments, and usually has better reliability, flexibility, adaptability, and robustness. Nevertheless, the solutions they reach are often suboptimal.

D. Fixed vs. variable number of the robot

The determination of the number of the robot in multiple mobile robot systems has two types. It can be fixed or variable depend on the mission. Many studies of multiple mobile robots coordination were carried out with the fixed number of the robot [12] [13] [14]. At least, they worked with two robots that is used from the beginning until finish of the mission [15]. The variable number of robots is the situation when the number of the robot can be changed at the execution of the mission [5].

III. Problem Setting

As stated above, this study is concerned with determining the number of the robot in multiple mobile robots object conveyance. Two wheeled mobile robots is used with disk shaped where the wheels are in the same direction to the center of the robot body. The illustration of object conveyance by multiple mobile robots on the world coordinate system Σ_W is shown in Fig. 1.

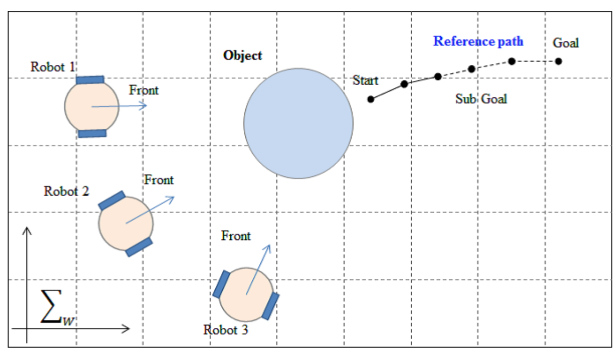


Fig. 1: Multiple mobile robots object conveyance.

A. Robot model

Figure 2(a) presents two-wheeled mobile robot movements. This mobile robot changes the direction depending on the velocities of each wheel. Radius of curvature, angular velocity, the velocity of the right wheel, the velocity of the left wheel, velocity of robot, and distance between the wheel are designated R_c, ω, v_R, v_L, v and l, respectively.

The simple abstraction of the radius R_i disk-shaped robot and its sensor configuration is shown in Fig. 2(b). The distance sensors are mounted on half front side of the robot that consists of eleven sensors s_0 to s_10.

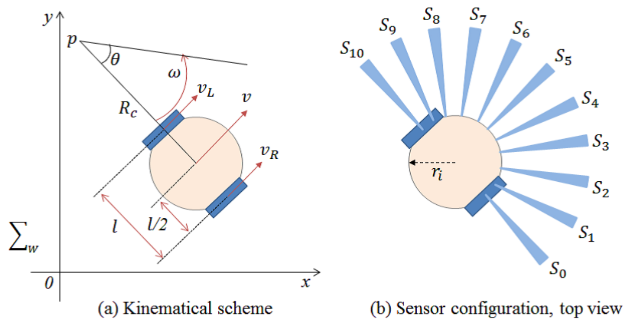


Fig. 2: Two-wheeled mobile robot model.

B. Assumptions

In order to specify the scope of our proposed algorithm, the assumptions about our task are made as follow.

- All robots have circle shape body which can contact with the object in any direction.
- Non-holonomic two wheeled robot is used.
- It has eleven distance sensors placed at frontside of halfbody to measure distance and direction on the robot direction toward any other object as shown in Fig. 2(a).
- A target object is placed on specified position.
- Actual environment consist of many objects as constraint or obstacle. However to simplify the process, robot and a target object are located on an obstacle free workspace.
- Each robot positions on the world coordination are known.
- Distance sensor reading is included with 0.2% random noise.
- Robot knows object position but have no information about object shape, size, and weight.
- Reference path consists of several subgoals. On the reference path, an object is required to be conveyed on the subgoals within 8 seconds.

C. Missions

Missions for the multiple mobile robot system is to convey the object on the reference path. In this study, to achieve this mission, multiple mobile robot systems are required to estimate object shape and size then determine push point for initial push. Regular and irregular shape object are used for demonstrating the algorithm. Object conveyance with minimum number of robot is also required to be demonstrated in order to verify the validity of the algorithm. These missions are carried out on a obstacle free workspace. Therefore, additional algorithm for avoiding obstacle is required for actual environment.

IV. Algorithm

In this object conveyance problem, an object is moved by pushing it. To make push mechanism, the position order is destination, object and robot. If the object is facing toward to the destination, then the robot position in the behind of the object. However, the robot will not able to determine where is the correct position for a push if object shape is unknown.

The flowchart of proposed algorithm for the dynamic team is shown in Fig. 3. The determination process of the robot number is first four blocks, and the rest three blocks are the conveyance process in general.

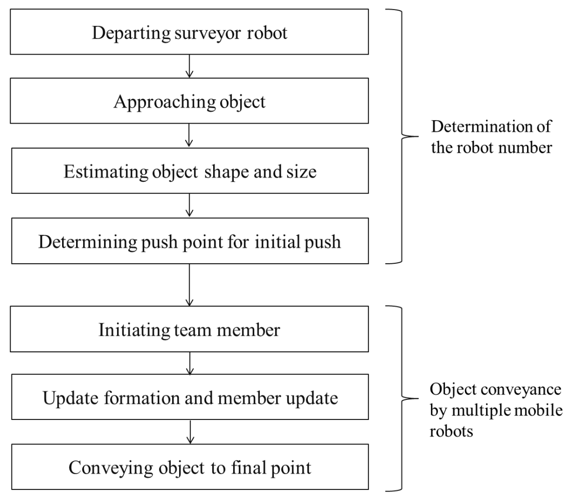


Fig. 3: Flowchart of dynamic team strategy.

A. Departing surveyor robot

A robot which is used for estimating object shape and size is called surveyor robot. The surveyor robot can be selected among the robots which have closest distance to the object. However, we choose random selection because the robots are located near each other. Also, the surveyor robot knows object position through positioning system. In practical, the positioning system can be an image, frequency, light or sound-based device which has advantages and disadvantages between each other. And in this simulation, position data can be obtained easily.

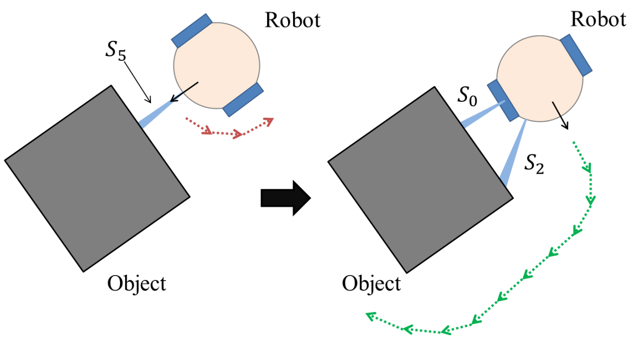
B. Approaching object

Since the surveyor robot is departed from the start position and it will approach directly to the object. During the travel, the robot can avoid the collision from other robots by detecting them using distance sensor. In order to determine robot avoidance movement for i-th robot, right side $D_r^{R_i}$ and left side $D_l^{R_i}$ sensing data are calculated and compared to get angular velocity ω by using equation (1) and (2). As the object position is known, surveyor robot is faced to its position by changing angular velocity. When the robot arrived at given distance from target object (60% of distance sensor S_5 reading) as shown in Fig. 4(a). The surveyor robot will stop and turn left to make it position in same direction with the object side. Then robot moves clockwise while keeping the distance by using sensor S_0 and S_2 as shown in Fig. 4(b). Moreover, the distance is measured by using sensor S_0. Start point of surrounding is recorded that will be used as stop point.

$$D_r^{R_i} = \sum_{k=0}^{4} \frac{1}{1+\exp(-1 \times ((12 \times {S_k}^{R_i}) - 6))} \quad (1)$$

$$D_l^{R_i} = \sum_{k=6}^{10} \frac{1}{1+\exp(-1 \times ((12 \times {S_k}^{R_i}) - 6))} \quad (2)$$

where ${S_k}^{R_i}$ is k-th sensor value of i-th robot.



(a) Approaching (b) Surrounding

Fig. 4: Approaching the object for measuring

C. Estimating object shape and size

The recorded distance data are converted to the coordinate and processed to get convex hull by using Graham Scan algorithm [9]. The first step in this algorithm is to find the point with the lowest y-coordinate on the world coordinate system Σ_W. If the lowest y-coordinate exists in more than one point in the set, the point with the lowest x-coordinate out of the candidates should be chosen as P_o. This step takes Q_n, where n is the number of points in question. Next, sort the remaining points of Q lexicographically by polar angle, measured in radians. Interior points on the ray cannot be convex hull points and remove these points during sort. Once the points are sorted, we connected them in counterclockwise order with respect to the anchor point P_o as shown in Fig. 5.

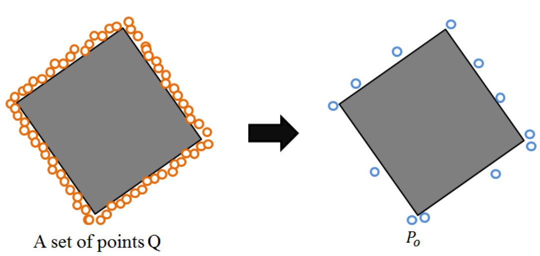


Fig. 5: Graham Scan algorithm

D. Determining push points

Convex hull is a set of recorded points which is simplified But it may contains unnecessary points. For example, a rectangular object consists of four corners and four straight lines. In rectangular convex hull, some points will appear between corner points and it is not necessary as shown in Fig. 6. Thus, such points must be eliminated to simplify next process. To eliminate these points, simple corner detection method is used. The method calculates the angle between two lines that composed by three sequences of points. If at the bottom corner of Fig. 6 is point Q_n, then point Q_{n+1} and Q_{n+2} are next points at clockwise direction. And then, the angle of first line (point Q_n and Q_{n+1}) to second line (Q_{n+1} and Q_{n+2}) are calculated. If the angle are larger or equal than minimum angle θ_e, point Q_{n+1} is considered as a corner. θ_e is angle threshold for detecting corner.

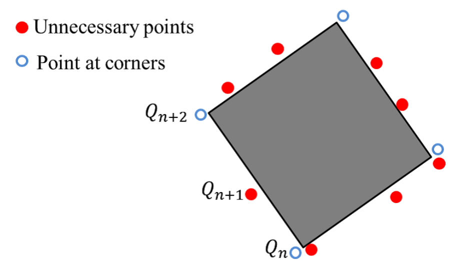


Fig. 6: Points of convex hull

After corner detection is finished, we have a set of points T_n that consist of corner points. However, push points are not in corner points because it has a small contact surface. Therefore, a position between corners is chosen as the push point. These points can be calculated by using average of T_n and T_{n+1}. Push points of the rectangular object are illustrated as the black circle in Fig. 7. The push point for the initial push is selected among them if its location is in the behind of the object.

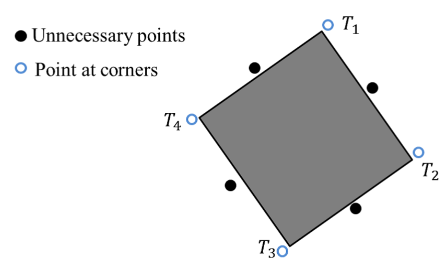


Fig. 7: Push points of rectangular object

E. Conveyance process

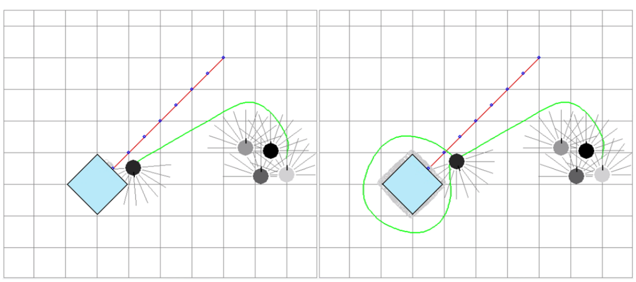
Object conveyance by multiple mobile robots is carried out by simple push. The surveyor robot performs initial push in determined push point. During conveyance, surveyor robot motion is controlled to be in same direction of the subgoal and object. If the surveyor robot can push the object to pass the subgoal within 8 second, another robot is not required. However, if the surveyor robot can not push by itself, another robot will come to help it. We defined formation into the line formation that upcoming robots will be in the right or left side of surveyor robot.

V. Simulation Results

A simulation system is developed using python 2.7 with two-dimensional dynamical physic module. Simulation platform consists of an obstacle free workspace, an object, a set of sub-goals, and five mobile robots. The workspace size is 500x600 pixels of y and x axis. It has grid with 60x60 pixels of x-y axis for user friendly. The object is configured in regular or irregular which is placed in the left area of the workspace. The object with regular shape are rectangular, pentagonal, hexagonal, heptagonal, and octagonal shape. For the irregular object, combination of curved side, straight side, and corner are used. Also, five disk-shaped mobile robots with eleven distance sensors are located in right area of the workspace. The object is required to be moved on a straight line that consists of eight sub-goals.

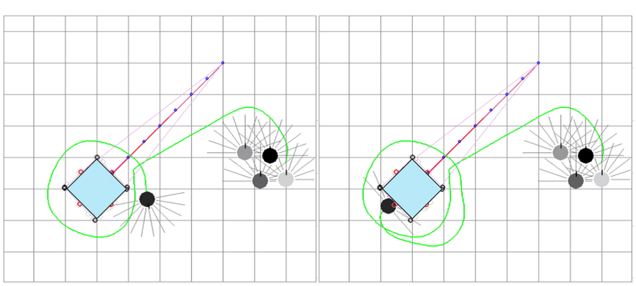
A. Regular polygon object

Approaching step is shown in Fig. 8(a) Robot departs from random start point to the object while avoiding any object. After arrived at 60% of distance sensor reading, robot stop and turn left to align the direction parallel to nearest object side. Then, robot moves one revolution to surround the object in clockwise as shown in Fig. 8(b). During surrounding, the robot measures the distance and convert it into a coordinate and record. The coordinates are illustrated using gray circle alongside the object. After the robot rotated the object and recorded the distance, collected data is processed using graham scan algorithm to make convex hull which consists of selected points. Such points consist of some unnecessary that have to be eliminated. Fig. 9(a) shows a rectangular object with surrounded convex hull points (gray circle). At the corners, there are four black circles which the points selected from convex hull. Then, red circles are push points that located between the corners. Fig. 9(b) shows that the robot is getting close to the selected push point as a push position. For rectangular object, push position can be determined satisfactory.



(a) Approaching (b) Surrounding

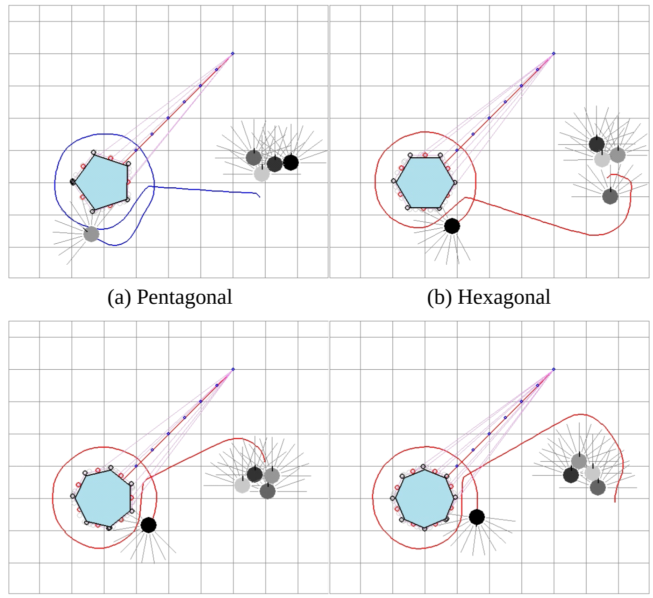
Fig. 8: Simulation result of rectangular object



(a) Creating convex hull (b) Get close to push point

Fig. 9: Simulation result of rectangular object

In the Fig. 10, pentagonal to octagonal object are used and also, good results are obtained.



(a) Pentagonal (b) Hexagonal

(c) Heptagonal (d) Octagonal

Fig. 10: Simulation results with different object shape

B. Irregular shaped object

Step from approaching to surrounding of irregular shape is similar to regular object. Surveyor robot departs from random start point to the object and arrived at specified distance, robot stop and turn left to align the direction parallel to nearest object side. Then, robot moves one revolution to surround the object in clockwise as shown in Fig. 11(b). During surrounding, robot measures the distance, convert it into a coordinate and record. The coordinates are illustrated using gray circle alongside the object.

The step of obtaining convex hull and push point are shown in Figs. 12(a) and 12(b), respectively. This irregular shaped object has three different sides. These are straight, curved and cornered sides. However, the algorithm can address this object and give the result as expected.

Another irregular shaped is also used to demonstrate the algorithm as shown in Fig. 13. In the figure 13(b), the step of obtaining push point where the surveyor robot needs to come for making an initial push is shown.

Next simulation result is the conveyance process with minimum robot number as the result of this algorithm.

C. Conveyancing an object

In this simulation, rectangle object is selected randomly. The algorithm step of approaching, surrounding, estimating

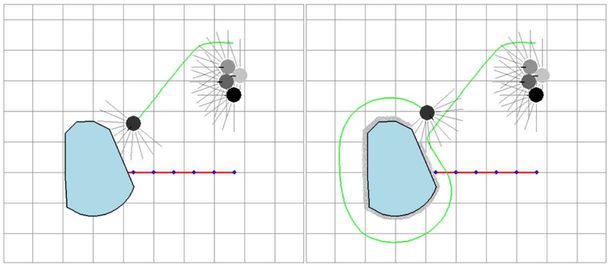


Fig. 11: Simulation result of irregular shape object

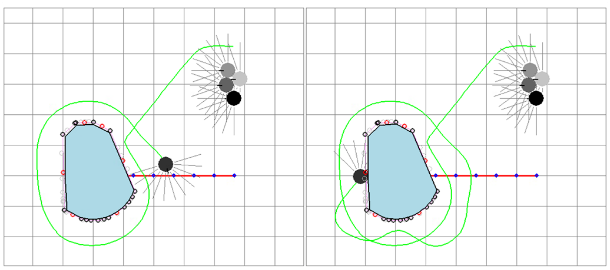


Fig. 12: Simulation result of irregular shape object

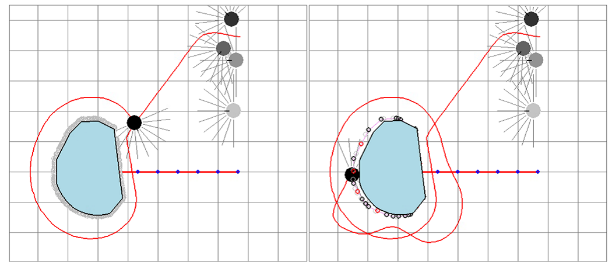


Fig. 13: Simulation result of another irregular shape object

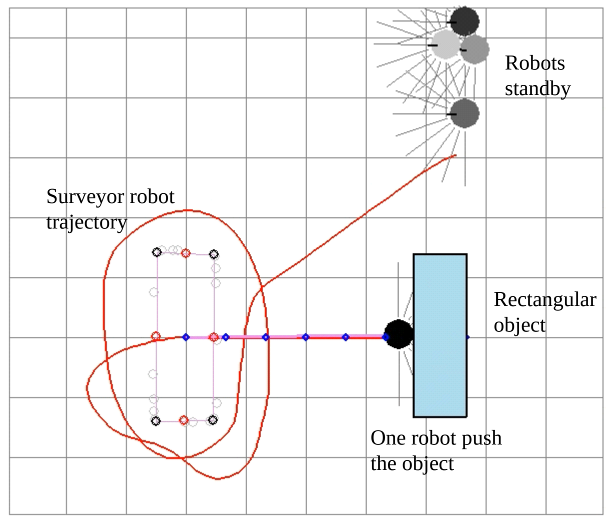


Fig. 14: Simulation result of one robot moves the object

and conveyancing object are shown in Fig. 14. This object conveyance is carried out by one robot which is surveyor robot because it can push the object in given distance within required time. As shown in Fig. 14, the desired trajectory consists of 8 sub goals which is indicated by blue circle on the red line. At that time, given mission includes a limitation of conveyance speed which the object are needed to be moved in 8 seconds for each subgoal. This value is known by considering simulation environment such as robot and object weight, friction, and pixel to actual distance ratio. Therefore, if the surveyor robot can not move the object within 8 seconds, the object is considered as heavy object.

In Fig. 15, the surveyor robot experiences difficult movement because the object is heavy as shown in Fig. 15. Therefore, the second robot comes and joins to support the first. Team member increment is decided by measuring movement time and elapsed distance. Next robot will join to the team if the performance less than the required setting value.

Object conveyance by dynamic team which consist of three robot is shown in Fig. 16. This team member are increased become three since the surveyor and second robot can not push the object within required time. Second and third robot joined to the support surveyor robot when the object can not achieve first subgoal within 8 seconds. The information of object movement is processed by simulation program and sent to the surveyor robot to make decision. The robot trajectories show the location where they start and come to approach the object.

VI. Conclusions

The algorithm presented has demonstrated its effectiveness in determining the number of the robot based on object shape and size for object conveyance with multiple mobile robot systems. It worked smoothly for both regular and irregular shaped object which carried out by the first robot as known as surveyor robot. The surveyor robot is able to determine the point to make initial push without knowing the information of object shape. This ability enables the robot to convey any kind object object. Moreover, the number of robot in the team is increased based on the object speed when it is moved. If the object is heavy and the surveyor robot can not move the object in given time, another robot will come to help.

It was verified in the simulation, that the each robot position is known, and the position can be obtained by using graphical positioning system for small indoor workspace to obtain precision position information. The challenge of the algorithm is the environment with lack of position information. The use of SLAM is important for this case.

References

[1] H. Suzuki, et, al. "Design and Implementation of Cooperative Conveyance Patterns for Multiple Mobile Robots Using Neural Network" Journal of Signal Processing, Vol. 11, No. 6, 2007

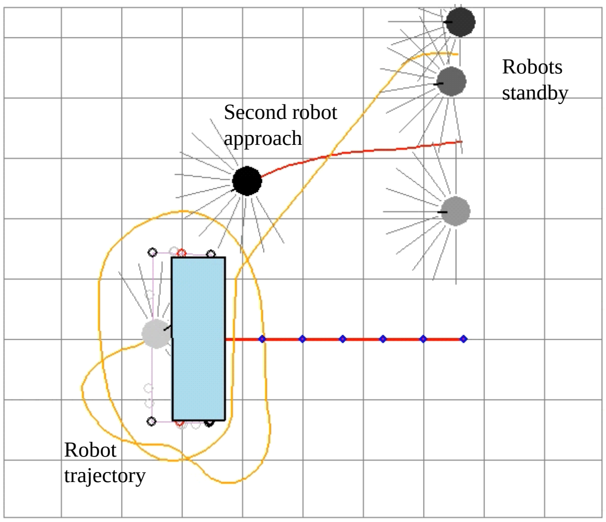


Fig. 15: Simulation result of second robot join

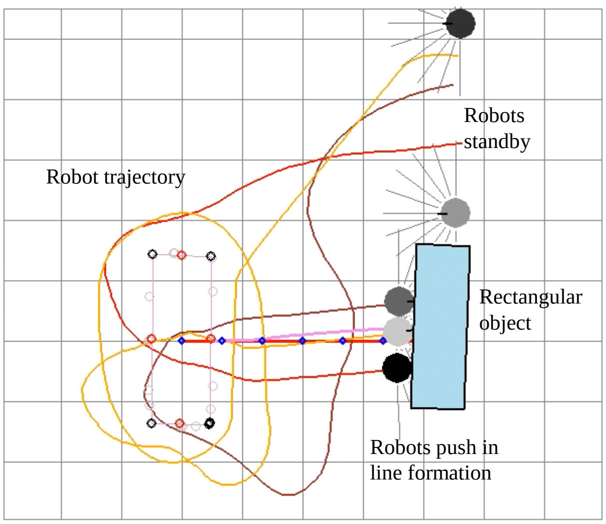


Fig. 16: Simulation results of three robots move the object

[2] H. Suzuki, et, al. "Cooperative Conveyance Control for Multiple-Mobile-Robot System Using Complex-Valued Neural Network and Indirect Cooperative Scheme" Journal of Signal Processing, Vol. 14, No. 4, 2010

[3] Jennings, J. and Kirkwood, C.W. "Distributed Mobile Robotics by the Method of Dynamic Teams" Distributed Autonomous Robotic System, pp 47-56, 1998

[4] Otte, M. and Correll, N. "Any-Com Multi-Robot Path-Planning with Dynamic Teams: Multi-Robot Coordination under Communication Constraints"

[5] Sejati, P. et, al. "Object Conveyance Control System By Multiple Mobile Robots With Dynamic Team" Proceedings of the 2015 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP'15), No. 28AM2-3-4, pp.118-121, Kuala Lumpur, Malaysia, February 28, 2015

[6] Wang, Z. and Kumar, V. "Object Closure and Manipulation by Multiple Cooperating Mobile Robots" Proceedings of the 6th International Symposium on Distributed Autonomous Robotic System, (2002)

[7] Bicchi, A. and Kumar, V. "Robotic Grasping and Contact" ICRA2000, pp.348-353, 2000 The 12th International Symposium On experimental Robotics, pp 743-753, 2014

[8] Sejati, P. et, al. "Push Position Estimation for Object Conveyance Problem Using Multiple Mobile Robots System" Academic Lecture of Society of Instrument and Control Engineer Shikoku Section, No.SO2-20, PS2-20, pp.172-175, 2014

[9] Petrovic, V and Ivetic, D. "Linearization of Graham's Scan Algorithm Complexity" The 5th International Conference on Information Technology, 2011

[10] T. Eduardo, et, al."Analysis and classification of multiple robot coordination methods" In Proceedings of ICRA'00, pages 3158-3163, San Francisco, CA, USA, April 2000.

[11] K. Shin, at, al. "Coordinating mobile robots by applying traffic rules" Proceedings of IROS'92, pages 1535-1541, Raleigh, NC, USA, July 1992.

[12] Y. Zhi, et, al. "A Survey and Analysis of Multi-Robot Coordination" International Journal of Advanced Robotic Systems, 2013, 10:399. doi: 10.5772/57313

[13] N. Michael A, et, al. "Object Manipulation through Explicit Force Control Using Cooperative Mobile Multi-Robot Systems" World Congress on Engineering and Computer Science, Vol. I, pp. 364-369, 2014

[14] V. Manuel Alejandro Molina, et, al. "Fuzzy logic controller for cooperative mobile robotics implemented in leader-follower formation approach" Revista Facultad de Ingeniera, Universidad de Antioquia, No. 76, pp. 19-29, 2015

[15] Y. Wentao, at, al. "A Cooperative Path Planning Algorithm for a Multiple Mobile Robot System in a Dynamic Environment " International Journal of Advanced Robotic Systems, 2014, 11:136. doi: 10.5772/58832

On the Use of Arabic Tweets to Predict Stock Market Changes in the Arab World

Khalid AlKhatib, Abdullateef Rabab'ah, Mahmoud Al-Ayyoub and Yaser Jararweh
Jordan University of Science and Technology
Irbid, Jordan

***Abstract*—Social media users nowadays express their opinions and feelings about many event occurring in their lives. For certain users, some of the most important events are the ones related to the financial markets. An interesting research field emerged over the past decade to study the possible relationship between the fluctuation in the financial markets and the online social media. In this research we present a comprehensive study to identify the relation between Arabic financial-related tweets and the change in stock markets using a set of the most active Arab stock indices. The results show that there is a Granger Causality relation between the volume and sentiment of Arabic tweets and the change in some of the stock markets.**

Keywords—Twitter; Sentiment Analysis; Granger Causality; Pearson Correlation; Arab Stock Market

I. Introduction

Social media provides us with a valuable source of the opinions of the public about different topics from different domains. One of the most interesting domains is the financial domain. Twitter nowadays is used as a platform enabling its users to read and write a large number of messages called tweets. The setting of tweets is public, which allows the researchers to fetch the tweets and perform their research using these data [1, 2].

The financial domain has special interesting characteristics since it most likely deals with money. Stock market prices and stock change prediction are the concern of many social media users who are interested in the financial market.

Several researchers are studying the relationship between tweets and financial markets. They focus on tweeting activity as well as tweets' contents (mainly tweets written in English). The task of studying the relationship between the sentiments conveyed in tweets (written in English) and stock market change is a very interesting one with many obvious applications [3, 4].

The same task can be much more difficult when the language under consideration is other than English. The reason behind this is very simple. The above task requires a mature set of tools and resources to efficiently analyze and accurately extract the sentiments conveyed in a large number of tweets, in what is known as Sentiment Analysis (SA). The field of English SA is the one that most likely fit this description. The same cannot be said about performing SA on other languages. Arabic, for example, is a very important language, but the field of Arabic SA is still at its early stages, which means that the tools and resources for Arabic SA are still not deep enough and are not yet tested thoroughly.

Another issue specific to studying the problem at hand outside the developed countries is the lack of proper data. E.g., for the Arab world, there is no commonly used stock market indices such as Dow Jones Industrial Average (DJIA). Instead of having a common stock index that is used across the Arab world (which can be very useful to conduct a study like ours), each country has its own index.

In this research, we study the relationship between the sentiments conveyed in Arabic tweets and the changes in stock market prices. The aim is to try to see whether there are causal relationships or not. This allows us to answer the question of whether we can use Twitter Arabic content to predict changes in the stock markets across the Arab world or not. To the best of our knowledge, there exist only one published paper addressing this problem for this specific part of the world [5]. However, that paper is restricted to only one country, Saudi Arabia, whereas, our work is not limited to any specific Arab country. Also, [5] has many shortcomings. For example, the kind of analysis the authors performed is rather simplistic and the dataset they used is very small.

The rest of this paper is organized as follows; Section II presents the related work. Section III presents methodology used. Section IV presents and discusses the results of this research. Section V presents a conclusion of our work.

II. Background and Related Works

In this section we present a set of related works that investigate the relationship between Twitter data and financial market. Before presenting the related works, we present a brief introduction about the tools and methods we employ to achieve our goal which include Sentiment Analysis, Granger Causality, and Pearson Correlation.

A. Background

Sentiment Analysis (SA) is a type of natural language processing (NLP) problems that is concerned with extracting the sentiment conveyed in a piece of text. I.e., SA is concerned with showing whether the author of the text feels positively or negatively about certain topics [6, 7, 8]. SA for the Arabic language is challenging for many reasons such as the triglossic nature of the language, the limited literature on Arabic NLP, and the scarcity of dependable resources including publicly available datasets that are collected and annotated for SA purpose, sentiment lexicons that support the Arabic Language, etc. [9, 10].

Granger Causality studies the relation between two time series in term of cause and effect. It uses the assumption that any event occurs at certain time must have had its causes occurring at an earlier time. Given two time series A and B, if A Granger causing B, then we can use the previous values of A to predict the value of B [11, 12].

Pearson Correlation measures the linear dependency between two variables by measuring the strength and direction of the relationship between them. The strength of the correlation is in the range [-1, 1] and the direction states whether the correlation is positive or negative [13].

B. Related Works

As mentioned earlier, using Twitter data to predict stock market changes is a relatively new field. Ranco et al. [14] studied the relation between Twitter users' posts and financial market. They studied the sentiment of tweets regarding the 30 stocks companies that form the Dow Jones Industrial Average (DJIA) index in a period of 15 months. Their results showed that the Pearson correlation and the Granger causality between the tweets' sentiments and the financial data time series are relatively low. Moreover, their results showed a significant dependency between the tweets' sentiment and abnormal returns, where the sentiment polarity of the tweets implies the direction of the abnormal returns.

Zang and Skiena [15] performed a comprehensive study to show the relation between the sentiment presented in companies' related news and the companies' stocks volumes and financial returns. Their results showed a significant correlation between the news sentiment and stock market indicators.

Another study conducted by Souza et al. [16] presented a case study for the relation between the sentiments conveyed in tweets about a set of retails companies and the stock return for these companies. Their results showed that the tweets' sentiments have statistically significant relation with the stock returns and there is a strong Granger causality between the number of tweets and the stock returns.

Bollen, Mao, and Zeng [3] studied the relation between the mood states derived from Twitter posts and Dow Jones Industrial Average (DJIA) values. They analyzed the content of the tweets by feeding text into two sentiment tools, OpinionFinder (which gives positive and negative moods) and Google-Profile of Mood States (GPOMS) (which gives mood classification into six classes: Calm, Alert, Sure, Vital, Kind, and Happy). These mood states are used to study the relation with the stock prices changes. The results showed a significant Granger causality between the sentiment moods and the values of DIJA. Moreover, they used a Self-Organizing Fuzzy Neural Network to predict the values of DIJA using the sentiment moods. A similar study was conducted by Mittal and Goel [4] but with different content. They used only four moods of sentiment classes (Calm, Happy, Alert, and Kind).

Zheuldev et al. [17] studied the effect of the sentiments presented in Twitter content of the future stock prices of the S&P 500 index. They studied both tweets volume and sentiment presented in tweets at hourly resolution, and related them to hourly price returns of 28 financial companies collected over a period of three months. Their results showed that the sentiment presented in tweets is more statistically significant in leading financial market than message volumes.

Mao et al. [18] study the relation between the daily numbers of tweets that mention the S&P 500 stocks and the closing stock prices. Their results showed that the daily number of tweets is significantly correlated with the daily closing stock prices.

Finally, one of the works that are most related to ours is that of AL-Rubaiee et al. [5]. However, this paper is restricted to only one country, Saudi Arabia, whereas, our work is not limited to any specific Arab country. Also, [5] has many shortcomings. For example, the kind of analysis the authors performed is rather simplistic (one-to-one model) whereas we perform causality analysis and statistical correlation. Finally, the dataset used in [5] is very small (less than 2K tweets, most of which are used for training) compared to our dataset of about 1.5M tweets. However, the last point is justified by the fact that the authors of [5] relied on manual annotation whereas we rely on automatic lexicon-based tool that required no training from our side.

III. Methodology

This sections discusses the details of the methodology used in this work. The first step is to obtain both the Twitter data and the financial data. Then, the tweets sentiment and volume are computed. Next we run Granger causality test to identify the causality relationship between the tweets and the financial market. After finding the causality relationship, we run Pearson correlation to determine the type and strength of the correlations. The following sections present the details of these steps.

A. Obtaining Twitter Data

The Twitter data are obtained using the tweepy API[1] by passing a set of keywords identified to be related to the financial market. We get this list of keywords from a financial expert who is aware of the social media and very familiar with Twitter posts. Table I presents the list of keywords used in Arabic along with their English translations. The tweets are collected starting from March 6 2016 till April 16 2016. The total number of collected tweets is 1,500,223. These tweets are collected from all days in the period of interest. As suggested in the literature, it is interesting in a problem like ours to consider the working days only. So, we filter out the tweets posted on weekends and we are left with 1,137,543 tweets. These numbers are large enough for our results to be trustworthy.

B. Obtaining Financial Data

The financial data are obtained from the most active Arabic Stock Market Indices.[2] Specifically, we focus on the following indices: Saudi Arabia (TASI), Abu Dhabi (ADI), Qatar (QSI), Dubai (DFMGI), Kuwait (KWSE), and Egypt (EGX30). For each of these indices we obtain the historical data in a daily manner. The historical data include information for the open, close, high, low, volume, and return values for a given day.

[1] http://www.tweepy.org/
[2] http://sa.investing.com/indices/ae-indices

TABLE I: Arabic financial key words list

English Translation	Arabic Keyword
Price	سعر
prices	أسعار
stocks	أسهم
monetary	نقدي
financial	مالية
market	سوق
stock exchange	بورصة
earnings	أرباح

C. Data Processing

In order to perform the causality testing, we need first to extract twitter data in daily manner and compute the volume and sentiment of the tweets. We then correlate tweets volume and sentiment to stock market returns. The stock market returns are available through the Arabic Stock Market Indices historical data under the name "التغير" (change) which can be computed based on the Daily Stock market Return (DSR) as shown in Equation (1).

$$DSR_d = \frac{C_d - C_{d-1}}{C_{d-1}} \quad (1)$$

where C_d is the closing price for a given stock at day d.

The stock market returns are available for working days only. In this research we handle the issue of weekend in two different ways. The first one is to completely ignore them. This requires excluding tweets posted during the weekends. The second way is to include weekends in the study. This requires an additional processing step to estimate/approximate the stock returns values for the weekend days. We follow the approximation process adopted by [14] which defined the Approximated Stock Returns (ASR) as shown in Equation (2).

$$ASR_d = \frac{DSR_{d+1} - DSR_{d-1}}{2} \quad (2)$$

This way of approximation simply takes the average of two days, one before and one after that day.

Figure 1 presents the Twitter data volume and sentiment daily distributions. The column bars in the figure are for the weekend days. The figure shows that the average numbers of total, positive, negative and neutral tweets are 37,506, 4,240, 7,106 and 26,159, respectively.

Figure 2 shows the stock market returns of the Arabic Stock Market Indices in daily manner. Figure 3 shows the same distribution but with an approximation of the weekend stock markets return values.

As shown in Figures 1, 2, and 3, one can see a clear relation between the daily distribution of the tweets and the stock return values for the Egypt stock market index (EGX30). The neutral tweets count and total tweets shows a higher degree of correlation compared with the positive and negative tweets counts. The correlation between Twitter data and the stock return values for EGX30 is positive which means that, as the volumes of tweets increases, the stock return values increase. Other stocks market indices show a negative correlation where

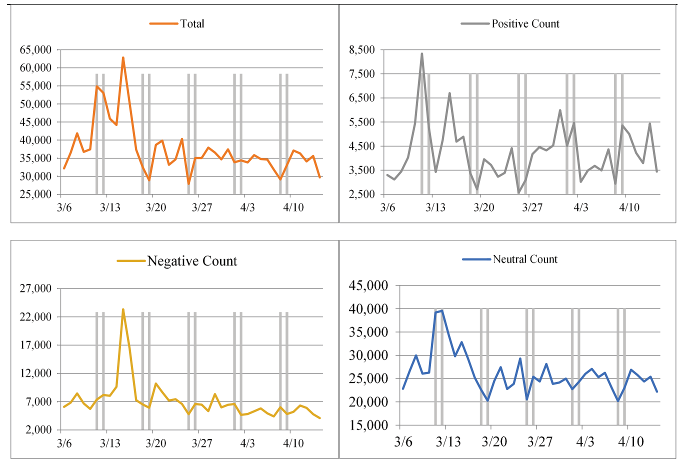


Fig. 1: Tweets distribution.

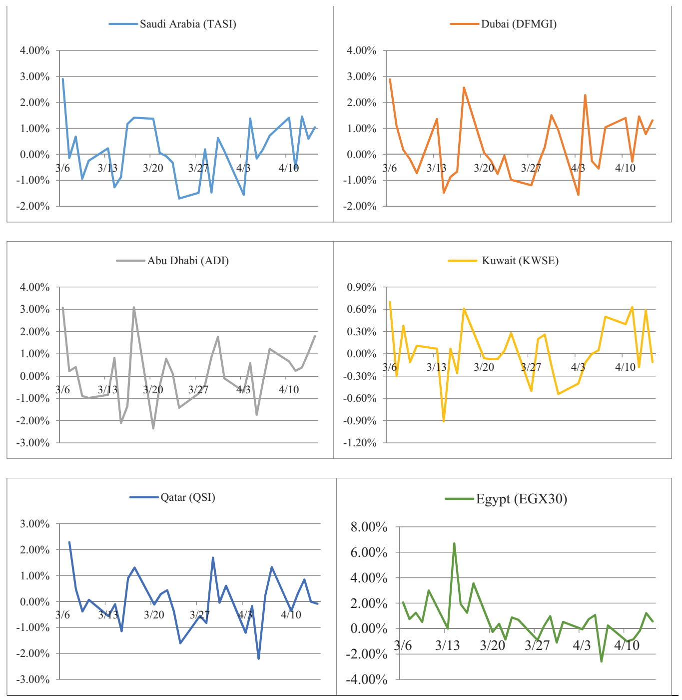


Fig. 2: Stock Markets Returns.

the movement of the stock return values opposes the movement of tweets volumes. The exact correlation values are presented later when discussing the result of Pearson correlation.

D. Applying SentiStrength

The well-known SentiStrength [19] tool for SA is used to detect and measure the strength of the sentiments expressed in the tweets. SentiStrength was originally developed for English and was later adapted for other languages including the Arabic [20]. To use SentiStrength for the tweets we collected, the default SentiStrength data files must be modified for the Arabic language.[3] The output of SentiStrength is a sentiment score

[3]http://sentistrength.wlv.ac.uk/SentStrength_ Data/arabic/

Fig. 3: Stock Markets Returns with weekend approximations.

that falls in one of two ranges: in the range [1, 5] for positive sentiment and in the range [-5, -1] for negative sentiment. The overall score is positive if the positive scores are greater than the negative ones. Similarly, the overall scoare is negative if the negative scores are greater than the positive ones. When the positive and negative scores are equals, the overall score is neutral.

The researchers in [19, 20] conducted a comparative study between different SA tools that support Arabic. Their results showed that SentiStrength is the best tool that can be used to measure the sentiment presented in Arabic text.

E. Computing tweets sentiments and volumes

To compute sentiments presented in tweets, we use the SentiStrength tool [19]. After computing the sentiment scores, we perform a set of statistical operations that result with the measures used in this research to perform the causality and correlation analysis. The measures values are computed in daily manner. Following are the list of measures used in this research: Total Tweets, Positive Tweets Count, Negative Tweets Count, Neutral Tweets Count, Net Values, and Sentiment Polarity.

The Sentiment Polarity measure was proposed by [14]. It is defined as Equation (3).

$$SentimentPolarity = \frac{Positive\ tweets\ count - Negative\ tweets\ count}{Positive\ tweets\ count\ + Negative\ tweets\ count} \quad (3)$$

The Net Value (NV) measure is defined as shown in Equation (4).

$$Net\ Value = \sum Positive\ tweets\ values + \sum Negative\ tweets\ values \quad (4)$$

F. Applying Granger Causality

Granger causality is applied separately on the two sets of financial data we have (the ones without weekend values and the ones with weekend values approximation). Remember that when applying causality analysis on the data including the weekends, we need to approximate the stock markets values for the weekend days since the stock markets values are not available for the weekends. We perform the approximation based on the approach adopted by [14, 4] using the average of the two days: before and after the weekends.

G. Computing Pearson Correlation

After identifying causality relationships between sentiment tweets and volumes and the stock return values, we run the Pearson correlation test to determine the direction and strength of the correlations between the sentiment measure used and the daily stock return values.

IV. Results and Discussion

In this section, we present and analyze the results of the Granger causality testing process between the sentiment measures we compute from the tweets we collect and the financial data. This section also discusses the results of applying the Pearson Correlation test.

The result of the Granger causality test between two variables X and Y is presented as a hypothesis that the probability that variable X does not Granger cause variable Y. Higher probability values represent higher significance. When the probability value is less than a certain threshold such as 0.05 or 0.1, the hypothesis is rejected and the opposite hypothesis becomes true. The Granger causality test presents the results for a specific Lag which indicates the number of passing days included in the causality testing process.

In the following, we discuss the results depicted as tables. In these tables, the colored cells in the tables refer to the significance level used; yellow color is used for significance level 0.1 and green color is used for significance level 0.05.

Table II presents the results of Granger causality test between tweets sentiment and volume and stock change values for the Arab stock market indices used in this research. The table shows that for the Saudi Arabia (TASI) index, there are Granger causality at significance level 0.1 for the measures negative count at lag 5, neutral count at lags 3 and 2, and net value at lag 5, and Granger causality with the total number of tweets at significance level 0.05 at lag 5. For the Abu Dhabi (ADI) index, there are Granger causality at significance level 0.1 for the measure positive count at lag 4, and Granger causality at significance level 0.05 with the total positive count at lags 2 and 3. For the Qatar (QSI) index there are Granger causality at significance level 0.1 for the measures positive count at lags 2 and 3. For the Dubai (DFMGI) index there are Granger causality at significance level 0.1 for the measures positive and total counts at lag 2, neutral counts at lags 1, and negative counts at lag 5. For the Egypt (EGX30) index there are Granger causality at significance level 0.1 for the measures positive counts at lag 4, neutral counts at lags 3 and 2, and net values at lag 1, and Granger causality with the total number of tweets, sentiment polarity and neutral count at significance level 0.05 at lag 1. The results show that the Kuwait (KWSE) index was not useful in studying the causality relationship.

Table III presents the Pearson correlation results for the measures that pass the Granger causality test. The table shows a negative correlation between the measures and the stocks indices for most of the indices. The only exception is the Egypt index which has a positive correlation with the measures.

Table IV presents the result of Granger causality test between tweets sentiment and volume and stock change values for the Arab stock market indices used in this research while including weekends and holidays. The table shows that for the Abu Dhabi (ADI) index there is Granger causality at significance level 0.05 for the measure negative count at lags 3, 4 and 5. For the Qatar (QSI) index there is Granger causality at significance level 0.1 for the measure negative count at lag 5. For the Dubai (DFMGI) index there is Granger causality at significance level 0.1 for the measure positive count at lag 2. For the Egypt (EGX30) index there is Granger causality at significance level 0.1 for the total measure at lag 3, and Granger causality with the total number of tweets at lags 1, 2, and 4, positive count at lags 4 and 5 and neutral at all lags at significance level 0.05 at lag 1. The result shows that the Kuwait (KWSE) and the Saudi Arabia (TASI) indices were not useful in studying the causality relationship.

Table V presents the Pearson correlation results for the measures that pass the Granger Causality test. The table shows a negative correlation between the measures and the stocks indices for most of the indices except for Egypt index which has a positive correlation with the measures.

Our results conform with those of [14] where it was shown that the sentiment polarity is not able to capture the causality relation of all indices. Our results also conform with those of [15, 16] where it was shown that the tweets sentiment and volume affect the stock prices change. The same can be said for [3, 4] even if they used different sentiment behaviors that are based on the emotion presented in tweets content. Finally, our results conform with some of the results presented in [17, 18] .

V. Conclusion

In this work, we study the relation between Twitter financial data and the stock prices change of a set of Arab stock market indices. The study is conducted with two sets of financial data. The first one includes only tweets for working days in studying the causality relationship between the tweets sentiment and volume and the stock prices changes. The result of this phase shows Granger causality relationship between the tweets sentiment and volume and stock change for most of the indices used in this research. Including weekends and holidays in the study has a negative effect on the causality relationship. The results of this study show that the best stock market index the can be used to study the relation between the twitter sentiment and volume and stock market change is the Egyptian index (EGX30).

References

[1] H. Kwak, C. Lee, H. Park, and S. Moon, "What is twitter, a social network or a news media?" in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 591–600.

[2] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment," *ICWSM*, vol. 10, pp. 178–185, 2010.

[3] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[4] A. Mittal and A. Goel, "Stock prediction using twitter sentiment analysis," Stanford University, Tech. Rep., 2011, http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf.

[5] H. AL-Rubaiee, R. Qiu, and D. Li, "Analysis of the relationship between saudi twitter posts and the saudi stock market," in *2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS)*. IEEE, 2015, pp. 660–665.

[6] E. Boiy, P. Hens, K. Deschacht, and M.-F. Moens, "Automatic sentiment analysis in on-line text." in *ELPUB*, 2007, pp. 349–360.

[7] G. Vinodhini and R. Chandrasekaran, "Sentiment analysis and opinion mining: a survey," *International Journal*, vol. 2, no. 6, 2012.

[8] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and trends in information retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.

[9] N. Abdulla, N. Mahyoub, M. Shehab, and M. Al-Ayyoub, "Arabic sentiment analysis: Corpus-based and lexicon-based," in *Proceedings of The IEEE conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, 2013.

[10] M. Al-Ayyoub, S. B. Essa, and I. Alsmadi, "Lexicon-based sentiment analysis of arabic tweets," *International Journal of Social Network Mining*, vol. 2, no. 2, pp. 101–114, 2015.

[11] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[12] R. Goebel, A. Roebroeck, D.-S. Kim, and E. Formisano, "Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping," *Magnetic resonance imaging*, vol. 21, no. 10, pp. 1251–1261, 2003.

[13] S.-D. Bolboaca and L. Jäntschi, "Pearson versus spearman, kendalls tau correlation analysis on structure-activity relationships of biologic active compounds," *Leonardo Journal of Sciences*, vol. 5, no. 9, pp. 179–200, 2006.

[14] G. Ranco, D. Aleksovski, G. Caldarelli, M. Grčar, and I. Mozetič, "The effects of twitter sentiment on stock price returns," *PloS one*, vol. 10, no. 9, p. e0138441, 2015.

[15] W. Zhang and S. Skiena, "Trading strategies to exploit blog and news sentiment." in *ICWSM*, 2010.

[16] T. T. P. Souza, O. Kolchyna, P. C. Treleaven, and T. Aste, "Twitter sentiment analysis applied to finance: A case study in the retail industry," *arXiv preprint arXiv:1507.00784*, 2015.

[17] I. Zheludev, R. Smith, and T. Aste, "When can social media lead financial markets?" *Scientific reports*, vol. 4,

TABLE II: Granger causality test results for working days only

Saudi Arabia (TASI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.5809	0.3301	0.3021	0.8287	0.975	0.6008
2	0.1127	0.3261	0.085	0.4144	0.9865	0.6732
3	0.135	0.5039	0.0698	0.2003	0.9763	0.8669
4	0.5187	0.151	0.2458	0.1428	0.6037	0.1645
5	0.7467	0.0798	0.2244	0.0241	0.4422	0.0564
Abu Dhabi (ADI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.431	0.6995	0.538	0.5137	0.8234	0.4968
2	0.0563	0.6689	0.3791	0.6441	0.2171	0.7103
3	0.0292	0.1436	0.5917	0.3664	0.315	0.4522
4	0.0624	0.2876	0.3937	0.3259	0.5636	0.7357
5	0.1771	0.2423	0.2388	0.1379	0.1868	0.4058
Qatar (QSI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.6396	0.1751	0.3283	0.1775	0.5113	0.2701
2	0.0741	0.282	0.6957	0.3457	0.8325	0.5045
3	0.091	0.2903	0.436	0.517	0.8877	0.5568
4	0.1695	0.3909	0.4817	0.6735	0.9283	0.7933
5	0.1693	0.6712	0.3298	0.8337	0.7512	0.8489
Dubai (DFMGI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.7258	0.6272	0.0702	0.2414	0.3972	0.4936
2	0.0953	0.225	0.1974	0.0799	0.5557	0.5483
3	0.2271	0.2318	0.3309	0.1551	0.5134	0.7203
4	0.2293	0.4063	0.3789	0.3545	0.7475	0.7873
5	0.4188	0.0741	0.3737	0.0244	0.1603	0.1179
Kuwait (KWSE)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.64	0.646	0.1102	0.286	0.5258	0.6536
2	0.8352	0.5329	0.2439	0.3362	0.448	0.4322
3	0.9307	0.5709	0.2648	0.5236	0.4716	0.5369
4	0.8623	0.7423	0.2625	0.5043	0.6726	0.6034
5	0.8436	0.9067	0.3663	0.7105	0.5943	0.7601
Egypt (EGX30)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.4404	0.1471	0.0025	0.0311	0.0343	0.0664
2	0.1005	0.6348	0.0168	0.3095	0.4514	0.6683
3	0.4121	0.5225	0.0687	0.7357	0.5886	0.678
4	0.0639	0.6667	0.1714	0.7954	0.7582	0.8383
5	0.1086	0.8966	0.251	0.9269	0.7048	0.7778

TABLE III: Pearson Correlation result

	Saudi Arabia (TASI)	Abu Dhabi (ADI)	Qatar (QSI)	Dubai (DFMGI)	Egypt (EGX30)
Positive Count	-0.031256	-0.074579	-0.048367	-0.02699	0.340172
Negative Count	-0.107913	-0.444323	-0.048833	-0.297334	0.288701
Neutral Count	-0.15561	-0.425684	-0.291694	-0.192645	0.592135
TOTAL	-0.140212	-0.472846	-0.16795	-0.271661	0.546638
Sentiment Polarity	0.14851	0.419403	-0.039095	0.309059	-0.115517
Net Value	0.086641	0.453577	0.053068	0.319605	-0.221203

2014.

[18] Y. Mao, W. Wei, B. Wang, and B. Liu, “Correlating s&p 500 stocks with twitter data,” in *Proceedings of the first ACM international workshop on hot topics on interdisciplinary social networks research*. ACM, 2012, pp. 69–72.

[19] M. Thelwall, “Heart and soul: Sentiment strength detection in the social web with sentistrength,” *Proceedings of the CyberEmotions*, pp. 1–14, 2013.

[20] N. Alsaedi and P. Burnap, “Arabic event detection in social media,” in *Computational Linguistics and Intelligent Text Processing*. Springer, 2015, pp. 384–401.

TABLE IV: Granger causality test results with weekend approximations

lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
Saudi Arabia (TASI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.5648	0.2013	0.6536	0.6721	0.7662	0.4717
2	0.5479	0.3423	0.7309	0.7262	0.6698	0.6345
3	0.7473	0.5494	0.8883	0.9063	0.8081	0.6538
4	0.8981	0.6363	0.1228	0.4114	0.9116	0.7351
5	0.7642	0.7068	0.2158	0.5726	0.6548	0.7878
Abu Dhabi (ADI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.9831	0.3942	0.9859	0.8046	0.9755	0.983
2	0.2892	0.3112	0.9405	0.5473	0.7509	0.5612
3	0.277	0.039	0.7205	0.3419	0.856	0.1486
4	0.2909	0.0416	0.8535	0.1602	0.7326	0.2597
5	0.4527	0.0348	0.8954	0.2086	0.8218	0.2007
Qatar (QSI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.8448	0.3942	0.7461	0.3579	0.5374	0.2437
2	0.4864	0.1825	0.8866	0.4733	0.8535	0.3612
3	0.1856	0.2057	0.9607	0.6391	0.5614	0.5579
4	0.3063	0.3773	0.9839	0.6668	0.6707	0.5885
5	0.094	0.6188	0.1393	0.2452	0.5698	0.798
Dubai (DFMGI)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.5082	0.7766	0.955	0.7831	0.5973	0.9082
2	0.1079	0.3605	0.9824	0.5264	0.4555	0.901
3	0.1902	0.5589	0.8266	0.5014	0.7147	0.9978
4	0.332	0.7906	0.6065	0.6072	0.7479	0.9998
5	0.5431	0.7855	0.7005	0.7018	0.6729	0.9977
Kuwait (KWSE)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.9864	0.8966	0.6531	0.741	0.7029	0.8449
2	0.8648	0.3672	0.7289	0.989	0.5013	0.3611
3	0.1137	0.6902	0.753	0.8328	0.3319	0.7305
4	0.1599	0.7765	0.4091	0.8204	0.2668	0.6028
5	0.1358	0.9611	0.3052	0.8404	0.2093	0.8048
Egypt (EGX30)						
lag	Positive Count	Negative Count	Neutral Count	Total	Sentiment Polarity	Net Value
1	0.4513	0.5382	0.002	0.0259	0.9434	0.4771
2	0.1121	0.2244	0.0032	0.0202	0.8666	0.39
3	0.1256	0.223	0.005	0.0691	0.641	0.5055
4	0.0336	0.3212	0.0063	0.0435	0.4996	0.7524
5	0.0378	0.6021	0.0305	0.141	0.5713	0.9512

TABLE V: Pearson Correlation result

	Abu Dhabi (ADI)	Qatar (QSI)	Dubai (DFMGI)	Egypt (EGX30)
Positive Count	-0.12238	-0.095965	-0.033781	0.340172
Negative Count	-0.377839	-0.028427	-0.238524	0.288701
Neutral Count	-0.339126	-0.222604	-0.105619	0.592135
TOTAL	-0.401113	-0.162766	-0.180936	0.546638
Sentiment Polarity	0.267611	-0.095124	0.183292	-0.115517
Net Value	0.364437	0.014813	0.239506	-0.221203

Optical Character Recognition System for Urdu Words in Nastaliq Font

Safia Shabbir* and Imran Siddiqi*
*Bahria University, Islamabad, Pakistan

***Abstract*—Optical Character Recognition (OCR) has been an attractive research area for the last three decades and mature OCR systems reporting near to 100% recognition rates are available for many scripts/languages today. Despite these developments, research on recognition of text in many languages is still in its early days, Urdu being one of them. The limited existing literature on Urdu OCR is either limited to isolated characters or considers limited vocabularies in fixed font sizes. This research presents a segmentation free and size invariant technique for recognition of Urdu words in Nastaliq font using ligatures as units of recognition. Ligatures, separated into primary ligatures and diacritics, are recognized using right-to-left HMMs. Diacritics are then associated with the main body using position information and the resulting ligatures are validated using a dictionary. The system evaluated on Urdu words realized promising recognition rates at ligature and word levels.**

Keywords—Optical Character Recognition; Urdu Text; Ligatures; Hidden Markov Models; Clustering

I. Introduction

Today, most of the information is available in digital form and can be accessed within a span of few clicks. This has resulted in a tendency to digitize the existing paper documents like books, newspapers, official notes etc. and make them available online. A number of libraries across the world have also scanned their books providing online access to the readers. These scanned versions of printed documents can be consulted more easily and efficiently as opposed to the paper format. However, mere scanning of huge amounts of printed papers is not sufficient. The job is only half done if these scanned documents are not searchable. Manual transcription of these huge collections of paper documents is naturally a tiresome, time consuming and very inefficient solution. This attracted researchers to develop automatic Optical Character Recognition systems which take scanned document images as input, apply image processing and pattern classification techniques and convert the image into text which is not only searchable and editable but also requires significantly lesser storage space as opposed to their image format.

Formally, Optical Character Recognition (OCR) is defined as a technique in which scanned image of (handwritten or printed) document is processed through machine, characters are recognized and extracted and rendered in a word editor [1], [2].

The most useful application of an OCR system is the concept of digital libraries where huge collections of books could be converted to text and made available online. Other typical applications include automatic processing of bank cheques, computerized validation of identification documents (passports, driving licenses etc.), processing of utility bills, text to speech applications, vehicle registration number recognition and text guided navigation of autonomous vehicles or visually impaired individuals.

Research in Optical Character Recognition is as old as the computerized document analysis and recognition itself. The ultimate objective of most of the document recognition problems is indeed a comprehensive OCR system which could eventually convert the huge collections of existing paper documents with all sorts of variations to digital form. The last few decades have witnessed extensive research on Optical Character Recognition systems for the Roman script. Today, commercial OCR software are available reporting near to 100% recognition rates for languages based on the Roman script. Research on Chinese and Arabic OCRs is also quite mature with acceptably good recognition rates. Few Multilingual OCRs have also been developed with the aim to propose techniques which are general in nature [3]. Despite these tremendous developments, OCRs for many languages around the globe are either non-existent or are witnessing early days of research, Urdu being one of them and makes the subject of our study. Research on Urdu and similar cursive scripts like Pashto, Farsi etc. is still in its early days with limited literature available till date.

This paper presents a segmentation free optical character recognition system for printed Urdu text in Nastaliq font. The proposed methodology relies on ligatures as units of recognition and is based on a semi-automatic clustering scheme which extracts ligatures from a given data set and clusters different instances of the same ligature into classes. Recognition is carried out using Hidden Markov Models (HMM) where a separate HMM is trained for each ligature (cluster). The recognition of ligatures is first carried out without dots which are later associated with the ligatures to recognize the complete word. Unlike most of the traditional approaches which either work on isolated Urdu characters or a fixed font size, the proposed technique works on complete Urdu words and is scale invariant. The sequential clustering employed for generation of training data makes the framework scalable which allows extension of the system to consider an increased number of ligatures.

The next section discusses the recognition techniques proposed for Urdu and similar scripts in the recent years along with a comparative analysis of these methods. Section III details the proposed methodology including training and recognition of Urdu words. Experimental evaluations carried out to validate the proposed methodology are presented in Section IV while Section V concludes the paper with a discussion on future research directions on this subject.

II. Background

This section presents an overview of notable contributions towards the development of an OCR for cursive scripts like Arabic, Farsi and Urdu etc. OCR systems for such scripts generally follow one of the two approaches: segmentation-based or segmentation-free. Another categorization of these OCRs is based on units of text used for recognition. Some of the techniques work on isolated characters only [2], [4], [5], [6] while others work on complete words or ligatures [7], [8], [9], [10], [11]. The systems developed to work with isolated characters naturally report much better results as opposed to those working on words or ligatures. Since the segmentation of Urdu and Urdu-like text images into its basic units (words or characters) is itself a challenging task, a significant research dedicated to segmentation of text has also been reported in the literature [12]. With the increase in the usage of tablets and other hand held devices, a new categorization of OCRs as offline or online OCR has also emerged. Online OCRs recognize the text on the fly as it is input by a user while offline OCRs work on the scanned images of text. Online OCRs have the advantage of having additional information on the sequence of strokes while offline OCRs have only the shape information in the form of text pixels making them more challenging.

Segmentation-based approaches for OCR work on individual characters which are extracted by segmenting the text into ligatures and further into characters. The main advantage of these approaches is that the number of classes to be recognized is the same as number of characters (and their different shapes). This number is much smaller when compared to the number of ligatures or words which are units of recognition in segmentation-free approaches. The segmentation of text into characters, however, is a complex and challenging problem in itslef [13], [7], [14].

Among well-known segmentation-based approaches, morphological processing followed by an analysis of contours is applied for character extraction in [15]. Chain codes computed from the character contours are employed as features while classification is carried out through feedback loop. In another study [16], moment invariant descriptors computed from Arabic characters are fed to a multilayer perceptron network for recognition. In [17], authors segment Arabic words into characters and use matching of edge points to recognize characters. Sarfraz et al. [18] employ horizontal and vertical projection profiles for segmentation of text into lines and characters respectively. Features based on moment invariant descriptors are extracted from segmented characters and are used to train a radial-basis function network which learns the different character classes. Authors in [13] present an effective methodology for segmentation of Urdu text into characters. A set of structural features is used for segmentation and many of the common problems causing over and under segmentation have been addressed in this study. Zaheer et al. [7] apply three levels of segmentation to Urdu text, line, word and character segmentation. The segmented characters are recognized using a neural network. Another segmentation based Urdu OCR is proposed in [19] where the authors exploit pixel strength to segment the text into words and subsequently into characters. Using neural network as classifier, the system reports an average recognition rate of 70% on 56 character classes.

Pal et al. [14] present an Urdu OCR which segments lines of text using projection profiles while characters are extracted using heuristics on runs of text and non-text pixels. Features including topological features, water reservoir features and contour features are used to recognize characters using a tree classifier. In another study [20], Arabic characters are segmented using a dynamic window sizing and are recognized using cross correlation. A segmentation based Urdu OCR for Nastaliq font is presented in [8] where the skeletonized text is segmented at branching points and each segment is framed. These segments are then used to train HMMs which are subsequently used for recognition. A similar approach for recognition of Noori Nastalique Urdu text in font size 36 is proposed in [11]. For segmentation, the authors first extract the baseline and categorize the components attached with the baseline as primary components (main body). Thinning of these primary components is then carried out and the stroke junctions in the thinned image are used a segmentation points. Each segment is then framed and DCT based features extracted from these frames are fed to HMMs for training/recognition.

In contrast to segmentaiton-based approaches, the segmentaiton-free approaches perform recognition at ligature or word levels. Segmentation-free techniques tend to be less complex than segmentation-based approaches in the sense that they do not require segmentation of text into individual characters. These methods are relatively easier to implement but the major challenge with these approaches is the larger number of classes to be recognized [21], [22], [4], [23], [24], [25]. This number is the same as the number of unique words (ligatures) in the vocabulary under study.

A segmentation-free approach for recognition of Urdu handwritten words is proposed in [9]. The authors extract gradient and structural features from Urdu words which are recognized using Support Vector Machine with radial basis function (RBF) as kernel. A relatively simple recognition system is presented in [21] where template matching using cross correlation is used to match ligatures. A font independent approach for offline and online Urdu OCR is presented in [10]. Each word is considered a composition of compound components and baseline information is used to associate the secondary components with the respective primary components. Features based on Hu moments are extracted by using four windows of different sizes while recognition is carried out using K-nearest neighbor (KNN) classifier [10].

In [26], the authors propose a multi-font Arabic and Persian OCR. The technique relies on a series of preprocessing steps including global thresholding, connected component extraction and skew detection and correction. Features extracted from the contours of connected components are used for recognition. A system for recognition of Arabic literal amounts is presented in [27]. The features investigated in this study include number of ascenders, number of descenders and number of loops etc. These structural features are fed to a number of classifiers for recognition.

A system for recognition of isolated Urdu characters is presented in [4]. The character images are binarized and the chain code of each character is saved in an xml file along with the respective class name and character code. For recognition, diacritics are removed and the chain code of query sample is compared with those in the database to find the nearest match [4]. Another effective Urdu OCR system based on Hidden Markov Models (HMM) is presented in [23]. Main body ligatures are separated from diacrtics and a separate HMM is trained on each ligature. Features based on Discrete Cosine Transform (DCT) computed from a sliding window on each ligature are fed to the HMMs for training. During recognition, the query word is separated into diacritics and main body and the features computed from each are fed to the recognizers. A set of rules is defined to associate the diacritics with the main body and recognize the complete ligature. Another segmentation-free recognition system for Urdu ligatures is presented in [24] where a set of shape descriptors is used to characterize the ligatures. A total of 10,000 Urdu ligatures in Nastaliq font and 20,000 Arabic ligatures in the Naskh font are used as training data. Recognition is carried out using k-nearest neighbor scheme to find the best match for the query ligature. The system evaluated on a custom developed UPTI (Urdu Printed Text Image Database) database realized promising recognition rates.

A font size independent Urdu OCR is proposed in [25]. Ligatures are extracted using connected component labeling and the stroke width information is exploited to distinguish primary and secondary (diacritics) ligatures. The secondary ligatures are categorized into four classes while a fifth class comprises all primary ligatures. Recognition is carried out using structural features mainly including end points, turning points, junction points, cross points stroke width and height etc. Shehzad et al. [28] developed a system for recognition of isolated Urdu characters. The technique rests on a set of primary and secondary stroke features. The primary stroke features include length of bounding box diagonal, angle of bounding box diagonal, and total length of primary stroke etc. while the secondary stroke features comprise the number of secondary strokes, total length of secondary strokes, positioning of the secondary strokes, and number of dots in secondary strokes. Among other ligature based recognition systems template matching [29] and Fourier descriptors have also been investigated [30].

The above paragraphs discussed few of the notable contributions towards development of an OCR for Urdu and Urdu-like cursive scripts. This discussion by no means is exhaustive but is intended to provide an overview of the wide variety of approaches that have been proposed in the recent years. Interested readers may consult the detailed survey papers on this subject [31].

After having discussed few of the well-known OCR techniques proposed for Urdu and similar cusrive scripts, we present the proposed recognition methodology in the next section.

III. Proposed Methodology

This section presents in detail the proposed approach for recognition of Urdu text in Nastaliq script. The technique relies on a segmentation-free method which employs ligatures as the basic units of recognition where each ligature comprises one or more characters (connected together using the joiner rules). As discussed earlier, segmentation-free approaches reduce the complexity of the system as segmentation at character level is not required but, on the other hand, it increases the number of classes to be recognized which is same as the number of unique ligatures. We extract the connected components in an image and separate the components into primary and secondary liguatres which correspond to main body and diacritics respectively. Each ligature is represented by a set of features and a clustering is carried out to group ligatures into clusters. These clusters serve as training data and a separate hidden Markov model (HMM) is trained on each ligature. Once the individual ligatures are recognized using these HMMs, the secondary ligatures are associted with the primary ligatures to recognize the complete word. These steps, distinguished into training and recognition phases, are presented in detail in the following.

A. Training

Training involves making the model(s) learn to discriminate between different (ligature) classes. The different steps invovled in this training phases are detailed in the following.

1) Preprocesing: Preprocessing is the first step in the development of any OCR system which prepares images for the subsequent phases. Depending upon the application and the type of input images, preprocessing may involve binarization [32], noise removal and, skew and slant detection and correction [33]. In our study, we intend to work on contemporary images of Urdu text which are not likely to suffer from noise or degradations. The preprocessing in our case, therefore, comprises binarization of image to segment text from the background. In our implementation, we have employed the well-known Otsu‘s global thresholding to binarize the text image. An example grayscale image and its binarized version are illustrated in Figure 1.

2) Extraction of Ligatures: Urdu, being a highly cursive language, makes the segmentation of Urdu text into characters not only challenging but also prone to errors. Urdu words are a composition of ligatures and diacritics. A ligature can be an isolated character or a combination of characters joined together while diacritics are the secondary components. In our study, we use ligatures as the basic recognizable units. Ligatures and diacritics are extracted from binarized words using connected component labeling. Figure 2 illustrates an example image and the corresponding connected components which are then fed to the next stage of feature extraction.

3) Feature Extraction: Feature extraction is the pivotal stage in any recognition/classification task. Representing shapes (ligatures in our case) by features not only allows reducing the dimension but also allows effective comparison of these ligatures as opposed to the pixel representation. Most of the features previously employed for Urdu OCR, however, work on fixed font sizes. In our methodology, we have chosen scale invariant global transformational features to represent ligatures and diacritics. These features include horizontal projection, vertical projection, upper profile and lower profile. These features have been effectively employed in a number of shape matching (word spotting) problems [34] and have shown promising performances. These features and their computational details are presented in the following.

a) Horizontal Projection: : is the sum of pixel values in each row of the input image. The projection is normalized to the range $[0-1]$ by dividing each value by the width (number of columns) of the input ligature image.

b) Vertical Projection: : is the sum of pixel values in each column of the input image. The sequence values are normalized between 0 and 1 by dividing them by the height (number of rows) of input ligature.

c) Upper Profile: : is calculated by finding, in each column, the distance of the first text pixel from the top of the bounding box of the ligature. Upper profile is then normalized by dividing it by the height (number of rows) of the input ligature.

d) Lower Profile: : is computed by finding, for each column, the distance of the last text pixel from the top of the bounding box of the ligature. Like upper profile, lower profile is also normalized by dividing it by the height of the ligature.

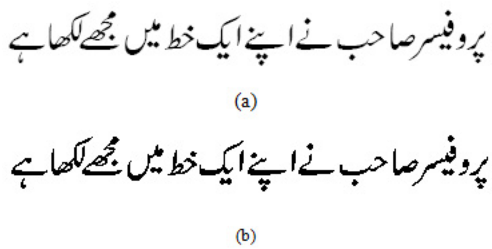


Fig. 1: a) Original image b) Binarized image with global thresholding

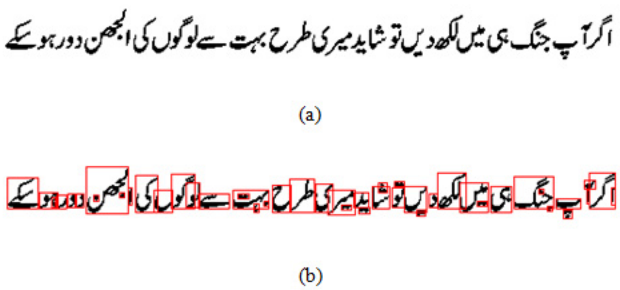


Fig. 2: a) Binarized image b) Connected component labeling

The projection and profile features extracted from an example ligature are illustrated in Figure 3.

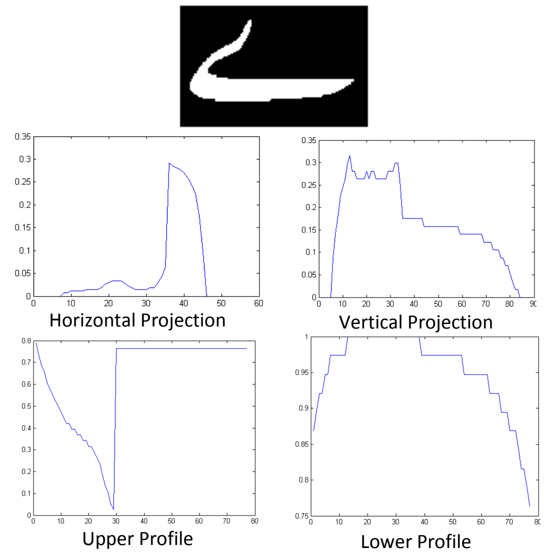


Fig. 3: Projection and Profile features extracted from a ligature 'Bari Yay'

4) Clustering of Ligatures: In order to train the classifiers to recognize ligatures, labeled ligature classes need to be established. Manual generation and labeling of training data, naturally, is a tedious task. We, therefore, chose to employ a semi-automatic scheme where clusters of ligatures are generated from a given set of document images. Each cluster is then labeled and the errors in clustering process are removed manually to ensure that the training data does not contain erroneous clusters. These clusters or classes serve as training data to train the recognizers.

To generate training data for recognizers, we take samples of 30 document images and extract the ligatures (main body as well as dots) as discussed earlier. In order to have scale invariance, some of the documents are resized using scale factors of 0.5, 0.75, 1.25, 1.5 and 2.0. The extracted ligatures from these images are grouped into clusters eventually to be used as training data. For clustering, we have employed a sequential clustering algorithm [35], [36] which does not require a priori the number of clusters. We start by randomly picking a ligature and assuming it to be the mean (representative) of the first class (cluster). For each subsequent ligature, we compute its distance (using Dynamic Time Warping) with the center of each cluster and chose the nearest cluster as a potential candidate. If the distance of the ligature in question to the nearest cluster is below an empirically determined threshold, the ligature is assigned to this cluster and the cluster mean is updated. In case the distance does not fall below the predefined threshold, a new cluster is created with the ligature in question as the mean of the newly generated cluster. This process is repeated until all the ligatures have been assigned to a cluster.

Naturally, the clustering algorithm used in our study has certain short comings. The most significant of these is that the generated clusters are sensitive to the order in which the ligtaures are presented to the algorithm. However, it should be noted that the objective of clustering is to generate an approximate set of ligature classes which are manually corrected prior to training of recognizers. Hence, the performance of the recognition system is not sensitive to this clustering step.

Executing the mentioned clustering algorithm on the sample images, we get a total of 246 clusters. These clusters, naturally, contain some errors which are corrected manually. After refinement, we come up with a total of 250 clusters. The idea of using a sequential clustering instead of traditional k-means clustering and similar algorithms is not to fix the number of clusters a priori. The 250 ligature clusters, representing the frequent Urdu ligatures have been generated using random Urdu texts. Using a larger collection of text images to generate the clusters will naturally produce a larger number of clusters.

It is to be noted that the total number of valid Urdu ligatures is more than twenty thousand. However, most of these ligatures occur very rarely in the text. Studies carried out by the Center of Language Engineering (CLE) at Lahore, Pakistan, have concluded that most of the Urdu words can be generated using the frequent Urdu ligatures. CLE has also compiled the frequencies of occurrence of Urdu ligatures from a huge corpus of text. In our study, we have extracted the frequencies of occurrence of our 250 ligatures from the statistics compiled by the CLE. These frequencies are sorted in descending order and are illustrated in Figure 4. It can be noticed that using only 250 frequent ligatures can allow recognition of a large number of frequent Urdu words.

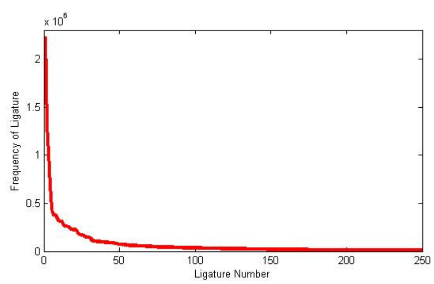


Fig. 4: Frequencies of 250 ligatures

Once the ligature clusters are generated, each cluster needs to be assigned its respective Unicode. Each ligature comprises one or more characters. Once the characters appear as a part of ligature, a number of characters may exhibit the same shape and can only be differentiated by the position and number of dots. Since in our implementation, the ligature classes do not include diacritics (which are treated separately), the number of character classes in the ligatures is less than the actual number of characters. To elaborate this idea Figure 5 illustrates two Urdu ligatures 'Ba' and 'Na'. In the absence of dots, both these ligatures are exactly the same so the characters 'Bay' and 'Noon' belong to the same class as they have exactly the same shape (for this combination). Analyzing different combinations of characters to form ligatures, a total of 20 character classes, listed with their respective codes in Figure 6, are identified. Each ligature has its unique Unicode string depending upon the classes of characters it comprises of. Referring again to the example of ligatures 'Ba' and 'Na', in the absence of dots, both these ligatures are in the same cluster and the Unicode associated with this cluster is the Unicode of 'Bay' + 'Alif' as both 'Bay' and 'Noon' belong to the character class 'Bay'. On the other hand, in isolated form, 'Noon' belongs to its own class.

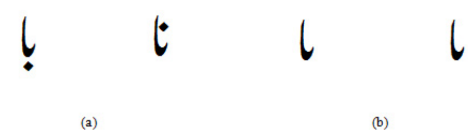


Fig. 5: Images of two ligatures 'ba' and 'na' (a): With dots (b): Without dots

Once the ligatures have been grouped into clusters, we need to choose a classifier which could be trained to recognize these ligatures. Classifiers like artificial neural networks, support vector machine or hidden Markov models (HMM) could be effectively employed for this purpose. In our study, we have used HMMs which have been successfully applied to problems like gesture recognition [37], [38], [39], speech recognition [40], handwriting recognition [41], [42], [43], [44], musical score recognition [45] and optical character recognition [23]. The training of HMMs to model the ligatures is discussed in the following.

5) HMM Training: A separate HMM is trained for each ligature (main body as well as diacritics) considered in our study. All ligatures are resized to a predefined height of 64 pixels and each ligature is scanned from right to left using a frame (sliding window) of size 64×7 with an overlap of 4 pixels (Figure 7). The sensitivity of overall system performance to these parameters is dicussed later in the paper. For each position of window, we extract a set of features which includes the total number of text pixels in the window, sum of horizontal edges, sum of vertical edges, horizontal projection and vertical projection. For edge based features, Sobel edge detection (horizontal/vertical) is applied to the pixels in the window and the numbers of 1s in the output image are counted. The horizontal (vertical) projection is computed by counting the number of pixels in each row (column) of the image (window). These features have been effectively applied to a number of recognition problems including character recognition [46], word spotting [47] and identification of writers [36]; consequently, they are likely to perform well for characterization of ligatures as well.

To train the models, the feature space is quantized to a 75 symbol codebook (since the HMMs are discrete). Right-to-left HMMs with 9 states (Figure 8) are used in our study and a separate HMM is trained for each ligature using

Class	Class Members	Class Unicode
ا	ا	0627
ب	ب, پ, ت, ٹ, ث, ن , ی	0628
ج	ج, چ, ح, خ	062C
د	د, ڈ, ذ	062F
ر	ر, ڑ, ز, ژ	0631
س	س, ش	0633
ص	ص, ض	0635
ط	ط, ظ	0637
ع	ع, غ	0639
ف	ف, ق	0641
ق	ق	0642
ک	ک, گ	06A9
ل	ل	0644
م	م	0645
ن	ن, ں	0646
و	و	0648
ہ	ہ	06C1
ھ	ھ	06BA
ی	ی	06CC
ے	ے	06D2

Fig. 6: Groups of character classes in the absence of diacrtics and the respective Unicodes

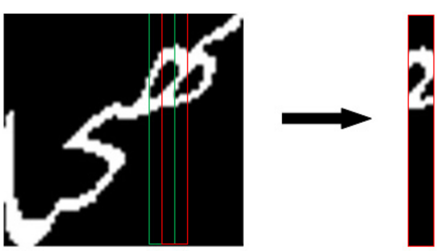

(a) (b)

Fig. 7: Sliding Windows a) Overlapping in Sliding Windows b) A single window

the standard Baum-Welch algorithm. Once the models are trained, the Unicode string of each ligature is associated with its respective HMM.

B. Recognition of Words

The basic unit of recognition in or study is ligature. However, since a word represents a semantically meaningful unit, the system is presented with words for recognition. Once a query word is presented to the system, we first segment it into ligatures using connected component labeling. Ligatures are then separated into primary ligatures and diacritics and the position information of diacritics with respect to the primary ligatures is also stored. Each ligature is recognized by finding the HMM which reports the maximum probability when the features of the respective ligature are fed to the trained HMMs. Diacritic information is then associated with the recognized primary ligature to find the true ligature. The Unicode of the complete word is then written to an output file after dictionary matching. The steps involved in the recognition phase are summarized in Figure 9 while each of these steps is discussed in detail in the following sections.

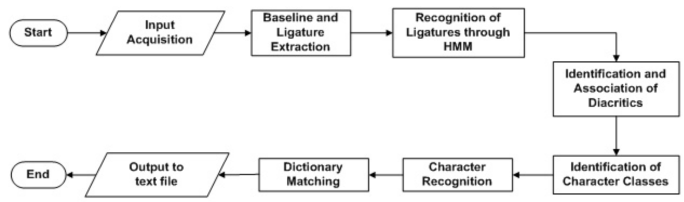


Fig. 9: Flow chart for Recognition of Urdu Words

1) Baseline and Ligature Extraction: The query word presented to the system is binarized using global thresholding and the baseline of the word is determined by finding the row with the maximum number of text pixels. Figure 10 illustrates a sample input word and its baseline. To extract the ligatures of a given word, connected component labeling is employed similar to the training phase. The extracted ligatures are then recognized using the trained models.


(a) (b)

Fig. 10: a) Input word 'Aiwan' b) Baseline detection

2) Recognition of Ligatures: Ligatures extracted from the query word are fed to the trained HMMs. Each ligature is recognized by the HMM which produces the maximum probability, the respective Unicode being the output. It should be noted that since the main body of the ligature is separate from the diacritics in the clusters of ligatures, the output of the HMM is actually the Unicode of the respective character class (as per Figure 6) rather than the true Unicode.

3) Association of Diacritics with Ligatures: Diacritics are differentiated from ligatures through their Unicodes. Diacritics are categorized depending upon their position with respect to baseline of word. When a diacritic is above the baseline it is categorized as 'above' and, if it is below the base line it is identified as a 'below' diacritic. In cases where the diacritic resides on the baseline we use the height information

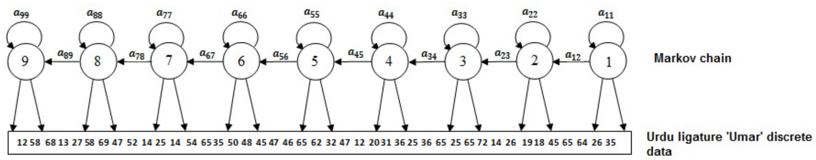


Fig. 8: Markov chain model generated for an example ligature

TABLE I: Diacritics with corresponding integer values and descriptions

Diacritic	Value	Description
No dots	-1	No dots
One dot above	1	one dot is above the baseline
One dot below	2	one dot is below the baseline
Two dots above	3	two dots are above the baseline
Two dots below	4	two dots are below the baseline
Three dots above	5	three dots are above the baseline
Three dots below	6	three dots are below the baseline

to determine whether more part of the diacritic is above or below the baseline. Based on the number and position of diacritics, an integer value is assigned to it as per Table I.

Within a word, diacritics are associated with ligatures depending upon their position information with respect to the ligature. Each diacritic is analyzed for each ligature and, if it lies within the width of ligature then it is associated to the respective ligature. In case of more than one candidate ligature, percentage of diacritic width overlap with each ligature is computed and the diacritic is associated with ligature which has maximum overlap.

Once the diacritics are associated with a main body ligature, they have to be further associated with the characters within a ligature. In some cases, multiple diacritics within a single main ligature can generate a large number of possible associations. This is illustrated in Figure 11 where we have ligature classes 'Bay' + 'Bay' + 'Alif'. The diacritic information associated with this ligature is 'one dot above' and 'one dot below'. This could result in two different ligatures 'Bana' and 'Naba'. In all such cases, the possible combinations of ligatures are matched with a valid ligature dictionary and the first valid instance is picked as the recognized ligature.

Fig. 11: (a) Ligature 'Bay' + 'Bay' + 'Alif' with 'one dot above' and 'one dot below' (b) possible words 'Bana' and 'Naba'

Once the individual ligatures in a word are recognized, the Unicodes of these ligatures are combined to generate the Unicode string for the query word. The Unicode string is written to a text file using UTF-8 encoding to save the output.

This concludes our discussion on the proposed recognition methodology. In the next section we present the experiments carried out to validate the proposed technique.

IV. Experiments and Results

This section presents the results of the experiments carried out to evaluate the effectiveness of the proposed methodology. We first present the data set used in different experiments followed by a discussion on the clustering of ligatures into classes. We then present the recognition rates reported by the system at ligature and word levels followed by some interesting analytical experiments.

A. *Dataset*

For generating the clusters of ligatures we used 30 full page length scanned images of Urdu documents from the CLE database (http://www.cle.org.pk). In order to have scale invariance, the original images were scaled by different factors as discussed earlier. The system was evaluated on 100 query words with a total of 351 ligatures. For comparison purposes, ligature recognition rate was also reported on 2,017 high frequency ligatures from the CLE database.

B. *Performance of Ligature Clustering*

As stated earlier, we employ a semi-automatic sequential clustering to generate ligature classes. The 30 document images used for clustering comprised a total of 10,364 ligatures out of which 8,800 were correctly categorized into 246 clusters realizing an accuracy of 85%. Since the clusters had to be used as training data, the errors in clusters were manually corrected after visual inspection making a total of 250 clusters.

C. *Performance of Recognition*

To evaluate the recognition performance we performed experiments with 100 query words. Recognition rate is computed at ligature as well as word level. Each of these is discussed in the following.

1) Ligature Recognition Rate: The 100 query words comprise a total of 351 ligatures which include isolated characters, two, three, four, and five character ligatures and diacritics. Figure 12 illustrates some example query words used in our study. For recognition, features extracted by sliding window on the query ligature are fed to all the (trained) HMMs. The HMM reporting the maximum probability identifies the ligature. Table II summarizes the recognition rates on each of these categories of ligatures. Over all the system correctly recognizes 335 out of 351 ligatures realizing an overall ligature recognition rate of 95%.

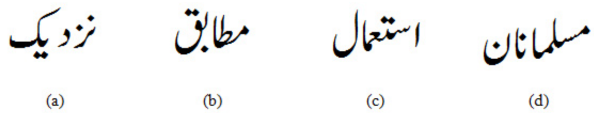


Fig. 12: Examples of query words

2) *Word Recognition Rate:* Once the ligatures are recognized, we apply the post processing steps to associate diacritics with their respective ligatures and recognize the complete word. This step naturally is error prone as association of diacritics with words is based on a set of heuristics and, in some cases, this association is not very straight forward. In addition, error in any one of the ligatures of a query word is considered an error at word level. Using the diacrtic association rules discussed earlier, the system is able to correctly recognize 89 out of 100 query words.

Comparing a word recognition rate of 89% with the results reported in the literature, it is important to mention that the high recognition rates reported in the literature are either on isolated characters or on ligatures. Some studies, which work at word level, totally ignore the diacritic information resulting in relatively high recognition rates. In our study, we consider the query words without any constraints to resemble the real world scenarios as closely as possible. A recognition rate of 89% at word leve, therefore, is very promising.

In order to compare the performance of our system with existing Urdu OCR techniques, we also perform a series of experiments on the CLE ligature database as discussed in the next section.

D. Performance on CLE High Frequency Ligature (HFL) Dataset

The experimental results discussed in the above sections are based on a total of 250 clusters of ligatures (independent of the font size). For comparison purposes and in order to study how the performance of the recognizer varies as the number of ligatures increases, we also train HMMs on the frequent ligature dataset of CLE. We compute the ligature recognition rate as a function of the number of ligatures by varying the number of ligatures from 50 to 2,017. The results of these evaluations are summarized in Figure 13 where it can be seen that there is a natural gradual decrease in the recognition rates as the number of ligatures rises. A ligature recognition rate of 92% on 2,017 ligatures is not only promising but it also demonstrates the scalability of the proposed recognition scheme. A comparison of our recognition rates with some notable studies on Urdu OCR is summarized in Table III. It can be seen that our results are comparable with those obtained by [23] on the same data set. The major drawback of this dataset, however, is the fixed font size of all the ligatures; a constraint which is hard to meet in real world problems. Other studies listed in Table III either work on isolated characters or ignore the diacritic information hence the high recognition rates realized in these studies may not be very signficiant.

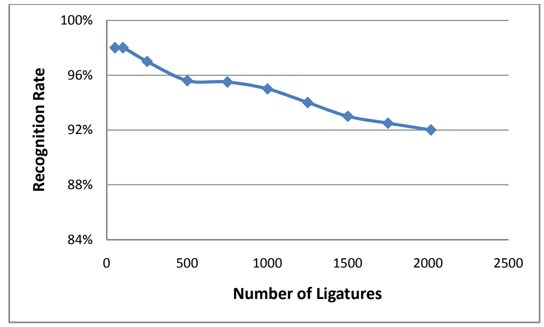


Fig. 13: Recognition rate as a function of number of ligatures (CLE Database)

In addition to the number of ligatures, we also study the sensitivity of the recognition rate to other parameters of the system. These include the number of states in the HMMs and the cell size for framing of ligatures. These experiments are carried out on the first 500 ligatures of the CLE high frequency ligature (HFL) database and the realized recognition rates are summarized in Figure 14. It can be observed that the recognition rates increase with the increase in number of states in the HMM and begin to stabilize from 9 states onwards. The recognition rates seem to be more sensitive to the frame size employed during feature extraction. For all experiments, for a frame width of n pixels, there is an overlap of $(n+1)/2$ pixels. Smaller frame sizes yield higher recognition rates which drop as the frame width increases. This observation is natural as larger windows include larger proportions of ligatures which may not be common across multiple samples of the same ligature. Smaller frame widths result in a larger number of windows per ligature and the extracted features are more effective in characterizing these ligatures.

V. Conclusion

This work presented a segmentation free Urdu OCR for printed text in Nastaliq font. The proposed technique considers ligatures as the basic units of recognition. Ligatures are extracted by performing connected component labeling on the binarized document images of Urdu text. A total of 250 ligature clusters are generated which serve as training data for ligature modeling. A separate right-to-left Hidden Markov Model (HMM) is trained for each of the ligatures (main body as well as diacritics) using features extracted by a sliding window. For recognition, ligatures in the query word are extracted and each ligature is recognized using the trained HMMs. Diacritics and main body ligatures are separately recognized. Diacritics are then associated with respective ligatures using the position information and the complete ligature is recognized using dictionary validation. Finally, the Unicode string of the word is written to a text file.

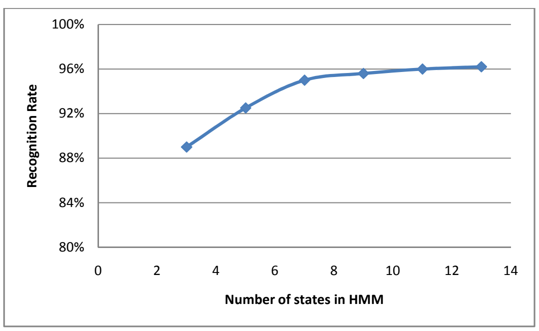
The proposed system realizes very promising ligature recognition rates. The relatively low word recognition rate is

TABLE II: Types of ligatures and corresponding recognition rates

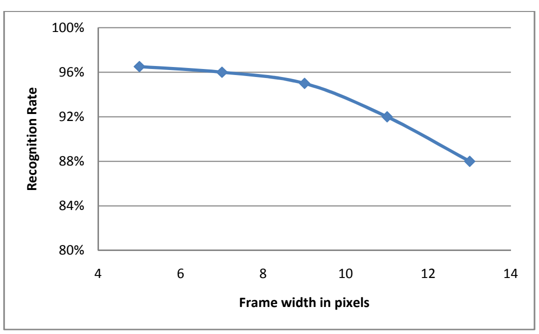
Type of ligature	Total Ligatures	Correctly Recognized	Recog. Rate
Isolated characters	117	112	96%
Two character ligatures	99	92	93%
Three character ligatures	22	20	91%
Four character ligatures	6	5	83%
Five character ligatures	4	4	100%
Diacritics	103	102	99%

TABLE III: Notable studies on Urdu OCR with recognition rates and limitations

Study	Dataset	Recognition Rate	Limitations
Z. Ahmad et al. [7]	Synthetic and real-world images of Urdu	93.4%	Assumes ligatures to be diacritic free
T. Nawaz et al. [4]	Isolated characters in different font sizes	89%	Tested only on isolated characters
N. Sabbour & F. Shafait [24]	UPTI Online Arabic e-book	Urdu:99% Arabic: 86%	Ignores diacritic information
D. Satti [25]	29,517 ligatures	97.10%	Ignores secondary components and ligature arrangement
Javed et al. [23]	1282 unique ligatures	92%	Fixed size ligatures only
Proposed Method	100 words 2,017 ligatures	89% 92%	Complete words, scale invariant Same dataset as in [23]



(a)



(b)

Fig. 14: Recognition rates on 500 ligatures as a function of: (a) Number of states in the HMM (Frame width fixed to 7) (b) Frame width for feature extraction (Number of states fixed to 9)

understandable due to the complexities involved in associating diacritics with the main body ligatures. It should also be noted that contrary to most of the existing Urdu OCRs, our approach is scale invariant and is not tuned to recognize words in a particular font size. The most obvious extension of the proposed system is to increase the number of ligatures to cover a major proportion of all Urdu words. The diacritics handled in our study include different numbers and positions of dots. Other diacritics like ‘choti toye’, ‘shad’, ‘mad’, ‘zabar’, ‘zair’ etc. can also be incorporated in the system. Real word Urdu documents may suffer from problems like noise, degradation and skew and consequently will require a preprocessing step prior to recognition.

REFERENCES

[1] F. Iqbal, A. latif, N. Kanwal, and T. Altaf, “Conversion of urdu nasstliq to roman urdu using ocr,” in *4th International Conference on Interaction Sciences (ICIS)*, 2011, pp. 19–22.

[2] I. Shamsher, Z. Ahmad, J. K. Orakzai, and A. Adnan, “Ocr for printed urdu script using feed forward neural network,” in *Proceedings of World Academy of Science, Engineering and Technology*, 2007, pp. 172–175.

[3] P. Natarjan, Z. Lu, R. Schwartz, I. Bazzi, and J. Makhou, “Multilingual machine printed ocr,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 15, no. 43, 2001.

[4] T. Nawaz, S. A. H. S. Naqvi, H. ur Rehman, and A. Faiz, “Optical character recognition system for urdu (naskh font) using pattern matching technique,” *International Journal of Image Processing*, vol. 3, pp. 92–104, 2009.

[5] N. B. Amor and N. E. B. Amara, “Multifont arabic character recognition using hough transform and hidden markov models,” in *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, 2005, pp. 285–288.

[6] K. Khan, R. Ullah, N. A. Khan, and K. Naveed, “Urdu character recognition using principal component analysis,” *International Journal of Computer Applications*, vol. 60, no. 11, pp. 1–4, 2012.

[7] Z. Ahmad, J. K. Orakzai, I. Shamsher, and A. Adnan, “Urdu nastaleeq optical character recognition,” in *Proceedings of world academy of science, engineering and technology*, 2007, pp. 249–252.

[8] S. T. Javed and S. Hussain, “Segmentation based urdu nastalique ocr,” in *Proceedings of 18th Iberoamerican Congress on Pattern Recognition*, 2013, pp. 41–49.

[9] M. W. Sagheer, C. L. He, N. Nobile, and C. Y. Suen, “Holistic urdu handwritten word recognition using support vector machine,” in *Proceedings of 20th International Conference on Pattern Recognition*, 2010, pp. 1900–1903.

[10] S. Sardar and A. Wahab, "Optical character recognition system for urdu," in *Proceedings of International Conference on Information and Emerging Technologies*, 2010, pp. 1–5.

[11] A. Muaz, "Urdu optical character recognition system," 2010.

[12] M. A. U. Rehman, "A new scale invariant optimized chain code for nastaliq character representation," in *Proceedings of 2nd International Conference on Computer Modeling and Simulation*, vol. 4, 2010, pp. 400–403.

[13] H. Malik and M. A. Fahiem, "Segmentation of printed urdu scripts using structural features," in *Proceedings of 2nd International Conference in Visualisation*, 2009, pp. 191–195.

[14] U. Pal and A. Sarkar, "Recognition of printed urdu script," in *Proceedings of 12th International Conference on Document Analysis and Recognition*, vol. 2, 2003, pp. 1183–1187.

[15] A. Cheung, M. Bennamoun, and N. W. Bergmann, "An arabic optical character recognition system using recognition-based segmentation," *Pattern Recognition*, vol. 34, no. 2, pp. 215–233, 2001.

[16] M. M. Altuwaijri and M. A. Bayoumi, "Arabic text recognition using neural networks," in *IEEE International Symposium on Circuits and Systems*, vol. 6, 1994, pp. 415–418.

[17] M. A. Abdullah, L. M. Al-Harigy, and H. H. Al-Fraidi, "Off-line arabic handwriting character recognition using word segmentation," *Computing Research Repository*, vol. 4, no. 3, pp. 40–44, 2012.

[18] M. Sarfraz, S. N. Nawaz, and A. Al-Khuraidly, "Offline arabic text recognition system," in *Proceedings on International Conference on Geometric Modeling and Graphics*, 2003, pp. 30–35.

[19] Z. Ahmed, J. K. Orakzai, and I. Shamsher, "Urdu compound character recognition using feed forward neural networks," in *Proceedings of 2nd IEEE International Conference on Computer Science and Information Technology*, 2009, pp. 457–462.

[20] A. Mesleh, A. Sharadqh, J. Al-Azzeh, Z. MazenAbu, N. Al-Zabin, T. Jaber, A. Odeh, and M. Hasn, "An optical character recognition," *Contemporary Engineering Sciences*, vol. 5, pp. 521 – 529, 2012.

[21] S. A. Sattar, S. Haque, and M. K. Pathan, "Nastaliq optical character recognition," in *Proceedings of the 46th Annual Southeast Regional Conference*, 2008, pp. 329–331.

[22] M. Akram and S. Hussain, "Word segmentation for urdu ocr system," in *Proceedings of the 8th Workshop on Asian Language Resouces*, 2010, pp. 88–94.

[23] S. Javed, S. Hussain, A. Maqbool, S. Asloob, S. Jamil, and H. Moin, "Segmentation free nastalique urdu ocr," in *Proceedings of World Academy of Science, Engineering and Technology*, vol. 46, 2010, pp. 456–461.

[24] N. Sabbour and F. Shafait, "A segmentation-free approach to arabic and urdu ocr," in *Proceedings of Document Recognition and Retrieval*, 2013.

[25] D. Satti, "Offline urdu nastaliq ocr for printed text using analytical approach," 2013.

[26] M. Kavianifar and A. Amin, "Preprocessing and structural feature extraction for a multi-fonts arabic/persian ocr," in *Proceedings of the 5th International Conference on Document Analysis and Recognition*, 1999, pp. 213–216.

[27] M. T. El-Melegy and A. A. Abdelbaset, "Global features for offline recognition of handwritten arabic literal amounts," in *Proceedings of the 5th International Conference on Information and Communications Technology*, 2007, pp. 125–129.

[28] N. Shahzad, B. Paulsonn, and T. Hammond, "Urdu qaeda: Recognition system for isolated urdu characters," in *Proceedings of IUI Workshop on Sketch Recognition*, 2009.

[29] Z. Shah and F. Saleem, "Ligature based optical character recognition of urdu nastaleeq font," in *Proceedings of International Multi Topic Conference*, 2002.

[30] S. M. Lodhi and M. A. Matin, "Urdu character recognition using fourier descriptors for optical networks," in *Photonic Devices and Algorithms for Computing VII, Proc. of SPIE*, 2005.

[31] S. Naz, K. Hayat, M. I. Razzak, M. W. Anwar, S. A. Madani, and S. U. Khan, "The optical character recognition of urdu-like cursive scripts," *Pattern Recognition*, vol. 47, no. 3, pp. 1229–1248, 2013.

[32] M. Michael and N. Papamarkos, "An adaptive layer-based local binarization technique for degraded documents," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 24, no. 245, 2010.

[33] P. Sanjoy, P. Bhowmick, S. Sural, and J. Mukhopadhyay, "Skew correction of document images by rank analysis in farey sequence," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 27, no. 1353004, October 2013.

[34] K. Khurshid, "Analysis and retrieval of historical document images," Ph.D. dissertation, 2009.

[35] A. Bensefia, T. Paquet, and L. Heutte, "A writer identification and verification system," *Pattern Recognition Letters*, vol. 26, no. 13, pp. 2080–2092, 2005.

[36] I. Siddiqi and N. Vincent, "Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features," *Pattern Recognition*, vol. 43, no. 11, pp. 3853–3865, 2010.

[37] C. W. Ng and S. Ranganath, "Real-time gesture recognition system and application," *Image and Vision Computing*, vol. 20, no. 13-14, pp. 993 – 1007, 2002.

[38] J. Triesch and C. von der Malsburg, "Classification of hand postures against complex backgrounds using elastic graph matching," *Image Vision Computing*, vol. 20, pp. 937–943, 2002.

[39] H.-S. Yoon, J. Soh, Y. J. Bae, and H. S. Yang, "Hand gesture recognition using combined features of location, angle and velocity," *Pattern Recognition*, vol. 34, no. 7, pp. 1491–1501, 2001.

[40] X. D. Huang, Y. Ariki, and M. A. Jack, *Hidden Markov models for speech recognition*. Edinburgh university press Edinburgh, 1990, vol. 2004.

[41] Thomas and G. A. Fink, "Markov models for offline handwriting recognition: A survey," *International Journal of Document Analysis and Recognition*, vol. 12, no. 4, pp. 269–298, 2009.

[42] E. Kavallieratou, E. Stamatatos, N. Fakotakis, and G. Kokkinakis, "Handwritten character segmentation using transformation-based learning," in *Proceedings of International Conference on Pattern Recognition*, vol. 2, 2000, pp. 2634–2634.

[43] H. Yasuda, K. Takahashi, and T. Matsumoto, "A discrete hmm for online handwriting recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, no. 675, 2000.

[44] J. J. Lee, J. Kim, and J. H. Kim, "Data-driven design of hmm topology for online handwriting recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 15, no. 107, 2001.

[45] B. Pardo and W. Birmingham, "Modeling form for on-line following of musical performances," in *Proceedings of 20th National Conference on Artificial Intelligence*, vol. 2, 2005, pp. 1018–1023.

[46] S. Al-Qahtani, M. Khorsheed, and M. AlSuliman, "Recognising cursive arabic script using hmms," in *NCC*, vol. 17, 2004, pp. 631–637.

[47] K. Khurshid, C. Faure, and N. Vincent, "Word spotting in historical printed documents using shape and sequence comparisons," *Pattern Recognition*, vol. 45, no. 7, pp. 2598–2609, 2012.

Semantic Feature Based Arabic Opinion Mining Using Ontology

Abdullah M. Alkadri[1], Abeer M. ElKorany[2]
Computer Science Department
Faculty of Computers and Information,Cairo University
Cairo, Egypt

***Abstract*—with the increase of opinionated reviews on the web, automatically analyzing and extracting knowledge from those reviews is very important. However, it is a challenging task to be done manually. Opinion mining is a text mining discipline that automatically performs such a task. Most researches done in this field were focused on English texts with very limited researches on Arabic language. This scarcity is because there are a lot of obstacles in Arabic. The aim of this paper is to develop a novel semantic feature-based opinion mining framework for Arabic reviews. This framework utilizes the semantic of ontologies and lexicons in the identification of opinion features and their polarity. Experiments showed that the proposed framework achieved a good level of performance compared with manually collected test data.**

Keywords—Opinion Mining; Sentimental Analysis; Ontology; Feature extraction; Polarity identification;

I. INTRODUCTION

As a result of dramatically increase of using the internet in the recent years; huge information of people opinions was produced on the web, people can post their views using Internet forums, discussion groups, product reviews and blogs. Analyzing this information manually is time consuming and maybe impossible mission. For example, if we wanted to judge the article or product positively or negatively according to the comments of the people, it is difficult to read all comments and classify them manually, so we need an automated technique to do such a task. Opinion mining is the appropriate way which automatically extracts knowledge from people comments.

Opinion mining (also called sentiment analysis, sentiment mining, sentiment classification, subjectivity analysis, and review mining or appraisal extraction) is a subtopic of text mining that it automatically extracts opinions, sentiments, and subjectivity from user-generated reviews [1]. Basic task in opinion mining is to determine the subjectivity, polarity (positive or negative) of a piece of text in other words: What is the opinion of the writer. Opinion mining has a wide range of applications from different domains such as commercial, governmental, political, educational and others [2].

Nowadays, there are three levels of Sentiment Classification in Opinion mining (document, sentence and feature). According to [3], the sentence and document level analyses do not discover what exactly people liked or not. However, studying the opinion text, especially feature level, is extremely challenging. For the ordinary user, it is too complex to analyze opinions about object and object features in the online review sites on the Web. To do such analysis it is necessary to perform some kind of opinion mining, feature-based opinion mining so as to identify the features in the review and classify the sentiments of the opinion for each of these features [4]. The feature-based opinion mining of object reviews is a difficult task, owing to both the high semantic variability of the opinions expressed, and the diversity of the features and sub- features that describe the products and the polarity of opinion words used to depict them [5]. In the last few years, new approaches based on both semantic web technologies and domain-dependent corpora for feature-based opinion mining have appeared [6]. In [7] Isidro et el. Believe that the already mature Semantic Web technology could be a valuable addition to traditional opinion mining approaches. More concretely, ontologies constitute the standard knowledge representation mechanism for the Semantic Web and can be used to structure information. The formal semantics underlying ontology languages enables the automatic processing of the information in ontologies and allows the use of semantic reasoners to infer new knowledge.

In the proposed work, an ontology is viewed as a formal and explicit specification of a shared conceptualization [8]. Ontologies provide a structured knowledge representation and a common vocabulary for a domain (e.g. hotel domain). In this work, the Web Ontology Language (OWL), the W3C standard used to represent ontologies in the Semantic Web, has been used to represent the concepts and features of the application domain (in our case hotel domain). The main contribution of the proposed framework is how to classify Arabic views of people about an entity (Object) in a specific domain to positive or negative opinions. We need a point of view about an entity through extracting the view about its features (attributes). For example, if the entity is a hotel its features will be a room, bar, lunch and so on. Most of researches done in this field were focused on English texts with very limited researches in an Arabic language. Limitation of research work in this area is due to the following reasons:

- Content found on Forums and Blogs is written in many forms of Arabic Dialect which makes the task of using a semantic approach for mining opinions very challenging. Moreover, the majority of the available preprocessing tools are mainly built for the modern standard Arabic.
- The limitation in availability of appropriate datasets, no opinion-related (or sentiment) Arabic Lexicon is present to assist in the task of measuring the polarity

of extracted subjective text.

The rest of this paper is organized as follows: Section II presents Related Work; Section III describes the Ontology and lexicon Development; Section IV details the Proposed Framework; experimental results and discussion are discussed in Section V; conclusions and future work are finally presented in Section VI.

II. RELATED WORK

Sentiment classification approaches can be divided into machine-learning approaches and semantic orientation approaches. Machine learning approaches are typically supervised approaches in which a set of data labeled with its class such as positive or negative are represented by feature vectors. Then, these vectors are used by the classifier as a training data inferring that combination of specific features yields a specific class [9]. Semantic orientation approaches (Dictionary-based approaches) are unsupervised approaches in which a sentiment lexicon is created. The sentiment lexicon performs classification based on positive and negative sentiment words and phrases contained in each evaluation text and mining the data requires no prior training. Several researches have been conducted in the opinion mining field. Researchers have proposed interesting approaches and developed various systems to solve this problem. Most of these systems are developed for English language and are not oriented to other languages. In this section, previous works of Arabic opinion mining systems based on sentiment classification techniques and levels is discussed. Some classic machine learning methods (Naive Bayes, Maximum Entropy, and SVM) have been experimented in [10] [11] [12] [13]. Other works, such as [14] [15] are based on dictionary-based algorithms. [16] Use the both techniques. All studies presented before are either on sentence or document level. Moreover, their datasets (reviews) are collected from general websites that are not interested in specific domain. These classification techniques are useful and improve the effectiveness of Arabic sentiment classification but cannot determine what the user opinion on each particular feature.

Here we will show some studies on feature-based level. In feature based opinion mining some researchers use lexicon to store the domain features such as [17] [18] [19] [20] [21]. While the others use the ontology to represent the domain features such as [7] [22] [23] [24] [25]. Furthermore they use a general, specific domain lexicon or both for the opinion words polarity. None of the previously mentioned works is concerned with Arabic feature based opinion mining. In general, none of the existing works efficiently addresses the task of Arabic feature based opinion mining based on ontology which we are going to address.

III. ONTOLOGY AND LEXICON DEVELOPMENT

The main objective of the proposed framework is provide a feature-based opinion mining for a specific domain. It is significant to mention that the proposed framework is generally applicable for any application by changing both ontology and lexicon. Thus, this section would provide details of construction of domain ontology in case study (hotel) as well as a large scale Arabic opinion lexicon.

A. Ontology Building

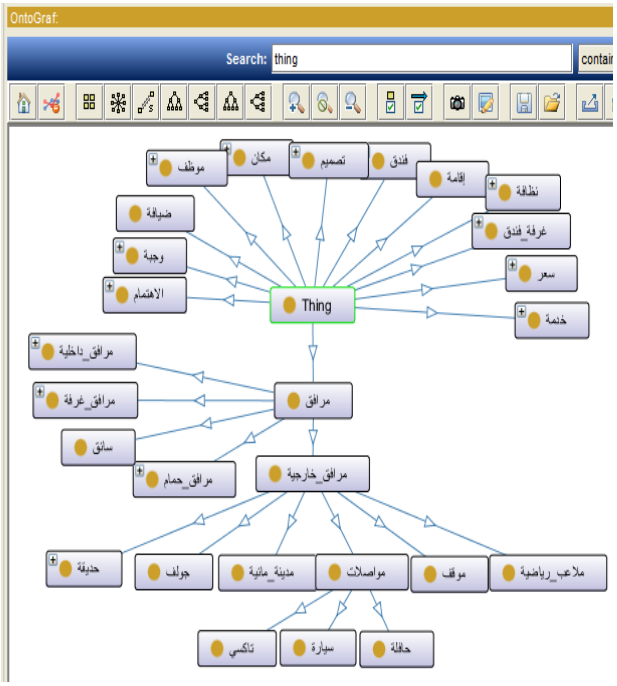
Domain ontology describes the concepts of special domain, including concepts, attributes of the concepts, relationship between concepts, and constraints among the relationships. The concepts refer to different entities that may be a product, or an organization. The aim of using ontology in feature-based opinion mining is to identify the main features of domain by defining the common terminologies in the domain, and giving the definition of the relationships among the terminologies [22].

1) Ontology development: In feature based opinion mining, ontology plays an important role. Thus, in the proposed model, ontology was used to identify the main features of domain (which in our case is hotel). In this section, we present the construction of the ontology. Generally, in order to design domain ontology two methods could be applied: (i) using an existing ontology, extending and adapting it to meet ones needs; and (ii) building one from scratch. Our work started by searching about the most well known ontology in our field in Arabic language but we did not find any. However, we found English hotel ontology Hontology [26] and manually translated it to Arabic. We have developed hotel ontology in Protg [27] with OWL 2. However, some of its concepts and relations do not satisfy what we need in Arabic hotel domain. Furthermore, lack of domain knowledge and ambiguity in the ontology hierarchy represent an obstacle that we have tackled through refinement phase.

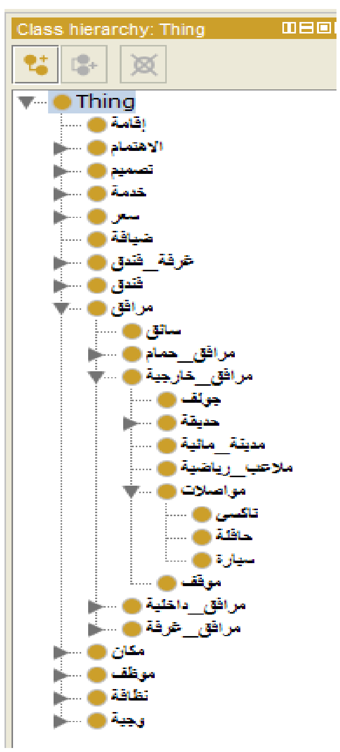
2) Ontology refinement: Due to lack of knowledge in Hontology, therefore, we decided to further extend and adapt this ontology before using it in order to meet our needs. The extension process has been performed manually by three annotators. We assigned set of reviews (690 reviews) to each one of them .The main goal of the annotators is to identify and extract the relevant concepts of hotel domain based on the existing reviews. Each annotator separately generated a list of relevant concepts, those lists are combined such that repeated concepts have been removed. Next, each concept should be aligned in the existing translated ontology. During the alignment step, some concepts have the same meaning, so one of them is manually aligned in the corresponding place and others are considered as synonyms to this concept. These synonyms were collected into a dictionary which called synonyms dictionary. A snapshot of our ontology is presented in Figure 1 (1a,1b). There are 12 top level classes which are associated to hotel domain. In total there are 242 classes in our current hotel ontology.

B. Building a large scale Arabic opinion lexicon

Most opinion mining approaches rely on opinion lexicons, such as English SentiWordnet (ESWN [28] and MPQA Lexicon [29] for identifying word polarity. In order to obtain higher accuracy, it is recommended to use a large scale Arabic opinion lexicon. Recently, some Arabic opining lexicon appear [30], however the availability of a large scale Arabic opinion lexicon is still limited and unavailable. Given the limitation of Arabic sentiment lexicons, we propose to address this limitation, by developing and refining a large-scale Arabic opinion lexicon (ArOpL). We integrate three opinion lexicons to build a common one that contains two lists: one for the



(a) View A



(b) View B

Fig. 1: Hotel Ontology

positive polarity words, and one for the negative polarity ones. The used lexicons were:

- The MPQA lexicon, which contains 8,000 English words that were manually annotated as (positive, negative, neutral, or both) [29]. We depend on Darwish et al. [31] translating; they used Bing online MT system to translate the MPQA lexicon into Arabic. Because The MPQA lexicon had many translation errors, we have refined it, removed these errors and selected only the positive and negative words.
- The ArabSenti lexicon [30] containing 3,982 adjectives that was extracted from news data and labeled as positive, negative, or neutral. We have selected only the positive and negative words.
- Amira et.al. [32] Lists of Sentiment Words, which contains 652 Arabic words labeled as positive and negative.

After refining these lexicons, we remove duplicated words and as a result of annotation phase, another set consists of 100 new words was produced by annotators. Accordingly, it has been added to our lexicon to produce a final set of lexicon with 4420 Arabic words (1825 positive and 2395 negative).

IV. PROPOSED FRAMEWORK FOR SEMANTIC FEATURE- BASED ARABIC OPINION MINING

In this section, the proposed approach is explained. Figure 2 illustrates the Architecture of this system. This framework is composed of five main components: preprocessing, Semantic Feature Identification, Polarity Identification, Feature Polarity Identification and Opinion Mining. These components are described in detail below.

A. Preprocessing

Several NLP techniques must be applied over the dataset to ensure the cleaning of data and remove noise that may affect the accuracy of the system. These techniques include Tokenizer, De-noiser, Normalizer and Stemmer by Althobaiti et.al. [33]. Finally Arabic Stanford Part-Of-Speech Tagger [34] is used to identify the nouns.

B. Semantic Feature Identification

Domain ontology is used for feature identification from hotel reviews. After preprocessed hotel review and determine the Nouns in the reviews using POS taggers, three steps were followed to find the corresponding feature using the domain ontology. According to workflow shown in Figure 3:

- First, the domain concepts is determined in the reviews: Using semantic information encoded in the ontology, our system determines which features are useful for extracting reviews by comparing the nouns of reviews with concepts in the ontology. If the noun does not exist in the ontology, we go to step 2.
- Second, compare the nouns of reviews with the synonyms dictionary which we created. If the noun does not exist in the dictionary, we go to step 3.

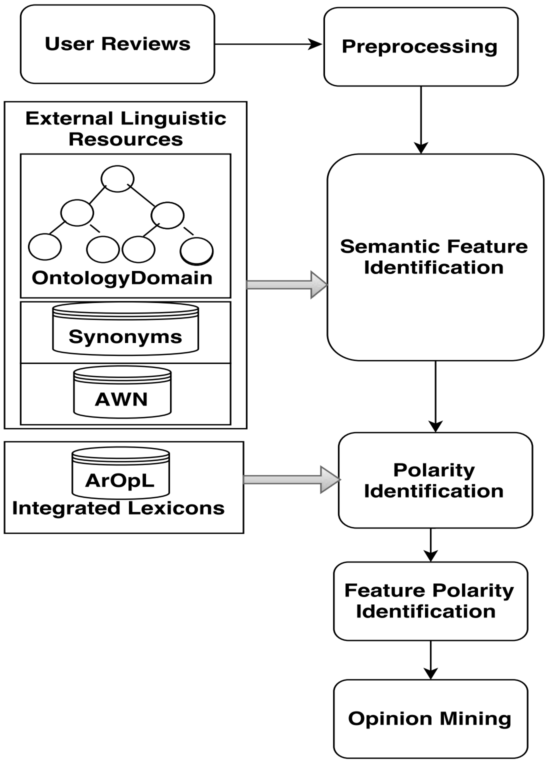


Fig. 2: Proposed system architecture.

- Third, compare the nouns of reviews with the synonyms in AWN [35]. If the noun does not exist, we ignore it.

Table I provides an example of feature identification from a review. According to the example in this table, the identified features within the given review will be (شامبو، فوطه، الغرفه، دورة المياه، السرير، التلفزيون ، الانترنت). More concretely, "الغرفه" ، "شامبو" ، and " السرير ، التلفزيون" are identified as subclasses in the ontology classes "غرفة الفندق" ، " منتجات نظافة "، " مرافق الغرفة" respectively. "فوطه" and "دورة المياه" are identified as features because they are synonymous of the "مناشف" and "حمام" classes from synonyms dictionary and AWN respectively.

C. Polarity Identification

For polarity identification, a list of opinion words is essential, i.e., an opinion lexicon. Opinion words are words that express positive or negative sentiments. Thus, we used the developed new large lexicon (ArOpL). During this phase we ignore the selected features which we already extracted in section IV-B. The polarity is determined by aggregating the polarity of the extracted words in reviews based on our new dictionary. In other words, for each review our method assigns the scores +1 and -1 to the positive and negative words respectively.

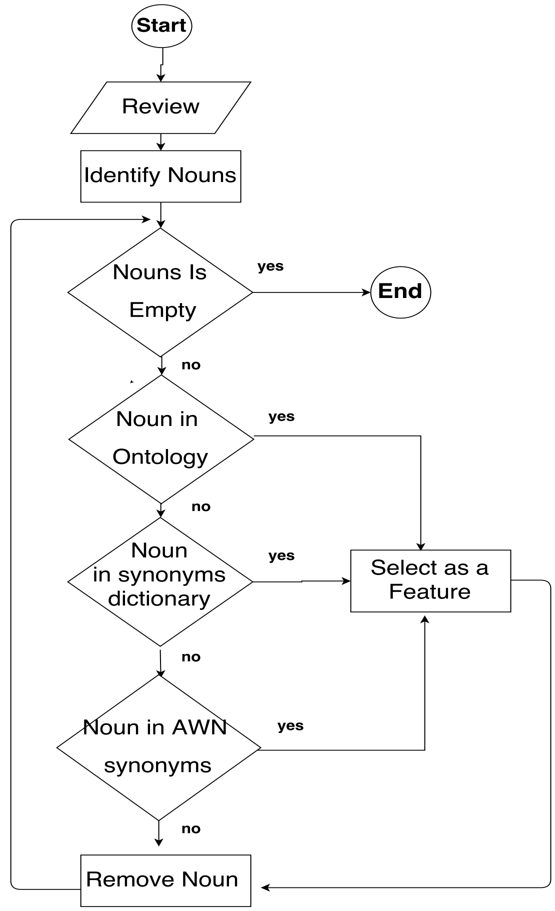


Fig. 3: Semantic Feature Identification Process.

TABLE I: Example of Feature Identification

review	الانترنت ضعيف التلفزيون حجمه صغير جدا السرير حجمه صغير دورة المياه مافيها شامبو ولا شيء فقط فوطه صغيره الغرفه نظيفه لكن كئيبه وصغيره
Part of speech	الانترنت/DTNNP ضعيف/JJ التلفزيون/DTNN السرير/DTNN حجمه/NN صغير/JJ جدا/NN المياه/DTNN حجمه/NN صغير/JJ دورة/NN فقط/RB مافيها/NN شامبو/NNS ولا/NN شيء/NN نظيفه/NN الغرفه/DTNN صغيره/JJ فوطه/NN وصغيره/NNP كئيبه/NNP لكن/VBP
Selected features	الانترنت, التلفزيون, السرير, دورة المياه, شامبو, فوطه, الغرفه

D. Feature Polarity Identification (Tuple Generation)

Using the extracted features and the lists of positive and negative words generated by previous phases, we identify the opinion orientation expressed for each feature. This step generates a set of tuples containing features and their polarities. In order to generate these tuples, it is necessary to obtain the words from around the feature. The words that are close to the feature can be obtained in different ways. The following three methods have been implemented to evaluate our solution:

- N-GRAM Before: this method obtains the N-GRAM words before the feature in the users review.
- N-GRAM After: this method obtains the N-GRAM words after the feature in the users review.
- N-GRAM Around: this method obtains the N-GRAM words before the feature and the N-GRAM words after the feature in the users review.

N-GRAM indicates the number of words near the feature that are to be selected in the polarity identification process. Also a Negative Words have been handled, the negation word usually reverses the polarity of the word in the sentence. Our proposed technique recognizes Negation words such as "غير ، ليس ، لا" and then reverses the opinion orientation. For example, the sentence, "الفندق غير جميل" conforms to the negation word "غير" then it is assigned the negative orientation, although "جميل" is a positive word.

E. Opinion Mining

The global polarity of a review is obtained by determining the majority of polarized features which our system already identified. If the major features are polarized as positive, then the global polarity is considered positive. Likewise, if the major features are polarized as negative, then it considered as negative. Otherwise the global polarity is considered as neutral.

V. EXPERIMENT AND DISCUSSION

A. Experiment setup

Since there was no tagged data in Arabic hotel domain, we collected the test reviews manually from a variety of related websites which have relevant to hotel domain. We crawled these reviews from different countries and three websites (www.tripadvisor.com), (http://www.booking.com) and (http://www.agoda.com) to find data related with the hotel domain. The total numbers of reviews that have been used are 890 reviews, 690 reviews were used for the sake of ontology extension and the rest 200 reviews, half of them are negative and the rest are positive were used for experiments. In order to measure the effectivnesss of the proposed model.

We manually tagged the reviews to collect the baseline results used to evaluate our proposed method. The details of the manual tagging are described as follows: The reviews have been shown to three educated annotators. They read the reviews and identified all features and associated polarities. According to features polarities they are classified into three categories: positive (1), negative (-1) and neutral (0). An example of the manual tagging is shown in Table II. Finally, the manual results and the output produced by our system are compared with each other. The following experiments have been conducted:

TABLE II: Example of Manual Tagging

Review	Manual Tagging
الفندق هادي ونظيف والاضاءة كانت جميلة	الفندق#هادي#1 الفندق#نظيف#1 الاضاءة#جميلة#1
الانترنت غير مجاني الفطور سئ	الانترنت #غير مجاني#-1 الفطور# سئ#-1

TABLE III: Feature identification accuracy

reviews	manually tagging	Proposed System Output	accuracy
positive	882	763	0.865
negative	757	655	0.866

1) Experiment1 measuring the feature identification accuracy: The aim of this experiment is to measure the accuracy of the correctly identified features (feature accuracy) using the manually tagged reviews. We take the labeled features as baseline in contrast to the results from the proposed method which described in section IV-B to obtain the number of features correctly identified feature accuracy. The results of this experiment are shown in Table III.

2) Experiment 2 measuring the Feature Polarity Identification and global polarity accuracy: During this experiment, two levels of accuracy measurements were tackled. On the first level, feature polarity identification where we applied different value of N-GRAM to identify correct polarity to correct feature. While the second level aims to measure accuracy of complete review. The three N-GRAM methods, which explained before, are used to compare the manual results and the output produced to obtain the number of features correctly classified (feature polarity identification accuracy) and the review global polarity (opinion mining accuracy). Different values for the N-GRAM parameter (between 1 and 4) have been used to discover the best result as presented in Table IV.

B. Discussion

The results are shown in Table IV have been divided into two different categories: the average accuracy of both feature polarity identification and opinion mining classification of the entire document for both positive and negative reviews. The results of the N-GRAM Around method for the opinion mining are shown in Table IV. As shown in the table, the best average success rate for the feature polarity identification process is obtained with N-GRAM = 4 with an accuracy of 72% in

TABLE IV: Feature Polarity Identification and Opinion Mining accuracy

		average feature polarity identification accuracy			opinion mining accuracy		
Method	Ngram	Positive Reviews	Negative Reviews	average	Positive Reviews	Negative Reviews	average
After	1	0.45	0.38	0.415	0.76	0.3	0.53
	2	0.56	0.48	0.52	0.94	0.71	0.825
	3	0.63	0.53	0.58	0.98	0.82	0.9
	4	0.66	0.56	0.61	0.99	0.85	0.92
Around	1	0.54	0.45	0.495	0.86	0.55	0.705
	2	0.65	0.57	0.61	0.99	0.88	0.935
	3	0.70	0.61	0.655	0.98	0.92	0.95
	4	**0.72**	**0.63**	**0.675**	**1**	**0.91**	**0.955**
Before	1	0.29	0.23	0.26	0.22	0.09	0.155
	2	0.38	0.33	0.355	0.55	0.22	0.385
	3	0.45	0.40	0.425	0.76	0.37	0.565
	4	0.50	0.44	0.47	0.86	0.59	0.725

the Positive Reviews and 63% in the Negative Reviews. This means that the feature based polarity calculated using 4 words before the feature and 4 words after the feature in the users review has achieved good accuracy. In fact, the worst results are obtained with N-GRAM = 1 with an accuracy of 54 in the Positive Reviews and 45% in the Negative Reviews. The best results for Opinion mining are obtained with N-GRAM=4 with an accuracy of 100% in the Positive Reviews and 91% in the Negative Reviews.

Table IV also shows the results obtained when using the N-GRAM After method. At first sight it will be noted that these results are worse than those obtained with the N-GRAM Around method. Here, the best average success rate for the feature polarity identification process is also obtained with N-GRAM = 4 with an accuracy of 66% in the Positive Reviews and 59% in the Negative Reviews. This means that the feature based polarity calculated using the next 4 words of the feature identified obtains good results. The best results for Opinion mining are obtained with N-GRAM=4 with an accuracy of 99% in the Positive Reviews and 85% in the Negative Reviews.

The results obtained with N-GRAM Before method are worse than those obtained with the "N-GRAM After method. More concretely, the best average success rate for the feature polarity identification process is also obtained with N-GRAM = 4 with a maximum accuracy of 50% in the Positive Reviews and 44% in the Negative Reviews. This means that the feature based polarity calculated using 4 words before of the feature identified obtains good results. The best results for Opinion mining are obtained with N-GRAM=4 with an accuracy of 86% in the Positive Reviews and 59% in the Negative Reviews.

Of the three proposed methods, the N-GRAM Around method is that which achieves the best results for both the feature polarity identification process and the Opining mining of users opinions in the Arabic language, obtaining accuracies of 67.5% and 95.5% in all Reviews, respectively.

VI. CONCLUSION AND FUTURE WORK

Arabic opinion mining is a challenging problem. It is concerned with analyzing the opinions that appear in users reviews, and determine whether these opinions are positive or negative. In this paper, a new methodology is proposed for feature-based Arabic opinion Mining. This approach is going through five different stages: Ontology and lexicon Development, Semantic Feature Identification, Polarity Identification, Feature Polarity Identification and Opinion Mining.

The main contributions of this work are: First, an ontology and lexicon development. Second, ontology-based feature identification, finally, three different configurable N-GRAM methods for feature polarity identification are proposed. These methods can be configured with different parameters to obtain the best polarity identification approach.

In spite of all the advantages and possibilities of the proposed method, it has several limitations that could be improved in the future. First, the proposed approach can be improved by incorporating opinion mining techniques based on machine learning. Second, since the current ontology is static and knowledge represented in it is not enough, it would be interesting to construct a semi-automatic ontology based on ontology learning techniques from the users reviews. Finally, we plan to apply the proposed approach in another domain such as product reviews.

REFERENCES

[1] A. Harb, M. Plantié, G. Dray, M. Roche, F. Trousset, and P. Poncelet, "Web opinion mining: How to extract opinions from blogs?" in *Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology*. ACM, 2008, pp. 211–217.

[2] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and trends in information retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.

[3] B. Liu, "Sentiment analysis and subjectivity." *Handbook of natural language processing*, vol. 2, pp. 627–666, 2010.

[4] R. Feldman, "Techniques and applications for sentiment analysis," *Communications of the ACM*, vol. 56, no. 4, pp. 82–89, 2013.

[5] A. Balahur and A. Montoyo, "Semantic approaches to fine and coarse-grained feature-based opinion mining," in *Natural language processing and information systems*. Springer, 2009, pp. 142–153.

[6] E. Cambria, B. Schuller, B. Liu, H. Wang, and C. Havasi, "Knowledge-based approaches to concept-level sentiment analysis," *IEEE Intelligent Systems*, no. 2, pp. 12–14, 2013.

[7] I. Peñalver-Martinez, F. Garcia-Sanchez, R. Valencia-Garcia, M. Á. Rodríguez-García, V. Moreno, A. Fraga, and J. L. Sánchez-Cervantes, "Feature-based opinion mining through ontologies," *Expert Systems with Applications*, vol. 41, no. 13, pp. 5995–6008, 2014.

[8] R. Studer, V. R. Benjamins, and D. Fensel, "Knowledge engineering: principles and methods," *Data & knowledge engineering*, vol. 25, no. 1, pp. 161–197, 1998.

[9] S. A. Morsy, "Recognizing contextual valence shifters in document-level sentiment classification," 2011.

[10] A. M. Azmi and S. M. Alzanin, "Aara–a system for mining the polarity of saudi public opinion through e-newspaper comments," *Journal of Information Science*, vol. 40, no. 3, pp. 398–410, 2014.

[11] J. B. Salamah and A. Elkhlifi, "Microblogging opinion mining approach for kuwaiti dialect," in *The International Conference on Computing Technology and Information Management (ICCTIM2014)*. The Society of Digital Information and Wireless Communication, 2014, pp. 388–396.

[12] S. B. Hamouda and J. Akaichi, "Social networks text mining for sentiment classification: The case of facebookstatuses updates in the arabic springera," *International Journal Application or Innovation in Engineering and Management*, vol. 2, no. 5, pp. 470–478, 2013.

[13] M. Abdul-Mageed, M. Diab, and S. Kübler, "Samar: Subjectivity and sentiment analysis for arabic social media," *Computer Speech & Language*, vol. 28, no. 1, pp. 20–37, 2014.

[14] N. A. Abdulla, N. A. Ahmed, M. A. Shehab, and M. Al-Ayyoub, "Arabic sentiment analysis: Lexicon-based and corpus-based," in *Applied Electrical Engineering and Computing Technologies (AEECT), 2013 IEEE Jordan Conference on*. IEEE, 2013, pp. 1–6.

[15] M. Elhawary and M. Elfeky, "Mining arabic business reviews," in *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1108–1113.

[16] A. El-Halees, "Arabic opinion mining using combined classification approach," 2011.

[17] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 168–177.

[18] M. N. Al-Kabi, I. M. Alsmadi, A. H. Gigieh, H. A. Wahsheh, and M. M. Haidar, "Opinion mining and analysis for arabic language," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 5, no. 5, pp. 181–195, 2014.

[19] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in *Proceedings of the 2008 International Conference on Web Search and Data Mining*. ACM, 2008, pp. 231–240.

[20] A. Muangon, S. Thammaboosadee, and C. Haruechaiyasak, "A lexiconizing framework of feature-based opinion mining in tourism industry," in *Digital Information and Communication Technology and it's Applications (DICTAP), 2014 Fourth International Conference on*. IEEE, 2014, pp. 169–173.

[21] E. Bjørkelund, T. H. Burnett, and K. Nørvåg, "A study of opinion mining and visualization of hotel reviews," in *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services*. ACM, 2012, pp. 229–238.

[22] L. Zhao and C. Li, *Ontology based opinion mining for movie reviews*. Springer, 2009.

[23] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades, "Ontology-based sentiment analysis of twitter posts," *Expert systems with applications*, vol. 40, no. 10, pp. 4065–4074, 2013.

[24] J. M. Ruiz-Martínez, R. Valencia-García, and F. García-Sánchez, "Semantic-based sentiment analysis in financial news," in *International Workshop on Finance and Economics on the Semantic Web (FEOSW 2012)*, 2012, p. 38.

[25] L. A. Freitas and R. Vieira, "Ontology based feature level opinion mining for portuguese reviews," in *Proceedings of the 22nd international conference on World Wide Web companion*. International World Wide Web Conferences Steering Committee, 2013, pp. 367–370.

[26] M. S. Chaves, L. A. de Freitas, and R. Vieira, "Hontology: A multilingual ontology for the accommodation sector in the tourism industry." in *KEOD*, J. Filipe and J. L. G. Dietz, Eds. SciTePress, 2012, pp. 149–154.

[27] H. Knublauch, R. W. Fergerson, N. F. Noy, and M. A. Musen, "The protégé owl plugin: An open development environment for semantic web applications," in *The Semantic Web–ISWC 2004*. Springer, 2004, pp. 229–243.

[28] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining." in *LREC*, vol. 10, 2010, pp. 2200–2204.

[29] R. Mihalcea, C. Banea, and J. M. Wiebe, "Learning multilingual subjective language via cross-lingual projections," 2007.

[30] M. Abdul-Mageed, M. T. Diab, and M. Korayem, "Subjectivity and sentiment analysis of modern standard arabic," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*. Association for Computational Linguistics, 2011, pp. 587–591.

[31] A. Mourad and K. Darwish, "Subjectivity and sentiment analysis of modern standard arabic and arabic microblogs," in *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis*, 2013, pp. 55–64.

[32] A. M. Shoukry, "Arabic sentence level sentiment analysis," Ph.D. dissertation, The American University in Cairo, 2013.

[33] M. Althobaiti, U. Kruschwitz, and M. Poesio, "Automatic Creation of Arabic Named Entity Annotated Corpus Using Wikipedia," in *Proceedings of the Student Research Workshop at the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Gothenburg, 2014, pp. 106–115.

[34] K. Toutanova, D. Klein, C. D. Manning, and Y. Singer, "Feature-rich part-of-speech tagging with a cyclic dependency network," in *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*. Association for Computational Linguistics, 2003, pp. 173–180.

[35] L. Abouenour, K. Bouzoubaa, and P. Rosso, "On the evaluation and improvement of arabic wordnet coverage and usability," *Language resources and evaluation*, vol. 47, no. 3, pp. 891–917, 2013.

The Information-Seeking Problem in Human-Technology Interaction

Mohammad Alsulami
College of Computer Science and Information Systems,
Najran University,
Najran, Saudi Arabia

Asadullah Shaikh
College of Computer Science and Information Systems,
Najran University,
Najran, Saudi Arabia
and
Faculty of Computer Science and Information Technology,
Institute of Business and Technology (IBT),
Karachi, Pakistan

***Abstract*—In the history of information-seeking, the intention of a query and the posed query have some level of distance between them. Because human query-responders are innately connected to times and trends and have the ability to understand natural language and human intention, they have often been the idealistic sources of knowledge-direction. As the quantity of depth of knowledge of humanity grows, technological systems have sought to utilize natural language, both spoken and written, as a format of accepted queries. Modern works seek to improve such systems utilizing distance-metrics of literal queries to understood questions with maps to knowledge-bases. However, these methods do not often take into account the value of information in terms of query interpretation for mapping and as such may have identifiable limitations compared with human responders. In this paper, a model for information value is proposed and existing works in speech and query recognition are discussed relative to their considerations of information value.**

***Keywords*—Information seeking; seeking problem in HCI; HCI Information seeking.**

I. Introduction

Gathering information has been an integral human task throughout all of history and the modern world is no exception to this. When asked the answer to a particular query, someone being interviewed may be able to venture an answer. However, in order for a technological system to respond to such a query, there are many factors that must be considered that might almost be automatically taken into account by a human responder. For example, the coherency and legitimacy of a question may be determinations that happen quite unconsciously. Even corrections to queries that may be interpreted as incoherent and illegitimate may be corrected. Such corrections are hereafter often referred to as interpretations because that is exactly what they are. All human interaction and response in way of communication of information can be seen as a transmission and interpretation process. Perhaps what is most important to consider when identifying the nature of a question is what an information-seeker intended to ask or, rather, what he intended to find.

As human communication in the form of querying is most logically-performed in the medium from which questions initially arise, that of natural language, it seems clear that the most direct and sound means of making and interpreting queries would happen in the form of natural language. However, natural language of humans tends to take two forms with almost equally-difficult challenges: spoken and written. Spoken language in the form of interpreted sound waves operates on such a challenging level that most technological applications choose to transform it into written language, which can be handled more compactly with specific construction. One of the major challenges in handling this aspect of human communication is that of correctly "hearing" what someone has said. Human capacity of correcting, distorting, and exaggerating what has been said is so profound that most people have heard entire sentences that were never actually stated! As such, it may be able to be stated that we as humans respond to what we "think" we hear and not what was actually said. Additionally, when responding to queries from other people, many human beings will construct the query within their own minds and perform many evaluations, distortions, and interpretations, perhaps even going so far as to rephrase the query in an entirely different (and hopefully desirable) way. Thus, the human ability to interpret the intention of a spoken query is quite profound, though modern advances in voice recognition, treated herein, have made great steps at approaching such a level of ability in technological applications. These devices though, it cannot be ignored, still translate from speech to text, thus transitioning into the other domain of human communication.

Written communication is perhaps arguably the clearest and most readily-processed form of human communication. In fact, "with difficult messages, both persuasion and comprehension of persuasive material were found to be greater when the message was written, compared to videotaped or audiotaped [4]". Even still, it cannot be ignored that intention may be lost if incorrect statement or interpretation of a query is made when searching for a particular piece of information. Herein the ways in which intention is attempted to be extracted from queries and the nature of information-seeking as it pertains to the value of sought-after information are discussed. Additional discussion includes understanding the gap that exists between a human responder and a technological responder in terms of function and ability and the way in which the "distance" between what has been requested and what is understood to have been requested. The latter factor has been addressed by many existing works in the domains of speech recognition and search engine implementation.

However, many of these systems attempt to pigeonhole an information-seekers queries using question-answer (QA) databases and other forms of context-limitation in order to ensure that an information-seekers queries are translated into something believed to be meaningful by the creators of such knowledge-bases. This is the approach taken in [1], where "we use information on the internet in the QA domain to find all possible question patterns, then use it to correct the queries that are partially recognized by speech-recognition software [1]". In this work, the value of information is considered as an extension to the simple aspect of quality of information and discussion is presented on how limiting systems and technologies attempt to produce value by distilling queries into a finite realm of knowledge, much like seeing a world full of nails when possessing a toolbox with only a hammer. Limiting approaches in interpreting queries, while pretending to be open oracles of knowledge, attempt to pigeonhole an information-seekers intended information-target into the knowledge-base of the oracle. Because the intention of a query lies within the information-seeker and the knowledge of humanity is the product of a history of communication, we first begin with a historical treatment of the history of searching and a background of the existing works in query-recognition.

II. Background: A Historical Perspective of Searching

In the world of modern technology, information is the currency. Todays world is about living and breathing vast amounts of information, ranging from simple data like emails to complex data such as answers to deep technical questions. It is obviously-clear that more searchable information exists now than two decades ago by a factor of many, many times. Perhaps at one point it might have been the task of researchers and even everyday people to put forth great effort in even attempting to locate the information that they sought. Perusing libraries, pouring over collections of statistics, seeking personal interviews, and even undertaking private investigations might have been necessary just to obtain information related to that which was being sought. At that time, finding too little information may have often been the norm. Those days are, for the most part, gone.

Today, one can access so much information with so little effort that the quantity can be mindboggling without even attempting to gather so much. One search of a popular search engine can yield so much data that the whole of the worlds libraries fifty years ago would collectively seem like an abbreviated encyclopedia in comparison. This information overload can make even the simplest searching tasks much more complex. The problem now is how to reduce the information that we collect to only the pertinent information that is desired. Maybe finding movie times at one point required a short stroll or a phone call. Now, a simple search for movie times on a smart phone might lead to confusion about what which of the multitudes of theatre results are the ones that match our intended destination. With this switch in paradigm new problems have arisen while lingering factors from the bygone days of searches in information shortage still remain.

It might be surprising to some that freedom comes at a cost. This statement may also lead to the possibly equally-surprising question: What kinds of freedom are there in searching for information? The answer to this question is quite simple: multitudes. Before delving into the academic nature of such freedom, it can be helpful to under-stand a brief history of what freedom means in terms of information-searching and why it has evolved with technology and culture. Fifty years ago, if one wanted the history of the rule of succession of British Empire, there were many choices available. The simplest choice might have been a trip to a reasonably-sized library, possibly of public, school, university, or private nature, and a consultation of a textbook of appropriate title. Perusing the index of such a textbook might or might not have yielded a reference to the information that was sought. It could and can be concluded that such a textbook was arranged with the intent on making the knowledge contained within reasonably complete and reasonably accessible.

However, it should also be observed that such assurances are based on the authors understandings of societal and cultural norms regarding the processes of searching for information as well as their own personal opinions on how information should be arranged and made available for search. Thus, searching multiple textbooks for the same knowledge might have led to similar ways of searching or perhaps completely different ways. As to the effectiveness of such a search, this vastly depended on ones own personal means and concepts of search compared with those arranging the information being sought. Perhaps the desired textbook was in the library, but finding this textbook may have required yet another search of a completely different nature, such as through a card catalog. This system may have been arranged in a completely different fashion than that of the organization of the desired textbook itself. If one knew the title of the textbook, finding it may have been relatively easy. If the title of the textbook was unknown or one did not know if such a textbook even existed and was attempting to determine the existence and location of such a textbook, this put the searcher at the mercy of the textbook-locating system. Of all textbook-locating systems, the one that is perhaps the closest to modern concerns was also often the most powerful: someone who knows, be they scholar, professor, or librarian.

Dating to ancient times, persons of knowledge have not only been the repositories of important knowledge, but also the means of seeking such knowledge. Wise-men, scholars, and teachers have ever played an important role in society. If they did not know the desired knowledge themselves, they could likely direct an information-seeker in an appropriate direction. Perhaps none was so famous a person in directing those to desired information than "the oracle" from which many modern information systems derive a basic term in describing their natures. One needed only ask the oracle and the desired information would be given or direction would be given that would reveal the desired information. Perhaps this "reveal rather than teach" approach allowed for a certain degree of information-isolation to the most appropriate and knowledgeable containers, a concept that also deserves academic treatment. Whatever the case, this human source of knowledge-direction had a crucial benefit over mechanical and artificial means of information search: the ability to speak ones own language.

It is hardly-debatable that many people of the past and of the present would almost exclusively prefer to query another human being regarding where to find certain information being sought. Modern call centers paired with automated hotlines are clear examples of such preference. Perhaps there is a clear emotional component of human-human interaction in this preference. However, if asked for a reason for preferring to query another human being, the logical response that someone might simply give is that it is faster than using a technological system. Modern technological systems attempt to mimic this behavior by allowing natural language queries that attempt to understand what an information-seeker is intending by literal interpretation and cross-matching of query terms. Several of the methods proposed for improving modern query and informational systems are discussed herein.

One such method is that of the RSVP system proposed in [1] that builds a QA knowledge-base of query questions by collecting a very large set of existing questions form QA websites. The authors of such work admit that there is very low coverage of all possible questions and propose that questions can be “clustered” into patterns for 99% coverage of all possible types of questions. This is the means in which they believe that can improve speech-recognition software in the QA domain. The methodic approach taken in [1] is that of attempting to minimize the distance between an interpreted query and the intended query and create a similar minimization to the database of QA question structures available. Related to the attempt of determining differences between queries and intended queries is the work in [2] in which the authors show that “information distance is a universal cognitive similarity distance [2]” and state that “information distance between individual objects is needed in pattern recognition where one wants to express effective notions of ‘pattern similarity’ or ‘cognitive similarity’ [2]”. Many of the same authors of this work also attempted to verify information-evolution algorithms based on genetic concepts through analysis of chain-letters in [3].

This relates in how information interpretation and meaning can evolve with time in response to trend. The algorithm proposed by the authors of [3], while originally developed for genomes, was applied in the realm of languages for detecting plagiarism in student assignments. This concept relates to how queries can be posed in different ways and mapped to similar information. In similar means to [1], [5] proposed a plagiarism-detection algorithm based on “a fundamental question in information theory and in computer science” on “how to measure similarity or the amount of shared information between two sequences [5]”. For performing this measurement, the authors proposed a metric based on Kolmogorov complexity and claim to have proven it to be universal. This, again, relates to posed queries and intended information in that evolution of similar intention is analyzed for similarity with trend identified. In terms of search-engine-oriented distance differences as a means of mapping queries [6], presents a theory of similarity between words and phrases based on information distance and Kolmogorov complexity. The authors therein present applications in hierarchical clustering, classification, and language translation based on the use of search results from a popular search engine based on the information trends of millions of independent web-knowledge-providers. The method proposed by the authors does take into account trending of information as the web evolves based on the evolution of search engine results from scouring the Internet over time.

In the work of [7], the authors attempt to correctly anchor in time and space a number of entities, implying a labelling system to reach such entities as factual objects. This may lead to certain prejudices of organization and labelling towards the designers ways of thinking. In contrast [8], proposes and tests a parameter-free algorithm that “would limit our ability to impose our prejudices, expectations, and presumptions on the problem at hand, and would let the data itself speak to us [8]”. The authors show that their approach is competitive or superior to state-of-the-art approaches in “anomaly/interestingness detection, classification, and clustering” utilizing empirical tests. This approach seems to allow for more independent expression of information-seeking that contradicts the more rigid means of many of the other works. The work of [9] relates to the continual seeking of proper metrics for information distance by addressing the fact that “the universality theorems for normalized information distances were only proved in a weak form.” The authors of this work expose some of the issues of information distance theory in attempting to establish a better means of measuring information distance. This is similar to the work of [10] in which the authors study “a new class of distances appropriate for measuring similarity relations between sequences, say one type of similarity per distance [10]”. In a completely different way [11], with its appropriate title of an incorrectly-interpreted voice-query “How to wreck a nice beach you sing calm incense”, takes the approach of solving the problem of speech-recognition systems inappropriately-ordering hypotheses for similar-sounding phrases. In their work, the authors proposed “a supplementary method for ordering hypotheses based on Commonsense Knowledge [11]”, the system they developed by “filtering acoustical and word-frequency hypotheses by testing their plausibility with a semantic network derived from 700,000 statements about everyday life.”

In response to these methods, herein a model is proposed of information value as a means of mapping queries to information. This model attempts to account for factors that may be overlooked by many of the literal-matching-and-correction schemes of current works.

III. Methods: Quantifying a Search

While quality of information is typically paramount on the mind of the searcher, the gaining of such excellent information is ultimately determined useless if the time spent obtaining such information were to exceed the reasonable scope of time when such information would be considered valuable. For example, finding out the time of showing of a movie in a theatre an hour after the movie ends is worthless. This scope may be lessened by the information-seekers own personal preference, meaning that the value of the sought-after information may be eclipsed by other demands on the information-seekers time-resources. For example, an information-seeker may be attempting to review products while standing in a store and eventually forgo obtaining additional information, or any information at all, about products to be purchased if the information-seeker determines that the benefits of such additional information are not worth the effort of obtaining it. This viewpoint of information-seeking leads one to the mindset that what is of importance in obtaining information is a matter of efficiency. Using the common definition of efficiency, referring to the ratio of output to input in a process, we can frame the efficiency of a process as

$Efficiency = OutputGained/InputPresented.(1)$

In considering the concept of efficiency, it is a basic economic fact that in prior centuries materials were extremely expensive compared with the cost of human time. Many important achievements of mankind have been historically achieved through raw effort of input of human time with most expense being considered in terms of the materials needed to reach such achievements. Thus, the price of food and travel may have been in the past of considerable more concern in the process of seeking of in-formation than the time spent by a person. The reverse of this paradigm has almost exactly corresponded with the paradigm reverse of information shortage versus overload. In the modern technological and business world, it is a known fact that fewer commodities are more expensive than human time. The correspondence of material-availability and information-availability is difficult to ignore. With so much information being available for ready-access, the definition of efficiency in information-seeking can be redefined as

$Information - SeekingEfficiency = Quality/SearchTime.(2)$

Considering that most human beings will likely not seek quality information that has subjectively-little value, what remains to be considered is how to determine the subjective value of sought-after information as it relates to quality. According to [12], the "value of search results as a whole, a utility measure, was found to be the best single measure of interactive information performance (success) among the 20 measures selected for study [12]".

The quality of information can be considered constant relative to the actual information desired. One reason for this is because at the time a query is posed, the ideal answer exists pertaining to the time that the query was posed. Posing the same query later results in searching a different knowledge-base than existed at the previous time. Thus, at any given time, there exists a response to any query that could be considered ideal for such a query given the knowledge available at the time. The quality of the response to a query at a given time is relative to the knowledge available at that time. However, the value of a particular piece of information relative to a query could depend greatly upon the quality of the information given changes to the knowledge base. As such, it seems clear that quality of information, again always relative to a particular query, can change with the knowledge-base, which can change with time. Given that quality has been established for a particular query at a particular time, it can be presumed that the mapping of the query to the knowledge base may change with time, leading to a single query mapping to different pieces of information at different times. For example, a query for the current movies showing at a particular theatre would result in very different information for equal qualities at different times. Thus, there must be a component of information value that relates quality to time based on the information-seeker and his particular query, a kind of pertinence. For example, the best way to treat a particular medical condition might be a piece of information of particular quality with value that changes with time based on the information-seeker that would warrant a level of pertinence. Even pertinence, however, is not enough to determine whether information has value at a particular time.

Even a pertinent piece of information may not be of significant value if it is not what is actually desired. The trend of information can vary personally and societally with time. A particular piece of information may be highly-valued during a particular scope of time simply due to its popularity as a trend. Popular search results on major search engines make good examples of this. Thus, when establishing the value of a piece of information, the trend of simple human interest must be taken into account. Tracing this trend can lead to another piece of information in itself, which is how preference for information evolves with time. Sometimes the value of information depends heavily upon the specific time that the information is being sought. Five days before a movie is to show at a theatre, the information regarding show times for that theatre may be of some value to the collective audience that will eventually see the shows on the destined day. This information will likely be of increasing value as the destined day approaches and then be of basically no value the next day. This means that the trend of information can grow as a time approaches and then fall to nothing shortly after that time has been reached. Under these premises, the definition of information value could be stated as

$InformationValue = Quality*Pertinence*Trend.(3)$

The definition in (3) establishes a model for information value though it does not yet determine what exactly constitutes quality of information. Basic academic thought might lead one to understand that what is normally intended with quality is correct information that is of significant enough quantity. In terms of the value of this quality of information, the pertinence and trend of information should act to determine its value. For example, an academic article or textbook entry might be of the highest quality for a particular query though an information-seeker might not desire this quality of information and thus a trend might be established towards a particular lower quality information that is either not as correct, not as voluminous, or both. This could explain the popularity of Internet search engine results and online encyclopedia articles being sought after. A particular model for information quality could be framed as $InformationQuality = Correctness * Quantity.(4)$

Combining definitions (2), (3), and (4) yields a more complex model for framing information value once rearranged as

$InformationValue = Correctness * Quantity * Pertinence * Trend.(5)$

The definition in (5) still does not seem to take into account the value of information as it relates to efficiency. Even given the availability of correct information of adequate quantity that is pertinent to the information-seekers query that has an established trend of desire, an information-seeker may still forgo obtaining such information. Returning to (3) and substituting efficiency for quality, (5) can be rewritten as

$InformationValue = Correctness * Quantity * Pertinence * Trend/SearchTime.(6)$

With so many of the terms in (6) depending on time, society, and personal consideration, one might be led to wonder how an information-seeker could possibly find the information being sought. The authors of [7] attempted to address this with a project undertaking based on ensuring that "entities, facts, and events are anchored in both time and space [7]". Perhaps there are some simple facts to consider in this. Historically, much of the information available at any given time has been through the devices, means, and trends of the day. This means that the pertinence and trend of the information are established within the resources of the day. Likewise, the information on how to find information was established by trend such that the correctness and quantity of such information to find other information was obtainable with reasonable search time. This could have occurred through how the people of the day were taught. By establishing the pertinence of such information-seeking sources, it may be that the value of information through history could have been carefully controlled and monitored on a societal level. In the modern world of search engines and QA databases, data tend to live forever, meaning that the evolution of data in the form of availability as a function of decay and growth relative to time is distorted.

In the modern Internet world, querying for information has been revolutionized and evolved by the sciences behind search engines and QA databases. The widespread, growing use of information in almost every aspect of life and business has driven these developments. As might seem a natural evolution, information-seeking technology has evolved towards natural-language systems allowing for queries in spoken and written form. These systems typically operate under the presumption that an information-seeker understands exactly the information he is seeking and that his query is worded to be directly-mapped to the intended information. This approach does not take into account many aspects of value, which will be discussed later herein. Although technology has managed to preserve and even widen the availability of data that would have previously decayed or become unavailable, such as recorded music and records, the value of data continues to change with society and the times. Modern information-seekers can find records, information, and cultural data from fifty years ago, which is an enormous contrast from the availability of one-hundred-year-old information at a point in time fifty years ago. What remains to be discussed is how modern query-response systems deal with this fact by attempting to understand the right questions and how those systems do or do not take into account pertinence and trend of information.

IV. Evaluation: Having the Right Questions Understood

Returning to the concept that an information-seeker may have many ways of attempting to obtain information, considering pertinence and trend to have been established by the information-seeker and utilizing (6) leads to the conclusion that there is a balance between the quality of information being sought and the search time. The pertinence and trend of the information may be considered as a kind of filter to limit the information-seekers search while the efficiency of search drives the information-seekers actions. However, the medium of querying for information, such as a library in the past or the Internet in the present, may not have been established with clear defining lines to allow pertinence and trend to be taken into account. This may lead to a misunderstanding of what the information-seeker is asking that might be obvious to a human oracle whose knowledge-base of information-seeking is established in the times of the information-seeker. One clear area of difference between an Internet search and an interview-search is the recognition of language in queries.

A human information-seekers own knowledge of how to seek information is firm-ly rooted in language, which is firmly rooted in the times. Thus, the diction and tendencies of the information-seeker may well be heavily-affected by the society and age in which he finds himself. This seems to lead to a point that language should be considered as a kind of knowledge-base that itself is dictated by information values. When considering the means of querying for information, human language is a natural choice as it is the basis in which thought and writing are established and likely the principally-understood form of most original queries. Considering that a human oracle whom an information-seeker may query has established values in language based on his upbringing and societal interaction, a human oracle may be able to interpret an information-seekers query utilizing his own value system of language. Whether reading a query or hearing it spoken, a human oracle has an advantage over a designed or programmed one: the ability to interpret a query based on his own querying ability.

When designing an oracle, a big challenge is one that changes its interpretations of queries with the values of information-seekers. As a programmed oracle, such as a search engine, does not inherently have the ability to establish itself in the times and determine its own values of information, it must rely on being given this information or deriving it. A major search engine may be able to notice trends in querying and information quite readily. Perhaps what the search engine may lack is the ability to understand the motivations behind a query. These motivations are likely established within the times and society of the information-seeker. Even a query of history comes from the present time. When communicating with a search engine, the interpretation of a query needs to be based on what the information-seeker intended to find. This can be rather difficult on several fronts of understanding.

When human information-seekers search for information using a search engine or other search technology, they often do not utilize the same protocols as they would if querying other humans. As such, queries for information posed to technological systems are often incorrect grammatically and structurally in a way that might be considered insulting to a human responder. An example might be "famous action movie 1966", which sounds more like a trivia statement than a query. One might be led to believe that the intended information here is the name of a particular famous action movie released in 1966, however the intention to this apparent clarification question is somewhat vague. Likewise, a query of "blue soccer ball" is more of a statement than a query and thus the intention is very unclear. However, it seems from the general usage of search engines in modern times that information-seekers have adapted to the way in which search engines operate and intentionally pose such queries with a belief that certain information will be returned. For modern speech-recognition and QA query-matching systems, this mode of operation can be potentially defeating of natural language processing and correction methods that they attempt to utilize. Several such current methods are herein discussed and analyzed.

The RSVP system proposed in [1] that builds a QA knowledge-base of query questions by collecting a very large set of existing questions form QA websites attempts to establish value by assuming a very broad range of possible questions that an information-seeker may ask and then literally-correcting queries to fit these questions. This is a bit like trying to record the thoughts of millions of people in order to learn a language. While trends of information may be captured if they evolving QA knowledge-bases are consistently kept in-sync with the times, the pertinence to individual queries is potentially lost by pigeonholing the information-seekers ways of posing queries. Additionally, this type of system can greatly suffer from ambiguously-posed queries that lack substantial qualification. The proposed system does attempt to equate related means of posing a query about a single piece of knowledge though perhaps in a crude, collective way. The statement of the authors in [2] that "information distance is a universal cognitive similarity distance [2]" as the basis of their efforts fails to take into account trends and localized pertinence of information. An example of such might be the profanities of one culture being typical speech in another. In their related work in [3] based on evolution of meaning and communication in chain-letter-generations, the authors do account for trending over time.

Thus, the intention of posed queries of similar semantic nature may evolve as time progresses, much as was noticed in the analyzed chain-letters. Works [1] and [5] used the application of plagiarism-detection in determining an appropriate way to measure similarity and shared information between two sequences of information. This literal interpretation of distance rather than a value-based interpretation of distance may be able to correct and identify certain semantic variations and facilitate a many-to-one mapping of knowledge with a certain degree of potential trending, but such as an approach may also suffer in being able to distinguish pertinence of information requests that are distinct in some way. The more vaguely a query is posed, the less effectively this method will likely be to discern subtle differences. The authors of these works proposed a metric based on Kolmogorov complexity and claim to have proven it to be universal. This, again, does not account for localized and temporal differences in trend and pertinence.

These approaches lead to the efforts of [6] regarding search-engine distances for mapping queries to information. The authors here use search results from a popular search engine based on the information trends of millions of independent web-knowledge-providers, which does in fact take into account trending of information as the web evolves based on the evolution of search engine results from scouring the Internet over time. This approach is unique in that it takes into account the ever-general population of online-user culture in regards to information-seeking. As an information-seeker is likely also an avid Internet user in modern times, these two cultures are likely closely-enough synchronized to allow for generality of trend and pertinence for time and space to be taken into account in interpreting an information-seekers query. This allows the establishment of value of information pertaining to a query given that enough web contributors make the associated mappings and intention-relationships. In a similar, but more rigidly-structured approach, the authors of [7] attempted to correctly anchor in time and space a number of entities, implying a labelling system to reach such entities as factual objects. This may lead to certain prejudices of organization and labelling towards the designers ways of thinking.

Additionally, such a controlled system that is not allowed to vary with trends and localized pertinence may be unable to maintain validity if significant deviation in knowledge-values occur on a broad enough spectrum. The approach of [7] is contrasted with that of [8] that proposes and tests a parameter-free algorithm that "would limit our ability to impose our prejudices, expectations, and presumptions on the problem at hand, and would let the data itself speak to us [8]". The authors show that their approach is competitive or superior to state-of-the-art approaches in "anomaly/interestingness detection, classification, and clustering" utilizing empirical tests. The effort here in allowing data to establish its own trend and pertinence is appropriate in terms of establishing value as it separates the labelling of searching for information from the underlying knowledge and allows for free-form association without utilizing the authors own personal knowledge-mapping concepts in the method. This approach seems to allow for more independent expression of information-seeking that contradicts the more rigid means of many of the other works. In returning to literal, distance-oriented metrics, [9] sought to establish proper metrics for information distances. The authors of this work expose some of the issues of information distance theory, but do not make direct conclusion of the need to establish underlying value of information in terms of determining information relationships.

The authors of [10] took similar direction in studying a new class of distances for measuring similarity. This work makes progress in literal-interpretations and has many of the same shortcomings of other related works. In contrast to the many literal methods analyzed, the authors of [11] take a direct, value-oriented approach to correcting voice-recognition queries by strongly sorting hypothesized query meanings and matching them against a system appropriately named Commonsense Knowledge. This designers of this method strived to eliminate the need for information-seekers to self-correct their own queries, creating more seamless and believable voice-only searching capability. What is particularly interesting about this approach is the somewhat indirect means of establishing information value by determining if a query "makes sense". This is interesting because a query "making sense" implies a general and localized understanding of information trend and pertinence. The fact that the authors used an extensive, temporally-oriented database of highly-pertinent phrases from "everyday life" inherently makes this system the strongest of those reviewed in terms of value-orientation of information-seeking.

V. Discussions

In the modern world of information-seeking, technological methods have evolved considerably from the indexes, libraries, and interviewing methods of old. Modern information-seekers have many more degrees of freedom in regards to how they interact with searching technology leading to many advantages and disadvantages. The trend of information-locating systems towards accepting queries in natural language forms close to the thoughts driving such queries of information-seekers has allowed faster, more-direct, and more-specific searching processes than the statically-arranged information-searching systems of history. Given the increasing excessive abundance of information, these processes are as much about filtering-out unwanted information as locating the wanted information. As such, it has become increasingly-important to understand the intended information that an information-seeker desires rather than assuming that by literally-interpreting a query. This is a matter of establishing the value of information to an information-seeker relative to the provided query and any other information known about the context of the user. Correspondingly many modern voice-recognition-based searching methods have been oriented to perform corrections based on established bounds for context in order to limit the degrees of freedom that a potentially-incorrect query may have to establish the appropriate query mapping to the information of value being sought.

This speech-to-text conversion has a similar counterpart in text-to-text conversion where calculated assumptions are made based on the context of the information-seekers relative location in the global knowledge-base. However, this attempt at arbitrarily-confining the pertinence and trend of sought-after information can inappropriately limit the degrees of freedom available to a user through the querying means. Establishing the correctness of information to a query by ensuring literal map-ping of query terms to a set of previously-mapped questions that supposedly have pertinence and trend toward information as a means of establishing value of information relative to a query assumes that the information-seeker's intentions and means of querying are similar to those used to derive the underlying mappings. Additionally, this method is complicated by the way in which information-seekers will pose queries to technological systems compared with similar query situations of actual humans.

When an information-seeker poses a question to another human, an established protocol requiring a certain, reasonable construction of the query as a question or command has already been instilled in both information-seeker and query-responder. Because the responder has an understanding of construction of queries and the ability to sense the supposed intention of the information-seeker in terms of the value of the information being sought relative to the information-seeker, the human responder can make corrections to misheard or misstated queries and even completely restate the query in order to arrive at the intended informational response. Technological systems do have not the luxury of being treated with human-human protocols and as such are often queried with incomplete, incoherent phrases that are often specifically-tailored to result in certain responses from technological systems and therefore may circumvent the ability of such systems to provide potentially-improved responses based upon complete and properly-constructed queries per human-human communication standards. Current technological QA systems lack the breadth of deriving intention from posed queries and instead rely upon literal interpretations of spoken or written queries with corrections being made by means of cross-examining existing databases of posed queries and mapping such queries to potentially-valuable information based on correct-query similarity. In this work, the current means of performing such context-limiting searches were reviewed with the perspective given of their attempts at providing informational-value through presuming pertinence and trend based on the context and content of posed queries and focusing efforts on correctness of mapping to information through creative, literal, mathematical interpretations of query content. The approach at identifying how different "what was said" was from "what it heard" seems at times in current technology to be identifying "what is your purpose here". Perhaps this is the price that is paid for specialization of knowledge-bases and a lack of understanding in technological devices of human intention. Two of the systems reviewed that allowed free-form interpretations of information based on evolving trends in daily life were perhaps the best oriented at maintaining information value in query responses through a loose connection between information sought and the means of querying. Whatever the case may be, great advancements have been made in natural language searching technologies, be they spoken or written, though there is thankfully still "Room" for improvement.

VI. Conclusions

In this paper, we have proposed, a model for information value and existing works in speech and query recognition for information value. In practical, information-seeking, the purpose of a query and the posed query has some level of distance. We have addressed the value of information in terms of query interpretation for mapping and as it may vary in limitations compared with human responders.

References

[1] Yang Tang, Di Wang, Jing Bai, Xiaoyan Zhu, and Ming Li. Information distance between what I said and what it heard. Commun. ACM 56 (July 2013), 70-77.

[2] Bennett, C.H., Gcs, P., Li, M., Vitnyi, P., and Zurek, W. Information distance. IEEE Transactions on Information Theory 44, 4 (July, 1998), 14071423.

[3] Bennett, C.H., Li, M., and Ma, B. Chain letters and evolutionary histories. Scientific American 288, 6 (June 2003), 7681.

[4] Chaiken, Shelly; Eagly, Alice H. Communication modality as a determinant of message persuasiveness and message comprehensibility. Journal of Personality and Social Psychology, Vol 34(4) (October 1976), 605-614.

[5] Chen, X., Francia, B., Li, M., McKinnon, B., and Seker, A. Shared information and program plagiarism detection. IEEE Transactions on Information Theory 50, 7 (July 2004), 15451550.

[6] Cilibrasi, R. and Vitnyi, P. The Google similarity distance. IEEE Transactions on Knowledge and Data Engineering 19, 3 (March 2007), 370383.

[7] Hoffart, J., Suchanek, F.M., Berberich, K., and Weikum, G. YAGO2: A spatially and tempo-rally enhanced knowledgebase from Wikipedia. Artificial Intelligence 194 (January 2013), 2861.

[8] Keogh, E., Lonardi, S., and Ratanamahatana, C.A. Towards parameter-free data mining. in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press, New York, 2004, 206215.

[9] Li, M. Information distance and its applications. International Journal on the Foundations of Computer Science 18, 4 (August 2007), 669681.

[10] Li, M., Chen, X., Li, X., Ma, B., and Vitnyi, P. The similarity metric. IEEE Transactions on Information Theory 50, 12 (December 2004), 32503264.

[11] Lieberman, H., Faaborg, A., Daher, W., and Espinosa, J. How to wreck a nice beach you sing calm incense. in Proceedings of the 10th International Conference on Intelligent User Interfaces (2005), 278280.

[12] Su, Louise T Su. Value of search results as a whole as the best single measure of information retrieval performance. in Information Processing & Management, Volume 34, Issue 5, September 1998, 557-579.

Virtual Heterogeneous Model Integration Layer

Muhammad Ali Memon
School of Information Technology,
Shaheed Benazir Bhutto University,
Shaheed Benazirabad, Pakistan

Khizer Hayat
Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology,
Hyderabad, Pakistan

Asadullah Shaikh
College of Computer Science and Information Systems,
Najran University,
Najran, Saudi Arabia

Mutiullah Shaikh
Faculty of Electrical, Electronic and Computer Engineering,
Mehran University of Engineering and Technology,
Jamshoto, Pakistan

***Abstract*—The classic way of building a software today simplistically consists in connecting a piece of code calling a method with the piece of code implementing that method. We consider these piece of code (software systems) not calling anything, behaving in a non deterministic way and providing complex sets of services in different domains. In software engineering reusability is the holly grail, and specially the reusability of code from autonomus tools requires powerful compostion/integration mechanisms. These systems are developed by different developers and being modified inceremently. Integrating these autonomous tools generate various conflicts. To deal with these conflicts, current integration mechanisms defines specific set of rules to resolve these conflicts and accompalish integration. Indeed still there is a big chance that changes made by other developers, or they update their changes in order to make them compliant with other developers cancel the updates done by others. The approach presented here claims three contributions in the field of Hetrogeneous Software Integration. First, this approach eliminate the need of conflicts resolving mechanism. Secondly, it provides the mechanism to work in the presence of conflicts without resolving them. Finally, contribution is that the integration mechanism does not affect if either of the system evolves. We do this by introducing an intermediate virtual layer between two systems that introduce a delta models which consist of three parts; viability that share required elements, hiding that hide conflicting elements and aliasing that aliases same concepts in both systems.**

***Keywords*—Model Driven Engineering; Co-evolution; Co-adaptation; Delta models; Model Integration**

I. Introduction

Scale changes everything. This is one of the main obstacles that software community has to conquer to build systems of the future. Issues that are not significant at smaller scale become significant when scale grows larger. Future systems will move far beyond the size of today's systems by every measure: lines of code; people employing the system for different purposes; data storage, accessibility and manipulation. We call these systems as Large Scale Systems (LSS) or Systems of Systems (SOS).

LSS systems will essentially be distributed and allow only limited possibilities of centralized control over data, development, evolution and operation. The reasons to their distributed nature are their development and usage by wide variety of stakeholders with conflicting needs and continuous evolution. More LSS system, does more conflicts are likely to engender.

LSS system will comprises of different variety of sub systems. These sub systems will have their own domain specific languages (DSLs) designed according to different methodologies and philosophies. Sub systems expand from different places that are created and modified by dispersed teams with different schedules, processes and goals. Some of the sub systems may originate from legacies, that were already designed before they were chosen to be part of LSS system. Moreover all new sub systems will be written in different languages and built on variety of platforms.

One way to integrate sub systems and share as much data between them is by resolving all of their conflicts. Resolving conflicts for these independently developed sub systems looks next to impossible. Second way is to live with these conflicts and still share the features and data. In this paper, we propose an approach that help us to integrate two independently developed subsystems without resolving their conflicts.

Our approach introduces an intermediate mechanism between two subsystems; that work as a virtual layer. We name it as Virtual Heterogeneous Model Integration Layer (VHMIL). VHMIL work independent of both sub systems and do not propose any modification in the structure of both the systems. As the matter of fact, it works as a new central independent layer to facilitate sharing of features and data between both the systems. We already stated that it is a virtual layer, which implies that it does not affect the structure of both sub systems. On the contrary, it takes domain specific language of two systems as input plus necessary requirements from their domain experts and defines a new composed language.

In other words, VHMIL works as a filter of features between both the systems. Domain experts from both systems identify the features that what should be and should not be shared. After identification of features when it comes to the step of sharing stage. Some of the features raise conflicts. Shared features of one system may conflict with the second system's existing structure and vice versa. For the simplification of the process of integration we create two categories; one is non conflicting features and other is conflicting features.

Non conflicting features are handled by the visibility

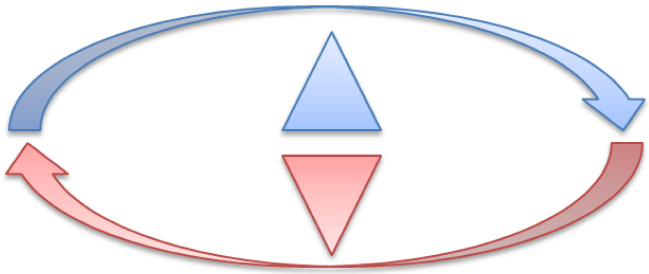


Fig. 1: Path between MM1 and MM2 meta-models

mechanism and conflicting features are handled by invisibility and masking mechanism. Visibility contain those features and data which are needed to be shared and do not generate any conflict. Invisibility contains those features which are not needed to be shared to avoid conflicts. Masking contains those features which are similar in nature and causes conflicts.

Non conflicting : Visibility
Conflicting: Invisibility and Masking

VHMIL creation is an automated activity except the identification of shared features, which domain experts propose manually. Domain experts have to decide which features need to be shared.

II. Model and Metamodel concept

When meta-model state every concept defined in the model and model uses the metamodel concepts according to rules stated by the metamodel then we call that model as instance of that meta-model. Conformance is described by a set of constraints between models and meta-models. When all constraints are validated, the model is said to conform to the meta-model. In this rest of paper we will be using MM for meta-model and M for instance model.

III. VHMIL: Virtual Heterogeneous Model Integration Creation

Consider two meta-models MM1 and MM2, both of them represent language for different systems. MM1 wants to use MM2's some features without modifying its own as well as MM2's language. Similarly MM2 wants to use MM1's some features without modifying its own as well as MM1's language. VHMIL layer consists of two paths. One path is from MM1 to MM2, and other is from MM2 to MM1. Both paths have one an intermediate meta-model as shown in figure 1. Path from MM1 to MM2 is represented as Δ(delta) and path from MM2 to MM1 is represented as $\triangledown$(nebla).

Δ MM12 metamodel share MM1's features for MM2. In the same way MM21 metamodel between path MM1 to MM2 provide MM2 features for MM1.

MM1 # MM2 = MM12 MM2 # MM1 = MM21

Delta and nebla consists of the same structure. Both consists of three sections (Visible, Hiding, Aliasing). Delta makes visible those elements for MM2, which need to be shared, hide those elements which need to hide and cause conflicts and alias those elements which has same syntax and semantics between the MM1 and MM2. Similarly nebla will hold those elements for MM1 which need to be shared for MM2, hide those elements which need to hide from MM2 and alias those elements which has same syntax and semantics between the MM1 and MM2.

Generally, these three sections are explained in the introduction section that what intermediate metamodel consists of and now we will explain in the perspective of metamodel.

We consider that it is more intuitive to represent a delta in operational terms; it is sequence of transformations that make visible , hide or mask elements from a metamodel. We have identified six elementary transformations in a metamodel that will be used as the basis for defining a delta. We assume that it is not possible to change the type of a model element, e.g., a UML Class cannot become a Package, and an element cannot change its UUID.

These operations are:

1) visibleElement: Make visible an element of type of the MM1 which MM2 require.
2) visibleLink: Make visible an MM1link which is required by the MM2.
3) hideElement: Hide an element which is not required by MM2
4) hideLink: Hide an association link which is not required by MM2.
5) aliasElement: Assign an alias(rename) to the element of MM1 required by MM2.
6) aliasLink: Assign an alias to the element of MM1 required by MM2.

IV. Methodology

VHMIL consists of several sub tasks. Figure 2 provides an overview of the proposed approach. We propose an approach to share common concepts between two metamodel variants with the impact of application of those common concepts on their instance models with emphasis on the minimization of manual effort. The envisioned steps are

1) **Requirements collection and writing pre directives:** Input gathering from domain experts of both metamodel variants and writing directives for these requirements manually .
2) **D2. Raw delta concepts:** Taking two meta-models variants plus directives as input and with transformation generate a list of direct (visible, hidden and alias) concepts. Etract information from paper to represent it.
3) **Generation of delta meta-model:** In this step is to generate an intermediary metamodel from list of direct concepts.
4) **Intermediate Directive Transformation (IDT) or Adapter**: This step have two options depending on the requirements of the system. One option is IDT, that take delta meta-model as input and generate a generic rule based transformation. Second option

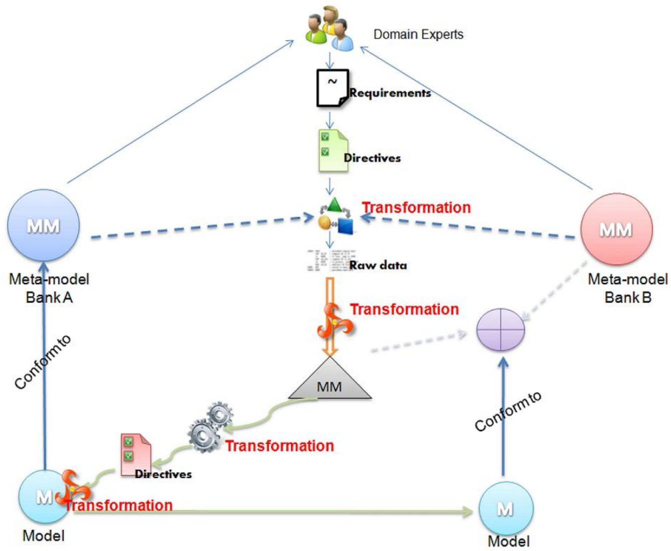


Fig. 2: VHMIL methodolgy diagram

is adapters to provide interface to existing instance model.

5) **Execution of IDT :**This step simply executes transformation on instance model generated from IDT. After applying these directives on the instance models of MM1 will yield new instance models that will work with the MM2 and delta meta-model.

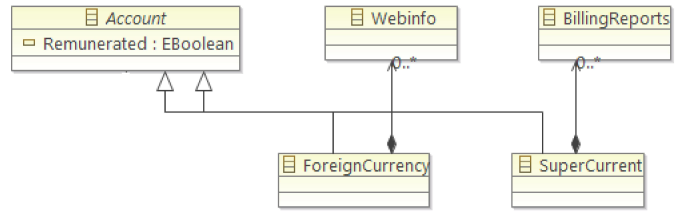


Fig. 3: Meta-model of bank A

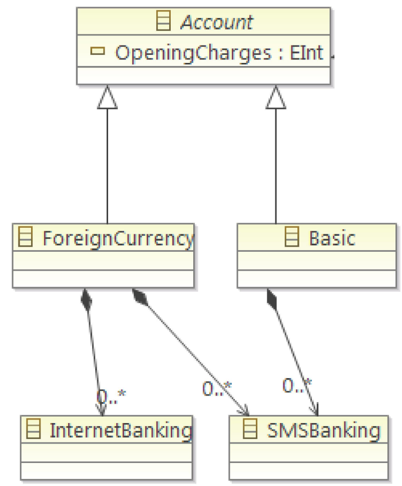


Fig. 4: Meta-model of Bank B

V. Running Example

We present here a small scenario to show that how VHMIL is created on a real time case study. We consider two systems of different banks. Both banks have different systems designed by different teams and built with different programming languages. Both the banks have different type of accounts and services and customers. Management of both banks decided to integrate and share some of the features between these banks in order to facilitate customers or we can call it as partial merger of banks.

Figure 3 shows a metamodel of bank A and figure 4 shows a metamodel of bank B. The ultimate goal is to facilitate customers of both banks to access their accounts and services from both the banks. Simply customer of bank A is able to access his account and associated services from bank B and vice versa.

Δ help customers of bank A to access

To avoid the repetition, I explain only calculation of Δ here, because calculation of $\bigtriangledown$ also follow the same process .

A. Requirements Collection and Writing Pre-directives

In this step, domain experts provide their view that what features have the same similar meaning between both meta-models. In our example, two accounts and one service have same semantics in both banks. Bank B will treat customer of Bank A ”SuperCurrent” account as their basic account and provide ”InternetBanking” service who requests for ”WebInfo” service. Similarly in Foreign Currency account is treated as ForeignCurr as shown in the table I. All other related features will be imported of Bank A meta-model that Bank B does not have. Delta model will import the service of BillingReports from Bank A. Similarly the concept that are not required from Bank A accounts will be make hidden for example ”SMSBanking”.

After identifying the similar concepts, developer writes pre directives for those concepts in directive language as shown below:

1) ABC::ForeignCurrency.name=ForeignCurrency
2) ABC::SuperCurrent.name=Basic
3) ABC::WebInfo.name=InternetBanking

B. Raw Delta Concepts

In this step an automated transformation takes both Bank A, Bank B meta-models and pre directives as an input and produce the raw input of all the concepts containing three parts: visible elements, alias elements and hidden elements for both delt and nebla.

Bank A	Bank B
Account : ForeignCurr	Account : ForeignCurrency
Account : SuperCurrent	Account : Basic
Service : WebInfo	Service :InternetBanking

TABLE I: Similar concepts between Bank A and Bank B

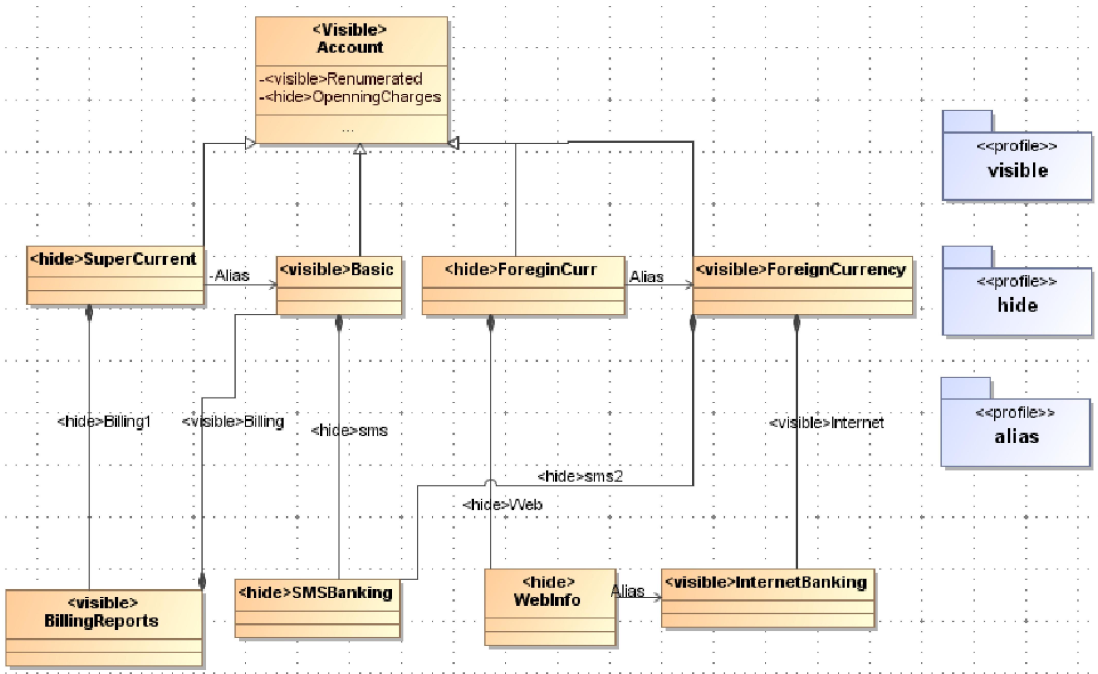


Fig. 5: Representation of delta meta-model for current example

Δ Dela meta-model
VisibleElement(Class, BillingReports)
VisibleElement(Parameter, Renumerated)
VisibleLink(Link,SuperCurrent,BillingReports)
VisibleLink(Link,ForeignCurrency,WebInfo)
HideElement(Class, SMSBanking)
HideAttribute(Class,Account, OpeningCharges)
HideElement(Containment, Basic, SMSBanking)
HideElement(Cotainment, ForeignCurrency, SMSBanking)
HideElement(Cotainment, ForeignCurrency, InternetBanking)
Alias(Class, SuperCurrent, Basic)
Alias(Class, Webinfo, InternetBanking)

∇ Nebla meta-model
VisibleElement(Class, SMSBanking)
VisibleElement(Parameter, OpeningCharges)
VisibleLink(Link,Basic,SMSBanking)
VisibleLink(Link,ForeignCurrency, SMSBanking)
VisibleLink(Link,ForeignCurrency, InternetBanking)
HideElement(Class, BillingReports)
HideAttribute(Class,Account, Renumerated)
HideElement(Containment, SuperCurrent, BillingReports)
HideElement(Cotainment, ForeignCurrency, WebInfo)
Alias(Class, Basic, SuperCurrent)
Alias(Class, InternetBanking, WebInfo)

C. Generation of Delta Meta-Model

In this step an intermediary delta meta-model is generated showing the visible, hidden and aliases concepts from list of direct concepts as shown in figure 5.

D. IDT Transformation or Adapter

In this step, the intermediary delta meta-model is translated into an executable generic transformation. The Intermediate Directive Transformation (IDT) takes as an input the final composed delta meta-model, and generate as output a model transformation written in a particular transformation language (e.g., ATL, XSLT, SQL-like, kermeta) as shown in figure 6.

```
import 'model.xmi';
Element : Webinfo
{
name ::=InternetBanking
}
Element : ForeignCurr
{
name ::=ForeignCurrency
}
Element : SuperCurrent
{
name ::=Basic
}
```

E. Execution of IDT

This step simply executes transformation on instance model. After applying transformation generated from IDT directives on the instance models of MM1 will yield new instance models that will work with the MM2 and delta meta-model.

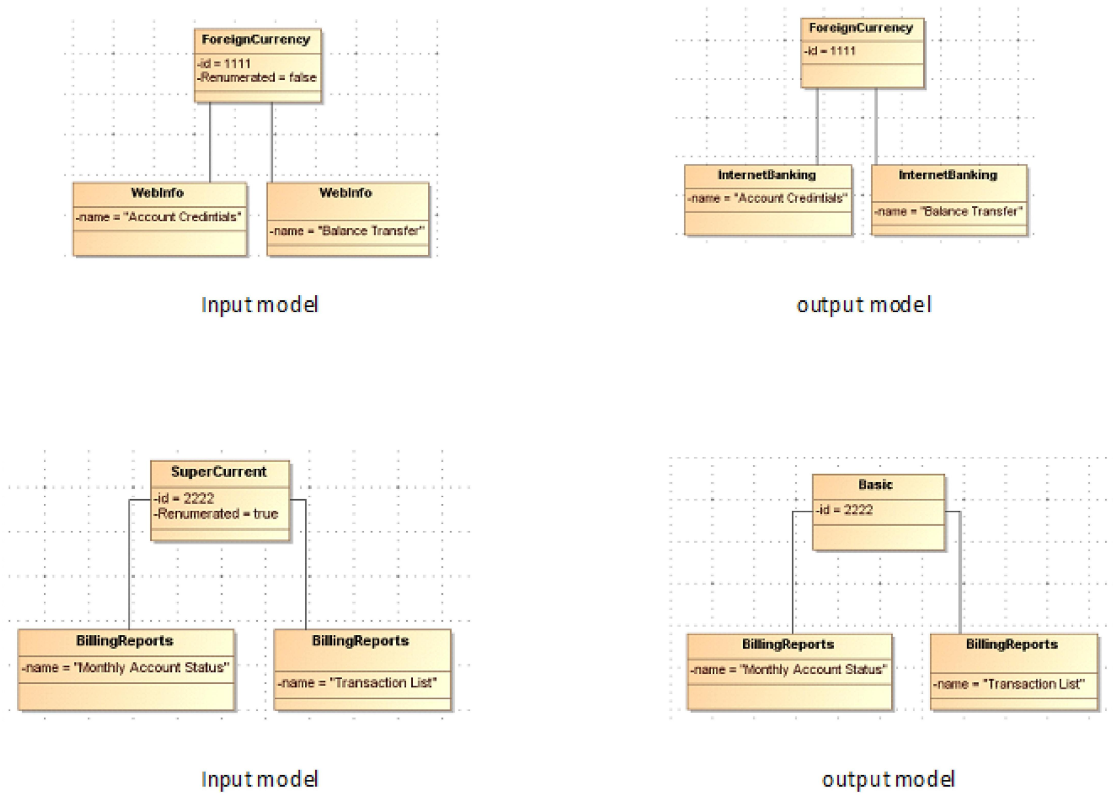


Fig. 6: Transformation of input models of one bank to another using detal and nebla meta-models

VI. RELATED WORK

We may divide the related approaches work on the basis of Software Integration and Software Evloution.

Xavier lanc et al. [2] presented a software integration mechanism called model bus. Model Bus is an approach to support data integration and control integration. ModelBus is based on the copy-modify merge mechanism. Models are split into several XML fils that are stored in a central storage, called the model base. Control integration facilitate heterogenous tools to invoke modeling services of each other. Each developer has his own workspace, called the model workspace, that stores only the XMI files modified by the developer. Developers can then modify their models locally and save their modifications to the model base. To manage the integration of developers' updates into the model base, ModelBus proposes dedicated delta extraction and delta integration mechanisms The section.When several developers modify a shared model and then generate different deltas on a same model, a delta integration mechanism has to combine together all the changes they do. Such a mechanism should automatically combine all possible changes but should also raise conflicts when changes cannot be automatically combined.to resolve these conflicts it also provides the conflicts resolving mechanism. In contrast to our approach we donot resolve these conflicts but try to live with these conflicts.[2]

Software Federations [13], [6], [5] approach provides the mechanism to build a large software application by composing Commercial Off The Shelf tools (COTS). This approach is based on two level programming paradigm.

Marcus et al. describes a generic approach for calculation of difference between two versions of models. When two developers change the same subset of the model, this lead to conflicts and it may not be possible to apply all changes in the model. On the contrary, In our approach, we calculate delta model between two model variants but not model versions and we do not compose this difference model to acquire new model. Instead delta model work with the base model in synchronization rather [1].EMF compare is a tool provided by the Eclipse Project. EMF Compare generate delta model, which reports simple changes between terminal models pairs or metamodel pairs [4]

Software Evolution is the most important factor to consider because if software evolves, it affects the validity of other software systems integrated with it. Hoessler et al. [10] identify transformation model for transition of M to M' conforming to the new Metamodel. This model consist of three types of patterns. First pattern is metmamodel extension that describes addition of class or property. Second pattern is metamodel projection for removal of a class or property. Third pattern is factoring for refinement of metamodel(Property Movement,Property Amalgamation, Class Splitting and Class Amalgamation). They do not develop any implementation for their proposed models.

Kelly et al. [7]have presented an approach for adapting model to evolving metamodels. This approach consists of three steps. In first step it computes similarities and differences between base metamodel and evolved metamodel by executing set of heuristics . In the second step HOT (Higher Order Transformation)translate takes all equivalences and differences and yield an executable transformation in ATL. Afterwards the transformation transform the model confirming to the base model to the newly evolved model while preserving the unchanged model elements.

Prawee Sriplakich et al. [18] shows a work of collaborating development and follows the copy-modify-merge approach to MDA models, where copy-modify-merge is consist of two steps delta calculation and delta integration. Further Author also propose a framework for automated to build programs for automated conflicts resolution resulted of deltas modifying same set of elements.

Gracia et al. [9] presents semi-automatic process for metamodel evolution that supports the co-evolution of transformation. The work is devided in two steps: (1) detection stage, where the changes in the metamodel are identified and classified (2) the action for the changes is being taken in co-evolution stage.

Schonbock et al. [17] presented an extensive survey for several co-evolution approaches from different areas such as software engineering like data, ontology, and grammar engineering by considering a certain roadmap for development in the field of co-evolution for agile MDE. Furthermore, the authors presented a conceptual co-evolution framework with help of running example for decrease in co-evolution effort and increase in co-evolution performance.

Garcés et al. [8] presented an approach for transformation that covers external transformation of metamodel. The approach works by identifying the transformation which deal with either refactoring/destruction changes or construction changes. It semi-automatically generates transformation with the help of AtlanMod matching language.

Yu et al. [21] proposed a framework for metamodel which is based on formal analysis for composition and adaptation. The framework is able to explore and evolve metamodel based on certain strategies. It also re-establish uniformity between existing models and the metamodels. The approach proves that model is able to accumulate information from the analysis.

Demuth et al. [3] presented a vision of co-evolution between metamodels and models that has an ability to adopt change propagation. The approach handles co-evolution issues without being dependent on specific metamodels or evolution scenarios. The approach can also detect failures that can occur during co-evolution process and generate appropriate suggestions for corrections.

Paige et al. [16] discussed key problems that occur in evolution process in the Model Driven Engineering (MDE) field. The authors discussed about up to date work of models and metamodels evolution and summarized the challenges it in this paper.

Taentzer et al. [19] proposed an approach which is related to graph transformation for model co-evolution. The models represent graphs with model relation and type-instance relations with respect to their metamodels.

Meyers et al. [15] presented a technique that help the user to solve the issues related to migration from one model to another. This approach simplifies migration specification and reduces the work of evolver. The formal framework of migration of the model which is independent of specific modeling approaches is presented by Mantz et al. [14]. While Wagelaar et al. [20] proposed a model transformation language for coupled evolution in metamodel. It shows executable semantics for EMFMigrate. Iovinoet al. [11] discussed the metamodel change impact of existing artifacts.

OCL plays an important role with respect to models and metamodels. Kusel [12] expressed the solution of actions for all metamodel that violate the changes of syntactic correctness of OCL expressions.

VII. Conclusions

It is a permanent effort in computer sciences, to find ways to integrate existing applications with new ones. This paper presents the current state of work in attempting to overcome the difficulties involved in application integration and reuse each other's services. While services in each application have different signature and structures, which may cause conflicts when trying to be integrate with each other. We propose an intermediatory layer to facilitate this integration of applications by eliminating the need of resolving these conflicts. In contrast conflicts are being handsled in a way that there conflicting behaviours effect the integration. Future work involves rigerous testing, verification and validation of our approach with large models.

References

[1] Marcus Alanen and Ivan Porres. Difference and union of models. pages 2–17, 2003.

[2] Xavier Blanc, Marie-Pierre Gervais, and Prawee Sriplakich. Model bus: Towards the interoperability of modelling tools. In *MDAFA*, pages 17–32, 2004.

[3] Andreas Demuth, Markus Riedl-Ehrenleitner, Roberto E Lopez-Herrejon, and Alexander Egyed. Co-evolution of metamodels and models through consistent change propagation. *Journal of Systems and Software*, 111:281–297, 2016.

[4] Eclipse.org. Emf compare, (2008). http://wiki.eclipse.org/index.php/EMF.

[5] J. Estublier, H. Verjus, and P. Y. Cunin. Modelling and managing software federations. volume 26, pages 299–300, New York, NY, USA, 2001. ACM.

[6] Jacky Estublier, Herv Verjus, and Pierre yves Cunin. Designing and building software federations. In *1st Conference on Component Based Software Engineering. (CBSE*, 2001.

[7] Kelly Garcés, Frédéric Jouault, Pierre Cointe, and Jean Bézivin. Managing model adaptation by precise detection of metamodel changes. In *ECMDA-FA '09: Proceedings of the 5th European Conference on Model Driven Architecture - Foundations and Applications*, pages 34–49, Berlin, Heidelberg, 2009. Springer-Verlag.

[8] Kelly Garcés, Juan M Vara, Frédéric Jouault, and Esperanza Marcos. Adapting transformations to metamodel changes via external transformation composition. *Software & Systems Modeling*, 13(2):789–806, 2014.

[9] Jokin García, Oscar Diaz, and Maider Azanza. Model transformation co-evolution: A semi-automatic approach. *Software Language Engineering*, 7745:144–163, 2013.

[10] Joachim Höler, Hajo Eichler, and Michael Soden. Coevolution of models, metamodels and transformations. Wissenschaft & Technik Verlag, June 2005.

[11] Ludovico Iovino, Alfonso Pierantonio, and Ivano Malavolta. On the impact significance of metamodel evolution in mde. *Journal of Object Technology*, 11(3):3–1, 2012.

[12] Angelika Kusel, Juergen Etzlstorfer, Elisabeth Kapsammer, Werner Retschitzegger, Johannes Schoenboeck, Wieland Schwinger, and Manuel Wimmer. Systematic co-evolution of ocl expressions. *11th APCCM*, 27:30, 2015.

[13] Tuyet Le-anh, Jorge Villalobos, and Jacky Estublier. Multi-level composition for software federations. 2003.

[14] Florian Mantz, Gabriele Taentzer, Yngve Lamo, and Uwe Wolter. Co-evolving meta-models and their instance models: A formal approach based on graph transformation. *Science of Computer Programming*, 104:2–43, 2015.

[15] Bart Meyers, Manuel Wimmer, Antonio Cicchetti, and Jonathan Sprinkle. A generic in-place transformation-based approach to structured model co-evolution. *Electronic Communications of the EASST*, 42, 2012.

[16] Richard F Paige, Nicholas Matragkas, and Louis M Rose. Evolving models in model-driven engineering: State-of-the-art and future challenges. *Journal of Systems and Software*, 111:272–280, 2016.

[17] J Schonbock, Juergen Etzlstorfer, Elisabeth Kapsammer, Angelika Kusel, Werner Retschitzegger, and Wieland Schwinger. Model-driven co-evolution for agile development. In *System Sciences (HICSS), 2015 48th Hawaii International Conference on*, pages 5094–5103. IEEE, 2015.

[18] Prawee Sriplakich, Xavier Blanc, and Marie-Pierre Gervais. Supporting collaborative development in an open mda environment. In *ICSM '06: Proceedings of the 22nd IEEE International Conference on Software Maintenance*, pages 244–253, Washington, DC, USA, 2006. IEEE Computer Society.

[19] Gabriele Taentzer, Florian Mantz, and Yngve Lamo. *Graph Transformations: 6th International Conference, ICGT 2012, Bremen, Germany, September 24-29, 2012. Proceedings*, chapter Co-transformation of Graphs and Type Graphs with Application to Model Co-evolution, pages 326–340. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[20] Dennis Wagelaar, Ludovico Iovino, Davide Di Ruscio, and Alfonso Pierantonio. *Theory and Practice of Model Transformations: 5th International Conference, ICMT 2012, Prague, Czech Republic, May 28-29, 2012. Proceedings*, chapter Translational Semantics of a Co-evolution Specific Language with the EMF Transformation Virtual Machine, pages 192–207. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[21] Ingrid Chieh Yu and Henning Berg. A framework for metamodel composition and adaptation with conformance-preserving model migration. In *Model-Driven Engineering and Software Development*, pages 133–154. Springer, 2015.

A Survey of Cloud Migration Methods: A Comparison and Proposition

Khadija SABIRI[1], Faouzia BENABBOU[1]
Information Technology and Modeling Laboratory[1]
University of science in Ben M'sik
Casablanca, Morocco

Mustapha HAIN[2], Hicham MOUTACHAOUIK[2]
Information Technology and Modeling Laboratory[2]
ENSAM
Casablanca, Morocco

Khalid AKODADI[3]
LAROSERI Laboratory[3]
Faculty of Sciences, CHOUAIB DOUKALLI[3]
El Jadida, Morocco

***Abstract*—Along with the significant advantages of cloud computing paradigm, the number of enterprises, which expect to move a legacy system towards a cloud, is steadily increasing. Unfortunately, this move is not straightforward. There are many challenges to take up. The applications are often written with the outdated technologies. While some enterprises redevelop applications with a specific Cloud provider in mind, others try to move the legacy systems, either because the organization wants to keep the past investments, or because the legacy systems hold important data. Migrating the legacy systems to the Cloud introduces technical and business challenges. This paper aims to study deeply and to compare existing Cloud migration methods, based on Model Driven Engineering (MDE) approach to highlight the strengths and weaknesses of each one. Finally, we have proposed a Cloud legacy system Migration Method relied on Architecture Driven Modernization (ADM), and explained its working process.**

***Keywords*—Application on premise; Migration methods; Cloud migration Method; PIM; PSM; ADM**

I. Introduction

Cloud computing has lately attracted full attention of many stakeholders due to its economic, business and technical benefits. Almost IT enterprises regard Cloud computing as a new way to adapt their business strategy to their business goals [1].Clouds are the model of the utility computing based on resources virtualization in terms of hardware, development platforms, and services[2]. Cloud computing has three service models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS)[2−4]. The pay-as-you-go is the Cloud business model, which the enterprises have the ability to decrease capital expenditure [5]. Additionally, Cloud computing provides plenty of advantages such as costing flexibility, high scalability, rapid elasticity, investment cost saving and so on [6]. The deployment models of Clouds are public, private, hybrid and community [7].

Many enterprises expect to move legacy systems to the Cloud to exploit the benefits of Cloud. However, many questions, as well as various challenges, are considered. Thus, through Cloud migration, the enterprises either have to start over or to adapt existing systems to Cloud. Despite outdating legacy systems technologies, it still works for the user's needs and is adaptable to business logic. In one hand, the cost of replacing them with the newly designed systems from scratch is often too high. In the other hand, every provider has their own standards that may impact on the way that applications are developed. As a result, users may be bound such a provider which hinders the decision of many customers to move to Cloud.

Many types of research in Cloud migration have been achieved. However, only a few have proposed a generic solution, even though there are more solutions to specific types of applications or at a given Cloud type. In this paper, we propose a comparative study between Cloud migration methods, notably those based on MDE, considered models in the centre of the software engineering process. The objective is to propose a Cloud migration Method based on ADM to ensure processes migration from legacy system on-premise to Cloud environment regardless of features applications and Cloud providers.

The rest of the paper is structured as follows: Section 2 illustrates literature survey of related works. A comparison of existed Cloud migration methods is presented in Section 3. Section 4 displays the proposed Cloud Migration Method. Section 5 contains a comprehensive conclusion of the paper findings.

II. Survey of Related Works

In the literature, there are several studies on cloud migration, such one has looked from a particular point of view which prevents to have a global vision of from the start of the cloud migration process. Basing on [8], those works are classified from the point of view discussed in the migration issue:

- Decision making support;
- Migration methods;
- Developments tools.

This section presents some work that focuses on decision-making support, migration methods and development tools for moving existing system to the Cloud services.

A. Decision Making Support

The works in references [9−10] present a support of a cloud migration decision in enterprises, analyzing the key factors to evaluate the application in terms of technical and non-technical constraint that must be considered before migration. Further aspects should be covered such as a study of migration feasibility, whether the application could be moved and choosing suitable Cloud provider.

Migration Methods

Other researchers have tackled migration methods [11−14]. the paper herein referenced, [11] proposes taxonomy of the migration tasks involved, which is based on experiences of migrating PetShop DotNet to Windows Azure, a cloud platform type, as well as Java PetStore to Amazon EC2, a Cloud infrastructure type, in order to enhance the taxonomy of migration tasks in a way to sum up a generic cloud migration process.

Other specialists [12] have analyzed experiences from moving Hackystat, an Open Source Software (OSS), to a Cloud. It aims to run the application on IaaS which will be consumed as services. Migrating to an SOA is required. Finally, the work referenced in [13] proposes Cloud Motion Framework (CMotion), which supports migrating composite applications into Clouds. The aim of this framework is to enable the different technologies to be compatible with each, by using adapters and runtimes to host each component.

B. Development Tools

Some works have been addressed to provide development tools for cloud migration process [15−17].

MoDisco [15] is an extensible framework defined as a support for several legacy technologies by providing a set of models to make easier the design and building of model-based solutions suited to legacy systems. Modisco is based on MDE paradigm to enhance the existed model-driven reverse engineering approaches.

Blu Age [16] is an agile solution with the aim of modernizing application, starting by extracting the business code of the legacy system to transform it into a PIM presentation independent of any technologies and regenerates it to a modernized system through MDA approach. Blu Age consists in designing a platform for supporting existing software modernization based on a model-driven development (MDD). It is focused on business level translation rather than translating legacy code at a technical level to a given technology.

Modelio is an extended open source tool supported UML standard. It is designed to support mechanisms extensibility by offering a set of modules to meet the specific needs [17]. Modelio starts with knowledge discovery from legacy system to recover legacy model in order to generate SOA components thanks to SOAtization methods.

C. Discussion

To summarize, this section enumerates current researches which are divided in decision-making support, migration methods and development tools. The first approach considered as the first step in migration process covers migration tasks with the aim of facilitating decision makers such as feasibility study, decision to move towards Cloud and the identification of the suitable Cloud provider. The second one is about migration methods that focused mainly on refactoring application code to leverage with such Cloud platform. The third approach defined some tools devoted to migrating a legacy system to Cloud.

The aim of this paper is to run comparative analysis tackling other Cloud migration methods that are based on MDE approach.

III. Presentation of Cloud Migration Methods

This section illustrates the existing Cloud migration methods, based mainly on MDE approach, its motivation, and its overall solution.

A. CloudMIG

The aim of CloudMIG [18] is summarized by the following points:

- Reduce over and under resources' provision;
- Evaluate additional expenditure;
- Discuss scalability issues;
- Align existing systems with a Cloud environment.

This approach consists of six steps to move a legacy system to a Cloud:

- Extraction: Reverse engineering step is needed to better understand legacy system for the purpose of generating result a software's representation, a current architecture of software, and software system's utilization model acting;
- Selection: aims at selecting a target platform for a given Cloud provider-specific during migration process, thanks to Cloud Environment Model (CEM), it describes shared characteristics of different Cloud environments;
- Generation: This activity produces three output models such as architecture of targeted platform, mapping model, and violations model of the target architecture which describes the constrains of Cloud environment;
- Adaptation: is a manual rearrangement of the target architecture and deployment adaptation;
- Evaluation: this step analyzes the produced target architecture using dynamic and static metrics;
- Transformation: comprises manual migration towards target architecture.

B. REMICS

The aim of REMICS (Reuse and Migration of legacy systems to Interoperable Cloud Services) is to support cloud migration of existing systems with using model-driven methodology and tools [19]. This work focuses on:

- Outdating technologies of legacy systems;
- Replacing the legacy system with systems designed from scratch;
- Reusing and integrating the legacy applications;

- Adapting existing services and increasing their interoperability to cope with the migration requirements.

To tackle these issues, this methodology helps a migration of the existing IT systems to Cloud through model-driven technologies, based on ADM which is the concept of modernization. Starting with recovering the architecture of the legacy system through the architectural model, in order to analyze the legacy system about the feasibility of modernizing different components parts of the legacy applications, and migration to SOA, to obtain the autonomous components. The migration activity identifies the different ways to modernize them by:

- The replacement of components by the online services;
- The utilization the Model-Driven Interoperability (MDI) helps to adapt the existing services using automatically generated mediators;
- The redevelopment of components by wrapping into SOA packages.

C. ARTIST

The project ARTIST (Advanced Software-based Service Provisioning and Migration of Legacy Software) covers a process to modernize legacy systems by providing migration method suitable to the Cloud [20]. The major outcomes have been treated by ARTIST:

- Tailor a legacy system to Cloud features;
- Migration of a legacy system is tied to using such technology and explicating the architecture of legacy systems.

To overcome these issues, this methodology composes of three main steps:

- Pre-migration phase consists of analyzing both technical and business aspects of legacy system which are considered as the result for the feasibility analysis , to identify the maturity-level assessment as well as target platform expected characteristics;
- Migration phase uses reverse engineering and forward engineering techniques to run the legacy system on Cloud environment. The migration process of the legacy system is tailor to the specific result gotten from the pre-migration analysis and the particular legacy application requirements. Additionally, this step takes into account the business model aspect. Lastly , checking and validating of resulted software are introduced in this phase;
- Post-migration phase relies on the applications components resulted, which are running in the cloud environment, in order to verify whether the technical and business objectives already defined in the pre-migration phase have been reached. Model-based testing and both of end-user functional and non-functional testing are the methods to approve the validation step.

D. Based on Services

Ref [21] presents a generic methodology, and displays the way to migrate legacy applications to Cloud service, to respond to these limitations:

- Outdated technologies and methodologies;
- Maintain system is hard;
- Tied migration methods to specific technologies;
- Less modification of legacy code.

This work consists of these following seven steps:

- Reconstructing architectural model of the legacy application: analyzes the legacy system and reconstructs a model of the legacy application architecture;
- Redesigning the architecture model: helps to determine services by redesigning the model of architecture's legacy system;
- MDA transformation: transforms the architecture to target;
- Generating Web-services: the Web service which are implemented in WSDL or JEE Annotation are generated as targeted web service;
- Invocating web services: The functionalities of the legacy system are identified in order to be invoked as the web service-base of legacy's functionalities;
- Selecting the suitable Cloud platform: it is chosen to be conformed to particular target platform needs , to support the web-services executed;
- Deploying web-services on Cloud platform: End user consumes the Web services running on Cloud. Service.

E. PaaS Migration

This work aims [22] at proposing a migration method between Cloud platforms, to deal with the vendor lock-in issue, resulted of locking application to specific Cloud features platform. Each cloud platform uses his specific tools, libraries, technology …etc.

In order to overcome this issue, this work has cited above attempts by offering a set of tools to re-engineer software, using the MDA and refactoring approach. This approach contains, besides Pattern Definition with Additional transformation rules, three steps:

- The first step is Discovery which aims at extracting the source platform to have a high level of sources representation.
- Transformation step aims at conforming the system to cloud platform through mapping rules, made to bridge the two representations.
- Migration step aims to test the resulted software deployed in cloud platform and to generate replies at runtime. The step checks the deployment at runtime and

the original platform utilizing rollback, used the old setting and code of the legacy system.

IV. Comparative Study of Migration Methods

In this section, the analysis is on the previous migration methods in terms of characteristics, drawing the comparative tabular to show strengths and weaknesses of each one. Firstly, we define some criterions used in this study.

A. Comparison criterions

- The Business and Technical Constraints: to show if the business and technical aspect take into consideration during the migration process.
- Level of automation: means the degree of migration methods automation: semi- or fully-automated approach;
- Architecture adaption: describes how to adapt legacy system architecture to Cloud environment, it composes of the following sub criterion:

a) Legacy system to PSM: explains process of recovering models with knowledge discovery;

b) Legacy PSM➔PIM: means transformation process from PSM to PIM through Reverse Engineering techniques;

c) Upgrade PIM: means describes the way architecture application was redefined, with aim of obtaining the autonomous components;

d) PSM➔PSMCloudified: is the process of adapting the software regarding the targeted platform through mapping rules;

e) PIMUp➔PSMCloudified: explains the transformation processes from PIMUp to PSMCloudified. It is the result of the compatibility with the cloud platform chosen and legacy platform specific provided as the migrated software.

- Migration layer: points out which layer takes place;
- Migration validation: aims to check whether the resulted system has worked as supposed to do, either by case study or by experimental experience.

B. Comparison

Each symbol "+" represents level of automation methods and describes the way used in it, namely:

[1]: Transforming legacy system to PSM;

[2]: Transforming legacy system to PSM and applying Forward Engineering techniques through PIM4Cloud;

[3:] Transforming a legacy system to PSM and PSM to PIM;

[4:] Transforming a legacy system to PSM and PIM to web services;

[5:] Transforming a legacy system to PSM and PSM to PSMCloudified.

According to this comparative study, we have noticed that each Cloud migration method has advantages and disadvantages, notably:

- The works in references [18, 21] didn't take business aspect into account rather they focus on running a legacy system in IaaS, PaaS Cloud. However, the other works stress out the need of adding the business concerns to technical aspects.
- Reference [19] is based on migrating the SOA autonomous components, either by replacing one or more legacy component by Cloud services, or wrapping legacy component.

As conclusion, this migration method presents three techniques to move to Cloud such as:

a) Replacing component(s) by Cloud offerings: This is the least invasive type of migration, where one or more (architectural) components are replaced by Cloud services [5]. Such migration type aims to resolve incompatibility issue through a set of reconfiguration to take a profit of the supported layer [23] and may be more expensive than rewriting.

b) Outsourcing and wrapping: the application functionality which is based on exposing them as service, hosting and running on Cloud environment. It requires the modernization of the legacy system by presenting them as autonomous components SOA in order to take a plenty of benefits of SOAs paradigm. Although the migration of a legacy system towards SOA is not an easy task, it must answer two major questions such as: what part of the legacy application can be moved[24], and how can be executed the migration process [25]. As a result, not all legacy system are enough mature to take up this transformation, and two studies required moving to Cloud, start with establishing a study to migrate to SOA, followed by another one to Cloud.

c) Cloudifying: is focused on transformation model to model based on Cloud blueprints, patterns, and best practices.

- In this study, the maturity aspect of the methods is measured by experiment experiences or case studies.
- In the Transformation PSM to PSMCloudified, users are not allowed to add more functionality or custom application responding to their needs.
- Automation level is referred to measure the level of automated cloud migration process. However, in this study, a fully automated process isn't recorded.

TABLE I. A COMPARATIVE TABULAR OF MIGRATION

Criterions / Methods	Constraints on migration		Architecture Adaptation					Migration Evaluation			Migration layer	Automation degree
	Technical	Business	Legacy system→ PSM		Upgrade PIM	Transformation: PIM→PSMCloudified	Transformation: PSM→PSMCloudified	Tools	Approach Migration validation	Test		
CloudMIG	Yes	No	Yes	Yes	No	No	No	Cloud SIM	Rating	Experiment experience	IaaS, PaaS	+[1]
REMICS	Yes	No	Yes	Yes	-Migration to SOA -Redesign architecture application by Soaml	-Replace legacy component -Wrap legacy component -Forward Engineering by PIM4Cloud	No	-	model-based testing model-based validation techniques	Case Study	IaaS, PaaS,	++[2]
ARTIST	Yes	Yes	Yes	Yes	No	-Transformation optimization patterns	No	-	Equivalence Test	Case study	IaaS, PaaS,	++[3]
Based on services	Yes	No	Yes	Yes	-Redesign architecture application -Identified service	-Transformation to web services -Hosting web service in environment Cloud	No	-	Not Mentioned	Case Study	IaaS, PaaS	++[4]
PaaS migration	-	-	Yes	No	No	No	Transformation rules	-	Test cases	Example	PaaS	++[5]

- Transformations focus on at the level of architectural models by redesigning the architecture of legacy system within a service-oriented architecture.

C. Discussion

As a conclusion, we find out that the underestimation of legacy system complexity is referred to the lack of architecture documentation or unknown internal structure. Additionally, trestriction technically related to the legacy systems nature such as huge of data and code, outdated legacy technologies and absence of separation concerns are considered as the main challenges to overcome so to better understand the legacy system. As shown, each migration process is tailored to a specific requirement of the legacy system.

To sum up, most of the presented Cloud migration methods are mainly based on migrating the SOA autonomous components, as it has potential advantages but also some significant drawbacks, mainly not all applications are enough mature to transform it into services and it may not be always possible to reuse legacy system's functionality by offering them as services due to technical concerns. In conclusion, replacing them could be less expensive than reused. In the next section, we propose our Cloud Migration Method and describe the processes migration independently of the type of legacy systems and Cloud Provider.

V. Proposed Cloud Migration Method

We summarize the objectives of this work showed in fig 1 in the following points:

- Moving legacy systems with keeping functionalities and data of the application in order to improves the past investments in software;
- Give users the possibility to add new business or Cloud functionalities if necessary;
- Target system becomes easy to understand, to maintain, and to make changes.

This method aims at presenting each process as a model to explore the benefits of Model Driven Engineering paradigm. The results of extracting Application Source Code, the PSM of Application, is modeled in three views, which enables to separate out the concerns in way to identify enterprise need and to match with Cloud model services such as:

a) Business Model: Identifies business requirement through business process of legacy application. In this step, the motivation for migartion and the goals to be achieved is elaborated. It also describes how the application will be behaved towards a cloud computing environment.

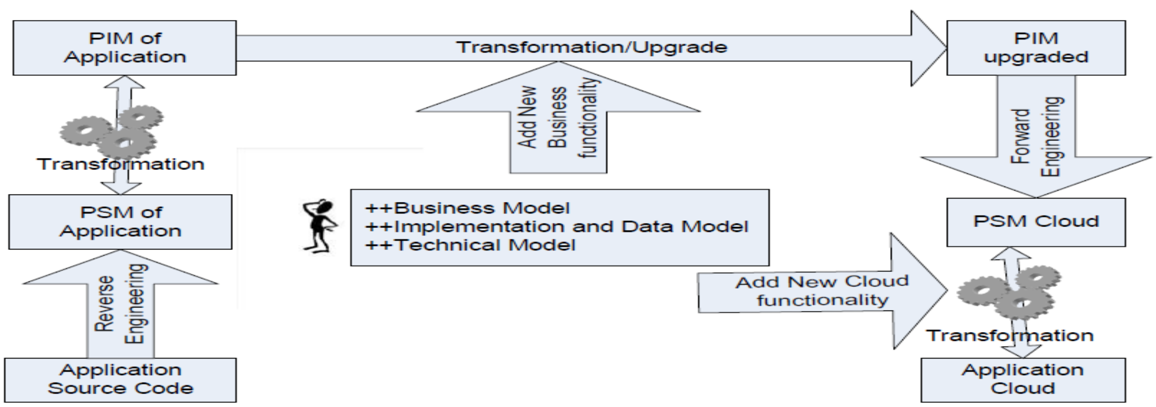


Fig. 1. Proposed Cloud Migration

b) Implementation and Data Model: Identifies business requirement through business process defines the architectural model of the application and their interrelationships and to identify the several components of the application in order to define which component of the architecture is hosted in which of the containers.

c) Infrastructure Model: describes the usage of the physical resource in a way to determine which resources are not properly used resulting in additional costs such as energy, maintenance, etc…

The difference between PIM Application and PIM Upgrade is that the first one describes the current functionality offered by the legacy application regardless of any technical details, and the second one enables users the possibility to add new business functionality if necessary. As a consequence, a new design artifact evolves from the original design. PSM Cloud is the results of Meta-model transformation between a Meta-model legacy system and a Meta-model Cloud application through transformation rules, which have to be conformed transformation Meta-model

VI. Conclusion and Perspectivses

In this paper, Cloud migration methods were reviewed based on Model Driven Engineering (MDE). As a result, we obtained a comparative tabular, which identifies weak and strong points for each one according to own comparison criterions. We also proposed own Cloud Migration Method, relied on ADM approach, which is focused on modernization activities on the architectural models rather than code artifact.

The future work aims to elaborate Meta models of the legacy system and Cloud application, which will be an abstract model to define Meta model transformation, in order to explore the benefit of models paradigm. Thus, legacy system model and Cloud application model have to be conformed to own Meta models to run Model-to-model transformations which translate between source and target models regardless of Cloud provider and requirement legacy system.

References

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica and M. Zaharia “A view of cloud computing”,ACM Communication. vol. 53, 2010.

[2] C. Y. Low, Y. Chen, M. C. Wu. Understanding the determinants of cloud computing adoption. Industrial Management & Data Systems, vol. 111, no. 7, 2011.

[3] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, M. Zaharia, “Above the clouds: A berkeley view of cloud computing” Technical Report, California University , 2009.

[4] L. M. Vaquero, L. Rodero-Merino, J. Caceres, M. Lindner. “A break in the clouds: towards a cloud definition”, ACM Communication, vol. 39, 2008.

[5] V. Andrikopoulos, T. Binz, F. Leymann,S. Strauch, “How to adapt applications for the cloud environment: challenges and solutions in Migrating Applications to the Cloud Computing”, vol. 95, no. 6, 2013.

[6] Q. Zhang, L. Cheng, R. Boutaba. “Cloud computing: Stateof-the-art and research challenges”. Journal of Internet Services and Applications, vol. 1, no. 1, 2010,pp. 7–18.

[7] C.Zhili, D. Zhihui, “A service-oriented virtualization management system with automated configuration”. IEEE International Symposium on Service-Oriented System Engineering, 2008, pp 251–256. IEEE.

[8] J.F.Zhao, J.T.Zhou. “Strategies and methods for cloud migration”, International Journal of Automation and Computing, DOI: 10.1007/s11633-014-0776-7, 11(2), 143-152, April 2014.

[9] A. K.Hosseini, D. Greenwood, J. W. Smith, L. Sommerville. “The Cloud Adoption Toolkit: Addressing the Challenges of Cloud Adoption in the Enterprise”.2012.

[10] M. Menzel, R. Ranjan. “CloudGenius: A Hybrid Decision Support Method for Automating the Migration of Web Application Clusters to Public Clouds”. Proceeding on IEEE Transactions on Computers, DOI: 10.1109/TC.2014.2317188 , 2014.

[11] V. Tran, J. Keung, A. Liu, A. Fekete. “Application migration to cloud: taxonomy of critical factors”. 2nd International Workshop on Software Engineering for Cloud Computing, ACM, USA, 2011.

[12] M. A. Babar, M. A. Chauhan. “A tale of migration to Cloud computing for sharing experiences and observations". 2nd International Workshop on Software Engineering for Cloud Computing, ACM, USA, 2011, pp. 50–56.

[13] T. Binz, F. Leymann, D. Schumm . “CMotion: A framework for migration of applications into and between Clouds”. IEEE International Conference on Service-Oriented Computing and Applications, USA, 2011.

[14] Q.H.Vu, R.Asal. “Legacy Application Migration to the Cloud : Practicability and Methodlogy ”. IEEE Eighth World Congress on Services, 2012.

[15] H. Bruneliere, J. Cabot, F. Jouault, F. Madiot. “MoDisco: A model driven reverse engineering”. Information and Software Technology Vol 56, 2014,pp. 1012–1032.

[16] Modernization legacy system Blue age Available:http://www.omg.org/mda/mda_files/Blu_Age_datasheet_en.pdf, August 12, 2015

[17] Modelio. Available: http://www.modelio.org, August 12, 2015

[18] S. Frey, W. Hasselbring. “Model-Based Migration of Legacy Software Systems to Scalable and Resource-Efficient Cloud-Based Applications: The CloudMIG Approach”.: The First International Conference on Cloud Computing, GRIDs, and Virtualization, CLOUD COMPUTING 2010.

[19] M.Parastoo, A. J. Berre, A.Henry, F. Barbier, A.Sadovykh “REuse and Migration of Legacy Applications to Interoperable Cloud Services”, REMICS Third European Conference, ServiceWave, pages 195–196, 2010.

[20] A.Menychtas, C.Santzaridou, L.O.Echevarria, J.Alonso “ARTIST Methodology and Framework: A novel approach for the migration of legacy software on the Cloud”. 15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 2013.

[21] W. Q. Zhang, A. J. Berre, D. Roman, H. A. Huru. “Migrating legacy applications to the service Cloud”. 14th Conference Companion on Object Oriented Programming Systems Languages and Applications, 2009.

[22] A. Beslic , R. Bendraou ,J. Sopena ,J-Y. Rigolet, “Towards a solution avoiding Vendor Lock-in to enable Migration Between Cloud Platforms”, 2nd International Workshop on Model-Driven Engineering for High Performance and Cloud computing (MDHPCL), 2013.

[23] P.Jamshidi, A.Ahmad, C.Pahl, Member, IEEE,. “Cloud Migration Research: A Systematic Review”, Cloud Computing, IEEE Transactions on (Volume:1 , Issue: 2), October 2013,pp 142 – 157, 08.

[24] Sneed, H. “Integrating legacy software into a Service Oriented Architecture”. 10th European Conference on Software Maintenance and Reengineering. IEEE CS Press, 2006, pp. 3–13.

[25] Lewis, G., Morris, E., Smith, D. , L. O'Brien “ Service-Oriented Migration and Reuse Technique”. 13th IEEE International Workshop on Software Technology and Engineering Practice (STEP'05).2005, pp. 222-229