IJCSE International Journal of Computer Sciences and Engineering Open Access

Review Paper Volume-3, Issue-10 E-ISSN: 2347-2693

An Effective Approach for Improving Anomaly Intrusion Detection

Kumar J S[*], Appa Rao S S V, Subha Sree M

[1*,3] *Department of CSE, DMS SVH College of Engg. , Machilipatnam ,AP*
[2] *Department of CSE, VISIT Engg. College , Tadepalliguem ,AP*

www.ijcseonline.org

Received: Sep/17/2015 Revised: Oct/01/2015 Accepted: Oct /16/2015 Published: Oct /31/ 2015

***Abstract*—** Intrusion Detection Systems (IDS) is a key part of system defense, where it identifies abnormal activities happening in a computer system. In general, the traditional intrusion detection relies on the extensive knowledge of security experts, in particular, on their familiarity with the computer system to be protected. To reduce this dependence, various data-mining, soft-computing and machine learning techniques have been proposed in recent years for the development of better intrusion detection systems. Many researchers used Conditional Random Fields and Layered Approach for purpose of intrusion detection. They also demonstrated that high attack detection accuracy can be achieved by using Conditional Random Fields and high efficiency by implementing the Layered approach. In the paper we explained a new method called fuzzy ARTMAP classifier (FAM) and clustering technique for effectively identifying the intrusion activities within a network. Processing huge data would make the system error prone, hence clustering the data into groups and then processing will result in having a better system. From each of the cluster, representative data is selected in the selective process for further processing. For classification process, layered fuzzy ARTMAP will have the better results when compared to other normal classifier algorithms. Finally the experiments and evaluations of the proposed intrusion detection system is using the KDD Cup 99 intrusion detection data set.

***Keywords*—**Intrusion Detection System, Layered approach, Clustering, FAM.

I. Introduction

Nowadays, networks in computer face an unprecedented range of threats and vulnerabilities which created a greater risk in computer security [1]. Also intrusion events to computer systems are growing due to the popularization of the Internet and local networks. Generally, the goal of threats and attacks is to subvert the traditional security mechanisms on the systems and execute operations in excess of the intruder's authorization. These operations could include reading protected or private data or simply doing malicious damage to the system or user files [2]. So only by building complex tool the system can be protected from malicious attacks. Intrusion detection systems are becoming increasingly important in maintaining proper network security. An intrusion detection system (IDS) monitors networked devices and looks for anomalous or malicious behavior in the patterns of activity in the audit stream. Intrusion Detection System is used to monitor the events occurring in a computer system or network, analyze the system events, detect suspected intrusion, and then raise an alarm

Various researchers used data mining as the key component to detect newly encountered attacks. Based on this research Clustering is considered as one of data mining that has been applied to many theoretical and practical problems. Also fuzzy clustering is considered as one of the best clustering method among hard clustering and fuzzy clustering which deals with "might be" situations in the real world. The different types of fuzzy clustering methods are fuzzy c-means (FCM), possibilistic c-means(PCM). However, most fuzzy clustering methods process propositional data, only a small portion can process relational data . Hence we present a new clustering and classification technique in order to improve the classification accuracy of the intrusion detection system. The clustering technique is the Possibilistic Fuzzy C-Means (PFCM) and the classification technique is Fuzzy ARTMAP neural network classifier (FAM).

On the other hand, many researchers have argued that artificial neural networks (ANNs) have been widely used as an intelligent classifier to identify the different categories based on learning pattern from empirical data modelling in complex systems. There are several types of neural architectures found which were successfully used for classification purpose. Several types of neural architectures found which were successfully used for classification purpose. The types are, back-propagation, associative neural networks, probabilistic neural networks, generalized regression neural networks, radial basis function networks, cascade correlation, neural networks trained via evolutionary algorithms, support vector machines, self-organizing maps, fuzzy neural networks (FNNs).

II. COMPONENTS OF INTRUSION DETECTION SYSTEMS

***A.** Data Preprocessor* – Data preprocessor is responsible for collecting and providing the audit data (in a specified form) that will be used by the next component (analyzer) to make a decision. Data preprocessor is, thus, concerned with collecting the data from the desired source and converting it into a format that is comprehensible by the analyzer.

***B.** Analyzer* (Intrusion Detector) – The analyzer or the intrusion detector is the core component which analyzes the audit patterns to detect attacks. This is a critical component and one of the most researched. Various pattern matching, machine learning, data mining and statistical techniques can be used as intrusion detectors. The capability of the analyzer to detect an attack often determines the strength of the overall system.

C. Response Engine – The response engine controls the reaction mechanism and determines how to respond when the analyzer detects an attack. The system may decide either to raise an alert without taking any action against the source or may decide to block the source for a predefined period of time. Such an action depends upon the predefined security policy of the network.

III. CLASSIFICATION OF IDS

One of the main approaches of IDS, namely anomaly detection is based on the assumption that an attack on a computer system will be noticeably different from normal system activity, and an intruder will exhibit a pattern of behavior different from that of the normal user. In the second leading approach, misuse detection, a collection of known intrusion techniques is kept in a knowledge base, and intrusions are detected by searching through the knowledge base.

A. Anomaly Based Approach

Anomaly detection techniques assume that all intrusive activities are necessarily anomalous. This means that if we could establish a "normal activity profile" for a system, we could, in theory, flag all system states varying from the established profile by statistically significant amounts as intrusion attempts. However, if we consider that the set of intrusive activities only intersects the set of anomalous activities instead of being exactly the same, we find a couple of interesting possibilities: (1) Anomalous activities that are not intrusive are flagged as intrusive. (2) Intrusive activities that are not anomalous result in false negatives (events are not flagged intrusive, though they actually are). This is a dangerous problem, and is far more serious than the problem of false positives [3]. The main issues in anomaly detection systems thus become the selection of threshold levels so that neither of the above 2 problems is unreasonably magnified, and the selection of features to monitor. Anomaly detection systems are also computationally expensive because of the overhead of keeping track of, and possibly updating several system profile metrics.

B. Misuse Based Detection

The concept behind misuse detection schemes is that there are ways to represent attacks in the form of a pattern or a signature so that even variations of the same attack can be detected [4]. This means that these systems are not unlike virus detection systems -- they can detect many or all known attack patterns, but they are of little use for as yet unknown attack methods. An interesting point to note is that anomaly detection systems try to detect the complement of "bad" behavior. Misuse detection systems try to recognize known "bad" behavior. The main issues in misuse detection systems are how to write a signature that encompasses all possible variations of the pertinent attack, and how to write signatures that do not also match non-intrusive activity[5].

IV. OBJECTIVES

1. To develop systems which have broad attack detection coverage and which are not specific in detecting only the previously known attacks.

2. To reduce the number of false alarms generated by anomaly and hybrid intrusion detection systems, thereby improving their attack detection accuracy [6].

3. To develop anomaly intrusion detection systems which can operate efficiently in high speed networks without dropping audit data.

V. LITERATURE SURVEY

In recent times, intrusion detection has received a lot of interest among the researchers because it is widely applied for preserving the security within a network. Here, we present some of the techniques for intrusion detection. G. Gowrisona *et al.* [7] designed an intrusion detection system to classify the network behaviour with less computational complexity of O (n). The KDD Cup99 is a bench mark data used here to achieve promising classification rate. To achieve high detection rate in Intrusion Detection System (IDS), Shingo Mabu et al [8], described a fuzzy class association rule mining method based on Genetic Network Programming (GNP). GNP is used to enhance the representation ability with compact programs derived from the reusability of nodes in a graph structure. The combined method is evaluated with KDD99Cup and DARPA98 databases and showed that it provides competitively high detection rates.

However, to overcome the network based anomalies detection issue, Latifur Khan*et al.* [9] has proposed a method which was the combination of SVM and DGSOT, which starts with an initial training set and expanded it gradually using the clustering structure produced by the DGSOT algorithm. They compared the proposed approach with the Rocchio Bundling technique and random selection in terms of accuracy loss and training time gain using a single benchmark real data set. Due to the necessity of misuse and anomaly detection in a single system, M. Bahrololum*et al.* [10] proposed an approach to design the system using a hybrid of misuse and anomaly detection for training of normal and attack packets respectively. The utilized method for attack training was the combination of unsupervised and supervised Neural Network (NN) for Intrusion Detection System. By misuse approach known packets were identified fast and unknown attacks were also be detected.

For the importance of an efficient Intrusion Detection System, K.S. Anil Kumar and V. NandaMohan [11] proposed a combination of three techniques comprising two machine-learning paradigms. K-Means Clustering [16], Fuzzy Logics and Neural Network techniques were deployed to configure an effective intrusion detection system. This approach revealed the advantage of converging K-Means-Fuzzy-Neural network techniques to eliminate the preventable interference of human analyst in such occasions. Also, to improve the accuracy as well as efficiency of the Intrusion Detection System, Shekhar R. Gaddam*et al*.[12] presented "K-Means+ID3," a method to cascade k-Means clustering and the ID3 decision tree learning methods for classifying anomalous and normal activities in a computer network, an active electronic circuit, and a mechanical mass-beam system. Results showed that the detection accuracy of the K-Means+ID3 method was as high as 96.24 percent at a false-positive-rate of 0.03 percent on NAD; the total accuracy was as high as 80.01 percent on MSD and 79.9 percent on DED.

Vipin Kumar *et al*,[13], have analyzed NSL-KDD dataset to using K-means clustering. Clustering algorithms proves to be very useful when we have huge amount of unlabelled dataset. The study analyses the different types of attacks present in NSL-KDD. K-means Clustering applied here is able to efficiently detect new type of attacks present in dataset. K-means clustering is able to cluster the attacks present in training dataset into four major categories giving a better representation of the clusters. The main objective of the paper was to provide a complete analysis of the NSL-KDD dataset and the attacks presented. We used K-means algorithm for this purpose and also represented the distribution of instances in clusters providing better representation of the instances and making it clearer to understand.

Security is always an important issue especially in the case of computer network which is used to transfer personal information's, ecommerce and media sharing. So, Rachnakulhare and Divakar Singh, [14], have presented an intrusion detection system based on fuzzy C-means clustering and probabilistic neural network which reduced the training time and increases the detection accuracy. They evaluated the designed method based on KDD99 dataset and the simulation results showed that by selecting effective characteristics and proper training the detection accuracy rate up to 99% was achievable .

VI. EXISTING SYSTEM

The existing system uses a novel approach for ANN-based IDS, FC-ANN, to enhance the detection precision for low-frequent attacks and detection stability. The general procedure of FC-ANN approach has the following three stages. In the first stage, a fuzzy clustering technique is used to generate different training subsets. Based on different training sets, different ANNs are trained in the second stage. In the third stage, in order to eliminate the errors of different ANNs, a meta-learner, fuzzy aggregation module, is introduced to learn again and combine the different ANN's results. The whole approach reflects the famous philosophy ''divide and conquer". By fuzzy clustering, the whole training set is divided into subsets which have less number and lower complexity. Thus the ANN can learn each subset more quickly, robustly and precisely, especially for low-frequent attacks, such as U2R and R2L attacks. To illustrate the applicability and capability of the system, the results of experiments on KDD CUP 1999 dataset gives better performance.

VII . BLOCK DIAGRAM OF OUR APPROACH

The principle aim of the study is to create powerful system network intrusion detection system by using data mining and artificial intelligence methods. In this paper, intrusion detection system which utilizes Possibilistic Fuzzy C-Means (PFCM) clustering techniques and Fuzzy ARTMAP classifier is offered to have effective difference between the relevant data and intruded data .The system consists two layers were preparing and testing of information is carried out.

The proposed strategy incorporates two layers of preparing and testing procedure to ascertain if the input data is attack or not. The dataset utilized here within our proposed system is the KDD cup 99 dataset. At first in proposed intrusion detection system the input KDD cup 99 dataset pre-processing is carried out with a specific to get better classification accuracy. Since the KDD cup 99 dataset utilized here has different attacks. It comprises of both symbolic and numeric esteemed qualities. So utilizing this sort of dataset will lessen the reliability of our intrusion detection system. So in pre-processing, we outline symbolic-valued attributes of KDD cup 99 dataset to numeric-valued attributes. Then the pre-processed input data is connected to a clustering method called PFCM. The PFCM cluster the input pre-processed data into N number of clusters. Further the N number clusters are prepared utilizing our proposed Fuzzy ARTMAP classifier which resolves the complex classification problems. At last after different phases of our system, the test data is given to the trained fuzzy ARTMAP neural network classifier and test whether the input information is attack or not.

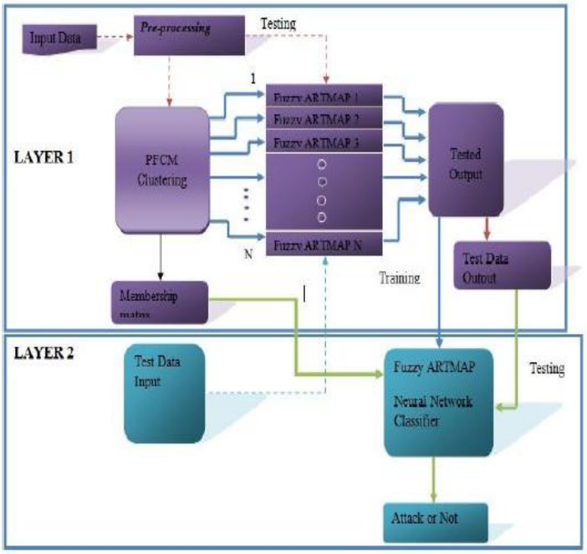

Figure 1: Block Diagram

A. *Possibilistic Fuzzy C-Means (PFCM) clustering*

Let the input pre-processed data be the huge dataset with a size of $M \times N$ .Within this processing, input large dataset is clustered utilizing our Possibilistic fluffy c-means clustering calculation. After using the pre-processed dataset to the PFCM clustering calculation N number of clusters are acquired C= {C_1, C_2, C_3…. C_N}, were N is the aggregate number of clusters. The particular single cluster C comprise of a vector of d estimations, were X= {X_1, X_2,X_3….X_D} . The capability of an individual data set is represented to as x_i and d denotes to the dimensionality of the vector. The PFCM is a just about the most effective parallel clustering technique.

The PFCM clustering algorithm includes various steps,

.

Step 1: Initialize $U = [u_{ik}] matrix, U^{(0)}$

Step 2: At k step: calculate the centers vectors $C^{(k)} = [v_i] with U^{(k)}$

$$v_i = \frac{\sum_{k=1}^{n} \left(u_{ik}^m + t_{ik}^\eta\right) x_k}{\sum_{k=1}^{n} \left(u_{ik}^m + t_{ik}^\eta\right)}, 1 \le i \le c.$$

Step 3: Update $U^{(k)}, U^{(k+1)}$

$$u_{ik} = \left(\sum_{j=1}^{c} \left(\frac{D_{ikA}}{D_{jkA}} \right)^{2/(m-1)} \right)^{-1}$$

Step 4: If $\left\| U^{(k)} - U^{(k+1)} \right\| < \varepsilon,$ then stop; otherwise return to step 2.

B. *Training stage*

Step: 1 the each obtained output cluster from the PFCM is trained using N number fuzzy ARTMAP neural Network classifier as shown in the fig below.

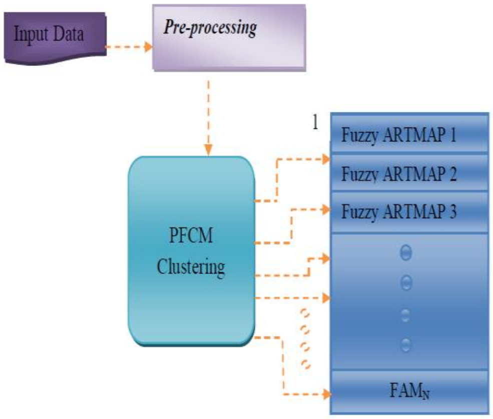

Figure 2: Training Process

Step 2: then a whole input data is given to the trained N number of neural network classifier and the output is

$$Z_K^{TR} = \left[z_{k1}^{TR}, z_{k2}^{TR}, z_{k3}^{TR}, ...z_{kN}^{TR} \right], k = 1,2,......n$$

Where, z_{kN}^{TR} is the output of fuzzy ARTMAP neural network N (FAM_N) and n is the number of training set.

Step 3: Then the input to the new fuzzy ARTMAP neural network classifier in layer 2 is formed with the membership function u_{ik} obtained from the PFCM clustering,

$$Z_{Input} = \left[z_1^{TR} * U_1^{TR}, z_2^{TR} .* U_2^{TR}, z_{k3}^{TR} * U_3^{TR}, ...z_N^{TR} * U_N^{TR}\right]$$

Where, U_N^{TR} is the membership function of the training dataset N belonging to cluster centre C ^{TR}

The input Z $_{Input}$ is trained using the new Fuzzy ATMAP neural network classifier in layer 2.

C. Fuzzy ARTMAP neural network classifier

Fuzzy ARTMAP is a neural network architecture which is depend on Adaptive Resonance Theory (ART). It has been used in supervised incremental learning, classification and reduction. The basic function of the fuzzy ARTMAP neural network classifier is operated by dividing the input space into a number of hyper boxes, which are mapped to an output space. The FAM has various advantages such as the classification task can be obtained by only one Fuzzy ART module and also computationally efficient. The structure of our proposed Fuzzy ARTMAP neural network classifier is shown in the figure below,

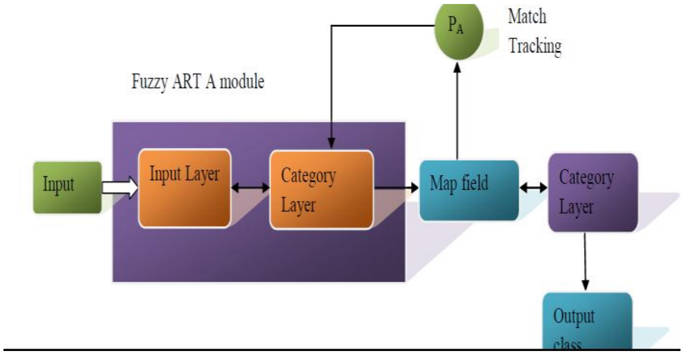

Figure 3: Fuzzy ARTMAP simplified structure

D. Testing Stage

In testing stage a new input test data is given as input to the n number of fuzzy ARTMAP neural network. Then the output of the test data from FAM is combined with the membership function u_{ik} obtained from the PFCM clustering. This input is applied to the new fuzzy ARTMAP neural network classifier in layer 2 which classifies whether the input is attack or not.

VIII. Comparisons and Results

A. Data preparation

The dataset contains about five million connection records as training data and about two million connection records as test data. And the dataset includes a set of 41 features derived from each connection and a label which specifies the status of connection records as either normal or specific attack type. These features have all forms of continuous, discrete, and symbolic variables, with significantly varying ranges falling in four categories:

(1) The first category consists of the intrinsic features of a connection, which include the basic features of individual TCP connections. The duration of the connection, the type of the protocol (TCP, UDP, etc.), and network service (http, telnet, etc.) are some of the features.

(2) The content features within a connection suggested by domain knowledge are used to assess the payload of the original TCP packets, such as the number of failed login attempts.

(3) The same host features examine established connections in the past two seconds that have the same destination host as the current connection, and calculate the statistics related to the protocol behavior, service,etc.

(4) The similar same service features inspect the connections in the past two seconds that have the same service as the current connection.

KDD'99 features can be classified into three groups:

1) Basic features: this category encapsulates all the attributes that can be extracted from a TCP/IP connection. Most of these features leading to an implicit delay in detection.

2) Traffic features: this category includes features that are computed with respect to a window interval and isdivided into two groups:

a) "same host" features: examine only the connections in the past 2 seconds that have the same destination host as the current connection, and calculate statistics related to protocol behavior, service, etc.

b) "same service" features: examine only the connections in the past 2 seconds that have the same service as the current connection[15].

The two aforementioned types of "traffic" features are called time-based. However, there are several slow probing attacks that scan the hosts (or ports) using a much larger time interval than 2 seconds, for example, one in every minute. As a result, these attacks do not produce intrusion patterns with a time window of 2 seconds. To solve this problem, the "same host" and "same service" features are re-calculated but based on the connection window of 100 connections rather than a time window of 2 seconds. These features are called connection-based traffic features .

3) Content features: unlike most of the DoS and Probing attacks, the R2L and U2R attacks don't have any intrusion frequent sequential patterns. This is because the DoS and Probing attacks involve many connections to some host(s) in a very short period of time; however the R2L and U2R attacks are embedded in the data portions of the packets, and normally involves only a single connection. To detect these kinds of attacks, we need some features to be able to

look for suspicious behavior in the data portion, e.g., number of failed login attempts. These features are called content features.

C. *Evaluation Metric*

The evaluation metrics used in our proposed method are true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Based on TP, TN, FP and FN, the performance of our intrusion detection system is evaluated by: a) Accuracy b) Sensitivity, c) Specificity.

a. Accuracy

The accuracy of our system is obtained by the following expression.
Accuracy = (TP+ TN) / (TP+TN+FP+ FN)

Accuracy means probability that our proposed system can correctly predict positive and negative examples.

b. Sensitivity
Sensitivity means probability that the algorithms can correctly predict positive examples.
Sensitivity = TP / (TP+FN)

c. Specificity
Specificity means probability that the algorithms can correctly predict negative examples
Specificity = TN/ (TN + FP)

IX. COMPARATIVE ANALYSIS

	Accuracy		**Sensitivity**		**Specificity**	
Cluster Size	**Proposed**	**Existing**	**Proposed**	**Existing**	**Proposed**	**Existing**
10	**0.92**	**0.84**	**0.8**	**0.76**	**0.8**	**0.76**
15	**0.9**	**0.81**	**0.83**	**0.74**	**0.83**	**0.74**
20	**0.88**	**0.8**	**0.76**	**0.61**	**0.76**	**0.61**
25	**0.87**	**0.78**	**0.73**	**0.65**	**0.73**	**0.65**

TABLE 1- Accuracy, Sensitivity and Specificity Obtained for Cluster Size 10, 15, 20, 25.

In this table for each cluster size, 10, 15, 20 and 25 the Accuracy, Sensitivity and Specificity values obtained for a both the proposed and existing method is noted.

From the above table we found that the accuracy obtained of our proposed method is, 0.92 for cluster size 10, 0.9 for cluster size 15 which is greater than the accuracy of the existing method which is, 0.84 for cluster size 10, 0.81 for cluster size 15 etc. Similarly the sensitivity obtained for our proposed method is 0.8 for cluster size 10, 0.83 for cluster size 15 which is greater than the existing method 0.76 for cluster size 10, 0.74 for cluster size 15. Likewise the values obtained for the specificity of our proposed method is greater than the specificity of the existing method and the respective plots for the Accuracy, Sensitivity and Specificity obtained for Cluster Size 10, 15 is plotted as given in the fig.4 and 5.

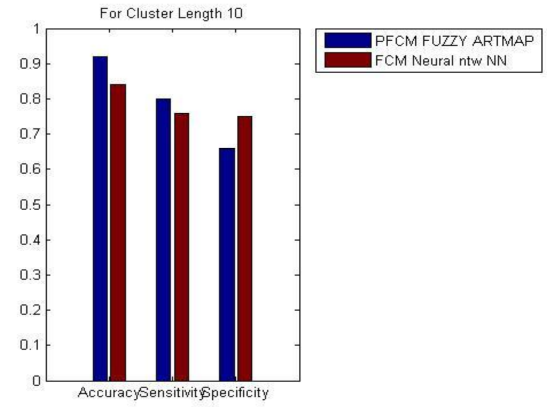

Figure 4: Accuracy, Sensitivity and Specificity plot Cluster Length 10

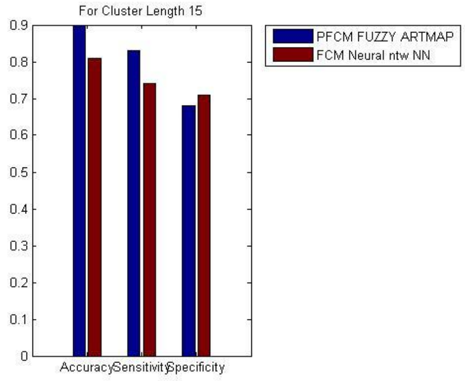

Figure 5: Accuracy, Sensitivity and Specificity plot Cluster Length 15

X. CONCLUSION

Nowadays network security is one of the major worry due to various attacks and vulnerabilities in internet. As a result, intrusion detection is an important component in network security. In this paper, we proposed a new intrusion detection system using PFCM clustering and layered fuzzy ARTMAP classifier. The PFCM is a clustering method used here which provides better clustering output compared to previously used clustering method. After this, the classification process is performed using the proposed fuzzy

ARTMAP neural network classifier. After various stages of training process test dataset is given as input and finally the classified output is obtained. The experimental results using the KDD CUP 1999dataset demonstrates the effectiveness of our new approach which provides better classification accuracy than the existing method. My Future work is to implement proposed fuzzy ARTMAP neural network classifier in various layers such as R2L, U2R, DOS and Probe layers for better performance.

REFERENCES

[1] Yao, J. T., S.L. Zhao, and L.V. Saxton, "A Study On Fuzzy Intrusion Detection", In Proceedings of the Data Mining, Intrusion Detection, Information Assurance, And Data Networks Security, SPIE, Vol. 5812, pp. 23-30 ,28 March - 1 April, Orlando, Florida, USA, 2005.

[2] Nivedita Naidu and Dr.R.V.Dharaskar, "An Effective Approach to Network Intrusion Detection System using Genetic Algorithm", International Journal of Computer Applications, Vol.1, No.3, pp.26–32, February 2010.

[3] Peyman Kabiri and Ali A. Ghorbani. Research on Intrusion Detection and Response: A Survey. International Journal of Network Security, 1(2):84–102, 2005

[4] B Mukherjee, L Todd Heberlein, K N Levitt, 1994. "Network intrusion detection. IEEE Network, Vol. 8, No. 3, pp.26–41,1994.

[5] J. Allen, A. Christie, and W. Fithen, "State Of the Practice of Intrusion Detection Technologies", Technical Report, CMU/SEI-99-TR-028, 2000.

[6] Kapil Kumar Gupta, Baikunth Nath and Ramamohanarao Kotagiri, "Layered Approach Using Conditional Random Fields for Intrusion Detection", IEEE Transactions on Dependable and Secure Computing, Vol. 7, No. 1, 2010.

[7]G. Gowrisona, K. Ramarb, K. Muneeswaranc, T. Revathic, " Minimal complexity attack classification intrusion detection system", Applied Soft Computing, vol 13, pp: 921–927, 2013.

[8]Shingo Mabu, Nannan Lu, Kaoru Shimada,KotaroHirasawa, " An Intrusion-Detection Model Based on Fuzzy Class-Association-Rule Mining Using Genetic Network Programming", IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, VOL. 41, NO. 1, PP: 130-139 , 2011

[9] Latifur Khan, MamounAwad, BhavaniThuraisingham, "A new intrusion detection system using support vector machines and hierarchical clustering", The International Journal on Very Large Data Bases, Vol. 16, no. 4, October 2007.

[10] M. Bahrololum, E. Salahi and M. Khaleghi "Anomaly intrusion detection design using hybrid of unsupervised and supervised neural networks", International Journal of Computer Networks & Communications, Vol.1, No.2, 2009.

[11] K.S. Anil Kumar and Dr. V. NandaMohan, " Novel Anomaly Intrusion Detection Using Neuro-Fuzzy Inference System ", IJCSNS International Journal 6 of Computer Science and Network Security, vol.8, no.8, pp.6-11 , August 2008.

[12] Shekhar R. Gaddam, Vir V. Phoha, Kiran S. Balagani, "K-Means+ID3: A Novel Method for Supervised Anomaly Detection by Cascading K-Means Clustering and ID3 Decision Tree Learning Methods", IEEE Transactions on Knowledge and Data Engineering, Vol. 19, No. 3, pp. 345-354, 2007.

[13] Vipin Kumar, Himadri Chauhan and Dheeraj Panwar, "K-Means Clustering Approach to Analyze NSL-KDD Intrusion Detection Dataset" International Journal of Soft Computing and Engineering (IJSCE), pp. 2231-2307, Volume-3, Issue-4, September 2013

[14] Rachnakulhare and Divakar Singh, "Intrusion Detection System based on Fuzzy C Means Clustering and Probabilistic Neural Network", International Journal of Computer Applications, Vol. 74, No.2, 2013.

[15]KDD Cup 1999. Available on: http://kdd.ics.uci.edu/databases/kddcup 99/kddcup99.html, Ocotber 2007.

[16] Jaskaranjit Kaur and Gurpreet Kaur, "Clustering Algorithms in Data Mining: A Comprehensive Study", International Journal of Computer Science and Engineering , vol. 3 Issue.7, pp 57-61, July 2015.

AUTHORS PROFILE

Kumar J S is working as an Asst. Professor in DMS SVH College of Engg., Machilipatnam. His areas of interest are Computer Networks, Network Security, and Data Mining.

Appa Rao S S V, working as an Associate Professor and HOD, Department of CSE, VISIT Engg. College, Tadepalligudem. His areas of interest are Data Mining, Network Security.

Subha Sree M is working as an Asst. Professor in DMS SVH Colleg of Engg.,Machilipatnam. Her areas of interest are Computer Networks, Network Security, and Data Mining, Aspect Oriented Programming.