Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Optimized Deep Learning with Binary PSO for Intrusion Detection on CSE-CIC-IDS2018 Dataset

Rawaa Ismael Farhan [a] ***, Abeer Tariq Maolood*** [b] ***, Nidaa Flaih Hassan*** [c]

[a] *Department of Computer Science, University of Technology, Wasit University, Iraq, Email:ralrikabi@uowasit.edu.iq*

[b] *Department of Computer Science, University of Technology, Iraq, Email: 110032@uotechnology.edu.iq*

[c] *Department of Computer Science, University of Technology, Iraq, Email: 110020@uotechnology.edu.iq*

ARTICLEINFO

Article history:
Received: 04 / 10/2020
Rrevised form: dd /mm/year
Accepted : 25 /10/2020
Available online: 11 /11/2020

Keywords:

Network Intrusion Detection System, Feature Selection, Deep Learning, Binary PSO, CSE-CIC-IDS2018

ABSTRACT

Anomaly detection is a term refer to any abnormal behaviors, comprise security breaches of network. Deep Learning (DL)has proven its outperformance compared to machine learning algorithms in solving the complex problems of real-world like intrusion detection. Though, this approach need more computational resources and consumes long time. Feature selection is play significant role of choosing the best features that describes the target concept optimally during a classification process. However, when handle large number of features the selecting of such relevant features becomes a difficult task. Thus, this paper proposes using Binary Particle Swarm Optimization (BPSO) to solve the feature selection problem. Then, features selected from BPSO are evaluated on Deep Neural Networks (DNN) classifiers and the CSE-CIC-IDS2018 dataset. The result of the proposed model has shown comparable performance based on processing time, detection rate and false alarm rate comparing with other benchmark classifiers. Experimental results have shown a high accuracy of 95%.

https://doi.org/10.29304/jqcm.2020.12.3.706

1. Introduction

The exponentially growing number of security breaches, cyberattacks on Internet of things IOT highly required reliable security solutions. Network Intrusion Detection System (NIDS) used as defense of network infrastructure by detecting malicious activities and preventing attacks [1]. NIDS can be divided into misuse detection is also called signature-based detection and anomaly detection NIDS that are monitoring the network pattern and learning the normal behavior of a system and distinguish each network activity detect it as an intrusion when deviate from the normal pattern [2].

∗Corresponding author :***Rawaa Ismael Farhan***

Email addresses*: ralrikabi@ uowasit.edu.iq*

Communicated by : ***Alaa Hussein Hamadi***

We focus on anomaly detection NIDS because its ability to detect unknown attacks despite it have high false alarm rate because inability to determine reasons of an abnormality.

Traditional machine learning (ML) approaches have been supplied for cyber security such as Bayesian Belief Networks (BBN), Random forest, Support Vector Machines (SVM) and others, but the generation of large scale data in IoT required a deep learning based approach which performs better with large data sizes and can learn representation of feature from raw data so it is adaptable to different attack scenarios [3]. They proposed a new malware prediction model that could detect the coming future malware by the implementing a deep learning method of Mal Generative Adversarial Network (Mal-GAN) [4]. showed that the LSTM classifier outperform over previously published results of other static classifiers on KDD Cup '99 dataset challenge for long time which prove the benefit of LSTM networks to intrusion detection, because the ability of LSTM to learn from look back in time and link connection records consecutively [5].The RNN, Stacked RNN, and CNN are supervised deep learning techniques applied to classify common five attack types using Keras .This technique used packet header information without need any user payload then compared its results with Snort IDS .The results showed that this technique gave superior results compared Snort [6]. Variant-Gated Recurrent Units (GRU) with encoders performed on ISCX2012 dataset to make preprocessing on packets of payload-aware intrusion detection. It could learn features of network packet header and payload automatically and improved the detection rate of the IDS [7]. proposed RNN-RBM model which take input data as byte-level without feature engineering. At first, RBM model used network packets to extract the feature vectors. Then RNN model extracted the flow feature vector which sent to the Softmax layer to detect result [8].

Recently, the large growth of data makes big challenge to the task of data classification. The feature selection is an option to solve this challenge by reducing the dimensionality of the data and achieve higher accuracy in data classification.

In terms of feature selection, it plays an important role in improving NIDS performance. This is because anomaly detection uses a large number of time-consuming features. Therefore, choosing the method for selecting the feature affects the improvement of the level of accuracy and the time required to check traffic behavior. There are three types of features selection: filter, wrapper and embedded techniques. The filter technique tries to classify a subset from the original set containing of several selected features based on the evaluation criteria. While the wrapper technique, chose the features that have high predictive accuracy from different learning algorithms. The embedded technique where the feature selection embeds into the training step [9]. Paper performed Experiments on NSL-KDD datasets using log2 and PCA on deep learning algorithms. Results proved the effect of dimensionality reduction on the accuracy ratio about 97.9%. Thus, minimizing features in dataset and select optimal subset of most relevant features for each class to reduce processing time, improve detection accuracy rate, reduce false alarm rate. As result, the efficiency for intrusion detection in IOT environment improved because the irrelevant and redundant features cause overfitting and poor generalization during the classification [10].

Optimization means finding the optimal solution from a set of choices with regards to an objective function and some conditions. Intelligent applications that using Swarm Intelligence algorithms are becoming famous because of their ability to handle any real time complex and uncertain situation. Swarm intelligence is kind of algorithms which simulate the behavior of living organisms such as birds, insects, and fish. These individuals able to complete complex tasks in real world when working in unity that would be very difficult to achieve it [11].

In today's world, application of Swarm intelligence and Deep learning have been provided in many fields successfully such as image classification, pattern recognition and intrusion detection system. this paper has designed seven layer CNN which commonest deep learning approach, called ConvNet performed to classification of handwriting digit. The Particle Swarm Optimization algorithm (PSO) is used to improve the input parameters of processing layers [12]. This work proposed approach of swarm intelligence for parameter setting in deep neural network. through providing this approach to the phishing websites classification. As a result, the proposed algorithm improves their detection compared to other algorithms [13]. The contributions of this paper as following:

- Using a swarm intelligence for features selection by implementing a Binary PSO algorithm.

- Improving the NIDS by optimized deep learning models with pre-processing phase employing a Binary PSO algorithm. This approach optimized detection rate (DR) of deep learning models while reducing false alarm rate(FAR)compared with corresponding values of deep learning models without preprocessing phase.
- Evaluating this approach by using new CSE-CIC-IDS2018 real datasets for classification tasks.
- Presenting four comparative analyses between our results and the literature best results. Also, employing several evaluation metrics to depict analysis performance of deep learning models on our approach.

However, Swarm intelligence are often limited by weak points of computation time and local solution for large and complex problems. While, Deep learning algorithms are often limited by weak points of data and parameters.

2. Related Work

The Paper proposed a new algorithm to optimize the structure of DBN network. At first designed a PSO next used the fish behavior to optimize the PSO and find the initial solution of optimization. Then, used the genetic operators (crossover probability and mutation)on the PSO to search the global solution for optimization which used to construct the network structure for intrusion detection on NSL-KDD[14].The researcher aimed to improving the performance of NIDSs on UNSW- N15 dataset by proposed four feature selection models based on the particle swarm optimization (PSO), firefly optimization (FFA),genetic algorithm (GA)and grey wolf optimizer (GWO).The derived features from this model are evaluated on the J48 ML and support vector machine(SVM) classifiers[15].A double PSO-based algorithm proposed to select subset of features and hyper parameters both in the same work . Three deep learning models (Deep Neural Networks (DNN), Deep Belief Networks (DBN) and Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN) utilized to show the differences in performance on CICIDS2017dataset [16]. PSO-RF is an intrusion detection mechanism based on binary particle swarm optimization (BPSO) and random forests (RF) algorithms to find best features set by BPSO and RF as a classifier for classifying intrusions of networkonKDD99Cup dataset [17].

3. Methodology

This section proposed a Swarm-based intrusion detection method. This method improves the previous work by applying Binary PSO-based algorithm for feature selection stage. The proposed algorithm will optimize the detection performance of deep learning.

3.1 Data Set Specification

The traditional NSL-KDD dataset and others dataset not reflect situations of real world. According to Gharib et al. [18], determined 11 essential criteria for each dataset to be reliable dataset, but none of previous NIDS dataset covered all criteria. While, our CSE-CIC-IDS2018 dataset covered all 11 criteria. CSE-CIC-IDS2018 dataset represents a shift from static data to dynamically generated data available on AWS cloud [19,20]. This paper evaluated the NIDS on a real traffic captured from AWS network and machines log files with 80 extracted features from 50 terminals represent Attacking infrastructure and 30 servers and 420 computers represent the infected organizations comprised. Seven types of attacks occurred: DOS, DDoS, Botnet, Web attacks, Brute-force, infiltration and Heartbleed [20].

3.2 Proposed Work

In this paper, the proposed work is given as framework from the following:

1) Data preparation: it is preprocessing phase of data where removing unused features and duplicate instances for classification.

2) Feature selection: using Binary PSO algorithm to determine the most relevant features subset that enter as input to the classification phase.

3) Classification: using DNN model to enhance the classification accuracy on CSE-CIC-IDS2018 dataset.

The general flowchart of the proposed Model depicts in Figure 1.

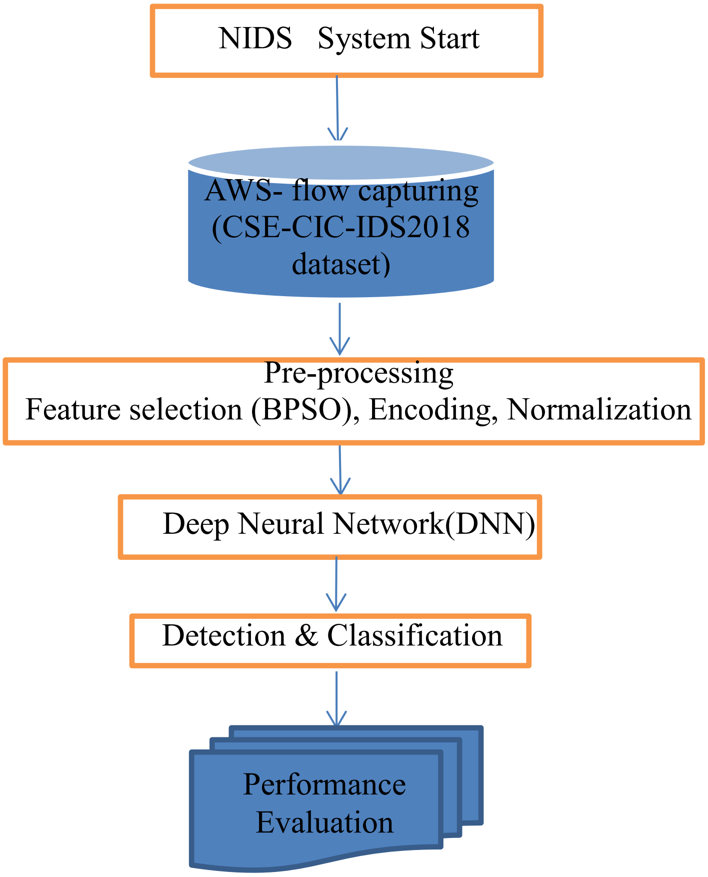

NIDS System Start
AWS- flow capturing (CSE-CIC-IDS2018 dataset)
Pre-processing
Feature selection (BPSO), Encoding, Normalization
Deep Neural Network(DNN)
Detection & Classification
Performance Evaluation

Figure 1. General Flowchart of proposed Model

3.2.1 Data Preparation

In data preparation, raw data transformed into more suitable form for modeling. Floating point numbers must be entered in range from 0 to 1 to the input layer of the DNN. There are three steps for data preprocessing:

- Data Cleaning: determine missing values and fix errors.
- Data Digitization: convert categorical features into numerical values.
- Data Transforms: such as normalization for changing scale, type, and probability distribution of variables in the dataset.

3.2.2 Feature Selection

Set of techniques used to select optimal subset of input features which are most relevant to the target variable need to predict. Dimensionality of the data mean the number of input features for a dataset, but the problem is more dimensions in space make dataset representation a very sparse and unrepresentative for samples in that space. So, this motivates feature selection that remove irrelevant and redundant input features which leading to lower predictive performance. Thus, models could develop by using the data that is required to make a prediction only.

this paper employed feature selection for reducing irrelevant attributes using random forest algorithm. Therefore, the task of NIDS became efficient. While, PSO algorithm performed on the selected features of the NSL-KDD dataset. This minimize the false alarm rate and increase the detection rate and the accuracy of the NIDS compared with machine learning classifiers such as SVM, KNN, DT and LR algorithms [21]. A new method has been presented based on particle swarm optimization with multiple criteria linear programming that improve attacks detection accuracy. During training phase, PSO used for tuning parameters to optimize the MCLP classifier performance [22]. Due to PSO has advantages that are simple implementation, fewer number of parameter, and no calculation of mutation. It is considered as best search algorithm for optimization. PSO classified into Standard PSO and binary PSO. Standard PSO assigns real numbers to particles, but binary PSO assigns binary numbers to particles.

A. Standard PSO

In [23] PSO after the population initialization each particle update its velocity and its position in each iteration based on their own experience (pbest) and the best experience of swarm (gbest) as in Equations (1) & (2). Then the performance of all particles evaluated by predefined cost functions at end of each iteration.

$$\mathrm{Vj}[\mathrm{st}+1] = \mathrm{W} * \mathrm{Vj}\,[\mathrm{st}] + \mathrm{F}\,1\,\mathrm{d}\,1(\mathrm{Pj\,best}\,[\mathrm{st}] - \mathrm{Pj}[\mathrm{st}]) + \mathrm{F}\,2\,\mathrm{d}2(\mathrm{G\,best}[\mathrm{st}] - \mathrm{Pj}[\mathrm{st}]) \qquad (1)$$

$$\mathrm{Pj}\,[\mathrm{st}+1] = \mathrm{Pj}\,[\mathrm{st}] + \mathrm{Vj}\,[\mathrm{st}+1] \qquad (2)$$

Where:
At each iteration st each particle j Acquire three vectors velocity, position and personal best all in length N which refer to the problem dimension. When either the improved value of the global best is smaller than stopping value (ε) or reached the maximum iteration number the stop condition is met and PSO terminates.

B. Binary Particle Swarm Optimization (BPSO)

The standard PSO used in continuous domains well, while in discrete space it gives poor effects on the results. Usually, the binary PSO in the feature selection problem outperforms the standard PSO because that the problem of feature selection occurring with a discrete search space. BPSO search space is seemed as a hypercube where a particle moves to nearer and farther corners of the hypercube through flipping bits into various numbers. The moving velocity represent changes of probabilities for the bit which may be in one state or the other. So, a particle in each dimension moves in a state space limited to 0 and 1 [24]. Therefore, we will exploit the binary PSO in our design for the feature selection method.

To implement the BPSO, the selected number of population is 100 and the number of iteration is considered to be 10, Initialize swarm randomly where X = (x1, x2, …., x n) is a particle as feature vector and y ϵ [0,1] represent class label which 0,1 respectively refer to normal and abnormal. Then, setting parameter as following:

W is the constant refer to Inertia weight that controls the velocity impact of particle during the current iteration it is usually ranged in [0.4,0.9]. **F 1** and **F 2** are acceleration coefficients constants ranged in [0.5]. while, **d 1** and **d 2** which are values ranged randomly in [0,1]. These parameter scale both of personal knowledge and swarm knowledge on the velocity changes. Consequently, calculate Activation Function to measure fitness value of each particle as in equation (3) to select particle with best value and called gbest.

$$\mathrm{F(X)} = \alpha(1 - \mathrm{Pr}) + (1 - \alpha)\left(1 - \frac{\mathbf{Ns}}{\mathbf{Nv}}\right) \qquad (3)$$

X is the input variable where **Pr** is the measure of classifier performance and **Ns** is the feature subset size have been tested and **Nv** is the total number of available input variables. The term on the left side of the equation refer to the total accuracy and the term on the right for the used features percentage.

BPSO is resulted by adapting equations in standard PSO to be suitable to binary space. The velocity vector in BPSO shows the probability of taking value 1 for element in the position vector. Moreover, the sigmoid function in Eq. (4) used to convert **V j(st+1)** to the range of [0,1].

$$S(V\ j\mathrm{st}+1) = \frac{1}{1+e^{-(V\ j\mathrm{st}+1)}} \qquad (4)$$

where rand () is a random selected value from range [0,1].

$$Pj\mathrm{st}+1 = \begin{cases} 1 & if \quad rand(\,) < S(V\ j\mathrm{st}+1) \\ 0 & otherwise \end{cases} \qquad (5)$$

Update position and velocity of each particle as equation (4) and (5). Finally, PSO output optimal solution that is global best vector next checks the stop condition when one met it the PSO will terminates.

Our DNN model implemented on Windows10 using Visual Studio 2019 contain python 3.7 and installed Keras on top of Tensorflow using (Numpy, Scikit-learn, Panda) libraries ,8GB Memory, CPU core i7,512GB Hard disk, seaborn library for visualization results.

Deep learning models classified into two types supervised and unsupervised learning models. comprise, deep neural networks (DNNs), deep brief networks (DBNs), recurrent neural networks (RNNs)and convolutional neural networks (CNNs) as supervised learning models. In other hand, restricted Boltzmann machines (RBMs), auto encoders and generative adversarial networks (GANs) as unsupervised learning models [25].

Deep learning methods plays a significant role for flow-based datasets compared with machine learning models because do not required manual feature engineering. Thus, it can learn feature representations automatically from raw data. The deep structure of deep learning represents comparable characteristic where used multiple hidden layers compared with shallow models, which contain one hidden layer or none [26].

The 55 optimal features selected by the BPSO algorithms from preprocessing phase will be provide to our DNN classifier to improve the performance of DNN contain three fully connected layers are used, they described as following:

dense1 layer with 55 neurons use ReLu Activation function.

dens2 layer with 64 neurons use ReLu Activation function.

dense3 layer with 10 neurons use Softmax Activation function.

Regularization method with two dropout ratio (0.2) are used to avoid overfitting.

Table 1 shows that good tuning of hyper parameter values is important to avoid overfitting.

Table 1: Experimental hyper parameter of proposed DNN model

Parameters	Value
Epoch	100
Batch size	500
Activation function	ReLu , Soft max
Loss function	categorical_crossentropy

Optimizer	Adam

In this paper used two types of non-linear Activation function are ReLu and softmax. ReLu is faster than other non-linear Activation function which maximize the deep learning efficient while Softmax used for multi classification as output layer in DNN model because it outputs the probability of each class then choose biggest value for accurate result.

Loss function represent the difference between the predicted and actual output. The Optimizer Adam used to minimize Loss function by calculate gradients of a loss after that apply gradients to update values and therefore enhance the DNN results.

4. Experimental Results and Discussion

We directed comparative analyses by comparing our results to the previous results in the literature. In addition to that, we proposed various evaluation Measurement in order to investigate the differences in performance in different approaches and focus on performance of deep learning models through using our approach.

4.1 Evaluation Measurements

We used for model evaluation various performance metrics to give powerful view on our model which based on BPSO with DNN as following:

1. **Confusion Matrix**

In the intrusion detection, Confusion Matrix is a good tool to predict the network attack type where TP normal data and TN refer to the abnormal data correctly classified, while FP the normal data and FN refer to abnormal data of the misclassification. Figure .2 show Confusion Matrix as resulted from the seaborn library in Python.

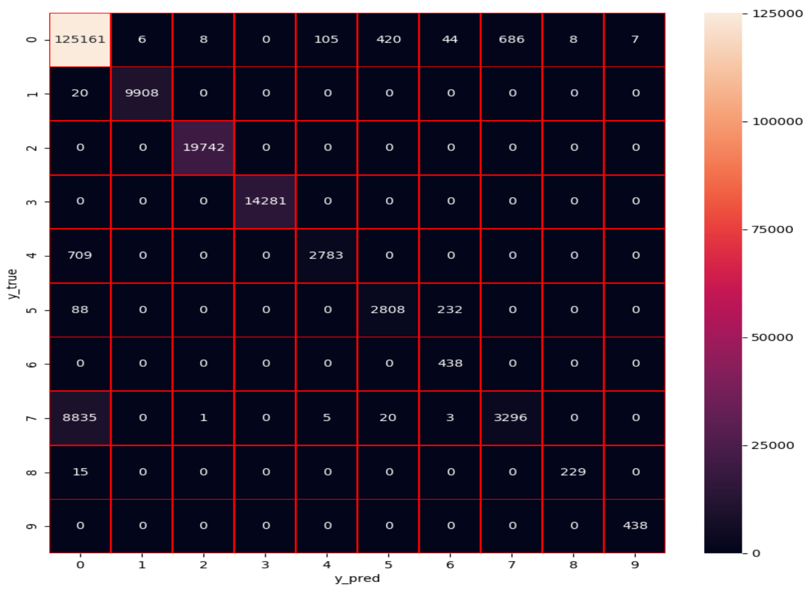
2. **Accuracy**: A percentage of positive detection of all data cases.

3. **Precision:** How many attacks are properly returned.

4. **Recall:** How many attacks the system returns.

5. **F1-score:** Rate of Precision and Recall in our model.

6. **Detection Rate (DR) and False Alarm Rate(FAR):** DR and FAR depicts how the classifier distinguishes well the positive and the negative classes, respectively.



	0	1	2	3	4	5	6	7	8	9
0	125161	6	8	0	105	420	44	686	8	7
1	20	9908	0	0	0	0	0	0	0	0
2	0	0	19742	0	0	0	0	0	0	0
3	0	0	0	14281	0	0	0	0	0	0
4	709	0	0	0	2783	0	0	0	0	0
5	88	0	0	0	0	2808	232	0	0	0
6	0	0	0	0	0	0	438	0	0	0
7	8835	0	1	0	5	20	3	3296	0	0
8	15	0	0	0	0	0	0	0	229	0
9	0	0	0	0	0	0	0	0	0	438

Figure2: Confusion Matrix of DNN based on BPSO

These metrics described in Table 2 from which We determine the detection rate (DR) and false alarm report (FAR).

Table 2 Performance analysis of our Model

Attack	PRECISION	RECALL	F1-SCORE
infiltrations	0.93	0.99	0.96
Benign	1.00	1.00	1.00
DDOS attack_HOLC	1.00	1.00	1.00
DDOS attack_LOTC_UDP	1.00	1.00	1.00
BOT	0.88	0.93	0.90
SQL Injection	0.74	0.89	0.84
FTP_Brutforce	0.00	0.00	0.00
SSH_Brutforce	0.87	0.27	0.41
DOS attack_slow HTTP Test	0.89	0.94	0.91
DOS attack_HULK	0.98	1.00	0.99

Accuracy, Precision, recall these criteria are limited, especially if one class among 10 classes is much larger than the other. With an imbalanced classification problem, the classification error in the minority class will not have much effect on the accuracy value. If the dataset is unbalanced, then in such cases, you only obtain very high accuracy by predicting the majority class, but you fail to capture the minority class, which is often the goal of creating the model in the first place .as shown in Table 3.

Table 3 Performance quality assessment

Attack type	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
DR %	0.144	2.217	7.035	0.0	0.002	0.005	0.0	0.003	0.000	5.267
FAR%	0.942	0.999	0.999	1.o	0.996	0.993	0.997	0.950	5.999	0.999

The relation between training accuracy and testing accuracy showing in Figure.3, where the model accuracy reached to 94% only with 10 Epoch, while when increased Epoch to 100 We noticed that accuracy has settled on 95%.

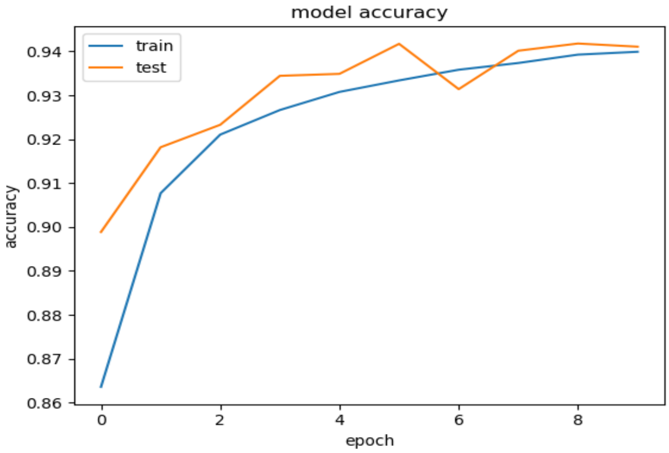


Figure .3 Model Accuracy

The Loss of Model showing in Figure.4, where in training phase beginning from o.6 and decreased through time, while The Loss of Model in testing phase beginning from o.4 and decreased through time. The elapsed Detection time about 580.27 sec.

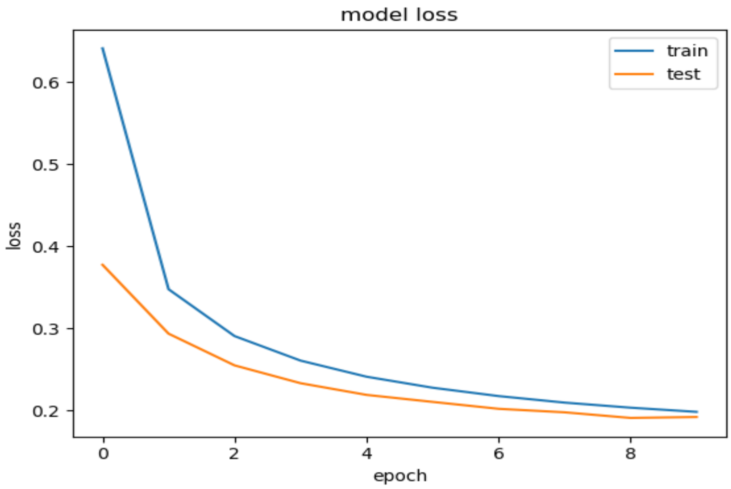


Figure .4 Loss of Model

4.2 Comparative Analysis

A comparative study in Table 4 is directed for showing the differences between our implemented Deep Neural Network (DNN) with Binary PSO and other previous methods. Our approach used Binary PSO as preprocessing phase that select only 55 best features from 80 features contained in real CSE-CIC-IDS2018 dataset. The selected features feed into DNN classifier. The results showed superiority over the previous classifier without feature selection. Moreover, no research used our dataset.

Table 4 Comparative Analysis

Classification algorithm	Feature selection method	Dataset	Accuracy	DR	FAR
Our DNN	Binary PSO	CSE-CIC-DS2018	95%	1.464	0.982
[14] DBN	At first PSO, Then Fish	NSL-KDD	83.86%	20.94	2.4
[15] SVM,J48	PSO,GWO,FFA,GA	UNSW-N15	89.01	80.84	2.817
			85.67	93.79	20.95
			86.03	96.58	22.59
			86.87	96.70	21.16
[16] three deep classifier DNN,LSTM,DBN	Double PSO	CICIDS2017	88.04	88.04	98.62
			92.41	92.41	99.31
			95.81	95.81	99.79
[27] Our Previous DNN	-	CSE-CIC-DS2018	90.25%	0.95	-

accuracy of proposed model is 95% after 100 Epoch which consider good, because we used Binary PSO which is one of the available feature selection methods that reduce the dataset dimensionality and select the most relevant features only. The proposed approach increase system accuracy, decrease false alarm report(FAR) and reduce the computation time where the elapsed time for Detection time about 580.27 sec.

Conclusion

Researches is still going on to improve NIDS because it is a challenge to reach the best results. The main goal is to increase Detection rate (DR) and reduce False alarm rate (FAR). Due to the increase in security threats at the present time on IOT networks and the resulting a massive amounts of data, Furthermore the heterogeneity of this data which takes a lot of processing time. It is necessary to choose the relevant features of the target class only, which improves the classification accuracy. This requires the use of a feature selection technique that improves the input features to the deep learning classifier. In this paper we used the BPSO algorithm for its ability to deal with real optimization problems by overcome the problem of falling into the local minima and fast implementation in large search area such as CSE-CIC-IDS2018 data set. Deep neural networks with swarm intelligence algorithms showed remarkably that it outperforms deep neural network alone with accuracy 95%.

References

[1] Sanju Mishra, Rafid Sagban, Ali Yakoob & Niketa Gandhi," Swarm intelligence in anomaly detection system: an overview", International Journal of Computers and Application, 2018, DOI:10.1080/1206212X.2018.152 1895

[2] Bruno Bogaz Zarpelao, et al., "A Survey of Intrusion Detection in Internet of Things, Journal of Network and Computer Applications, http://dx.doi.org/10.1016/j.jnca.2017.02.009

[3] Hongyu Liu and Bo Lang, "Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey", applied science,2019,9,439, DOI:10.3390/app9204396

[4] Shuqiang Lu, et al., New Era of Deep learning -Based Malware Intrusion Detection: The Malware Detection and Prediction Based On Deep Learning" ,2019, ArXiv, abs/1907.08356.

[5] Ralf C. Staudemeyer,"Applying long short-term memory recurrent neural networks to Intrusion detection", SACJ, No. 56, July 2015.

[6] Navaporn Chockwanich, VasakaVisoottiviseth, "Intrusion Detection by Deep Learning with TensorFlow", International Conference on Advanced Communications (ICACT),2019.

[7] Y. Hao et al., "Variant-Gated Recurrent Units with Encoders to Preprocess Packets for Payload-Aware Intrusion Detection", IEEE, VOLUME 7, 2019.

[8] Chaopeng Li et al., "Using a Recurrent Neural Network and Restricted Boltzmann Machines for Malicious Traffic Detection", Neuro Quantology, Volume 16, Issue 5 | Page 823-831, 2018, DOI: 10.14704/nq.2018.16.5.1391

[9] Antonia Nisioti, Alexios Mylonas, Paul D. Yoo, Vasilios Katos, "From Intrusion Detection to Attacker Attribution: A Comprehensive Survey of Unsupervised Methods", DOI: 10.1109/COMST.2018.2854724, IEEE.

[10] Murooj Khalid Ibraheem, et al., " Network Intrusion Detection Using Deep Learning Based On Dimensionality Reduction", REVISTA AUS 26-2, DOI:10.4206/ Aus. 2019.n26.2.23.

[11] Constantinos Kolias,Vasilis Kolias, Georgios Kambourakis,"TermID : a distributed swarm intelligence-based approach for wireless intrusion detection", Int. J. Inf. Secur.,Springer, 2016,DOI:10.1007/s10207-016-0335-z.

[12] Mujahid H. Khalifa et al., "Particle Swarm Optimization for Deep learning of Convolution Neural Network ", Sudan Conference on Computer Science and Information Technology (SCCSIT), IEEE, 2017.

[13] Grega Vrbančič, Iztok Fister Jr. and Vili Podgorelec,"Swarm Intelligence Approaches for Parameter Setting of Deep Learning Neural Network: Case Study on Phishing Websites Classification", International Conference on Web Intelligence, Mining and Semantics, 2018, Novi Sad, Serbia. ACM, https://doi.org/10.1145/3227609.3227655

[14] Peng Wei, et al., "An Optimization Method for Intrusion Detection Classification Model based on Deep Belief Network", IEEE, doi:10.1109/ACCESS.2019.2925828.

[15] Omar Almomani, "A Feature Selection Model for Network Intrusion Detection System Based on PSO, GWO, FFA and GA Algorithms", Symmetry 2020, 12, 1046; doi:10.3390/sym12061046.

[16] Wisam Elmasry, Akhan Akbulut, Abdul Halim Zaim, "Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic", Computer Networks, 2019,doi:https://doi.org/10.1016/j.comnet.2019.107042 .

[17] Arif. J. Malik, W. Shahzad and F. A. Khan, "Network intrusion detection using hybrid binary PSOand random forests algorithm", Security An Communication Networks
Security Comm. Networks 2015; 8:2646–2660, DOI: 10.1002/sec.508.

[18] Sharafaldin, I., Lashkari, A. and Ghorbani, A.," Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018), pages 108-116, DOI: 10.5220/0006639801080116.

[19] I. Sharafaldin et al.," Towards a Reliable Intrusion Detection Benchmark Dataset ", Journal of Software Networking, 177–200, doi: 10.13052/jsn2445-9739.2017.009.

[20] https://www.unb.ca/cic/datasets/ids-2018.html

[21] Nilesh Kunhare , Ritu Tiwari And Joydip Dhar,"Particle swarm optimization and feature selection for intrusion detection system", Sadhana (2020) 45:109 , https://doi.org/10.1007/s12046-020-1308-5Sadhana(0123456789).

[22] Seyed Mojtaba Hosseini Bamakan et al., "A New Intrusion Detection Approach using PSO based Multiple Criteria Linear Programming ", Elsevier, Procedia Computer Science 55(2015) 231 – 237.

[23] D. Asir Antony Gnana Singh et al.," Enhancing the Performance of Classifier Using Particle SwarmOptimization (PSO) - based Dimensionality Reduction",International Journal of Energy, Information and Communications,Vol.6, Issue 5 (2015), pp.19-26,http://dx.doi.org/10.14257/ijeic.2015.6.5.03.

[24] H. Nezamabadi-pour, M. Rostami-shahrbabaki, M.M. Farsangi, “Binary Particle Swarm Optimization: challenges and New Solutions”, The Journal of Computer Society of Iran (CSI) On Computer Science and Engineering (JCSE), vol. 6, no. (1-A), pp. 21-32, 2008.

[25] Samaneh Mahdavifar, Ali A. Ghorbani, " Application of deep learning to cybersecurity: A survey", Elsevier, Neuro computing 347, pp. 149–176, 2019.

[26] Mehdi Mohammadi et al., " Deep Learning for IoT Big Data and Streaming Analytics: A Survey", IEEE Communications Surveys & Tutorials,2018, doi:10.1109/comst.2018.2844341.

[27] Rawaa Ismael F., Abeer T. M., Nidaa F. H.," Performance Analysis of Flow-Based Attacks Detection on CSE-CIC-IDS2018 Dataset Using Deep Learning", Indonesian Journal of Electrical Engineering and Computer Science, Vol20, No.3,2020.