An Effective Fusion of PSO-Neural Network Technique for Intrusion Detection

[1]Ms. Lissy P K, [2]Ms. Jesna Anver

M.Tech Scholar, Associate Professor,

TOC H INSTITUTE OF SCIENCE AND TECHNOLOGY, ARAKKUNNAM
ERNAKULAM, KERALA, INDIA

Abstract— **Nowadays it has become harder to identify a new type of attack and anomaly behavior, because the flows of modern network to analyze are huge. A variety of algorithms are available to resolve this major issue, but, the performance of those algorithms is not much better. To improve the performance we encompass two areas such as Computational Intelligence (CI) and Swarm Intelligence (SI). An Effective Fusion of PSO-Neural Network Technique for Intrusion detection is an algorithm to enhance the performance of the IDS. The proposed algorithm is composed of two major phases: (i) the classification phase uses artificial neural network to create the classifier from the known training data set and (ii) the clustering phase uses a PSO algorithm to classify newly incoming data patterns, which may contain known and unknown network attacks.**

Index Terms—**Classification, Clustering, Neural Network, PSO.**

I. Introduction

Intrusion Detection System (IDS) is an important type of security software used in Internet. Intrusion detection means to monitor and possibly prevent attempts to intrude into or otherwise compromise your system and network. The IDS prevents the unauthorized access of system or Network. We can use a firewall system or authentication system to prevent unauthorized access, but both can be broken. Monitor, detect, and respond to any unauthorized activity are the adages of Intrusion detection systems. If you have a computer system, you want to protect your data and system integrity. The fact that you cannot always protect that data integrity from outside intruders in today's internet environment using mechanisms such as ordinary password and file security, leads to a range of issues. The first step to ensure data protection is providing adequate security. Firewalling and other access prevention mechanisms should always be put in place. Intrusion detection takes that one step further. A network based intrusion detection system placed between the firewall and the system can provide an extra layer of protection to that system. The IDS system monitor access from the internet to the **s**ensitive data ports of the secured system and can determine whether the firewall has perhaps been compromised, or whether an unknown mechanism has been used to bypass the security mechanisms of the firewall to access the network being protected [1].

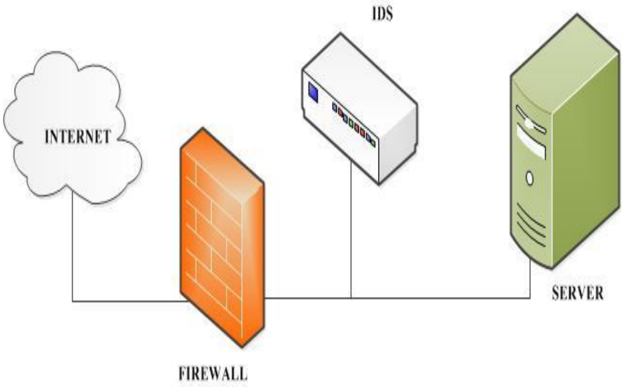

Fig 1.Typical locations for an intrusion detection system

Due to a growing number of intrusion events and also because the Internet and local networks have become so ubiquitous, organizations are increasingly implementing various systems that monitor IT security breaches [2]. As the threat becomes a serious matter year by year, intrusion detection technologies are essential for network and computer security. A variety of intrusion detection approaches are available to resolve this severe issue, but the main problem is performance. For getting better performance, it is important to increase the detection rates and reduce false alarm rates in the area of intrusion detection. In order to detect the intrusion, various methods have been developed and proposed over the last decade.

Intrusion Detection System is in two types:

- Network –Based Intrusion Detection System
- Host –Based Intrusion Detection System

IDSs can be categorized as misuse detectors or anomaly detectors by sorting out broadly based on their models of detection. An intrusion detection system is capable of detecting various types of malicious network traffic and computer usage. The performance of the IDS is measured by evaluating the computation time and the accuracy rate. The computation time is also called response time. If a system can provide highly accurate detection of abnormal behavior with minimum computational speed, then the system is efficient.

II. Related Works

The quantity and types of the intrusions increase radically as the speed and difficulty of networks expand quickly, particularly when these networks are unlocking to the public Web. Therefore, it is becoming hard for any presented intrusion detection system to suggest a trustworthy repair with the varying technology and the exponential growth of Internet traffic it has been created that a behavioral model exists in the attacks that can be educated from former study. Various algorithms such as Support vector machine, genetic algorithms, Fuzzy logic, Data mining, neural networks, etc have been broadly used to the huge volume of complex and active data set to detect known and unknown activities so as to identify intrusion activities.

In this section a brief review of support vector machine classifier related intrusion detection is discussed .In the period 2007-2012, a lot of papers have been presented based on the Support vector machine for intrusion detection. Some of the papers have been discussed below. A revise for improving the training time of SVM has been presented by Latifur Khan et al. [3], particularly when contracting with large data sets using hierarchical clustering analysis in 2007.They utilized the Dynamically Growing Self-Organizing Tree (DGSOT) algorithm for gathering since it had verified to triumph over the disadvantages of traditional hierarchical clustering algorithms (e.g., hierarchical agglomerative clustering). The advanced research done in the year 2011, Iftikhar Ahmad,Azween et al. have proposed a paper to surmount presentation issues an optimized interference detection mechanism by means of soft computing techniques [4]. An intelligent multi level classification technique for effective intrusion detection proposed by S. Ganapathy et al. [5] in Mobile Ad-hoc Networks in the same year. The algorithm is a combination of a tree classifier which used a labeled training data and an Enhanced Multiclass SVM algorithm. An advantage of SVM as a classifier for an IDS is that they are highly accurate and are able to model complex non linear decision boundaries. Support Vector Machines are also less prone to over more fitting than other methods. The disadvantages of SVM as a classifier for IDS are its algorithmic complexity and memory requirements. High algorithmic complexity and extensive memory requirements leads to slower the training and testing process.

Here, we discuss the different papers that utilize k-means algorithm. In 2003 -2004 some papers presented about the K-means algorithm based intrusion detection. Some of the papers have been discussed below. In the year 2003, Yu Guan et al. [6] presented a K-means based clustering algorithm, named Y-means, for intrusion detection. One of the benefits of the Y-means algorithm for intrusion detection is a suitable number of clusters were divided by a data set routinely. The other advantage was the unprocessed log data of information systems can directly be applied as training data with-out being physically labeled. In 2004, K. M. Faraoun and A. Boukelif [7], using the K-means clustering algorithm to improve the learning capacities and decrease the computation strength of a competitive learning multi-layered neural network.

In 2005, Jiu-Ling Zao et al. [8] proposed a novel approach of using clustering genetic algorithms is put forward to solve the computer network intrusion detection problem. The algorithm can not only cluster the cases automatically, but also detect the unknown intruded action. S.Janakiraman et al. [9] proposed an intelligent distributed intrusion detection system using genetic algorithm in 2009. The proposed system presents an intelligent learning approach using Genetic Algorithm (GA) for distributed intrusion detection system (DIDS), which uses a simple representation of rules and an effective fitness function. M.Sadiq Ali Khan [10] presented a paper in 2011 .The paper describes Rule Based Network Intrusion Detection Using Genetic algorithm. The study showed that GA can be effectively used for formulation of decision rules in intrusion detection through the attacks which are more common can be detected more accurately. It is also determined that an increasing number of iterations of application of an algorithm contributes in accuracy of data. However initial iteration converges to result more quickly as compared to later iteration. The genetic algorithm designed was successfully able to generate a model with the desired characteristics of a high correct detection rate and low false positive rate of learning over training data. The genetic algorithm was able to perform the mutation and evolution strategies according to the fitness function. The implementation of such a model generated by the clustering genetic algorithms for real time intrusion detection is difficult. In this section a brief review of two techniques related to neural network based intrusion detection is discussed.

Marjan Bahrololum et al.[11] improved the concept of anomaly detection and use both neural network (NN) and decision tree (DT) for intrusion detection. M. Bahrololum et al. published a paper in the same year to plan the system using a hybrid of misuse and irregularity detection for training of normal and attack packets respectively [12]. It was the mixture of unsupervised and supervised Neural Network (NN) for Intrusion Detection System. Attacks were categorized into smaller categories taking into consideration their similar features by the unsupervised NN based on Self Organizing Map (SOM), and followed by unsupervised NN based on Back propagation was utilized for grouping. Known packets were recognized fast by misuse approach and unknown attacks will be able to detect by this method.

III. Particle Swarm Optimization

Particle swarm optimization (PSO) is a robust stochastic optimization technique based on the movement and intelligence of swarms developed in 1995 by Dr. Eberhart and Dr. Kennedy, inspired by social behavior of bird flocking or fish schooling. It uses a number of agents or particles that constitutes a swarm moving around in the search space looking for the best solution. In the PSO algorithm, each particle is treated as a point in a N-dimensional space which adjusts its "flying" according to its own flying experience as well as the flying experience of other particles. The basic concept of PSO algorithm is to create a swarm of particles which move in the space around them (the problem space) searching for their goal, the place which best suits their needs given by fitness function [13].

This value is called personal best (pbest). Another best value that is tracked by the PSO is the best value obtained so far by any particle in the neighborhood of that particle. This value is called global best (*gbest).*

Each particle is associated with a position pi and velocity V_i and the modified position of particle P_i at iteration t+1 can be mathematically modeled as:

$V_i^{t+1}=wV_i^t+C_1*\text{rand}\ (0,1)*Pb_i-P_i^t+C_2*\text{rand}\ (0,1)*gb-P_i^t+..(1)$

Where

- V_i^t : velocity of particle i at iteration t,
- W: weighting function,
- $C_{i\,:}$Weighting factor,
- Rand: uniformly distributed random number lies between 0 and 1,
- Pb_i : Pbest of particle i ,
- gb_i: Global best of particle i ,
- P_i^t: Current position of particle i at iteration t.

The weighting factor calculated using the equation given below:

$$W=W_{max}-[W_{max}-W_{min}]*t/t_{max} \qquad (2)$$

- W_{max}: Initial weight
- W_{min} : final weight
- t_{max} : Maximum number of iteration

$$P_i^{t+1}=P_i^{\,t}+V_i^{t+1} \qquad (3)$$

IV. Artificial Neural Networks

Artificial neural network is one of the core methods of computational Intelligence. Artificial Intelligence is a most powerful tool available for detecting precise relationship in massive amounts of unrelated data. The neural network can be used to solve a variety of problems such as optimization, pattern recognition, prediction etc. The neural networks are characterized by a set of nodes and the connection between the nodes. The connection can be unidirectional or bidirectional which determines the flow of information. Artificial neural networks are relatively crude electronic networks of "neurons" based on the neural structure of the brain.

A neuron in an artificial neural network is

- A set of input values (xi) and associated weights (wi)
- A function (g) that sums the weights and maps the results to an output (y).

A typical neural network has millions of artificial neurons known as units arranged as layers, each layer is connected to either side. The input units are to receive various forms of information from the outside world that the network will attempt to learn about, recognize, or otherwise processed. The output units respond to the information it's learned. In between the input units and output units are one or more layers of hidden units together, form the majority of the artificial brain.

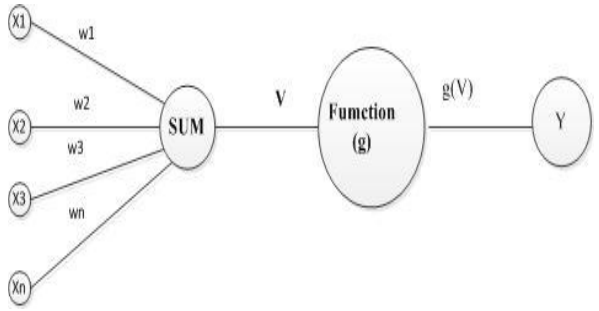

Fig 2.Neuron Structure

Fully connected neural networks mean each hidden unit and each output unit is connected to every unit in the layers either side. The weight is a number representing the connections between one unit and another, which can be either positive (if one unit excites another) or negative (if one unit suppresses or inhibits another). If the weight is high, which means the more influence one unit has on another. The artificial neural networks data classification mainly has the following advantages:

- The stronger generalization through the elimination of redundant nodes will avoid saturation; the better of robustness;
- The small deviation; fast and accurate real-time processing after fast hardware;
- High classified and predicted precision;
- Scalable algorithm.

V. Proposed Algorithm

The computation time and the accuracy rate are the two main measurements for evaluating the performance of the IDS. So our goal is to focus on reducing the computation time and finding an efficient detection mechanism to improve the accuracy rate of the proposed algorithm for recognizing the network traffic.

A. Overview

The algorithm consists two phases: *classification phase* and *clustering phase*. In the classification phase, the artificial neural network, one of the core methods of Computational Intelligence is used to create the classifier from the known network traffic data. The neural network it identifies the input pattern and tries to output the corresponding class. It can map

input patterns to their associated output patterns using their mapping capabilities. They can be trained with known examples of a problem before they are tested for their inference capability on unknown instances of the problem. Therefore, they can identify new objects previously untrained. The clustering phase uses a PSO algorithm to classify the newly incoming patterns to through the network, which may contain known and unknown network attacks. The PSO algorithm requires no previous knowledge of the dataset to be partitioned, and can determine the optimal number of classes dynamically. Clustering aims at representing large data sets by a fewer number of prototypes or clusters. Here we encompass the two areas such as computational intelligence and swarm intelligence. Characteristics of computational intelligence (CI) systems, adaptation, fault tolerance, high computational speed and error resilience in the face of noisy information fit the requirements of building a good intrusion detection model. A family of bio-inspired algorithms known as Swarm Intelligence (SI) is good in the field of pattern recognition and clustering and which has gained huge popularity in these days.

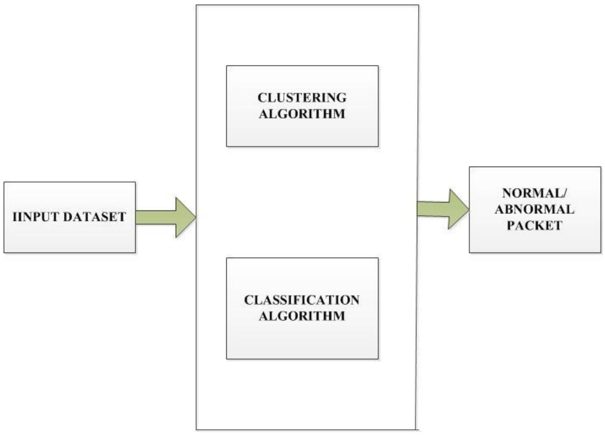

Fig.3.Overview of the algorithm

B. Clustering Algorithm

Till now, many approaches have been proposed to improve the performance of IDS. Clustering is the most important techniques of data mining which has been widely used and acceptable. It is an unsupervised method and takes a different approach by grouping objects into meaningful sub classes so that members from the same cluster are quite similar and different to the members of different cluster. The unsupervised anomaly detection algorithm clusters the unlabeled data together into clusters using a simple distance-based metric. Normal instances should form large clusters and all of the instance that appear in small clusters are labeled as anomalies. Because normal instance are qualitatively different so they do not fall into same cluster.

In clustering phase, preprocess the data from the input dataset. In particle swarm optimization algorithm the input data is considered as particles. The outline of the classification algorithm is as follows:

Algorithm 1: Clustering Algorithm

Input: Data packet as Particle

Output: The best particle

Description:

```
1. The first assignment procedure
2. Initialize all the particles by the centroids obtained by classification algorithm
3. For each particle i
4. Let $Td = 0$
5. For each unlabeled pattern x in $Du$
6. Calculate the distance between x and all the centroids
7. Find the centroid $Ci$ that is closest to x and its distance to $xdx$
8. IF $dx < di$ or $dx < Td$ then
9. Assign pattern x to the cluster i
10. Else
11. Add a new cluster and assign x to it.
12. Let $Td = 3dx$
13. End for
14. Create $pa$ particles first from classification then create another $pb$ particles each of which consists of $Kmin$ upto $Kmax$ randomly generated centroids.
15. Perform the calculate, update and change operators to adjust the velocities and positions.
16. Perform one-iteration k-means to adjust the particles [option].
17. If the stop criterion is satisfied, then stop and out the best particle.
18. Else goes to 15.
```

The algorithm clusters the particle into different clusters according to their characteristics.

C. Classification Algorithm

The classification and recognition of individual characteristics and behaviors of data packets are the main function of the intrusion detection system. When performing classification analysis with a set of existing data, is to split the into a larger data set for training and a smaller data set for testing the model.

- Training an artificial neural network means is the correct class for each pattern is (supervised training), and the output nodes can therefore be assigned "correct" values , "1" for the node corresponding to the correct class, and "0" for the others.
- The iterative learning process in which the input data are presented to the network one by one, and the weights associated with the input values are adjusted each time. After all input values are present, the process starts over again. During this phase, the

network learns by adjusting the weights so as to be able to predict the correct class label of input samples.

The very first thing that classification algorithm does is to determine attributes (fields) of the labeled patterns that are to be used in the Classification Phase. This is because each pattern may contain a large number of attributes, classification based on all of them is a time consuming task, especially when the number of patterns becomes large. Then, the major procedure of the classification is performed to create the initial classifier based on the set of labeled patterns given.

Algorithm 2: Classification Algorithm

Input: Labeled pattern

Output: Set of weights as the initial classifier

Description:

1. Create an initial network with as many input neurons as required by the specific problem described. Randomly initialize the connection weights of the network within a certain range.
2. Partially train the network on the training set for a certain number of training epochs using a training algorithm.
3. For Eliminate the redundant weights by using a weight elimination algorithm.
4. Test this network. If the accuracy of the network falls below an acceptable range, then add one more hidden unit and go to step 2.
5. If there any input node $x_1$with x_1^m= 0, for m=1to h, then remove this node.
6. Test the generalization ability of the network with the test set. If the network successfully converges it terminate, otherwise, go to step 1.
7. Return the set of weights as the initial classifier.

In order to apply the techniques to information security, we needed datasets. We used a commonly applied dataset in information security research: The network intrusion dataset from the KDD archive popularly referred to as the KDD 99 Cup dataset. The training dataset has 19.69% normal and 80.31% attack connections. KDD CUP 99 has been most widely used in attacks on networks. The simulated attack categories are: Denial of Service (DOS) attack, User to Root (U2R), Remote to local (R2L), Probing. The protocols that are considered in KDD dataset are ICMP, TCP and UDP.

VI. Conclusion

The proposed algorithm is the combination of neural network and PSO algorithm. When the neural network is used, it identifies the input pattern and tries to output the corresponding class. It can map input patterns to their associated output patterns using their mapping capabilities. They can be trained with known examples of a problem before they are tested for their inference capability on unknown instances of the problem. Therefore, they can identify new objects that are previously untrained. The PSO algorithm used in clustering phase is classifying the newly incoming patterns, which may contain known and unknown network flow types. It requires no previous knowledge of the dataset to be partitioned, and can determine the optimal number of classes dynamically. Clustering aims at representing large data sets by a fewer number of prototypes or clusters. It brings simplicity in modeling data and thus plays a central role in the process of knowledge discovery and data mining. The proposed algorithm can provide a highly accurate classification rate even using a small number of training patterns. So we can reduce the computation time and can improve the accuracy rate for recognizing a great deal of network traffic.

References

[1] http://www.symantec.com/connect/articles/intrusion-detection-theory-and-practice

[2] http://www.windowsecurity.com/articles- utorials/intrusion_detection /IDS-Part2-Classification-methods-techniques.html

[3] Latifur Khan, Mamoun Awad, Bhavani Thuraisingham, “”A new intrusion detectionsystem using support vector machines and hierarchical clustering'', Journal of VLDB Journal, vol.16, pp.507-521, 2007

[4] Iftikhar Ahmad,Azween Abdullah,Abdullah Alghamdi,Muhammad Hussain, ``Optimized intrusion detection mechanism using soft computing techniques'', Telecommunication System,2011.

[5] S. Ganapathy, P. Yogesh, and A. Kannan,``An Intelligent Intrusion Detection System for Mobile Ad-Hoc Networks Using Classification Techniques'', Advances in Power Electronics and Instrumentation Engineering, Communications in Computer and Information Science Vol.148, pp 117-122,2011.

[6] Yu Guan, Nabil Belacel and Ali A. Ghorbani, `` Y-Means: A Clustering Method for Intrusion Detection'',Canadian Conference on Electrical and Computer Engineering, vol.2, pp. 1083-1086, 2003.

[7] K. M. Faraoun and A. Boukelif,``Neural Networks Learning Improvement using the K -Means Clustering Algorithm to Detect Network Intrusions'',International Journal of Computational Intelligence, Vol.3, no.2, 2005.

[8] Jiu-Ling Zhao, Jiu-Fen Zhao, Jian –Jun Li,``Intrusion detection Based on Clustering Genetic Algorithm'',Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 2005.

[9] S.Janakiraman, V.Vasudevan,``An Intelligent Distributed Detection System using Genetic Algoritm'',Journal of Convergence Information Technology, Volume 4, Number 1, March 2009.

[10] M.Sadiq Ali Khan,``Rule Based Network Intrusion Detection Using Genetic algorithm'', International Journal of Computer Applications (0975 –8887)Volume 18–No.8, March 2011.

[11] Marjan Bahrololum, Elham Salahi, Mahmoud Khaleghi, ``An Improved Intrusion Detection Technique based on two Strategies Using Decision Tree and Neural Network''Journal of Convergence Information Technology,Vol.4, No.4, December 2009.

[12] M. Bahrololum, E. Salahi and M. Khaleghi, ``Anomaly intrusion detection design Using Hybrid of Unsupervised and supervised neural Network'',International Journal of Computer Networks and Communications (IJCNC), Vol.1, No.2, July 2009.

[13] S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” *IEEE Trans. on Neural Networks*, vol. 4, pp. 570-578, July 1993.

[14] J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility,” *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34-39, Jan. 1959.

[15] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, “Rotation, scale, and translation resilient public watermarking for images,” *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 767-782, May 2001.

Ms.Lissy P K is currently a scholar of Master of Technology in Computer Science and Engineering at Toc H Institute of Science and Technology, Arakkunnam, Ernakulum (Kerala). She obtained her graduation in Computer Science and Engineering from The Institution of Engineers (INDIA). She is an Associate member of IEI. Her special fields of interest are computer networking and data security.

Ms.Jesna Anver is currently an Associate Professor in Department of Computer Science and Engineering at Toc H Institute of Science and Technology, Arakkunnam, Ernakulum (Kerala). She obtained her graduation from CUSAT. She obtained her post graduation from Amritha University, Coimbatore. Her special field of interest and research are data mining, machine learning, image processing, and computer networks. She has more than 10 years teaching experience.