International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-12 E-ISSN: 2347-2693

Intruder Detection Using Fuzzy Min-Max Neural Network and A Principal Component Analysis (PCA) in Network Data

A. F. Aldubai[1*], V. H. Humbe[2], S.S. Chowhan[3], Y.F. Aldubai[4]

[1*]Department of Computational Sciences and Technology, S.R.T.M University, Nanded, India
[2] School of Technology, Sub-Center Latur S.R.T.M University, latur, India
[3]Department of Computational Sciences and Technology, S.R.T.M University, Nanded , India
[4]Faculty of Administrative Sciences, Albaydaa University, Radaa , Yemen

[*]*Corresponding Author: Ahmed_Aldubai86@yahoo.com, Tel.: +91 8669384165*

Available online at: www.ijcseonline.org

Received: 11/Nov/2017, Revised: 23/Nov/2017, Accepted: 12/Dec/2017, Published: 31/Dec/2017

***Abstract*—** To improve classification performance and prediction time of (IDS) system when a few number of large hyperboxes are formed in the network. We have proposed FMM NN and PCA features extraction algorithms. The process of a system consisting of the pre-processing dataset based on interquartile ranges filter and features extraction. Followed by the Fuzzy min-max neural network (FMM NN), which define as supervised learning classifier, that utilizes a fuzzy set hyberbox as pattern classes for learning and classification. Each hyberbox consists of min and mix points of opposite corners of hyberbox with corresponding to membership function. Two real-time and faster streaming datasets (KDD99 and NSL-KDD) are used to empirically evaluate the effectiveness of the proposed FMM NN system. The results are analyzed and compared with others existing systems and published results.

Keywords— Intrusion Detection, Fuzzy min-max neural network, Principal Component Analysis (PCA), Machine Learning

I. INTRODUCTION

Intrusions are deliberate illegal acts on the system or network, and they appear in specific forms such as illegal intrusion of the system and network, modification or leakage of important information, illegal use, damage, computer viruses and denial of service. IDS can be divided into network-based IDS and host-based IDS depending on the target of monitoring. Host-based IDS is a software installed on the system that monitors the activity of users inside a system and detects hacking attempts. Network-based IDS detects intrusions by analyzing packets passing through the network based on packet capture of the network [1]. An intrusion detection system is a software that detects and responds to these illegal intrusions quickly. There are a variety of software from simple log file analysis to complex real-time intrusion detection systems. Intrusion detection schemes can be roughly divided into abnormal intrusion detection methods and misuse intrusion detection methods [2]. IDS will monitor suspicious activity in a way that can be taken if unauthorized users fail to prevent unauthorized attempts to access or attempt to intercept, misuse, abuse, It is a system that aims to discover and real-time processing. It's the next generation security solution following the Firewall. The main reason for this is that if the intrusion prevention system fails effectively, it minimizes the damage, There is a growing demand for security solutions that can respond. An intrusion detection system complements the security limitations of an intrusion prevention system by blocking illegal intrusion according to simple rules [3]. Firewalls allow certain rules to be applied to allow and deny special services and hosts. If the network traffic matches the allowed packet. Allow access to the service. It does not care what packets are included. IDS detects hackers' intrusion based on hacking method, so the application of new technology is fast. It can block not only attacks from outside but also hacker by insiders. Accordingly, the internal attacker's hacking through the ID theft which is not

supported by the existing firewall is also blocked. The existing firewall blocks the attack by authenticated IP, so if the hackers attack with the authenticated IP, the firewall becomes useless[4]. However, the IDS checks for all the packets attributes regardless of IP. It is more secure. Immediately responds to system intrusions, when a hacking fact is discovered, information about the hacking can be immediately sent to a mobile phone, pager, e-mail, etc., and system security can be maintained even in the absence of a network administrator[5].

Machine learning and data mining can do prediction, classification, clustering, and association rule mining [6,7,8] between the data items. So the field of intrusion detection system based on the data mining is the best encapsulation of the technological framework to overcome the intrusion in networking. Our major focus in this paper is the Fuzzy Min-Max neural network as the pattern classification system. FMM NN can be used for tackling classification (supervised) problems. which utilized hyperbox fuzzy sets, each fuzzy set is an n-dimensional box defined by a set of minimum and maximum points. Each input pattern is classified based on the degree of membership to the corresponding boxes [9]. The size of hyperboxes is a user-defined value bounded between zero and one. A smaller hyperbox size means that the hyperbox can contain only a smaller number of patterns, which will increase the network complexity. A larger hyperbox size means that the hyperbox can contain a larger number of patterns, and will decrease the network complexity. However, large hyperboxes may lead to a low classification performance as the hyperboxes may not capture salient features of the input patterns.The processing of our proposed IDS consists various steps. The first step is preprocessing of the KDD dataset by PCA features extraction algorithms, removing the duplicate entries, and extreme values. The second step is training and learning the system using FMM NN algorithm. lastly, classification and testing the system to evaluate the performance of the system. All steps of proposed IDS have their own significance. Which cannot be ignored? IDS, must be able to keep pace with development and enhance the performance. Due to duplicate entries, unbalanced data distribution, large numbers of the attributes, high volume of the data, availability of the noise, and extreme values. So there is a need for network IDS which can automatically extract useful information from a huge volume of the data which is previously unknown.[10, 11].

The main contributions of this paper are to improve the classification performance of intruder detection over network traffic, and reducing the predicting time using the FMM NN and PCA features extraction algorithms. To test the efficiency of the proposed methods with a comparison to the existing methods.

This paper is planned as follows Section II related work. Section III gives a brief description of FMM NN. Section IV describes the Fuzzy Min-Max Learning Algorithm. Section V Learning Algorithm for IDS. The experimental studies and the results obtained are explained and discussed in section VI. The conclusion of the paper is presented in the section.

II. RELATED WORK

Wang et al. [12] a fuzzy clustering and ANN based intrusion detection are proposed. In this first fuzzy clustering method is used to generate the training subsets and the different ANN are learned on those subsets and finally fuzzy aggregation module to aggregate the results from the learned ANN's. The system is learned and evaluated using the KDD CUP dataset. Chandrashekhar Azad et al. [13] They proposed a hybrid IDS which is based on the fuzzy min-max neural network and the particle swarm optimization. The proposed system is tested with the help of preprocessed KDD CUP data set. Classification accuracy and classification error are taken as a performance evaluation parameter to test the effectiveness of the system. The proposed system is compared with MLP classifiers, multilayer perceptron, RBF classifier, RBFN Classifiers, SMO, naïve Bayes, LibSVM, KDD Cup Winner, KDD Cup Runner UP, FMM and FMM GA etc., the results show that the proposed system performed well as compared to the other systems. Ghorbani et al. [14] proposed Feedforward neural network and the back propagation network-based IDS for unusual traffic in the computer network. The system is evaluated using the DARPA dataset. Basant Subba et al. [15] They proposed IDS model uses the feedforward and the backpropagation algorithms along with various other optimization techniques to minimize the overall computational overhead, while at the same time maintain a high-performance level. Experimental results on the benchmark NSL-KDD dataset shows that the performance (accuracy and detection rate) of the proposed ANN-based IDS model is at par and in some

cases even better than other IDS models. Azad Chandrashekhar et al. [16] Enhanced intrusion detection system has been proposed, which is based on the fuzzy min-max neural network. The main component of the fuzzy min-max neural network is the fuzzy hyperbox to find the decision boundary between the various overlapping classes. The proposed system has been the online adaption facility, nonlinear separability. The most important property of the proposed system is the lesser time requirement for the learning against the tradition neural network like backpropagation neural network, and Boltzmann neural network. The traditional network takes the multiple passes in the learning process. The system is trained and evaluated using two datasets (KDD Cup 99 and NSL-KDD). The proposed system provides superior performance as compared to multilayer perceptron classifiers, Naïve Bayes classifier, SMO classifier, logistic regression classifier, LibSVM classifier and other published results. The result of existing systems is evaluated on WEKA with default values.

III. FUZZY MIN-MAX NEURAL NETWORK

Fuzzy sets were proposed by Lofti A. Zadeh as a paper entitled Fuzzy sets only in 1965. This paper was the establishment of all fuzzy logic that followed by a means of representing and manipulating data that were not precise, but rather fuzzy [17]. This paper describes a Fuzzy min max classifier neural network (FMM NN) that creates classes by aggregating several smaller fuzzy sets into a single fuzzy set class. This technique can learn pattern classes in a single pass through the data, it can add new pattern classes on the fly, it can refine existing pattern classes as new information is received [18].

Fuzzy min max classification neural network (FMM NN) is based on the concept of the hyper-box fuzzy sets. The FMM neural network is a collection of the N input vectors as an input to the neural network, B number of the hyper box as a hidden layer and the M class labeled as an output to the neural network. Each input is connected to all the hyperboles and each hyper box is connected to a particular class label. A hyper box represents a region in the n-dimension that contain pattern vectors those who have the full class membership value. The hyper box is defined by its min–max point and the hyper box membership function. The membership function is defined with respect to this hyperbox min-max. In this neural network, the learning is performed by the series of expansion and contraction process and the testing of the patterns is carried out by the finding the membership value of each input pattern corresponding to each hyper box fuzzy sets. The membership value is lies between the 0 and 1, 1 means the full class membership and the 0 mean no membership. Figure 1 shows the architecture of the FMM NN.

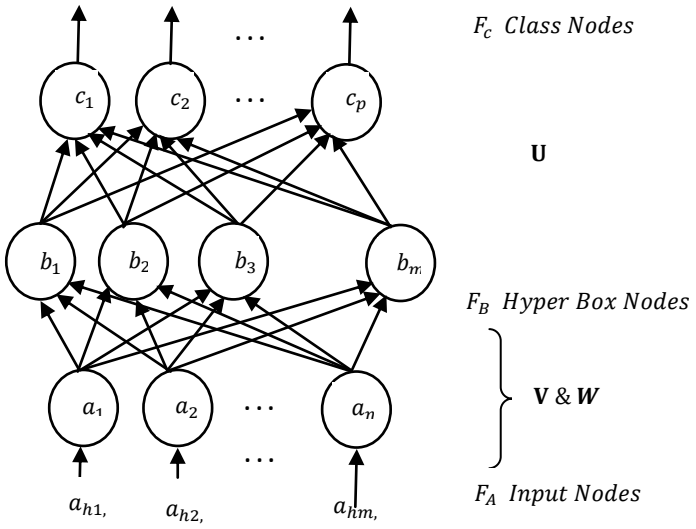

Figure. 1. Topology of Fuzzy Min - Min Neural Network

Fuzzy min max classification neural network (FMM NN) is an example of the last type of neural network classifier. Implementing the fuzzy min-max classifier as a neural network, it is possible to immediately exploit the parallel nature of the classifier and provide a mechanism for fast and efficient implementations. The main target of implementing Fuzzy min max classification neural network (FMM NN) for intruder detection is real-time adaption, parallel classifier, single pattern scanning of the training set, less training time requirement, and nonlinear separability etc.

A. Fuzzy set hyperbox

A hyperbox defines a region of the n-dimensional pattern space that has patterns with full class membership. The hyperbox is completely defined by its minimum and maximum points. Fuzzy hyperbox classification of data has been shown to be a powerful algorithmic approach. The quality of coverage, measured as a misclassification rate, depends on the maximum size of hyperboxes. The smaller is the

maximum size of hyperboxes the more accurate coverage can be obtained.

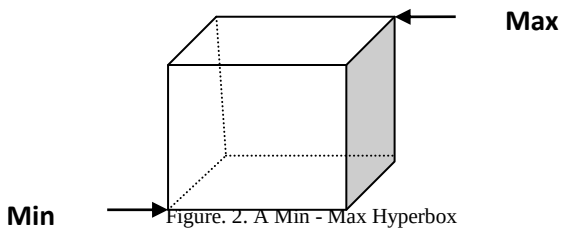

Figure. 2. A Min - Max Hyperbox

The definition of each hyperbox fuzzy set B_i is

$$B_{j=\{X,V_j,W_j,F(X,V_j,W_j)\}} \quad \forall X \in I^n \qquad (1)$$

Here, B_i is the J^{th} hyperbox, A is the n-dimensional input vector or pattern, Where, X is the input, In is a unit cube with n dimensions and V_j and W_j are the min and max points, respectively.

The size of the hyperboxes θ can have a value between 0 and 1. A small value of θ will produce Where, X is the input, In is a unit cube with n dimensions and V_j and W_j are the min and max points, respectively small-size hyperboxes, and vice versa.

The min and the max points of hyperbox B_i are defined as:

$$V_j = (V_{j1}, V_{j2}, V_{j3}, V_{jn}) \qquad (2)$$

$$W_j = (W_{j1}, W_{j2}, W_{j3}, W_{jn}) \qquad (3)$$

Using equation 1 the aggregate fuzzy set defines the K^{th} pattern class C_k is defined as

$$C_K = \bigcup_{j \in k} B_j \qquad (4)$$

where K is the index set of those hyperboxes associated with class k. Note that the union operation in fuzzy sets is typically the max of all of the associated fuzzy set membership functions.

B. Membership function

A membership function for a hyperbox B_i on input pattern A_j is defined as μ_A: $A_j \rightarrow [0,1]$, where each element of A_j is mapped to a value between 0 and 1. This value, called membership value or degree of membership, quantifies the grade of membership of the element in A_j to the hyperbox B_i [19]

The input patterns are classified depending on how much they are "contained" by a hyperbox. When the membership value is closer to 1, it means that the pattern is more contained in the hyperbox; 1 represents the full membership. The membership function for the J^{th} hyperbox $b_j(A_h)$, $0 \le b_j(A_h) \le 1$, must measure the degree to which the h_{th}input pattern $A_h$falls outside of the hyperbox B_i formed by the min point V_j and the max point W_j. As A_h approaches 1, and when the point is contained within the hyperbox, $b_j(A_h, V_j, W_j) = 1$. The resulting membership function is defined as

$$b_{j(A_h)} = \frac{1}{2n}\sum_{i=1}^{n}\left[\max\left(0, -\max\left(0, \gamma \min\left(1, a_{hi} - w_{ji}\right)\right) + \max\left(0, -\max\left(0, \gamma \min\left(1, v_{ji} a_{hi}\right)\right)\right)\right] \qquad (5)$$

$A_h = (a_{h1}, a_{h2}, \ldots a_{hn}) \in I_n$ is the h_{th} input pattern, $V_j = (v_{j1}, v_{j2}, \ldots v_{jn})$ is the Min point and Max point for b_j is $W_j = (w_{j1}, w_{j2}, \ldots w_{jn})$ and γ is the sensitive parameter $0 \le \gamma \le 1$, which governs how fast the membership value decreases outside the hyperbox as A_h and b_j increases.

IV. FUZZY MIN-MAX LEARNING ALGORITHM

The supervised FMM learning algorithm is the combination of the three general steps, which are an expansion, overlap test, and the contraction process. The learning process begins by selecting and preprocessing the given data set. In our experimental, we have used KDD dataset for classification of intruder detection. Next, is Identify the hyperboxes, determine if any overlap exists between hyperboxes from different classes, and remove the overlap by minimally adjusting each of the hyperboxes. The following section describes the three general steps in detail.

A. Hyperbox Expansion:

The goal of the hyperbox expansion process is to Identify the hyperbox that can expand and expand it. If an expandable hyperbox cannot be found, add a new hyperbox for that class.

Given the h_{th} training pair (A_h, d_h), find the hyperbox B_j which provides the highest degree of membership, allows expansion if necessary and represents the same class as d_h the degree of membership is measured using equation 5. For the

hyperbox B_j to expand to include A_h the following constraint must be met.

$$n\theta \geq \sum_{i=1}^{n} (\max(w_{ji} a_{hi}) - \min(v_{ji} a_{hi})) \qquad (6)$$

If the expansion criterion has been met for hyperbox B_j min point of the hyperbox is adjusted using the equation.

$$v_{ji}^{new} = \min(v_{ji}^{old}, a_{hi}) \qquad (7)$$

and the max point is adjusted using the equation.

$$w_{ji}^{new} = \max(w_{ji}^{old}, a_{hi}) \qquad (8)$$

B. Hyperbox Overlap Test:

The extension of the hyperboxes can create hyperbox overlap. The overlap of hyperboxes that have the same class labels does not present any problem but the overlap of hyperboxes with different class labels must be prevented since it would create ambiguous classification [18]. So the main function of overlap test is to determine if any overlap exists between hyperboxes from different classes.

Considering that the hyperbox Bj was expanded from the above step and that the hyperbox Bk represents another class and it is to be tested for possible overlap class. Assuming $\delta^{old} = 1$ initially, the four test cases is satisfied for each of the n-dimension.

$$\begin{aligned} &\textbf{case 1}: v_{ji} < v_{ki} < w_{ji} < w_{ki}, \\ &\textbf{case 2}: v_{ki} < v_{ji} < w_{ki} < w_{ji}, \\ &\textbf{case 3}: v_{ji} < v_{ki} < w_{ki} < w_{ji}, \\ &\textbf{case 4}: v_{ki} < v_{ji} < w_{ji} < w_{ki}. \end{aligned} \qquad (9)$$

C. Hyperbox Contraction:

If overlap between hyperboxes that represent different classes does exist, remove the overlap by minimally adjusting each of the hyperboxes.

If $\Delta > 0$, then the Δth dimensions of the two hyperboxes are adjusted. To determine the proper adjustment to make, the same four cases are studied.

$$\textbf{Case 1}: v_{ji\Delta} < v_{k\Delta} < w_{j\Delta} < w_{k\Delta}, \quad w_{j\Delta}^{new} = v_{k\Delta}^{new} = \frac{w_{j\Delta}^{old} + v_{k\Delta}^{old}}{2}$$

$$\textbf{Case 2}: v_{k\Delta} < v_{j\Delta} < w_{k\Delta} < w_{j\Delta}, \quad w_{k\Delta}^{new} = v_{j\Delta}^{new} = \frac{w_{k\Delta}^{old} + v_{j\Delta}^{old}}{2}$$

$$\textbf{Case 3i}: v_{j\Delta} < v_{k\Delta} < w_{k\Delta} < w_{j\Delta}, \textbf{and } (w_{k\Delta} - v_{j\Delta}) < (w_{j\Delta} - v_{k\Delta}), \quad v_{j\Delta}^{new} = w_{k\Delta}^{new}.$$

$$\textbf{Case 3ii}: v_{j\Delta} < v_{k\Delta} < w_{k\Delta} < w_{j\Delta}, \textbf{and } (w_{k\Delta} - v_{j\Delta}) > (w_{j\Delta} - v_{k\Delta}), \quad w_{j\Delta}^{new} = v_{k\Delta}^{old}.$$

$$\textbf{Case 4i}: v_{k\Delta} < v_{j\Delta} < w_{j\Delta} < w_{k\Delta}, \textbf{and } (w_{k\Delta} - v_{j\Delta}) < (w_{j\Delta} - v_{k\Delta}), \quad w_{k\Delta}^{new} = v_{j\Delta}^{old}.$$

$$\textbf{Case 4ii}: v_{k\Delta} < v_{j\Delta} < w_{j\Delta} < w_{k\Delta}, \textbf{and } (w_{k\Delta} - v_{j\Delta}) > (w_{j\Delta} - v_{k\Delta}), \quad v_{k\Delta}^{new} = w_{j\Delta}^{old}. \qquad (10)$$

V. LEARNING ALGORITHM FOR IDS

A. Training phase

Input: Training set, Testing set, L(lamda), G (sensitivity parameter(Gamma)), V, W, N (No of classes), n (No. elements in pattern vector), P (Index for counting hyperbox modification & No of patterns in the feature matrix)

Output: Save min_max_pt.mat W V U RH lamda gamma n N P zz;

Process:

Step 1. Start

Step 2. Load the Training set

Step 3. Specify the initial parameters L(lamda), and G (sensitivity parameter(Gamma)).

Step 4. Appending class index RH

Step 5. Initialization of hyperboxes W V

Step 6. Learning of Fuzzy Min Max Neural Network begins.

Step 7. 1. Flag = 1 to indicate the expansion of hyperbox

Step 7. 2. Flag = 2 to indicate the result of overlap test of hyperboxes that indicates need of Contracton process

Step 7. 3. Find membership values of hyperboxes corresponding to a class of applied pattern and store in vector b

Step 7.4. Calculate index of hyperbox to be tested for expansion and store the index in the variable max_val_index

Step 7. 5. Test the hyperbox for "Expansion test"

Step 7. 6. Expand the hyperbox if condition is satisfied else create new hyperbox

Step 8. Test the hyperbox for "Overlap test" to be carried if flag1=1

Step 8.1. While testing for overlap find the smallest overlap along any dimension and index of dimension

Step 9. Save min_max_pt.mat W V U RH lamda gamma n N P zz;

Step 10. Training ends here

B. Testing phase

Input: Testing set, P (Index for counting hyperbox modification & No of patterns in the feature matrix)

Output: FMM neural Networks

Process:

Step 1. Start

Step 2. load min_max_pt_1000.mat W V U lamda gamma n N zz%RH;

Step 3. Load the Testing set

Step 4. Appending class index RH

Step 5. Count for counting efficiency of recognition

Step 6. Calculate output of hyperboxes

Step 7. Calculate output of nodes in third (FD)layer that gives soft (fuzzy) decision

Step 8. Program segment to get hard decision on the output

Step 9. Save record.mat p cnt

Step 10. Testing ends here

VI. SIMULATION RESULTS AND DISCUSSIONS

This section describes the process and result of the fuzzy min-max Neural Network-based on intrusion detection system (FMM NN IDS). This simulation is carried out in MATLAB R2016a with two datasets (KDD CUP and NSL-KDD). The training data include the two types of the network data, normal and attack patterns. The attack pattern is of the different categories like DoS, U2R, R2L, Probe, and Normal. The test dataset used which is the subset of the KDD Cup dataset that includes some attacks that are not in the training set. The dataset contains a total of 41 attributes and one as a class attribute or the decision attribute, which describes the category of the traffic pattern whether it is normal or attack. The total number of instances in both datasets are (126887 and 11850) instances, 70% of instances are used for training and the remaining of instances are used as a testing dataset. Those datasets first are preprocessed through a number of steps. Started by converting nominal attributes to numeric using the grp2idx function. follow by removing the duplicate entries and extreme values using filtering based on interquartile ranges concept of weka. lastly is featured extractor using PCA algorithm. Due to PCA, the features are reduced in both datasets into (9 and 12) features. The FMM-based IDS is tested on sensitivity parameter Gamma = 1 and on the different lamda values. Table.1 shows the result of the effectiveness, classification accuracy, classification error, training time, and testing time of FMM NN using different lamda Sigma values and full features and reduced features of KDD dataset. The performance in term of accuracy for both full features and reduced features dataset is perfect in both but in term of training time and testing time the KDD dataset with reduced features dataset is match better than the full features dataset.

Table.1 shows the result of the effectiveness, classification accuracy, classification error, training time, and testing time of FMM using different lamda values and full and reduced features of KDD dataset.

KDD dataset with Full features				
Θ	***Training Time Sec***	***Testing Time Sec***	***Accuracy %***	***Classification Error %***
0.00001	**455**	**372**	**98.96**	**01.04**
0.000045	**474**	**382**	**96.90**	**03.01**
0.000075	**466**	**342**	**92.03**	**07.97**
KDD dataset with 9 features				
Θ	***Training Time Sec***	***Testing Time Sec***	***Accuracy %***	***Classification Error %***
0.001	**125**	**76**	**99.08**	**00.92**
0.0020	**102**	**71**	**97.03**	**02.97**
0.0030	**109**	**75**	**93.00**	**07.00**

Figure 3 and 4 shows the comparison result of KDD dataset, in terms of accuracy and training time against different values of lamda.

The performance in term of accuracy for both is almost same and perfect but in term of training time, the reduced features are totally matched better than the full features dataset.

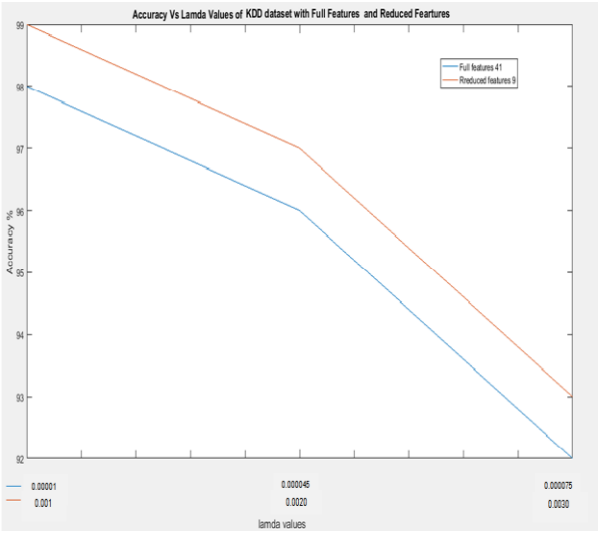

Fig. 3. Accuracy Vs Lamda Values of KDD dataset with Full Features and Reduced Features

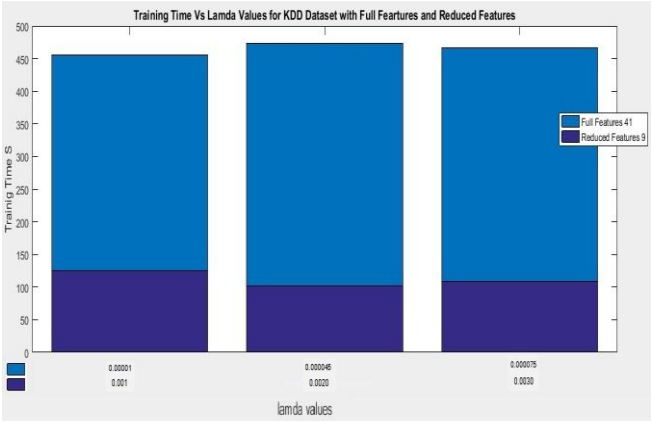

Fig. 4. Training Time Vs Lamda Values of KDD dataset with Full Features and Reduced Features

Table.2 shows the result of the effectiveness, classification accuracy, classification error, training time, and testing time of FMM NN using NSL-KDD dataset and different lamda values.The performance in term of accuracy and training time for reduced features is match better than the full features dataset.

Table.2 shows the result of the effectiveness, classification accuracy, classification error, training time, and testing time of FMM using different lamda values and full and reduced features of the NSL-KDD dataset.

NSL-KDD dataset with Full features				
Θ	***Training Time Sec***	***Testing Time Sec***	***Accuracy %***	***Classification Error %***
0.0001	19	29	98.04	01.96
0.00045	17	28	93.09	06.91
0.00078	15	25	91.03	08.97
NSL-KDD dataset with 12 features				
Θ	***Training Time Sec***	***Testing Time Sec***	***Accuracy %***	***Classification Error %***
0.0001	9	14	99.05	00.95
0.00015	8	12	96.09	03.91
0.00030	7	12	92.00	08.00

Figure 5 and 6 shows the comparison result of NSL KDD dataset, in terms of accuracy and training time against different values of lamda. The performance in term of accuracy for both full features and reduced features dataset is perfect and almost same but in term of training time, the reduced features dataset is match better than the full features dataset.

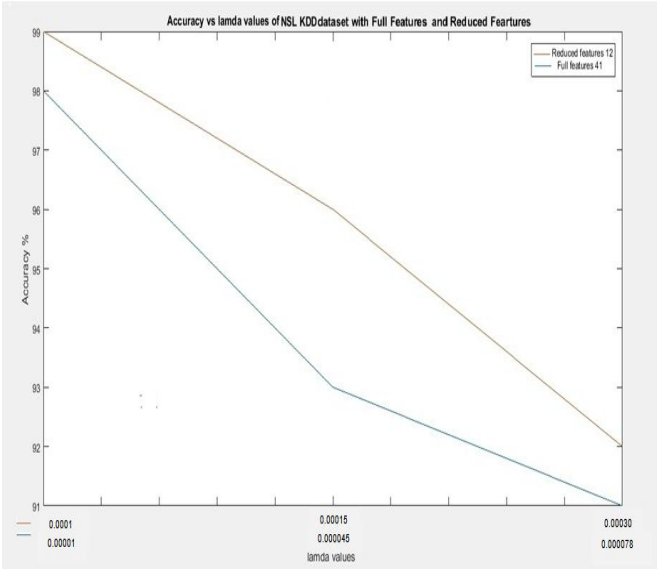

Fig. 5. Accuracy Vs Lamda Values of NSL KDD dataset with Full Features and Reduced Features

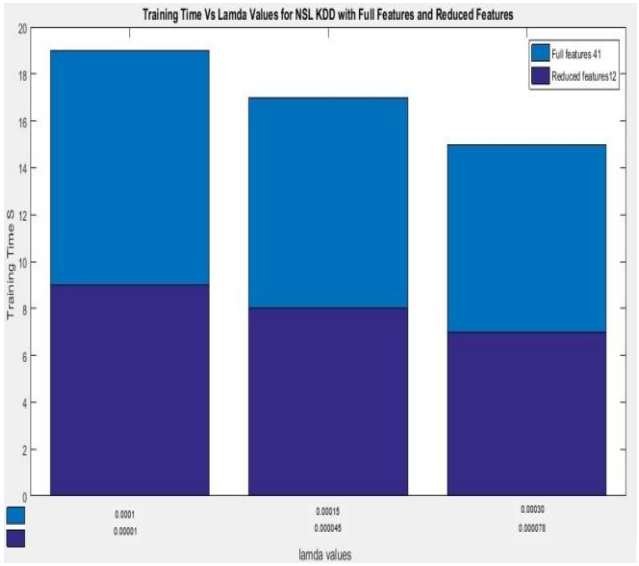

Fig. 6. Training Time Vs Lamda Values of KDD dataset with Full Features and Reduced Features

Table.3 shows the comparison result of the proposed system with the other existing systems published results. The FMM-NN based system is compared with the multilayer perceptron classifiers, Naïve Bayes classifier, SMO classifier, logistic regression classifier, LibSVM classifier. The result of existing systems is evaluated on WEKA with default values.

Our proposed system demonstrates the ability of fuzzy min-max classification neural network to find reasonable decision boundaries in overlapping classes, learn highly nonlinear decision boundaries, and provide best results that were equivalent to other neural and traditional classifiers.

Table.3 shows the comparison result of the proposed system with the other existing systems and published results.

Classifier	*Accuracy %*	*Misclassification %*
Multilayer Perceptron MLP	**93.59**	**06.41**
Naïve Bayes	**77.28**	**22.72**
SMO	**91.63**	**08.37**
Logistic Regression	**91.56**	**08.44**
LibSVM	**91.48**	**09.52**
FMM GA (13)	**95.93**	**04.07**
FMM NN (16)	**95.17**	**04.83**
KDD cup winner (20)	**91.80**	**08.20**
KDD cup runner up (21)	**91.50**	**08.50**
FMMNN IDS KDD dataset (proposed)	**99.96**	**00.04**
FMMNN IDS NSL-KDD dataset (proposed)	**99.05**	**00.95**

VII. CONCLUSION

Our paper demonstrates the ability of fuzzy min-max classification neural network to find reasonable decision boundaries in overlapping classes, learn highly nonlinear decision boundaries, and provide best results that were equivalent to other neural and traditional classifiers. The proposed system required less timing for the learning and predicting compared to the tradition neural network like backpropagation neural network, and Boltzmann neural network. The system is trained, tested, and evaluated using two datasets (KDD Cup 99 and NSL-KDD). The total number of instances included in training and the testing dataset is (126887, 11850) instances. 70% of instances are used for training and the remaining of instances are used as a testing dataset for FMM NN IDS. Those datasets first are preprocessed through a number of steps. Started by converting nominal attributes to numeric using the grp2idx function. follow by removing the duplicate entries and extreme values using filtering based on interquartile ranges concept of weka. lastly is featured extractor using PCA algorithm. Due to PCA, the features are reduced in both datasets into (9 and 12) features. The performance in term of accuracy for both datasets is good but in term of training time and testing time both datasets with reduced features is match better than the full features dataset.The proposed system provides best results comparing to other existing classifiers and published results.

VIII. REFERENCES

[1] Ahmed Fuad Mohammed, Vikas T. Humbe, Santosh S. Chowhan, "*Analytical Study of Intruder Detection System in Big Data Environment* " Soft Computing: Theories and Applications SoCTA 2016, Volume 2, 2016.

[2] Khattab M. Ali Alheeti, Anna Gruebler, Klaus D. McDonald-Maier, "*An Intrusion Detection System against Black Hole Attacks on the Communication Network of Self-Driving Cars*", Emerging Security Technologies (EST), 2015 Sixth International Conference on, Publisher: IEEE, 10 March 2016.

[3] JABEZ Ja , Dr.B.MUTHUKUMAR, *"Intrusion Detection System (IDS): Anomaly Detection using Outlier Detection Approach "*, International Conference on Intelligent Computing, Communication & Convergence (ICCC-2014), Procedia Computer Science 338 – 346 Publisher: Science Direct, 2015.

[4] Modi, Chirag N., Dhiren R. Patel, Avi Patel, and Rajarajan Muttukrishnan. *"Bayesian Classifier and Snort based network intrusion detection system in cloud computing"*, 2012 Third International Conference on Computing Communication and Networking Technologies (ICCCNT 12), 2012.

[5] Ahmed Fuad Mohammed, Vikas T. Humbe, Santosh S. Chowhan, *"Data Mining Based Traffic Classification Using Low-Level Features"* International Journal of Computer Applications (0975 – 8887), 2017.
[6] Azad C, Jha VK , *"Data mining based hybrid intrusion detection system"*, Indian J Sci Technol 7(6):781–789, 2014.
[7] Chen T, Zhang X, Jin S, Kim O , *"Efficient classification using parallel and scalable compressed model and its application on intrusion detection"*, Expert Syst Appl 41(13):5972–5983, 2014.
[8] Gu B, Guo H, *"The intrusion detection system based on a novel association rule*", In: International conference on information science, electronics and electrical engineering (ISEEE), vol. 2, pp 1313–1316, 2014.
[9] Anas Quteishat, Chee Peng Lim. *" A modified fuzzy min–max neural network with rule extraction and its application to fault detection and classification"*, Elsevier, Applied Soft Computing 8 (2008) 985–995.
[10] Liao S-H, Chu P-H, Hsiao P-Y, "*Data mining techniques and applications–a decade review from 2000 to 2011"*, Expert Syst Appl 39(12):11303–11311, 2012.
[11] Julisch K, *"Data mining for intrusion detection. Applications of data mining in computer security*". Springer, US, Fuzzy Min-Max Neural Network-Based Intrusion Detection System pp 33–62, 2002.
[12] Wang G, Hao J, Ma J, Huang L *"A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering.Expert Syst"* Appl 37(9):6225–6232.
[13] Chandrashekhar Azad, Vijay Kumar Jha, *"Fuzzy min–max neural network and particle swarm optimization based intrusion detection system"*, Springer-Verlag Berlin Heidelberg 2016.
[14] . Lei JZ, Ghorbani AA , *"Improved competitive learning neural networks for network intrusion and fraud detection"*, Neurocomputing 75(1)135–145.
[15] Basant Subba , Santosh Biswas, Sushanta Karmakar, *"A Neural Network Based System for Intrusion Detection and Attack Classification"*, 978-1-5090-2361-5/16/$31.00 c 2016 IEEE, 08 September 2016.
[16] Azad Chandrashekhar, *"Fuzzy Min-Max Neural Network-Based Intrusion Detection System"*, © Springer Nature Singapore Pte Ltd. PP 191-201, 2017.
[17] L. A. Zadeh, "*Fuzzy Sets"*, Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley, California.
[18] Simpson PK, *"Fuzzy min-max neural networks. I. classification*", IEEE Trans Neural Networks 3(5), 776–786, 1992.
[19] Anjay Krishnankutty Alonso, *eMath teacher for MAMBANI'S FUZZY INFERENCE MMETHOD*, Retrievedfrom,http://www.dma.fi.upm.es/recursos/aplicaciones/logica_borrosa/web/fuzzy_inferencia/funpert_en.htm, (2017, Oct, 19)
[20] Pfahringer B Winning, *" the KDD99 classification cup: bagged boosting"*, ACM SIGKDD Explor Newsl 1(2):65–66, 2000.
[21] Levin I, KDD-99, "*classifier learning contest: LLSoft's results overview*", SIGKDD Explor 1(2):67–75, 2000.

Authors Profile

Mr. Ahmed Fuad Mohammed Al-Dubai, Research Scholar At School of Computational Sciences, S.R.T.M University Nanded, Maharashtra.

Dr. Vikas T. Humbe Assistant Professor At School of Technology, S.R.T.M University Sub-Campus Latur, Maharashtra.

Dr. Santosh S. Chowhan Assistant Professor At School of Computational Sciences, S.R.T.M University Nanded, Maharashtra.

Dr. Yaser Fuad Mohammed Al-Dubai, Assistant Professor At 4Faculty of Administrative Sciences, Albaydaa University, Radaa City, Yemen Country.