A Hybrid Intrusion Detection Model for Identification of Threats in Internet of Things Environment

Nsikak Pius Owoh[1]
Department of Cyber Security, School of Information and Communication Technology, Federal University of Technology, Owerri, Imo State, Nigeria

Manmeet Mahinderjit Singh[2], Zarul Fitril Zaaba[3]
School of Computer Sciences
Universiti Sains Malaysia, 11800 USM
Penang, Malaysia

***Abstract*—Internet of Things (IoT) has transcended from its application in traditional sensing networks such as wireless sensing and radio frequency identification to life-changing and critical applications. However, IoT networks are still vulnerable to threats, attacks, intrusions, and other malicious activities. Intrusion Detection Systems (IDS) that employ unsupervised learning techniques are used to secure sensitive data transmitted on IoT networks and preserve privacy. This paper proposes a hybrid model for intrusion detection that relies on a dimension reduction algorithm, an unsupervised learning algorithm, and a classifier. The proposed model employs Principal Component Analysis (PCA) to reduce the number of features in a dataset. The K-means algorithm generates clusters that serve as class labels for the Support Vector Machine (SVM) classifier. Experimental results using the NSL-KDD and the UNSW-NB15 datasets justify the effectiveness of our proposed model in detecting malicious activities in IoT networks. The proposed model, when trained, identifies benign and malicious behaviours using an unlabelled dataset.**

***Keywords*—Internet of things; intrusion detection system; k-means; principal component analysis; support vector machine**

I. Introduction

Internet of Things (IoT) is a self-organizing and adaptive network that interconnects uniquely identifiable "Things" to the internet via communication protocols [1]. The "Things" (also known as devices) are capable of sensing data from humans and the environment. IoT devices collect and sometimes store information that can be accessed pervasively and at any time. The Internet of Things (IoT) is a proliferating technology that offers many advantages in many areas of life [2]. However, the IoT is faced with several information security vulnerabilities and threats. Considering the intrinsic computational limitations of IoT devices and their vulnerabilities and the increasing rate of unauthorized access to these devices [3], IoT risks increase exponentially. Threats to the IoT network are similar to a traditional network, which threatens confidentiality, integrity, and availability. Such threats, when exploited, may lead to eavesdropping, data leakage/loss, and denial-of-service attacks [4].

The connection of IoT devices to the internet through vulnerable networks such as 6LoWPAN and IPv6 makes them susceptible to various intrusions. Nevertheless, these intrusions can be detected by intrusion detection systems (IDS) [5]. Intrusion detection systems (IDS) can identify internal and external attacks [6]. Though a post-active security measure, Intrusion detection systems can identify attacks in networks using adaptive network detection algorithms and act as a multilayer security mechanism to cryptographic solutions in a network. The different types of IDS are signature-based (misuse), anomaly-based, and specification-based detection systems.

In signature-based detection systems, predefined attack patterns are modelled and stored in a database. IDSs of this type accurately detect known intrusions. Also, low false-positive rates and minimal computation overhead are experienced with signature-based IDS. However, they ignore unknown intrusions, making them ineffective in detecting network attacks [7]. On the other hand, anomaly-based detection systems employ statistical or machine learning approaches to identify unusual (possible threats) from normal behaviours in network traffic or system activities. Detection, in this case, is based on the features and labels in each data. Detection rates are higher with the anomaly-based system since they can detect new and unseen attacks. Nevertheless, increased computation overhead and false alarms are some drawbacks of anomaly-based IDSs [7]. Specification-based detection systems are like anomaly-based detection systems but require involvement of users in obtaining valid network traffic to develop a normal behaviour model [5].

A significant problem with anomaly detection systems is that they require unlabelled data. This approach is challenging because of the difficulty of acquiring large datasets that are labelled as "normal" or "malicious." Detecting anomalies in IoT becomes even more complicated when applied to high-dimensional data with large features. High-dimension datasets often reduce the accuracy of anomaly detection systems due to the presence of irrelevant features, exponential search space, and data bias [8]. To this end, there is a need for a detection system capable of detecting threats (such as anomalies and attacks) in an IoT network with high accuracy using unlabelled data. Achieving the proposed high accuracy would require the removal of irrelevant and redundant data through feature reduction.

This paper proposes a hybrid intrusion detection system for IoT, which relies on PCA for dimension reduction, K-means for threats clustering, and SVM for anomaly classification. To the best of our knowledge, this is the first paper to apply these algorithms to detect anomalies in both unlabelled and labelled datasets. The contributions of this paper are summarized as follows:

1) To develop an intrusion detection model that performs feature reduction and anomaly detection in unlabelled and labelled datasets.

2) To build a classification model using the generated cluster labels from the unsupervised learning phase.

3) To evaluate the performance of the anomaly detection model when trained with different number of clusters and features.

The rest of this paper is structured as follows: Section II presents a review of related works on attacks in the IoT and intrusion detection systems used in identifying such threats. In Section III, we present our proposed hybrid intrusion detection model. Furthermore, datasets and methods used for data clustering and classifier training are also discussed in this section. The hybrid model results, including feature reduction, data clustering, and binary and multi-class classification, are shown in Section IV. In Section V, we discuss obtained results and conclude the paper in Section VI.

II. Related Work

Akin to the desired security requirements in traditional networks, IoT networks need to ensure confidentiality, integrity, availability, non-repudiation, and privacy. It is worthy to note that, in IoT networks, a breach in any of these requirements can be life-threatening because of its applicability and peculiarity [9]. The availability of sensitive data in IoT devices makes them an attractive target for cyber-attacks. Threats on IoT networks are increasing massively, especially as IoT devices can automatically join and leave sensor networks [10]. Another reason for the increasing number of successful IoT attacks is their limited resources (power, storage, and computational capabilities). These constraints make it challenging to implement sophisticated security and privacy mechanisms [11].

A. Attacks on the Internet of Things (IoT)

There are several possible attacks on IoT networks. Among these attacks, distributed denial of service (DDoS) attack has grown to become one of the most severe. Even so, its detection and prevention have also been a security challenge. DDoS exploits compromised devices (zombie or botnet) to flood IoT devices or communication channels with bogus requests and eventually rendering their services unavailable to legitimate users. Solving this problem has brought about several proposed solutions in different applications and networks. However, detecting and preventing DDoS attacks is tasking due to the difficulty of differentiating attack packets from legitimate ones. Even more troubling is that DDoS attacks can be perpetuated over any of the four layers of the IoT [11]. In what follows, we enumerate some attacks at each layer of the IoT.

The perception layer, also referred to as the sensing layer, handles the data gathering from users and the environment. It employs technologies such as wireless sensor networks (WSNs), radio frequency identification (RFID), mobile crowdsensing (MCS), and micro-electro-mechanical (MEMS) [12]. Eavesdropping, tag cloning, spoofing, unauthorized access, and Radio Frequency jamming are some of the attacks in this layer. These attacks compromise devices by affecting vital architectural components of the IoT system. Memory corruption and misconfiguration of IP addresses are reasons for these attacks [13].

The network layer transmits sensor data between the information processing system and sensor devices using communication infrastructures such as wired and wireless connections. Attacks in the network layer include sinkhole, Man-In-The-Middle, Sybil, and DDoS attacks [14]. In the network attack, an adversary targets intercommunication among devices by causing latency or dropping sent messages. Such attacks destroy computational processes within the IoT configuration systems. The middleware layer guarantees and oversees services needed by applications or clients. Furthermore, service management and database connection are handled in this layer. DoS and unauthorized access are possible attacks in this layer [14].

The application layer consists of interaction techniques of users and applications, and it conveys application services to users. Attacks such as phishing, sniffing, code injection, and DoS are possible threats in the application layer. These attacks compromise system applications (Mobile and Web applications) [13]. Table I summarizes the different attack types at the different layers of the IoT.

B. Intrusion Detection Systems in the Internet of Things (IoT)

Predicting threats or detecting them at their initial stages effectively prevents successful attacks on IoT devices [15]. Interestingly, several cybersecurity tasks can be performed using machine learning. These tasks include anomaly detection, spam filtering, user monitoring, risk analysis, and zero-day exploit identification [16]. Machine learning algorithms have been used widely in developing intrusion detection systems for IoT networks. Its adoption in this area is justified in its ability to detect anomalies in network traffic. Based on their properties, data usage patterns, and learning style, machine learning algorithms are classified into three groups: supervised, unsupervised, and semi-supervised algorithms [17]. The algorithm is trained using training data (labelled input) in supervised learning, often called ground truth [18].

TABLE I. Attack Classification in the different Layers of the Internet of Things

	Perception layer	**Network layer**	**Middleware layer**	**Application layer**
Components	GPS, RFID tags, RFID reader/writers, Barcodes, BLE devices	WLAN, Social networks, WSNs, Cloud network	Database, Service APIs, Service management	Interface, Smart applications
Possible Attacks	Code injection, Noisy data, Unauthorized access	Routing attacks, DoS, DDoS, Network congestion	Spoofing, DoS, Malicious information, Unauthorized access, Data manipulatio	Phishing, Misconfigurations, Code injection.

On the other hand, unsupervised learning algorithms do not require labels in the training datasets as they can infer from the input data. They can reveal the hidden structure and distribution in data which provides more information about the data. A typical example of this category of algorithms is clustering (K-means). With clustering, structures or patterns in an unlabelled dataset are identified by grouping the data of interest into *k* number of clusters [18].

The work proposed by Li et al. [19] presents an approach that employs deep belief networks and Autoencoder for intrusion detection. The authors evaluated their proposed system using the KDD-CUPP 99 dataset. The authors' results from the 2000 records show that the proposed hybrid system can accurately detect anomalies in data but takes too long to pre-process data. Similarly, an unsupervised hybrid architecture for anomaly detection in large-scale high-dimensional is proposed by Erfani, Rajasegarar [8]. This work also evaluated the performance of deep belief networks against one-class SVMs when detecting anomalies in high-dimensional data. The DBN in the proposed system extracts only relevant features in the dataset, while the ISVM is trained using the extracted features. However, the datasets used for the evaluation of the proposed model do not ideally simulate real-world scenarios. In Nskh, Varma [20], a dimension reduction and classifier model relies on the KDD Cup 99 dataset is proposed. The model employs Principal Component Analysis for dimension reduction and Support Vector Machine for attack classification. However, the model is non-trivial, and the computing complexity of the model is not provided.

Meanwhile, Pajouh, Javidan [21] proposed a two-layer dimension reduction and two-tier classification model for intrusion detection in IoT. The model uses Principal Component Analysis and Linear Discriminant Analysis for feature extraction, while Naïve Bayes and K-nearest Neighbour algorithms are used for attack classification. The authors show that the model is trivial as it uses fewer computing and memory resources. Zhao, Li [22] present a model for anomaly-based intrusion detection in IoT. The model is based on PCA for dimension reduction and SoftMax Regression for classification. Low computing complexity was obtained with the reduced dimension, while accurate detection was accomplished with small training sets. Accuracy results obtained from the SoftMax regression model are 84.9%, 84.4%, and 84.4% for 3, 6, and 10 features, respectively. SVM classifier, on the other hand, produced slightly better results when tested with similar features.

A malware detection model for IoT devices that employ KNN and Random Forest classifiers were developed in Narudin, Feizollah [23]. KNN used in the proposed system allocates network traffic to a class with the most objects among its K-nearest neighbours. On the other hand, the random forest uses the labelled network traffic from the KNN classifiers to develop decision trees that identify malware in network traffic. Obtained results from the experiments performed with the MalGenome dataset show a true positive rate (TPR) of 99.7% and 99.9% for KNN and Random Forest, respectively. A Host-based Intrusion Detection and Mitigation framework for home-based IoT is proposed in Nobakht, Sivaraman [24]. The framework uses software-defined networks (SDN) and machine learning techniques to ensure security in IoT devices. The authors of this work also proposed an attack simulation model that collects data then distinguishes malicious actions from normal activities.

A machine learning framework that detects DDoS attacks in the IoT by collecting data, extracting its features, and performing binary classification is shown in Doshi, Apthorpe [25]. The proposed framework has four steps: traffic capture, packet grouping, feature extraction, and binary classification. The authors also evaluated several classifiers, including support vector machine, K-Nearest Neighbour (KNN), Decision Trees (DT), Neural Networks (NN), and Random Forests. Furthermore, Abeshu and Chilamkurti (2018) proposed an intrusion detection system that uses deep learning. The proposed IDS can detect zero-day attacks in a fog-to-things computing environment using the NSL-KDD dataset for evaluation. The IDS model uses 150 neurons in the first layer, 120 in the second, 50 in the third, and a SoftMax layer in the last layer. Also, the model was compared with shallow models, and an accuracy score of 99.20% was obtained with a FAR of 0.85% against a FAR of 6.57% in shallow models. However, detecting attack types such as probing, DoS and U2R were omitted in the presented work.

Few works have been proposed on anomaly detection with the capability of dimension reduction and attack classification. These works mostly rely on labelled data for accurate attack classification in IoT networks. Zhao, Li [22] presented an anomaly detection system that employs PCA and SoftMax regression algorithms. However, the proposed method is based on a supervised learning model and only functions as a binary classifier that detects only normal or malicious attacks, leaving out other attack vectors. Furthermore, the authors evaluated their proposed system on the KDD-CUP 99 dataset, which contains old records. Considering this, we propose an anomaly detection system that employs an unsupervised learning technique with a classifier capable of detecting up to four classes of attacks present in the NSL-KDD dataset. We also evaluate our proposed hybrid model using the UNSW-NB15 dataset, a more recent dataset with new attack activities.

III. Methodology

This section presents the architecture of the proposed model, including the datasets and techniques employed for the detection of anomalies in the IoT.

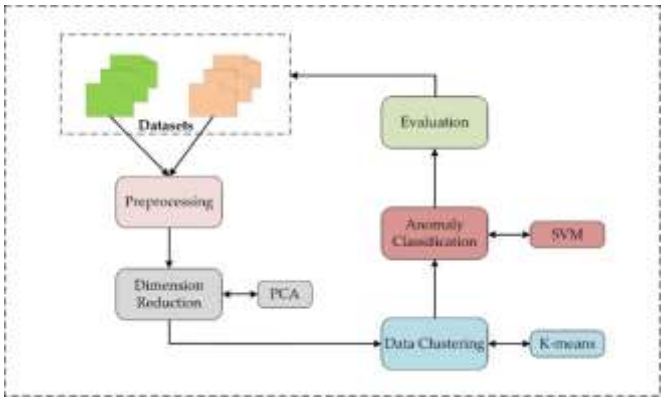

Fig. 1. Architecture of the Proposed Hybrid Detection Model.

A. Architecture

The architecture for our proposed model, as shown in Fig. 1, consists of three parts: dimension reduction, data clustering, and anomaly classification. The model is implemented in Python using available libraries such as SciKit-Learn, Pandas, Numpy [26], and Matplotlib [27]. The experiments which involved the implementation of all three components of the proposed model (i.e., PCA, K-means, and SVM) were performed on an Intel(R) Core (TM) i7500U CPU@2.70GHz laptop with a 12 GB RAM and running Windows 10 Home edition.

1) Dataset: The first dataset used in the proposed model is the NSL-KDD dataset [28]. The dataset is commonly used for the simulation of anomaly detection systems and models. Most of the inherent issues with the earlier KDD-CUP 99 dataset are resolved in the NSL-KDD dataset, and it is a preferred choice for baseline evaluation of IDSs. The dataset consists of training and testing datasets with 41 features: duration, protocol, service, flag, source bytes, destination bytes, and normal/attack labels. Furthermore, the dataset consists of 125,973 records for the training data and 22,544 records for the test data. The labels in the dataset can be categorized into four attack classes, which are Denial of Service (DoS) attack, User to Root (U2R) attack, Probing attack, and Remote to Local (R2L) attack. Table II presents the details of these attack classes.

a) Probing Attack: This attack involves scanning IoT targets and serves as a starting point for other attacks. Scanning programs are used to discover vulnerabilities in IoT applications. Tools such as mscan and saint can be used for this purpose.

b) Remote-to-Local (R2L): After a successful scan, the attacker may employ a remote-to-local ((R2L) attack to access the local system from remote ports, thereby escalating system privileges. Examples of this attack include *ftp-write*, *guest-exploit*, which either exploit poorly configured security policies or network programs.

c) User to Root (U2R) Attack: This attack originates from the R2L attacks and exploits unsecured programs running as roots. This attack-type leads to a buffer overflow caused by ffbconfig, fdformat, and eject.

d) Denial of Service (DoS) Attack: A denial-of-service (DoS) attack is successfully launched on a target machine or device by flooding such device with overloaded requests to stop legitimate requests from getting access to the device(s) [29].

Though the NSL-KDD dataset [28] solved most issues, such as data imbalance among normal and malicious records associated with the earlier KDDCUP dataset, the NSL-KDD dataset still does not depict present-day attack activities. To ascertain the effectiveness of our proposed hybrid model on recent malicious activities, we also evaluate the proposed model on the UNSW-NB15 dataset [30]. The UNSW-NB15 dataset consists of 49 features, including the class label. Table III shows the different features and categories in the dataset [30].

TABLE II. ATTACK CLASSIFICATION IN THE NSL-KDD DATASET

Probing	Remote to Local (R2L)	User to Root (U2R)	Denial of Service (DoS)
ipsweep	ftp_write	buffer_overflow	back
nmap	guess_passwd	Loadmodule	land
portsweep	imap	Perl	neptune
satan	Multihop, Phf	Rootkit	Pod, smurf
	spyware_client		teardrop
	spyware _master		

TABLE III. RECORD DISTRIBUTION OF THE UNSW-NB15 DATASET

Type	No. of Records	Description
Normal	2,218,761	The name of each attack category. In this data set, nine categories (e.g., Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms)
Fuzzers	24,246	0 for normal and 1 for attack records
Analysis	2,677	It contains different attacks of the port scan, spam, and HTML file penetrations.
Backdoors	2,329	A technique in which a system security mechanism is bypassed stealthily to access a computer or its data.
DoS	16,353	A malicious attempt to make a server or a network resource unavailable to users, usually by temporarily interrupting or suspending the services of a host connected to the internet
Exploits	44,525	The attacker knows of a security problem within an operating system or a piece of software and leverages that knowledge by exploiting the vulnerability.
Generic	215,481	A technique that works against all block-ciphers (with a given block and key size) without considering the block-cipher structure.
Reconnaissance	13,987	It contains all Strikes that can simulate attacks that gather information
Shellcode	1,511	A small piece of code is used as the payload in the exploitation of software vulnerability.
Worms	174	The attacker replicates itself to spread to other computers. Often, it uses a computer network to spread itself, relying on security failures on the target computer to access it.

2) Data pre-processing: For machine learning algorithms to perform optimally, feature scaling is necessary since the range of values may vary in the input data. The range of data of some features in the NSL-KDD and UNSW-NB15 datasets is enormous, and such dimensions determine the distance variance; hence the need for data normalization. Similar to the work proposed by Zhao, Li [22], we adopted the Min-Max normalization method to ensure that all the data values come under the range of 0 and1. This approach is presented mathematically in equation 1.

$$A_j^{(i)} = \frac{A_j^{(i)} - min}{max - min} \quad (1)$$

3) Dimension reduction: Dimension reduction was chosen in the proposed model to solve the problems faced with high dimensional data, typical with anomaly-based datasets such as the NSL-KDD and UNSW-NB15 datasets [28, 30]. The high dimensional data contain redundant and irrelevant features, which degrade the performance of the detection model. In the proposed model, the 41 features present in the NSL-KDD dataset and the 49 features in the UNSW-NB15 dataset are reduced using PCA to 3, 6, and 10 features. To reduce the dimension of the features, the covariance matrix is calculated to obtain the matrix for projection using equation 2 [22]:

$$\sum = \frac{1}{m}\sum_{i=1}^{m}\left(A^{(i)}\right)\left(A^{(i)}\right)^T \quad (2)$$

Three different components were used to evaluate the proposed model. When developing the model with three features from the dataset, 75% were retained, while 89% were kept from the original data when the features were reduced to six. When the features were reduced to10, 96% of the data were retained from the original dataset. Furthermore, categorical features were encoded into discrete features via the 1-to-n encoding method, and the class labels were dropped before clustering was performed.

4) Data clustering: Clustering algorithms search for groups of similar data vectors in a dataset. This unsupervised approach does not require labelled data to ascertain which class or cluster data inputs should be assigned. It is also a non-parametric technique requiring no prior knowledge of data parameters [31]. The K-means algorithm [32] is a clustering algorithm based on the similarity measure between data inputs. In our hybrid model, the algorithm was employed to accept both random observations N and a parameter showing the number of clusters (i.e., their centroids) $C_i \leq i \leq k$. An observation is assigned to a cluster in each iteration using the shortest distance between the observation and the centroids. The algorithm reassigns the centroids by reducing the mean distance of all observations in the cluster to its centroids after each iteration. The algorithm converges when the position of the centroids no longer changes. The aim is to find a set of k cluster centres, represented as $\{C_1, \dots, C_k\}$ such that there is minimization in the distance between data points and their nearest centre. Assigning data points to a cluster centre requires a set of binary variables $\lambda_{nk} \in \{0,1\}$, such that if cluster centre C_k contains data point a_n, then $\lambda_{nk} = 1$ as captured in the algorithm in Table IV. Two different experiments were conducted using the K-means algorithm. The first involved generating two clusters (k=2), representing normal and malicious. The second generated four clusters (k=4), representing normal data and the different attack types in the NSL-KDD dataset (Normal, DoS, Probing, U2R, and R2L). Meanwhile, for the UNSW-NB15 dataset, only two clusters are generated (i.e., normal and malicious).

TABLE IV. THE ALGORITHM FOR THE PROPOSED MODEL

Algorithm: Dimension Reduction and Data Clustering
Inputs: Unlabelled dataset $\{a_1, \dots, a_N\}$; Number of clusters N = Number of samples X, Y
Output: Principal components , cluster centres $\{C_k\}$ and assigned data points $\{\lambda_{nk}\}$

Set $P_k = \{3,6,10\}$ *(Reduction)*
Initialize
$C_k = \{2,4\}$ (Number of clusters)
for $n = 1$ to N **do**
For $K=1$ to K **do**
$if k = (C_i - a_i, \dots, C_n - a_n)$ then
$\lambda_{nk} = 0$
Else
$\lambda_{nk} = 1$
end if
end for
end for
For $n=1$ to K do
$\lambda_{nk} = 4$
end for
$\lambda_{nk} \mid C_k$ converges

5) Anomaly classification: The proposed model uses the Support Vector Machine algorithm for anomaly classification. The SVM is a supervised learning model used for data classification, regression, and outlier detection. SVM, which is most suitable for non-linear data used in this paper, can be represented formally in equation 3 [33].

$$\Theta = \sum_{i=1}^{m} \alpha_i c_i x_i \quad (3)$$

Where x is the given input, c is the Class label, α is the LaGrange multiplier, and $\Theta$is the weight vector.

In this paper, the class labels used by the SVM classifier are cluster labels generated from the K-means algorithm. The classification task incorporates both binary and multi-class classification. The binary classification trains the classifier to predict unseen data from IoT network traffic as either normal or malicious. Meanwhile, the multi-class classification implements a more detailed classification, where the classifier was trained to predict unseen data into the normal, DoS, Probing, U2R.R2L classes. The U2R.R2L class is a merged class due to its low occurrence as captured in the NSL-KDD dataset. For the UNSW-NB15 dataset, only binary classification is performed.

IV. RESULTS AND DISCUSSION

The first results obtained are from the data clustering task. Fig. 2 shows the normal and malicious clusters when k is set to 2. Fig. 3, on the other hand, displays the four different clusters when k is set to four. These clusters illustrate the similarity of data points in the same group (normal traffic data) from the malicious cluster.

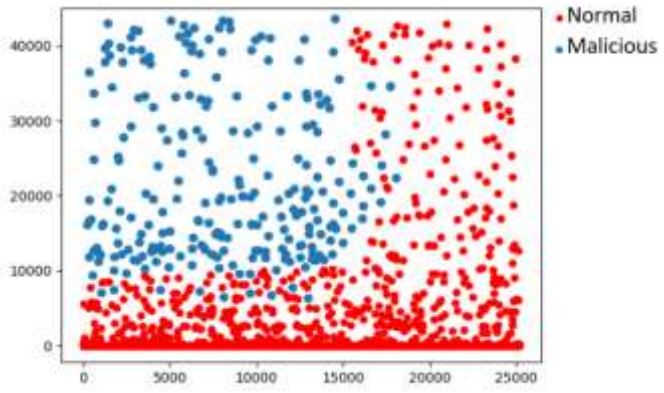

Fig. 2. Data Clusters when k=2.

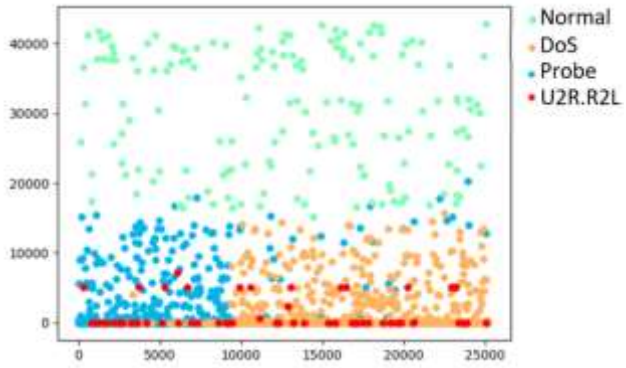

Fig. 3. Data Clusters when k=4.

As stated earlier in this paper, the generated clusters from the first phase of the detection model are used to train the SVM classifier. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are performance indicators used to evaluate the proposed anomaly detection model to ascertain its accuracy, precision, and recall, as shown in equation 4 to 6, respectively). TP shows that normal behaviours are classified correctly as normal behaviours; TN shows that malicious activities are classified correctly as malicious. FP demonstrates that malicious activities are incorrectly classified as normal behaviours, while FN shows that normal behaviours are incorrectly classified as malicious activities. In addition to the above performance metrics, the Detection Rate (DR) of the classifier in identifying malicious activities was also evaluated using equation 7. False Alarm Rate (FAR) (incorrectly detecting normal behaviour as malicious activities) was also examined using equation 8. The classification summaries for the NSL-KDD and the UNSW-NB15 datasets are presented in Table V.

$$Accuracy = \frac{TN+TP}{FN+FP+FN+TP} \quad (4)$$

$$Precision = \frac{TsP}{FP+TP} \quad (5)$$

$$Recall = \frac{TP}{FN+TP} \quad (6)$$

$$DR = \frac{TP}{(FN+TP)} \quad (7)$$

$$FAR = \frac{FP}{(FP+TN)} \quad (8)$$

When two clusters were used as class labels for the SVM classifier, accuracy scores of 97.82%, 97.58%, and 97.01% were obtained for the 3, 6, and 10 features NSL-KDD dataset as depicted in Table VI. There was no significant difference in the accuracy scores recorded across the different number of features. However, DR was remarkably higher with three features than with six features. Nevertheless, FAR was significantly lower with six features with 0.95% (less than one per cent) against 2.81% observed with three features. In this experiment, data were classified either as normal or malicious. This result proves that high-dimension features do not necessarily equal high accuracy and detection rate in datasets used for the experiment.

Furthermore, with four clusters employed as class labels (normal, DoS, Probing, U2R.R2L) for the SVM classifier, accuracy scores of 93.96%, 95.03%, and 91.79% were recorded for features reduced to 3, 6, and 10, respectively. These accuracy scores are lower compared to those observed with two class labels. The results show that the model performs better when predicting data into a binary class. However, reasonably high detection rates were recorded when detecting data as normal, DoS, Probing, U2R.R2L. The performance of the model based on accuracy, precision, recall, DR, and FAR when trained with two and four clusters is presented in Fig. 4.

TABLE V. CLASSIFICATION DISTRIBUTION FROM SVM CLASSIFIER

NSL-KDD Dataset					
3 Features		**6 Features**		**10 Features**	
TN=1730	FP=47	TN=5520	FP=50	TN=5525	FP=53
FN=87	TP=4434	FN=99	TP=629	FN=94	TP=626
UNSW-NB15 Dataset					
3 Features		**6 Features**		**10 Features**	
TN=19844	FP=2	TN=19840	FP=6	TN=23941	FP=3
FN=5	TP=23995	FN=3	TP=23997	FN=4	TP=19888

TABLE VI. THE PERFORMANCE OF SVM CLASSIFIER WITH DIFFERENT NUMBER OF CLASSES AND FEATURES ON THE NSL-KDD DATASET

K=2				**K=4**		
Metrics	**3 Features**	**6 Features**	**10 Features**	**3 Features**	**6 Features**	**10 Features**
Variance	75%	89%	96%	75%	89%	96%
Accuracy	97.82%	97.58%	97.01%	93.96%	95.03%	91.79%
Precision	98.88%	92.19%	94.55%	92.77%	94.09%	87.82%
Recall	98.07%	86.34%	90.76%	91.07%	93.30%	83.95%
Detection Rate	98.07%	86.34%	90.76%	96.13%	97.54%	94.21%
False Alarm Rate	2.81%	0.95%	1.36%	3.24%	2.93%	3.79%

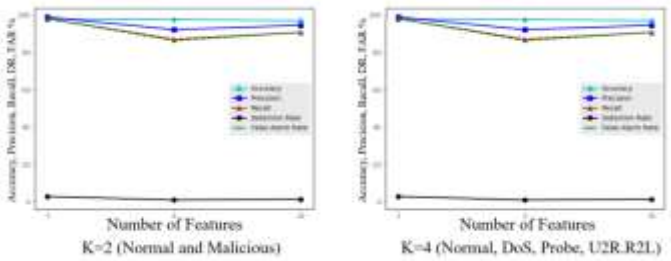

Fig. 4. Performance of Proposed Model on Two different Clusters.

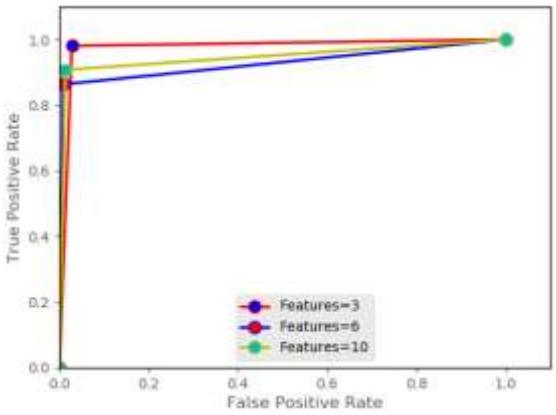

Fig. 5. Data Clusters when k=4.

Fig. 5 shows a ROC curve for the three experiments, where features were reduced to 3, 6, and 10. The trained classifier obtained from the cluster labels was applied to the NSL-KDD dataset, which contained different features. We then analysed how accurately the model detects anomalies from normal traffic data because the initial process was achieved using unsupervised learning.

TABLE VII. THE PERFORMANCE OF THE SVM CLASSIFIER ON THE UNSW-NB15 DATASET

UNSW-NB15 (K=2)			
Metrics	**3 Features**	**6 Features**	**10 Features**
Variance	**75%**	**89%**	**96%**
Accuracy	99.95%	99.98%	99.99%
Precision	99.93%	99.97%	99.98%
Recall	99.92%	99.95%	99.97%
Detection Rate	99.98%	99.98%	99.99%
False Alarm Rate	0.4%	0.5%	0.42%

Similarly, Table VII presents accuracy results from the evaluation of the proposed model on the UNSW-NB15 dataset. An accuracy of 99% was obtained when the model was tested using 3, 6, and 10 features. These results show the effectiveness of the proposed model in detecting malicious activities in recent datasets.

Apart from evaluating the classification accuracy, precision, DR, and FAR of the proposed intrusion detection model, we also identified features selected by the PCA algorithm after feature dimension reduction. These features were chosen from the 41 available features in the NSL-KDD dataset. Table VIII shows the most important features after dimension reduction to 3, 6, and 10. The features are presented in descending order of importance, and the top three features are DST_HOST_SRV_SERROR_RATE, SRV_RERROR_ RATE, and DST_HOST_SAME_SRC_PORT_RATE.

On the other hand, Table IX presents the most relevant features in the INSW-NB15 dataset after dimension reduction using the proposed model. The model also captures the associated weight of each feature. To accurately compare results obtained from our model with an earlier work presented in Zhao, Li [22], we adopted the same number of dimensions after dimension reduction (i.e., best 3, 6, and 10 dimensions of the singular vector). With a variance of 75%, dimensions were reduced to 3, 6 dimensions were obtained with a variance of 89%, while a variance that retained 96% of the data produced ten dimensions from the available 41 features. The two experiments conducted in this paper are based on the reduced features and are used to generate clusters (i.e., k-2 and k=4), which served as cluster labels for the classifier. A comparison of the results presented in Zhao, Li [22] shows that our proposed model performs better accuracy using 3 and 6 features, as demonstrated in Fig. 6.

TABLE VIII. THE MOST RELEVANT FEATURES IN THE NSL-KDD DATASET

PCs =3		
S/N	**Features**	**Weights**
1	DST_HOST_SRV_SERROR_RATE	0.508
2	SRV_RERROR_RATE	0.212
3	DST_HOST_SAME_SRC_PORT_RATE	0.068
PCs =6		
S/N	**Features**	**Weights**
1	DST_HOST_SRV_SERROR_RATE	0.508
2	SRV_RERROR_RATE	0.212
3	DST_HOST_SAME_SRC_PORT_RATE	0.068
4	DST_HOST_COUNT	0.052
5	DST_HOST_SAME_SRV_RATE	0.043
6	SRV_DIFF_HOST_RATE	0.021
PCs =10		
S/N	**Features**	**Weights**
1	DST_HOST_SRV_SERROR_RATE	0.508
2	SRV_RERROR_RATE	0.212
3	DST_HOST_SAME_SRC_PORT_RATE	0.068
4	DST_HOST_RERROR_RATE	0.052
5	IS_GUEST LOGIN	0.043
6	IS_HOST_LOGIN	0.021
7	DST_HOST_SRV_DIFF_HOST_RATE	0.018
8	DST_HOST_SRV_COUNT	0.017
9	WRONG_FRAGMENT	0.012
10	DST_HOST_SAME_SRV_RATE	0.010

TABLE IX. THE MOST RELEVANT FEATURES IN THE UNSW-NB15 DATASET

PCs =3		
S/N	**Features**	**Weights**
1	*dwin*	0.588
2	*sttl*	0.146
3	*ct_srv_dst*	0.078
PCs =6		
S/N	**Features**	**Weights**
1	*dwin*	0.588
2	*sttl*	0.146
3	*ct_srv_dst*	0.078
4	*dttl*	0.034
5	*stcpb*	0.024
6	*dtcpb*	0.023
PCs =10		
S/N	**Features**	**Weights**
1	*dwin*	0.588
2	*sttl*	0.146
3	*ct_srv_dst*	0.078
4	*dttl*	0.034
5	*stcpb*	0.024
6	*dtcpb*	0.023
7	*dmeanz*	0.019
8	*ct_srv_src*	0.014
9	*smeanz*	0.012
10	*swin*	0.011

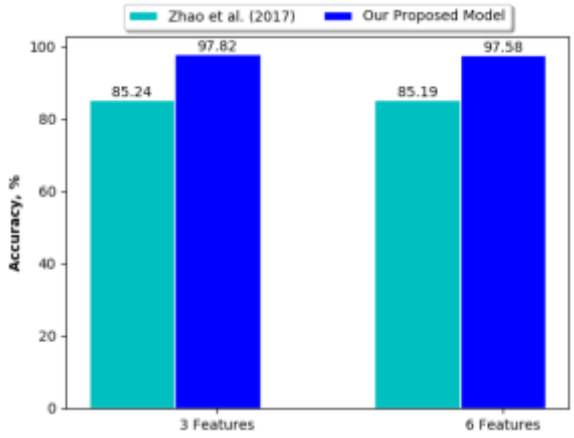

Fig. 6. Performance Comparison of the Proposed Model.

Cybersecurity has become an important research area in the Internet of Things, especially with the vast amount of sensitive data stored and transmitted by IoT devices. IoT devices have several security threats such as eavesdropping, data leakage/loss, denial-of-service attacks, etc. In tackling these issues, this paper presented a hybridized machine model that detects several anomalies. The proposed hybrid detection model detects anomalies in two most common communication models in IoT devices (i.e., direct and gateway-based communication models). One of the proposed model features is learning and detecting malicious patterns in IoT traffic data. Such functionality involves learning the benign and detecting the anomalies that do not conform to the normal patterns.

The model presented in this paper detects threats in the network layer of the IoT. The need for such a detection model in this layer of the IoT cannot be overemphasized since the network layer is most vulnerable to attacks due to the large amount of data it transmits. The proposed model accurately detects the denial of Service (DoS) attack in the IoT network layer with a low false alarm rate. Another threat in the IoT network layer detected by the model proposed in this paper is the routing attack (Probing attack). Such attacks are used to scan the network for possible vulnerabilities. Attacks used to escalate privileges (such as U2R and R2L attacks) come under this category. The model can detect normal and malicious behaviours and identify four different attack types in the IoT (using its multi-classification feature).

The uniqueness of the proposed hybrid intrusion detection model is in its ability to be trained with unlabelled data. The model ensures a quality experience for users and security experts as manual data identification and labelling are not needed. This attribute is required in detection models in IoT networks since the acquisition of labels in big data from IoT devices can be time-consuming and laborious. Furthermore, the high accuracy score of the model guarantees that malicious data (threats) in IoT traffic can be detected, thereby reducing zero-day exploits in IoT networks. The dimension reduction performed on the features ensures the low complexity of the model desired when dealing with IoT devices with limited resources such as memory and processing power. The model, when accurately deployed, can alert security experts to initiate preventive measures from the identified threats. Providing prior warnings aids administrators, stakeholders in IoT and minimizes exploitable vulnerabilities. Consequently, the security of sensitive data is enhanced, which preserves the privacy of IoT users.

V. CONCLUSION

This paper proposed a hybrid model for the detection of anomalies in the network layer of the IoT. The proposed system performs dimension reduction (using PCA algorithm), data clustering (using K-means algorithm), and a data classification based on the Support Vector Machine (SVM) algorithm. The proposed hybrid model was evaluated on both the NSL-KDD and the UNSW-NB15 datasets. Performance evaluation of the proposed model shows that dimension reduction improves the detection rate of attacks since irrelevant features that increase noise are removed from the new dataset (with reduced features). The conducted experiments also revealed that classification accuracy is higher with binary classification than with multi-class, mainly when classes are generated from cluster labels (i.e., unsupervised learning). Also, the classifier was benchmarked with the classifier presented by Zhao, Li [22]. Our proposed model outperforms the model shown by Zhao, Li [22] in terms of detection rate and accuracy. As future work, we will employ the proposed

hybrid anomaly detection model to detect different categories of IoT attacks that are not covered in this paper (i.e., from other datasets that simulate various attack activities).

REFERENCES

[1] Zhang, Z.-K., M.C.Y. Cho, and S. Shieh. Emerging security threats and countermeasures in IoT. in Proceedings of the 10th ACM symposium on information, computer and communications security. 2015. ACM.

[2] Singh, S. and N. Singh. Internet of Things (IoT): Security challenges, business opportunities & reference architecture for E-commerce. in 2015 International Conference on Green Computing and Internet of Things (ICGCIoT). 2015. IEEE.

[3] Gartner, Gartner Says 6.4 Billion Connected. (2015). Retrieved September 14, 2017 from http://www.gartner.com/newsroom/id/ 3165317. 215.

[4] Baig, Z.A., et al., Future challenges for smart cities: Cyber-security and digital forensics. Digital Investigation, 2017. 22: p. 3-13.

[5] Sheikhan, M. and H. Bostani. A hybrid intrusion detection architecture for internet of things. in 2016 8th International Symposium on Telecommunications (IST). 2016. IEEE.

[6] Desai, A.S. and D. Gaikwad. Real time hybrid intrusion detection system using signature matching algorithm and fuzzy-GA. in 2016 IEEE international conference on advances in electronics, communication and computer technology (ICAECCT). 2016. IEEE.

[7] Sedjelmaci, H., S.M. Senouci, and M. Al-Bahri. A lightweight anomaly detection technique for low-resource IoT devices: A game-theoretic methodology. in 2016 IEEE International Conference on Communications (ICC). 2016. IEEE.

[8] Erfani, S.M., et al., High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recognition, 2016. 58: p. 121-134.

[9] Cherian, M. and M. Chatterjee. Survey of Security Threats in IoT and Emerging Countermeasures. in International Symposium on Security in Computing and Communication. 2018. Springer.

[10] Adat, V. and B. Gupta, Security in Internet of Things: issues, challenges, taxonomy, and architecture. Telecommunication Systems, 2018. 67(3): p. 423-441.

[11] Lohachab, A. and B. Karambir, Critical Analysis of DDoS—An Emerging Security Threat over IoT Networks. Journal of Communications and Information Networks, 2018. 3(3): p. 57-78.

[12] Khan, R., et al. Future internet: the internet of things architecture, possible applications and key challenges. in 2012 10th international conference on frontiers of information technology. 2012. IEEE.

[13] Tweneboah-Koduah, S., K.E. Skouby, and R. Tadayoni, Cyber security threats to IoT applications and service domains. Wireless Personal Communications, 2017. 95(1): p. 169-185.

[14] Choraś, M., R. Kozik, and I. Maciejewska, Emerging cyber security: Bio-inspired techniques and MITM detection in IoT, in Combatting Cybercrime and Cyberterrorism. 2016, Springer. p. 193-207.

[15] Sapienza, A., et al. Discover: Mining online chatter for emerging cyber threats. in Companion Proceedings of the The Web Conference 2018. 2018. International World Wide Web Conferences Steering Committee.

[16] Harel, Y., I.B. Gal, and Y. Elovici, Cyber security and the role of intelligent systems in addressing its challenges. ACM Transactions on Intelligent Systems and Technology (TIST), 2017. 8(4): p. 49.

[17] Berral-García, J.L. A quick view on current techniques and machine learning algorithms for big data analytics. in 2016 18th international conference on transparent optical networks (ICTON). 2016. IEEE.

[18] Shanthamallu, U.S., et al. A brief survey of machine learning methods and their sensor and IoT applications. in 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA). 2017. IEEE.

[19] Li, Y., R. Ma, and R. Jiao, A hybrid malicious code detection method based on deep learning. International Journal of Security and Its Applications, 2015. 9(5): p. 205-216.

[20] Nskh, P., M.N. Varma, and R.R. Naik. Principle component analysis based intrusion detection system using support vector machine. in 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). 2016. IEEE.

[21] Pajouh, H.H., et al., A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks. IEEE Transactions on Emerging Topics in Computing, 2016.

[22] Zhao, S., et al. A dimension reduction model and classifier for anomaly-based intrusion detection in internet of things. in 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech). 2017. IEEE.

[23] Narudin, F.A., et al., Evaluation of machine learning classifiers for mobile malware detection. Soft Computing, 2016. 20(1): p. 343-357.

[24] Nobakht, M., V. Sivaraman, and R. Boreli. A host-based intrusion detection and mitigation framework for smart home IoT using OpenFlow. in 2016 11th International conference on availability, reliability and security (ARES). 2016. IEEE.

[25] Doshi, R., N. Apthorpe, and N. Feamster. Machine learning ddos detection for consumer internet of things devices. in 2018 IEEE Security and Privacy Workshops (SPW). 2018. IEEE.

[26] McKinney, W., Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. 2012: " O'Reilly Media, Inc.".

[27] Hackeling, G., Mastering Machine Learning with scikit-learn. 2017: Packt Publishing Ltd.

[28] Tavallaee, M., et al. A detailed analysis of the KDD CUP 99 data set. in 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications. 2009. IEEE.

[29] Rehim, R., Python Penetration Testing Cookbook: Practical recipes on implementing information gathering, network security, intrusion detection, and post-exploitation. 2017: Packt Publishing Ltd.

[30] Moustafa, N. and J. Slay. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). in 2015 military communications and information systems conference (MilCIS). 2015. IEEE.

[31] Zheng, Y., et al. Smart car parking: temporal clustering and anomaly detection in urban car parking. in 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP). 2014. IEEE.

[32] MacQueen, J. Some methods for classification and analysis of multivariate observations. in Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967. Oakland, CA, USA.

[33] Hasan, M., et al., Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet of Things, 2019. 7: p. 100059.