Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 7348 - 7361
Received 16 February 2021; Accepted 08 March 2021.

An Efficient Intrusion Detection Scheme Using Revised Equality Constraints based Lagrange's Multiplier for Cloud Applications

[1]Dr.Raju Rameshkumar, Professor/CSE, Sphoorthy Engineering College,
Nadergul(V),Balapur(M), Ranga Reddy Dist., Telangana.

[2]Mr.Thotakura Veeranna, Asso.Professor/CSE,SaiSpurthi Institute of Technology, B.Gangaram,
Sathupalli(M), Khammam Dist., Telangana.

[3]Mr.Shaik Yakoob,Asso.Professor/CSE, SaiSpurthi Institute of Technology, B.Gangaram,
Sathupalli(M), Khammam Dist., Telangana.

[4]Dr.G.Kavithaa, Assistant Professor, Department of Electronics and Communication
Engineering, Government College of Engineering, Salem, TamilNadu, India.

ABSTRACT

Cloud computing has become a major part of the IT industry in recent years. The end-users via the Internet easily provide a framework that connects to powerful services and applications in the cloud. An intrusion detection system is the most common mechanism for detecting an attack and is inefficient to be deployed in the cloud for unknown attacks.It is challenging to observe and perceive malicious activity within any manipulative system or web, or cloud. In the existing method,the intrusion detection system is designed based on the Stacked Contractive Auto-Encoder (SCAE)classification algorithm. SCAE classificationalgorithmin existing work requires high-dimensional representations method whichprovidesa deprivedpresentation for massive datasets.In the proposed scheme,the dataset is preprocessed using the Improved Feature Scaling (IFS) algorithmwith effective rules. The resultant preprocessed data is then trained with a training process. The trained information is then compared with the test dataset, and then the classification processis done using Revised Equality Constraints Based Lagrange's Multiplier (RECLM). The classification using RECLMprovides an outstanding classification result of 99.68% accuracy for attack typedetection for large datasets and multiple groupings in the clouds.

Keywords: Cloud Computing, Intrusion Detection System (IDS), Improved feature scaling preprocessing, Revised Equality Constraints Based Lagrange's Multiplier (RECLM).

1. INTRODUCTION

The intrusion detection framework turns into a significant component in framework security that controls constant information and prompts a huge dimensional issue. Hence, information pre-handling is essential to diminish errorsand clean the system information.

Another strategy to take care of complex advancement issues has been proposed to lessen the false positive rate and build recognition productivity. The Single sign-on (SSO) is the session user authentication service that uses a set of login credentials.SSO makes it easier to create, remember and use strong passwords, as users only need to use a password.

Accuracy issues must be resolved to reduce false alarm rates and attack recognition rates. This concept has the potential to drive this exploration activity. Thus, Support Vector Machines (SVM), Random Forests (RF), and Extreme Learning approaches are applied to the task. These methods have been proven to address the problem of efficient classification. The Fixed Data Insertion Detection Agency sets the KDD to be validated. For this task, KDD is considered a benchmark to evaluate the use of an advanced knowledge discovery Data Mining Methods (DMM) dataset.

Intrusion detection methods based on Manifold Kernel Support Vector Machines (MK-SVM) have been analyzed and can calculate kernel functionality simultaneously. The semi-infinite linear and programmable Lagrange multiplier weights achieve kernel functionality and classification optimization selection. The preprocessing technique is usedfor time and space requirements [3].

For large datasets and multiple classifications,the classification using Revised Equality Constraints Based Lagrange's Multiplier algorithm is used in the proposed technique.SCAE method in the existing process deliverslow performance on large datasets and various classifications.The proposed Revised Equality Constraints Based Lagrange's Multiplier classification has overcome the limitations of the current approach. The revisedequality constraints-based Lagrange multiplier with SVM is used in the classification processextends the marginal distances between two data groups(i.e., standard or attack).

2. RELATED WORKS

To accomplish a steady pace of counterfeit alerts and simultaneously recognize changes in the measurable model, the system depends on the change-point recognition hypothesis and test point statement [1]. Propelled calculations are self-discovering that empower them to adjust to various system burdens and utilization designs [2]. They permit minors to have a normal postponed affirmation of a specific bogus alert rate. It is anything but difficult to ascertain, and it very well may be accomplished on the web [3] [4].

To distinguish new mysterious attacks, the half and half framework consolidate the upsides of an Acquired Demyelinating Syndrome (ADS) with anIntrusion Detection System (IDS) and a mark-based interruption location framework. Weightedsignature generation program has emerged from the waves of integrating signatures into disgusting adverts. HIDS across internet connection episodesapproach Simultaneous detection of automatic data mining and signature generationthat demonstrate the vitality of intrusions and anomalies [5].

Hybrid detection systems can improve detection performance by merging the benefits of the fault and anomaly detection data discovery efforts in the knowledge discovery dataset. Unsupervised anomaly detection methods achieve a worse false alarm rate and a sophisticated detection rate to close other reports [6].

A multi-model anomaly detection algorithm has been proposed in many works. Anomalydetectionwas introducedto overcome classifiers' shortcomings based on intelligent Hidden Markov Models designed to distinguish real attacks.A complete examination of the acknowledgment precision and real execution of the existingattack identification framework is done. These attacks are dependent on a joined reproduction stage utilizing upgraded execution organize building devices [7].

The Bayesian game model's existing systemprecisely detected these two security problems and solved this work with less cost. The exchange of intrusion detection rate and overhead is measured in this work. Whenever it shows a bad sign of harm, this code may be temporary, so it immediately pops up or noises or chants. This is because passive communication is not the best strategy. Finding the incorrect rates is considered [8].

In the existing method’sintrusion detection framework, the graph's matrix is estimated using the Gaussian-Markov probability field models to capture the captured data's size without measuring the adaptive sensor. The existing intrusion detection method is based on the Naïve Bayesian likelihood ratio test design, compiled and the formula closed-form test point is obtained. A complete analysis of network characters isfinally generated by calculating the distance between subsequent instants and the measured distribution [9].

The interruption location framework is utilized to examine enormous traffic information. In this manner, a valuable machine learningmethod is used to avoid the issue. This issue is viewed as acceptable and is a notable method utilized in SVM, lopsided woodlands, and thick learning machines. These strategies have presumptions of notable order capacities. It is viewed that highlights ought to be utilized to assess information revelation and informational collections and information entrance location strategies [10].

Supply Chain management, such as Block chain technology, has proven its compatibility in many areas of international charges mutual banking. Block chain is an area suitable for intrusion detection that protects data storage integrity, recognizes cyber network threats and possible events, and ensures transparency of the process [11].If the ID is to be reliably checked on simple network topology and then installed on a random walk, the new Navigation Detection System Verification Can-FT network topology is proposed to increase the Can-FT network's security. If an intrusion is detected, the security mode is triggered to prevent further attacks on the system [12].

The Wireless Sensor Network IDS system is provided for a complete study of the use of the machine and deep learning solution. To track the latent deep learning-based IDS system for critical infrastructure and wireless sensor networks, the Restricted Boltzmann Machine-Based Clustered IDS (RBS IDS)is achieved [13].

With the advent of the Internet of Things (IOT) and the network layer's security, this work has attracted more and more attention. Traditional intrusion detection technology cannot adapt to things complex Internet environment.Consider the impact of attack strength and node movement on attack detection performance and get high attack detection accuracy by analyzing locally available information without increasing packet overhead [14].The infiltration detection mechanism is measuredas the primary source of information and telecommunicationssecurity. Challenge in intrusion detection framework is identified when the number of instances references class is small to deal with unbalanced intrusion data set[15].

The intrusion detection is based on an improved genetic algorithm and a robust confidence network model. With the hidden number of neurons of all layers, the optimal amount of multiple loci of GA is adaptively generated, depending on the attack and the penetration detection model to achieve a high detection rate. Finally, it evaluates the KDD file 99 used to simulate the data set and algorithm [16].

An Intrusion Detection System (IDS) detects malicious activity and it plays an important role in network security by preventing it [17]. Complex and time-varying network environment and mass network intrusion samples of normal samples bring a high false-positive rate of soaked insufficient sample training and detection. The existing algorithm One-Side Selection (OSS) is used to minimize the noise based on Synthetic Minority Over-sampling Technique (SMOTE) algorithm [18].

With the spread of network technology and the development of the Internet, it can identify the attack IDS that has been developed. The traditional intrusion detection algorithm is usually used to identify the invasion by using the mining industry association rules.The second layer, the receptor immune system anomaly detection, has been studied to determine susceptible network sets with probabilities [19].

3. MATERIALS AND METHODS

The Intrusion Detection System (IDS) plays an important role in the perseverance of active defense systems against malicious attacks by an intruder of security with any company and the IT organization. The IDS implementation in cloud computing needs an efficient and scalable virtualization approach.The proposedRECMalgorithm is used to effectively solve the explosion parameters' problem to ensure the accuracy of the classification.

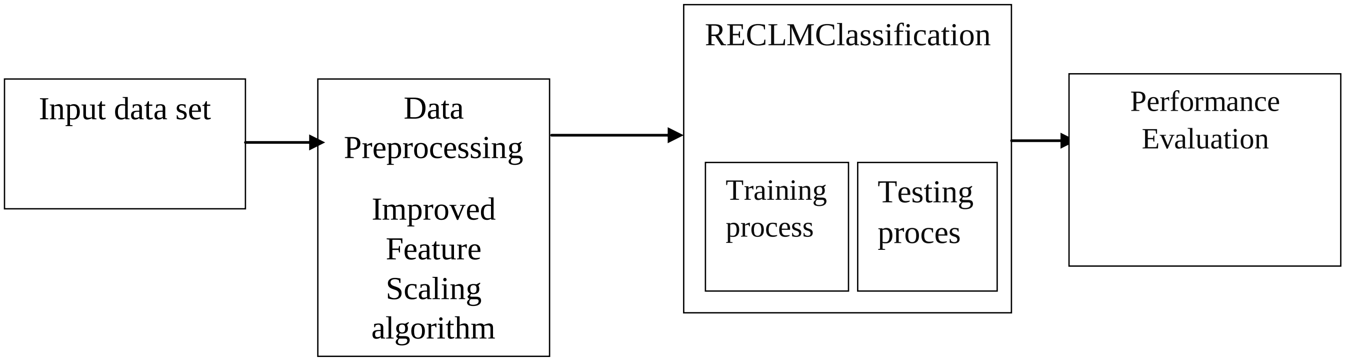

Figure1. System Architecture for a proposed Intrusion detection system

The input dataset is a sample dataset in the organization design, namely KDD Cup 99 datasets (Where KDD is Knowledge Discovery and Data Mining dataset) is an intrusion detection dataset. These datasets were first preprocessed using the preprocessing algorithm to transform the features by scaling each element into a given range. Transform each function individually, that this algorithm scale is within the range given for the training set, between 0 and 1. The training process is done to produce a comparison of training data with test data. The trained data is then compared with the test data, and then the classifier detects and is classified into normal or an attack. The accuracy of classifiers, False Positive Rate, and Area underCurve arethe major performance evaluation parameters.

3.1. Preprocessing of the datasetUsing Improved Feature Scaling algorithm:

One-hot encoders are used to encode features in categories, which increase the number of essential elements. The next step in operation is functional normalization.Improved feature scaling is used to preprocess a rescaled dataset independently of other samples. Every example with at least one non-zero element has its normalized value equal to 1. These datasets were first preprocessed using the preprocessing algorithm to transform the features by scaling each feature to a given range. The proposed IFS algorithm transforms each function individually, so that proposed algorithm is within the range given for the training set i.e. between 0 and 1.

3.1.1.Improved Feature Scaling Algorithm Steps:

Input:tuple (min, max)

Output: (default = (0, 1)) →the desired range of transformed data.

Step1: Obtain feature scaling with Base value and Transformer value;

Step2: Define the initial with self, sample range like 0,1;

Step3: Calculate type. Samplerange = sample range

Step4: Define the fit function of values such as self, X=0, y=1;

Step5: return self-value;

Step 6: Define the transform values such as self, X;

Step7: Initialize X value and checkarray with data typeas float64;

Step8: Perform improved feature scaling with parameters as X and Yas self-samplerange

Step9: After preprocessing function, then perform transform function;

Step10: Return the result of the transform function as X-new.

3.2. Training process:

As a well-known Navigation Rating Data Package, the KDD Cup 99 datasetspackage is a sizeable location-based information set that runs under grinding conditions. This data set includes more than five million quick tests on a single trial of 2,000,000. In such an extensive data set range, the classic network configuration and testing procedures are obstructed. Because of the disappointing performance of the architecture brought by the insufficient network client memory as the traditional network is irrelevant. Also, data sets contain a broad range of information leaks that, as a rule, are noisy or do not provide the fundamental difficulty of information.

3.3 Classification using RECLM:

Once the dataset has been trained, this trained set is brought into the classification phase, where Revised Equality Constraints Based Lagrange's Multiplier is used. It can adoptRevised Equality Constraints Based Lagrange's Multiplierclassifieris the only binary classification problem that can be handled to select the 15 best data sets. The KDD Cup 99 is an intrusion detection dataset. Each category is different from other types of records. For example, a common type of attack differs based on information received by the system. The DoS attack class distinguishes communications based on data collected by non-DoS instance (R2L (including the Remote to Local User) R2L and U2R (Root to User) Instances service (denial of service). Then the RECLM classifiers are all combined to create an intrusion detection model that distinguishes different classes.

3.3.1. RECLMALGORITHM STEPS

Input: Input data set for the training process, test data from online

Output: Classification result as Normal or Attack

```
    Start
Step 1:  Acquire input data.
Step 2:  for each sample, do
Step 3: Train the input data;
Step 4:Compare the test data;
Step 5:If (test data == train data)then status= "normal";
    Else status= "attack";
Step 6:  if (status =="attack") then
Step 7:Feed input feature to classifier
        {
        getattack type
        }
Step 9:  attack type = classifier’s output;
        else
Step 10:  attack type="normal";
        end
Step 11:Acquireattack type;
Step 12:Send the classification result;
        end
```

Initially, the cloud IDS collects sample input data from various hosts and performs preprocessing and classification on these data into attack types.These results of the classificationare shared with the network by the administrator. The executive's system actualizes grouping by taking the effect of prepared information and afterward contrasting the test information as assault inputs and characterizing the test information as normal or an attack. During classification, the administrator's network classifies the data to detect if the feature vector is susceptible. If the feature vector is considered fragile, then it will attack network-type administrators positively. Besides, the administrator's network organization environment is interactive. The hierarchical condition is constrained by the manager, who is liable for the real dynamic approaches and eventually settling on the framework.

4. RESULTS AND DISCUSSION

Implementing the proposed intrusion detection systemusingRevised Equality Constraints Based Lagrange's Multiplier (RECLM) is simulated in Java language. The Existing classifierssuch as K-Nearest Neighbors (KNN)and Stacked Contractive Auto-Encoder (SCAE),are compared with the proposed method Revised Equality Constraints Based Lagrange's Multiplier(RECLM).The various factors like Accuracy, False Positive Rate, and Area under Curve values of all the classifiers are compared and evaluated the performance values.The proposedRECLM model is usedto detect attacks on the network and implement them on the cloud.

Table 1. Implementation parameter used in the proposed method

Processed Parameter	**Value processed**
Type of data	Datasets for IDS in the cloud
Name of dataset	KDD Cup 99 datasets
Service provider	CSP
Language	Python
Tool	Anaconda

Above table 1 describes the proposed implementation and classification analysis parameters. The most common type of data used for evaluating intrusion detection methods is a KDD file 99 datasets.The KDD Cup 99 dataset contains approximately 5 million training data records and 2 million test data records. Here, about 10% of the records, namely 494021 training data records and 311029 test data records are used to evaluate theRECLM classifier. By removing the records shown in the training data rather than the test data, we have 292300 test records left. Each record consists of 41 different features and is labeled as either normal or attack.The efficiency of classification using Revised Equality Constraints Based Lagrange's

Multiplier classifier has been analyzed and the results are shown below in the detailed tables and graphs.

- True Positive (TP): Actual positive anticipated as positive
- True Negative (TN): Actual negative anticipated as negative
- False Positive (FP): Actual negative anticipated as positive
- False Negative (FN): Actual positive anticipated as negative

Accuracy (ACC): The proportion of genuine qualities to add up to perceptions, determined as follows:

$$ACC = \frac{TP+TN}{TN} \dots \quad (1)$$

The proposed work aims to maintain a low False Positive Rate (FPR) and thePerformance evaluation of several classifiers are compared and the proposed Revised Equality Constraints Based Lagrange's Multiplierclassificationoutperforms the existing classifiers in the cloud environment.

Table 2. Performance Analysis of variousClassifiers

Classifiers	**Accuracy in %**	**FPR in %**	**Area Under a Curve** **AUC in per unit**
KNN	78.95	2.96	0.8054
SCAE	76.65	2.801	0.8087
RECLM	99.68	0.32	0.8098

The proposed Revised Equality Constraints Based Lagrange's Multiplier classification methodprovides 99.68% accuracy than the existing methods. The False Positive Rate is 0.32 %, which is the percentage of detecting non-attack as an attack, which is remarkably lower than the

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 7348 - 7361
Received 16 February 2021; Accepted 08 March 2021.

current arrangements. Thearea undera ROC Curve is 0.8098is used to classify the attacks in the network model. The performance of other classifiers is deficient when compared with the proposed method.

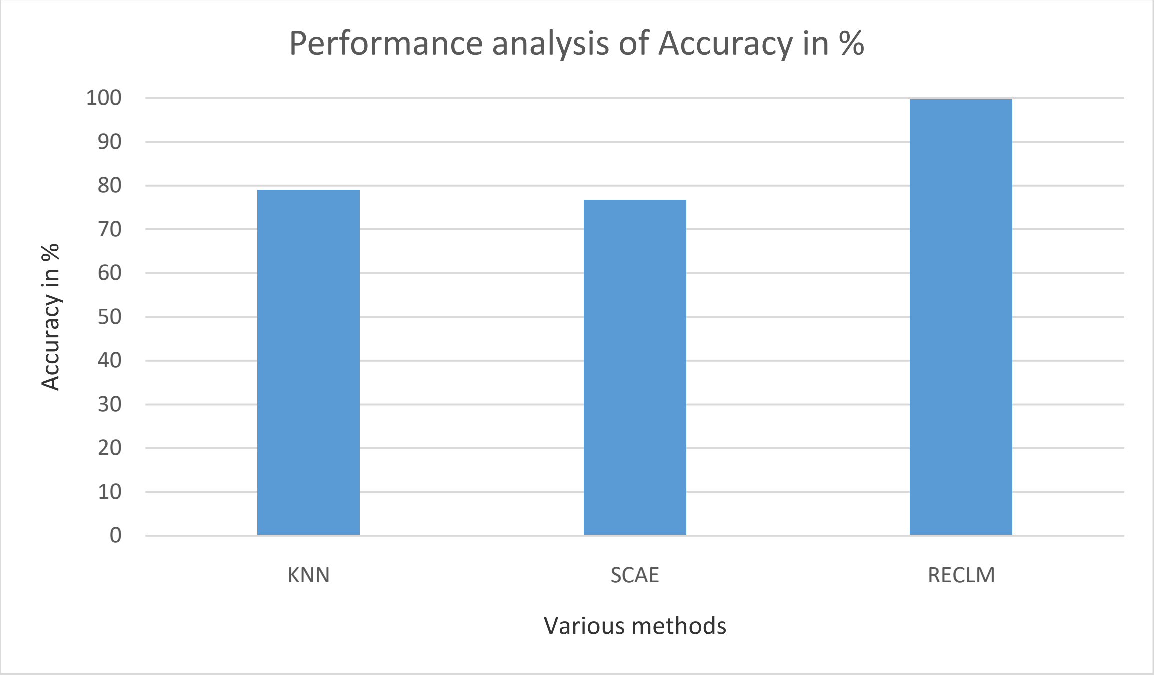

Figure2.Performance Comparison of accuracy for various classifiers

Figure2indicates the performance comparison of the accuracy of classifiers. The proposed classifier performance is higher when comparing to the existing methods. The accuracy of Revised Equality Constraints Based Lagrange's Multiplier is 99.68%.

False Positive Rate: The proportion of false positives to the total of false positives and true negatives. Also known as fallout.

$$FPR = \frac{FP}{FP + TN} \dots (2)$$

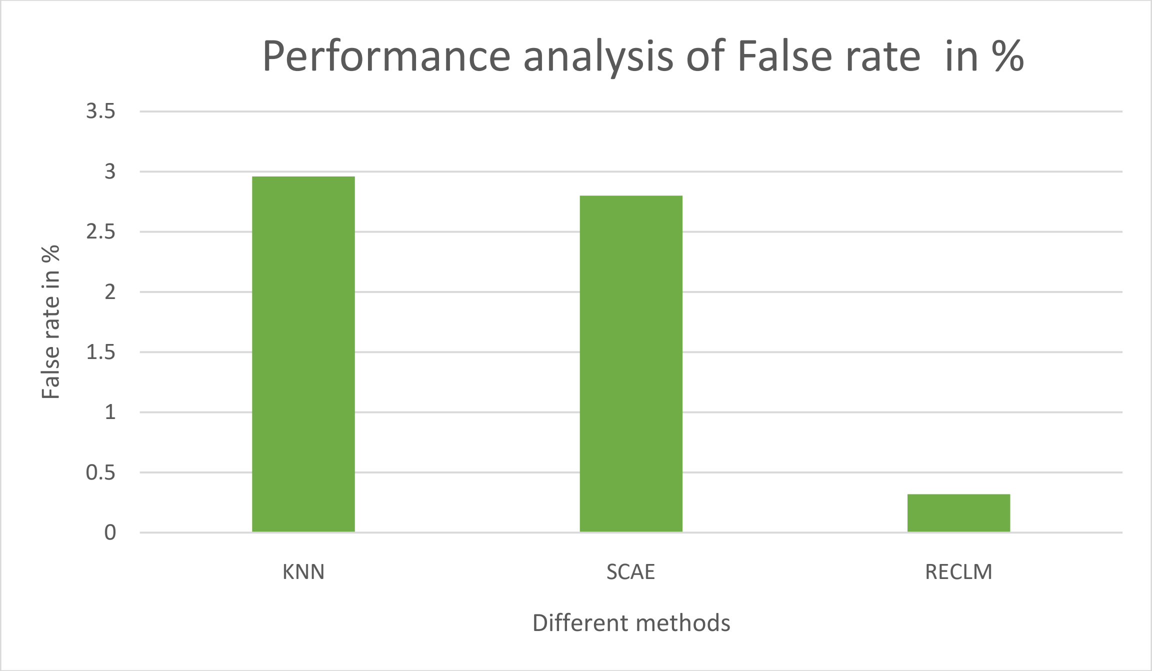

Figure 3.Performance Comparison of False Positive Rate

Figure 3 shows the False Positive Rate which is the percentage of detecting non-attack as an attack, which is remarkably lower than the current methods. The proposed classifier'sFalsePositiveRate is low when comparing to the existing methods. The FPR of Revised Equality Constraints Based Lagrange's Multiplier is 0.32.

5. CONCLUSION

In this proposed work,an Intrusion Detection Systemwith a combination of improved feature scaling for Preprocessing andRevised Equality Constraints Based Lagrange's Multiplier feature selection for efficient classification of IDS in the cloud is used. Most advanced intrusion detection systems are incapable of responding to new attacks and preserve a balance between accuracy and cluster. The proposed Revised Equality Constraints Based Lagrange's Multiplier classification method providesa 99.68% accuracy than the existing methods. The False Positive Rate is 0.32 %, which is the percentage of detecting non-attack as an attack, which is remarkably lower than the current arrangements. The Area underCurve (AUC) is 0.8098 which is defined as the model whichcan distinguish and classify the attacks in the network. The performance of other classifiers is deficient when compared with the proposed method.

REFERENCES

1. Tartakovsky, A. G., Rozovskii, B. L., Blazek, R. B., &Hongjoong, Kim. (2006). A novel approach to detecting intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods. IEEE Transactions on Signal Processing, 54(9), 3372–3382.

2. Hwang, K., Cai, M., Chen, Y., & Qin, M. (2007). Hybrid Intrusion Detection with Weighted Signature Generation over Anomalous Internet Episodes. IEEE Transactions on Dependable and Secure Computing, 4(1), 41–55.

3. Jiong Zhang, Zulkernine, M., &Haque, A. (2008). Random-Forests-Based Network Intrusion Detection Systems. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38(5), 649–659.

4. Zhou, C., Huang, S., Xiong, N., Yang, S.-H., Li, H., Qin, Y., & Li, X. (2015). Design and Analysis of Multimodel-Based Anomaly Intrusion Detection Systems in Industrial Process Automation. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45(10), 1345–1360.

5. Sadreazami, H., Mohammadi, A., Asif, A., &Plataniotis, K. N. (2018). Distributed-Graph-Based Statistical Approach for Intrusion Detection in Cyber-Physical Systems. IEEE Transactions on Signal and Information Processing over Networks, 4(1), 137–147.

6. Ahmad, I., Basheri, M., Iqbal, M. J., & Rahim, A. (2018). Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection. IEEE Access, 6, 33789–33795.

7. Meng, W., Tischhauser, E. W., Wang, Q., Wang, Y., & Han, J. (2018). When Intrusion Detection Meets Blockchain Technology: A Review. IEEE Access, 6, 10179–10188.

8. Yang, H., & Wang, F. (2019). Wireless Network Intrusion Detection Based on Improved Convolutional Neural Network. IEEE Access, 7, 64366–64374.

9. Yu, T., & Wang, X. (2019). Topology Verification Enabled Intrusion Detection for In-Vehicle CAN-FD Networks. IEEE Communications Letters, 1–1.

10. Otoum, S., Kantarci, B., &Mouftah, H. T. (2019). On the Feasibility of Deep Learning in Sensor Network Intrusion Detection. IEEE Networking Letters, 1–1.

11. Zuniga-Mejia, J., Villalpando-Hernandez, R., Vargas-Rosales, C., &Spanias, A. (2019). A Linear Systems Perspective on Intrusion Detection for Routing in Reconfigurable Wireless Networks. IEEE Access, 7, 60486–60500.

12. Jan, S. U., Ahmed, S., Shakhov, V., & Koo, I. (2019). Towards a Lightweight Intrusion Detection System for the Internet of Things. IEEE Access, 1–1.

13. Camacho, J., Theron, R., Garcia-Gimenez, J. M., Macia-Fernandez, G., & Garcia-Teodoro, P. (2019). Group-Wise Principal Component Analysis for Exploratory Intrusion Detection. IEEE Access, 1–1.

14. Zhang, Y., Li, P., & Wang, X. (2019). Intrusion Detection for IoT Based on Improved Genetic Algorithm and Deep Belief Network. IEEE Access, 1–1.

15. Khan, I. A., Pi, D., Khan, Z. U., Hussain, Y., & Nawaz, A. (2019). HML-IDS: A Hybrid-Multilevel Anomaly Prediction Approach for Intrusion Detection in SCADA Systems. IEEE Access, 7, 89507–89521.

16. Duan, T., Tian, Y., Zhang, H., Liu, Y., Li, Q., Jiang, J., & Shi, Z. (2020). Intelligent Processing of Intrusion Detection Data. IEEE Access, 8, 78330–78342.

17. Jiang, K., Wang, W., Wang, A., & Wu, H. (2020). Network Intrusion Detection Combined Hybrid Sampling With Deep Hierarchical Network. IEEE Access, 8, 32464–32476.

18. Dutt, I., Borah, S., &Maitra, I. K. (2020). Immune System Based Intrusion Detection System (IS-IDS): A Proposed. IEEE Access, 8, 34929–34941.

19. Xiao, Y., Xing, C., Zhang, T., & Zhao, Z. (2019). An intrusion detection model based on feature reduction and convolutional neural networks. IEEE Access, 1–1.