Back

RESEARCH ARTICLE | Full Access

Intrusion detection framework using auto-metric graph neural network optimized with hybrid woodpecker mating and capuchin search optimization algorithm in IoT network

Shanthi Govindaraju, Wilson Vimala Rani Vinisha, Francis H. Shajin, D. Adhimuga Sivasakthi

First published: 21 September 2022
https://doi.org/10.1002/cpe.7197
Citations: 1

Summary

Intrusion detection systems (IDSs) are the major component of safe network. Due to the high volume of network data, the false alarm report of intrusion to the network and intrusion detection accuracy is the problem of these security systems. The reliability of Internet of Things (IoT) connected devices based on security model is employed to protect user data and preventing devices from engaging in malicious activity. In this article, intrusion detection framework using auto-metric graph neural network optimized with hybrid woodpecker mating and capuchin search optimization algorithm in IoT Network (IDF-AGNN-HYB-WMA-CSOA- IoT) is proposed. Initially the attacks affected in the IoT data is taken from the dataset such as CSIC 2010 dataset, ISCXIDS2012 dataset, then these data are preprocessed and the features are extracted to remove the redundant information using improved random forest with local least squares. Then the malicious and the normal attacks are classified using the auto-metric graph neural network. At last hybrid woodpecker mating and capuchin search optimization algorithm (Hyb-WMA-CSOA) is utilized to optimize the weight parameters of AGNN. The performance of ISCXIDS2012 dataset of the proposed method shows higher accuracy 25.37%, 29.57%, and 18.67%, compared with existing methods, such as IDF-ANN-IoT, IDF-BMM-IoT and IDF-DNN-IoT respectively.

1 INTRODUCTION

With the rapid increase of IoT applications, there is an increase of new attacks day by day,[1] which distributed in the form of email, messages, advertisement, tweets, videos and so forth.[2-4] This information affects the perception of people in public and industry, which created people in abiased opinion.[5-7] Researchers use various mechanisms and technologies to predict the hybrid attacks, like Denial of Service,[8-13] poisoning, adversarial, phishing and

Back

Though the approaches used in these models are hard to attack, there are no measures provided once an attack is detected.[17] Generally, DoS attacks would violate the availability and integrity of data either in targeted or widespread mode.[18] Detection and prevention of DoS attacks, particularly adversarial attacks are critical,[19] since it is nonintrusive in nature, required perfect adversary knowledge (white box)[20] about the structure, targeted community and their social network.[21] Moreover, these attacks are initiated by the adversary utilizing similar channel (URL, port, etc.) as valid users.[22]

Typically, an adversary with perfect knowledge can attack the system successfully at any time[23] Hence, analyzing the datasets using digital DNA techniques helps to identify key behavioral features of the attacks (spam)[23] at the time of feature selection. But the models are needed to be designed in proactive rather than reactive manner,[24] since the reaction of detecting attacks would never prevent the future attacks.[25] Hence, anticipating and investigating of adversarial activities proactively enables some possible defense strategies.[26-28]

The task of ensuring security and network services availability to the users becomes a difficult task, because increases the count of hosts connected with internet.[29] Over the past two decades, numerous tools and strategies have been created by organizations towards the security of networks and systems against different security threats, like access control mechanisms, user authentication, firewalls.[30] Despite the fact that these methods prevent illegal access from outsiders, they are vulnerable to insider attacks. The attacks are created to activate second line of defense as well as protect information loss for intruders. It detects malicious network traffic and system activity that is undetectable by typical fi[31] This involves network attacks against vulnerable services, application attacks, hos namely, unauthorized logins and access to confidential data. Generally, the attacks are categorized into two types: (i) host-based, (ii) network based. Host-based attacks analyze data from a single system, like system calls, application logs, file access logs, file-system modifications. Besides, network-based attacks collect raw network packets through various network segments and analyze them for signs of intrusions in systematic manner.

The reliability of IoT connected devices is depending on security model for protecting user data along preventing devices from engaging in malicious activity. The malicious attacks are entered into the system due to hybrid attacks, like Denial of Service attack, poisoning attacks, adversarial attacks, phishing attacks, potential attacks, mirai attack, and so forth which interrupts and blocks legitimate users to use the network resources by flooding the targeted system with traffic. There are several methods are used to predict and detect these attacks from the IoT devices but does not provide sufficient accuracy, also increased the error rate. To overcome these issues, an auto-metric graph neural network optimized with hybrid

Back

The major contributions of this article are summarized below,

- An auto-metric graph neural network optimized with hybrid woodpecker mating and capuchin search optimization approach is proposed to detect and classify the attacks from the IoT devices.
- Initially, the attacks affected in the IoT data is taken from the two dataset such as CSIC 2010[32] dataset, ISCXIDS2012 dataset,[33] then these data are preprocessed and the features are extracted to remove the redundant information using improved random forest along Local least squares.
- Then the malicious attacks and the normal attacks are classified using the auto-metric graph neural network.[34]
- Hybrid woodpecker mating[35] and capuchin search optimization algorithm[36](Hyb-WMA-CSOA) utilized to optimize the weight parameters of auto-metric graph neural network (AGNN).
- The proposed work is executed in MATLAB platform.
- To validate the proposed method, the performance metrics namely, accuracy, F-score, precision are analyzed.
- Then the efficiency of the CSIC 2010 dataset are analyzed and compared with existing methods, like an unsupervised learning-based network threat situation assessment model for Internet of Things (IDF-ResNet-IoT)[37] a distributed deep learning system for web attack detection on edge devices (IDF-GAN-IoT)[38] and web attack detection system for internet of things via ensemble classification (IDF-EDL-IoT)[39] respectively.
- Then the efficiency of the ISCXIDS2012dataset is analyzed and compared with existing methods, like Attack detection with deep learning analysis (IDF-ANN-IoT),[40] statistical learning-enabled botnet detection framework for protecting smart cities networks (IDF-BMM-IoT)[41] and deep neural network in IoT intrusion detection (IDF-DNN-IoT)[42] respectively.

The remaining article is structured as: Section 2 presents the recent investigation works. Section 3 describes the proposed model. Section 4 demonstrates the results with discussion. Section 5 concludes the article.

2 RELATED WORKS

Some of the literature survey regarding intrusion detection framework for detecting security threats in IoT is reviewed here,

cloud deals the above challenges in the paradigm of the Internet of Things (IoT). Multiple concurrent deep models were enhanced the system stability and convenience in uIDFting. It provides higher accuracy conversely with lower precision.

Yang et al.,[38] have presented a deep learning-base network threat situation assessment model for IoT. Initially, it combines the encoder of variational autoencoder with the generative adversarial networks discriminator to create V-G network. The reconstruction error was obtained for every network layer by training the network collection layer of V-G network including normal network traffic. Moreover, the group threat testing was carried out with the test dataset and scale the threat probability of every test group. It provides lower accuracy with higher F-score.

Luo et al.,[39] have presented the web attack detection system for internet of things via ensemble classification. In this ensemble deep learning based web attack detection system to lessen the issues. Specially, three deep learning models detect web attacks individually. Use the ensemble classifier to make the final decision based on the results obtained from the three deep learning methods. The effectiveness of the proposed method is evaluated using CSIC 2010 dataset. Experimental outcomes display the presented system exactly detects the web attacks with high false positive and negative rates.

Pecori et al.,[40] have presented IoT attack identification along deep learning analysis. A large integrated dataset of IoT traffic flows was created utilizing four network scenarios. Such datasets verify the effectiveness of a deep learning network and compared its outcomes with ones derived through traditional deep learning methods. It provides lower F-score with higher accuracy.

Ashraf et al.,[41] have presented IoT based statistical learning-enabled botnet identification framework for securing smart cities networks. Where, IoTBoT-IDS was presented to smart networks depending on secure IoT against botnet attacks. IoTBoT-IDS capture the normal behavior of IoT networks by utilizing statistical learning-base strategies, beta mixture model, and Correntropy model. It provides lower detection accuracy with higher F-score.

Tsimenidis et al.,[42] have presented intrusion detection depending on deep neural network in IoT. The deep learning technique offers cutting edge solutions for internet of things intrusion detection and its data-driven, anomaly-base method and capability to detect emerging, unknown attacks. Solutions were categorized by model, structured analysis of how deep learning was employed for internet of things cyber security and its unique contributions for effectual IoT intrusion detection solutions development. It provides higher detection accuracy with lower precision.

3 PROPOSED WORK

preprocessing stage, in which redundancy removal and missing value replacement are carried out. Then preprocessed output is given to the auto-metric graph neural network for classification that categorizes the data as normal and anomalies. Figure 1 shows the overall block diagrams of proposed IDF-AGNN-HYB-WMA-CSOA-IoT method.

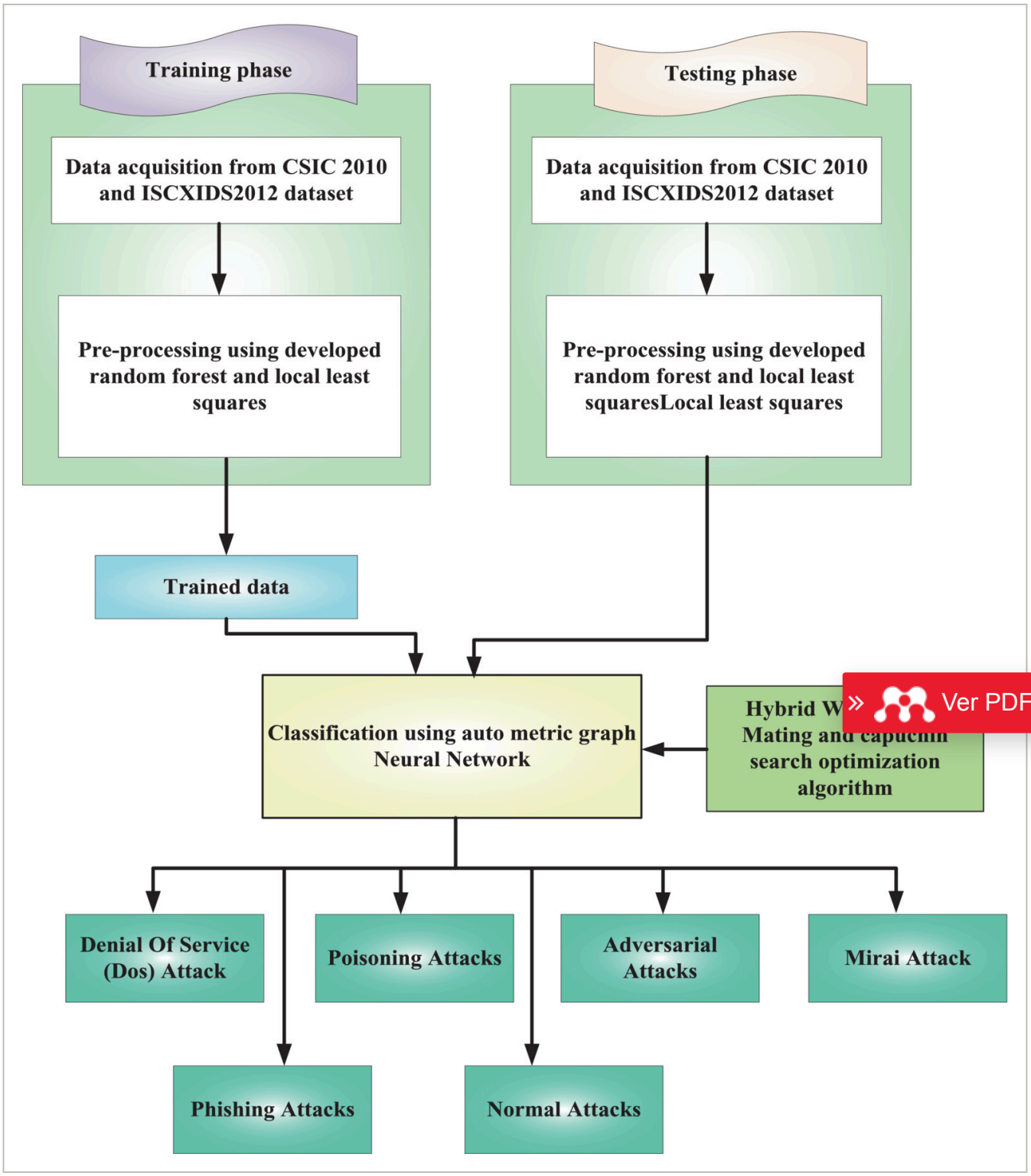


FIGURE 1

Back

3.1 Data set

For detecting attacks from the IoT devices, in this, two datasets are taken; are CSIC 2010 dataset, ISCXIDS2012 dataset. Here CSIC 2010 dataset consists of 43,000 data, ISCXIDS2012 is a binary class dataset gathered from New Brunswick University for examining intrusion detection system10, 000 data.

3.2 Preprocessing using improved random forest (IRF) with Local least squares (LLS)

The preprocessing depending on IRF with LLS is described in this section. The LLS consist of two stages: (i) utilize pearson correlation coefficients in the dataset to lessen duplicate with redundant data, (ii) Pearson correlation coefficients output is fed to subsequent stage, whereby, replaced the lost values through deterioration and assessment.

Let $d1$ record has m features together with x missing values. For recovering entire number of x on any location, found the nearest neighbor record vector for $d1$. The x components for each record has similar lost values location in $d1$ are ignored when similar records identifying procedure. The $M \in S^{p*(l-x)}$ matrix, here M indicates 2 dimensional matrix, by this, a number of rows equivalent to adjacent neighbors $p_i where \{i = 1, 2, \ldots, 7\}$, in which, number of columns similar to number of entire features l less the number of columns having lost values x.

Consider $N \in S^{p*x}$ matrix, where, N implicates 2 D matrix, the number of rows similar to number of adjacent neighbors $p_i where \{i = 1, 2, \ldots., 7\}$. The number of col… number of columns has missing values x. Consider $y = S\,(l-x)^{*1}$ vector, he… implies y, in which, number of columns similar to number of entire features l less the number of columns having missing values x. The least square problem is exhibited as,

$$\min_{q} \|M^T Q - y\|_2. \tag{1}$$

The vector $a = (\alpha_1, \alpha_2, \alpha_3, \ldots., \alpha_x)^T$ of x lost values is assessed by Equation (2),

$$a = \begin{pmatrix} \alpha_1 \\ . \\ . \\ \alpha_x \end{pmatrix} = N^T Q = N^T \left(M^T\right)^t y, \tag{2}$$

$$M^t = [v_1, v_2] \begin{bmatrix} \sum_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} [W_1 \ W_2]^T \\ = v_1 \sum_1^{-1} W_1^T, \tag{3}$$

where $v_1 \in S^{n-1*rank}$, $\sum_1^{-1} \in S^{rank*rank}$, $W_1^T \in S^{p*rank}$. By utilizing the given equation, the y known element is calculated.

$$y \approx q_1 m_1 + q_2 m_2 + \ldots . + q_p m_{pi}, \tag{4}$$

where q_i specifies linear combination coefficient initiates minimal squares that is articulated in Equation (2). As a result, the multi regressions categorize the target record in terms of dissimilar missing values features due to the linear combination, its nearest neighbors are delineated utilizing Equation (5)

$$Target = q_1 n_1 + q_2 n_2 + \ldots \ldots + q_p n_p, \tag{5}$$

where n_p is the p^{th} closer neighbor, the regression coefficients similar to neighbor implicates q_p. Wherein, it averts the classifier to bias towards recurrent records. This preprocessing state eliminates all superfluous data, also replaces missing values and eliminating every uncertainties from the dataset.

3.3 Classification using auto-metric graph neural network (AGNN)

In this section, auto-metric graph neural network (AGNN) used to classify the intrusion detection. In this the preprocessed data are given to the IoT environment for detect and classify the attacks from the IoT devices as normal and anomaly. Moreover, the proposed AGNN is directly used for categorizing the normal and anomaly attacks that involved initialization of graph structure, construction of layer, loss function, and IoT device type identification using meta-learning training process.

3.3.1 Initialization of graph structure

The sample selected from the network traffic dataset serve as the input of the AGNN model, which involves normal and anomaly attacks. Here, the input data is denoted as (R) includes (k) normal and (M) anomaly, which is represented in Equation (6)

$$R = \{[(p_1 l_1), \ldots \ldots, (p_{n-1}, l_{n-1})], [\bar{p}]; l_n \in (1, C)\}, \tag{6}$$

each category with anomaly attack. Moreover, the graph structure $g = (N, E, W)$ is initialized, where node set is (N), edge probability matrix is (E) and the weights are mentioned as (W). Additionally, the features of node p_i is calculated using Equation (7),

$$p_i = [l_i, \eta_i, r_i, m_i], \tag{7}$$

where risk factor is mentioned as η_i, label encoding is l_i, cognitive score is r_i and the features are represented as m_i. Here, l_i represent the zero vectors for the anomaly detection. Moreover, the initialization of graph is fully connected together with weights of the nodes in the network are equal to the value 1 that means weights and edge probability matrix for every element value is 1.

3.3.2 Construction of graph neural network layer

After the initialization of graph structure, this is given to the AGNN layer input through probability constraint. Here, the node features are uIDFted by transferring information between the nodes, which is used for obtaining the normal and anomaly attack in the IoT network. The probability of edge matrix is computed based on the risk factor, in which the weight matrix is automatically identified by features. Moreover, the calculation of probability matrix with risk factors $(\eta_1, \eta_2, \eta_3 \ldots\ldots \eta_N)$ is mentioned in Equation (8),

$$e_{i,j}^{k} = \begin{cases} 1 & if \quad \eta_i^k - \eta_j^k \ \leq \varphi \\ 0 & otherwise \end{cases}, \tag{8}$$

where the edge weights among nodes p_i and p_j is belongs to [0,1], the element $(e_{i,j})$ is mentioned as $e_{i,j} = \frac{1}{\left(K+1-\sum_{k=1}^{K} e_{i,j}^{k}\right)}$, (K) denotes the quantity of risk factors, risk factor feature is mentioned as (k) and the threshold value is mentioned as (φ) that is used for measuring the similarity of risk factor features among two different nodes, which is mentioned as $(f_1, f_2, f_3 \ldots\ldots\ldots\ldots f_N)$, $f \in S^d$ where S^d is the feature dimension. Additionally, the weights matrix between nodes with features that is calculated using Equation (9),

$$W_{i,j}^{(x)} = e_{i,j}\,(ab(f_i - f_j), \tag{9}$$

where $W_{i,j}^{(x)}$ weights between nodes with (x) number of layers in the network using absolute difference ab and f_i, f_j represents the feature set between the nodes. Subsequently, the nodes in the network is uIDFted for attaining accurate outcome using Equation (10),

where gn is the node uIDFting factor, P^x is the total number of nodes in the (x) number of network layers, $L-\mathrm{Re}LU$ denotes the nonlinear activation function and β_B^x denotes the training parameters used for changing the dimension of features of every node. Moreover, the outcome of this model is concatenated through P^x to preserve the input node features and the attained outcome of the AGNN layer is mentioned in Equation (11),

$$P^{(x+1)} = [P^x, gn(P^x)]. \tag{11}$$

After changing the feature dimension, the output of the device categories is inserted directly into the softmax layer to normalize the output, instead of being attached to the input nodes for the last layer.

3.3.3 Loss function

Additionally, the calculation of loss function of the AGNN model is used for uIDFte the parameters that is given in Equation (12),

$$lf(\widehat{M}, M) = -\sum_l m_l \log P\left(\widehat{M} = m_c \mid R\right), \tag{12}$$

where M denotes the prediction outcome of the known device categories, $\widehat{M}$ denotes the outcome of the unknown device categories, (R) implies input data, (C) denotes the count of classes.

The proposed AGNN model is developed based on graph structure, GNN lay function, which are initialized and creates the AGNN model. Here, the sample data is extracted to create the graph structure that is given to the layer, in which the data is transferred between the network nodes and that are uIDFted by AGNN model. Finally, the IoT devices type as known and unknown are identified. Additionally, the network parameters are uIDFted based on the loss function of the model that is given in Equation (13),

$$\beta_{B(t+1)} = S\left(lf_t, \beta_{B(t)}\right), \tag{13}$$

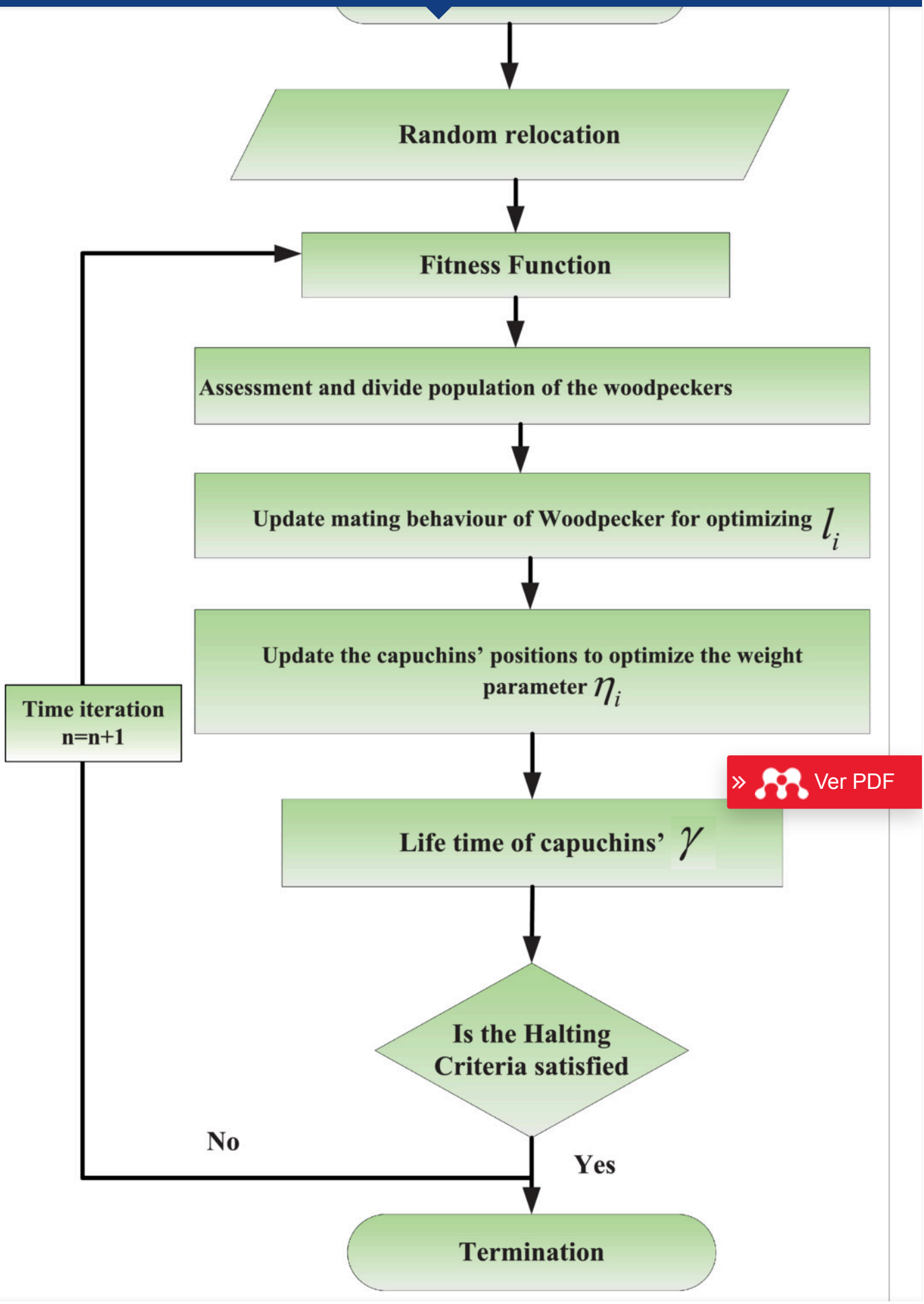
where S denotes the factor for uIDFting parameters β_B based on loss function $lf_{(t)}$ with training epoch (t). After performing the training epoch, the parameters are uIDFted and obtain the final outcome that is used for identifying the normal and anomaly in the new graph. The attained outcomes indicate that the proposed AGNN model has effectively identify the type of IoT devices. To get more accurate classification the weight parameters l_i and η_i of the AGNN method are optimized using hybrid woodpecker mating and capuchin

Back

3.4 Step by step procedure for hybrid woodpecker mating and capuchin search optimization

The woodpecker mating algorithm (WMA) and capuchin search optimization (CSO) is employed to optimize the parameters of AGNN model for getting the optimum parameters. These parameters are optimized for computing the optimal parameters for assuring accurate classification. WMA algorithm is metaheuristic algorithm which motivated via red-bellied woodpeckers mating behavior. The metaphor is the drumming sound of male woodpeckers attracts female. Woodpecker mating algorithm is prompted by the impression of sound wave intensity. This is defined as a cross algorithm. Woodpeckers are the noteworthy search factors, while males are the best positions they have already explored. In fact, female woodpeckers are not impressed by the sounds of drumming raised by males. So, Capuchin search optimization is used. Capuchin search optimization algorithm is one of the meta-heuristic methods focused on capuchin monkeys, like integrated primary method for social behavior of foraging animals. Typically, Capuchins goes far away for searching food sources jumping from one place to another by trees, also it slide to search food from the river banks. Therefore, jumping mechanism is same as global search. In this article, hybrid woodpecker mating algorithm and Capuchin search optimization algorithm is selected since it has its own improvement; good performance in solving complex problems in high dimensions, due to large count of control parameters, also find out optimum hyper parameters value based on the behavior of capuchin monkeys at various ages. Figure 2 shows the flowchart for hybrid woodpecker mating and capuchin search optimization algorithm. The step-by-step process of woodpecker mating algorithm (WMA) and Capuchin search optimization are delineated below,

Back


Random relocation
Fitness Function
Assessment and divide population of the woodpeckers
Update mating behaviour of Woodpecker for optimizing l_i
Update the capuchins' positions to optimize the weight parameter η_i
Time iteration n=n+1
Life time of capuchins' γ
Is the Halting Criteria satisfied
No
Yes
Termination

Back

Open in figure viewer | PowerPoint

Flowchart for hybrid woodpecker mating and capuchin search optimization

Step 1: Initialization.

Initialize the process of woodpecker and capuchin to find the attacks. The populations of the Woodpecker and capuchin are considered as $q = 1, 2, 3, \ldots\ldots\ldots, M$ and the location of Woodpecker and capuchin is considered as $n = 1, 2, 3, \ldots\ldots\ldots, N$, which are initialized.

Step 2: Random generation.

The input parameters generated randomly after the initialization process. Hence, the values of best fitness for each woodpecker and capuchin are selected based on the explicit hyper-parameter situation.

Step 3: Fitness function.

Generate the random solution from the initialized values. Fitness function solution is determined, then the objective function indicates optimal parameter value l_i as well as η_i

$$\text{Fitness function} = \text{optimization}\left[l_i \text{ and } \eta_i\right]. \tag{14}$$

Step 4: Assessment and divide population of the woodpeckers.

The male woodpecker with the uttermost amount of fitness function is prefe… male woodpecker. The best male woodpecker is the maximum appealing wo… every female woodpecker would perceive the level of its drum that is qualified with their departure, advance on the way to it. The best male woodpecker fitness function is determined with the help of Equation (15)

$$\text{Best male woodpecker fitness function} = \text{Target value} - \text{Current output value}. \tag{15}$$

Step 5: UIDFte mating behavior of woodpecker for optimizing l_i.

In this step, the woodpecker position is uIDFted for optimizing l_i minimizing weight including biases parameters.

$$Z^{a}_{FW_{ra}} = l_i{}^{*}\, Random^{*}\left[Z_{BMW}\left(t\right) - Z_{random}\right], \tag{16}$$

$Random$ represents the random number and its values lies between −1 to +1, t represents present iteration number.

Step 6: UIDFte the capuchins' positions to optimize the weight parameter η_i.

Here, the parameter η_i is introduced for reinforcing balance among the search space exploration and exploitation. Therefore, capuchins' position is upgraded based on Equation (17),

$$y_G = y_j + u_0\, s + \frac{1}{2}\, c s^2, \tag{17}$$

where y_G and y_j denotes the final and initial displacement, s represents the time instance, u_0 specifies the initial velocity and c denotes acceleration. Hence, the acceleration is calculated using Equation (18),

$$c = \frac{\Delta u}{\Delta s} = \frac{u_G - u_0}{s_1 - s_0}, \tag{18}$$

where, s_1 and s_0 denotes final and initial time instance and parameter u_G expresses final velocity.

Step 7: Life time of capuchins'.

Here, lifetime exponential function γ is introduced in capuchins' to occur bal exploration and exploitation through global search and local search process. Hence, it is expressed in Equation (19),

$$\gamma = \chi_0\, e^{-\chi_1\left(\frac{F}{f}\right)\chi_2}. \tag{19}$$

From Equation (37), F and f specifies current and high iteration values, the parameter χ_0, χ_1, χ_2 denotes the arbitrarily selected values.

Step 8: Termination.

In termination, the optimum hyper-parameter l_i and η_i are selected in AGNN using hybrid woodpecker mating and capuchin search optimization algorithm alternatively repeat step 3 till the halting criteria n = n + 1 is met. At last, AGNN classifies the normal and anomaly attack in IoT environment by utilizing Hyb-WMA-CSOA.

This segment describes the auto metric graph neural network using hybrid woodpecker mating and capuchin search optimization algorithm fostered IDS to secure the IoT environment. The simulations are executed in MATLAB in core i7CPU. The performance metrics, like accuracy, F-score, precision are examined. Then the efficiency of the CSIC 2010 dataset are analyzed and compared with existing methods, like unsupervised learning-based network threat situation assessment model for Internet of Things (IDF-ResNet-IoT),[37] a distributed deep learning system for web attack detection on edge devices (IDF-GAN-IoT)[38] and web attack detection system for internet of things via ensemble classification (IDF-EDL-IoT)[39] respectively. Then the efficiency of the ISCXIDS2012dataset are analyzed and compared with existing methods Attack detection with deep learning analysis (IDF-ANN-IoT)[40] a novel statistical learning-enabled botnet detection mode for protecting smart cities networks (IDF-BMM-IoT)[41] and deep neural network in IoT intrusion detection (IDF-DNN-IoT)[42] respectively. Table 1 tabulates the simulation parameter.

TABLE 1. Simulation parameter

Parameter	Value
Number of nodes	25
Number of malicious node	6
Maximum of iteration	500
Number of training sample	80%
Number of testing sample	20%
Maximum velocity	5
Objective function	Maximize the detection accuracy

4.1 Dataset description

For detecting attacks from the IoT devices, in this, two datasets are taken, they are CSIC 2010 dataset, ISCXIDS2012 dataset. Here, CSIC 2010 dataset consists of 43,000 data, ISCXIDS2012 is a binary class dataset gathered through New Brunswick University for examining intrusion detection system 10,000 data. Out of total images, 80% data for training, 20% for testing.

4.2 Performance metrics

The performance of proposed IDF-AGNN-HYB-WMA-CSOA-IoT approach is analyzed under accuracy, precision, F-score. True positive, false positive, true negative, and false negative are

Back

- True negative (tn): Normal exactly detected as Attack.
- False positive (fp): Attack exactly detected as Normal.
- False negative (fn): Attack exactly detected as Attack.

4.2.1 Accuracy

This is computed to identify the performance of the proposed IDF-AGNN-HYB-WMA-CSOA-IoT approach while classifying attacks, which is calculated using the below formula. It is a ratio of exact predictions to a total count of proceedings in dataset. It is given in Equation (20)

$$\text{Accuracy} = {tp + tn}/{tp + tn + fp + fn}. \tag{20}$$

4.2.2 Precision

This is determined using Equation (21),

$$\text{Precision} = {tp}/{tp + fp}. \tag{21}$$

4.2.3 F-measure

This is scale the robustness of proposed model during attack classification. It represents harmonic mean of recall, precision.

$$F - \text{Measure} = {2tp}/{2tp + fp + fn}. \tag{22}$$

4.3 Simulation results

This segment describes the simulating outcomes of IDF-AGNN-HYB-WMA-CSOA-IoT Framework espoused Intrusion Detection for Protecting IoT Network for evaluating CSIC 2010 dataset and ISCXIDS2012 dataset.

4.3.1 Simulation result for CSIC 2010 dataset

Table 2–4 depicts the performance of accuracy, precision, F-measure, for CSIC 2010 dataset are analyzed and compared with existing methods, like IDF-ResNet-IoT,[37] IDF-GAN-IoT[38] and IDF-EDL-IoT[39] respectively.

TABLE 2. Comparison of accuracy (%) for CSIC 2010 dataset

Back

Denial Of Service (Dos) attack	93.1	90.31	92	99.35
Poisoning attacks	89.1	88.41	87	96.2
Adversarial attacks	85.1	84.31	85	95.3
Phishing attacks	88	83.4	84.3	99
Normal attacks	86	82.1	85.2	98
Mirai attack	75	74.65	78	99.77

TABLE 3. Comparison of precision (%) for CSIC 2010 dataset

Attack name	IDF-ResNet-IoT	IDF-GAN-IoT	IDF-EDL-IoT	IDF-AGNN-HYB- WMA-CSOA-IoT (proposed)
Denial of Service (DoS) attack	53.1	60.31	72	98.45
Poisoning attacks	59.1	68.41	77	92.2
Adversarial attacks	55.1	74.31	65	98.3
Phishing attacks	78	63.4	54.3	97
Normal attacks	56	62.1	75.2	94
Mirai attack	55	64.65	68	98.77

TABLE 4. Comparison of F-score (%) for CSIC 2010 dataset

Attack name	IDF-ResNet-IoT	IDF-GAN-IoT	IDF-EDL-IoT	IDF-AGNN-HYB- WMA-CSOA-IoT (proposed)
Denial of Service (DoS) attack	56.54	67.97	55.97	97.76
Poisoning attacks	65.98	76.87	72.87	97.96
Adversarial attacks	59.08	75.86	55.95	92.07

Back

Phishing attacks	67.85	60.97	54.97	90.75
Normal attacks	78.97	87.77	79.97	93.77
Mirai attack	76.88	59.98	68.97	92.88

Table 2 shows the comparison of accuracy for CSIC 2010 dataset. The proposed IDF-AGNN-HYB-WMA-CSOA-IoT method provides 24.87%, 11.56%, and 22.78% higher accuracy for DoS attack; 34.87%, 36.56%, and 24.78% higher accuracy for poisoning attack; 23.87%, 35.56%, and 46.78% higher accuracy for adversarial attack; 26.87%, 29.56%, and 42.78% higher accuracy for phishing attacks; 43.87%, 47.56%, and 32.78% higher accuracy for potential attacks; 13.87%, 47.56%, and 52.78% higher accuracy for normal attack; 25.29%, 63.25%, and 16.08% higher accuracy formirai attack compared to the existing IDF-ResNet-IoT, IDF-GAN-IoT, and IDF-EDL-IoT methods respectively.

Table 3 shows the comparison of precision for CSIC 2010 dataset. The proposed IDF-AGNN-HYB-WMA-CSOA-IoT method provides 32.75%, 16.43%, and 26.87% higher precision for DoS attack; 32.86%, 25.87%, and 20.98% higher precision for poisoning attack; 26.98%, 33.76%, and 20.76% higher precision for adversarial attack; 23.86%, 34.87%, and 45.75% higher precision for phishing attacks; 32.86%, 33.87%, and 32.86% higher precision for potential attacks; 31.65%, 21.55%, and 32.87% higher precision for normal attack; 15.13%, 31.74%, and 11.98% higher precision formirai attack compared to the existing IDF-ResNet-IoT, IDF-GAN-IoT, and IDF-EDL-IoT methods respectively.

Table 4 shows the comparison of F-score for CSIC 2010 dataset. The proposed IDF-AGNN-HYB-WMA-CSOA-IoT method provides 33.65%, 27.97%, and 43.86% higher F-score for DoS attack; 32.76%, 31.76%, and 32.43% higher F-score for poisoning attack; 26.87%, 34.87%, and 26.86% higher F-score for adversarial attack; 25.97%, 32.37%, and 26.98% higher F-score for phishing attacks; 21.87%, 23.77%, and 25.86% higher F-score for potential attacks; 24.86%, 25.55%, and 20.97% higher F-score for normal attack; 13.03%, 21.69%, and 9.21% higher F-score formirai attack compared to the existing IDF-ResNet-IoT, IDF-GAN-IoT, and IDF-EDL-IoT methods respectively.

4.4 Simulation result for ISCXIDS2012 dataset

Table 5–7 depicts the performance of accuracy, precision, F-measure, for ISCXIDS2012dataset are analyzed and compared with existing methods Attack detection with deep learning analysis (IDF-ANN-IoT),[40] statistical learning-enabled botnet identification framework for securing networks of smart cities (IDF-BMM-IoT)[41] and deep neural network in IoT intrusion detection (IDF-DNN-IoT)[42] respectively.

Back

Attack name	IDF-ANN-IoT	IDF-BMM-IoT	IDF-DNN-IoT	IDF-AGNN-HYB- WMA-CSOA-IoT (proposed)
Denial of Service (DoS) attack	67.86	75.97	73.97	98.08
Poisoning attacks	65.97	74.97	67.97	97.98
Adversarial attacks	76.97	74.97	68.97	99.88
Phishing attacks	79.76	68.99	59.97	97.08
Normal attacks	79.77	68.98	75.97	97.97
Mirai attack	68.88	58.9	69.88	97.77

TABLE 6. Comparison of precision (%) for ISCXIDS2012 dataset

Attack name	IDF-ANN-IoT	IDF-BMM-IoT	IDF-DNN-IoT	IDF-AGNN-HYB- WMA-CSOA- IoT (proposed)
Denial of Service (DoS) attack	57.86	75.86	75.86	94.86
Poisoning attacks	69.97	75.87	75.97	95.97
Adversarial attacks	56.97	75.99	76.97	95.97
Phishing attacks	78.97	68.94	77.96	94.86
Normal attacks	74.97	50.97	78.95	93.97
Mirai attack	75	74.65	67.77	96.65

TABLE 7. Comparison of F-score (%) for ISCXIDS2012 dataset

Attack name	IDF-ANN-IoT	IDF-BMM-IoT	IDF-DNN-IoT	IDF-AGNN-HYB- WMA-CSOA-IoT (proposed)
Denial of Service (DoS) attack	54.33	69.96	59.97	93.97
Poisoning attacks	69.99	79.67	76.97	94.97

Back

Adversarial attacks	60.97	86.87	68.9	95.88
Phishing attacks	69.88	76.97	75.88	95.97
Normal attacks	79.06	86.97	79.08	93.88
Mirai attack	76.99	69.99	76.99	96.97

Table 5 shows the performance analysis of accuracy for ISCXIDS2012 dataset. Here, the proposed IDF-AGNN-HYB-WMA-CSOA-IoT method provides 56.84%, 35%, and 27.86% higher accuracy for DoS attack; 52.88%, 43.86%, and 33.53% higher accuracy for poisoning attack; 43.86%, 32.87%, and 27.97% higher accuracy for adversarial attack; 26.97%, 25.07%, and 33.03% higher accuracy for phishing attack; 35.99%, 26.97%, and 52.86% higher accuracy for normal attack; 26.86%, 38.84%, and 29.04% higher accuracy for Mirai attack compared to the existing IDF-ANN-IoT, IDF-BMM-IoT, and IDF-DNN-IoT methods respectively.

Table 6 shows the performance analysis of precision for ISCXIDS2012 dataset. Here, the proposed IDF-AGNN-HYB-WMA-CSOA-IoT method provides 35.86%, 37.86%, and 37.88% higher precision for DoS attack; 37.87%, 43.86%, and 39.88% higher precision for poisoning attack; 37.87%, 43.98%, and 37.97% higher precision for adversarial attack; 38.97%, 27.88%, and 32.99% higher precision for phishing attack; 28.97%, 26.99%, and 42.86% higher precision for normal attack; 28.97%, 26.99%, and 43.75% higher precision for Mirai attack compared to the existing IDF-ANN-IoT, IDF-BMM-IoT, and IDF-DNN-IoT methods respectively.

Table 7 shows the performance analysis of F-score for ISCXIDS2012 dataset. Here, the proposed IDF-AGNN-Hyb-WMA-CSOA-IoT method provides 54.97%, 32.97%, and 37.87% higher F-score for DoS attack; 67.88%, 43.86%, and 32.77% higher F-score for poisoning attack; 31.65%, 21.66%, and 53.64% higher F-score for adversarial attack; 32.86%, 27.97%, and 41.86% higher F-score for phishing attack; 25.86%, 26.97%, and 29.77% higher F-score for normal attack; 32.86%, 20.88%, and 32.87% higher F-score for Mirai attack compared to the existing IDF-ANN-IoT, IDF-BMM-IoT, and IDF-DNN-IoT methods respectively.

5 CONCLUSIONS

In this article, intrusion detection framework using auto-metric graph neural network optimized with hybrid woodpecker mating and capuchin search optimization algorithm is successfully implemented for IoT Network. The real network traffic dataset CSIC 2010 dataset and ISCXIDS2012 dataset are deemed to estimate the efficiency of the proposed model. Here, the performance metrics, like accuracy, precision, F-score are analyzed. The proposed work is activated in MATLAB platform. The CSIC 2010 dataset performance provides higher accuracy

Back

FUNDING INFORMATION

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

Open Research

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

1 Ibor AE, Okunoye OB, Oladeji FA, Abdulsalam KA. Novel hybrid model for intrusion prediction on cyber physical Systems' communication networks based on bio-inspired deep neural network structure. *J Inf Secur Appl*. 2022; **65**:103107.

Web of Science® | Google Scholar

2 Esmaeilpour M, Cardinal P, Lameiras KA. A robust approach for securing audio classification against adversarial attacks. *IEEE Trans Inf Forens Secur*. 2020; **15**: 2147-2159.

Web of Science® | Google Scholar

3 Ma Z, Xia L, Gong X, Kokogiannakis G, Wang S, Zhou X. Recent advances and development in optimal design and control of ground source heat pump systems. *Renew Sustain Energy Rev*. 2020; **131**:110001.

Web of Science® | Google Scholar

4 Ahmad Z, Shahid Khan A, Wai Shiang C, Abdullah J, Ahmad F. Network intrusion detection system: a systematic study of machine learning and deep learning approaches. *Trans Emerg Telecommun Technol*. 2021; **32**(1):e4150.

Back

5 Saranya T, Sridevi S, Deisy C, Chung TD, Khan MA. Performance analysis of machine learning algorithms in intrusion detection system: a review. *Proc Comput Sci*. 2020; **171**: 1251-1260.

Google Scholar

6 Bi S, Huang L, Wang H, Zhang YJ. Lyapunov-guided deep reinforcement learning for stable online computation offloading in mobile-edge computing networks. *IEEE Trans Wirel Commun*. 2021; **20**(11): 7519-7537.

Web of Science® Google Scholar

7 Kasongo SM, Sun Y. A deep learning method with wrapper based feature extraction for wireless intrusion detection system. *Comput Secur*. 2020; **92**:101752.

Web of Science® Google Scholar

8 Almiani M, AbuGhazleh A, Al-Rahayfeh A, Atiewi S, Razaque A. Deep recurrent neural network for IoT intrusion detection system. *Simul Model Pract Theory*. 2020; **101**:102031.

Web of Science® Google Scholar

9 Shajin FH, Rajesh P, Raja MR. An efficient VLSI architecture for fast motion estimation exploiting zero motion prejudgment technique and a new quadrant-based search algorithm in HEVC. *Circuits Syst Signal Process*. 2022; **41**(3): 1751-1774.

Web of Science® Google Scholar

10 Jin D, Lu Y, Qin J, Cheng Z, Mao Z. SwiftIDS: real-time intrusion detection system based on LightGBM and parallel intrusion detection mechanism. *Comput Secur*. 2020; **97**:101984.

Web of Science® Google Scholar

11 Rajesh P, Shajin FH, Rajani B, Sharma D. An optimal hybrid control scheme to achieve power quality enhancement in micro grid connected system. *Int J Numer Model Electron Netw Dev Fields*. 2022;e3019.

Web of Science® Google Scholar

12 Shajin FH, Rajesh P. FPGA realization of a reversible data hiding scheme for 5G MIMO-OFDM system by chaotic key generation-based Paillier cryptography along with LDPC and its side channel estimation using machine learning technique. *J Circuits Syst Comput*. 2022; **31**(05):2250093.

Back

13 Rajesh P, Muthubalaji S, Srinivasan S, Shajin FH. Leveraging a dynamic differential annealed optimization and recalling enhanced recurrent neural network for maximum power point tracking in wind energy conversion system. *Technol Econom Smart Grids Sustain Energy*. 2022; **7**(1): 1-5.

Google Scholar

14 Abhishek N, Tandon A, Lim T, Sikdar B. A GLRT-based mechanism for detecting relay misbehavior in clustered IoT networks. *IEEE Trans Inf Forens Secur*. 2020; **15**: 435-446.

Web of Science® Google Scholar

15 Zhou Y, Cheng G, Jiang S, Dai M. Building an efficient intrusion detection system based on feature selection and ensemble classifier. *Comput Netw*. 2020; **174**:107247.

Web of Science® Google Scholar

16 Li X, Chen W, Zhang Q, Wu L. Building auto-encoder intrusion detection system based on random forest feature selection. *Comput Secur*. 2020; **95**:101851.

Web of Science® Google Scholar

17 Eskandari M, Janjua ZH, Vecchio M, Antonelli F. Passban IDS: an intelligent anomaly-based intrusion detection system for IoT edge devices. *IEEE Internet Things J*. 2020; **7**(8): 6882-6897.

Web of Science® Google Scholar

18 Thamilarasu G, Odesile A, Hoang A. An intrusion detection system for internet of medical things. *IEEE Access*. 2020; **8**: 181560-181576.

Web of Science® Google Scholar

19 Liu G, Yang Q, Wang H, Liu AX. Trust assessment in online social networks. *IEEE Trans Depend Secure Comput*. 2019; **18**(2): 994-1007.

Web of Science® Google Scholar

20 Aledhari M, Di Pierro M, Hefeida M, Saeed F. A deep learning-based data minimization algorithm for fast and secure transfer of big genomic datasets. *IEEE Trans Big Data*. 2018; **7**(2): 271-284.

Web of Science® Google Scholar

Back

Web of Science® Google Scholar

22 Lv L, Wang W, Zhang Z, Liu X. A novel intrusion detection system based on an optimal hybrid kernel extreme learning machine. *Knowl Based Syst*. 2020; **195**:105648.

Web of Science® Google Scholar

23 Kumar V, Sinha D, Das AK, Pandey SC, Goswami RT. An integrated rule based intrusion detection system: analysis on UNSW-NB15 data set and the real time online dataset. *Clust Comput*. 2020; **23**(2): 1397-1418.

Web of Science® Google Scholar

24 Alazzam H, Sharieh A, Sabri KE. A feature selection algorithm for intrusion detection system based on pigeon inspired optimizer. *Expert Syst Appl*. 2020; **148**:113249.

Web of Science® Google Scholar

25 Safaldin M, Otair M, Abualigah L. Improved binary gray wolf optimizer and SVM for intrusion detection system in wireless sensor networks. *J Ambient Intell Humaniz Comput*. 2021; **12**(2): 1559-1576.

Web of Science® Google Scholar

26 Hanselmann M, Strauss T, Dormann K, Ulmer H. CANet: an unsupervised intrusion detection system for high dimensional CAN bus data. *IEEE Access*. 2020; **8**: 58194-58205.

Web of Science® Google Scholar

27 Zhang J, Ling Y, Fu X, Yang X, Xiong G, Zhang R. Model of the intrusion detection system based on the integration of spatial-temporal features. *Comput Secur*. 2020; **89**:101681.

Web of Science® Google Scholar

28 Mighan SN, Kahani M. A novel scalable intrusion detection system based on deep learning. *Int J Inf Secur*. 2021; **20**(3): 387-403.

Web of Science® Google Scholar

Back

Web of Science® Google Scholar

30 Devan P, Khare N. An efficient XGBoost–DNN-based classification model for network intrusion detection system. *Neural Comput Appl*. 2020; **32**(16): 12499-12514.

Web of Science® Google Scholar

31 Sarhan M, Layeghy S, Portmann M. Towards a standard feature set for network intrusion detection system datasets. *Mobile Netw Appl*. 2022; **27**(1): 357-370.

Web of Science® Google Scholar

32 https://www.impactcybertrust.org/dataset_view?idDataset=940

Google Scholar

33 https://www.unb.ca/cic/datasets/ids.html

Google Scholar

34 Song X, Mao M, Qian X. Auto-metric graph neural network based on a meta-learning strategy for the diagnosis of Alzheimer's disease. *IEEE J Biomed Health Inform*. 2021; **25**(8): 3141-3152.

PubMed Web of Science® Google Scholar

35 Karimzadeh Parizi M, Keynia F, Khatibi BA. Woodpecker mating algorithm (WMA): a nature-inspired algorithm for solving optimization problems. *Int J Nonlinear Anal Appl*. 2020; **11**(1): 137-157.

Google Scholar

36 Braik M, Sheta A, Al-Hiary H. A novel meta-heuristic search algorithm for solving optimization problems: capuchin search algorithm. *Neural Comput Appl*. 2021; **33**(7): 2515-2547.

Web of Science® Google Scholar

37 Tian Z, Luo C, Qiu J, Du X, Guizani M. A distributed deep learning system for web attack detection on edge devices. *IEEE Trans Ind Inform*. 2019; **16**(3): 1963-1971.

Web of Science® Google Scholar

Back

Google Scholar

39 Luo C, Tan Z, Min G, Gan J, Shi W, Tian Z. A novel web attack detection system for internet of things via ensemble classification. *IEEE Trans Ind Inform*. 2020; **17**(8): 5810-5818.

Web of Science® Google Scholar

40 Pecori R, Tayebi A, Vannucci A, Veltri L. IoT attack detection with deep learning analysis. Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN); 2020: 1-8; IEEE.

Google Scholar

41 Ashraf J, Keshk M, Moustafa N, et al. IoTBoT-IDS: a novel statistical learning-enabled botnet detection framework for protecting networks of smart cities. *Sustain Cities Soc*. 2021; **72**:103041.

Web of Science® Google Scholar

42 Tsimenidis S, Lagkas T, Rantos K. Deep learning in IoT intrusion detection. *J Netw Syst Manag*. 2022; **30**(1): 1-40.

Web of Science® Google Scholar

Citing Literature