DETERMINATION OF ERROR RATE THROUGH FCM ALGORITHM IN AN INTRUSION DETECTION SYSTEM

S.Vijaya Rani [1], Dr.G.N.K.Suresh Babu[2]
[1]Assistant Professor, MCA Department, Brindavan College, DwarakaNagar, Bangalore.
[2]Associate Professor, MCA Department, Acharya Institute of Technology, Bangalore.

Abstract
In the present world due to the advancement of internet technology, there are threats in network security. Strategies should be adopted to secure the network and to find out the violations occurred in the security policies. Intrusion detection system is one of the better techniques to monitor and detect the security violations in a computing system. The attack pattern can be learned through supervised and unsupervised learning. Clustering method is one of the unsupervised learning used to detect an attacked packet. Also ABC, PSO algorithms could be materialized for detecting the intrusions. But in FCM algorithm, the cluster centers minimize the dissimilarity function. Hence it reduces the error rate. This paper elucidates the error rate optimization of various algorithms in the detection of attacked packets.

Keywords: FCM, Intrusion Detection, Error rate, Cluster, Attack, KDD cup99, Network, Normalization, LAND attack, misuse detection, anomaly detection, malicious packets.

I INTRODUCTION

A neural network consists of many processing units in distributed manner and can store information and make it readily for use. It behaves like human brain and acquires knowledge from the environment through some learning process. Several algorithms are available for learning process. Here artificial neural algorithm could be used for optimizing the error rate in detection of malicious packets.

Intrusion detection system is one of the device for security of network. Juniper, Radware, cisco, D-Link, Axent, cybersafe and shadow are premiers in manufacturing IDS. Network security has to be addressed properly due to the mounting security concern in today's network. The vulnerabilities of computer systems might be by unauthorized individuals and the misuse of system resources by authorized system users. The need for effective intrusion detection system is for prevention of attacks and vulnerabilities caused by legitimate and illegal users and hackers.[6]

II TYPES OF COMPUTER NETWORK ATTACKS

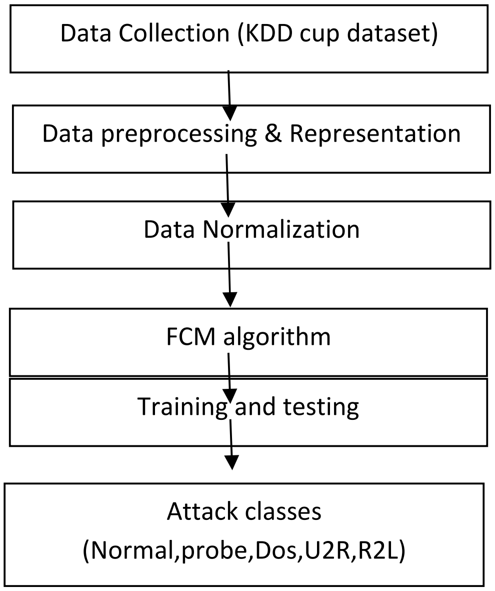
The network attacks can be divided in to four categories such as DoS , Probe, U2R, R2L attacks[2]. LAND attack, Back attack, synflooding attack, smurf attack, teardrop attack falls under Dos attack. Portsweep, IPsweep,[5] NMAP, Satan are Probe attacks. Buffer overflow, rootkit, perl attack, loadmodul are user to root attacks. Similarly remote to local attacks are spy, phf, multihop, imap, warezclient, warezmaster etc.

III METHODS OF INTRUSION DETECTION SYSTEM

The basic principle of intrusion detection lies on the deviations of normal data to that of unknown functions. It can be of misuse detection, anomaly detection and specification based detection. Pattern matching, rule based techniques, state based techniques and data mining are types of misuse detection[3]. Advanced statistical models, Biological models and learning models are examples of anomaly detection. Specification based detection is a combination of both misuse and anomaly detection. Clustering is an unsupervised learning process [8]where the samples are divided into categories which the members are having same qualities. Classic clustering deals with the fact that the sample belongs to one cluster may not be the member of several clusters. Hence, if a sample is having qualities of more than one cluster, it would be very difficult to recognize the behavior of cluster. In fuzzy clustering, the sample can belong to more than one cluster.

IV PROPOSED MODEL OF INTRUSION DETECTION SYSTEM

The dataset used here is KDD cup 1999 dataset, which are available in the MIT Lincoln labs and in KDD website. Before training the dataset, it should be preprocessed to remove the redundancy[1]. Further the non numerical entities must be represented in numerical form correctly. After normalization the instances are assigned to every cluster and the center of clusters are calculated. The best solution is calculated by repeated equations until the best value id fixed among the clusters .


Data Collection (KDD cup dataset)
Data preprocessing & Representation
Data Normalization
FCM algorithm
Training and testing
Attack classes
(Normal,probe,Dos,U2R,R2L)

V FCM ALGORITHM

The FCM algorithm was introduced by J.C.Bezdek. The fuzzy logic is a set of rules which are often used to increase the detection accuracy. This approach is justified in the stage of design of any IPS to protect information in the conditions of uncertainty of data processing algorithms. The amount of incoming and outgoing traffic, number of incorrect packets, number of flags in packets, number of flows in the network decides the nature of attack. The analysis may be non linear and difficult to predict network traffic forecasting due to fluctuations in network traffic, complexity of application, system software, data transmission protocols and non homogeneity of information flows.

However, the prediction of the network traffic can be evaluated in a certain time interval, such as ,

t_0+T

$Y_i = \Sigma\, \alpha_t \mid x_t \mid$

$t=t_0$

where x_t - Measured value at the time of moment t

α_t- Weight coefficients, characterizing the significance of measurement at the moment t,

T - Measurement period.

The discrete chance quantity of independent attack occurrence can be represented as ,

$Tr_A=\{A_1 \,/\, P_1,\ A_2 \,/\, P_2,\ \ldots, A_n \,/\, P_n\ \}$ where $A_1, A_2, \ldots, A_n$ are fuzzy values that the chance quantity posses with probabilities $P_1, P_2, \ldots, P_n$.

Sensor fusion is defined a the process of collecting information from multiple and possibly heterogeneous sources and combining them to obtain a more descriptive, meaningful and solid results. The fusion can occur at the various levels

The algorithm uses the weights that minimize the total weighed mean square error. [4]

$J(W_{qk},z^{(k)})= \Sigma_{(k=1,k)} \Sigma_{(k=1,k)}(W_{qk}) \| x^{(q)} - z^{(k)} \|^2$

$\Sigma_{(k=1,k)}(W_{qk}) = 1$ for each q

$W_{qk} = (1/ D_{qk})^2)^{1/(p-1)} / \Sigma_{(k=1,k)} (1/ D_{qk})^2)^{1/(p-1)}$, $p>1$

[7]FCM permits each feature vector in every cluster with fuzzy truth value varying from 0 to 1, it is calculated by above equations. The algorithm gives a feature vector to the each cluster according to the maximum weight of the feature vector with respect to all the clusters. Its advantages are

a) Gives better result for overlapped data set than k-means algorithm.
b) Unlike k-means data point may belong to more than one cluster center.

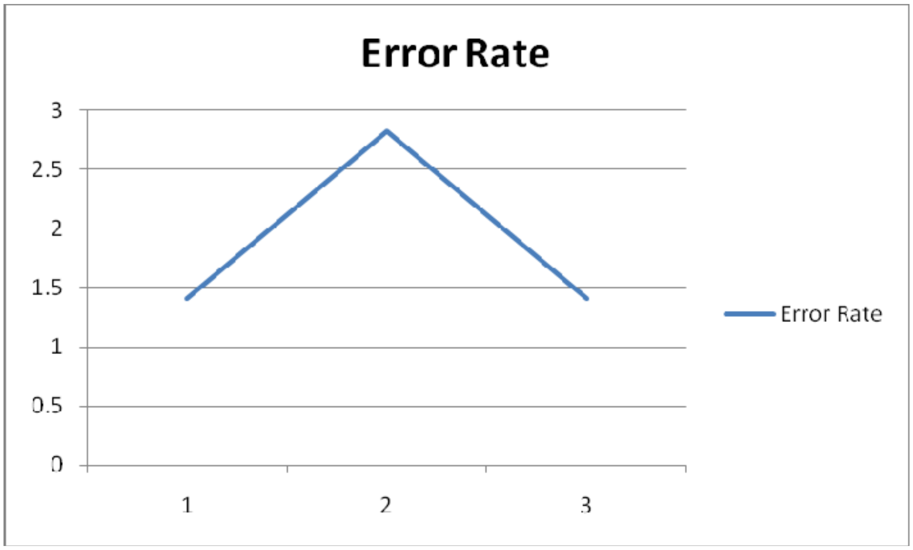
The disadvantages of FCM are A priori specification of the number of clusters and with lower value of β we get the better result but at the expense of more number of iteration.

The error rate is checked for three instances, with KDD dataset of all attack types varying from 25000 to 68000. The error rate is directly proportional to the number of input data set taken. The result is given as below.

Error Rate	Iteration	Mean	Standard Deviation
1.4128	327	163	94.68
2.825	901	450	260.38
1.4128	330	165	95.551

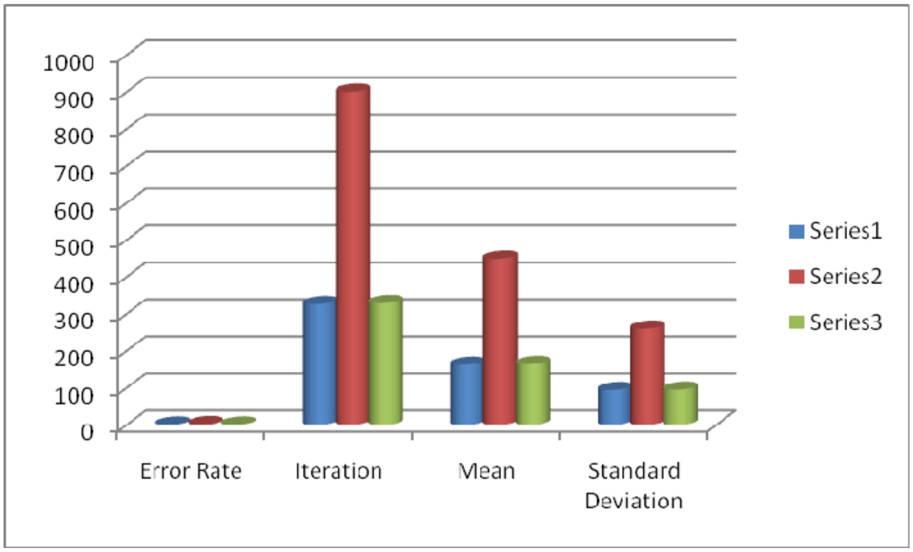
Table1. Result obtained by FCM Algorithm

The error rate is graphically represented as below,



Graph1. Representation of error rate.

The error rate, iteration, mean and standard deviation is depicted here under,



Graph2. Representation of output parameters by FCM algorithm

VI CONCLUSION AND FUTURE WORK

FCM is a clustering aspect to form clusters for entities having the same qualities. By optimization we can achieve minimum number of replicates and clusters with the similar entities. As the number of entities are increasing, the error rate is also increasing, which is the drawback of this algorithm, Anomaly based intrusion detection is also a problem to be addressed. It is a scope for modifying the cluster function to reduce the error rate.

VII REFERENCES

[1] "Intrusion Detection using Artificial Neural Network" By Poojitha G, 978-1-4244-6589-7/10/$26.00@2010 IEEE.

[2] "Classifying Attacks in a Network Intrusion Detection System based on Artificial Neural Networks" by Mohammad Reza Norouzian, ISBN 978-89-5519-155-4, Feb. 13-16,2011 ICACT2011.

[3] "A New Network Intrusion Detection Identification Model Research" by WenJie Tian, 978-1-4244-5194-4/10/$21.00@2010 IEEE.

[4] "Comparison of Fuzzy c-means algorithm and New Fuzzy Clustering and Fuzzy Merging Algorithm" by Liyan Zhang, Professional Paper, May 2001.

[5] "Evaluation of Fuzzy K-Means and K-Means clustering algorithms in intrusion detection systems" by Farhad, Neda, Zeinab published in IJSTR@2012 , pages 62 to 72, www.ijstr.org.

[6] "Intrusion Detection System : A Review", by Sanjay Sharma and R.K.Gupta, published in International Journal of Security and its Applications, vol.9.No.5(2015).pp.69-76.

[7] 'Some clustering algorithms to enhance the performance of the network intrusion detetion system" by Mrutyunjaya Panda, Manas Ranjan Patra in Journal of Theoretical and Applied Information Technology, 2005-2008 ,pp 710 to 716,www.jatit.org

[8] "Fuzy Clustering as an Intrusion Detection Technique", by Disha Sharma in International Journal of Computer Science & Communication Networks, vol 1(1), sep- Oct 2011