Enhancement of Network Attack Classification using Particle Swarm Optimization and Multi Layer-Perceptron

Ibraim M. Ahmed
College of Science
University of Mosul
Mosul, Iraq

ABSTRACT

Network intrusion detection systems (NIDSs) give classification for all data passing during these systems and produce an alarm report whether these data are normal or abnormal. Many researchers have used various techniques to solve classification problems in IDSs but these techniques still have some vulnerability by getting imperfect classification for attacks. In this study, a proposed system has been developed that achieves classification technique by using hybrid soft computing technique which is Multi Layer-Perceptron (MLP) with Particle Swarm Optimization (PSO). The PSO has been used to improve the learning capability of the MLP by setting up the linkage weights in an attempt to enhance classification accuracy of the MLP. Simulation results conducted over three forms of experiments show that the proposed system gives high classification compared with other methods. The results show also that the percentages of classification has been reached to (98.9%) with (1.1) false alarm.

General Terms

Networks Security, Intrusion Detection, Neural Networks Applications

Keywords

Network Intrusion Detection (NIDS), Multi Layer-Perceptron (MLP), Particle Swarm Optimization (PSO), NSL-KDD99**.**

1. INTRODUCTION

Network applications growth in various fields was accompanied with some obstacles such as attacks on networks. Attackers adopt several methods to penetrate a system such as Denial of service, malware, worms, and etc. [1]. There are many types of virus or worms each one have techniques to work to made Legitimate operation on any services as types of attack defined. Attacks can be divided by into active and passive and each one have several types. The four well known class of attacks are Denial of service (DOS), Probing (probe), User to Root (U2R) and Remote to Local (R2L).The first one is active attack whereas the other are classified into passive attack. Each attack has many types of virus or worms to get efficient work in network [2]. Thus the security system must have the ability to determine types of attack to select best mechanism to prevent the damage from this attack or to repel to these attacks. Several researchers have been worked to solve this problem by using several security solutions such as firewall, antivirus, encryption and intrusion detection system. Intrusion detection systems (IDS) whether hardware or software are used to decide where threat is normal or abnormal depending on some rules of analysis engine in the IDS [3]. An analysis engine of IDS can be built by many types of techniques such as soft computing, statically module, and etc. Soft computing techniques have many aspects such as artificial neural network (ANN), Genetic Algorithms (GA), swarm intelligence. Attempt to combine many techniques to get efficient IDS were also adopted [1,2, 3, 4]. Efficient IDS can be achieved by combining many techniques and using standard international dataset for training and evaluation such as IDS KDD99 and NSL-KDD [4].

2. MULTI LAYER-PERCEPTRON NEURAL NETWOK BASED IDS

Multilayer Layer-Perceptron neural network is a feed forward learning machine and supervised learning machine that consists of more than two layers. Each layer is connected with each other layers by linkage weights and having connection with balancing node called bias node as shown in Figure 1 [5].

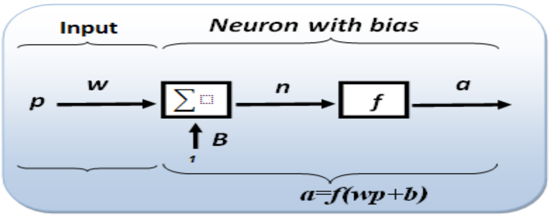

Fig 1: Neuron activity of MLP [5]

Each output of a neuron is produced by applying one of activation function on the input by its weight such as (sigmoid, bipolar and etc.) [5]. Stopping criteria of this network depends on mean square error between actual output and target of neuron or reached the target of training [5].

Neural network concepts have been used for intrusion detection system. Building neural network in intrusion detection system takes many aspects of applications. Each one is different from the other in order to get better performance [6, 7]. Some focuses on the application on data processing and feature selection other application focuses on analysis of traffic to distinguish between normal and abnormal connection record. Most researchers who used neural networks in intrusion detection systems uses Standard international Dataset KDD99 or NSL-KDD99 (2009) for training and evaluation intrusion detection system[8].

ANN method can be used in the analysis part of the intrusion detection system and work to train system to detect intrusion in network traffic. Some neural networks work with misuse approach because it uses pattern matching with pattern of a standard database and other work with anomaly detection because it uses threshold between two states[5, 8, 9]. NID systems usually achieve data clustering via classification and self-organizing maps whereas neural networks the MLP is used for detection [3]. On the other hand, the hybrid combination of neural network learning algorithms can be used for normal behavior modeling. It has been shown that the IDS adopting neural networks modeling algorithm has the ability to detect all intrusion attempts with lease false alerts [4].

Other models of NIDS are PCA-SVM model, DBN-SVM model, GA-HNB model and GA-IEM-C4.5 model [7]. These models involve data preprocessing, data reduction and intrusion classification. These models use different intelligent algorithms and feature selection and extraction techniques. The experimental results of the four NID models show the model's advantages of enhancing the detection accuracy and testing speed can be achieved by reducing the feature dimension space [7].

Other approach for the design and implementation of NIDS is based on the genetic algorithm. With this approach, researchers reduce the redundancy and selected appropriate features by using Principle Component Analysis (PCA) of NSL-KDD99 dataset [8]. Other researchers concluded that system the IDS can be further speed up the process via the usage of GA and PCA to select features not only the usage of the NSL-KDD99 [9, 10, 11]. This will reduces the CPU time, the time of training and the testing time [9]. Furthermore, other researchers designed and implemented NIDS by using Artificial Bee Colony with Multi Layer-Perceptron (ABC-MLP) on NSL-KDD99 [4].

3. PARTICLE SWAM OPTIMIZATION

Particle Swarm Optimization is an adaptive optimization algorithm based on problem environment with velocity update and position update primary operators. Within each iteration, and for each particle a new velocity value is calculated using current particle velocity, the distance from its previous best position and the distance from the global best position. Then the next position in the search space next position of the particle in the search space is updated using the new velocity as shown in PSO algorithms steps shown in Figure 2 [10].

NSL-KDD99 (2009) [11, 12] is a modern standard dataset for training and evaluation network intrusion detection system. NSL-KDD99 contains 125973 traffic records for training and 22544 traffic records for testing NIDS. These dataset records contain 41 features for each connection. NSL-KDD99 datasets were derived from KDD99 data set of DARPA [12]. These dataset are classified into 9 basic features, 13 content based features and 19 time based feature. NSL-KDD99 dataset different from the KDD99 Data set about many features such as number of pattern, number of redundancy records, number of duplication and less complexity. These dataset contains many types of attack categories such as Dos, Prob, U2R and

R2L [12].

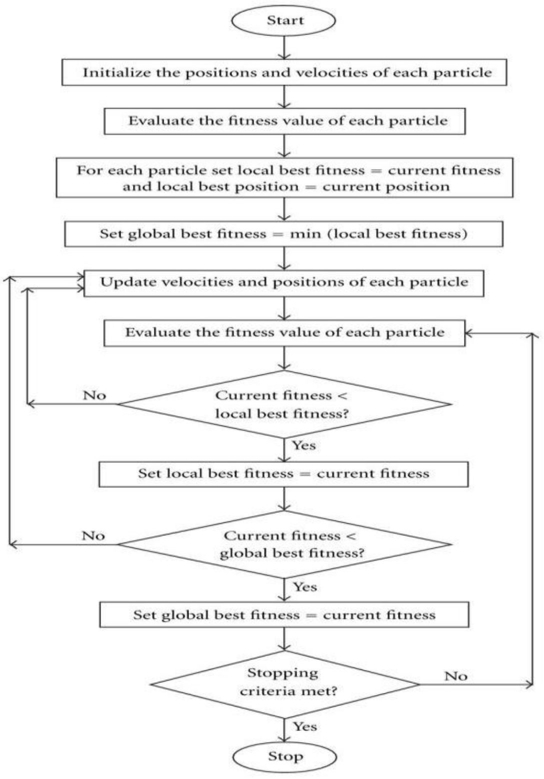
Start
Initialize the positions and velocities of each particle
Evaluate the fitness value of each particle
For each particle set local best fitness = current fitness and local best position = current position
Set global best fitness = min (local best fitness)
Update velocities and positions of each particle
Evaluate the fitness value of each particle
No
Current fitness < local best fitness?
Yes
Set local best fitness = current fitness
No
Current fitness < global best fitness?
Yes
Set global best fitness = current fitness
Stopping criteria met?
No
Yes
Stop

Fig 2: PSO algorithm steps

4. PROPOSED NETWORK INTRUSION DETECTION SYSTEM

The proposed ID system architecture layout shown in Figure 3 contains several parts which are data collection, analysis engine, storage and classifier. Data collection part job is to enter traffic data to the system (NSL-KDD99) data set. Analysis engine with PSO technique has given the multi political solution to Multi–layer Perceptron network to produce a best particle after training the system to use it in classifier part of testing dataset.

The NSL-KDD99 dataset in its form is not suitable for the analysis engine. Therefore, it has been converted to the numeric form. The symbolic feature value is converted to numeric value depending on the number of occurrence value in dataset. Each value will be in the range from [*1, maximum no of occurrence*]. These parameter values are as shown in Table 1. Next, normalization method is conducted on the numeric dataset by min_max methods using the following equation [13];

$$X\_value = x\_value - min\_x / max\_x - minx \quad \ldots\ldots \qquad (1)$$

Having completed normalization data became suitable to enter to PSOMLP structure. With the proposed system the MLP NN architecture consists of three layers which are input, hidden and output layer. Input layer consists of 41 node represented features of each connection record. Hidden layer consists of 5 node.output layer consists of 5 layer represent classes of connection record. Connection weights between layers represented by set of particles updating depend on PSO condition as shown in Figure 3.

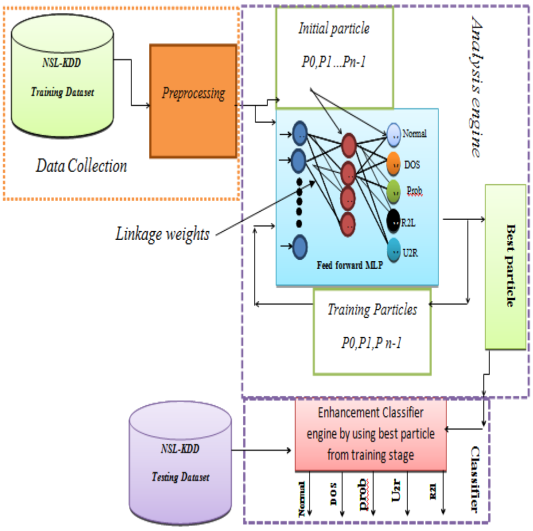

Fig 3: Proposed Hybrid PSOMLP system layout

Table 1. Features representation

Symbolic feature	No. of value	Min	Max	
Protocol type		3	1	3
service		66	1	66
Flag		11	1	11
Other feature		Numeric range		

5. SYSTEM SIMULATION AND EVALUATION

The proposed system has two units, one for training the system and the other for testing the system. The training is achieved on selected training data set (NSL-KDD99) and initial particle set that generated randomly. Particle set represented weight set of the MLP are selected for each connection records of data set and feed forward for each particle are examined. The value of the fitness for each output class of the MLP is used to select best particle (PBest) and when selected a new particle test global particle (GBest) is also achieved. This nested operation enhances the ability of detected different types of attacks. Having exceeded the number of training step, the best particle is stored. Testing the system uses the best particle and testing data set. The mechanism adopted in this proposed system can be well understood with flow chart of operation activity shown in Figure 4.

Three different experiments have been conducted using different parameter, the number of particle and the number of training steps. All experiments are executed on a 2.6GHZ core i5 processor and 4GB of RAM running windows 8.1 The processing is done using Microsoft visual C# 2012. The results of all experiments differ from other experiments by certain percentages in the diagnosis of the attacks. However, when increasing the training steps and the number of particle, the time spent in the training process and accuracy of detection will be increased.

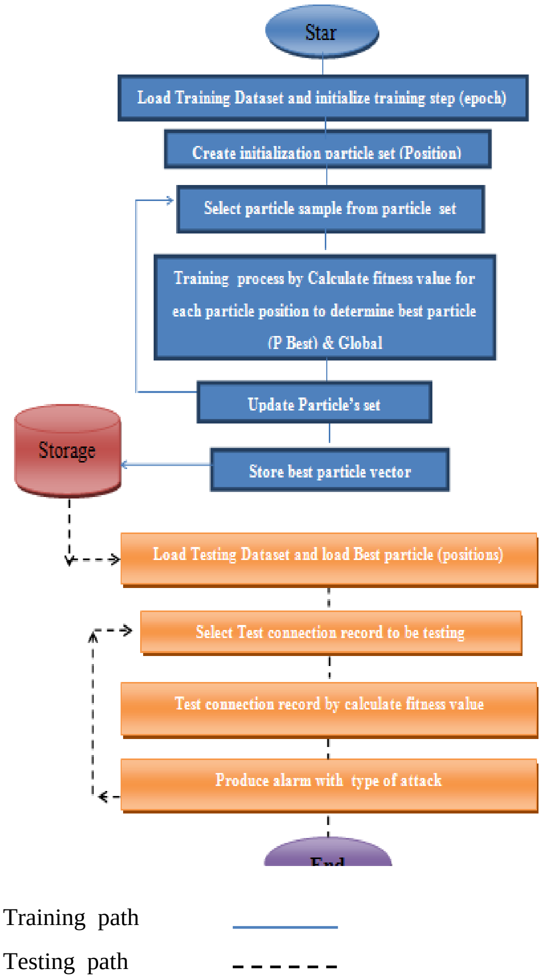
Star
Load Training Dataset and initialize training step (epoch)
Create initialization particle set (Position)
Select particle sample from particle set
Training process by Calculate fitness value for each particle position to determine best particle (P Best) & Global
Update Particle's set
Storage
Store best particle vector
Load Testing Dataset and load Best particle (positions)
Select Test connection record to be testing
Test connection record by calculate fitness value
Produce alarm with type of attack
End
Training path
Testing path

Fig 4: Operation activity of the proposed (PSOMLP) system

To evaluate the performance of IDS we needed to use some criteria. The True Positive (TP) and True Negative (TN) can be used as correct classification criteria. A False Negative (FN) occurs when the outcome is incorrectly predicted as negative when it is actually positive. A False Positive (FP) occurs when the outcome is incorrectly predicted as yes (or positive) when it is actually no (negative) as shown in Table 2 [11].

The accuracy of the detection (detection rate) DR, and the error rate ER, have been calculated using the followings equations [6];

$$DR = \frac{TN+TP}{Total\ Number\ \ of\ Samples} *100 \qquad …….(2)$$

$$ER= \ = \sqrt{(DesiredOutput - ActualOuput)^2} \qquad ….(3)$$

The proposed system have been tested and evaluated on the complete universal dataset NSL-KDD99 with 125973 patterns for training and 22544 patterns for testing in each experiment. Three different experiments have been conducted based on the changing the values of some parameters in the PSO as shown in Table (3).

The first experiment result produces the number of false alarm rate in all classes of connection record as shown in Table 3. It is clear that the result tends to increase the training step (no. of epoch) and the number of particles to enhance the result.

The result as with experiments 2 and 3 preview that the U2R result have high false alarm rate. Furthermore, the three experiments results show that the increase in number of training step to 100 training step with the increase in the no. of particle's to 50 will produce high detection rate and lower false alarm in the proposed system as shown in Tables 4-6 for experiments 1-3, respectively.

Table 2. Confusion Matrix representation

Predicate Actual	Normal (Negative)	Attack (Positive)
Normal(Negative)	TN	FP
Attack(Positive)	FN	TP

Table 3. Applied experiments on the proposed system

Parameter	Experiment 1	Experiment 2	Experiment 3
No of epoch	20	50	100
No. of particles	10	25	50
C1	2	2	2
C2	2	2	2
R	[0,1]	[0,1]	[0,1]

Table 4. The result of the first experiment

Class	D.R	FAR
Normal	82.15323	17.846772
DoS	78.34499	21.65501
Prob	74.40819	25.59181
U2R	55.22388	44.77612
R2L	86.54179	13.458213

Table 5. The result of the second experiment

Class	D.R	FAR
Normal	97.71393	2.286067
DoS	97.42133	2.578671
Prob	96.10429	3.895713
U2R	85.07463	14.92537
R2L	98.70317	1.29683

Table 6. The result of the third experiment

Class	D.R	FAR
Normal	99.21738	0.782618
DoS	99.03846	0.961538
Prob	99.29623	0.703775
U2R	98	2.0
R2L	99.27954	0.720461

In testing phase, the comparison between the detection rate for the three experiments has been conducted and presented as shown in Figure 5. It is clear that the third experiment gives the best detection rate of 98.9%. On the other hand the results of false alarm rate conducted with the three experiments show that experiment three gives the lowest rate as shown in Figure (6). The results also show that the accuracy of detection is increased when increasing the number of particle's and epoch. Furthermore, when we comparing the result of the proposed system with other recently proposed methodology from literature survey it can be noticed that the performance of the proposed system is better in terms of detection rate as shown in Table 7.

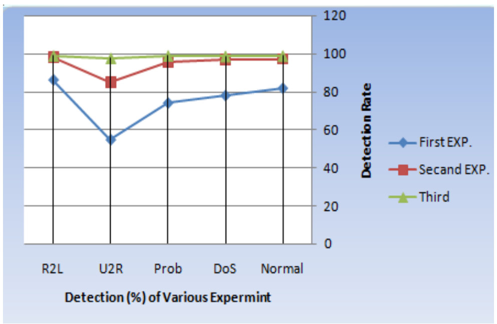

Fig 5: Detection Rates of the three experiments

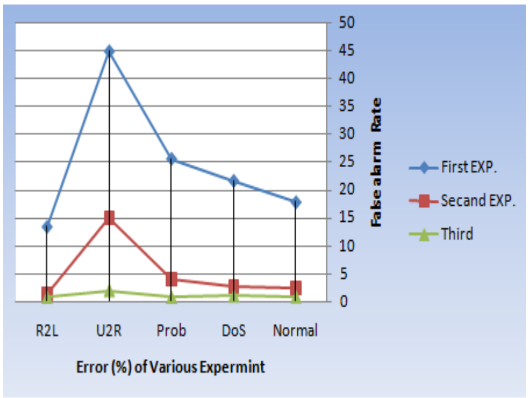

Fig 6: False alarm rate of the three experiments

Table 7. Comparative results of PSOMLP

Proposed methods	No of pattern	No of feature	DR	FA
RBF-SVM[9]	/	41	98.5	1.5%
Bagging[13]	22544	41	61.8	38.2%
Stacking [13]			81	19%
Naïve Bayes[13]	22544	41	72.6	27.4
FC-ANN[8]	/	41	96.75	2.5%
ABC+MLP[4]	22544	41	87.27%	12.73%
Proposed methods	22544	41	98.9	1.1%

6. CONCLUSIONS

The research proposed a new hybrid intrusion detection that adopts multi layer-perceptron neural network for classification with particle swarm optimization (PSO) to enhance the learning of MLP, named MLPPSO. The PSO is used to set up the linkage weights in an attempt to enhance classification accuracy of the MLP. With this proposed system, NSL-Kdd99 data set was used to evaluate the proposed system with three different experiments. The result of classification using MLPPSO enhancement produce a high detection rate because of the use of multi weight set (particles) in the learning phase. The percentages of classification reached is (98.9%) with (1.1) false alarm**.** The experiments results show that the proposed system gives high classification result compared with other methods. The results also show that increasing the numbers of practices and numbers in the training step will enhance the accuracy of classification**.** Future suggestion can be in the design and development of the proposed network intrusion system on a real time environment.

7. REFERENCES

[1] Toosi A. N. et al., 2007, A Novel Soft Computing Model Using Adaptive Neuro-Fuzzy Inference System for Intrusion Detection, Proceedings of the IEEE International Conference on Networking, Sensing and Control, London, UK.

[2] Chou T. N. Kristopher K., 2009, Hybrid Classifier Systems for Intrusion Detection, IEEE Seventh Annual Communication, Networks and Services Research Conference, P 286 – 291, Moncton.

[3] Linda, O. J. Volmer, T. and Manic, M., 2012, Neural Network Based Intrusion Detection System for Critical Infrastructures", International Joint Conference on Neural Networks.

[4] Mahmod S. M. and Alnaish Z. A. H. ,2015, Hybrid Intrusion Detection System Using Artificial Bee Colony Algorithm and Multi-Layer Perceptron, IJCSIS, , Pittsburgh PA, USA.

[5] Daniel Graupe, 2007, Principles of Artificial Neural Network, 2nd Edition Advanced Series on Circuits and Systems , Vol. 6 ,University of Illinois, Chicago, World Scientific Publishing Co. Pte. Ltd.,USA

[6] Kemmerer R. A. and Vigna G., 2002,Intrusion Detection: A Brief History and Overview, .IEEE Computer Journal, vol. 35, issue 4, pp. 27-30.

[7] Eid H. F., 2012, Computational Intelligene in Intrusion Detection System, MSc Thesis, Al-Azhar University, 2013.

[8] Ibraheem N. B. and Osman, H. M. ,2013, Principle Components Analysis in network Intrusion Detection System using NSl-KDD, Rafidain J. of Comp. & Math's., vol. 10, no. 1.

[9] Nagle M. K., Chaturved, S. K., 2013, Feature Extraction Based Classification Technique for Intrusion Detection System", International Journal of Engineering Research and Development, ijerd, Volume 8, Issue 2, PP. 23-38.

[10] Yuan K., Shu Y.; Wei W. and Wang D., 2014, Particle Swarm Optimisation Clustering for Cement Kilning System Fault Recognition, International Journal of Industrial and Systems Engineering (IJISE), Vol. 17, No. 4.

[11] Charles E, 2000, Results of the KDD99 Classifier Learning, ACM SIGKDD Explorations News letter, Vo. 1. Issue 2, pp. 63-64.

[12] Tavallaee M., E. Bagheri, W. Lu, and A. Ghorbani, 2009, A Detailed Analysis of the KDD CUP 99 Data Set, Second IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA).

[13] Levin I. , 2000, KDD-99 Classifier Learning Contest LLSoft's Results Overview, SIGKDD Explorations, ACM SIGKDD, vol. 1 issue 2, pp. 67-75.

[14] Wang G., Hau J. , Ma J., and Huang L., 2010, A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering, Expert Systems with Applications, Elsevier, vol. 37, issue 9, pp. 6225–6232.