A Hybrid Network Anomaly Detection system using Glowworm Swarm Optimization with Principal Component Analysis

Rashmita Khilar
Saveetha School of Engineering
K. Mariyappan
Jain University Faculty of Engineering & Technology
Mary Subaja Christo
SRM Institute of Science and Technology
J Amutharaj
RajaRajeswari College of Engineering
Anitha T
Saveetha School of Engineering
Rajendran Thavasimuthu (rajendran@makeittech.in)
Makeit Technologies

Research Article

Keywords: Anomaly Detection, IoT, Intrusion Detection System, Glowworm Swarm Optimization, Principal Component Analysis, NSL-KDD dataset

Posted Date: May 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-408246/v1

A Hybrid Network Anomaly Detection system using Glowworm Swarm Optimization with Principal Component Analysis

Rashmita Khilar[1], K. Mariyappan[2], Mary Subaja Christo[3], J Amutharaj[4], Anitha T[1] and Rajendran T[5]

[1]Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India.

[2]Department of Computer Science and Engineering, Jain University, Bangalore.

[3]Department of Computer Science, School of Computing, SRM Institute of Science and Technology, Kattankulathur, India.

[4]Department of Information Science & Engineering, RajaRajeswari College of Engineering, Bangalore, India.

[5]Makeit Technologies (Center for Industrial Research), Coimbatore, India.

Abstract

The security of the network is a significant issue in any distributed system. For that intrusion detection system (IDS), have been proposed for securing the network from malicious activities. This research is proposed to design and develop an anomaly detection model for detecting attacks and unusual activities in IoT networks. The primary objective of this research is to design efficient IDS for IoT network. The intrusion detection plays an essential role in detecting different attacks on IoT and enhances the performance of the IoT. In this research, anomaly detection in IoT networks using glowworm swarm optimization (GSO) algorithm with principal component analysis (PCA) is proposed. However, the proposed model is metaheuristic algorithm-based anomaly detection model to identify attacks by using the NSL-KDD dataset. The GSO algorithm based on PCA is implemented to perform the anomaly detection. For feature extraction, the PCA is used, and for classification, the GSO algorithm is used. For performance analysis, various parameters like accuracy, precision, recall, detection rate and FAR are evaluated. For normal class the proposed model achieved 94.14% accuracy, for DoS 95.52%, for R2L 93.15%, for probe 93.50% and for U2R 88.62% accuracy. Overall the detection rate was 94.08% and FAR was 3.41%.

Keywords: Anomaly Detection; IoT; Intrusion Detection System; Glowworm Swarm Optimization; Principal Component Analysis; NSL-KDD dataset;

1. Introduction

The concept of Internet of Things (IoT) was created by a member from the Radio Frequency Identification Development (RFID) group in 1999, and it has presently turned out significant to the practical world for development of mobile phones, embedded and universal communication, cloud computing and data analysis [1]. The IoT assumes a significant part in all aspects of our day-to-day lives. It covers several domains including industrial appliances, automobiles, healthcare, sports, entertainment, smart homes, and so on. The prevalence of IoT facilitates some daily activities, enhances the manner in which humans collaborate with the world and environment, and expands our social communications with others and objects [2-3].

The concept behind the IoT is to connect not just humans and computers as well as day-to-day objects to the Internet. This could be accomplished by outfitting things with computing and communication capacities hence altogether mapping the physical world to the digital one [4]. The initial years of the IoT mainly included data communication through machine to machine (M2M) communication. Though, the idea has developed quickly to incorporate human communication also, introducing a generation of Internet-of-Everything (IoE) [5]. However, this comfort comes at the cost of privacy and security challenges: the private, customized data, if access to an unapproved, malicious operator, can lead to critical harm to our wealth, status, and personal security. Moreover to our very own information, these devices additionally incorporate resources presented by their producers at different stages during their production distribution chain. These comprises fuses, firmware, and troubleshoot modes. Unapproved access to these resources can lead to the loss of a million of dollars in stolen copyrights, just as possibly critical exploitation of the resources. With the worldwide implementation of these devices, such security vulnerabilities could be disastrous [6-7].

In this research, anomaly detection in IoT networks using glowworm swarm optimization (GSO) algorithm with principal component analysis (PCA) is proposed. Normally GSO algorithm is mainly used for image classification and optimization. In this work, it is used for intrusion detection in IoT networks. For feature extraction process of dataset the PCA is used. However, in real cases, every feature differently affects the decision. Due to this fact, this work initially extracted primary features from the dataset using PCA, for decreasing the data dimension, and reducing the duration of classification.

This research is proposed to design and develop an anomaly detection model for detecting attacks and unusual activities in IoT networks. However, the proposed model is metaheuristic algorithm-based anomaly detection model to identify attacks by using the NSL-KDD dataset. The GSO algorithm based on PCA is implemented to perform the anomaly detection. For feature extraction, the PCA is used, and for classification, the GSO algorithm is used.

The remaining part of this work as follows: section 2 discusses the related work done on anomaly detection in IoT using different techniques, section 3 discusses the proposed methodology, section 4 discusses the obtained outcomes, and section 5 is conclusion of the work.

2. Related Works

Dang Hai Hoang and Ha Duong Nguyen proposed an anomaly detection model for IoT network traffic using PCA method. PCA method was used for reducing higher data dimension. A new distance formula was proposed and implemented to derive formulas from past works. Based on those derivations a new technique for anomaly detection in network traffic was implemented and obtained appropriate results using new distance formula by reducing the computational overhead [8].

Ren-Hung Hwang et al. proposed an early network traffic anomaly detection using unsupervised deep learning model that included both CNN and autoencoder. Anomaly traffic detection scheme called D-PACK was used for auto-profiling the patterns of traffic and abnormal traffic filtering. But the D-PACK inspected just the initial few bytes of the initial few packets in every flow for early detection. USTC-TFC and Mirai-CCU were the datasets used for experiments. Finally by examining less packets and total bytes from every packet feasible, the model majorly reduced the traffic volume for processing with 100% accuracy and less than 1% false positive and false negative rates [9].

Zhaomin Chen et al. presented the network anomaly detection model using the autoencoder method. Convolutional autoencoder (CAE) was additionally used here for the dimensionality reduction. As we know, this CAE require less training time compared with regular autoencoder. This model can obtain the non-linear correlations among features in order to improve the accuracy of detection. Here NSL-KDD data set was utilized for evaluation and the CAE model performed better in detection [10].

Yansen Zhou and Jinwei Li proposed Multilevel Autoregression technique for network traffic anomaly detection model on the basis of information entropy. DDOS attack detection was the main consideration using this autoregression model. For improving the detection rate of network traffic, information entropy and multilevel autoregression models were used. The model initially calculated the network traffic's information entropy per unit time and used zero-mean for obtaining the information entropy's time series, and then used a multi-level autoregression approach for predicting sequence entropy, divide the residual by the residual average value between actual values of entropy and prediction, and assessed whether an anomaly existed. This model can be used to identify unknown anomaly traffic [11].

Maryam A and Keivan B used metaheuristic algorithms like genetic, particle swarm and glowworm optimization for IDS. DoS and DDoS were the attacks applied to these algorithms and lifetime of nodes in WSN was assessed. In terms of energy consumption, genetic algorithm was better, and in terms of permittivity, the PSO algorithm was higher. The GSO algorithm has low performance in permittivity and achieved higher energy consumption [12].

3. Proposed Methodology

The security of the network is a significant issue in every distributed system. For that IDS have been used to protect the network against malicious activities. For detecting unauthorized intrusion scenarios, IDS is used to help by adding a protection layer over the networks. For IDS two main techniques are anomaly-based and signature-based detection. Most of the IDS are signature-based systems that use detection rules. However, for IDS a large distributed network should need a lot of rules, which can be time-consuming and costly. In addition, IDS utilizes signatures of an attack to identify malicious activities. If the signatures were not described adequately, then intruders may gain the network access. Anomaly-based systems are not dependent on human intercession which has been presented to overcome these problems.

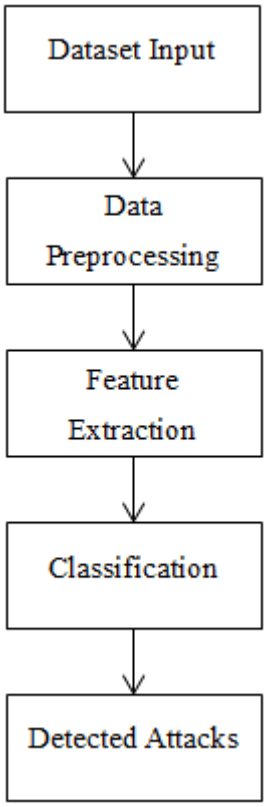

Figure.1. Proposed Model

The Network-based IDS (NIDS) and Host-based IDS (HIDS) are the two types based on the concept of intrusion detection. The HIDS is not appropriate for some of the IoT device applications that have limited resources and functionalities. The NIDS can monitor total network traffic and identify known, and unknown attacks depended on a hybrid technique that had both anomaly-based and signature-based methods.

In general, the anomaly-based NIDS techniques are very important for monitoring network traffic and for detecting new attacks. The signature-based NIDS is not suitable to detect a new attack in future network traffics. Therefore the proposed model is developed to detect anomaly-based NIDS-based attacks using a GSO algorithm with PCA.

3.1 Principal Component Analysis

PCA is a technique for dimension reduction, which converts a set of actual correlative variables into a group of some uncorrelated variables, namely Principal Components (PC). These PCs are the original variable's linear combinations. The total PCs obtained are smaller than the total actual variables, or equal. PCA thus enables for a lower complexity. It was considered as uncomplicated but efficient technique for the detection of network-based anomalies.

The fundamental concept of PCA was to create a linear transformation series for the actual data which has specific related features. Hence, the principal component load (PCL) matrix changed into the group of new data with minimum attributes to signify the real data. It is appropriate for the process of multidimensional data dimensionality reduction.

Step1: Determine the observation matrix Z for real data. After N times, observation for x variables θ_1, θ_2, ... θ_x determine matrix Z. Every row depicts a numerical estimation of sample data of the dataset; column number n regards n samples were discovered as per Equation (1).

$$Z = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1x} \\ Z_{21} & Z_{22} & \cdots & Z_{2x} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n1} & Z_{n2} & \cdots & Z_{nx} \end{bmatrix} \quad (1)$$

Step2: For the observation matrix, centralize the data process. Compute both the sample mean

$$\overline{z_b} = \frac{1}{n}\sum_{a=1}^{n} z_{ab} \quad (2)$$

And standard deviation

$$S_b = \sqrt{\frac{1}{n}(z_{ab} - \bar{z_b})^2} \quad (3)$$

Based on the formula

$$\widetilde{z_{ab}} = \frac{z_{ab} - \bar{z_b}}{S_b} \ (a = 1,2,\dots,n, b = 1,2,\dots,x) \quad (4)$$

Execute the centralized data processing and form the standardized matrix $\tilde{Z}$.

Step3: Compute sample correlation matrix based on the formula,

$$W = \frac{1}{n}\tilde{Z}^T\tilde{Z} \quad (5)$$

The computation of every element in W is as per Equation (6),

$$w_{ab} = \frac{\sum_{k=1}^{n}(z_{ka} - \bar{z}_a)(z_{kb} - \bar{z}_b)}{\sqrt{\sum_{k=1}^{n}(z_{ka} - \bar{z}_a)^2 \sum_{k=1}^{n}(z_{kb} - \bar{z}_b)^2}} \quad (6)$$

Step4: Compute the eigenvector and eigenvalue of W. Obtain x characteristics values $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_x \geq 0$ of W. Hence compute every principal component's rate of contribution using the formula

$$r_a = \frac{\lambda_a}{\lambda_1 + \lambda_2 \dots + \lambda_x}(a = 1,2,\dots,x) \quad (7)$$

And choose highest p principal component that attain 85% and encounter $\lambda_{p+1} < 1$ as PCA outputs. Descending order characteristics values are $\lambda_1, \lambda_2, \dots, \lambda_x$, compute the related eigenvectors $e_1, e_2, \dots, e_x$. Choose top p feature vectors for forming the PCL

$$L_{x \times p} = (e_1, e_2, \dots, e_p) \quad (8)$$

Step5: Create linear transformation for the real data following the PCL matrix $L_{x \times p}$ following Equation (9), to create new variables of principal component $w_1, w_2, \dots, w_p$.

$$\begin{bmatrix} w_1 \\ \vdots \\ w_p \end{bmatrix} = L_{x \times p}^T \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{bmatrix} \quad (9)$$

The matrix dimension decreased from x to p after the linear transformation and hence, significantly decreases the sample data volume.

3.2. Glowworm Swarm Optimization

The GSO was created from the concept of glowworms courtship behaviour or foraging, which glows to draw companions. Each glowworm in GSO has a specific amount of luciferin that determines the luminance intensity. In a process of operation with an algorithm, every glowworm moves towards its neighbour, which was luminous than itself. These activities were only based on local data that allows the swarm to divide into disjoint sub groups which combine to multimodal function through multiple optima. GSO has four primary stages: glowworm distribution, luciferin updation, glowworm movement, and neighborhood range updation [13].

Glowworm Distribution

A group of n glowworms are distributed in various locations of the search space randomly. Each glowworm carries the same amount of luciferin l_o.

Luciferin Updation

This stage is based on the value of past individual luciferin level and objective function. The luciferin updation rule was

$$l_g(i) = (1 - \rho)l_{g-1}(i - 1) + \gamma H(x_g(i)) \quad (10)$$

where $l_g(i)$ indicates the glowworm g's luciferin value at the i^{th} iteration, γ and ρ were the luciferin enhancement factor and decay, and $H(x_g(i))$ was the objective function value at the glowworm g's location.

Glowworm Movement Stage

During this stage, every glowworm, based on a probabilistic method, draw near a neighbour that has more luminance than itself. For every glowworm g, the moving toward a neighbour h probability was

$$p_{gh}(i) = \frac{(l_h(i) - l_g(i))}{\sum_{j \in N_g(i)}(l_j(i) - l_g(i))} \quad (11)$$

where, $h \in N_g(i)$, $N_g(i)$ is a set, which could be confirmed by

$$h \in N_g(i), N_g(i) = \{h: d_{gh}(i) < r_d^g(i); l_g(i) < l_h(i)\} \quad (12)$$

where $d_{gh}(i)$ indicates the Euclidean distance among glowworms g and h at the i^{th} iteration and $r_d^g(i)$ indicates the variable neighbourhood range related with glowworm g at the i^{th} iteration. Thus the glowworm movement model was

$$x_g(i+1) = x_g(i) + si * \left(\frac{x_h(i) - x_g(i)}{\|x_h(i) - x_g(i)\|}\right) \quad (13)$$

where $x_g(i) \epsilon R^m$ was the glowworm g's location at the i^{th} iteration in the m-dimensional real space R^m, $\|x_h(i) - x_g(i)\|$ indicates the Euclidean norm operator, and $si(>0)$ was the step size.

Neighborhood Range Updation

Considering r_0 as a first neighbourhood domain for every glowworm, the neighbourhood domain of every glowworm was presented in the generation:

$$r_d^g(i+1) = \min\left\{r_s, max\{0, r_d^g(i) + \beta(n_i - N_g(i))\}\right\} \quad (14)$$

where β was a constant, n_i was a control parameter for the number of neighbourhoods, r_s indicates the glowworms sensory radius, and $N_g(i)$ indicates neighbourhoods set.

GSO Algorithm

Initialize

Set the generation $A = 1$; problem dimension = m; population size = n;

Step size = $si(0)$; initial luciferin = l_0; initialization parameter β; γ and $r0$

Glowworms distribute

Glowworms are distributed randomly in search space.

Each glowworm carries same level of luciferin l_0 and on the same initial neighbourhood domain radius r_0

While the $A <$ max *generation* do

for $g = 1$: n (all glowworms) do

Update luciferin based on Equation (10)

Verify neighbours set based on Equation (12);

Calculate the probability of movement based on Equation (11);

Choose a neighbour h using probabilistic mechanism;

Glowworm g moves toward h based on Equation (13);

Update neighbourhood range based on Equation (14);

End for

End while

Output and algorithm end

In this research, anomaly detection in IoT networks using glowworm swarm optimization (GSO) algorithm with principal component analysis (PCA) is proposed. Normally GSO algorithm is mainly used for image classification and optimization. In this work, it is used for intrusion detection in IoT networks. For feature extraction process of dataset the PCA is used. However, in real cases, every feature differently affects the decision. Due to this fact, this work initially extracted primary features from the dataset using PCA, for decreasing the data dimension, and reducing the duration of classification.

4. Performance Analysis

In MATLAB 2017a, the implementation and assessment of the proposed model are carried out on a personal computer with a Core I5 3.20 GHz CPU and 4 GB RAM. Using the result parameters such as precision, recall, accuracy, FAR, and detection rate, the proposed method will be assessed. The performance analysis will be compared with other techniques such as ANN, SVM, BPNN, and PSO for the proposed GSO-PCA technique.

4.1 Dataset Description

National Security Lab-Knowledge Discovery and Data-mining (NSL-KDD) was improved from KDD99 dataset, designed to surpass KDD99's restrictions. The dataset is open to the public from its official site.

In the training and test sets, duplicated records were initially removed. Then, numerous records are selected from the actual KDD99 for obtaining accurate results for classifier systems. Third, it eliminated the problem of an unbalanced distribution of probabilities. In the NSL-KDD data collection, there are 125,973 training cases and 22,544 test cases, with 41 attributes, 38 consistent, and 3 categorical (discrete-valued). Six continuous variables have been dismissed as being significantly 0s. There are 23 possible labels in the training data set (Normal + 22 labels linked to different forms of intrusion); at the same time, there are 38 groups in the test data set, indicating that there is no intrusion into the test information during training. Overall, there are 21 classes in the 23 training classes and 38 testing classes; two classes only take place on the training and 17 classes are odd for test knowledge. Approximately 16.6% test data set samples are classes that are excellent for the test dataset and are not present during the training. This distribution variation of class gives the classification more complexity. The training/testing classes are linked to one of five possible classes: Normal, PROBE, R2L, U2R, and DoS. Each classification, apart from the Normal class, is equivalent to an intrusion that infers that no anomalies have been presented. These classes are still useful in the concept of IDS and still highly unbalanced and contain a number of cases large enough to generate more significant results in each class.

Probe: when an intruder attempts to acquire information inside the objective network through the network and host scanning behaviour (i.e., ports scanning).

DoS: when an intruder impedes the login of actual users to the particular service or machine.

U2R: when an intruder attempts to extend a minimal user's advantage to root access (by malware injection or stolen data).

R2L: when an intruder gains access to a target computer wirelessly imitating current local users. Among the most offensive attacks to be described as mimicking usual user actions were the U2R and R2L attacks.

Table.1. NSL-KDD Traffic Distribution

Traffic	Training	Test
Normal	67343	9711
DoS	45927	7458
Probe	11656	2754
R2L	995	2421
U2R	52	200
Total	125973	22544

4.2 Performance Metrics

The accuracy was just the performance subset of the model. It is one of the performance metrics for evaluating classification techniques. The accuracy computation is calculated using the following expression (15).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

Precision is a positive predictive rate. It is the ratio of properly predicted positive observation of the total predicted positive value. The calculation of precision is calculated using the following expression (16):

$$Precision = \frac{TP}{TP+FP} \quad (16)$$

The recall is also termed as sensitivity. It is the ratio of properly predicted positive value to each observation in the actual class. The computation of recall is calculated using the following expression (17):

$$Recall = \frac{TP}{(TP+FN)} \quad (17)$$

The detection rate (DR) represents the level of intrusion instances. It represents the total proper positive class predictions made as to the proportion of all the predictions made. The calculation of the DR is computed using the accompanying condition (18):

$$DR = \frac{TP}{TP+FN} \quad (18)$$

The Equation (19), False Alarm Rate (FAR) is the ratio of normal data identified as activity of attack falsely. It is also called as ratio of false detection.

$$FAR = \frac{FP}{TN+FP} \quad (19)$$

The evaluation of the proposed model is evaluated based on the above criteria depended on the identification of the attacks from the dataset. In each case, accuracy is the measure of correct identification; DR implies the classifier’s detection rate

of attacks; FAR means the proportion of normal instances misclassified; recall reflects the attack returns of the model. The precision determines the attacks returned are right. The performance evaluation of various outcome metrics must be tested in terms to validate the GSO-PCA model and to compare with other current methods.

Table.2. Proposed model result on traffic distribution of dataset

Attack	ACC	FAR	Precision	Recall
DoS	92.22	1.2	94.22	95.71
Probe	94.89	1.4	93.47	93.45
R2L	90.41	8.7	89.74	91.12
U2R	92.68	3.1	92.30	93.79

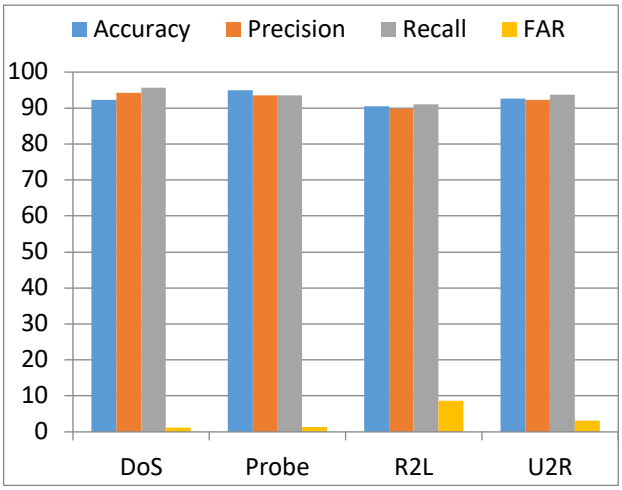

Figure.2. Performance Analysis of Proposed Model

Table.2 represents the performance evaluation of the GSO-PCA method executed on the NSL-KDD dataset. Where, every class of attack is assessed in terms of accuracy, precision, recall, and FAR. Fig.2 presents the graphical plot of the performance analysis of proposed model.

As mentioned earlier, the performance analysis of the proposed model was compared with other existing techniques like artificial neural network (ANN), support vector machine (SVM), back-propagation neural network (BPNN), and particle swarm optimization (PSO) as shown in table.3. Here, classification accuracy is compared and validated following the graphical plot in fig.3.

Table.3. Comparison of Classification Accuracy

Method	Normal	DoS	R2L	Probe	U2R
ANN	90.40	89.23	80.87	89.10	62.02
SVM	91.72	85.37	88.21	86.75	78.85
BPNN	93.24	91.07	92.80	93.98	84.38
PSO	92.45	94.30	90.53	93.27	86.26
GSO-PCA	94.14	95.52	93.15	93.50	88.62

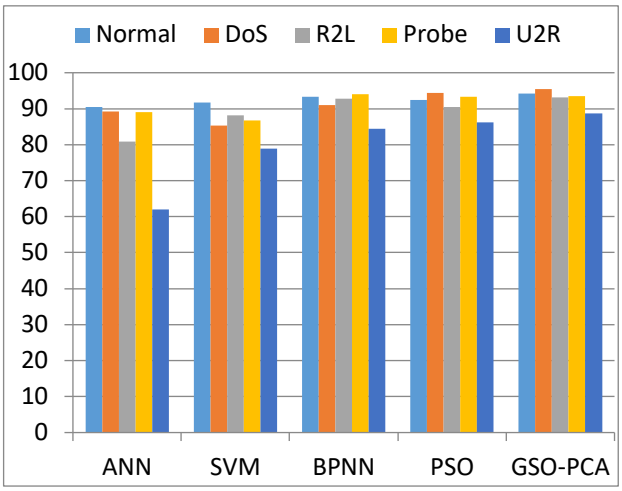

Figure.3. Classification Accuracy Comparison

The proposed GSO-PCA model outperforms in most of the performance compared with the other techniques. For normal class, the proposed model obtained 0.9 to 3.7% more detection, in DoS class 1.2 to 10.15% higher detection, in R2L 2.7 to 14% more detection, in probe class BPNN performed better than proposed method, other than BPNN the GSO-PCA model achieved 0.2 to 6.75% more accurate detection, and in U2R class 2.3 to 26% higher detection obtained.

Table.4. Detection Rate and FAR Comparison

Methods	Detection Rate (%)	FAR (%)
ANN	86.52	5.06
SVM	88.30	4.73
BPNN	91.84	4.55
PSO	93.60	3.98
GSO-PCA	94.08	3.41

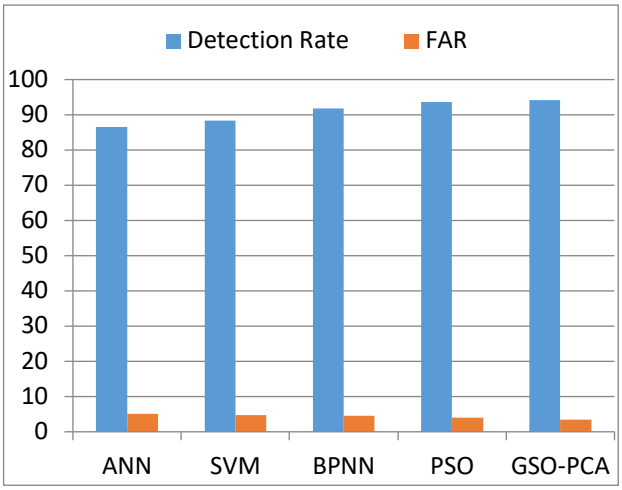

Figure.4. Comparison of Detection Rate and FAR

The comparison of DR and FAR is evaluated in table.4. Fig.4 represents the graphical plot of the comparison respectively. The proposed model achieved better detection rate and FAR rate compared with other techniques. The PSO performed close to the results of the proposed model in both the detection rate and FAR. The lowest performance was obtained by ANN with 86.52% DR and 5.06% FAR and the highest performance was obtained by the proposed model with 94.08% DR and 3.41% FAR. The GSO-PCA model has achieved 0.4 to 7.5% more detection rate and 0.5 to 1.6% lesser FAR.

5. Conclusion

In this research, anomaly detection in IoT networks using glowworm swarm optimization (GSO) algorithm with principal component analysis (PCA) was proposed. Generally, GSO algorithm was mainly used for image classification and optimization. In this work, it was used for intrusion detection in IoT networks. For feature extraction process of dataset the PCA was used. The proposed model was developed to achieve better performance in detecting anomalies in IoT networks. For attack classification and detection the NSL-KDD dataset was used. By using this dataset the proposed model was trained and tested for performance evaluation. The features of the dataset are effectively extracted by the PCA and the GSO was used to classify and detect the different class of attacks present in the dataset. For performance analysis various parameters like accuracy, precision, recall, detection rate and FAR are evaluated. For normal class the proposed model achieved 94.14% accuracy, for DoS 95.52%, for R2L 93.15%, for probe 93.50% and for U2R 88.62% accuracy. Overall the detection rate was 94.08% and FAR was 3.41%. It is not the best performance achieved overall, but compared with the other current methods like ANN, SVM, BPNN, and PSO, the proposed model performed well in every parameter. In future, the proposed model can be implemented using different dataset in different network platforms like WSN, MANET, etc. and to improve the detection rate and reduce the FAR, a new hybrid model using deep learning technique can be developed.

***FUNDING**

NIL

***CONFLICTS OF INTEREST/COMPETING INTERESTS**

DECLARED NONE

***AVAILABILITY OF DATA AND MATERIAL**

ALL THE DATA ARE AVAILABLE WITHIN THE MANUSCRIPT CONTENTS

***CODE AVAILABILITY**

ALL THE DATA ARE AVAILABLE WITHIN THE MANUSCRIPT CONTENTS

***AUTHORS' CONTRIBUTIONS**

[1]Rashmita Khilar – is responsible for data collection, and proof reading [2]K. Mariyappan – is responsible for data curation and programming, [3]Mary Subaja Christo – is responsible for programming and survey, [4]J Amutharaj - is responsible for programming and survey, [1]Anitha T - is responsible for programming, data validation, and writing and [5]Rajendran T – is responsible for programming, data validation and survey

REFERENCES

[1] Mehdi N, Vijay S, and Roksana B, A Host-based Intrusion Detection and Mitigation Framework for Smart Home IoT using OpenFlow, 11th International Conference on Availability, Reliability and Security (ARES) IEEE, 2016, pp. 147-156.

[2] Mansour S and Hamid B, A Hybrid Intrusion Detection Architecture for Internet of Things, International Symposium on Telecommunications (IST'2016), IEEE, 2016, pp.601-606.

[3] Mohammed H A, Bahaa A D A M, Alyani I, and Mohamad F Z, A New Intrusion Detection System Based on Fast Learning Network and Particle Swarm Optimization, IEEE Access, Vol.6, 2016, pp.20255-20261.

[4] Sailaja M, Kiran K R, Sita R M P, Krishna P P.E.S.N, A Novel Approach for Intrusion Detection Using Swarm Intelligence, Proceedings of the InConINDIA, Springer, Vol.132, 2012, pp.469-479.

[5] Nilesh K S and Indrajit M, Machine Learning based anomaly detection for IoT Network, Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020), IEEE, 2020, pp.787-794.

[6] Sohaila E, May B, Noora A N, Zina C, and Aiman E, Machine Learning Techniques for Network Anomaly Detection: A Survey, IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), 2020, pp.156-162.

[7] Andrew C, Goksel M, and Zhong F, Anomaly Detection for IoT Time-Series Data: A Survey, IEEE Internet of Things Journal, Vol.7, No.7, 2020, pp.6481-6494.

[8] Dang H H and Ha D N, A PCA-based Method for IoT Network Traffic Anomaly Detection, International Conference on Advanced Communications Technology(ICACT), 2018, pp.381-386.

[9] Ren-H H, Min-C P, Chien-W H, Po-C L, and Van-L N, An Unsupervised Deep Learning Model for Early Network Traffic Anomaly Detection, IEEE Access, Vol.8, 2020, pp.30387-30399.

[10] Zhaomin C, Chai K Y, Bu S L, and Chiew T L, Autoencoder-based Network Anomaly Detection, Wireless Telecommunications Symposium (WTS), 2018, pp.1-5.

[11] Yansen Z and Jinwei L, Research of Network Traffic Anomaly Detection Model Based on Multilevel Autoregression, IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT), 2019, pp.380-384.

[12] Maryam A and Keivan B, Using Metaheuristic Algorithms of Genetic, Particle Swarm Optimization and Glowworm in the Intrusion Detection System, International Journal of Computer Science and Network Security, Vol.16 No.10, 2016, pp.78-86.

[13] Zhonghua T and Yongquan Z, A Glowworm Swarm Optimization Algorithm for Uninhabited Combat Air Vehicle Path Planning, Journal of Intelligent Systems, Vol.24, No.1, 2015, pp.69-83.

Figures

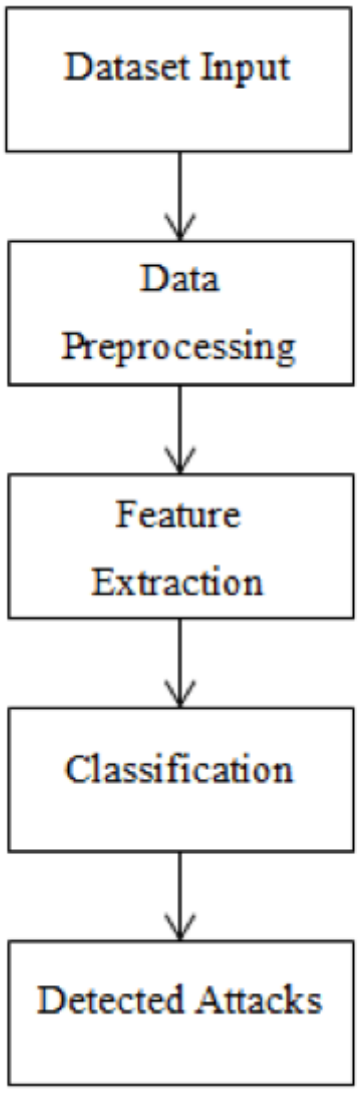

Figure 1

Proposed Model

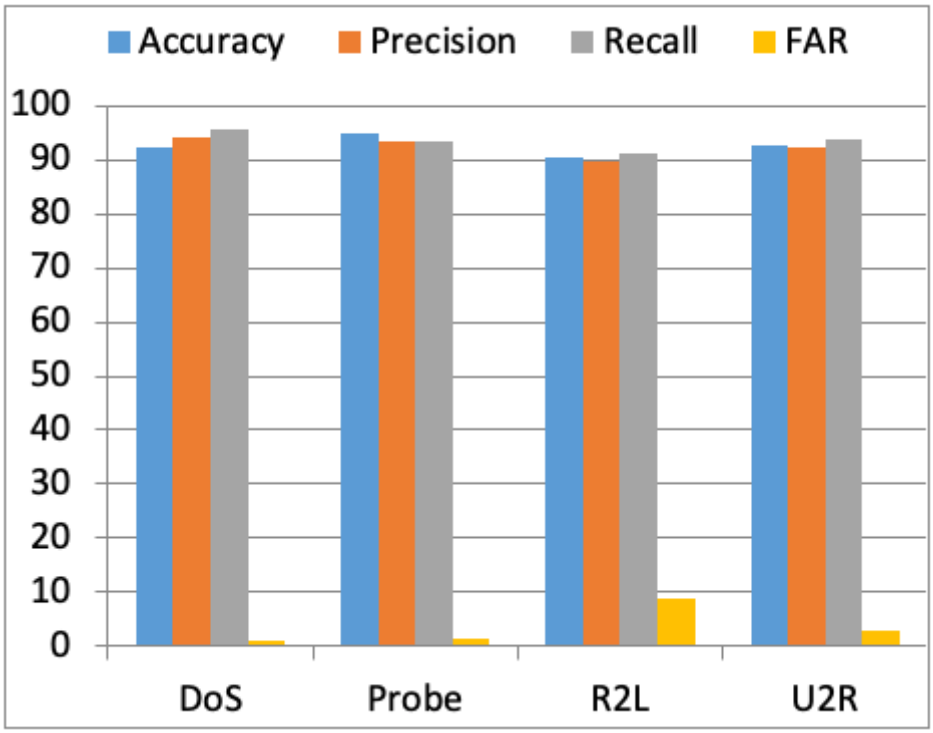

Figure 2

Performance Analysis of Proposed Model

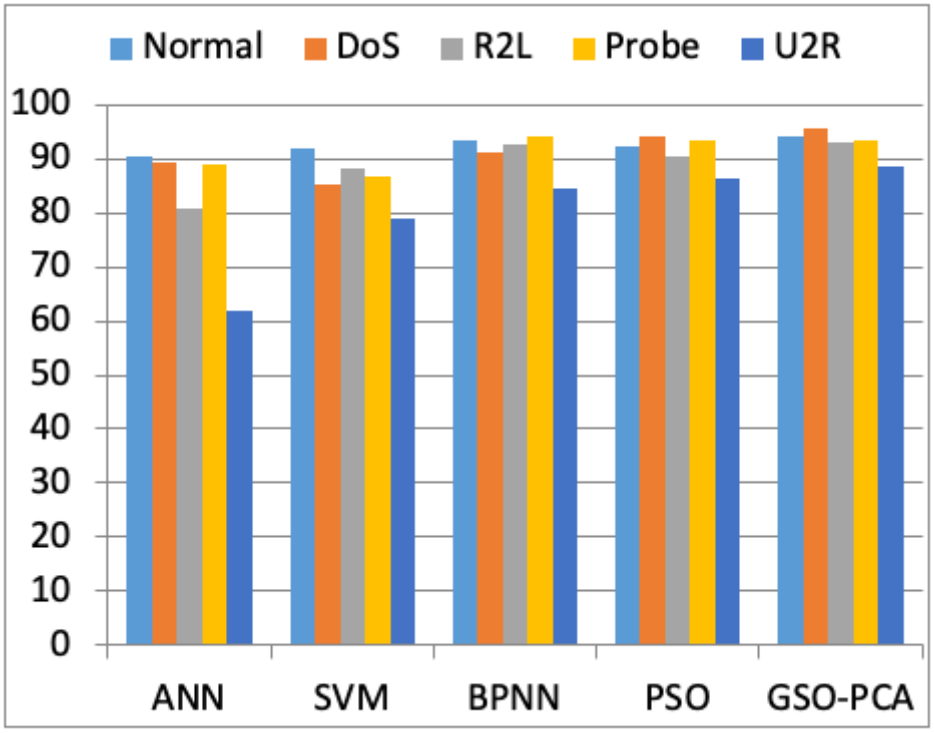

Figure 3

Classification Accuracy Comparison

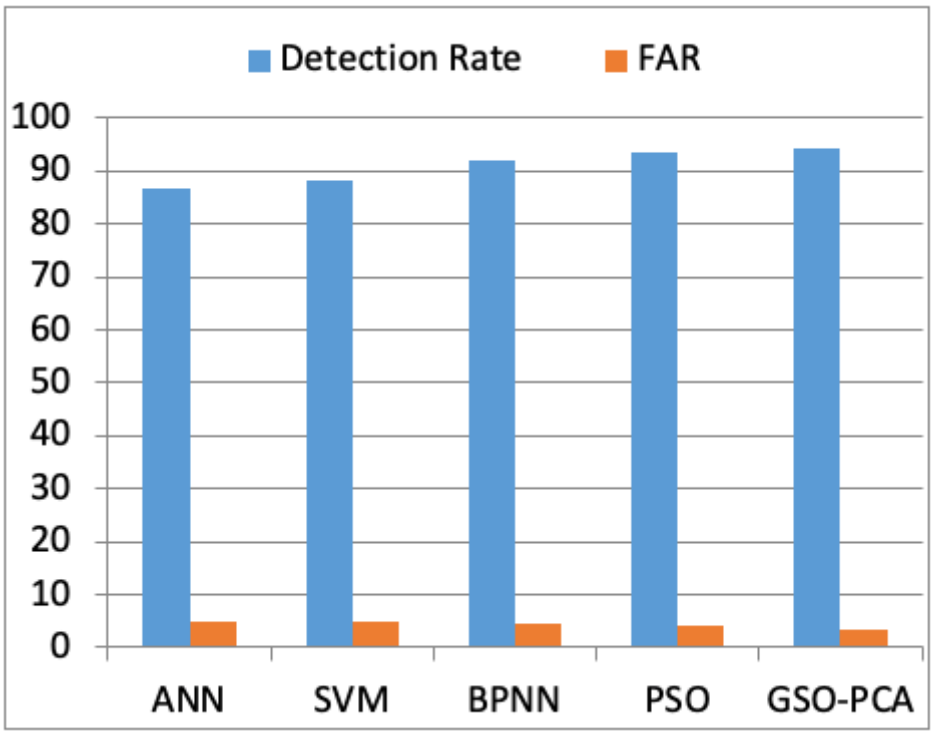

Figure 4

Comparison of Detection Rate and FAR