ISSN : 2165-4069(Online)
ISSN : 2165-4050(Print)

IJARAI

International Journal of
Advanced Research in Artificial Intelligence

Volume 4 Issue 2

www.ijarai.thesai.org

A Publication of
The Science and Information Organization

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN ARTIFICIAL INTELLIGENCE

THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org

Editorial Preface

From the Desk of Managing Editor…

Artificial Intelligence is hardly a new idea. Human likenesses, with the ability to act as human, dates back to Geek mythology with Pygmalion's ivory statue or the bronze robot of Hephaestus. However, with innovations in the technological world, AI is undergoing a renaissance that is giving way to new channels of creativity.

The study and pursuit of creating artificial intelligence is more than designing a system that can beat grand masters at chess or win endless rounds of Jeopardy!. Instead, the journey of discovery has more real-life applications than could be expected. While it may seem like it is out of a science fiction novel, work in the field of AI can be used to perfect face recognition software or be used to design a fully functioning neural network.

At the International Journal of Advanced Research in Artificial Intelligence, we strive to disseminate proposals for new ways of looking at problems related to AI. This includes being able to provide demonstrations of effectiveness in this field. We also look for papers that have real-life applications complete with descriptions of scenarios, solutions, and in-depth evaluations of the techniques being utilized.

Our mission is to be one of the most respected publications in the field and engage in the ubiquitous spread of knowledge with effectiveness to a wide audience. It is why all of articles are open access and available view at any time.

IJARAI strives to include articles of both research and innovative applications of AI from all over the world. It is our goal to bring together researchers, professors, and students to share ideas, problems, and solution relating to artificial intelligence and application with its convergence strategies. 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.

We hope that this journal will inspire and educate. For those who may be enticed to submit papers, thank you for sharing your wisdom.

Editor-in-Chief
IJARAI
Volume 4 Issue 2 February 2015
ISSN: 2165-4069(Online)
ISSN: 2165-4050(Print)

Editorial Board

Reviewer Board Members

CONTENTS

Blocking Black Area Method for Speech Segmentation

Dr. Md. Mijanur Rahman
Dept. of Computer Science & Engineering
Jatiya Kabi Kazi Nazrul Islam University
Trishal, Mymensingh, Bangladesh

Fatema Khatun
Dept. of Electrical & Electronic Engineering
Hamdard University Bangladesh
Sonargoan, Narayanganj, Bangladesh

Dr. Md. Al-Amin Bhuiyan
Dept. of Computer Engineering
King Faisal University
Al Ahssa 31982, Saudi Arabia

***Abstract*—Speech segmentation is an important sub problem of automatic speech recognition. This research is concerned with the development of a continuous speech segmentation system using Bangla Language. This paper presents a dynamic thresholding algorithm to segment the continuous Bngla speech sentences into words/sub-words. The research uses Otsu's method for dynamic thresholding and introduces a new approach, named blocking black area method to identify the voiced regions of the continuous speech in speech segmentation. The developed system has been justified with continuously spoken several Bangla sentences. To test the performance of the system, 100 Bangla sentences have been recorded from 5 (five) male speakers of different ages and 656 words have been presented in the 100 Bangla sentences. So, the speech database contains 500 Bangla sentences with 3280 words. All the algorithms and methods used in this research are implemented in MATLAB and the proposed system has been achieved the average segmentation accuracy of 90.58%.**

Keywords—Blocking Black Area; Boundary Detection; Dynamic Thresholding; Otsu's Algorithm; Speech Segmentation

I. INTRODUCTION

Automated Speech Recognition (ASR) is a popular and challenging area of research in developing human computer interactions. The main challenge of speech recognition lies in modeling the variations of the uttered speech, such as different geographical boundaries, social background, age, gender, occupation etc. Automated segmentation of speech signals has been under research for over 30 years [1]. It is a necessity for phonetic analysis of speech [2, 3], audio content classification [4] and many applications in the field of automatic speech recognition (ASR), including word recognition [5, 6]. Speech Recognition system requires segmentation of Speech waveform into fundamental acoustic units [7]. Segmentation is the very basic step in any voiced activated systems like speech recognition system and speech synthesis system. The set of fundamental acoustic units into which the speech waveform can be segmented are words, phonemes or syllables. Word is the preferred and natural unit of speech, because word units have well defined acoustic representation. So, this research chooses word as the basic unit for segmentation. Speech segmentation was done using wavelet [8], fuzzy methods [9], artificial neural networks [10] and Hidden Markov Model [11].

This paper will present the proposed dynamic thresholding algorithms for segmenting continuous Bangla speech sentences into words/sub-words. For speech segmentation, this research introduces a new approach, named *blocking black area method* to properly detect word boundaries in continuous speech segmentation. The paper is organized as follows: Section I describes the introduction of speech processing and the organization of this paper. In Section II, we will discuss about speech segmentation and types of segmentation. Section III will describe thresholding. In Section IV, Otsu's thresholding method will be discussed. Section V will present the blocking black area method. The implementation of the proposed system will be described in Section VI. Sections VII and VIII will describe the experimental results and conclusion, respectively.

II. SPEECH SEGMENTATION

Speech segmentation is the process of identifying the boundaries between words, syllables, or phonemes in spoken natural languages. The general idea of segmentation can be described as dividing something continuous into discrete, non-overlapping entities [12]. In speech segmentation, the basic idea of segmentation is to divide a continuous speech signal into smaller parts, where each of these segments has phonetical or acoustical properties that distinguishes it from neighboring segments. Segmentation can be performed, for example, at the *segment*, *phone*, *syllable*, *word*, *and sentence* or *dialog turn* level. In isolated word recognition systems, accurate detection of the endpoints of a spoken word is important for two reasons, namely: Reliable word recognition is critically dependent on accurate endpoint detection and the computation for processing the speech is less, when the endpoints are accurately located [13]. Automatic speech segmentation methods can be classified in many ways, but one very common classification is the division to blind [14] and aided segmentation algorithms [15]. A central difference between aided and blind methods is in how much the segmentation algorithm uses previously obtained data or external knowledge to process the expected speech.

III. DYNAMIC THRESHOLDING

In general, thresholding is the simplest method of image segmentation.

This research proposes thresholding techniques on speech segmentation. From a grayscale image, thresholding can be used to convert binary image [16]. In order to convert the image into a binary representation, the technique first converts the image into a grayscale representation and performs a particular threshold analysis process in order to determine which pixels are turned into black or which are white. This research proposes *dynamic thresholding* to convert 256 gray-levels images into monochrome ones. Two important thresholding techniques are fixed or static thresholding and dynamic thresholding. In fixed or static thresholding, the systems usually uses 127 (say) as default threshold value, but you could change this value and obtain darker or lighter images. In dynamic thresholding, the system uses a different threshold value for each pixel of the image. This value is selected automatically, analyzing the sub-image area around each pixel and finding the local contrast. If the contrast of this area is low, the pixel is binarized using a global pre-calculated threshold value, otherwise, when the contrast is high, the local threshold value is calculated and used. In thresholding technique, the output image replaces all pixels in the input image with luminance greater than a threshold with the value of 1 (white) or 0 (black). The problem is how to choose the desired threshold value. Different dynamic thresholding techniques have been used to compute the threshold value [17]. Hence, the research proposes Otsu's thresholding algorithm to compute the desired threshold.

IV. OTSU'S ALGORITHM

Otsu's method is a simple and effective automatic thresholding method, used in image segmentation [18], invented by Nobuyuki Otsu in 1979 [19], also known as binarization algorithm. It is used to automatically perform histogram shape-based image thresholding (i.e. the reduction of a grayscale image into a binary image). The algorithm assumes that the image is composed of two basic classes; such as foreground and background [19]. It then computes an optimal threshold value that minimizes the weighted within class variance; also maximizes the between class variance of these two classes. The algorithmic steps for calculating the threshold is given in Figure-1.

The mathematical formulation of the algorithm for computing the optimum threshold will explain in this section. Let *P(i)* represents the image histogram of speech spectrogram. The two class probabilities $w_1(t)$ and $w_2(t)$ at level *t* are computed by:

$$w_1(t) = \sum_{i=1}^{t} P(i) \qquad (1)$$

and $$w_2(t) = \sum_{i=t+1}^{I} P(i) \qquad (2)$$

The class means, $\mu_1(t)$ and $\mu_2(t)$ are:

$$\mu_1(t) = \sum_{i=1}^{t} \frac{iP(i)}{w_1(t)} \qquad (3)$$

and $$\mu_2(t) = \sum_{i=t+1}^{I} \frac{iP(i)}{w_2(t)} \qquad (4)$$

Individual class variances:

$$\sigma_1^2(t) = \sum_{i=1}^{t} [i - \mu_1(t)]^2 \frac{P(i)}{w_1(t)} \qquad (5)$$

and $$\sigma_2^2(t) = \sum_{i=t+1}^{I} [i - \mu_2(t)]^2 \frac{P(i)}{w_2(t)} \qquad (6)$$

The within class variance (σ_w) is defined as a weighted sum of variances of the two classes and given by:

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \qquad (7)$$

Now we will calculate the *between class* variance. The between class variance (σ_b) is defined as a difference of total variance and within class variance and given by:

$$\sigma_b^2(t) = \sigma^2(t) - \sigma_w^2(t)$$
$$= w_1(t)[\mu_1(t) - \mu]^2 + w_2(t)[\mu_2(t) - \mu]^2$$
$$= w_1(t)w_2(t)[\mu_1(t) - \mu_2(t)]^2 \qquad (8)$$

$$where\ \mu = w_1(t)\mu_1(t) + w_2(t)\mu_2(t) \qquad (9)$$

These two variances $\boldsymbol{\sigma_w}$ and $\boldsymbol{\sigma_b}$ are calculated for all possible thresholds, $t = 0 \ldots I$ (max. intensity). Otsu finds the best threshold that *minimizes the weighted* ***within class variance*** **(σ_w)**, *also maximizes the weighted* ***between class variance*** **(σ_b)**. Finally, the pixel luminance less than or equal to threshold is replaced by 0 (black) and greater than threshold is replaced by 1 (white) to obtain the binary or B/W image.

V. BLOCKING BLACK AREA METHOD

For speech segmentation, this research introduces a new approach, named *blocking black area method*. This method is used to block the voiced regions of the continuous speech, so that we can easily separate the voiced parts of the speech from silence or un-voiced parts in the continuous speech. The edges of the block are used as word boundaries in the continuous speech. The main task of speech segmentation is to detect the boundaries of speech units (i.e., start and end points detection). The algorithm is applied in the thresholded spectrogram image that produces rectangular black boxes in the voiced regions of the speech sentence, as shown in Figure-2. Each black box represents a speech unit (i.e., word or sub-word) of a speech sentence. The method works as follows:

- Summing the column-wise intensity values of thresholded spectrogram image.
- Find the image columns with fewer white pixels based on summing value and replace all pixels on this column with luminance 0 (black).
- Find the image columns with fewer black pixels based on summing value and replace all pixels on this column with luminance 1 (white).

- Detect the boundaries of voiced block and separate the voiced block as speech units.

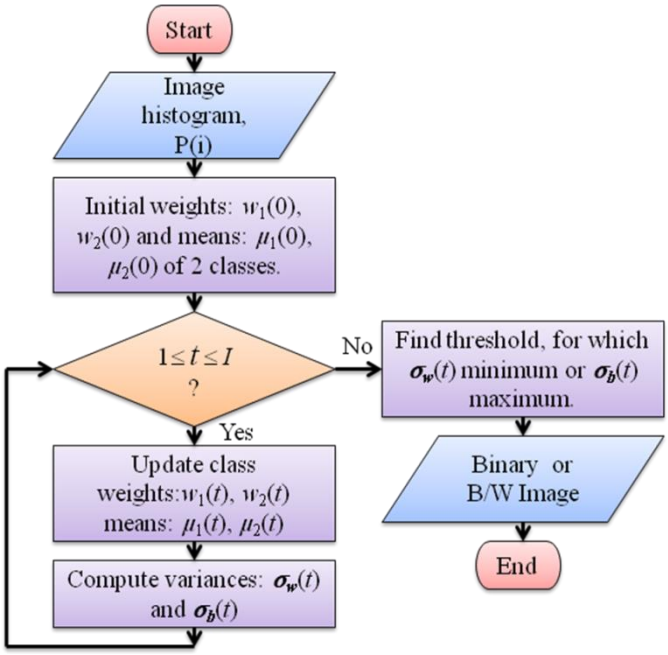


Fig. 1. Otsu's Thresholding Algorithm

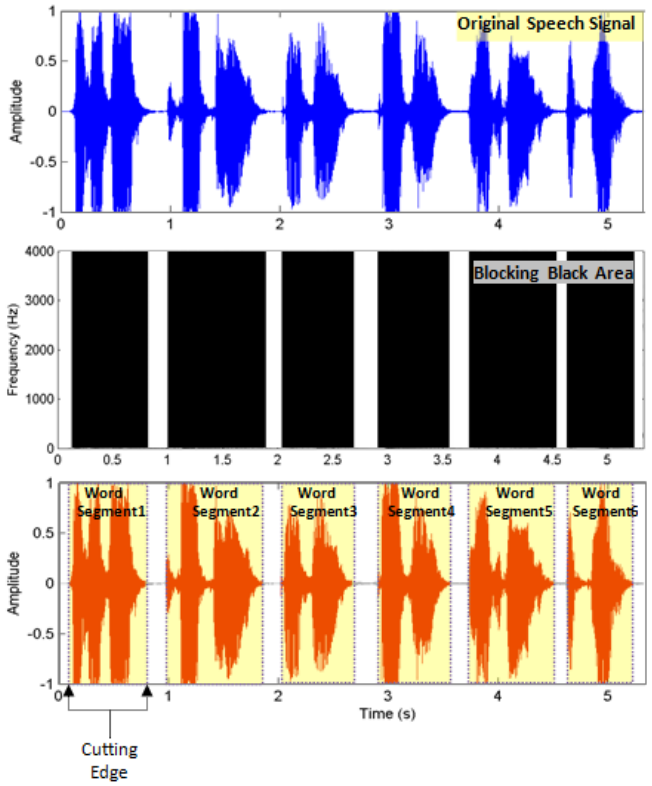


Fig. 2. Blocking the voiced regions by using blocking black area method

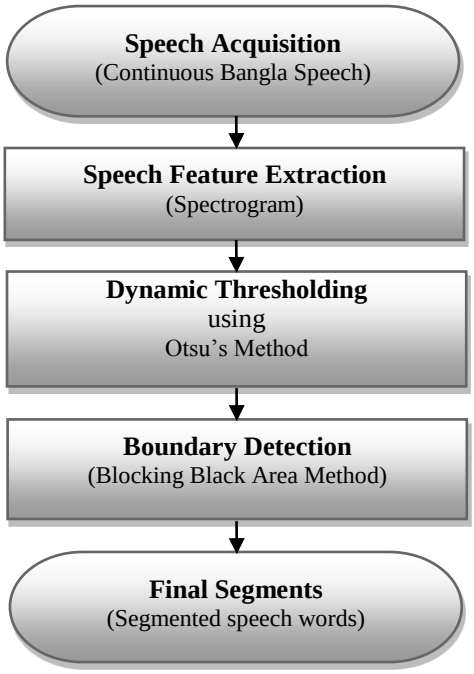


Fig. 3. Proposed Speech Segmentation Procedure

VI. IMPLEMENTATION

The proposed segmentation system, shown in Figure-3, has the following major steps and will discuss in the following sub-sections.

A. Speech Acquisition

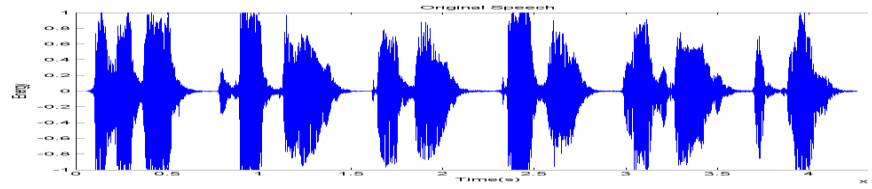
B. Feature Generation and Thresholding

C. Word Boundary Detection

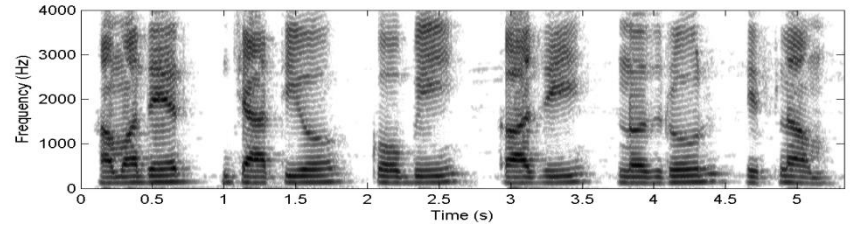
D. Speech Segment Separation

A. Speech Acquisition

Speech acquisition is acquiring of continuous Bangla speech sentences through the microphone. Recording was done by 5 (five) native male speakers of Bengali. The sampling frequency is 16 KHz; sample size is 8 bits, and mono channels are used. The time-domain plot of a speech sentence ('আমাদের জাতীয় কবি কাজী নজরুল ইসলাম') is shown in Figure-4(a).



(a) Original Speech Signal



(b) Spectrogram Image (grayscale)

(c) Thresholded Image

Fig. 4. Thresholded Spectrogram Images of the Speech Sentence 'আমাদের জাতীয় কবি কাজী নজরুল ইসলাম'

B. Speech Feature Generation and Thresholding

The feature extraction process generates spectrogram features from Bangla speech sentences. The grayscale spectrogram image of the speech sentence ('আমাদের জাতীয় কবি কাজী নজরুল ইসলাম') is shown in Figure-4(b). Spectrograms can be used to identify spoken words phonetically. For further processing of the spectrogram image, the labels of the image, such x-label, y-label and tile of the image, have been omitted, that's why label or title of the image is not shown in Figure-4(c). The thresholding algorithm is used to separate voiced regions from silence/un-voiced on continuous speech. The Matlab's '`graythresh`' function is used to implement the Algorithm-3. This algorithm returns a level (i.e., threshold) value for which the intra-class variance of the black and white pixels is minimum. The output image replaces all pixels in the input image with luminance greater than or equal to the threshold with the value of 1 (fully white) and less than threshold with 0 (fully black) to get fully black/white image (i.e., thresholded image). The thresholded image of the above speech sentence is shown in Figure-4(c).

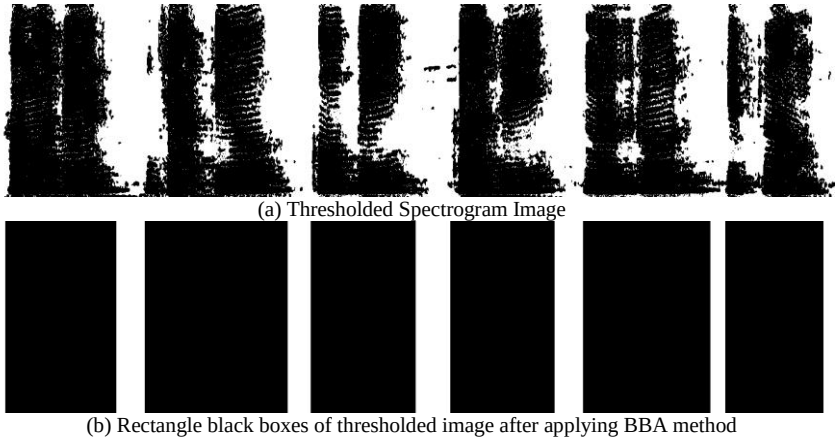
C. Word Boundary Detection

The newly introduced *blocking black area method* and *shape identification* techniques to properly detect word boundaries in continuous speech and label the entire speech sentence into a sequence of words/sub-words. The *block black area* method is applied in the thresholded spectrogram image that produces rectangular black boxes in the voiced regions of the speech sentence, as shown in Figure-5. Each rectangular black box represents a speech word or sub-word.

The method uses Matlab's '`regionprops`' function to identify each rectangular object in the binary image that represents speech words/sub-words. The function '`regionprops`' measures the properties of each connected object in the binary image. Different shape measurements properties, such as '`Area`', '`BoundingBox`', '`Centroid`' are used to identify each rectangular object in the binary image. The '`Extrema`' measurement, which is a vector of `[top-left top-right right-top right-bottom bottom-right bottom-left left-bottom left-top]`, is used to detect the start (`bottom-left`) and end (`bottom-right`) points of each rectangular object, as shown in Figure-6.

D. Word Segment Separation

Each rectangular black box represents a speech segment, such as a word or sub-word. After detecting the start and end points of each black box, the word boundaries in the original speech sentence are marked automatically by these two points and separated each speech segment from the speech sentence. Figure-7 shows that 6 (six) black boxes represent 6 (six) word segments in the speech sentence 'আমাদের জাতীয় কবি কাজী নজরুল ইসলাম'.



(a) Thresholded Spectrogram Image

(b) Rectangle black boxes of thresholded image after applying BBA method

Fig. 5. Effect of applying Blocking Black Area (BBA) Method – Producing rectangle black boxes in voiced regions. (a) Before applying Blocking Black Area Method and (b) After applying Blocking Black Area Method – Each black box represents a word/sub-word of the continuous speech

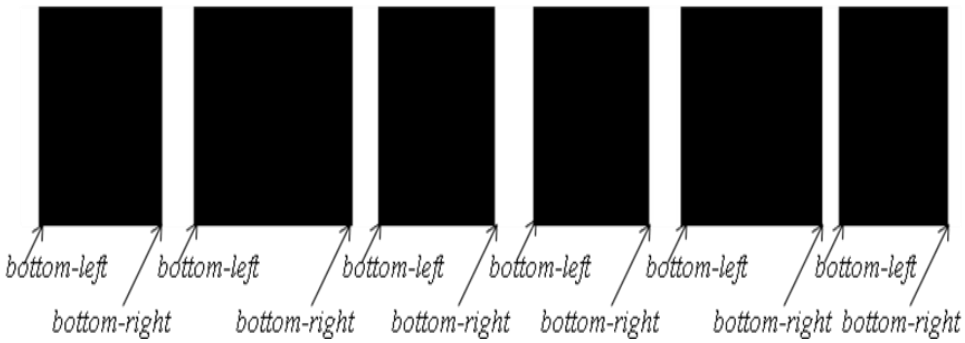


Fig. 6. Star and End point Detection of rectangular object

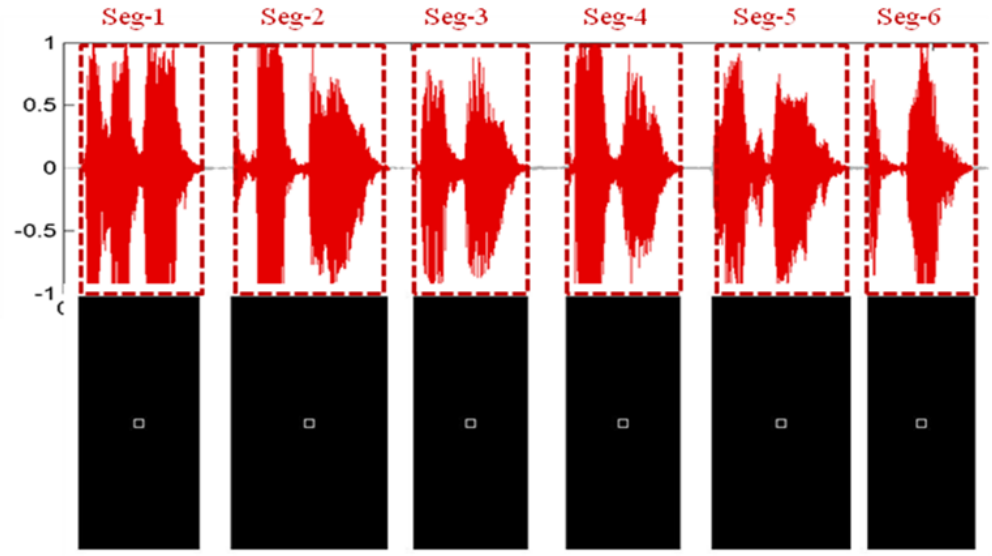


Fig. 7. Word Segments - 6 word segments in speech sentence 'আমাদের জাতীয় কবি কাজী নজরুল ইসলাম'

VII. Experiments and Results

For speech segmentation, this research proposes the dynamic thresholding algorithm with newly introduced *blocking black area method* to segment the continuously spoken Bangla sentence into words or sub-words. All the programs related to the speech segmentation approaches have been implemented in Matlab. The '*myspectrogram.m*' program computes spectrogram image from the original speech signal.

This research uses MATLAB's '*graythresh*' function to implement modified Otsu's algorithm that returns the desired threshold. The output image replaces all pixels in the input image with luminance greater than or equal to the threshold with the value of 1 (fully white) and less than threshold with 0 (fully black). The '*Blocking Black Area*' method has been implemented in the program '*blockingBlackArea.m*' that produces rectangular black boxes in the thresholded spectrogram image. The research uses MATLAB function '*regionprops*' to identify each rectangular object and the function's '*Extrema*' is used to detect the start and end points of each black box. The word boundaries of the original speech sentence are marked automatically by these two points and cut the word segments from the speech sentence and finally, the speech segments are save as .wav file format.

The developed system has been justified with continuously spoken several Bangla sentences. To test the performance the system, 100 Bangla sentences have been recorded from 5 (five) male speakers of different ages and 656 words have been presented in the 100 Bangla sentences. So, the speech database contains 500 (100x5) Bangla sentences with 3280 (656x5) words. Each sentence has been recorded separately and saved as .wav file format to make the speech database. In segmentation this research expects only properly segmented words as segmentation output, but the program produced some sub-words. The developed system achieved the average segmentation accuracy of **90.58**%; the details result of segmentation is given in Table-1.

TABLE I. SPEECH SEGMENTATION RESULTS

Speaker ID	No. of Sentences	No. of Words Present	No. of Properly Segmented Words	Accuracy (%)
S1	100	656	517	78.81
S2	100	656	601	91.62
S3	100	656	612	93.29
S4	100	656	619	94.36
S5	100	656	622	94.82
Total	**500**	**3280**	**2971**	**90.58**

VIII. CONCLUSION

The main objective of this research is to develop an efficient system that can automatically segments words from the continuously spoken Bangla sentences. This research introduces some ideas to develop the system. This research proposes dynamic thresholding algorithm a new approach, named "Blocking Black Area" method to detect proper word/sub-word boundaries in speech segmentation. Some words are not properly segmented. No or very little gap between two successive words causes two or more words in a single segment. Also the gap within a word causes sub-word segmentation. This is due to some sources of variability is speech, such as, Phonetic identity (two samples might correspond to different phonetic segments), Pitch and Amplitude, Speaker (based age, sex, emotion, etc.), Microphone and Media, and Environment (including background noise, room acoustics, distance from microphone, etc).

For further improvements and expansions of the speech segmentation developed system, this research can be employed by using noise reduction algorithms in a noisy environment. Also a fuzzy logic based speech segmentation approach can be employed.

REFERENCES

[1] Okko Rasanen, "Speech Segmentation and Clustering Methods for a New Speech Recognition Architecture", M.Sc Thesis, Department of Electrical and Communications Engineering, Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, Espoo, November 2007.

[2] Mermelstein P, "Automatic segmentation of speech into syllabic units", Journal of Acoustical Society of America, Vol. 58, No. 4, pp. 880-883, Oct. 1975.

[3] S L Mattys, P W Jusczyk, "Phonotactic cues for segmentation of fluent speech by infants", Cognition 78, 91–121, 2001.

[4] Zhang T and Kuo C C J, "Hierarchical classification of audio data for archiving and retrieving", Proceedings of the Acoustics, Speech, and Signal Processing 1999 on 1999 IEEE International Conference, Vol. 6, pp. 3001-3004, 1999.

[5] Antal M, "Speaker Independent Phoneme Classification in Continuous Speech", Studia Univ. Babes-Bolyal, Informatica, Vol. 49, No. 2, 2004.

[6] D Dahan and M R Brent, "On the discovery of novel word like units from utterances: an artificial-language study with implications for native-language acquisition", J. Exp. Psychol. 128 (1999) 165–185.

[7] Thangarajan R and Natarajan A M, "Syllable Based Continuous Speech Recognition for Tamil", South Asian Language Review VOL.XVIII, No.1, 2008.

[8] Hioka Y and Namada N, "Voice activity detection with array signal processing in the wavelet domain", IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, 86(11):2802-2811, 2003.

[9] Beritelli F and Casale S, "Robust voiced/unvoiced classification using fuzzy rules", In 1997 IEEE workshop on speech coding for telecommunications proceeding, pages5-6, 1997.

[10] Qi Y and Hunt B, "Voiced-unvoiced-silence classification of speech using hybrid features and a network classifier", IEEE Transactions on Speech and Audio Processing, I(2):250-255, 1993.

[11] Basu S, "A linked-HMM model for robust voicing and speech detection", In IEEE international conference on acoustics, speech and signal processing (ICAASSP'03), 2003.

[12] Kvale K, "Segmentation and Labeling of Speech", PhD Dissertation, The Norwegian Institute of Technology, 1993.

[13] Lawrence Rabiner and Biing-Hwang Juang, "Fundamentals of speech Recognition", Prentice Hall, Englewood Cliffs, N.J., 1993.

[14] Sharma M and Mammone R, "Blind speech segmentation: Automatic segmentation of speech without linguistic knowledge", Spoken Language, 1996. ICSLP 96. Proceedings. Vol. 2, pp. 1237-1240, 1996.

[15] Schiel F, "Automatic Phonetic Transcription of Non-Prompted Speech", Proceedings of the ICPhS 1999. San Francisco, August 1999. pp. 607-610, 1999.

[16] Shapiro, Linda G. and Stockman, George C., "Computer Vision", Prentice Hall, ISBN 0-13-030796-3, 2002.

[17] Md. Mijanur Rahman and Md. Al-Amin Bhuiyan, "Dynamic Thresholding on Speech Segmentation", IJRET: International Journal of Research in Engineering and Technology, Volume: 02 Issue: 09, Sep-2013.

[18] Gonzalez, Rafael C. & Woods, Richard E, "Thresholding", In Digital Image Processing, pp. 595–611. Pearson Education, 2002.

[19] Nobuyuki Otsu, "A threshold selection method from gray-level histograms", IEEE Trans. Sys., Man., Cyber. 9 (1): 62–66, 1979.

Innovative Processes in Computer Assisted Language Learning

Khaled M. Alhawiti
Computer Science Department, Faculty of Computers and Information technology
Tabuk University, Tabuk, Saudi Arabia

***Abstract*—Reading ability of an individual is believed to be one of the major sections in language competency. From this perspective, determination of topical writings for second language learners is considered tough exam for language instructor. This mixed i.e. qualitative and quantitative research study aims to address the innovative processes in computer-assisted language learning through surveying the reading level and streamline content of the ESL students in the classrooms designed for students. This study is based on empirical research to measure the reading level among the ESL students. The findings of this study have revealed that using the procedures of language preparing such as shortened text as well as assessed component tools used for automatic text simplification is profitable for both the ESL students and the teachers.**

Keywords—Natural Language Processing; Computer Assisted Language Learning; Syntactic Simplification Tools

OUTLINE

This paper will encompass various sections such as introduction to the topic, materials and methods, results, discussion, conclusions, and future work. In the first part named as "introduction," it will describe the different levels of English as Second language (ESL) learners in the United States as well as their specific needs to be successful immigrant students. The second part of this paper will encompass the prior researches conducted relevant to the topic of interest. The subsequent part i.e. results and discussion will focus upon the provision of results of the data gathered through primary sources. The final part of this paper will conclude the study along with the provision of future work.

I. INTRODUCTION

The Educational system of the United States is confronted with the testing assignment of instructing developing quantities of understudies for whom English is a second language. Washington had 72,215 understudies (7.2% of all understudies) between the school year 2001 and 2002. These understudies were related to the LEP, known as Limited English Proficient. From a year, 2003 onwards more than 2.9 million understudies got English language learner (ELL) administrations, including 19% of all government funded school understudies in California and 20% of all understudies in Texas [1]. In any case, in 2001-2002, 21% of LEP understudies had been in the project for more than three years.

On the other hand, reading is considered to be the basic piece of language and educational advancement, yet discovering appropriate reading material for LEP understudies is considered frequent upsetting. To help the learners who lie below the evaluation level, the educators with bilingual instruction search out "high investment level" writings at "low reading levels." For example, writings at a first or 2nd grade level back the 5th grade science educational program. Evaluated course readings and different materials are accessible, yet these do not meet the high investment/low reading level model.

In addition, learners also need to be engaged in supplemental reading outside of evaluated reading material for class ventures. Educators additionally require finding material, which consists of a blend of levels, since understudies need distinctive writings to peruse autonomously vs. with assistance from the instructor. This study address the issue by creating computerized instruments to help instructors and understudies discover reading-level fitting writings matched to a specific point to assist furnish these understudies with more intelligible reading material [1].

The term "Natural Language Processing" (NLP) refers to the innovation that is a perfect asset for computerizing the errand of selecting fitting material for reading understudies for bilingual. Data recovery systems effectively find "topical materials" and many of them answers complex questions in content databases on the World Wide Web. A compelling mechanized approach used for evaluating the reading level of the recovered content is still required [2]. Notwithstanding understudies in bilingual instruction, these devices will likewise be valuable for individuals with inabilitiy for learning and education understudies for adults. In both of these circumstances, the understudy's level of reading does not match their educated level as well as investments.

A. BILLINGUAL EDUCAITON

The term "Bilingual education" refers to the different approaches of teaching to the students who have the ability to communicate in multiple languages. From this perspective, "English as a Second Language" (ESL) programs are designed to prepare immigrant students for English-speaking classrooms. Some schools also offer bilingual programs recognized as "dual language immersion" in which throughout the school day two languages are used. In various ESL classrooms, students are ranging from various levels. They require various text levels to review with the assistance and individually involving instructors to locate or generate a great text variety. It would not be wrong to state that teachers working for dual language programs have to face the

challenge. This is mainly due to that, students may learn a few subjects in either of the language.

B. COMPUTER ASSISTED LANGUAGE LEARNING

Combined with advanced information technologies (IT), communication systems are considered key to the information society. IT is a baked information revolution, as it gives new human intelligence and vast capabilities. It would not be wrong to state that IT provides resources, and change the way people work and live. It is a new way of living and working together, a new means for communication and interaction in 21st century. The training needs of citizens extend beyond the first studies leading to a degree and extend throughout their lives. The introduction of computers and ICT in the classrooms is mainly due to three perspectives. These perspectives may include initially that students learn about computers because they focus their interest on the technological components. Secondly, they learn computers to employ a range of tailored programs for teaching. Finally, learning with computers and using them as tools give the students benefits of all their applications and connectivity [3].

The computer-assisted language learning is described as incorporation of unit CPU or process and peripherals (monitor, keyboard, mouse, microphone, speaker, video camera, and printer). In the same way, it is also described as a series of texts-processor software, browsers, and educational games. It is an educational application of Information and Communications Technology (ICT). These new technologies help the student to learn more effectively, by allowing them to learn or practice the target language at any time and communicate with the teacher remote, enabling e-learning [3]. Modern multimedia programs often offer an attractive presentation to the user/student to combine text with animated sequences of images and sound. Today the computer is already part of the service of instrumental ELE student at the University, in language schools in self-learning centers [4].

The language learning and computer-assisted instruction (computer-assisted language) are the two complementary facets of the same phenomenon, one from the perspective of the learner and the other from the perspective of the teacher. Computer programs and materials designed for ALAO have some drawbacks. By its very nature, it can hardly pick up all the nuances of human language and the negotiation of meaning that occurs in verbal interaction, such as gestures and intonation [4]. Moreover, these programs do not always cover the various language skills as naturally happens in the human language. Yet, ALAO offers the students a number of advantages, such as:

1) The student chooses a menu among various options for the item, the level of language difficulty, etc.

2) Each student chooses their study schedule; work at their pace, repeating a difficult activity and overlooking another who is not interested

3) The student is frequently the center of the learning process; this helps the student to take responsibility for his or her learning, while contributing to the formation as an independent learner

4) Students get help and correction. A machine usually afflicts to correct the error except when an individual edits it. The computer encourages self-evaluation

C. IMPACT OF TECHNOLOGY ON LEARNING PROCESS

Technology has positively affected the learning process of students in today's society. It has shown that there are many benefits of education in our society. The application of information technology is a feasible and necessary activity. These new technologies are affecting the educational world, particularly the students in the field of training. It is particularly because the media plays a powerful role in the learning process of educational training in term of multiplication. On the other hand, the notion known as "natural language processing" serves as an essential tool to teach the students, specifically the ones with special needs [4].

In a world where distance shrinks and borders disappear, growing mobility of people has a great significance, as they provide an opportunity to work together and to solve problems. The field of education enjoys this scenario with new forms. It appears that the inclusive schools try to solve and address the educational needs of different social groups or sectors of informal education. They find the ICT as a general contributor to the solution of such problems [5].

D. NLP TOOLS FOR LANGUAGE LEARNERS

Currently, there exist few Natural Language Processing (NLP) tools to support teachers as well as language learners. In this way, the existing systems concentrate mainly on constructing vocabulary and involve a large amount of human intervention. The observation of Horst et.al reveals that whereas the reading is a useful technique for vocabulary building, it also helps the readers to recognize new words in accordance to the context. Text Ladder system of Ghadirian classifies selected articles by teachers for the optimum arrangement of vocabulary. At the same time as functional, it is a requisite for the teacher to locate all the stories physically; the tool mainly sorts them [5].

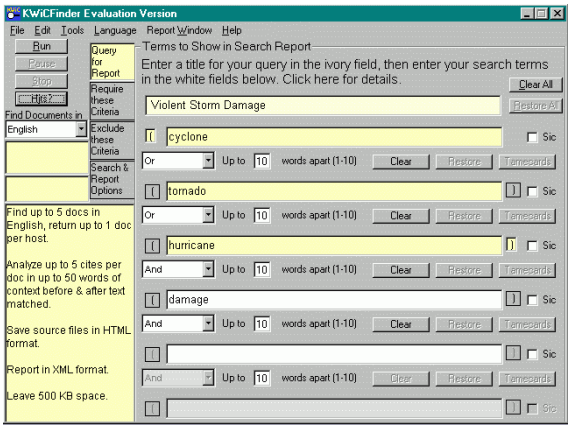


Fig. 1. Fletcher`s KWiC Finder

Fletcher has successfully developed a concordance tool to be used for "foreign language teaching" (FLT). He often makes use of web in the EFL classes in order to find examples related to the FLT.

In this way, he easily checks the questionable usage of various words. On the other hand, he also finds new words, particularly, which have not yet made it into dictionaries. His tool named as "KWiCFinder" finds examples of "keywords in context" (KWiC) and routinely constructs synopsis documents. It was intended mainly as a filter to pace up the job of locating examples of particular words. The research conducted at the "Carnegie Mellon University" contributed in the development of REAP. It is an intelligent system meant for tutoring and constructing to identify online reading material particularly for ESL students. It is particularly based on grade level unigram models, curriculum, and the level of reading of students individually with the incorporation of grammar rules. For example, the rule known as handcrafted is present in the most up-to-date version of the system [6].

For the adult ESL students at grade 6-8, REAP is mainly targeted. A pre-processed database of articles is constructed by the system that entails several words from the Academic World List (570 words). It is a fact that the exact choice of word list is not a primary element of the system. The approach of REAP mainly focuses on students acquisition at an individual level, facilitated through a gram model usage. Our Study considering the same context believes that it would be useful to revisit the complex syntactic structure within the test as well as the short phrases.

In contrast, acquisition for vocabulary remains the vital part of language learning, the content material however needs to be learned by the students. Brunelle & Boonthum-Denecke (2012) explain how their work differentiates from previous studies where they have focused on more features. These features are structural in nature of reading level used particularly to permit the specification for the users regarding the topic discretely from the level [20].

1) Goals and Contribution

This study underlines two main objectives. Firstly, the application of natural language processing and its existing technology to the issues faced in bilingual education. These issues can be related to either teachers or students. Secondly, enabling and enhancing the modern approaches in the area of NLP [18].

The mentioned purpose is also used for the development of the tool that would prove to be helpful in assisting instructors and teachers to locate appropriate level. This assessment can also be done through the development of new techniques of assessment at the reading level and simplification of text to make significant contributions. Such contributions would be mentioned further in this study [7].

2) Reading level detection

Reading level detection is the primary example that would be constructed for corpus of articles that are clean text. It thereafter extends these detectors to be applied to web retrieved pages by the use of a standard search engine. That is how the developing such reading level detectors that are trainable for plain text [7]. These detectors act as SVM classifiers (Support vector machine. They include the usual grade level features including "parser" functions, "n gram" language model scores, and so on.

3) Extension of Read in Level Detectors for Web Pagesdetection

The static group of excellent quality test to the dynamic recourses of test was found to enable to produce additional challenges. It was eventually found that those "web pages" returned inclusive of several pages that trained through detectors on "clean text" merely are not constructed to hold [8]. To discriminate the "web pages" along with text, narrative in nature from those that chiefly have advertisements, links, or other unnecessary content, that would substantially reduce the amount of discarded pages by approximately 50% [8].

4) Investigation of Extension of Algorithm for adjustment Detectors for Individuals

The assessment at the reading level is a variable and subjective issue. Various annotators have diverse insights of the suitability of "articles for a particular grade level." This is partially due to the inconsistency between the students while working. One of the major aims is to develop the detectors that may be modified in the active learning style or in the significance feedback. In this way, one can learn the application of existing processes of SVM to this task. In order to meet the requirements of each user, it is necessary to adapt the reading level detectors via observations from each instructor [9]. It was impossible to attain developments using existing techniques what were known as SVM adaptation. Nevertheless, these were created on non-text-based tasks.

5) Theory/Calculation

The main purpose of the study was to present an investigation of corpus of manual and original news articles that are simplified. The main goal of the analysis is to gain an insight about the practices of people to simplify text to frame enhanced development tools. Burstein et.al approached to choose synonyms for the words they saw challenging. They had the opinion that the use of synonyms could be used to simplify the vocabulary [7]. Instead of the concept of synonym, this section aims to be presenting the analysis of corpus that is related to the manually simplified and original news-articles. The research focused on gaining answer to following questions:

- Suggest the differences in usage of phrase types and POS, found in simplified in comparison to original sentences?
- Identify the traits of the dropped sentences, in resultant of simplified article?
- Identify the traits of split sentences, in resultant of simplified article?

Unlike other studies, our study laid emphasis on corpus of manual and original sentences that are simplified. This study incorporated corpus of paired articles, however it is to be noted that each sentence selected may not have resultant simplified sentence. The corpus of this study makes it possible to discover where the sentences have been dropped and simplified by the rewriters.

6) Aligned Corpus of News Articles

This study took 108 authentic news articles that had parallel-abridged editions framed by literacy works. These

literacy works consists of websites for instructors and learners. The target audience selected for articles was native speakers accompanied with poor reading skills.

TABLE I. CORPUS OF 108 PAIRS(ARIBDGED/ORIGNAL)

	Original	Abridged
# of Sentences(Total)	2439	2359
Words (Total)	40282	28584
Length of Sentence (Avg. Words)	15.5	13.0

7) Corpus Statistics

Upon analysis, it was deduced that number of abridged sentences was nearly equal to original sentences. It was also deduced that there were 29% fewer words in the abridged article set, and the average length of the sentences was 16% shorter in the used set as shown in table 1. In order to explore the differences among abridged and original sentence the study made use of automatic parser. The main purpose of the automatic parser was to acquire tags for parts of speech and parses for sentences.

Table 2 signifies the average length of abridged sentences was 16% shorter; therefore, fewer POS tags and words per sentence were fewer. It was noticed that there was a percentage decrease in the average frequency for adjectives, coordinating conjunctions, and adverbs. There was a 31% decrease for nouns, and 45% for pronouns, which denoted that nouns are unlikely to be replaced with pronouns and deleted less often.

TABLE II. AVERAGE FREQUENCY, SELECTED POS TAGS (ORIGNAL/ABRIDGED/DIFFERENCE)

Tag	Original	Abridged	Difference (%)
Adjective	1.3	0.8	38%
Adverb	1.1	0.5	55%
CC	0.6	0.4	33%
Determiner	1.8	1.3	27%
IN	1.7	1.4	18%
Noun	3.2	2.2	31%
Proper Noun	1.4	1.0	28%
Pronoun	1.1	0.6	45%
Verb	2.1	1.5	28%

8) Original and Aligned Sentences

The original sentences were distributed in categories based on alignment explained in the above section. This categorization allowed us to drop or align sentences to "one or more abridged sentences". The sentences, that were aligned to precisely the other sentence. In this way, the study calculated the length of the abridged sentence. The study calculated whether the abridged sentence is 19.5% shorter, longer, or roughly equal to the length of the original sentence.

The Sentence is hypothesized to be split that is associated to more than single sentence that is abridged [10]. Similarly, sentences that are aligned to a single and a shorter sentence are assumed as split with one part of the sentence that is dropped. However, it is to note that the average length of these sentences is longer than that of sentences in other categories. Nevertheless, the standard sentence length in such categories is comparatively longer [22].

TABLE III. ALIGNMENT (ORIGINAL TO ABR. SENTENCES)

Category	# of Sentences (%)	Avg. length
Total	2439 (100%)	16.5
1 - 0 (dropped)	663 (30%)	14.1
1 - >=2 (split)	370 (19%)	24.6
1 - 1 (total)	988 (47%)	15.8
1 - 1 (shorter abr.)	320 (14%)	21.0
1 - 1 (same length abr.)	525 (29%)	14.4
1 - 1 (longer abr.)	103 (4%)	9.1
2 - 1 (merged)	127 (7%)	14.6

9) Annotating True Split Sentences

In this study, around 20% original sentences are adjusted to more than single abridged sentence. The study also assumed that sentences with one part dropped that are aligned to the shorter abridged sentences could be split. On the other hand, sentences having no split points were categorized as "edited," and the sentences conveying same information were marked as "different" [11]. As shown in Table 4, the original sentences were spread among 3 categories i.e. the hypothesized, one-to one splits, and one-to-many splits. In addition, it is not surprising that making a new sentence that is somewhat shorter seems more plausible as compared to the sentence that is changed into two new sentences having no obvious split points [12].

TABLE IV. HYPOTHESIZED SPLIT- SENTENCE, (DISTRIBUTION)

	# of Sentences	
Category	1 to Many (%)	1 to 1 (%)
Total	441 (100%)	365 (100%)
True split	356 (80%)	198 (54%)
Edited	16 (4%)	162 (45%)
Different	69 (16%)	5 (1%)

10)Analysis of Split vs. Unsplit Sentences

The first and foremost step in simplification that is automatically done via sentence selection to split. The study selected long sentences to be split, incorporating other characteristics as well. For the purpose of in-depth analysis, 1675 sentences were used. The can be identified as 356 "true splits" to 1319 un split sentences mentioned in Table 4. For this study, the different sentences as well as the edited sentences are assumed split sentences as they are measured unsplit. As a matter of fact, the average amount of phrases

"identified by the parser S, NP, etc." as well as the average length of the phrases is longer. Therefore, it is assumed that the split of sentence depends upon the syntactic features in addition to the length of a particular sentence. [13]

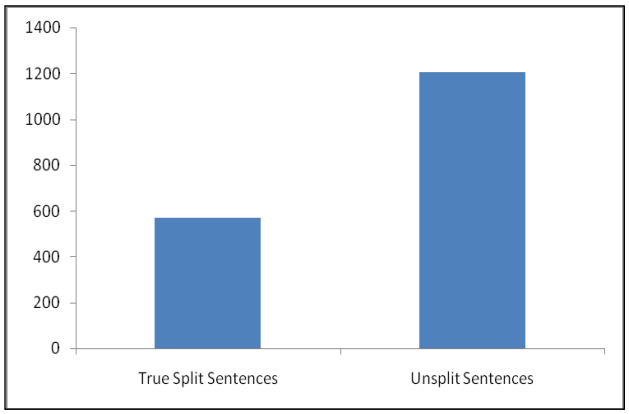


Fig. 2. True Split Sentences vs Unsplit Sentences

In order to examine the most important feature for splitting sentence the study made use of "C4.5 decision tree learner" to build a classifier for unsplit and split sentences. From this perspective, the rule generators of C4.5 were selected for this study considering the easiness in the emerging results. This can be described in other words that in this section the major focus was on analysis instead of classification. From this perspective, the study incorporated a few features of sentence. The features, on the other hand, may include "Sentence length in words," "POS", and "Phrase." In other words, POS include number of adverbs, adjectives, CC, determiners, IN, nouns, pronouns, proper nouns, and verbs. On the other hand, the term "Phrase" may include average length and number of SBAR, S, PP, NP, and VP.

11)Analysis of Dropped Sentences

The researcher tends to attempt, in this section, to evaluate the "dropped sentences" in the comparison with the other original sentences. As "C4.5 rule generator" was used in the previous section in order to see the most important feature of sentence, "the basis for analyze the sentences to drop" seems more credible to see the dropped sentences. On the other hand, it is more credible to be content-based as compared to the syntactic. In this way, content-based is considered to be a quite different group of feature in the present section. It is forecasted instinctively that the replicated sentences can be dropped as well.

II. Results and Discussion

Based on the prior work above, the study decomposed the problem, what is known as simplification, into the following four component problems: sentence selection, sentence splitting, sentence compression, and lexical replacement.

The decoupling of such processes is useful for recognition their role in simplification, and for leveraging existing text processing tools. The key aspect is Sentence selection, furthermore, component of summarization systems and major research addresses this problem. Summarization systems targets to extract the most vital information from an article but majority select longer and more complicated sentences since those sentences convey major information [14]. By the same token, the articles that are simplified in the Literacy works corpus analyzed in paper above are seventy percent as long as the original articles.

In this study, the Literacy works corpus hand-aligned was studied to train the models for the selection of sentence and then splitting. The use of selected set developed from "Literacy works corpus" as well as "two test sets selected from two sources on the web" were used to evaluate the mentioned models as well as the processing tools of the existing language. The assessment of "automatic simplification" is considered to be quite similar to the compression tasks and closely-related summarization, as it generally involves the evaluation that is to be handwritten. This way is also associated with the human decision-making or gold standard as it is related to the "quality of the results." On the other hand, a huge variety of articles "from the Literacy works corpus" has been selected for the present study in order to use the development set as well as for the experiments.

The study created two additional test sets. The initial consists of usage of 5 articles from the online Principles of Aeronautics textbook written by Cislunar Aerospace, Inc.

These articles selected were about the same length as the Literacy works articles, and pointed out the factor that the "average sentence length" is longer for the corpus comparatively. Nevertheless, the "average sentence length" that is longer will represent the different challenge for the purpose of simplification though the mentioned articles are included in the similar "news domain" as the Literacy works corpus are.

Several of other systems do not use the quotation feature, rather than that it appears to be useful for the mentioned corpus. It is believed that it is helpful in general for such application. The content features represent the related purpose of the "tf-idf features used by Nobata and Sekine". Nevertheless, such features are to offer "information about content at the level of a single document instead of requiring a corpus of related documents." On the other hand, some features are mainly used in the extractive summarization. Nevertheless, these features are not quite relevant to the unimpeded domain.

Using the IND "decision tree package", the instructor may train "the classifier" with the mentioned features that are mainly based on dropped and not dropped sentences in the interpreted corpus. The IND usage in the place of "C4.5" is thus because of the package that provides probability to estimate the classifications. Eventually, to apply the classifier to the new feature known as "vector x for a sentence", all of the B trees 7) may be applied to "the vector and the resulting posterior probabilities" that are initially averaged and then regularized to the account for "re-sampling", i.e.

$$p(drop|x) = \frac{q_d}{0.5B} \sum_{i=1}^{B} p_i(drop|T_i(x))$$

where B = 48 and = 0.28 is the previous prospect of the "dropped" class. The application of the classifier to developing the set gives way performance, which is not much better than probability. From this perspective, the present paper did not follow any additional evaluation of the classifier on the set test. As a matter of fact, the present study is intended to compare to the set results of the development along with the "summarizer results" occurring in the subsequent section. The manual review related to the "various decision trees" identify that "all of the types of feature used, with the "quote" feature that often occur near to the top of tree [15]. The both feature categories i.e. redundancy and position emerge significant. It is unclear that what category is used consistently than the other one. The presence of complexity in using commands a language that is more rewarding than GUI [17]. However, they are still being used in various applications, typically without the users' intentions. People, however, do not notice simple control line interfaces that have been integrated into systems, which include a search box on the Web and page range box in Microsoft Word's printing options. Unnoticeable interface is a sign of a good interface. An extensive variety of functions can be fluently provided by a text interface in comparison to GUI, which lack scalability. The major cause is that the text is so much lighter than graphics [16]. The aptitude for pure texts to be effortlessly view, copy, paste, edit, stored, and share is accessible in almost every user interface and application

III. CONCLUSIONS

The purpose of the present research was to apply as well as to extend the existing "NLP technology" to be used for the problems faced by students as well as teachers within the context where bilingual education takes place. In this way, the study advanced "the state of the art in the relevant NLP areas". The paper has discussed in detail about the tools that support the instructors in order to choose appropriate-level and topical texts to be used for their students by the use of innovations and techniques of reading assessment. This study also explored the "characteristics of abridged text" as well as it "assessed component tools for automatic text simplification". It would not be wrong to state that this study contributed to the area where "reading level assessment" takes place and includes the "development of reading level detectors for clean text and extending them to a most varied text found on the World Wide Web". The researcher found that the combining "SVM classifiers" within the traditional grade-level, the scores of the models named as n-gram language, and the features that are parser-based are based on the LMs alone in form of n-gram.

Through the study, it was established that various methods of SVM adaptation have been developed for the "non-text-based tasks with well-separable components do not essentially apply to this task". Nevertheless, the study finds that "other classifier/adaptation combinations will provide better results in the future". With the use of Literacy works as well as "news article corpus", it was suggested that the further studies will characterizes the texts to expand insight into what people generally do when performing the type of text-adaptation. The study concluded that the hand-aligned sentences were to illustrate that how sentences in the original description and abridged description relate to one another other. From this perspective, the "split points" were marked at the places where an original sentence is mapping to two "abridged sentences". The resulting corpus in such observations will be the contributing factor to the "field of NLP research." Therefore, it could be used for the purpose of research on the topic that includes sentence alignment, simplification, and summarization [19].

IV. FUTURE WORK

There are various possible paths for the work that will be conducted in the future that might be based on this research. In this section, the author discusses the four most important directions such as future work for simplification, future work for adaptation, creating the system that can be used for interaction between teachers and students, and application to languages that are other than target language. The basic focus of the author was to develop the tools that can be used for English language. Nevertheless, the research would be equally helpful for students as well as teachers to find as well as simplify the texts available in the target language. On the other hand, the other then English languages English may have advantage from "the use of additional features" as well. For example, it can be used to capture the richer morphology. In the same way, extending and modifying the tools that have been developed in the research is the additional area for future work, as it is to create a system that may work in real-time.

REFERENCES

[1] J. Amaral, L. A., & Meurers, D. (2011). On using intelligent computer-assisted language learning in real-life foreign language teaching and learning. ReCALL,23(01), 4-24.

[2] Chapelle, C. A. (2010). The spread of computer-assisted language learning.Language Teaching, 43(01), 66-74.

[3] Cohen, K. B. (2013). Biomedical Natural Language Processing and Text Mining.Methods in Biomedical Informatics: A Pragmatic Approach, 141.

[4] Tyagi, D., Joshi, T., Ghule, D., & Joshi, A. (2014). An Interactive Answering System using Template Matching and SQL Mapping for Natural Language Processing. International Journal, 2(2).

[5] Kamath, R. S. (2013). Development of Intelligent Virtual Environment by Natural Language Processing. Special issue of International Journal of Latest Trends in Engineering and Technology.

[6] Field, D., Richardson, J. T., Pulman, S., Van Labeke, N., & Whitelock, D. (2014). An exploration of the features of graded student essays using domain-independent natural language processing techniques. International Journal of e-Assessment, 4(1).

[7] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[8] Reeves, T., & McKenney, S. E. (2013). Computer-assisted Language Learning and Design-based Research: Increased Complexity for Sure, Enhanced Impact Perhaps.

[9] Song, P., Shu, A., Zhou, A., Wallach, D., & Crandall, J. R. (2012). A pointillism approach for natural language processing of social media. arXiv preprint arXiv:1206.4958.

[10] Denny, J. C., Choma, N. N., Peterson, J. F., Miller, R. A., Bastarache, L., Li, M., & Peterson, N. B. (2012). Natural language processing improves identification of colorectal cancer testing in the electronic medical record.Medical Decision Making, 32(1), 188-197.

[11] Nakata, T. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: a critical investigation of flashcard software.Computer Assisted Language Learning, 24(1), 17-38.

[12] Esit, O. (2011). Your verbal zone: an intelligent computer-assisted language learning program in support of Turkish learners' vocabulary learning. Computer Assisted Language Learning, 24(3), 211-232.

[13] Kennewick, R. A., Locke, D., Kennewick, M. R., Kennewick, R., & Freeman, T. (2011). U.S. Patent No. 8,015,006. Washington, DC: U.S. Patent and Trademark Office.

[14] Gorjian, B., Moosavinia, S. R., Ebrahimi Kavari, K., Asgari, P., & Hydarei, A. (2011). The impact of asynchronous computer-assisted language learning approaches on English as a foreign language high and low achievers' vocabulary retention and recall. Computer Assisted Language Learning, 24(5), 383-391.

[15] Litman, D., Moore, J. D., Dzikovska, M., & Farrow, E. (2010). Using natural language processing to analyze tutorial dialogue corpora across domains and modalities.

[16] Meurers, D. (2012). Natural language processing and language learning. The Encyclopedia of Applied Linguistics.

[17] Jarvis, S. (1984). Language Learning Technology and Alternatives For Public Education. CALICO Journal, 1(4), 11-16.

[18] Michos, S. E., Fakotakis, N. & Kokkinakis, G. (1996). Towards an adaptive natural language interface to command languages.. Natural Language Engineering, 2, 191-209.

[19] Lewis, D. D., & Jones, K. S. (1996). Natural language processing for information retrieval. Communications of the ACM, 39(1), 92-101.

[20] Kerr, D., Mousavi, H., & Iseli, M. (2013). Automatic Short Essay Scoring Using Natural Language Processing to Extract Semantic Information in the Form of Propositions. CRESST Report, 831.

[21] Brunelle, J. F., & Boonthum-Denecke, C. (2012). Natural Language Processing Tools. Cross-disciplinary Advances in Applied Natural Language Processing: Issues and Approaches, 9.

[22] Murff, H. J., FitzHenry, F., Matheny, M. E., Gentry, N., Kotter, K. L., Crimin, K., ... & Speroff, T. (2011). Automated identification of postoperative complications within an electronic medical record using natural language processing. Jama, 306(8), 848-855.

Comparative Analysis of Improved Cuckoo Search(ICS) Algorithm and Artificial Bee Colony (ABC) Algorithm on Continuous Optimization Problems

Shariba Islam Tusiy[1], Nasif Shawkat[2], Md. Arman Ahmed[3], Biswajit Panday[4], Nazmus Sakib[5]
Ahsanullah University of Science & Technology (AUST), Dhaka, Bangladesh

***Abstract*—This work is related on two well-known algorithm, Improved Cuckoo Search and Artificial Bee Colony Algorithm which are inspired from nature. Improved Cuckoo Search (ICS) algorithm is based on Lévy flight and behavior of some birds and fruit flies and they have some assumptions and each assumption is highly observed to maintain their characteristics. Besides Artificial Bee Colony (ABC) algorithm is based on swarm intelligence, which is based on bee colony with the way the bees maintain their life in that colony. Bees' characteristics are the main part of this algorithm. This is a theoretical result of this topic and a quantitative research paper.**

Keywords—Artificial Bee Colony (ABC) algorithm; Bioinformatics; Improved Cuckoo Search (ICS) algorithm; Lévy flight; Meta heuristic; Nature Inspired Algorithms

I. INTRODUCTION

Beautiful nature is full of surprises and mystery. People have learnt a lot from the Mother nature. By analyzing symptoms people manage to reveal the mystery of nature. As time changes, humans also change their characteristics and their behavior to the nature. Now a day's people find solutions of their daily life problems with the help of nature and that is known as meta-heuristic solutions. The bee colony and the improved cuckoo search algorithm elevate the eco-life system in a new level. On the basis of key functions and iteration number, the comparison between Artificial Bee Colony and Improved Cuckoo Search algorithm is done. Artificial Bee Colony works on the optimization algorithm introduced by D. Karaboga[1]. And the Improved Cuckoo Search algorithm is extended to more complicated cases in which each nest has multiple eggs representing a set of solutions[2][3][4]. Within last few decades, dozens of meta-heuristic algorithms are published and still been publishing. Among them Bat [5][6], Firefly [7][8], Flower Pollination [9], Artificial Bee Colony [10], Improved Artificial Bee colony [11], Ant Colony [12], Cuckoo search [13] is highly recommended algorithms. The algorithms which have mentioned above are upgrading day by day. So, here it has been focused on the implementation and the operations of the iteration number, and the tested functions for both algorithms that mentioned above are same. For preparing this research, first of all, the data of mean and median for improved cuckoo search have been measured and the algorithm is obtained. Then the comparison makes them different from each other. By producing graphical outcome, it is observed that improved cuckoo search is good enough. Improved cuckoo search (ICS) & its algorithm is being described in section II. Then in section IV the artificial bee colony (ABC) is being described with its algorithm. After that in section V the simulation & analysis part is being described and then the findings in section VI. Finally, in section VII the total work is being summarized in short in the conclusion.

II. CUCKOO SEARCH

A. Basic Ideas of Cuckoo Search

Cuckoo Search (CS) is used to solve optimization problems which are a meta–heuristic algorithm, developed by 'Xin-She Yang' that is based on the manner of the cuckoo species with the combination of Lévy flight behavior of some birds and fruit flies [14][15]. The inspiration behind developing Cuckoo Search Algorithm is the invasive reproductive strategy and the obligate brood parasitism of some cuckoo species by laying their eggs in the nest of host birds [16]. Some female cuckoo like Guira and Ani can copy the patterns and colors of few chosen host species. This imitates power is used to increase the hatching probability which bring their next generation. The cuckoo has an amazing timing of laying eggs. Parasitic cuckoos used to choose a nest where the host birds lay their own eggs and it takes less time to hatch cuckoo's egg than the host bird's eggs. After hatching the first egg, the first instinct, action is to throw out the host eggs or to propel the eggs out of the nest to ensure the food from the host bird.

B. Basic Points of Cuckoo Search

Each Cuckoo's egg in a nest illustrates a new solution. The aim of Improve Cuckoo Search is to serve the new and potentially better solutions to replace the previous solutions in the Cuckoo Search. The algorithm can be extended to more complicated cases in which each nest has multiple eggs that represent a set of solutions. The CS is based on three idealized rules that are given bellow:

1) Each cuckoo lays one egg at a certain time, and dumps it in a nest which is randomly chosen [17].

2) The best nests provide high quality of eggs (solutions) that will carry over to the next generations [17].

3) A host bird can discover an alien egg from his nest with probability of Pa ∈ [0, 1]. In this case, the host bird can

either throw the egg away or abandon or can completely build a new nest in a new location [17].

C. Lévy Flights

Generally, the foraging path of an animal is successful a random walk as the next step is based on both the current location and the transition probability to the next location. The chosen direction implicitly depends on a probability, which can be modeled mathematically. The flight behavior of many animals and insects demonstrates the typical characteristics of Lévy flights. A Lévy flight is a random walk in which the step-lengths are distributed according to a heavy probability distribution. After a large number of steps, the distance from the origin of the random walk tends to a stable distribution [17][18].

III. Improved Cuckoo Search (ICS)

A. Characteristics of Improved Cuckoo Search

The parameters Pa, λ and α introduced in the CS, help the algorithm to find globally and locally improved solutions, respectively. The parameters Pa and α is very important parameters in fine-tuning of solution vectors, and can be potentially used to adjust the convergence rate of the algorithm. The traditional CS algorithm uses a fixed value for both Pa and α. The key difference between ICS and CS is the way of adjusting Pa and α. To improve the performance of CS algorithm and eliminate the drawbacks lies with fixed values of Pa and α, the ICS algorithm uses variables Pa and α. The values of Pa and α dynamically change with the number of generations and have been expressed in equations 1-3, where NI and gn are the number of total iterations and the current iteration respectively.

$$P_a(g_n) = P_{a\,max} - \frac{g_n}{NI}\,(P_{a\,max} - P_{a\,min}) \quad (1)$$

$$\alpha\,(g_n) = \alpha_{max}\exp(c.\,g_n) \quad (2)$$

$$c = \frac{1}{NI} L_n\,(\frac{\alpha_{min}}{\alpha_{max}}) \quad (3)$$

B. Algorithm of ICS

Begin

Objective function f(x), $x = (x_1, ..., x_d)^T$;
Initial a population of *n* host nests xi (i = 1, 2, ..., n);
while (t<MaxGeneration) or (stop criterion)
Get a cuckoo (say i) randomly by *Lévy* flights;
Evaluate its quality/fitness F_i;
Choose a nest among *n* (say j) randomly;
if(F_i>F_j) Replace j by the new solution; *end*
A fraction (P_a) of worse nests are abandon and new once are built.
Keep the best solutions (or nests with quality solutions);
Rank the solutions and find the current best;
end while
Post-process results and visualization;

End

When generating new solutions X_i (t +1) for the i^{th} cuckoo, the following Lévy flight is performed

$$X_i(t+1) = X_i(t) + \alpha \oplus \text{Lévy}\ (\lambda) \quad (4)$$

Where α>0 is the step size, which should be related to the scale of the problem of interest. The product ⊕ means an entry-wise multiplications. According to Yang's research work, it has considered that a *Lévy* flight in which the step-lengths are distributed according to the following probability distribution

$$\text{Lévy u} = \text{t} - \lambda,\ 1 < \lambda \leq 3 \quad (5)$$

This has an infinite variance. Here, the consecutive steps of a cuckoo essentially form a random walk process which obeys a power law step length distribution with a heavy tail.

It is worth pointing out that, in the real world, if a cuckoo's egg is very similar to a host's egg, then this cuckoo's egg is less likely to be discovered, thus the fittest should be related to the difference in solutions. Therefore, it is a good idea to do a random walk in a biased way with some random step sizes.

IV. Artificial Bee Colony (ABC)

A. Basic Ideas of Artificial Bee Colony (ABC)

The ABC algorithm is of wide range of insects that are dependent and meta-heuristic algorithm that is developed on the provision behavior of honey bee colonies [19]. The ABC is an algorithm which describes the intelligent provision behavior of honey bee swarms. It is simple, vigorous, strong and healthy and population dependent randomly determined optimization algorithm [20]. The ABC algorithm which may be used for explanation of multidimensional and multimodal optimization matters [21].

B. Some Common Mistakes

- In ABC, honey bees are classified into three groups that are named as employed bees, onlooker bees and scout bees.
- The employed bees are the bee which searches for the food source and gather the information about the quality of the food source.
- Onlooker bees stay in the hive and search the food sources on the basis of the information gathered by the employed bees.
- The scout bee, searches new food sources randomly in places of the abundant food sources.

C. Algorithm of ABC

1) Algorithm 1 Artificial Bee Colony Algorithm

Initialize the parameters;

While Termination criteria is not satisfied do

Step 1: Employed bee phase for computing new food sources.

Step 2: Onlooker bees phase for updating the location of food sources based on their amount of nectar.

Step 3: Scout bee phase for searching about new food sources in place of rejected food sources.

Step 4: Memorize the best food source identified so far.

End of while

Output The best solution obtained so far.

2) Algorithm 2 Solution update in Employed bee phase

Input: solution xi, probi and j ∈ (1, D);

for j ∈ {1 to D} do

if U (0, 1) >probi then

vij = x ij + φ ij (x ij – x kj) + ψ ij (x best j – x ij) ;

else

vij = x ij ;

end if

end for

V. Simulation and Analysis

A. GRAPH with Parameter settings

In this paper, 70 independent runs on each algorithm to get the result from the test functions which are rowed in Table-1. The population for each function is set for 14. Maximum cycle has been used 70 for both algorithms. And the Dimension for each function for each algorithm are set for D=5, 10, 15, 25 respectively. Of ICS, we use Pa, α, λ for improved the result for ICS globally and locally. Here in ABC Pa=0.25, αmin=.05, αmax=0.5. Finding the best, worst, mean, median and Standard deviation value for both algorithms is the main goal. On the basis of the result of finding the best and new place as well as nest or colony the3D surface and mesh are simulated for Rosen rock function, Ackley functions have shown in two views. In X-axis the objective value and in Y-axis two variable values is plotted. For this MATLAB R2013a version is used for simulation with 4th generation Intel i5 processor 2.7GZ with 4GB RAM of PC.

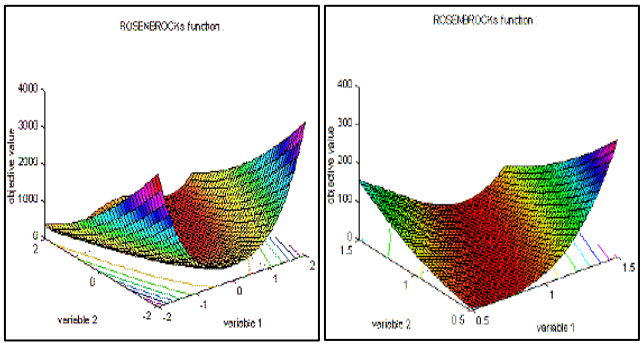


Fig. 1. 3D surface plots (2 view) of Rosenbrock function that best for ICS

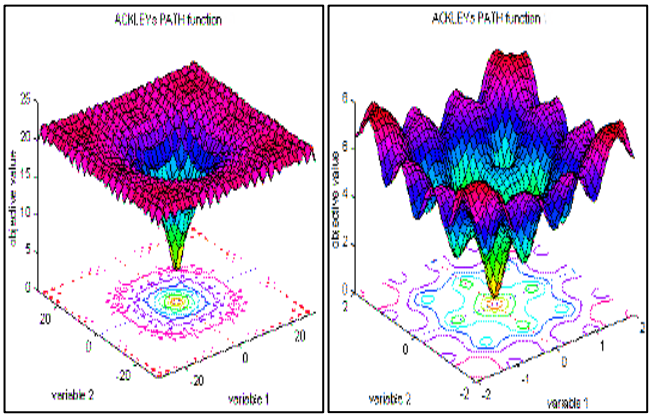


Fig. 2. 3D surface plots (2 view) of Ackley function that best for ICS.

VI. Findings

In ABC algorithm the fitness and global min is compared with ICS algorithm. The work has mainly focused on the differences between these two algorithms and its basis on the mean value and on the basis of time. It is true that, if the exploitation is too high and the exploration is too low, then algorithms may trap into locally optimal points. So, some methods are followed and tried to avoid any kinds of trap that may cause trouble. Because this could affect find the global optimum. Or if the exploration is too high and exploration is too low than exploration then the convergence speed will decrease. In the ABC algorithm we use population for about 14 and tested in f1, f2, f3, f4, f5, f6 functions with runtime 70 for both ABC and ICS algorithm. When the dimension increases ABC gives poorer results than ICS but gives good result in lower dimensions. That means ICS gives the best result in high Dimension. So, it can be said that ABC works well in exploitation, but in the exploration it works poorly. But ICS works better in exploration. In this experiment ICS shows better results for dimension 10 and 25.For dimension 10, the ICS gives better result than ABC for f1, f3, f4, f5 functions on the basis of the mean value. And for dimension 25 ICS gives better result for f1, f3, f4, f5. And for other two dimensions, it works equally as ABC. Among these functions Rosenbrock gives the best result for ICS. Basically, in ICS cuckoo search his food within a wide range of area, not in a limited range of area. That means its food area is large. On the other hand In ABC a bee only finds its honey on its own place where the least and maximum capacity honey holder bees are present .If the bee fails to find honey from other sources that has the maximum capacity of honey than the bee turns back and looks for other bees that has the maximum amount of honey [22][23][24].

TABLE I. BENCHMARK FUNCTIONS USED IN THE EXPERIMENTAL STUDIES. HERE, D: DIMENSIONALITY OF THE FUNCTION, S: SEARCH SPACE, C FUNCTION CHARACTERISTICS WITH VALUES — U: UNIMODAL AND M: MULTIMODAL

func	Name	D	C	S	Function Definition	f_{min}
f1	Sphere	5,10,15,25	U	[-5.12, 5.12]D	$f(x) = \sum_{i=1}^{d} x_i^2$	0.0
f2	Griewank	5,10,15,26	M	[-15, 15]D	$f(x) = \frac{1}{4000}\sum_{i=1}^{d} x_i^2 - \prod_{i=1}^{d} cos\frac{x_i}{\sqrt{i}} + 1$	0.0
f3	Rastrigin	5,10,15,27	M	[-15, 15]D	$f(x) = \sum_{i=1}^{d} [x_i^2 - 10\cos(2\pi x_i) + 10]$	0.0
f4	Rosenbrock	5,10,15,28	U	[-15, 15]D	$f(x) = \sum_{i=1}^{d-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	0.0
f5	Ackley	5,10,15,29	M	[-32, 32]D	$f(x) = -20\exp\left(-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n} cos\, 2\pi x_i\right) + 20 + e$	0.0
f6	Schwefel	5,10,15,25	M	[-500, 500]D	$f(x) = 418.9829 * d - \sum_{i=1}^{n} -x_i \sin\left(\sqrt{\lvert x_i \rvert}\right)$	0.0

TABLE II. COMPARISON BETWEEN ABC & ICS ON 6 STANDARD BENCHMARK FUNCTIONS. ALL ALGORITHMS ARE RUN 24 DIFFERENT TIMES ON EACH OF THE FUNCTIONS. THE BEST RESULT FOR EACH FUNCTION WITH EACH DIMENSION IS MARKED BOLD

func	Name	Algorithm	Dim	Best	Worst	Mean	Median	SD
f1	Sphere	ABC	5	1.42E-15	4.02E-12	**6.78E-13**	4.04E-14	2.10E-12
		ICS		6.90E-02	2.00E+00	1.00E+00	9.32E-01	7.92E-01
		ABC	10	0.0011	396.638	59.95378571	0.6839	137.54161
		ICS		0.0562095	6.667	**2.399571888**	0.475504	3.02238008
		ABC	15	8.15E-07	0.406	**0.077864353**	0.0018	0.21592723
		ICS		0.0515647	11.4665	3.957276326	0.353732	5.31127886
		ABC	25	2.8065	4808.4	1494.409743	882.56	2418.11741
		ICS		0.053437	9.13084	**3.194834213**	0.400225	4.19977765
f2	Griewank	ABC	5	0.0099	0.1052	**0.041857143**	0.0296	0.04682883
		ICS		0.0622457	3.70203	1.476183907	0.664273	1.59298316
		ABC	10	0.0197	0.2396	**0.0757**	0.0557	0.0693187
		ICS		0.0569547	6.15002	2.234685552	0.497087	2.77257803
		ABC	15	0.0322	0.1881	**0.073914286**	0.0375	0.09214291
		ICS		0.0596114	4.6883	1.775403562	0.578299	2.0705851
		ABC	25	0.9752	4.1287	**1.953914286**	1.1097	1.76112841
		ICS		0.0516801	11.3016	3.903275186	0.356502	5.23291644

f3	Rastrigin	ABC	5	8.79E-10	1.0087	**0.428446591**	0.00012567	0.75568014
		ICS		0.05697	6.13833	2.230957736	0.497574	2.76877797
		ABC	10	3.57E-11	7.44E+09	1298.930031	0.0777	3180.44622
		ICS		0.0669566	2.39661	**1.103110736**	0.845762	0.96833078
		ABC	15	2.0855	10.7955	6.661642857	7.5693	2.8464014
		ICS		0.0556383	7.09699	**2.537333517**	0.45937	3.22837601
		ABC	25	0.9989	9.4733	4.462328571	3.7386	4.27762977
		ICS		0.0553482	7.34006	**2.61550469**	0.451101	3.34467278
f4	Rosenbrock	ABC	5	4381.7	95705	19810.21429	8493.5	47395.5977
		ICS		0.0534698	9.09728	**3.183930407**	0.401039	4.18377811
		ABC	10	570.316	2.27E+07	4704766.516	767.163	8153588.31
		ICS		0.0547036	7.88175	**2.790039516**	0.433669	3.60370266
		ABC	15	25.9912	9.64E+09	1376914539	254.39	5151941350
		ICS		0.0518323	11.1223	**3.844661369**	0.359818	5.14762771
		ABC	25	0.2996	6459.9	962.3464143	68.6211	3428.69951
		ICS		0.051982	10.914	**3.77656329**	0.363682	5.04855246
f5	Ackley	ABC	5	1.67E+01	20.0102	19.53458571	20	1.74778989
		ICS		0.935065	12.2685	**5.024255357**	1.86915	5.13666397
		ABC	10	20	21.2061	20.31675714	20.0979	0.44159039
		ICS		0.0689939	2	**3.002425739**	0.933432	2.15179115
		ABC	15	20	20.4971	20.09742857	20.0045	0.26100873
		ICS		0.0563243	6.58119	**2.372110814**	0.478821	2.98126045
		ABC	25	20.0354	21.5963	20.4319	20.3623	0.77635362
		ICS		0.0530613	9.54667	**3.330121363**	0.39063	4.39792509
f6	Schwefel	ABC	5	-318.175	14.2288	**-238.4171**	-252.299	166.583262
		ICS		0.0615128	3.94352	1.548055762	0.639132	1.71018618
		ABC	10	-577.248	-599.325	**-690.0644574**	-607.041	88.8336281
		ICS		0.0602585	4.39303	1.684305824	0.599625	1.92797598
		ABC	15	-917.593	-864.837	**-931.2613**	-944.684	45.1666012
		ICS		0.0535746	8.98318	3.146848581	0.403787	4.12938848
		ABC	25	-1590.9	-693.199	**-1457.342729**	-1581.6	476.567607
		ICS		0.0515618	0.353507	3.968892143	0.353507	5.32786014

On the X-axis the number of generations and in Y-axis fitness is set and plot this graph (fig-3). From this graph the comparison is clarified clearly.

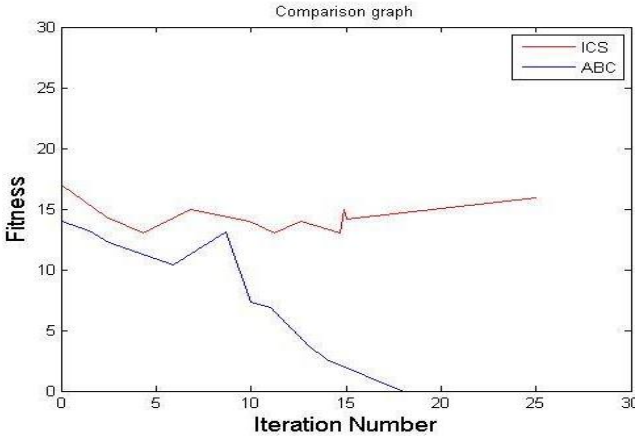


Fig. 3. 2D plot of ICS vs ABC algorithm

VII. CONCLUSION

a) This paper represents the comparative study between swarm intelligence base and Lévy flight behavior base algorithms the Artificial Bee Colony (ABC) algorithm [1] and the Improved Cuckoo Search (ICS) algorithm [2][3][4].Optimization results in the standard benchmark problems for Artificial Bee Colony (ABC) algorithm and Improved Cuckoo Search (ICS) algorithm exhibit the effective results and competitive results of the algorithms. The main reason of the performance difference is basically in ICS where cuckoo search his food within a vast area rather than limited. On the contrary, in ABC a bee only finds its own light in its own place even though the light holder bees are present which have the minimum and maximum intensity of light. And if the bee fails to find from others that hold the highest capacity of light, then it turns back and search for other bees which have the larger intensity of light. And its place is limited not wide while searching for bees. So, that means cuckoo works on a wide range of area and it needs more dimension than ABC. On the other side, ABC needs fixed area to search its best. Last of all, it can be assumed that ICS and ABC can be improved more than before in the future.

REFERENCES

[1] D. Karaboga, "An Idea Based On Honey Bee Swarm For Numerical Optimization," Technical Report-TR06,Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.

[2] X. S. Yang, "Nature-Inspired Metaheuristic Algorithms," LuniverPress,UK, 2008.

[3] M. Gendreau, "Handbook of Metaheuristics," in An introduction to tabu search, Kluwer Academic Publishers, 2003, p. 37–54.

[4] X. Yang and S. Deb, "Engineering Optimisation by Cuckoo Search," in Int. J. Mathematical Modelling and Numerical Optimisation, 2010, p. 330–343.

[5] X.-S. Yang, "Bat algorithm: literature review and applications," Int. J. Bio-Inspired Computation, vol. 5, p. 141–149, 2013.

[6] I. J. Fister, D. Fister and X.-S. Yang, "A hybrid bat algorithm," in Elek-trotehniski vestnik, press, 2013.

[7] X.-S. Yang and s. H. Xing, "Firefly Algorithm: Recent Advances and Applications," Int. J. Swarm Intelligence, vol. 1, p. 36–50, 2013.

[8] X.-S. Yang, "Firefly algorithms for multimodal optimization," Stochastic Algorithms: Foundations and Applications, SAGA 2009, Lecture Notes in Computer Sciences, vol. 5792, pp. 169-178, 2009.

[9] X.-S. Yang, "Flower pollination algorithm for global optimiza-tion," Unconventional Computation and Natural Computation 2012,Lecture Notes in Computer Science, vol. 7445, pp. 240-249, 2012.

[10] D. Karaboga, "An Idea Based on Honey Bee Swarm for Numerical Optimization.Technical Report-TR06," in Erciyes University, Computer Engineering Department, 2005.

[11] M. Kiran and A. Babalik, "Improved Artificial Bee Colony Algorithm for Continuous Optimization Problems," Journal of Computer and Communications, vol. 2, pp. 108-116, 2014.

[12] A. Farzindar and V. (. Keselj, Canadian AI 2010, LNAI 6085, 2010.

[13] I. J. Fister, X. Yang, D. Fister and I. Fister, "Cuckoo search: A brief literature review," Cuckoo Search and Firefly Algorithm: Theory and Applications, Studies in Computational Intelligence, vol. 516, pp. 49-62, 2014.

[14] X.-S. Yang, Optimization Problem, Department of Engineering,University of Cambridge, Trumpinton Street, Cambridge CB2 1PZ, UK.

[15] M. Gendreau, "An introduction to tabu search," in Handbook of Metaheuristics, Kluwer Academic Publishers, 2003, p. 37–54.

[16] E. Valian, S. Mohanna and S. Tavakoli, "Improved Cuckoo Search Algorithm for Global Optimization," University of Sistan and Baluchestan, Dec. 2011.

[17] X.-S. Yang, Cuckoo Search via L´evy Flights, Department of Engineering,University of Cambridge, Trumpinton Street, Cambridge CB2 1PZ, UK.

[18] Clifford T. Brown, "Lévy Flights in Dobe Ju/'hoansi Foraging Patterns," 777 Glades Road, Boca Raton, FL 33431, USA, Department of Anthropology, Florida Atlantic University , 6 December 2006, p. 129–138.

[19] B. Kumar and D. Kumar, A review on Artificial Bee Colony algorithm, Dept. of Computer Science & Engineering, Guru Jambheswar University, Hisar, Haryana, India.

[20] D. Karaboga and C. Ozturk, "A novel clustering approach: Artificial Bee Colony (ABC) algorithm," Kayseri, Turkey, Erciyes University, Intelligent Systems Research Group, Department of Computer Engineering.

[21] G. YAN and Chuangqin, "Effective Refinement Artificial Bee Colony Optimization algorithm Based On Chaotic Search and Application for PID Control Tuning," Taiyuan, China, College of Information Engineering, Taiyuan University of Technology.

[22] IJCIT, International Journal of Communications and Information Technology, vol. 1, Dec. 2011.

[23] IJAIA, "International Journal of Artificial Intelligence & Applications," vol. 2, July 2011.

[24] IJAIS, International Journal of Applied Information Systems, vol. 7, no. ISSN: 2249-0868, September 2014.

Speech emotion recognition in emotional feedback for Human-Robot Interaction

Javier G. Rázuri*, David Sundgren*, Rahim Rahmani*, Aron Larsson*, Antonio Moran Cardenas‡ and Isis Bonet§

*Dept. of Computer and Systems Sciences (DSV)
Stockholm University, Stockholm, Sweden

‡Pontifical Catholic University of Peru´ (PUCP)
Lima, Peru

§Antioquia School of Engineering (EIA)
Antioquia, Colombia

***Abstract*—For robots to plan their actions autonomously and interact with people, recognizing human emotions is crucial. For most humans nonverbal cues such as pitch, loudness, spectrum, speech rate are efficient carriers of emotions. The features of the sound of a spoken voice probably contains crucial information on the emotional state of the speaker, within this framework, a machine might use such properties of sound to recognize emotions. This work evaluated six different kinds of classifiers to predict six basic universal emotions from non-verbal features of human speech. The classification techniques used information from six audio files extracted from the eNTERFACE05 audio-visual emotion database. The information gain from a decision tree was also used in order to choose the most significant speech features, from a set of acoustic features commonly extracted in emotion analysis. The classifiers were evaluated with the proposed features and the features selected by the decision tree. With this feature selection could be observed that each one of compared classifiers increased the global accuracy and the recall. The best performance was obtained with Support Vector Machine and bayesNet.**

Keywords*—*Affective Computing; Detection of Emotional Information; Machine Learning; Speech Emotion Recognition.

I. Introduction

Traditionally, emotions in machines have been presented as dissociated from any type of rationality having virtually no role in their internal decision systems. However, recent discoveries in neurosciences, together with the extension of notions like emotional intelligence and multilevel intelligence, has led to the emergence of the new framework "Affective Computing" [1], according to which, the main aim is to build machines that recognize, express, model, communicate and respond to users emotion indicators. In the new framework, emotions hold a key role in machines which could impact positively their future decisions, bringing closer to taking part in a more sociable loop of human-machine interaction. As main field of application, the research shall implement the connection between robots and humans that will involve an emotional feedback framework, in which robots can understand emotions from some cues from human speech. The idea is to use robots which may understand emotions, and take part in the society cooperatively, according to the emotional state received from humans. Improving the communicative behavior of robots is urgent if people are to accept and integrate them in their world representation [2]. Robots have to be spontaneous, polite and must learn how to react according to the human being emotional charge, providing a friendly environment. Without the emotional feedback from humans, it will be very difficult for robots to interact with humans in a natural way [3], [4]. Within the context of human natural language, automatic emotional speech recognition by machines will expand the possibilities of interaction, since human speech provides a natural and intuitive interface for interaction with machines.

Emotions are visualized through various indicators in humans, many of these indicators have been previously analyzed to provide affective knowledge to machines, focusing on facial expressions [5], [6], vocal features [7], [8], [9], body movements and postures [10], [11], [12], [13] and the integration of all of them in emotion analysis systems [14], [15], [16]. But human beings cannot always hope that robots may be able to react in a timely and sensible manner, especially if they haven't be able to recover all the affective information through their sensors. Not always are the emotional features that the robot must capture provided by different sources from the human body at the same time. Maybe, all the information collected lacks robustness or, because the robot lacks the specific sensor to extract the emotional feature. Along the way to this goal, this research is based on the possible effects of some crucial speech features on the inference of emotions in communication with humans. It is known that emotions cause mental and physiological changes which are also reflected in uttered speech [17], [18], [19]. It is possible to find connections between emotional cues in speech and they can be utilized to learn about human emotions. Once such links are learned, theoretically, one can calculate the features and then automatically recognize the emotions present in human speech utterances, taking into account that the emotional content of speech does not depend on the speaker or the lexical content. Decrypting emotions in speech through several features has been a challenging research issue and one that has been of growing importance

in robotics, because of the emotional factors that the robot can handle and learn in social situations. In emotional classification from speech a multitude of different features have been used and a rule to follow is not yet established.

The fields of psychology and psycholinguistics provided interesting results about how prosodic cues, fundamental frequencies and the intensity of the voice can show variability levels across different speakers [20]. Short-term spectral features and sound quality can reveal emotional indicators [21], [22]. To delimit the scope of features selection, the research focus on the most useful group of them. Prosodic features, like pitch, loudness, speaking rate, durations, pause and rhythm show have strong correlations between them, providing valuate emotional information. In the case of the analysis of entire segment of voice, statistical functions like mean, median, minimum, maximum, standard deviation, or more seldom third or fourth standardized moment are applied to the fundamental frequency (F0) base contour [23], [24], [25]. The speech signal contains other frequency related characteristics that are spectral features. Mel Frequency Cepstral Coefficients (MFCCs) are generally used in speech recognition with great accuracy in emotion detection [26]. Predictive Cepstral Coefficients (LPCC) or Mel Filter Bank (MFB) features have a more common use [27]. The same performance displayed by MFCCs, is showed by RASTA-PLP (Relative Spectral Transform - Perceptual Linear Prediction) [28]. Through the analysis of voice quality [29], [30] and linguistic features, it can clearly be seen that there is a strong correlation between voice, pronounced words and emotions [31], [32], [33]. Different levels of voice could be depicted by neutral, whispery, breathy, creaky, and harsh or falsetto voice. In the case of features extracted from chains of words, the relation is depicted by the affective states associated with specific words; many of them are related to the probability of one emotion giving a certain sequence of words.

The machine learning framework shows several classifiers used in several tasks related to emotion recognition. Each classifier has advantages and disadvantages in order to deal with the speech emotion recognition problem. The more common group used are composed of Hidden Markov Model (HMM) [34], [35] regarded as the simplest dynamic Bayesian networks, Gaussian Mixture Models (GMM) [36], Nearest-Neighbour classifiers [37], artificial neural networks (ANN) [38], support vector machine (SVM) [39], k-NN [40], Decision Trees [41] and many others. The vast majority of emotion recognition systems over speech have employed a highdimensional speech grouped in a big vector of features, so the main goal will be to handle the dimensionality in order to improve the emotion recognition performance.

In this paper, the most commonly used features in several researches for capturing emotional speech characteristics in time and frequency were selected. The performance of different well known classifiers was compared in order to select the best result to predict the emotion, based on speech emotional data. To effectively reduce the size of speech features and improve the results obtained by the classifiers, the output from a decision tree classifier like feature selection method was used.

This paper is organized as follows. Section II describes the data set used in the research, the features extracted to represent the emotions from human speech, the machine learning techniques to perform the emotion classification experiments and the measures to evaluate the performance of classifiers. Section III describes the experimental results of all the several classification tests. Some conclusions are presented in Section IV.

II. Methods

1) Dataset: The emotional speech characteristics were extracted from the eNTERFACE05 audio-visual emotion database [42]. The data base is based on six universal emotions [43] like anger, disgust, fear, joy, sadness, and surprise. The voice data are provided by 44 non-native English speakers from 14 nations. The individuals expressed six basic emotions through five different sentences portrayed in 1320 videos, with a duration ranging from 1.2 to 6.7 seconds. For this research only one sentence per each emotion was used, which leads to a total of 264 videos. Each video is subsequently converted to a Waveform Audio File Format, for this task the MultimediaFileReader object from the DSP System Toolbox Library of MATLAB [44] was used to read the group of audio frames from each multimedia file. Fig. 1 shows the process applied to each video to build the emotional data set.

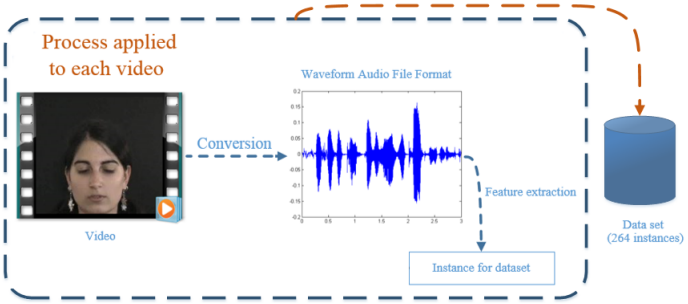


Fig. 1: Flowchart of the construction of dataset

2) Features: The data were acquired directly from the group of Waveform Audio files and they were transformed in 264 vectors of features. A wide range of possibilities exist for parametrically representing a speech signal and its content in a vector, with the intention of the extraction of relevant information from it. A variety of choices for this task can be applied to represent the speakers speech in a large number of parameters, in which the changes in these parameters will result in corresponding change in emotions. Taking into account that the system could be useful for a companion robot, the efforts should be focused on a system that contributing to a gradual gain of controllability and robustness that might save a substantial cost in computational efficiency. Not all the features that the robot can capture could be helpful and essential for its emotional feedback loop. Using all the features is not a guarantee to arrive the best performance, it could be better that the robot localizes the own best features and discard useless features from the

data base.

The kind of extracted features used in the research have been commonly used in music information retrieval (MIR), much of the research is based on the extraction mechanism from musical pieces, retrieval methodologies covered in various tasks related to different music representation media. It is attainable that the variability of emotions can be explained by a small set of acoustic features, for this task in order to identify objective acoustic features MATLAB was used, most of them developed in [45]. The spectral change of a signal is measured by the Spectral Flux (SF) feature [46]; the value is calculated through differences between each magnitude spectrum bin in the current frame to the corresponding value related to the magnitude spectrum of the previous frame. The result is the sum of the squares of the differences. Spectral Centroid (SC) [47] measures the center of mass of the power spectrum, it weighs the mean of the frequencies present in the speech signal. The SC uses the highest concentration point of energy in the spectrum and is correlated with the dominant frequency over the signal. Spectral Roll off Point [48] is often used as an indicator of the slant of the frequencies depicted in a frame. It is represented by a measure of the right-skewedness of the power spectrum. It increases with the bandwidth of a signal. Root Mean Square (RMS) [49] measures the power of a signal over a frame; the squares of each sample are summed and divided by the number of samples contained in frames. The value is square root of the total sum. Spectral Centroid Variability (SCV) [50] is the standard deviation of the magnitude spectrum, it measures the variability of the speech signal. Zero Crossing rate (ZCR) [50] provides an approximate estimation of dominant frequency and the spectral centroid and is described as the number of zero crossings during one second in the temporal domain. Compactness [51] is an indicator of the levels of noise in a signal; it is calculated by comparisons of components in a magnitude spectrum of a frame and the magnitude spectrum of its neighboring frames. Mel-Frequency Cepstral Coefficients (MFCCs) [52] is used to describe a spectrum frame, its first and second derivative in time are used to reflect dynamic changes. The first 8-13 MFCC coefficients are commonly used to describe the shape of the spectrum. They represent the information of the spectral envelope of the signal. Method of Moments [53] is composed for the first five statistical moments (area, mean, power spectrum density, spectral skew and spectral kurtosis) describing the shape of the spectrograph of a given frame. Linear predictive coding (LPC) is used to estimate the basic parameters into a speech signal, such as the vocal tract transfer function and the formant frequencies. It has good intuitive interpretation both in time domain and in frequency domain. The cepstral representation (Linear Prediction Cepstral Coefficients (LPCC)) of its coefficients is more used because of its higher retrieval efficiency [54]. 2D Method of Moments (2DMM) [55] gives a spectrograph description and the variation of it during a short time frame. The feature is composed by spectral data in frames analyzed with two-dimensional method of moments. Strongest Frequency Via Zero Crossings [48] is an estimation of the highest frequency of the component of a signal, found through the number of zero-crossings. 2D Method of Moments of MFCCs [47] is the 2D statistical computation of the Mel Frequency Cepstral Coefficients (MFCCs), this feature composed for a group of coefficients, allows recognizing the part of mid-frequencies from the signal. Fraction of Low Energy frame [56] is an indicator of the variability of the amplitude of frames; it is a fraction of previous frames, in which the Root Mean Square of each frame is less than the mean Root Mean Square. Strongest Frequency via FFT Maximum [48] is strongest frequency component in Hz of a signal. This is found by finding the highest bin (observations that fall into each of the disjoint categories) in the power spectrum. Strongest Frequency Via Spectral Centroid [48] is the strongest frequency in Hz in a signal related to the spectral centroid. The group of features conformed by Mel-Frequency Cepstral Coefficients, Linear Prediction Cepstral Coefficients, Method of Moments, 2D Method of Moments and 2D Method of Moments of MFCCs are matrices of, 4x13, 4x9, 4x5, 4x10, 4x10 respectively, that they will be transformed to vectors. Spectral Centroid, Spectral Roll off Point, Spectral Flux, Compactness, Spectral centroid Variability, Root Mean Square, Fraction of Low Energy frame, Zero Crossing rate, Strongest Frequency Via Zero Crossings, Strongest Frequency Via Spectral Centroid and Strongest Frequency Via FFT Maximum are conformed by 11 vectors of 8 features each one. Thus, the total feature vector contains 276 attributes that will be evaluated by the classifiers.

3) Machine learning techniques: The binary classification algorithm Support Vector Machine (SVM) which originated in statistical learning theory, offers robust classification to a very large number of variables and small samples [57]. SVM is capable of learning complex data from classification models applying mathematical principles to avoid overfitting. The more used kernels in SVM are polynomial and linear.

Another relatively fast classification model is the Decision tree, it works with a group of simple classification rules that are easy to understand. The rules represent the information in a tree based in a set of features. The classic decision tree is named ID3 based on growing and pruning [58], although C45 is other topdown decision trees inducers for continuous values [59], the last one is named as J48 in WEKA [66] and it uses the information gain as measure to select and split the nodes.

Within the connectionist techniques is also found the Artificial Neural Network (ANN). The ANN has a structure comparable to human neural networks where neurons located in layers process information. They have a graphical representation of an interconnected group of artificial neurons, in which the information resides in the weights from the arcs that connect the neurons. The ANN has two algorithms: feed-forward and recurrent neural network, in FF networks are supported over a directed acyclic graph, while RR networks have cycles. The most used feed-forward training algorithm is the Multilayer Perceptron named backpropagation [60]. The learning process covers two steps, the first step is a forward processing of input data by the neurons that produces a forecasted output, the second step is the adjustment of weights within the neuron layers, in order to minimize the errors of the forecasted solution compared with the correct output.

A graphical model (GMs) for probabilistic relationships

among a set of variables is bayesNet (Bayesian Network), it is used to represent knowledge the uncertainty [61]. The graph depicted in bayesNet is composed by nodes that represent random variables. In the graph, the edges between the nodes represent probabilistic dependencies among the corresponding random variables. Per each node there is a probability table specifying the conditional distribution of the variable given the values of its predecessors in the graph. These conditional dependencies in the graph are generally calculated by using known statistical and computational methods.

k Nearest Neighbors (kNN) is one of the simplest of classification algorithms available for supervised learning. The algorithm classifies unlabeled examples based on their similarity with examples in the training set. It is a lazy learning method that searches the closest match of the test data in feature space, based on distance function [62]. In this work is applied the Euclidean metric.

The supervised learning method naive Bayes [63] is a statistical method for classification, it is based on the well-known Bayes theorem with strong assumptions. The naive Bayes allows capturing the uncertainty about the model in a principled way by determining probabilities of the outputs. One of the advantages is the robustness to noise in input data. The classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature, given the class variable.

4) Validation techniques: Machine learning techniques have several measures in order to evaluate the performance of classifiers, which are principally focused on handling two-class problems. The performance of classifiers can be evaluated through several measures of machine learning techniques, which are principally focused in handling two-class problems. This research has faced a classification problem of six classes formed by six universal emotions. Most of the measures to evaluate binary problems could also apply to multi-class problem. In a problem with m classes, the performance of classifiers can be assessed based on an $m \times m$ confusion matrix, as shown in Table I. The groups of rows that describe the matrix represent the actual classes, while the columns are the predicted classes.

TABLE I: Confusion Matrix

	$Predicted\,Class_1$	$\cdots$	$Predicted\,Class_m$
$True \quad Class_1$	CM_{11}	$\cdots$	CM_{1m}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$True\,Class_m$	CM_{m1}	$\cdots$	CM_{mm}

For example, the accuracy is the percentage of correctly classified cases of the dataset. Based on the confusion matrix, the accuracy can be computed as a sum of the main values from the diagonal of the matrix, which represents the correctly classified cases divided by the total number of instances in the dataset (Eq. 1).

$$Accuracy = \frac{\sum_{i=1}^{m} CM_{ii}}{\sum_{i=1}^{m} \sum_{j=1}^{m} CM_{ij}} \tag{1}$$

where CM_{ij} represents the elements in the row i and column j of the confusion matrix.

Some measures like accuracy do not represent the reality of the number of cases correctly classified per each class. In order to make a deeper analysis, the measure of recall has been calculated for each class. Recall provides the percentage of correctness of classification into each class. Eq. 2 represents the recall for class [64].

$$Recall_i = \frac{CM_{ii}}{\sum_{j=1}^{m} CM_{ij}} \tag{2}$$

A k-fold cross-validation with $k = 10$ was used to make validations over the classifiers. This technique allowed the evaluation of the model facing an unknown dataset. The group of data is randomly divided in k equal parts, one part of the group is used as a validation set and the rest $k-1$ will be the training set. The process is repeated k times using a different group as a validation set, this process continues until each group can used once as validation test. Then, the k results obtained by groups can be averaged to a single result. The advantage of 10-fold cross-validation is that all examples of the database are used for both, training and testing stages [65].

III. Results

The intent of this study was to provide the best classification of emotions contained in a speech signal, which might serve to feed the decision support system of a synthetic agent capable of supporting the societal participation of persons deprived of conventional modes of communication, in the context of socially intelligent systems. A 10-fold crossvalidation scheme was employed in the speech dataset for all the emotion classification experiments; this was done to validate the performances of the classifiers selected. Six classifiers were tested, the Support Vector Machine (SVM) has used three kernels, linear and polynomial (with degrees 2 and 3), k Nearest Neighbors (kNN) has used k from 1 to 15 (showed the best result with k=5), Multilayer Perceptron (MLP) with hidden neurons from 2 to 20 (showed the best result with 10 neurons), bayesNet (BN), NaiveBayes (NB) and decision tree (J48). All the several classification tests were conducted using the WEKA [66] toolbox. The best performance was achieved with the decision tree (J48) reaching a 96.21 % of accuracy facing the other classifiers, as shown in Fig. 2. The accuracy and the recall results were compared. As you can see in Fig. 3, the percentage of most relevant results per emotion positively classified (recall) was raised by the decision tree (J48).

The decision tree has reached the best result in accuracy and recall facing the remaining classifiers; therefore, this result is obtained with only a few features selected, taking into account their information gain. As can be seen the decision tree is composed of six nodes, which correspond to six features of the dataset, which means the tree only needs these six features to predict the emotions. The features selected of the tree are

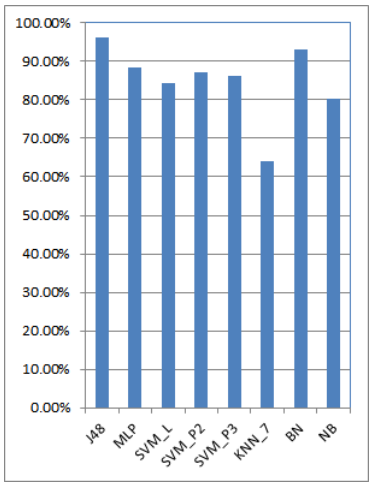


Fig. 2: Comparison of Accuracy of different classifiers

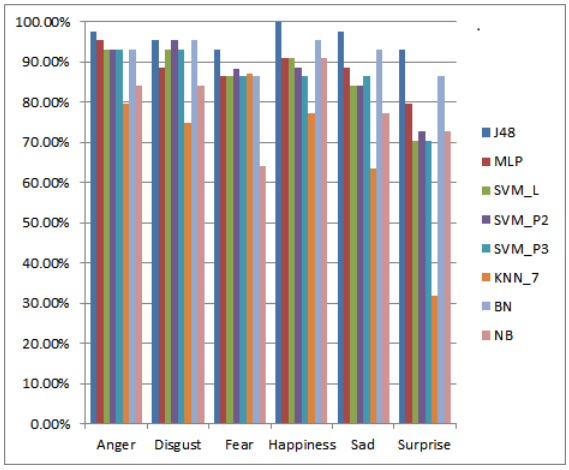


Fig. 3: Comparison of Recall of different classifiers

based on the 2D Method of Moments $(2DMM)$ and the 2D Method of Moments of MFCCs $(2DMM - MFCCs)$. Fig. 4 illustrates the graphical rendition obtained of the classification tree in which $2DMM_n$ is the n-th element within $n = 1, 2, 3...10$ from the feature vector of 2D Method of Moments. Similarly, $2DMM - MFCCs_m$ is the m-th element within $m = 1, 2, 3...10$ from the feature vector of 2D Method of Moments of MFCCs.

The confusion matrix of J48, as depicted in Table II, shows a balanced distribution of misclassifications rates in the group of emotions. For all six emotions, "happiness" is not confused at all with the rest of emotions and it is recognized with 100 %. Further analysis of the confusion matrix shows that the emotions "fear" and "surprise" attained a higher number of misclassifications and lower percentages of recall (both of them 93.20 %), as shown in Table II and Table III respectively. In case of emotions "disgust" and "fear", this speech signals could be interpreted from the psychophysiological framework, some acoustic cues in discrete emotions could lead listeners

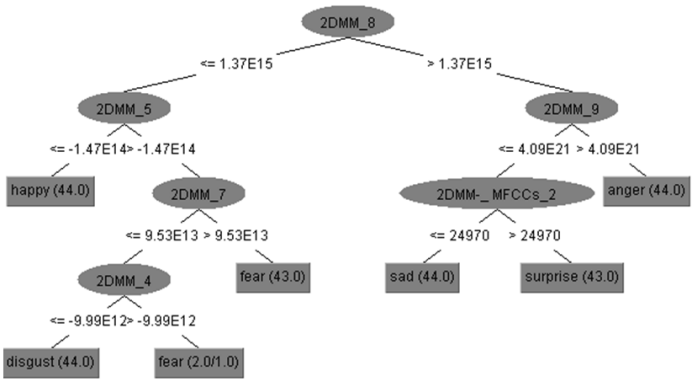


Fig. 4: Decision tree

to misclassification with poorer decoding accuracy [67].

TABLE II: Confusion Matrix J48

	Anger	Disgust	Fear	Happiness	Sad	Surprise
Anger	43	0	0	0	0	1
Disgust	0	42	2	0	0	0
Fear	0	2	41	1	0	0
Happiness	0	0	0	44	0	0
Sad	0	0	0	0	43	1
Surprise	1	1	0	0	1	41

Comparing the results of the decision tree with the remaining classifiers, it seems likely that the learning mechanism in the tree is essential in this problem. An important process in the algorithm, is how to determine which attribute to split on. The attributes are selected based on information gain, resulting in a set of selected relevant features. This can only lead to conclude that dataset probably has noisy and redundant features. Then, the information gain is visualized as a heuristic to select features as is done for the decision tree. Taking into account the features selected by the tree, the data set were reconstructed with the selection of the 2D Method of Moments and the 2D Method of Moments of MFCCs. The same classifiers with the same parameters were trained and compared. Also it is shown the best result for each classifier, where the best result for MLP was with 12 neurons and kNN for k=6. Comparisons between the accuracy previously obtained and the results are showed in Fig. 5. As can be seen, the features selected by the decision tree have highlighted improvements in performance of all classifiers in a range of 3.5 % to 31.8 %. The accuracies of MLP, BN and SVM (with polynomial kernel of degree 2) were superior to the decision tree. The MLP and BN have achieved 96.97 % and the SVM 96.59 %.

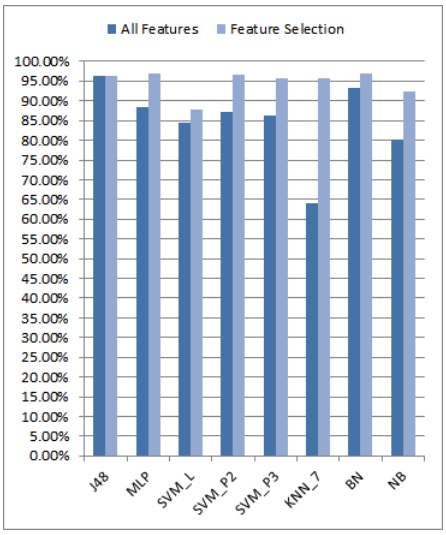


Fig. 5: Comparison of Accuracy of different classifiers with all features and with feature selection

In order to analyze the results in each emotion, the recall can be analyzed in Fig. 6. It is clear that all emotions show improvements over the results obtained before. The best results also are obtained by MLP, BN and polynomial SVM with degree 2. In order to see more detail, Table III shows the recall of the four classifiers (J48, MLP, SVM with degree 2 (SVM-P2) and BN). The three last rows illustrate the average of recall for each classifier (Average) and the range of the recall (Min and Max). The differences between the classifiers are not significant in accuracy. However, the three classifiers are superior to J48 based on the range of the recall and the average. BN and MLP show similitudes in average, while MLP has a better range. The range of SVM is equal to MLP, while the average is lower than MLP. This comparison means that MLP has achieved the best results.

TABLE III: Comparison of recall for J48, MLP, SVM with degree 2 (SVM-P2) and BN

	J48	MLP	SVM-P2	BN
Anger	97.70 %	100.00 %	100.00 %	97.70 %
Disgust	95.50 %	95.50 %	95.50 %	97.70 %
Fear	93.20 %	95.50 %	95.50 %	95.50 %
Happiness	100.00 %	97.70 %	97.70 %	97.70 %
Sad	97.70 %	97.70 %	95.50 %	97.70 %
Surprise	93.20 %	95.50 %	95.50 %	95.50 %
Average	96.22 %	96.98 %	96.62 %	96.97 %
Min	93.20 %	95.50 %	95.50 %	95.50 %
Max	100.00 %	100.00 %	100.00 %	97.70 %

Keeping in mind that the building of a system for real time probably is applicable to a robot, the time consumption of the algorithm is relevant. Taking into account that the classifiers (SVM, BN and MLP) have a little difference in the results, it can be suggested the use of BN or SVM instead of MLP to consume less computational resources.

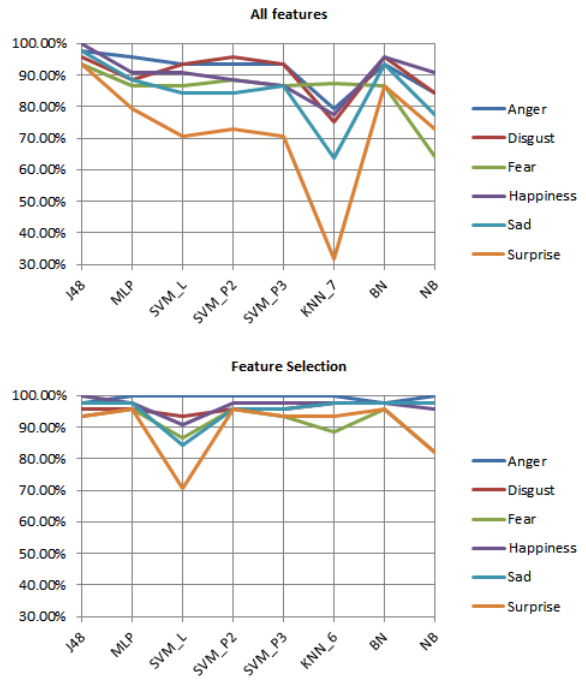


Fig. 6: Comparison of Recall of different classifiers with all features and with feature selection for each emotion

The emotions “anger” and “happiness” have achieved the best performance from the beginning, while surprise is the lowest results and shows the same behavior for all classifiers.

IV. Conclusions

The purpose of this research was to perform parameterization of audio data for the purpose of automatic recognition of emotions in speech. A collection of audio data from several videos related to human emotional expressions were gathered and turned into a data set. A group of six classifiers in order to identify the best of them to predict emotions in humans were selected. The outputs from a decision tree have been used as a feature selection technique to remove redundant and noisy features. The features provided by the decision tree were 2D Method of Moments and 2D Method of Moments of MFCCs. The feature selection increases the efficiency of the accuracy and the recall. The feature selection also allows reduction of the dimensionality of the data in turn leading to less computation processes in the robot memory.

After the selection of features, a group of experiments in order to select the best classifiers were conducted. Multilayer Perceptron, Support Vector Machine and bayesNet have achieved the best results. Support Vector Machine and bayesNet could be good candidates to build the emotional recognition system of a robot, because of their easily implementation and the less computational complexity.

This simple system with the classifiers is easy to understand and implement because of the utilization from a small group

of features would work remarkably well on real-world data, making it possible to develop a real-time system in which the robot can make a fast decision in accordance with the emotional feedback provided from humans. As a real application, it could be considered a real-time system that can serve like a motor of emotional knowledge in order to understand the autistic children, to describe accurately their internal state and show the real content of their emotions. The system is not only applied to companion robots it could also be applicable to diverse smart sources (smart devices), this could be the case of healthcare, telemedicine or smart well-being systems that can be seen more often. This type of emotional devices working with emotional feedback will have the potential to reveal more about emotional state and the early detection of crisis, balanced lifestyle including and regulated stress level.

ACKNOWLEDGMENT

The authors greatly appreciate the financial support provided by the institution VINNOVA Swedish Governmental Agency for Innovation Systems through the ICT project The Next Generation (TNG). We also grateful to Antioquia School of Engineering "EIA" (Colombia) and Pontifical Catholic University (Perú) in a joint effort for collaborative research.

REFERENCES

[1] R. Picard, Affective Computing. The MIT Press, United States, 1998.

[2] T. Ziemke and R. Lowe, "On the role of emotion in embodied cognitive architectures: From organisms to robots," Cognitive computation, vol. 1, no. 1, pp. 104–117, 2009.

[3] H.A. Samani and E. Saadatian, "A Multidisciplinary Artificial Intelligence Model of an Affective Robot," international Journal of Advanced Robotic Systems, vol. 9, pp. 1–11, 2012.

[4] J.G. Rázuri, P.G. Esteban and D.R. Insua, "An adversarial risk analysis model for an autonomous imperfect decision agent," In T.V. Guy, M. Kárný and D.H. Wolpert, Eds. Decision Making and Imperfection. SCI, vol. 474, pp. 165–190. Springer, Heidelberg, 2013.

[5] M. Pantic and L.J.M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1424–1445, 2000.

[6] D. Filko and G. Martinovic, "Emotion recognition system by a neural network based facial expression analysis, AutomatikaJournal for Control," Measurement, Electronics, Computing and Communications, vol. 54, no. 2, 2013.

[7] R. Cowie and E. Douglas-Cowie, "Automatic statistical analysis of the signal and prosodic signs of emotion in speech," In Proc. International Conf. on Spoken Language Processing, pp. 1989–1992, 1996.

[8] T. Sobol-Shikler, P. Robinson, "Classification of complex information: Inference of co-occurring affective states from their expressions in speech," IEEE Trans. Pattern Anal. Mach. Intell, vol. 32, no. 7, pp. 1284–1297, 2010.

[9] K. Han, D. Yu and I. Tashev, "Speech Emotion Recognition Using Deep Neural Network and Extreme Learning Machine," Interspeech 2014, pp. 223–227, 2014.

[10] N. Bianchi-Berthouze and A. Kleinsmith, "A categorical approach to affective gesture recognition," Connection Science, vol. 15, no. 4, pp. 259–269, 2003.

[11] G. Castellano, S.D. Villalba and A. Camurri, "Recognising Human Emotions from Body Movement and Gesture Dynamics," In Proc. of 2nd International Conference on Affective Computing and Intelligent Interaction, Berlin, Heidelberg, 2007.

[12] K. Schindler, L. van Gool, and B. de Gelder, "Recognizing emotions expressed by body pose: a biologically inspired neural model," Neural Networks, vol. 21, no. 9, pp. 1238–1246, 2008.

[13] A. Kleinsmith and N. Bianchi-Berthouze, "Recognizing affective dimensions from body posture," In Affective Computing and Intelligent, Lecture Notes in Computer Science, pp. 48–58, Springer, Berlin, Germany, 2007.

[14] H. K. M. Meeren, C. van Heijnsbergen and B. de Gelder, "Rapid perceptual integration of facial expression and emotional body language," Proc. National Academy of Sciences of the USA, vol. 102, no. 45, pp. 16518–16523, 2005.

[15] A. Metallinou, A. Katsamanis and S. Narayanan, "Tracking changes in continuous emotion states using body language and prosodic cues," In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2288–2291, 2011.

[16] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C.M. Lee, A. Kazemzaeh. S. Lee, U. Neumann and S. Narayanan, "Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal information," In Proc. of ACM 6th int'l Conf. on Multimodal Interfaces (ICMI2004), State College, PA, pp. 205–211, 2004.

[17] J.Q. Wang, N. Trent, E. Skoe, M. Sams and N. Kraus, "Emotion and the auditory brainstem response to speech," Neuroscience Letters, vol. 469, no. 3, pp. 319–323, 2010.

[18] D.A. Abrams, N. Trent, S. Zecker and N. Kraus, "Rapid acoustic processing in the auditory brainstem is not related to cortical asymmetry for the syllable rate of speech," Clinical Neurophysiology, vol. 121, no. 8, pp. 1343–1350, 2010.

[19] M. Drolet, R.I Schubotz and J. Fischer, "Authenticity affects the recognition of emotions in speech: behavioral and fMRI evidence," Cognitive, Affective, and Behavioral Neuroscience, vol. 12, no. 1, pp. 140–150, 2012.

[20] J. Ang, R. Dhillon, A. Krupski, E. Shriberg and A. Stolcke, "Prosody-based automatic detection of annoyance and frustration in human-computer dialog," Proc. International Conference on Spoken Language Processing (ICSLP 2002), pp. 2037–2040, 2002.

[21] V. Hozjan and Z. Kacic, "Context-independent multilingual emotion recognition from speech signals," International journal of Speech Technology, vol. 6, pp. 311–320, 2003

[22] B. Schuller, D. Seppi, A. Batliner, A. Maier, and S. Steidl, "Towards more reality in the recognition of emotional speech," In International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 941–944, Honolulu, HI, USA, 2007.

[23] R. Cowie,E. Douglas-Cowie, N. Tsapatsoulis, S. Kollias, W. Fellenz and J. Taylor, "Emotion recognition in humancomputer interaction," IEEE Signal Process, vol. 18, pp.32–80, 2001.

[24] I. Murray, J. Arnott, "Toward a simulation of emotions in synthetic speech: A review of the literature on human vocal emotion," J. Acoust. Soc. Am, vol. 93, no. 2, pp. 1097–1108, 1993.

[25] D. Ververidis and C. Kotropoulos, "Emotional speech recognition: Resources, features, and methods," Speech Communication, pp. 1162–1181, 2006.

[26] D. Neiberg, K. Elenius, I. Karlsson and K. Laskowski, "Emotion Recognition in Spontaneous Speech," pp. 101–104, 2006.

[27] C. Busso, S. Lee, and S.S. Narayanan, "Using neutral speech models for emotional speech analysis," In Interspeech 2007-Eurospeech, pp. 2225–2228, 2007.

[28] K.P. Truong and D.A. van Leeuwen, "Automatic discrimination between laughter and speech," Speech Commun, vol. 49, no. 2. pp. 144–158, 2007.

[29] P. Alku, "Glottal inverse filtering analysis of human voice production A review of estimation and parameterization methods of the glottal excitation and their applications," Sadhana, vol. 36, no. 5, pp. 623–650, 2011.

[30] M. Lugger and B. Yang, "The relevance of voice quality features in speaker independent emotion recognition," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007), Honolulu, HI, USA, vol. 4, pp. 17–20, 2007.

[31] C.M. Lee and S.S. Narayanan, S. S, "Toward Detecting Emotions in Spoken Dialogs," IEEE Transactions on Speech and Audio Processing, vol. 13, no. 2, pp. 293–303, 2005.

[32] S. Steidl, Automatic Classification of Emotion-Related User States in Spontaneous Children's Speech, Logos-Verlag, 2009.

[33] B. Schuller, A. Batliner, S. Steidl and D. Seppi, "Emotion Recognition

from Speech: Putting ASR in the Loop," Proc. ICASSP 2009, IEEE, Taipei, Taiwan, pp. 4585–4588, 2009.

[34] D. Le and E. M. Provost, "Emotion recognition from spontaneous speech using Hidden Markov models with deep belief networks," in Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on, pp. 216–221, 2013.

[35] J. Wagner, T. Vogt, and E. André, "A systematic comparison of different HMM designs for emotion recognition from acted and spontaneous speech," in Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction (ACII), Lisbon, Portugal, pp. 114–125, 2007.

[36] T. Hao, S.M. Chu, M. Hasegawa-Johnson and T.S. Huang, "Emotion recognition from speech VIA boosted Gaussian mixture models," in Multimedia and Expo, 2009. ICME 2009. IEEE International Conference, pp. 294–297, 2009.

[37] S.A. Rieger, R. Muraleedharan and R.P. Ramachandran, "Speech based emotion recognition using spectral feature extraction and an ensemble of kNN classifiers," in Chinese Spoken Language Processing (ISCSLP), pp. 589–593, 2014.

[38] S.A. Firoz, S.A. Raj and A.P. Babu, "Automatic Emotion Recognition from Speech Using Artificial Neural Networks with Gender-Dependent Databases," in Advances in Computing, Control and Telecommunication Technologies, ACT '09, pp. 162–164, 2009.

[39] C. Yu, Q. Tian, F. Cheng and S. Zhang, "Speech Emotion Recognition Using Support Vector Machines," in Advanced Research on Computer Science and Information Engineering. vol. 152, G. Shen and X. Huang, Eds., ed: Springer Berlin Heidelberg, pp. 215–220, 2011.

[40] M. Feraru and M. Zbancioc, "Speech emotion recognition for SROL database using weighted KNN algorithm," in Electronics, Computers and Artificial Intelligence (ECAI) , pp. 1–4, 2013.

[41] C.-C. Lee, E. Mower, C. Busso, S. Lee and S. Narayanan, "Emotion recognition using a hierarchical binary decision tree approach," Speech Commun, vol. 53, pp. 1162–1171, 2011.

[42] O. Martin, I. Kotsia, B. Macq and I. Pitas, "The eNTERFACE' 05 Audio-Visual Emotion Database," in Data Engineering Workshops, Proceedings. 22nd International Conference, pp. 8–8, 2006.

[43] P. Ekman and W.V. Friesen, "A new pan-cultural facial expression of emotion," Motivation and Emotion, vol. 10, no. 2, pp. 159–168, 1986.

[44] MathWorks, (2014). DSP System Toolbox: User's Guide (R2014b). http://fr.mathworks.com/help/pdf_doc/dsp/dsp_ug.pdf, 2014.

[45] T. Giannakopoulos and A. Pikrakis. Introduction to Audio Analysis. Elsevier Academic Press, 2014.

[46] P. Masri, "Computer modelling of sound for transformation and synthesis of musical signal," Ph.D. dissertation, University of Bristol, UK, 1996.

[47] G. Peeters, "Large Set of Audio Features for Sound Description," Technical report published by IRCAM, 2004.

[48] C. McKay and I. Fujinaga, "Automatic music classification and similarity analysis," International Conference on Music Information Retrieval, 2005.

[49] K.V. Cartwright, "Determining the Effective or RMS Voltage of Various Waveforms without Calculus," Technology Interface, vol. 8, no. 1, pps. 20, 2007.

[50] J. Kim, E. Andre, M. Rehm, T. Vogt and J. Wagner, "Integrating information from speech and physiological signals to achieve emotional sensitivity," In Proc. Interspeech, Lisbon, Portugal, pp. 809812, 2005.

[51] F. Pachetand P. Roy, P, "Analytical features: A knowledge-based approach to audio feature generation," EURASIP Journal on Audio, Speech, and Music Processing, 2009.

[52] B. Bogert, M. Healy, and J. Tukey, "The quefrency alanysis of time series for echoes: cepstrum, pseudo-autocovariance, cross- cepstrum, and saphe-cracking," Proceedings of the Symposium on Time Series Analysis, Wiley, 1963.

[53] I. Fujinaga, "Adaptive Optical Music Recognition," Ph.D. thesis, Department of Theory, Faculty of Music, McGill University, Montreal, Canada, 1997.

[54] X. Changsheng, M. C. Maddage and S. Xi, "Automatic music classification and summarization," Speech and Audio Processing, IEEE Transactions on, vol. 13, pp. 441-450, 2005.

[55] R. Mittra and V. Varadarajan, "A technique for solving 2D methodof-moments problems involving large scatterers," Microwave and Optical Technology Letters, vol. 8 no. 3, 2007.

[56] C. McKay and I. Fujinaga, jMIR: Tools for automatic music classification. Ann Arbor, MI: MPublishing, University of Michigan Library, 2009

[57] V. Vapnik, The Nature of Statistical Learning Theory ed.; Springer-Verlag, New York, 1995.

[58] J. R. Quinlan, "Induction of decision trees," Mach. Learn, vol. 1, no. 1, pp. 81-106, 1986.

[59] J. R. Quinlan, C4.5: Programs for Machine Learning, 1st ed.; Morgan Kaufmann Publishers: San Francisco, CA, USA, 1993.

[60] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Parallel distributed processing: explorations in the microstructure of cognition," D. E. Rumelhart and J.L. McClelland, eds, MIT Press: Cambridge, MA, USA, vol. 1, pp. 318-362, 1986

[61] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Mateo, CA, 1988.

[62] T. M. Mitchell, Machine Learning; McGraw-Hill: New York, NY, p. 432, 1997.

[63] H. Zhang, "The Optimality of Naive Bayes," Proc. the 17th International FLAIRS conference, Florida, USA, pp. 17-19, 2004

[64] P. Flach, "Machine Learning: The Art and Science of Algorithms that Make Sense of Data," Cambridge University Press, 2012.

[65] B. Efron and R. J. Tibshirani, "An introduction to the Bootstrap," Chapman and Hall: New York, USA, 1993.

[66] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," ACM SIGKDD Explorations Newsletter, vol. 11, no. 1, pp. 10-18, 2009.

[67] P. Laukkaand P. Juslin, "Similar patterns of age-related differences in emotion recognition from speech and music, " Motivation and Emotion, vol. 31, no. 3, pp. 182-191. 2007.

A Trust-based Mechanism for Avoiding Liars in Referring of Reputation in Multiagent System

Manh Hung Nguyen
Posts and Telecommunications Institute of Technology (PTIT)
Hanoi, Vietnam
UMI UMMISCO 209 (IRD/UPMC), Hanoi, Vietnam

Dinh Que Tran
Posts and Telecommunications Institute of Technology (PTIT)
Hanoi, Vietnam

***Abstract*—Trust is considered as the crucial factor for agents in decision making to choose the most trustworthy partner during their interaction in open distributed multiagent systems. Most current trust models are the combination of experience trust and reference trust, in which the reference trust is estimated from the judgements of agents in the community about a given partner. These models are based on the assumption that all agents are reliable when they share their judgements about a given partner to the others. However, these models are no more longer appropriate to applications of multiagent systems, where several concurrent agents may not be ready to share their private judgement about others or may share the wrong data by lying to their partners.**

In this paper, we introduce a combination model of experience trust and experience trust with a mechanism to enable agents take into account the trustworthiness of referees when they refer their judgement about a given partner. We conduct experiments to evaluate the proposed model in the context of the e-commerce environment. Our research results suggest that it is better to take into account the trustworthiness of referees when they share their judgement about partners. The experimental results also indicate that although there are liars in the multiagent systems, combination trust computation is better than the trust computation based only on the experience trust of agents.

***Keywords*—Multiagent system, Trust, Reputation, Liar.**

I. Introduction

Many software applications are open distributed systems whose components are decentralized, constantly changed, and spread throughout network. For example, peer-to-peer networks, semantic web, social network, recommender systems in e-business, autonomic and pervasive computing are among such systems. These systems may be modeled as open distributed multiagents in which autonomous agents often interact with each other according to some communication mechanisms and protocols. The problem of how agents decide with whom and when to interact has become the active research topic in the recent years. It means that they need to deal with degrees of uncertainty in making decisions during their interaction. Trust among agents is considered as one of the most important foundations based on which agents decide to interact with each other. Thus, the problem of how do agents decide to interact may reduce to the one of how do agents estimate their trust on their partners. The more trust an agent commits on a partner, the more possibility with such partner he decides to interact.

Trust has been defined in many different ways by researchers from various points of view [7], [15]. It has been being an active research topic in various areas of computer science, such as security and access control in computer networks, reliability in distributed systems, game theory and multiagent systems, and policies for decision making under uncertainty. From the computational point of view, trust is defined as a quantified belief by a truster with respect to the competence, honesty, security and dependability of a trustee within a specified context [8].

These current models utilize the combination of experience trust (confidence) and reference trust (reputation) in some way. However, most of them are based on the assumption that all agents are reliable when they share their private trust about a given partner to others. This constraint limits the application scale of these models in multiagent systems including concurrent agents, in which many agents may not be ready to share with each other about their private trust about partners or even share the wrong data by lying to their opponents.

Considering a scenario of the following e-commerce application. There are two concurrent sellers S_1 and S_2 who sell the same product x. An independent third party site w is to collect the consumer's opinions. All clients could submit their opinions about sellers. In this case, the site w could be considered as a reputation channel for clients. It means that a client could refer the given opinions on the site w to select the best seller. However, since the site w is a public reputation and all clients could submit their opinions. Imagining that S_1 is really trustworthy, but S_2 is not fair, some of its employments intentionally submit some negative opinions about the seller S_1 in order to attract more clients to them. In this case, how a client could trust on the reputation given by the site w? These proposed models of trust may not be applicable to such a situation.

In order to get over this limitation, our work proposes a novel computational model of trust that is a weighted combination of experience trust and reference trust. This model offers a mechanism to enable agents take into account the trustworthiness of referees when they refer the the judgement about a given partner from these referees. The model is evaluated experimentally on two issues in the context of the e-commerce environment: (i) It is whether necessary to take into account the trust of referees (in sharing their private trust about partners) or not; (ii) Combination of experience trust

and reputation is more useful than the trust based only on the experience trust of agents in multiagent systems with liars.

The rest of paper is organized as follows. Section II presents some related works in literature. Section III describes the model of weighted combination trust of experience trust, reference trust with and without lying referees. Section IV describes the experimental evaluation of the model. Section V is offered to some discussion. Section VI is the conclusion and the future works.

II. Related Works

By basing on the contribution factors of each model, we try to divide the proposed models into three groups. Firstly, The models are based on *personal experiences* that a truster has on some trustee after their transactions performed in the past. For instance, Manchala [19] and Nefti et al. [20] proposed models for the trust measure in e-commerce based on fuzzy computation with parameters such as cost of a transaction, transaction history, customer loyalty, indemnity and spending patterns. The probability theory-based model of Schillo et al. [28] is intended for scenarios where the result of an interaction between two agents is a boolean impression such as good or bad but without degrees of satisfaction. Shibata et al. [30] used a mechanism for determining the confidence level based on agent's experience with Sugarscape model, which is artificially intelligent agent-based social simulation. Alam et al. [1] calculated trust based on the relationship of stake holders with objects in security management. Li and Gui [18] proposed a reputation model based on human cognitive psychology and the concept of direct trust tree (DTT).

Secondly, the models combine both personal experience and reference trusts. In the trust model proposed by Esfandiari and Chandrasekharan [4], two one-on-one trust acquisition mechanisms are proposed. In Sen and Sajja's [29] reputation model, both types of direct experiences are considered: direct interaction and observed interaction. The main idea behind the reputation model presented by Carter et al. [3] is that "the reputation of an agent is based on the degree of fulfillment of roles ascribed to it by the society". Sabater and Sierra [26], [27] introduced ReGreT, a modular trust and reputation system oriented to complex small/mid-size e-commerce environments where social relations among individuals play an important role. In the model proposed by Singh and colleagues [36], [37] the information stored by an agent about direct interactions is a set of values that reflect the quality of these interactions. Ramchurn et al. [24] developed a trust model, based on confidence and reputation, and show how it can be concretely applied, using fuzzy sets, to guide agents in evaluating past interactions and in establishing new contracts with one another. Jennings et collegues [12], [13], [25] presented FIRE, a trust and reputation model that integrates a number of information sources to produce a comprehensive assessment of an agent's likely performance in open systems. Nguyen and Tran [22], [23] introduced a computational model of trust, which is also combination of experience and reference trust by using fuzzy computational techniques and weighted aggregation operators. Victor et al. [33] advocate the use of a trust model in which trust scores are (trust, distrust)-couples, drawn from a bilattice that preserves valuable trust provenance information including gradual trust, distrust, ignorance, and inconsistency. Katz and Golbeck [16] introduces a definition of trust suitable for use in Web-based social networks with a discussion of the properties that will influence its use in computation. Hang et al. [10] describes a new algebraic approach, shows some theoretical properties of it, and empirically evaluates it on two social network datasets. Guha et al. [9] develop a framework of trust propagation schemes, each of which may be appropriate in certain circumstances, and evaluate the schemes on a large trust network. Vogiatzis et al. [34] propose a probabilistic framework that models agent interactions as a Hidden Markov Model. Burnett et al. [2] describes a new approach, inspired by theories of human organisational behaviour, whereby agents generalise their experiences with known partners as stereotypes and apply these when evaluating new and unknown partners. Hermoso et al. [11] present a coordination artifact which can be used by agents in an open multi-agent system to take more informed decisions regarding partner selection, and thus to improve their individual utilities.

Thirdly, the models also compute trust by means of combination of the experience and reputation, but consider unfair agents in sharing their trust in the system as well. For instances, Whitby et al. [35] described a statistical filtering technique for excluding unfair ratings based on the idea that unfair ratings have some statistical pattern being different from fair ratings. Teacy et al. [31], [32] developed TRAVOS (Trust and Reputation model for Agent-based Virtual OrganisationS) which models an agent's trust in an interaction partner, using probability theory taking account of past interactions between agents, and the reputation information gathered from third parties. And HABIT, a Hierarchical And Bayesian Inferred Trust model for assessing how much an agent should trust its peers based on direct and third party information. Zhang, Robin and collegues [39], [14], [5], [6] proposed an approach for handling unfair ratings in an enhanced centralized reputation system.

The models in the third group are closed to our model. However, most of them used Bayes network and statistical method to detect the unfairs in the system. This approach may result in difficulty when the number of unfair agents become major.

This paper is a continuation of our previous work [21] in order to update our approach and perform experimental evaluation of this model.

III. Computational Model of Trust

Let $A = \{1, 2, ...n\}$ be a set of agents in the system. Assume that agent i is considering the trust about agent j. We call j is a *partner* of agent i. This consideration includes: (i) the direct trust betwwen agent i and agent j, called *experiment trust* E_{ij}; and (ii) the trust about j refered from community called *reference trust (or reputation)* R_{ij}. Each agent l in the community that agent i refers for the trust of partner j is called a *referee*. This model enables agent i to take into account the trustworthiness of referee l when agent l shares its private trust (judgement) about agent j. The trustworthiness of agent l on the point of view of agent i, in sharing its private trust about partners, is called a *referee trust* S_{il}. We also denote T_{ij} to be the overall trust that agent i obtains on agent j. The following sections will describe a computational model to estimate the values of E_{ij}, S_{il}, R_{ij} and T_{ij}.

TABLE I: Summary of recent proposed models regarding the fact of avoiding liar in calulation of reputation

Models	Experience Trust	Reputation	Liar Judger
Alam et al. [1]	✓		
Burnett et al. [2]	✓	✓	
Esfandiari and Chandrasekharan [4]	✓	✓	
Guha et al. [9]	✓	✓	
Hang et al. [10]	✓	✓	
Hermoso et al. [11]	✓	✓	
Jennings et al. [12], [13]	✓	✓	
Katz and Golbeck [16]	✓	✓	
Lashkari et al.[17]	✓	✓	
Li and Gui [18]		✓	
Manchala [19]	✓		
Nefti et al. [20]	✓		
Nguyen and Tran [22], [23]	✓	✓	
Ramchurn et al. [24]	✓	✓	
Sabater and Sierra [26], [27]	✓	✓	
Schillo et al. [28]	✓		
Sen and Sajja's [29]	✓	✓	
Shibata et al. [30]	✓		
Singh and colleagues [36], [37]	✓	✓	
Teacy et al. [31], [32]	✓	✓	✓
Victor et al. [33]	✓	✓	
Vogiatzis et al. [34]	✓	✓	
Whitby et al. [35]	✓	✓	✓
Zhang, Robin and collegues [39], [14], [5], [6]	✓	✓	✓
Our model	✓	✓	✓

A. Experience trust

Intuitively, experience trust of agent i in agent j is the trustworthiness of j that agent i collects from all transactions between i and j in the past.

Experience trust of agent i in agent j is defined by the formula:

$$E_{ij} = \sum_{k=1}^{n} t_{ij}^{k} * w_k \tag{1}$$

where:

- t_{ij}^{k} is the transaction trust of agent i in its partner j at the k^{th} latest transaction.
- w_k is the weight of the k^{th} latest transaction such that

$$\begin{cases} w_{k_1} \geqslant w_{k_2} \text{ if } k_1 < k_2 \\ \sum_{k=1}^{n} w_k = 1 \end{cases}$$

- n is the number of transactions taken between agent i and agent j in the past.

The weight vector $\overrightarrow{w} = \{w_1, w_2, ..w_n\}$ is decreasing from head to tail because the aggregation focuses more on the later transactions and less on the older transactions. It means that the later the transaction is, the more its trust is important to estimate the experience trust of the correspondent partner. This vector may be computed by means of Regular Decreasing Monotone (RDM) linguistic quantifier Q (Zadeh [38]).

B. Trust of referees

Suppose that an agent can refer all agents he knows (referee agents) in the system about their experience trust (private judgement) on a given partner. This is called *reference trust* (this will be defined in the next section). However, some referee agents may be liar. In order to avoid the case of lying referee, this model proposes a mechanism which enables an agent to evaluate its referees on sharing their private trust about partners.

Let $X_{il} \subseteq A$ be a set of partners that agent i refers their trust via referee l, and that agent i has already at least one transaction with each of them. Since the model supposes that agent always trusts in itself, the trust of referee l from the point of view of agent i is determined based on the difference between experience trust E_{ij} and the trust r_{ij}^{l} of agent i about partner j referred via referee l (for all $j \in X_{il}$).

Trust of referee (sharing trust) S_{il} of agent i on the referee l is defined by the formula:

$$S_{il} = \frac{1}{\mid X_{il} \mid} * \sum_{j \in X_{il}} h(E_{ij}, r_{ij}^{l}) \tag{2}$$

where:

- h is a *referee-trust-function* $h : [0,1] \times [0,1] \to [0,1]$, which satisfies the following conditions:

  $$h(e_1, r_1) \leqslant h(e_2, r_2) \text{ if } \mid e_1 - r_1 \mid \geqslant \mid e_2 - r_2 \mid .$$

  These constraints are based on the following intuitions:
  - The more the difference between E_{ij} and r_{ij}^{l} is large, the less agent i trust on the referee l, and conversely;
  - The more the difference between E_{ij} and r_{ij}^{l} is small, the more agent i trusts on the referee l.
- E_{ij} is the experience trust of i on j
- r_{ij}^{l} is the reference trust of agent i on partner j that is referred via referee l:

$$r_{ij}^{l} = E_{lj} \tag{3}$$

C. Reference trust

Reference trust (also called reputation trust) of agent i on partner j is the trustworthiness of agent j given by other referees in the system. In order to take into account the trust of referee, the reference trust R_{ij} is a combination between the single reference trust r_{ij}^l and the trust of referee S_{il} of referee l.

Reference trust R_{ij} of agent i on agent j is a non-weighted average:

$$R_{ij} = \begin{cases} \dfrac{\sum\limits_{l \in X_{ij}} g(S_{il}, r_{ij}^l)}{\mid X_{ij} \mid} & \text{if } X_{ij} \neq \varnothing \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where:

- g is a *reference-function* $g : [0,1] \times [0,1] \rightarrow [0,1]$, which satisfies the following conditions:

$$\begin{array}{ll} (i) & g(x_1, y) \leqslant g(x_2, y) \text{ if } x_1 \leqslant x_2 \\ (ii) & g(x, y_1) \leqslant g(x, y_2) \text{ if } y_1 \leqslant y_2 \end{array}$$

  These constraints are based on the intuitions:
  - The more the trust of referee l is high in the point of view of agent i, the more the reference trust R_{ij} is high;
  - The more the single reference trust r_{ij}^l is high, the more the final reference trust R_{ij} is high
- S_{il} is the trust of i on the referee l
- r_{ij}^l is the single reference trust of agent i about partner j referred via referee l

D. Overall trust

Overall trust T_{ij} of agent i in agent j is defined by the formula:

$$T_{ij} = t(E_{ij}, R_{ij}) \quad (5)$$

where:

- t is a *overall-trust-function*, $t : [0,1] \times [0,1] \rightarrow [0,1]$, which satisfies the following conditions:

$$\begin{array}{ll} (i) & min(e,r) \leqslant t(e,r) \leqslant max(e,r); \\ (ii) & t(e_1, r) \leqslant t(e_2, r) \text{ if } e_1 \leqslant e_2; \\ (iii) & t(e, r_1) \leqslant t(e, r_2) \text{ if } r_1 \leqslant r_2. \end{array}$$

  This combination satisfies these intuitions:
  - It must neither lower than the minimal and nor higher the maximal of experience trust and reference trust;
  - The more the experience trust is high, the more the *overall trust* is high;
  - The more the reference trust is high, the more the *overall trust* is high.
- E_{ij} is the experience trust of agent i about partner j.
- R_{ij} is the reference trust of agent i about partner j.

E. Updating trust

Agent i's trust in agent j can be changed in the whole its life-time whenever there is at least one of these conditions occurs (as showed in Algorithm 1, line 2):

- There is a new transaction between i and j occurring (line 3), so the experience trust of i on j changed.
- There is a referee l who shares to i his new experience trust about partner j (line 10). Thus the reference trust of i on j is updated.

1: **for all** agent i in the system **do**
2: **if** (there is a new transaction $k-$th with agent j) or (there is a new reference trust E_{lj} from agent l about agent j) **then**
3: **if** there is a new transaction k with agent j **then**
4: $t_{ij}^k \leftarrow$ a value in interval [0,1]
5: $t_{ij} \leftarrow t_{ij} \cup t_{ij}^k$
6: $t_{ij} \leftarrow Sort(t_{ij})$
7: $w \leftarrow GenerateW(k)$
8: $E_{ij} \leftarrow \sum_{h=1}^{k} t_{ij}^h * w_h$
9: **end if**
10: **if** there is a new reference trust E_{lj} from agent l about agent j **then**
11: $r_{ij}^l \leftarrow E_{lj}$
12: $X_{il} \leftarrow X_{il} \cup \{j\}$
13: $S_{il} \leftarrow \frac{1}{\mid X_{il} \mid} * \sum_{j \in X_{il}} h(E_{ij}, r_{ij}^l)$
14: $R_{ij} \leftarrow \frac{\sum_{l \in X_{ij}} g(S_{il}, r_{ij}^l)}{\mid X_{ij} \mid}$
15: **end if**
16: $T_{ij} \leftarrow t(E_{ij}, R_{ij})$
17: **end if**
18: **end for**

Algorithm 1: Trust Updating Algorithm

E_{ij} is updated after the occur of each new transaction between i and j as follows (lines 3 - 9):

- The new transaction's trust value t_{ij}^k is placed at the first position of vector t_{ij} (lines 4 - 6). Function $Sort(t_{ij})$ sorts the vector t_{ij} in ordered in time.
- Vector w is also generated again (line 7) in function $GenerateW(k)$.
- E_{ij} is updated by applying formulas 1 with the new vector t_{ij} and w (line 8).

Once E_{ij} is updated, agent i sends E_{ij} to its friend agents. Therefore, all i's friends will update their reference trust when they receive E_{ij} from i. We suppose that all friend relations in system are bilateral, this means that if agent i is a friend of agent j then j is also a friend of i. After having received E_{lj} from agent l, agent i then updates her/his reference trust R_{ij} on j as follows (lines 10 - 15):

- In order to update the individual reference trust r_{ij}^l, the value of E_{lj} is placed at the position of the old one (line 11).

- Agent j will be also added into X_{il} to recalculate the referee trust S_{il} and recalculate the reference trust R_{ij} (lines 12 - 14).

Finally, T_{ij} is updated by applying the formulas 5 from new E_{ij} and R_{ij} (line 16).

IV. Experimental Evaluation

This section presents the evaluation of the proposed model by taking emperimental data. Section IV-A presents the setting up our experiment application. Section IV-B evaluates the need of avoiding liars in refering of reputation. Section IV-C evaluates the need of combination of experience trust and reputation even if there are liars in refering reputation.

A. Experiment Setup

1) An E-market: An e-market system is composed of a set of seller agents, a set of buyer agents, and a set of transactions. Each transaction is performed by a buyer agent and a seller agent. A seller agent plays the role of a seller who owns a set of products and it could sell many products to many buyer agents. A buyer agent plays the role of a buyer who could buy many products from many seller agents.

- Each seller agent has a set of products to sell. Each product has a quality value in the interval $[0, 1]$. The quality of product will be assigned as the transaction trust of the transaction in which the product is sold.
- Each buyer agent has a transaction history for each of its sellers to calculate the experience trust for the corresponding seller. It has also a set of reference trusts referred from its friends. The buyer agent will update its trust on its sellers once it finishes a transaction or receives a reference trust from one of its friends. The buyer chooses the seller with the highest final trust when it want to buy a new product. The calculation to estimate the highest final trust of sellers is based on the proposed model in this paper.

2) Objectives: The purpose of these experiments is to answer two following questions:

- First, is it better if buyer agent judges the sharing trust of its referees than does not judge it? In order to answer to this question, the proposed model will be compared with the model of Jennings et al.'s model [12], [13] (Section IV-B).
- Second, what is better if buyer agent uses only its experience trust in stead of combination of experience and reference trust? In order to answer this question,the proposed model will be compared with the model of Manchala's model [19] (Section IV-C).

3) Initial Parameters: In order to make the results comparable, and in order to avoid the effect of random aspect in value initiation of simulation parameters, the same values for input parameters of all simulation scenarios will be used: number of sellers; number of products; number of simulations. These values are presented in the Table.II.

TABLE II: Value of parameters in simulations

Parameters	Values
Number of runs for each scenario	100 (times)
Number of sellers	100
Number of buyers	500
Number of products	500000
Average number of bought products/buyer	100
Average number of friends/buyer	300 (60% of buyers)

4) Analysis and evaluation criteria: Each simulation scenario will be ran at least 100 times. At the output, the following parameter will be calculated:

- The average quality (in %) of brought products for all buyers. A model (strategy) is considered better if it brings the higher average quality of brought products for all buyers in the system.

B. The need of avoiding liar in reputation

1) Scenarios: The question need to be answerd is: is it better if buyer agent uses reputation with trust of referees (agent judges the sharing trust of its referees) or uses reputation without trust of referees (agent does not judge the sharing trust of its referees)? In order to answer this question, there are two strategies will be simulated:

- *Strategy A - using proposed model*: Buyer agent refers the reference trust (about sellers) from other buyers *with* taking into account the trust of referee.
- *Strategy B - using model of Jennings et al. [12], [13]*: Buyer agent refers the reference trust (about sellers) from other buyers *without* taking into account the trust of referee.

The simulations are launched in various values of the percentage of lying buyers in the system (0%, 30%, 50%, 80%, and 100%).

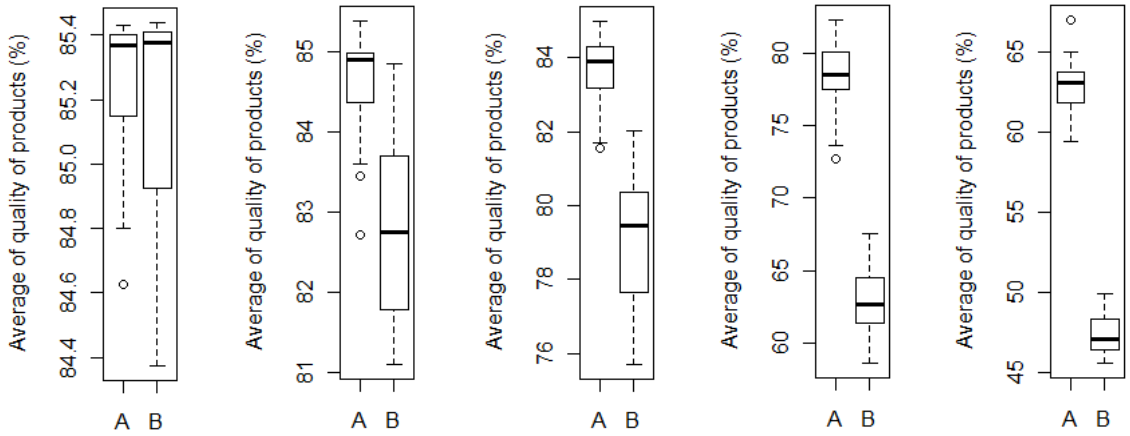
2) Results: The results indicate that the average quality of bought products of all buyers in the case of using reputation with considering of trust of referees is always significantly higher than those in the case using reputation without considering of trust of referees.

When there is no lying buyer (Fig.1.a). The average quality of bought products for all buyers in the case using strategy A is not significantly different from that in the case using strategy B ($M(A) = 85.24\%$, $M(B) = 85.20\%$, significant difference with $p\text{-}value > 0.7$)[1].

When there is 30% of buyers is liar (Fig.1.b). The average quality of bought products for all buyers in the case using strategy A is significantly higher than in the case using strategy B ($M(A) = 84.64\%$, $M(B) = 82.76\%$, significant difference with $p\text{-}value < 0.001$).

When there is 50% of buyers is liar (Fig.1.c). The average quality of bought products for all buyers in the case using strategy A is significantly higher than in the case using strategy

[1]We use the *t-test* to test the difference between two sets of average quality of bought products of two scenarios, therefore if the probability value $p\text{-}value < 0.05$ we could conclude that the two sets are *significantly different*.



(a) 0% liars (b) 30% liars (c) 50% liars (d) 80% liars (e) 100% liars

Fig. 1: Significant difference of average quality of bought products of all buyers from the case using proposed model (strategy A) and the case using Jennings et al.'s model (strategy B)

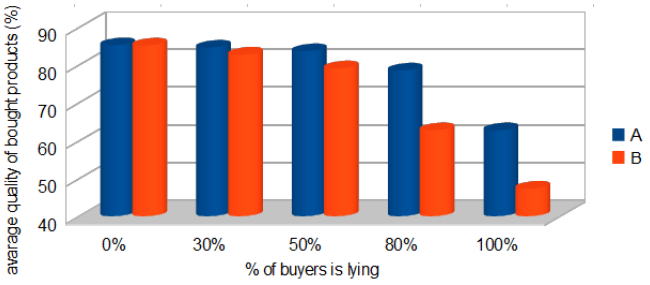


Fig. 2: Summary of difference of average quality of bought products of all buyers between the case using our model (A) and the case using Jennings et al.'s model (B)

B ($M(A) = 83.68\%$, $M(B) = 79.11\%$, significant difference with $p\text{-}value < 0.001$).

When there is 80% of buyers is liar (Fig.1.d). The average quality of bought products for all buyers in the case using strategy A is significantly higher than in the case using strategy B ($M(A) = 78.55\%$, $M(B) = 62.76\%$, significant difference with $p\text{-}value < 0.001$).

When all buyers are liar (Fig.1.e). The average quality of bought products for all buyers in the case using strategy A is significantly higher than in the case using strategy B ($M(A) = 62.78\%$, $M(B) = 47.31\%$, significant difference with $p\text{-}value < 0.001$).

In summary, as being depicted in the Fig.2, the more the percentage of liar in buyers is high, the more the average quality of bought products of all buyers in the case using our model (strategy A) is significantly higher than those in the case using Jennings et al.'s model [12], [13] (strategy B).

C. The need of combination of experience with reputation

1) Scenarios: The results of the first evaluation suggest that using reputation with considering of trust of referees is better than using reputation without considering of trust of referees, especially in the case there are some liars in sharing their private trust about partners to others. And in turn, another question arises: in the case there are some liars in sharing data to their friends, is it better if buyer agent use reputation with considering of trust of referees or use only experience trust to avoid liar reputation? In order to answer this question, there are two strategies also simulated:

- *Strategy A - using proposed model*: Buyer agent refers the reference trust (reputation) from other buyers by taking into account their considering of trust of referees.
- *Strategy C - using Manchala's model [19]*: Buyer agent does not refer any reference trust from other buyers. It bases only on its experience trust.

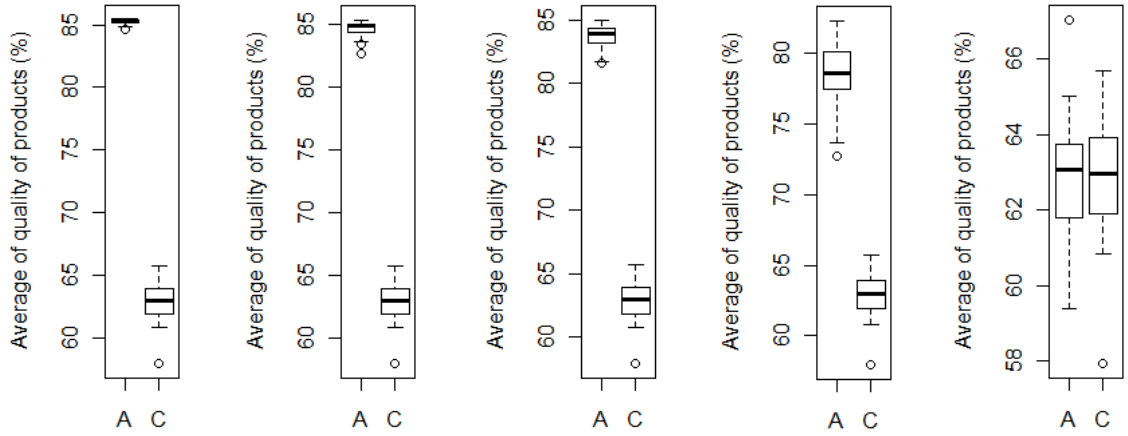
The simulations are also launched in various values of the percentage of lying buyers in the system (0%, 30%, 50%, 80%, and 100%).

2) Results: The results indicate that the average quality of bought products of all buyers in the case with considering of trust of referees is almost significantly higher than those in the case using only the experience trust.

When there is no lying buyer (Fig.3.a). The average quality of bought products for all buyers in the case using strategy A is significantly higher than in the case using strategy C ($M(A) = 85.24\%$, $M(C) = 62.75\%$, significant difference with $p\text{-}value < 0.001$).

When there is 30% of buyers is liar (Fig.3.b). The average quality of bought products for all buyers in the case using strategy A is significantly higher than the in case using strategy C ($M(A) = 84.64\%$, $M(C) = 62.74\%$, significant difference with $p\text{-}value < 0.001$).

When there is 50% of buyers is liar (Fig.3.c). The average quality of bought products for all buyers in the case using strategy A is significantly higher than in the case using C ($M(A) = 83.68\%$, $M(C) = 62.76\%$, significant difference with $p\text{-}value < 0.001$).



(a) 0% liars (b) 30% liars (c) 50% liars (d) 80% liars (e) 100% liars

Fig. 3: Significant difference of average quality of bought products of all buyers between the case using proposed model (strategy A) and the case using Manchala's model (strategy C)

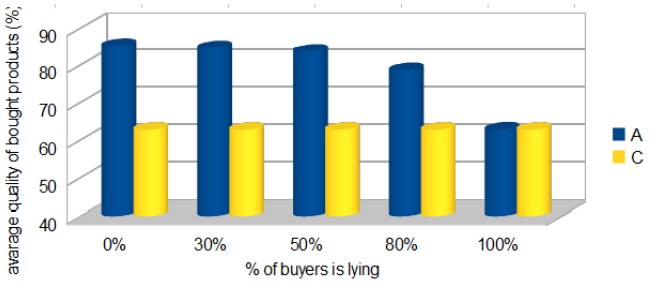


Fig. 4: Summary of difference of average quality of bought products of all buyers between the case using our model (A), and the case using Manchala's model (C)

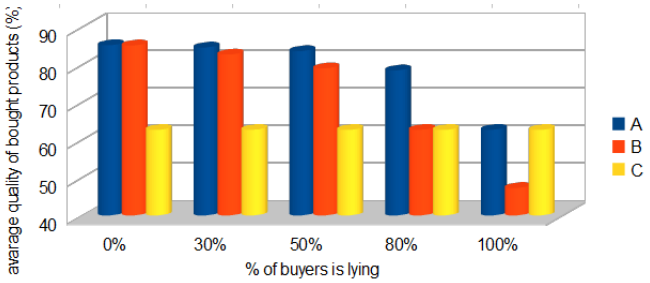


Fig. 5: Summary of difference of average quality of bought products of all buyers among the case using our model (A), the case using Jennings et al.'s model (B), and the case using Manchala's model (C)

When there is 80% of buyers is liar (Fig.3.d). The average quality of bought products for all buyers in the case using strategy A is significantly higher than in the case using strategy C ($M(A) = 78.55\%$, $M(C) = 62.78\%$, significant difference with $p\text{-}value < 0.001$).

When all buyers are liar (Fig.3.e). There is no significant difference between the case using strategy A and the case using strategy C ($M(A) = 62.78\%, M(C) = 62.75\%$, significant difference with $p\text{-}value > 0.6$). It is intuitive because in our model (strategy A), when almost referees are not trustworthy, the trustor tends to trust in himself instead of other. In other word, the trustor has the tendency to base on its won experience rather than others.

The overall result is depicted in the Fig.4. In almost cases, the average quality of bought products of all buyers in the case of using our model is always significantly higher than those in the case of using Manchala's model [19]. In the case that all buyers are liar, there is no significant difference of the average quality of bought products from all buyers between two strategies.

In summary, Fig.5 illustrates the value of average quality of bought products of all buyers in three scenarios. In the case there is no lying buyer, this value is the highest in the case using our model and Jennings et al.'s model [12], [13] (there is no significant difference between two mosels in this situation). Using Manchala's model [19] is the worst case in this situation. In the case there are 30%, 50% and 80% buyers to be lying, the value is always highest in the case of using our model. In the case that all buyers are liar, there is no significant difference between agents using our model and agents using Manchala's model [19]. Both of these two strategies win a much more higher value compared with the case using Jennings et al.'s model [12], [13].

V. Discussion

Let us consider a scenario of an e-commerce application. There are two concurrent sellers S_1 and S_2 who sell the same product x, there is an independent third party site w which collects the consumer's opinions. All clients could submit its opinions about sellers. In this case, the site w could be considered as a reputation channel for client: a client could refer the given opinions on the site w to choose the best seller. However, because the site w is a public reputation: all clients could submit their opinions. Imagining that S_1 is really trustworthy, but S_2 is not fair, some of its employments

intentionally submit some negative opinions about the seller S_1 in order to attract more clients from S_1 to S_2.

Let consider this application in two cases. Firstly, the case without mechanism to avoid liars in the applied trust model. If an user i is considering to buy a product x that both S_1 and S_2 are selling. User i refers the reputation of S_1 and S_2 on the site w. Since there is not any mechanism to avoid liars in the trust model, the more negative opinions from S_2's employments are given about S_1, the lower the reputation of S_1 is. Therefore, the lower the possibility that user i chooses buying the product x from S_1.

Secondly, the case of our proposed model with lying against mechanism. User i will refer the reputation of S_1 and S_2 on the site w with considering the sharing trust of the owner of each opinion. Therefore, the ones from S_2 who gave negative opinions about S_1 will be detected as liars. Their opinion weights thus will be decreased (considered as unimportant ones) when calculating the reputation of S_1. Consequently, the reputation of S_1 will stay high no matter how many people from S_2 intentionally lie about S_2. In other word, our model helps agent to avoid some liars in calculating the reputation of a given partner in multiagent systems.

VI. Conclusion

This paper presented a model of trust which enables agents to calculate, estimate and update trust's degree on their partners based not only on their own experiences, but also based on the reputation of partners. The partner reputation is estimated from the judgements from referees in the community. In which, the model taken into account the trustworthiness of the referee in judging a partner.

The experimental evaluation of the model has been set up for multiagent system in the e-commerce environment. The research results indicate, firstly, that it is better to take into account the trust of referees to estimate the reputation of partners. Seconly, it is better to combine the experience trust and the reputation than using only the experience trust in estimating the trust of a partner in the multiagent system.

Constructing and selecting a strategy, which is appropriate to the context of some application of a multiagent system, need to be investigated furthermore. These research issues will be presented in our future work.

References

[1] Masoom Alam, Shahbaz Khan, Quratulain Alam, Tamleek Ali, Sajid Anwar, Amir Hayat, Arfan Jaffar, Muhammad Ali, and Awais Adnan. Model-driven security for trusted systems. *International Journal of Innovative Computing, Information and Control*, 8(2):1221–1235, 2012.

[2] Chris Burnett, Timothy J. Norman, and Katia Sycara. Bootstrapping trust evaluations through stereotypes. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1*, AAMAS '10, pages 241–248, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems.

[3] J. Carter, E. Bitting, and A. Ghorbani. Reputation formalization for an information-sharing multi-agent sytem. *Computational Intelligence*, 18(2):515–534, 2002.

[4] B. Esfandiari and S. Chandrasekharan. On how agents make friends: Mechanisms for trust acquisition. In *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, pages 27–34, Montreal, Canada, 2001.

[5] Hui Fang, Yang Bao, and Jie Zhang. Misleading opinions provided by advisors: Dishonesty or subjectivity. IJCAI/AAAI, 2013.

[6] Hui Fang, Jie Zhang, and Nadia Magnenat Thalmann. A trust model stemmed from the diffusion theory for opinion evaluation. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, pages 805–812, Richland, SC, 2013. International Foundation for Autonomous Agents and Multiagent Systems.

[7] D. Gambetta. Can we trust trust? In D. Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*, pages 213–237. Basil Blackwell, New York, 1990.

[8] Tyrone Grandison and Morris Sloman. Specifying and analysing trust for internet applications. In *Proceedings of the 2nd IFIP Conference on e-Commerce, e-Business, e-Government*, Lisbon, Portugal, October 2002.

[9] R. Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Propagation of trust and distrust. In *Proceedings of the 13th international conference on World Wide Web*, WWW '04, pages 403–412, New York, NY, USA, 2004. ACM.

[10] Chung-Wei Hang, Yonghong Wang, and Munindar P. Singh. Operators for propagating trust and their evaluation in social networks. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, AAMAS '09, pages 1025–1032, Richland, SC, 2009. International Foundation for Autonomous Agents and Multiagent Systems.

[11] Ramón Hermoso, Holger Billhardt, and Sascha Ossowski. Role evolution in open multi-agent systems as an information source for trust. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1*, AAMAS '10, pages 217–224, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems.

[12] Dong Huynh, Nicholas R. Jennings, and Nigel R. Shadbolt. Developing an integrated trust and reputation model for open multi-agent systems. In *Proceedings of the 7th Int Workshop on Trust in Agent Societies*, pages 65–74, New York, USA, 2004.

[13] Trung Dong Huynh, Nicholas R. Jennings, and Nigel R. Shadbolt. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 13(2):119–154, 2006.

[14] Siwei Jiang, Jie Zhang, and Yew-Soon Ong. An evolutionary model for constructing robust trust networks. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, pages 813–820, Richland, SC, 2013. International Foundation for Autonomous Agents and Multiagent Systems.

[15] Audun Josang, Claudia Keser, and Theo Dimitrakos. Can we manage trust? In *Proceedings of the 3rd International Conference on Trust Management, (iTrust)*, Paris, 2005.

[16] Yarden Katz and Jennifer Golbeck. Social network-based trust in prioritized default logic. In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI-06)*, volume 21, pages 1345–1350, Boston, Massachusetts, USA, jul 2006. AAAI Press.

[17] Y. Lashkari, M. Metral, and P. Maes. Collaborative interface agents. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*. AAAIPress, 1994.

[18] Xiaoyong Li and Xiaolin Gui. Tree-trust: A novel and scalable P2P reputation model based on human cognitive psychology. *International Journal of Innovative Computing, Information and Control*, 5(11(A)):3797–3807, 2009.

[19] D. W. Manchala. E-commerce trust metrics and models. *IEEE Internet Comp.*, pages 36–44, 2000.

[20] Samia Nefti, Farid Meziane, and Khairudin Kasiran. A fuzzy trust model for e-commerce. In *Proceedings of the Seventh IEEE International Conference on E-Commerce Technology (CECÓ5)*, pages 401–404, 2005.

[21] Manh Hung Nguyen and Dinh Que Tran. A computational trust model with trustworthiness against liars in multiagent systems. In Ngoc Thanh Nguyen et al., editor, *Proceedings of The 4th International Conference on Computational Collective Intelligence Technologies and Applications (ICCCI), Ho Chi Minh City, Vietnam, 28-30 November 2012*, pages 446–455. Springer-Verlag Berlin Heidelberg, 2012.

[22] Manh Hung Nguyen and Dinh Que Tran. A multi-issue trust model

in multiagent systems: A mathematical approach. *South-East Asian Journal of Sciences*, 1(1):46–56, 2012.

[23] Manh Hung Nguyen and Dinh Que Tran. A combination trust model for multi-agent systems. *International Journal of Innovative Computing, Information and Control (IJICIC)*, 9(6):2405–2421, June 2013.

[24] S. D. Ramchurn, C. Sierra, L. Godo, and N. R. Jennings. Devising a trust model for multi-agent interactions using confidence and reputation. *International Journal of Applied Artificial Intelligence*, 18(9–10):833–852, 2004.

[25] Steven Reece, Alex Rogers, Stephen Roberts, and Nicholas R. Jennings. Rumours and reputation: Evaluating multi-dimensional trust within a decentralised reputation system. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '07, pages 165:1–165:8, New York, NY, USA, 2007. ACM.

[26] Jordi Sabater and Carles Sierra. Regret: A reputation model for gregarious societies. In *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, pages 61–69, Montreal, Canada, 2001.

[27] Jordi Sabater and Carles Sierra. Reputation and social network analysis in multi-agent systems. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02)*, pages 475–482, Bologna, Italy, July 15–19 2002.

[28] M. Schillo, P. Funk, and M. Rovatsos. Using trust for detecting deceitful agents in artificial societites. *Applied Artificial Intelligence (Special Issue on Trust, Deception and Fraud in Agent Societies)*, 2000.

[29] S. Sen and N. Sajja. Robustness of reputation-based trust: Booblean case. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02)*, pages 288–293, Bologna, Italy, 2002.

[30] Junko Shibata, Koji Okuhara, Shogo Shiode, and Hiroaki Ishii. Application of confidence level based on agents experience to improve internal model. *International Journal of Innovative Computing, Information and Control*, 4(5):1161–1168, 2008.

[31] W. T. Luke Teacy, Jigar Patel, Nicholas R. Jennings, and Michael Luck. Travos: Trust and reputation in the context of inaccurate information sources. *Journal of Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006.

[32] W.T. Luke Teacy, Michael Luck, Alex Rogers, and Nicholas R. Jennings. An efficient and versatile approach to trust and reputation using hierarchical bayesian modelling. *Artif. Intell.*, 193:149–185, December 2012.

[33] Patricia Victor, Chris Cornelis, Martine De Cock, and Paulo Pinheiro da Silva. Gradual trust and distrust in recommender systems. *Fuzzy Sets and Systems*, 160(10):1367–1382, 2009. Special Issue: Fuzzy Sets in Interdisciplinary Perception and Intelligence.

[34] George Vogiatzis, Ian Macgillivray, and Maria Chli. A probabilistic model for trust and reputation. AAMAS, 225-232 (2010)., 2010.

[35] Andrew Whitby, Audun Josang, and Jadwiga Indulska. Filtering out unfair ratings in bayesian reputation systems. In *Proceedings of the 3rd International Joint Conference on Autonomous Agenst Systems Workshop on Trust in Agent Societies (AAMAS)*, 2005.

[36] B. Yu and M. P. Singh. Distributed reputation management for electronic commerce. *Computational Intelligence*, 18(4):535–549, 2002.

[37] B. Yu and M. P. Singh. An evidential model of distributed reputation management. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02)*, pages 294–301, Bologna, Italy, 2002.

[38] L. A. Zadeh. A computational approach to fuzzy quantifiers in natural languages. pages 149–184, 1983.

[39] Jie Zhang and Robin Cohen. A framework for trust modeling in multiagent electronic marketplaces with buying advisors to consider varying seller behavior and the limiting of seller bids. *ACM Trans. Intell. Syst. Technol.*, 4(2):24:1–24:22, April 2013.