A Two-layer Dimension Reduction and Two-tier Classification Model for Anomaly-Based Intrusion Detection in IoT Backbone Networks

Hamed HaddadPajouh, Reza Javidan, Raouf Khayami, Ali Dehghantanha and Kim-Kwang Raymond Choo

***Abstract*—With increasing reliance on Internet of Things (IoT) devices and services, the capability to detect intrusions and malicious activities within IoT networks is critical for resilience of the network infrastructure. In this paper, we present a novel model for intrusion detection based on two-layer dimension reduction and two-tier classification module, designed to detect malicious activities such as User to Root (U2R) and Remote to Local (R2L) attacks. The proposed model is using component analysis and linear discriminate analysis of dimension reduction module to spate the high dimensional dataset to a lower one with lesser features. We then apply a two-tier classification module utilizing Naïve Bayes and Certainty Factor version of K-Nearest Neighbor to identify suspicious behaviors. The experiment results using NSL-KDD dataset shows that our model outperforms previous models designed to detect U2R and R2L attacks.**

***Index Terms*— Anomaly Detection, CF-KNN, Intrusion Detection System, IoT, Multi-layer Classification**

I. Introduction

Internet of Things (IoT) technologies are becoming increasingly prevalent across different industry sectors such as health care, personal and social domains, and smart cities [1]. Similar to most consumer technologies, IoT technologies are not designed with security in mind, which are now emerging as a key barrier in the wider adoption of IoT networks and services [2]. Intrusion detection is one of several security mechanisms for managing security intrusions [3], which can be detected in any of four layers of IoT architecture shown in Figure 1 [4]. The Network layer not only serves as a backbone for connecting different IoT devices, but also provides opportunities for deploying network-based security defense mechanisms such as Network Intrusion Detection Systems (NIDS) [5],[6],[7]. According to the analysis of KDD99 [3] and its latter version NSL-KDD [9], malicious behaviors (attacks) in network-based intrusions can be classified into the following four main categories [7]:

- Probe: when an attacker seeks to only gain information about the target network through network and host scanning activities (i.e. ports scanning).
- DoS (denial of service): when an attacker interrupts legitimate users' access to the given service or machine.
- U2R (User to Root): when an attacker attempts to escalate a limited user' privilege to a super user or root access (e.g. via malware infection or stolen credentials).
- R2L (Remote to Local): when an attacker gains remote access to a victim machine imitating existing local users.

User to Root (U2R) and Remote to Local (R2L) attacks are among the most challenging attacks to detect as they mimick normal users behavior [10] [11].

IDS are categorized into signature-based and anomaly-based detection based on their technique in detecting an intrusion [12]. Signature-based IDS relies on a set of pre-defined malicious activates patterns and attack signatures to detect intrusions while anomaly-based IDS relies on deviations from normal behaviors to detect intrusions [6]. Signature-based IDSes generally outperform anomaly-based IDSes in detecting previously known attacks, but the former is ineffective against unknown or polymorphic attacks [13]. On the other hand, anomaly-based IDSes are capable of detecting unknown attacks in the absence of a predefined pattern. Due to the diversity of devices deployed in IoT networks, it would be unrealistic and impractical to rely on pre-defined attack patterns for intrusion detection, which limits signature-based IDS utilization in IoT networks [14].

In this paper, we present a network anomaly-based model for intrusion detection, hereafter referred to as Two-layer Dimension Reduction and Two-tier Classification (TDTC) model.

H. H. Pajouh, R. Javidan, and R. Khatami are with the Department of Computer Engineering and Information, Technology, Shiraz University of Technology, Iran (e-mail: hp@sutech.ac.ir, reza.javidan@sutech.ac.ir, khayami@sutech.ac.ir).

A. Dehghantanha is with the School of Computing, Science and Engineering, University of Salford, UK. (e-mail: A.Dehghantanha@salford.ac.uk).

K-K R Choo, is with the Department of Information Systems and Cyber Security, University of Texas at San Antonio, USA (e-mail: raymond.choo@fulbrightmail.org).

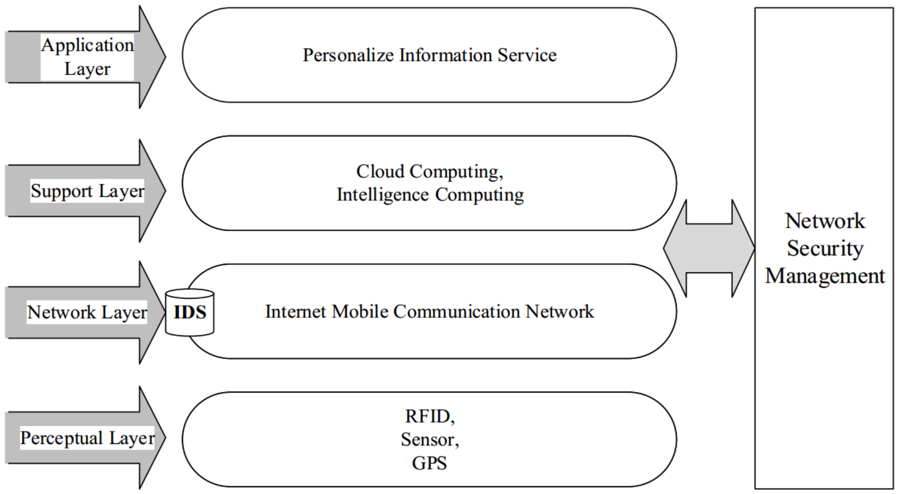

Fig 1. IoT Network Security Architecture [4]

The proposed model, designed for anomaly-based intrusion detection in IoT backbone networks, uses two-layer dimension reduction and two-tier classification detection techniques to detect "hard-to-detect" intrusions, such as U2R and R2L attacks. We also demonstrate that the proposed model has the following characteristics:

- Higher overall detection rates due to the deployment of a multi-layer classifier
- Lower false positive due to deployment of a refinement feature
- Accurate detection of U2R and R2L attacks, without reducing performance
- Lower computational complexity due to deployment of dimension reduction in the two layers.

In the next section, we present related work. The proposed model is presented in Section 3, and evaluation of the model is presented in Section 4. Section 5 concludes this paper and outlines future research topics.

II. Related work

Existing intrusion detection and prevention models generally use statistical approaches [15] such as Hidden Markov Model (HMM) [15], Bayes theory [16], cluster analysis [17], signal processing [18] and distance measuring [19] to detect anomalous activities. Anomaly detection approaches can be broadly categorized into supervised and unsupervised learning [6]. In supervised anomaly detection approach, normal behavior of a system or networks is constructed using a labeled dataset [20]. Unsupervised technique assumes that normal behaviors are more frequent and, thus, the model is built based on this assumption; thus, no training data is required [21].

Casas et al. [22] proposed an unsupervised NIDS based on subspace clustering and outlier detection and demonstrated that their approach performs well against unknown attacks. In [23], a feature section filter module is proposed, which utilizes Principal Component Analysis and Fisher Dimension Reduction to filter noises. In the approach, Self-Organizing Maps (SOMs) neural model is also used to filter out normal activities. However, this approach has a high false positive rate. Bostani and Sheikhan [24] proposed an unsupervised framework based on Optimum-path forest algorithm and K-Means clustering technique. This framework models malicious and normal behavior of networks.

The supervised anomaly detection approach in [25] leverages both distance measure and density of clusters for intrusion detection. Zhaung et al [26] proposed a model based on random forest algorithm to discover anomaly patterns with a high accuracy yet low false negative rate.

Guo et al. [27] proposed a two-level intrusion detection approach which first detects misuse and then uses KNN algorithm to reduce false alarms. Toosi et al. [28] proposed a multi attack classifier model, which implements a mix of fuzzy neural network, fuzzy inference approach, and genetic algorithms for intrusion detection. Despite a high accuracy rate in identifying normal behaviors and detecting simpler attacks such as DoS attacks and probe, the model performs poorly in detecting low frequency and distribution attacks such as R2L. Horng et al [29] proposed a multi-classification attack model consisting of support vector machines (SVM) and BRICH hierarchical clustering technique to extract significant attributes from KDD99 dataset. Their proposed model has a high detection rate for DoS and Porbe attacks, but is ineffective against U2R and R2L attacks.

Tan et al. [30] proposed a system for DoS detection using multivariate correlation analysis (MCA) to improve the accuracy of network traffic characterization. In [31], a two-layer classification module was used to detect U2R and R2L attacks with low computational complexity due to its optimized feature reduction. Osanaiye et al. [13] proposed an ensemble-based multi-filter feature selection method to detect distributed DoS attacks in cloud environments using four filter methods to achieve an optimum selection over NSL-KDD dataset. Iqbal et al. [32] presented an attack taxonomy for cloud services and suggested a cloud-based intrusion detection system.

Ambusaidi et al. in [33] proposed a mutual information based IDS that selects optimal feature for classification based on feature selection algorithm. Their approach was evaluated using three benchmark data set (KDD Cup 99, NSL-KDD and Kyoto 2006+).

Intrusion detection systems have also been used for managing security risks in industrial control systems [14]. For example, Pan et al. [34] proposed a systematic and automated approach to build a hybrid IDS that learns temporal state-based specifications for electric power systems to accurately differentiate between disturbances, normal control operations, and cyber-attacks. Zhou et al. [35] presented an industrial anomaly and multi model driven IDS based on Hidden Markov Model to filter attacks from actual faults.

Security issues can be a barrier to widespread adoption of IoT devices [36]. Whitmore et al., [37] showed that wide range of techniques could mitigate cyber threat targeting IoT systems. Ning et al. [38] proposed a hierarchical authentication architecture to provide anonymous data transmission in IoT networks. Cao et al. [39] highlighted the impact and importance of ghost attacks on ZigBee based IoT devices. Chen et al. [40] proposed an autonomic model-driven cyber security management approach for IoT systems, which can be used to estimate, detect, and respond to cyberattacks with little or no human intervention. Teixeira et al. [41] proposed a scheme for thwarting insiders attacks in IoT networks by crosschecking data transformation of every IoT node.

III. Proposed TDTC Model

The proposed model comprises a dimension reduction module and a classification module, to be discussed in sections III.A and III.B, respectively.

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \xrightarrow{\text{linear feature extraction}} \begin{bmatrix} y_1 \\ y_2 \\ \\ y_M \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & & w_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

Fig 2. In PCA, linear transformation is used to reduce high dimension dataset to a low dimension dataset

A. Dimension Reduction Module

The dimension reduction module is deployed to address limitations due to dimensionality that may lead to making wrong decisions while increasing computational complexity of the classifier. We deployed both Linear Discriminant Analysis (LDA) (i.e. a supervised dimension reduction technique) and Principal Component Analysis (i.e. an unsupervised dimension reduction technique) in order to address the high dimensionality issue. Principal Component Analysis (PCA) can be used to perform feature selection and extraction [42]:

a) Feature selection: choose a subset of all features based on their effectiveness in higher classification (i.e. choosing more informative features)
b) Feature extraction**:** create a subset of new features by combining existing features.

In TDTC, we used PCA as a feature extraction mechanism to map the NSL-KDD dataset, which consists of 41 features to one with a lower feature space by removing less significant features. Feature extraction technique is commonly limited to linear transforms**:** $y = Wx$ as shown in in Figure 2.

Let **X** be an N-dimensional random vector in the original dataset, and the new feature space consists of lower M-dimensions (M is the number of new dataset features that are transformed) where $(M < N)$. For the transformation operation, we will need to compute Eq. 1 to Eq.3:

Covariance matrix:

$$\Sigma_x = \sum_{k=1}^{n} (x_k - m)(x_k - m)^T, \qquad \text{(Eq.1)}$$

Where m (mean vector) is:

$$m = \frac{1}{n} \sum_{k=1}^{n} x_k \qquad \text{(Eq.2)}$$

Eigenvector-eigenvalue decomposition:

$\Sigma v = \lambda v$ Where v=Eigenvector λ=Eigenvalue (Eq.3)

PCA will then sort the eigenvectors in descending order. In other words, eigenvectors with lower eigenvalues have the least information about the distribution of the data and these are the eigenvectors we wish to drop. A common approach is to rank the eigenvectors from the highest to the lowest eigenvalue and choose the top $\boldsymbol{k}$ eigenvectors based on eigenvalues. Similarly, in TDTC, one may decide which eigenvalues are more useful; thus, the ideal feature mapping matrix W can be concluded and used for linear transformation of training and test dataset.

At this layer of dimension reduction, Imbedded Error Function (IEF) factor analysis measure [43] is used to select the principal [44] as shown in Eq.4, where *l*, *m* denotes the number of Principal Components (PCs). Both *l* and *m* are used to represent the data and number of dimension, respectively. N and λ denote the number of samples and Eigenvalues, respectively.

$$\text{IEF(l)} = \left[\frac{l \sum_{j=l+1}^{m} \lambda}{Nm(m-l)}\right]^{1/2} \qquad \text{(Eq.4)}$$

Cross Validation (CV) is used to evaluate optimum principals with minimum errors as shown in Figure 3. Applying selection criteria would reduce some features and help the next layer of dimension reduction module to compute lower dimension matrix and spreadable objects.

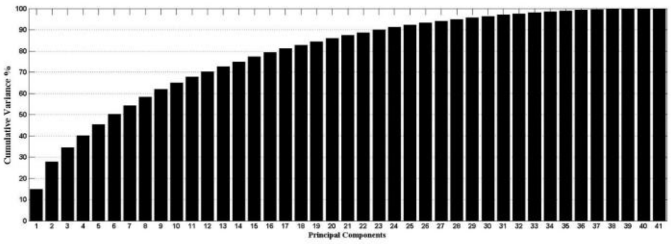

Fig 5. Imbedded Error Function measure of NSL-KDD train data set to select optimum number of dimension with minimum error and information loss.

As observed in Figure 5, Cumulative Percent Variance (CPV) measure with 95% threshold is also examined to justify the selection of optimum dimensions.

$$\mathrm{CPV}(l) = 100\left[\frac{\sum_{j=1}^{l}\lambda_j}{\sum_{j=1}^{m}\lambda_j}\right]\% \qquad \text{(Eq.5)}$$

B. Linear Discriminant Analysis

Linear computation can be used to achieve a reasonable speed in intrusion detection systems [31].

Since objects (samples) in the PCA-transformed dataset are not ideal for classification, the proposed model utilized another feature reduction module to apply the labeled data in an optimal transformation to new dimensions. LDA examines the class labels to reduce the dimension of large working datasets and LDA is widely used in different domains such as image processing and stock analysis [45]. LDA chooses an After the transformation using LDA, the new mapped features will have only four dimensions {lda1, ..., lda4}.

Figure 4 shows the two-dimension of the newly mapped original data set transformed by LDA. In other words, the dataset has been converted into $C-1$ dimensions, where C is number of class labels that exist in the original dataset. optimal projection matrix to map a higher dimensional feature space to a new lower dimensional space while preserving the required information for data classification [46].

There are two scatter matrices that need to be obtained in LDA, namely: S_B which is the between-class scatter matrix, and S_W the within-class scatter matrix. In TDTC, the LDA dimension reduction module transforms the NSL-KDD dataset to a lower dimension. It is assumed that there is a set of *n* d-dimensional vectors of x_i, ..., x_n belonging to *k* different class labels of C_i, where each i = 1, 2, 3,...,k has n_i samples (in TDTC k = 5 e.g. normal, DoS, Probe, U2R, L2R).

The projection matrix W is calculated to maximize $S_{\mathbf{B}}$ – see Eq. 6, and minimize S_W – see Eq. 7.

$$S_B = \sum_c (\mu_c - \bar{x})(\mu_c - \bar{x})^T \qquad \text{(Eq.6)}$$

$$S_W = \sum_c \sum_{i \in c} (x_i - \mu_c)(x_i - \mu_c)^T \qquad \text{(Eq.7)}$$

μ_c is the mean value of class C_i samples, and is given by Eq.8.

$$\mu_c = \frac{1}{n_i} \sum_{x \in C_i} x \qquad \text{(Eq.8)}$$

Since the ratio J in Eq.9 is within the range of S_B and S_W, it can be easily maximized as an optimization problem using the projection matrix W_r (see Eq.9).

$$J = \frac{W_r^T S_B\, W_r}{W_r^T S_W\, W_r} \qquad \text{(Eq.9)}$$

All these operations will be conducted on the training dataset (see Section IV) to obtain an ideal transformation matrix that can be applied to future test sets or unknown instances.

Table 1. Transformed Features Dependency Of Train+ Data Set After Applying Two Level Of Reduction Due To Correlation Coefficient Measure.

features	LDA1	LDA2	LDA3	LDA4
LDA1	1	-3.76E-17	4.73E-16	1.06E-16
LDA2	-3.76E-17	1	-6.69E-17	-3.52E-16
LDA3	4.73E-16	-6.69E-17	1	-1.65E-15
LDA4	1.06E-16	-3.52E-16	-1.65E-15	1

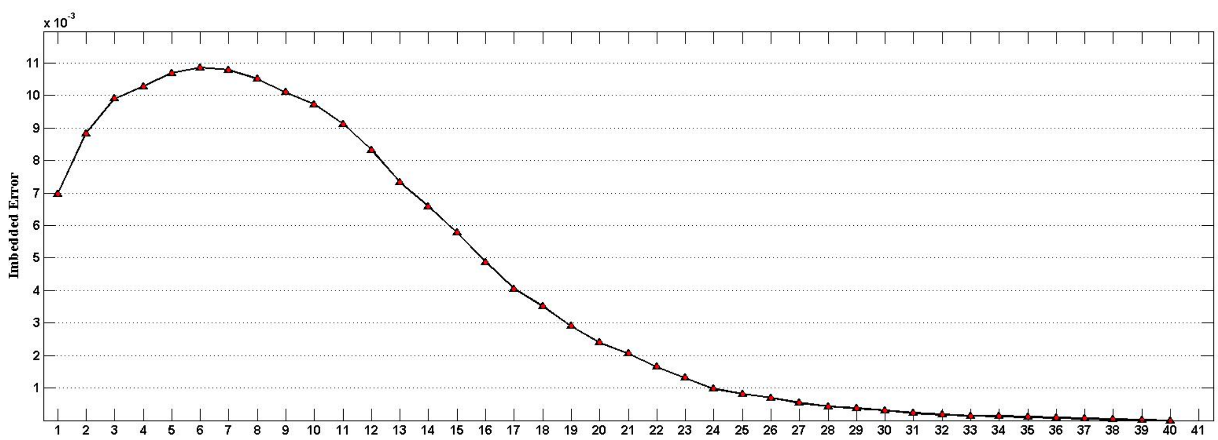

Fig 3. Imbedded Error Function measure of NSL-KDD train data set to select optimum number of dimension with minimum error and information loss.

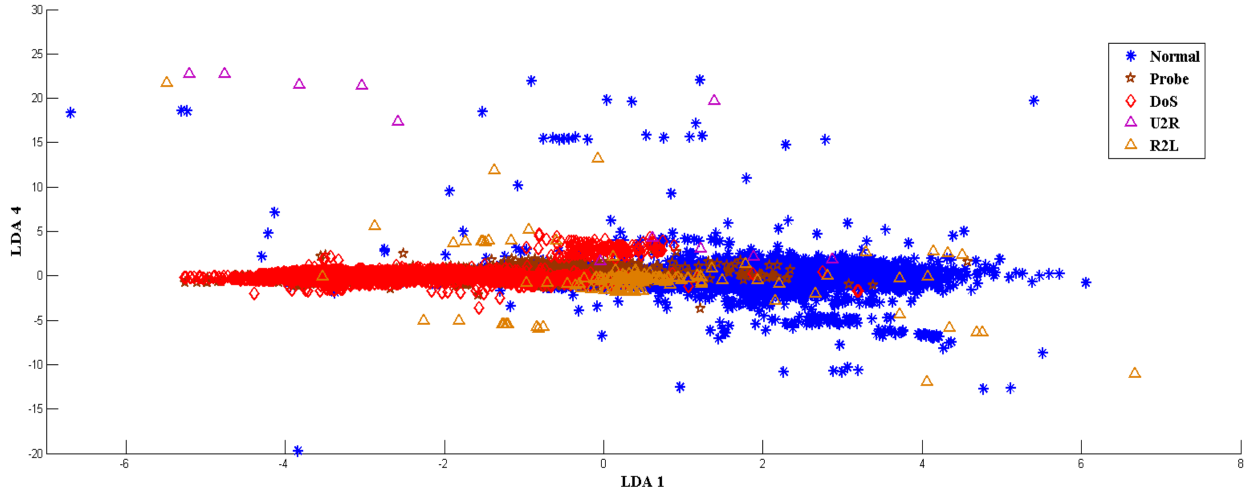

Fig 4. Two-dimensions of new mapped dataset processed by dimension reduction module

Table 2. Transformed Features Dependency Of Train_20% Data Set After Applying Two Level Of Reduction Due To Correlation Coefficient Measure.

features	**LDA1**	**LDA2**	**LDA3**	**LDA4**
LDA1	1	-8.37E-17	-4.20E-17	2.49E-16
LDA2	-8.37E-17	1	-1.89E-16	-4.88E-16
LDA3	-4.20E-17	-1.89E-16	1	4.81E-16
LDA4	2.49E-16	-4.88E-16	4.81E-16	1

C. *Classification Module*

At this stage, TDTC is already trained using the transformed dataset and classified incoming traffic utilizing a multilayer classifier (introduced in [31]) to detect anomalies. The choice of the classifier is due to its capability in detecting abnormal behaviors due to the use of:

- Two embedded classifiers for assigning exact class labels;
- Simple classifier techniques such Naïve Bayes [47] and K-Nearest Neighbor (KNN);
- Good similarity measure for rare instances to handle imbalanced datasets; and
- Bucketing technique to speed up classification tasks.

Figure 6 illustrates how classification modules are applied on incoming labeled instances. The Naïve Bayes classifier is used to classify anomalous behavior, which is then refined to normal instances using the Certainty-Factor version of K-Nearest Neighbor (CF-KNN). Naïve Bayes is an efficient classification method since it presumes independence of all features of each sample in the given class-label (conditional independence assumption).

The transformed features are assessed using correlation coefficient parameter. This measure [48] shows the relation between variables (features) by giving a number in the [-1, 1] interval, where 1 indicates a positive linear correlation, 0 no linear correlation, and −1 a negative linear correlation. The Correlation Coefficient assessments of the final features shows that the transferred features at two layers of dimension reduction module are mostly independent, since $\rho=0$.

This measure indicates that there is no strict dependency among the classifier input features – see also Tables 1 and 2. The figures dependency among the features also significantly decreases, in comparison to the findings reported in [31]. The certainty-factor similarity measure in the classification module is based on the distribution proportion of classes in the training dataset to resolve imbalance data set issue. Certainty-Factor (CF) is a number that lies in [-1, 1] interval and specifies the amount of certainty for a given incoming sample [49].

CF measure is included in the KNN [50] classification module:

- Let N (S, k) be k closest adjacent of S;
- P (C= c_i |D) be the ratio of c_i in training set D; and
- P (C= c_i |N (S, k)) be the ratio of c_i in the query result.

Now, CF measure can be computed using Eq. 10 and Eq. 11:

if (p(C= c_i |N (S, k)) ≥ p(C= c_i |D))

$$CF\big(C= c_i, N(S,k)\big) = \frac{p(C=c_i \mid N(S,k)) - p(C=c_i \mid D)}{1 - p(C=c_i \mid D)} \qquad \text{(Eq.10)}$$

Else

$$CF\big(C= c_i, N(S,k)\big) = \frac{p(C=c_i \mid N(S,k)) - p(C=c_i \mid D)}{p(C=c_i \mid D)} \qquad \text{(Eq.11)}$$

The values of CF(C= c_i, N (Q, k)) are in the range of [-1, 1]. The CF strategy for KNN classification is defined as:

$$S_{CF} = \operatorname{argmax}\left\{CF\big(C= c_i, N(Q,k)\big)\right\} \qquad \text{(Eq.12)}$$

At this tier, KNN classifier uses a bucketing technique called K-d tree [51] to accelerate the nearest neighbor searching process of KNN.

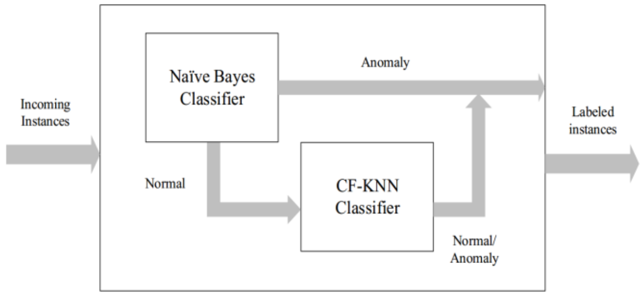

FIG. 6. APPLIED CLASSIFICATION MODULE IN TDTC

TABLE 3. NSL-KDD DATA SET CLASSES DISTRIBUTION

Datasets	Total Records	Normal	Probe	DoS	U2R	R2L
Train_20%	25192	13449	2289	9234	11	209
Train^{+}	125973	67343	11656	45927	52	995
Test^{+}	22544	9711	2421	7458	67	2887

IV. EXPERIMENT SETUP

A. NSL-KDD

In the NSL-KDD dataset, flaws reported in the original KDD99 dataset [52] were removed. Although there are still known issues in the NSL-KDD dataset [53], this does not affect the application of the dataset in this research or the validity of the findings. Each NSL-KDD record consists of a network connection with 41 defined attributes (e.g. protocol type, service and flag), which are labeled as normal or one of the 24 type of attack classes (e.g. Probe, DoS, U2R and R2L). NSL-KDD has two training sets and one test set with different distribution – see Table 1. Since the test set contains 17 new attack types not included in the training set, we can evaluate the effectiveness of TCTD in detecting unknown or uncommon attacks.

B. Data transformation

Before the dataset is applied, each feature vector is normalized to a positive integer value within the range of [1,100] in order to improve the performance of the classifier and dimension reduction module.

Each nominal feature value is specified with a unique integer number (e.g. TCP = 1, UDP = 2, ICMP = 3). The result value of each feature is mapped into an integer number, to avoid any bias, as shown in Eq.13 for each continuous-valued z. Continuous-valued features is normalized using logarithm to base 2 and then casting into an integer value.

$$\text{if } (z > 2)\ z = \int (\log_2(z) + 1) \qquad \text{(Eq.13)}$$

C. Performance indicators

The four common performance indicators for the intrusion detection systems are as follows [54]:

- True Positive (TP): indicates that benign behavior is correctly predicted as benign;
- True Negative (TN): indicates that malicious behavior is correctly detected;
- False Positive (FP): indicates that malicious behavior is identified as benign; and
- False Negative (FN): indicates that benign behavior is wrongly detected as malicious.

The Detection Rate (DR) is a measure of the classifier correctly detecting malicious samples of all malicious objects, and is computed as: DR = $\frac{TP}{FN+TP}$.

The False Alarm Rate is a measure of the classifier wrongly detecting benign samples as malicious of all benign objects, and is computed as: $FAR = \frac{FP}{FP + TN}$

V. FINDINGS

The experiment was conducted using MATLAB R2015a running on a personal computer (PC) powered by AMD Phenom II X6 3.8GHz and 12 GB RAM. TDTC is trained with both training sets and then evaluated using the test set (Test+). TDTC's classification module is adopted from [31], with the same the parameter setting. Thus, k = 3 was used for CF-KNN classifier.

Figure 7 shows the mapped test dataset into new feature space, after applying the dimension reduction module. TDTC only uses two features of new mapped data (instead of all four features of lda1 to lda4, based on detection rates) – see Figure 8.

Table 4 NSL-KDD Data Set Attacks Label Taxonomy And Their Existence In Train And Test Set Respectively.

Main Class	Sub Class (Attacks) in Train set	New Subclass (Attacks) in Test set
DoS	back, land, neptune, pod, smurf, teardrop	Apache2, Mailbomb, Processtable
Probe	ftp write, guess passwd, imap, multhop, phf, spy, warezclient, warezmaster	Mscan, Saint
User-to-Root (U2R)	Buffer overflow, perl, loadmodule, rootkit.	Httptunnel, Ps, Sqlattack, Xterm
Remote-to-Local (R2L)	ipsweep, nmap, portsweep, satan	Sendmail, Named, Snmpgetattack, Snmpguess, Xlock, Xsnoop, Worm

In addition, TDTC has an improved performance in detecting U2R and R2L attacks as shown in Table 5, as well as achieving a higher detection rate against probe attacks. The detection rate for DoS attacks in TDTC is also better than the two-tier model proposed in [16] and [55]. False alarm rate shows a reduction to 5.56% from 6.3% reported in [55].

A. Computation complexity

In TDTC, the complexity overhead was reduced to half since only 35 out of 41 data set features were used. TDTC two dimension reduction module performance is an offline task, which is applied once to obtain the transform vectors for incoming samples.

The first dimension reduction module is completely unsupervised while the generated class labels were added to the training dataset for another transformation based on the (supervised) LDA technique. The two-tier classification module (defined in [31]) embedded into TDTC reduces the computational complexity.

The computational complexity of Naïve Bayes classifier of the classification module is determined as $O(e \times f)$, where e is the count of samples in dataset and f represents number of features. Therefore at this level, due to LDA optimum transformation, the first classifier of TDTC is equipped with only four features instead of 35. Thus, the computation overhead decreases by approximately ten times. In the second tier of classifier where KNN classifier was implemented, TDTC maintains only two attributes of the training dataset with the highest detection rate, as shown in Fig 8. Therefore, KNN consumes less memory space than the original dataset. In addition KNN classifier is equipped with k-d tree [51] for searching nearest samples. K-d tree is a data structure which keeps the data sample based on their distances; thus, this technique helps KNN to search faster than using the traditional approach.

According to the second tier of classifier, searching nearest samples will take $O(\log n)$ time on average.

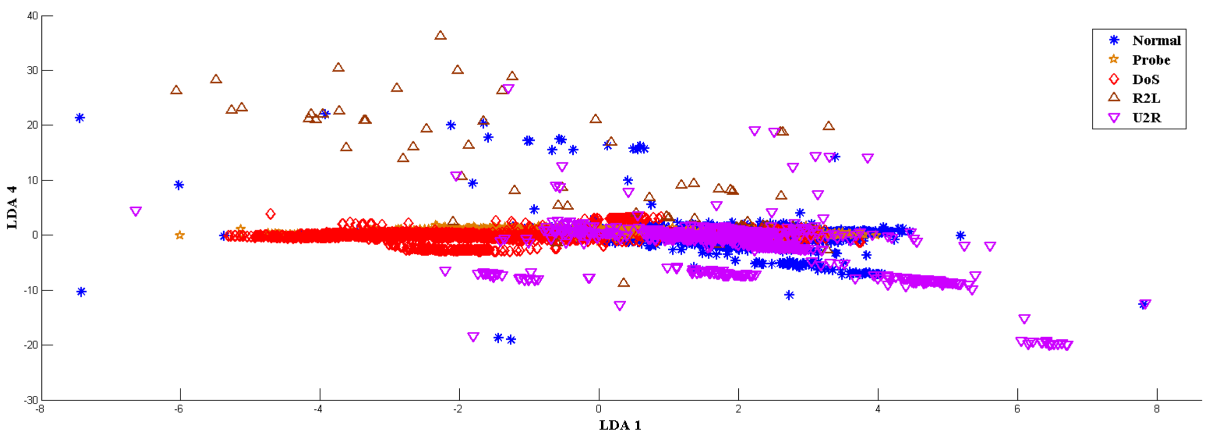

FIG 7. TWO-DIMENSIONS OF TRANSFORMED TEST SET WITH OBTAINED PROJECTION MATRIX

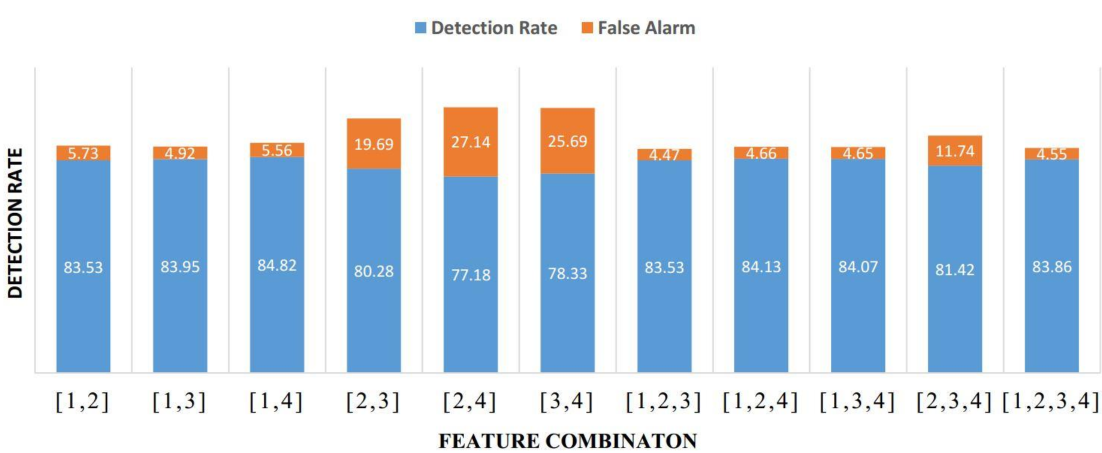

Fig 8. TDTC KNN classifier feature evaluation

B. Real-world Applications

Since TDTC has a higher performance yet relatively lower resource requirements, it can be deployed to detect intrusion attempts in IoT backbone networks and their infrastructure services. TDTC also can be deployed as an auxiliary service for digital forensics in IoT ecosystem, such as those discussed in [56] to detect residual attack patterns of IoT network layer.

Due to the increases in low frequency, low profile IoT-based attacks [39], TDTC capabilities in detection of U2R and R2L attacks are useful in incident detection and handling.

TABLE 5. A COMPARATIVE SUMMARY

Method	Normal	Probe	DoS	U2R	R2L
TDTC	94.43	87.32	88.20	**70.15**	**42**
Two-tier[31]	94.56	79.76	84.68	67.16	34.81
SVM with BIRCH [29]	99.3	**99.5**	**97.5**	28.8	19.7
ESC-IDS[28]	98.2	**99.5**	84.1	31.5	14.1
Association rule IDS [57]	**99.5**	96.8	74.9	0.79	0.38
HFR-MLR Method [55]	93.70	80.2	**89.70**	29.50	34.20

TABLE 6. TWO CLASSES CLASSIFICATION COMPARISON RESULT IN PERCENT USING TRAIN_20% AND TEST+

Method	Train set	Detection Rate	False Alarm Rate
TDTC	Train_20%	**84.82**	5.56
Two-tier [31]	Train_20%	83.24	**4.83**
Naïve Bayes [9]	Train_20%	76.56	N/A
Random forest [9]	Train_20%	80.67	N/A
SVM [9]	Train_20%	69.52	N/A
Decision trees (J48) [9]	Train_20%	81.05	N/A

TABLE 7. BINARY CLASSIFICATION COMPARISON RESULT IN PERCENT USING TRAIN+ AND TEST+

Method	Train set	Detection Rate	False Alarm Rate
TDTC	Train+	**84.86**	**4.86**
Two-tier [31]	Train+	81.97	5.44
SOM IDS [58]	Train+	75.49	N/A
Fuzzy Classification by Evolutionary Algorithms [59]	Train+	82.74	3.92
Feature selection with SVM IDS [26]	Train+	82	15

VI. CONCLUDING REMARKS

With the widespread adoption of IoT devices and services in our data-centric and Internet-connected societies, ensuring the security of IoT infrastructure is important to ensure a secure and stable society. A successful attack on the IoT infrastructure can have crippling effects. For example, compromise of IoT services in smart cities could easily lead to a major chaos or even life threatening situations (see [58], [59], [60]).

In this paper, a model with two-layer dimension reduction and classification was proposed. This model is designed to detect intrusive activities in IoT backbone networks, particularly in detecting low frequency attacks (e.g. U2R and R2L) that could have potentially damaging consequences. Our proposed model outperformed existing similar models in terms of detection rate for both low frequency and common attacks. Since TDTC uses both unsupervised (PCA) and supervised (LDA) feature extraction methods, we were able to accurately distinguish between different attack types and normal behaviors, thanks to utilized classification algorithms.

Future research includes exploring the potential of non-parametric methods such as dimension reduction module and fuzzy clustering to achieve a better classification against U2R, R2L and other attacks. Another interesting future work could be extension of the proposed model to detect intrusions at other layers of the IoT architecture such as application and support layers, as well as other protocols running in the network layer.

REFERENCES

[1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.

[2] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in Internet of Things: The road ahead," *Comput. Netw.*, vol. 76, pp. 146–164, Jan. 2015.

[3] E. M. Shakshuki, N. Kang, and T. R. Sheltami, "EAACK - A Secure Intrusion-Detection System for MANETs," *IEEE Trans. Ind. Electron.*, vol. 60, no. 3, pp. 1089–1098, Mar. 2013.

[4] H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the Internet of Things: A Review," 2012, pp. 648–651.

[5] H.-J. Liao, C.H. Richard Lin, Y.C. Lin, and K.Y. Tung, "Intrusion detection system: A comprehensive review," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 16–24, Jan. 2013.

[6] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: methods, systems and tools," *IEEE Commun. Surv. Tutor.*, vol. 16, no. 1, pp. 303–336, 2014.

[7] I. Butun, S. D. Morgera, and R. Sankar, "A Survey of Intrusion Detection Systems in Wireless Sensor Networks," *IEEE Commun. Surv. Tutor.*, vol. 16, no. 1, pp. 266–282, 2014.

[8] S. Hettich and S. D. Bay, "Kdd cup 1999 data," *UCI KD Arch. Irvine CA Univ. Calif. Dep. Inf. Comput. Sci.*, 1999.

[9] M. Tavallaee, E. Bagheri, W. Lu, and A.A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009*, 2009.

[10] Y. Bouzida and F. Cuppens, "Neural networks vs. decision trees for intrusion detection," *IEEEIST Workshop Monit. Attack Detect. Mitig. MonAM Tuebingen Ger.*, vol. 28, p. 29, 2006.

[11] R. Beghdad, “Efficient deterministic method for detecting new U2R attacks,” *Comput. Commun.*, vol. 32, no. 6, pp. 1104–1110, 2009.
[12] S. Dua and X. Du, *Data Mining and Machine Learning in Cybersecurity*. CRC Press, 2011.
[13] O. Osanaiye, K.K. R. Choo, and M. Dlodlo, “Distributed denial of service (DDoS) resilience in cloud: Review and conceptual cloud DDoS mitigation framework,” *J. Netw. Comput. Appl.*, vol. 67, pp. 147–165, May 2016.
[14] F. Daryabar, A. Dehghantanha, N. I. Udzir, S. bin Shamsuddin, and others, “Towards secure model for scada systems,” in *Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012 International Conference on*, 2012, pp. 60–64.
[15] D. Ariu, R. Tronci, and G. Giacinto, “HMMPayl: An intrusion detection system based on Hidden Markov Models,” *Comput. Secur.*, vol. 30, no. 4, pp. 221–241, Jun. 2011.
[16] L. Koc, T. A. Mazzuchi, and S. Sarkani, “A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier,” *Expert Syst. Appl.*, vol. 39, no. 18, pp. 13492–13500, 2012.
[17] W.C. Lin, S.W. Ke, and C.F. Tsai, “CANN: An intrusion detection system based on combining cluster centers and nearest neighbors,” *Knowl.-Based Syst.*, vol. 78, pp. 13–21, 2015.
[18] M. Thottan and C. Ji, “Anomaly detection in IP networks,” *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2191–2204, 2003.
[19] D. J. Weller-Fahy, B. J. Borghetti, and A. A. Sodemann, “A Survey of Distance and Similarity Measures Used Within Network Intrusion Anomaly Detection,” *IEEE Commun. Surv. Tutor.*, vol. 17, no. 1, pp. 70–91, 2015.
[20] J. P. Theiler and D. M. Cai, “Resampling approach for anomaly detection in multispectral images,” in *AeroSense 2003*, 2003, pp. 230–240.
[21] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv. CSUR*, vol. 41, no. 3, p. 15, 2009.
[22] P. Casas, J. Mazel, and P. Owezarski, “Unsupervised network intrusion detection systems: Detecting the unknown without knowledge,” *Comput. Commun.*, vol. 35, no. 7, pp. 772–783, 2012.
[23] E. De la Hoz, E. De La Hoz, A. Ortiz, J. Ortega, and B. Prieto, “PCA filtering and probabilistic SOM for network intrusion detection,” *Neurocomputing*, vol. 164, pp. 71–81, 2015.
[24] H. Bostani and M. Sheikhan, “Modification of supervised OPFbased intrusion detection systems using unsupervised learning and social network concept,” *Pattern Recognit.*, vol. 62, pp. 56–72, Feb. 2017.
[25] Z.S. Pan, S.C. Chen, G.B. Hu, and D.Q. Zhang, “Hybrid neural network and C4. 5 for misuse detection,” in *Machine Learning and Cybernetics, 2003 International Conference on*, 2003, vol. 4, pp. 2463–2467.
[26] J. Zhang, M. Zulkernine, and A. Haque, “Random-forests-based network intrusion detection systems,” *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 38, no. 5, pp. 649–659, 2008.
[27] C. Guo, Y. Ping, N. Liu, and S.-S. Luo, “A two-level hybrid approach for intrusion detection,” *Neurocomputing*, 2016.
[28] A. N. Toosi and M. Kahani, “A new approach to intrusion detection based on an evolutionary soft computing model using neuro-fuzzy classifiers,” *Comput. Commun.*, vol. 30, no. 10, pp. 2201–2212, Jul. 2007.
[29] S.J. Horng *et al.*, “A novel intrusion detection system based on hierarchical clustering and support vector machines,” *Expert Syst. Appl.*, vol. 38, no. 1, pp. 306–313, Jan. 2011.
[30] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, “A system for denial-of-service attack detection based on multivariate correlation analysis,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 2, pp. 447–456, 2014.
[31] H. H. Pajouh, G. Dastghaibyfard, and S. Hashemi, “Two-tier network anomaly detection model: a machine learning approach,” *J. Intell. Inf. Syst.*, pp. 1–14, 2015.
[32] S. Iqbal *et al.*, “On cloud security attacks: A taxonomy and intrusion detection and prevention as a service,” *J. Netw. Comput. Appl.*, vol. 74, pp. 98–120, Oct. 2016.
[33] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, “Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm,” *IEEE Trans. Comput.*, vol. 65, no. 10, pp. 2986–2998, Oct. 2016.
[34] S. Pan, T. Morris, and U. Adhikari, “Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems,” *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3104–3113, Nov. 2015.
[35] C. Zhou *et al.*, “Design and Analysis of Multimodel-Based Anomaly Intrusion Detection Systems in Industrial Process Automation,” *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 45, no. 10, pp. 1345–1360, Oct. 2015.
[36] A. Whitmore, A. Agarwal, and L. D. Xu, “The Internet of Things—A survey of topics and trends,” *Inf. Syst. Front.*, vol. 17, no. 2, pp. 261–274, Mar. 2014.
[37] Q. M. Ashraf and M. H. Habaebi, “Autonomic schemes for threat mitigation in Internet of Things,” *J. Netw. Comput. Appl.*, vol. 49, pp. 112–127, Mar. 2015.
[38] H. Ning, H. Liu, and L. T. Yang, “Aggregated-Proof Based Hierarchical Authentication Scheme for the Internet of Things,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 3, pp. 657–667, Mar. 2015.
[39] X. Cao, D. M. Shila, Y. Cheng, Z. Yang, Y. Zhou, and J. Chen, “Ghost-in-ZigBee: Energy Depletion Attack on ZigBee-Based Wireless Networks,” *IEEE Internet Things J.*, vol. 3, no. 5, pp. 816–829, Oct. 2016.
[40] Q. Chen, S. Abdelwahed, and A. Erradi, “A Model-Based Validated Autonomic Approach to Self-Protect

Computing Systems,” *IEEE Internet Things J.*, vol. 1, no. 5, pp. 446–460, Oct. 2014.
[41] F. A. Teixeira *et al.*, “Defending Internet of Things against Exploits,” *IEEE Lat. Am. Trans.*, vol. 13, no. 4, pp. 1112–1119, Apr. 2015.
[42] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.
[43] E. R. Malinowski, “Determination of the number of factors and the experimental error in a data matrix,” *Anal. Chem.*, vol. 49, no. 4, pp. 612–617, 1977.
[44] S. Valle, W. Li, and S. J. Qin, “Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods †,” *Ind. Eng. Chem. Res.*, vol. 38, no. 11, pp. 4389–4401, Nov. 1999.
[45] T. Li, S. Zhu, and M. Ogihara, “Using discriminant analysis for multi-class classification: an experimental investigation,” *Knowl. Inf. Syst.*, vol. 10, no. 4, pp. 453–472, 2006.
[46] Z. Tan, A. Jamdagni, X. He, and P. Nanda, “Network Intrusion Detection based on LDA for payload feature selection,” in *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*, 2010, pp. 1545–1549.
[47] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.
[48] P. D. J. Benesty, J. Chen, Y. Huang, and P. I. Cohen, “Pearson Correlation Coefficient,” in *Noise Reduction in Speech Processing*, Springer Berlin Heidelberg, 2009, pp. 1–4.
[49] J. C. Giarratano and G. Riley, *Expert systems*. PWS Publishing Co., 1998.
[50] S. Zhang, “KNN-CF Approach: Incorporating Certainty Factor to kNN Classification.,” *IEEE Intell. Inform. Bull.*, vol. 11, no. 1, pp. 24–33, 2010.
[51] J. H. Friedman, J. L. Bentley, and R. A. Finkel, “An algorithm for finding best matches in logarithmic expected time,” *ACM Trans. Math. Softw. TOMS*, vol. 3, no. 3, pp. 209–226, 1977.
[52] M. Panda, A. Abraham, and M. R. Patra, “Discriminative multinomial naive bayes for network intrusion detection,” in *Information Assurance and Security (IAS), 2010 Sixth International Conference on*, 2010, pp. 5–10.
[53] R. P. Lippmann *et al.*, “Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation,” in *DARPA Information Survivability Conference and Exposition, 2000. DISCEX’00. Proceedings*, 2000, vol. 2, pp. 12–26.
[54] A. L. Buczak and E. Guven, “A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection,” *IEEE Commun. Surv. Tutor.*, vol. 18, no. 2, pp. 1153–1176, 2016.
[55] “A Novel Anomaly Detection System Based on HFR-MLR Method,” *Mobile, Ubiquitous, and Intelligent Computing*, vol. 274, pp. 279–286, 2014.
[56] S. Watson and A. Dehghantanha, “Digital forensics: the missing piece of the Internet of Things promise,” *Comput. Fraud Secur.*, vol. 2016, no. 6, pp. 5–8, Jun. 2016.
[57] W. Xuren, H. Famei, and X. Rongsheng, “Modeling Intrusion Detection System by Discovering Association Rule in Rough Set Theory Framework,” in *International Conference on Computational Intelligence for Modelling, Control and Automation, 2006 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, 2006, pp. 24–24.
[58] C. D'Orazio, K.-K. R. Choo, and L. T. Yang, "Data Exfiltration from Internet of Things Devices: iOS Devices as Case Studies," *IEEE Internet of Things Journal* [In press]. doi: 10.1109/JIOT.2016.2569094.
[59] N. D. W. Cahyani, B. Martini, K.-K. R. Choo, and M. H. Al-Azhar, "Forensic data acquisition from cloud-of-things devices: Windows smartphones as a case study," *Concurrency and Computation: Practice and Experience* [In press]. doi: 10.1002/cpe.3855.
[60] K.-K. R. Choo, "A conceptual interdisciplinary plug-and-play cyber security framework". In Kaur H & Tao X, editors, ICTs and the Millennium Development Goals – A United Nations Perspective, pp. 81–99, New York, USA: Springer, 2014.

Hamed HaddadPajouh is Ph.D. Student of Information Technology Engineering (Computer Networks) at Shiraz University of Technology, Shiraz, Iran. He earned his Ph.D. scholarship position as talented student. He is also Information Security lecturer at Shiraz University of Technology and University of Applied Sciences. He founded Computer security and forensics community for talented students in his hometown. His research interests are Network Security like intrusion detection systems and malware analysis, Machine learning and enterprise security architecture.

Reza Javidan is graduated from MSc Degree in Computer Engineering (Machine Intelligence and Robotics) from Shiraz University in 1996. He received PhD degree in Computer Engineering (Artificial Intelligence) from Shiraz University in 2007. He has many Ph.D and MSc students working together as a team. His major fields of interest are Network security, Underwater Wireless Sensor Networks (UWSNs), Software Defined Networks (SDNs), artificial intelligence, image processing and SONAR systems. Reza is now member of faculty and lecturer in Department of Computer Engineering and Information Technology in Shiraz University of Technology.

Raouf Khayami received his BS degree in computer engineering (hardware systems) from Shiraz University in 1993, the MS degree in artificial intelligence and robotics from the same university in 1996, and the Ph.D. degree in software systems from Shiraz University in 2009. Raouf is currently an assistant professor in the Computer Engineering and Information Technology Department, Shiraz University of Technology, Shiraz, Iran, and there, he is the Head of the Department. His research interests include data mining, business intelligence, and enterprise architecture, on which he has published a number of refereed articles, surveys and technical reports in prestigious national and international conferences and journals. He also is active in consulting and industrial projects.

Ali Dehghantanha (SM'16) is a Marie-Curie International Incoming Fellow in Cyber Forensics and a fellow of the UK

Higher Education Academy (HEA). He has served for many years in a variety of research and industrial positions. Other than Ph.D. in Cyber Security he holds many professional certificates such as GXPN, GREM, CISM, CISSP, and CCFP. He has served as an expert witness, cyber forensics analysts and malware researcher with leading players in Cyber-Security and E-Commerce.

Kim-Kwang Raymond Choo (SM'15) received the Ph.D. in Information Security in 2006 from Queensland University of Technology, Australia. He currently holds the Cloud Technology Endowed Professorship at The University of Texas at San Antonio, and is an associate professor at University of South Australia. He is the recipient of various awards including ESORICS 2015 Best Paper Award, Winning Team of the Germany's University of Erlangen-Nuremberg (FAU) Digital Forensics Research Challenge 2015, and 2014 Highly Commended Award by the Australia New Zealand Policing Advisory Agency, Fulbright Scholarship in 2009, 2008 Australia Day Achievement Medallion, and British Computer Society's Wilkes Award in 2008. He is a Fellow of the Australian Computer Society, and a Senior Member of IEEE.