ISSN 2278-3091

Volume 8, No.1.4, 2019

International Journal of Advanced Trends in Computer Science and Engineering

Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse0281.42019.pdf

https://doi.org/10.30534/ijatcse/2019/0281.42019

An Improved Novel ANN Model for Detection Of DDoS Attacks On Networks

Bilal Hikmat Rasheed[1], M.Sivaram[2], D.Yuvaraj[3] A. Mohamed Uvaze Ahamed[0]

[0]Department of Computer Science, Cihan University - Erbil, Kurdistan Region, Iraq
Email:mohamed.sha33@gmail.com
[1]Department of computer science, Cihan university - Duhok , Kurdistan Region- Iraq
Bilal.h@duhokcihan.edu.krd. Email:bilal.rasheed11@gmail.com
[2]Department of Computer networking, Lebanese French University – Erbil, Kurdistan Region, Iraq
Email: phdsiva@gmail.com,sivaram.murugan@lfu.edu.krd
[3]Department of computer science, Cihan university - Duhok , Kurdistan Region- Iraq
Email: yuva.r.d@gmail.com, yuvaraj.d@duhokcihan.edu.krd

ABSTRACT

Attacks over the internet have become an increasing menace in recent time which tries to hack or illegally tamper with the data available over the networks. On the other hand, there has been an increase in volume in research contributions to effectively counter attack these attacks and implement a strong defence mechanism. There have been numerous algorithms and frameworks implemented in recent times which are intelligent and soft computing based. These evolution based algorithms play a vital role in self adapting the system under attack towards increasing and new types of attacks which are increasing day by day. One such area of soft computing algorithms investigated in this chapter is the artificial neural network or popularly known as ANNs. They work analogous to the biological neurons in the human body. The chapter is organized in a systematic manner to give an insight in to ANN based network models to counter attack DDoS attacks which has been the primary focus of this thesis, architecture and implementation of ANNs, the experimental investigations and findings which help in drawing an inference of ANN based defence models.

Key words: Network attacks, distributed denial of service attacks, ANN, training and confusion matrix.

1. INTRODUCTION

DDoS attacks are a special class of attacks prevalent in online communication among networks utilizing internet services for storage, processing and utility [2]. These attacks increase the congestion of the network by introducing zombie packets which are infected packets which contaminate good packets as they progress down the communication layers. In a DDoS attack, the attacker or hacker sends an array of infected packets which cause flooding [10] due to which the target system gets occupied to service the flooded requests which cause a severe degradation in the network bandwidth and increases the computation system overhead. Although almost all attacks could be detected by existing techniques, there are some techniques which cannot be detected and are known as zero attacks.

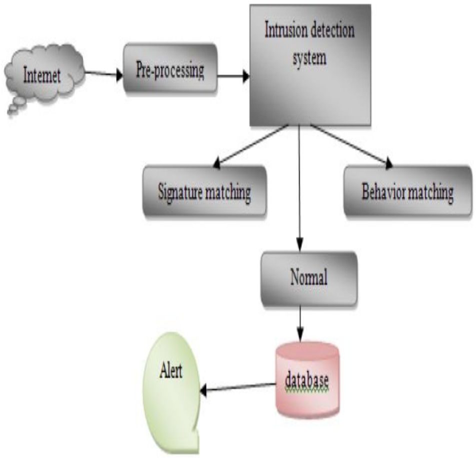


Figure 1: General scheme of intrusion detection mechanism

A simple illustration of an intrusion detection scheme is depicted in figure 1 where the packets of information are pre-processed and the conditioned input is given to the IDS system. The signature and behaviour patterns are extracted from the packets regarding abnormalities in terms of file size, frequency of repetition, bandwidth etc., Based on the features extracted and the adaptation of IDS system, the given packet of information is classified as normal or an infected packet. The essential motivation for going in for ANN based DDoS attack detection in this thesis is that they are capable of detecting these zero attacks using a specific pattern. These patterns distinguish the normal attacks from these DDoS attacks. These features could be given as vector set to train the neural networks to improve the detection accuracy. Before the actual implementation, it is quite necessary to have knowledge about the essential components in the intrusion detection framework which define the efficiency of the implemented system. These components or parameters are elaborated in the next section.

2. RELATED WORK

Rule based methods have been found in the literature [4], [8-9] for detection of DDoS attack detection wireless sensor networks and the process occurs in three phases. In the first phase, monitoring of information to filter the vital data from the mainstream network is carried out with the help of monitor nodes. The second phase is known as rule application stage where predefined rules are applied to the information that has been filtered in the first stage by the monitor nodes. In the last phase, a detection alarm is raised if any of the information packets fail the rule application test. A slight variant [12] of this technique is proposed by extracting the behaviour patterns of the neighbouring node in the senor network.

A cumulative sum algorithm [15] has been reported in the literature which continuously monitors the incoming and outgoing packets of information for any behavioural and pattern changes. A similar three phase scheme has been reported in the literature [7] where the nodes are first monitored and detected whether they originate from legitimate or illegitimate nodes. If they originate from legitimate nodes, the normal tasks of the system are carried out. Information packets originating from illegitimate nodes are extracted for their features and the rule base is applied on this information. The set of predefined rules as mentioned above are framed by observing the network protocol patterns in a normal communication network. Any deviation from the normal pattern classifies the incoming packet as illegitimate and filtered off in succeeding processes. Another variant of the rule based method is applicable for effectively detection DDoS attacks in IoT networks [4] based on an event processing model. The experimentations have been done using SQL as reference and the rule codes are stored in the repository. The experimental results indicate least utilization memory but suffers a drawback of utilized more system resources at the cost of minimal processing time.

There has been a migration of detection schemes towards cluster computing and soft computing due to the nature of intelligence and capability to handle huge volumes of incoming as well as outgoing data and at the same time to produce a quick response time. A wide range of soft computing algorithms [14] like principal component analysis [6], linear discriminant analysis [12] [15], local binary pattern [10], particle swarm optimization and greedy search algorithms have been found in the literature.

Utilization of support vector machine [9] based methods have also been found in the literature which is useful for classification of malicious packets of information. Support vector machine strategies have been imposed on mobile agent models for detection of known attacks. Two other mobile agents are identified in the literature namely collector agent who gives feedback from the wireless sensor network and the other agent is known as misuse detection agent who detects known malicious patterns in the network. Support vector machines have been integrated with Gaussian kernel and experimented with three type of data sets with a 98.7% accuracy being reported in the literature.

PCA based techniques have been reported [11] to be bring about reduction in data dimensionality especially when dealing with huge volume of traffic with a wide range of feature based attacks incident on the mainstream of the network. PCA based techniques have been found to exhibit not only dimension reduction feature but also exhibit a high degree of classification attributed to their ability to differentiate the malicious packets from normal data packets using multiple attribute values.

3. PROPOSED WORK

3.1 Artificial Neural Networks

Artificial neural networks are essential networks in almost any data processing and computing applications and are function analogous by the neurons in our central nervous system. Similar to biological connectivity of neurons, ANNs are also featured with high interconnected elements in perfect coordination to meet an objective function. They serve several applications which include pattern recognition and classification, detection problems, adaptation and control applications etc., A simple scheme of artificial neural network is depicted in figure 2

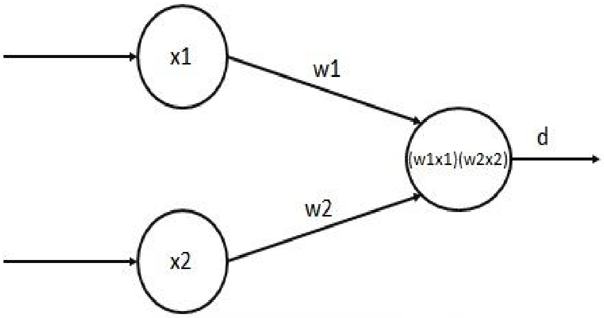


Figure 2: Illustration of a simple neural network model

The above figure illustrates a two input neural network scheme whose operation is quite straightforward in approach. The two inputs x_1 and x_2 along with an associated weight function are given as inputs to the function block denoted by the bigger circle where a simple product between these two quantities is performed to obtain the desired output. In a feedback system, the neural networks are able to adjust the weights w_1 and w_2 in successive iterations based on the error signal generated from the difference of obtained and desired outputs. The overall objective of the above network shown above lies in minimizing the error in the least time possible.

The above single layer neural network could be extended further by interconnection of many other multiple nodes with the objective function of transforming inputs to desired outputs. A most widely used configuration for intrusion detection systems are the multilayer perceptron models (MLP) as multiple inputs of differing pattern are incident on the target system under attack.

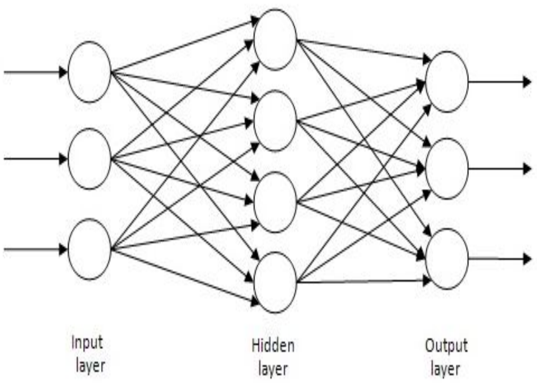


Figure 3: Architecture of Multi-layer perceptron model

The above architecture depicts a three layer ANN namely input, hidden and output layer. The number of nodes in the input layer correspond to feature vector in the given problem definition. The number of nodes in the output layer corresponds to the number of sets or classes to which the desired throughput may be designated. For example, the output nodes may be two to three for a brain tumour detection and classification problem where the output needs to known as either the brain tumour being malignant or benign. In the given problem of attack detection, the number of output layer may be limited to two comprising of infected or valid packet of information.

The input and output layers are connected to one another by methods for the shrouded layer and the procedure of update starts with some random loads allocated to every node. On completion of the first iteration, the weights are updated according to a weight update equation defined in (1) to minimize the error at the output as defined in (2).

$$WI(x,y) + \alpha \qquad (1)$$

and

$$E\{e2[n]\} = E\{(d[n] - y[n])^2\} \qquad (2)$$

where E{ e^2 [n] } denotes the expectation of mean squared error function, d[n] indicates the desired output and y[n] denotes the obtained output.

This iteration process follows a learning algorithm which may follow a propagation rule mechanism.

3.2 Neural Network Training

The process of making the implemented neural network learn the feature vector patterns and thereby decide upon categorizing the given inputs into a class of designated outputs is known as training and forms the backbone of neural network efficiency and working. The overall objective of neural network learning follows a minimization rule to reduce the loss function as depicted in figure 4.

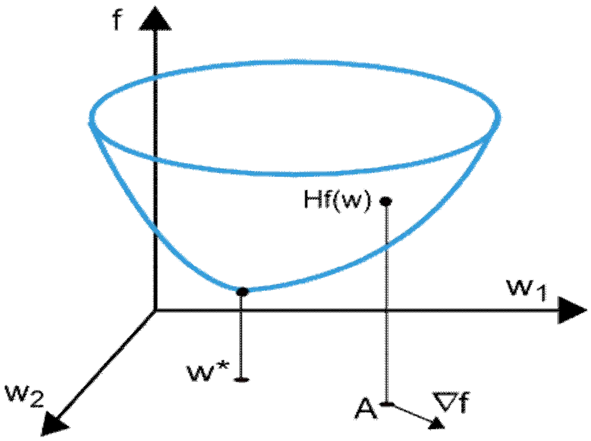


Figure 4: Illustration of overall loss function for given problem

From the above figure it could be clearly seen that the point w* denotes the minimal value of the given loss function and solution to the given problem objective could be achieved by computing the first and second derivatives of this loss function as shown in equations (3 &4)

$$\nabla_i f(w) = df/dw_i \ (i = 1,\dots,n) \qquad (3)$$

The second derivative could be generalized using the Hessian matrix as

$$H_{i,j} f(w) = d^2f/dw_i \cdot dw_j \ (i,j = 1,\dots,n) \qquad (4)$$

These mathematical formulations could be mapped onto one dimensional search spaces to obtain the solution to the given minimization problem as shown in figure 5.

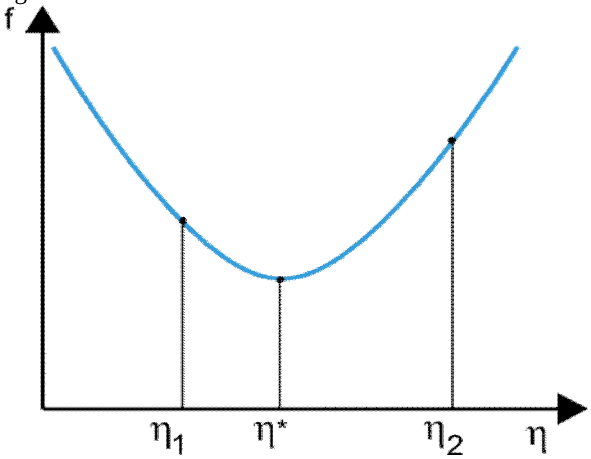


Figure 5: One dimensional mapping of minimum loss function

In the above figure the minimal function of loss is present in between the point's η1and η2. Golden section and Brent's method is popularly used algorithm for one dimensional loss minimization function. However, most of the real time problem definitions in real time require a multi-dimensional search and minimization strategy including the problem objective of this thesis. The next section deals with multi-dimensional optimization techniques and methodologies in detail which greatly aid in implementing the proposed architecture for defence against attacks.

(ii) Levenberg – Marquardt Method

It is based on sum of squares iteration and is also alternately known as damped least squares method. This algorithm also does not have any interference with Hessian matrix and its inverse values but rather works on another matrix known as Jacobian Matrix. The Jacobian matrix is defined using a loss function of the form

$$f = \sum {e_k}^2, k = 0, 1, 2, \ldots.. l \quad (5)$$

$$J_{i,j} f(q) = \frac{de_i}{dq_j} \quad (6)$$

The weight update equation is given as

$$W_{i+1} = W_i - (J^T J + \delta I)^{-1} (2J^T \cdot e_i), \ i = 0,1,\ldots \quad (7)$$

In the above equation, δ denotes the damping function and it becomes a conventional Newton's method when it becomes a zero. I is the identity matrix and the term $J^T J + \delta I$ denotes the approximation of the Hessian matrix. The damping factor value is decreased or increased to get to the convergence value and acceleration towards minima is very fast in this method making it suitable for training of medium size neural models. The problem arises for large sized networks when the size of the Jacobian matrix gets doubled affecting cost and complexity.

(iii) Proposed ANN model

based on the examinations and discoveries identified with ANN usage models and their benefits, a three layer ANN feed forward model has been proposed and executed in this paper with modification of principal component analysis (PCA) for dimensionality reduction. The reason behind providing a dimensionality reduction is to reduce the storage and computation complexity of a LM learning rule when deriving Jacobian matrix for large NN models as used in the proposed thesis. Figure 11 depicts the ANN model utilized in this implementation in which the decreased measurement input features are given to the ANN model before training.

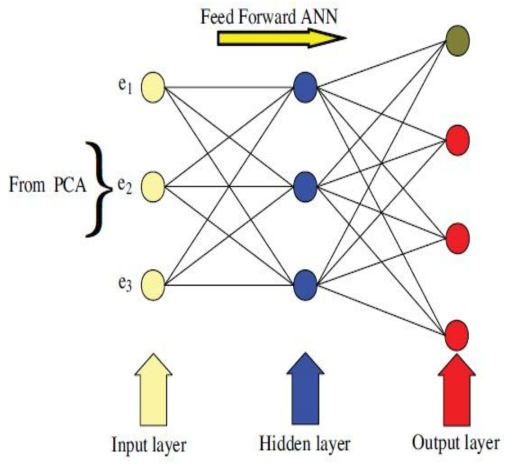


Figure 6: Feed forward ANN architecture for DDoS.

Applying PCA to decrease the dimension of the high dimensional data incorporates the accompanying stages outlined below.

S1: make a framework M utilizing the data set.

S2: Normalize the data set utilizing -value.

S3: Calculate the deterioration estimation of the data matrix.

S4: Calculate the variance utilizing the diagonal elements

S5: Sort variance in the diminishing method.

S6: Choose the principal components from the biggest variance.

S7: Form the transformation matrix comprising of those qualities

S8: Find the decreased anticipated data set in another arrange pivot by applying to

.The first step as indicated in the algorithm above involves computation of the initial centroid value for the input data set with dimensionality $M.$ subsequently, the variance is computed for each data in dimension. The column with maximum variance is M_{max} and is sorted in the descending order of magnitudeIn the complete k subset groups, every median is instated in the cluster centers. In straight measurement decrease procedure PCA is the best, the mean-square error is a dependent on the covariance matrix of the variables, it is a second-request technique. The proposed algorithm is given as a pseudo code which accepts the features of the DDoS attacks incident at various time instants on the network or system under attack and the output of trained NN is twofold depicting a normal and infected packet to prevent further infection of true nodes.

The input has x features and m number of instances sampled and τ is the adjustable relevance threshold

```
Initialize the matrix
For $i = 1\ ton$
$begin$
$randomlyselectaninstanceI$
$findthenearesthitHandmissJ$
        $forj = 1\ toN$
$s(i) = S(j) - \frac{|j, I, H^2|}{N} + \frac{diff(j, H^2)}{N}$
$end$
Initialize the matrix $M = \{\}, fori = 1\ toN$
$for\ each\ subset\ of\ S\ with\ size\ j$
$ifdvar(s, m) = 0$
$returnM$
```

$the\ minimum\ subset\ satisfies\ the\ \delta$

The dimensionality dimension algorithm provides the data output about the dimension for the input and for a given sample n = [n1, n2, ...n] $\in \varphi_n$ where n is sample size

```
For $u = (n,k)$
{
//Initialization
        Generation of class label encoding $L_1 L_2, L_3 \dots \dots, L_c$
Initialization $u(1) = (n,k)$
// C is Classification
        Set the maximum number of iterations $MaxIteration$
        Set precision $\vartheta$, set counter $C = 1$
//Dimensionality reduction
        While
$(r < MaxItera\&\&|u(r') - u(r-1)^2| < \varphi_n)$
Do
{
        r= r+1
Obtain an approximate solution with gradient iteration method V(r)
}
//print result        $u = (n,k)$ }
Import=(l1, l2,……ln)T,ω=ΣlixiyiT
//projection matrix $M'$

$i = 1$ {

Calculate the symmetric positive semi definite matrix M,
elements $Aij = |yiTyj|xiTxj_n$, for
$i = 1,2, \dots \dots, n;$ use the PBB method to solve the
optimization problem $\frac{minlTAl}{2} - lTl,$ Constraint
conditions 0$<= l < \eta l$ ,

$getl = \{l1, l2, \dots \dots ln, \}$T  Given
$l(1) = \{l1(1), l2(1), \dots \dots ln(1)\}T \in Rn, \lambda 1 > 0$
If l(1),   l(1) replace$\rho$ (l(1))Calculate the projection vector
$gk = Al(k) - 1$
,$If\ |P(l(k) - gk)^2 - l(k)\ |2 < \tau$

Stop the cycle and jump to the final output statement

Calculate $l(k+1) = P(l(k) - \lambda k \cap gk)$

$sk = l(k+1) - l(k), \lambda k = \sum_{gk}^{T} skt / (skT(gk + 1T - gk))$

//long of the step

$k = k + 1$,

The following is the final output statements

Import : $l = (l1, l2, \dots \dots ln)T, M' = \Sigma lixiyiT$

}
```

4. EXPERIMENTAL ANALYSIS AND FINDINGS

The proposed work flow is depicted in figure 12 in which the input parameters from the packets of data with respect to their size, transfer speed inhabitancy, personal conduct standard are arranged in the database.

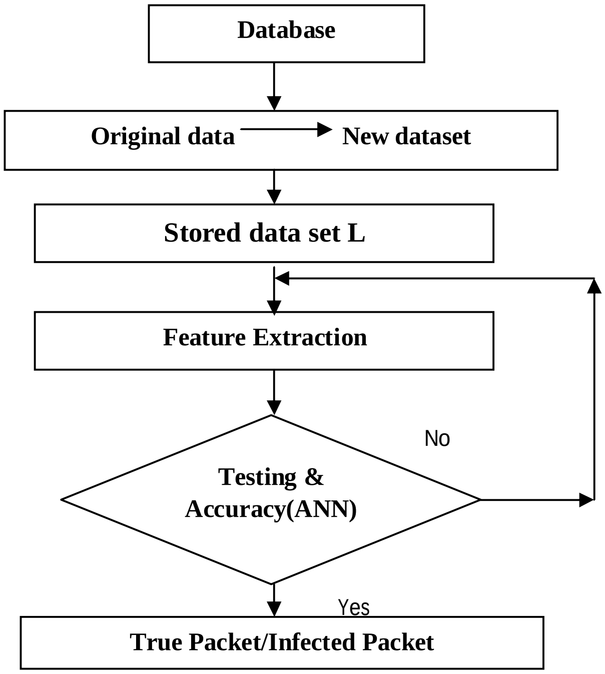
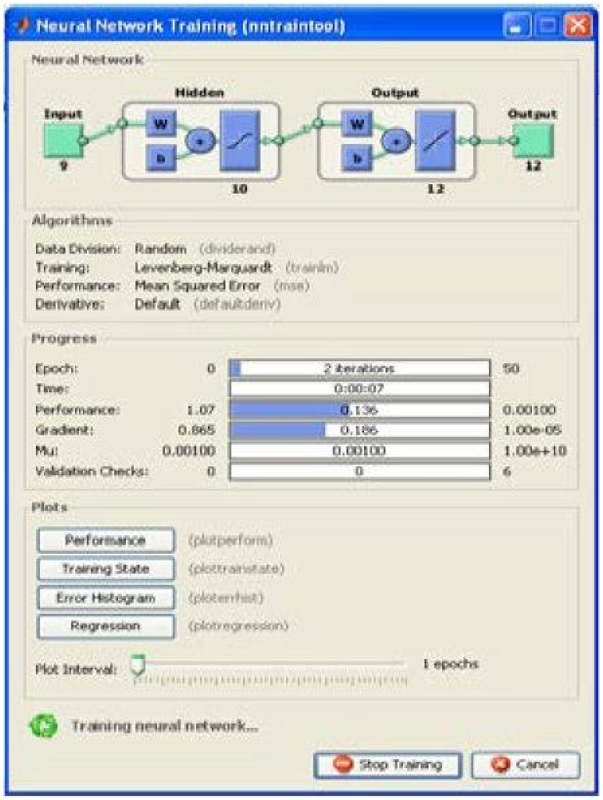


Figure 7: Flow process of proposed DR- ANN

The given list of capabilities is changed into the neural database and prepared to utilize Levenberg-Marquardt training algorithm. The proposed framework is tested on a Celeron processor 1.85 GHz with 2GB RAM running Windows XP and coded by Matlab 6.5. The KDD Cup99 dataset is gotten from the DARPA98 arrange traffic dataset by gathering singular TCP packets into TCP associations has been utilized for benchmarking the proposed ANN model. very TCP connection has 60 highlights with a mark which indicates the status of an association as either being typical or a particular assault type.

In the above flow process, the abbreviation denotes dimension reduced ANN model which is the proposed model of ANN for DDoS attack detection in this thesis. The proposed network is fit for taking care of three classes of attacks namely the DDoS, DoS and Probe which are found from the experimental investigations. The snapshot of the 'nntool' utilized in the execution utilizing LM learning rule and MSE criteria for blunder intermingling is portrayed in figure 8.


Neural Network Training (nntraintool)
Neural Network
Hidden
Output
Input
9
W
b
10
W
b
12
Output
12

Algorithms

Data Division: Random (dividerand)
Training: Levenberg-Marquardt (trainlm)
Performance: Mean Squared Error (mse)
Derivative: Default (defaultderiv)

Progress

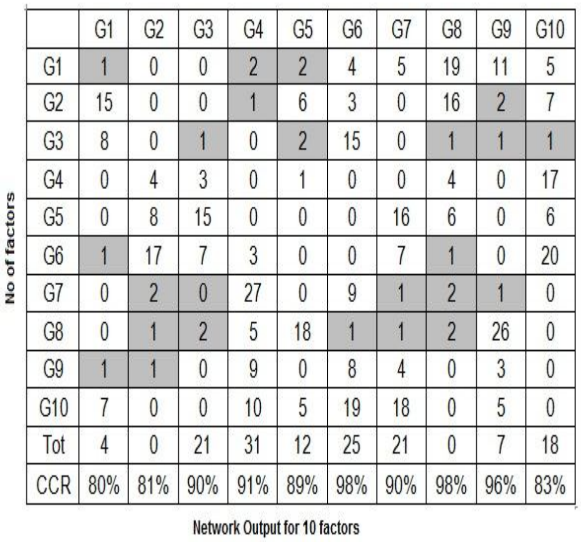
Epoch:	0	2 iterations	50
Time:		0:00:07	
Performance:	1.07	0.136	0.00100
Gradient:	0.865	0.186	1.00e-05
Mu:	0.00100	0.00100	1.00e+10
Validation Checks:	0	0	6

Plots
Performance (plotperform)
Training State (plottrainstate)
Error Histogram (ploterrhist)
Regression (plotregression)
Plot Interval: 1 epochs
Training neural network...
Stop Training
Cancel

Figure 8: Snapshot of nntrain tool used in the proposed work using LM

The network is prepared for various estimations of epochs and error objective, where epochs and error goals are training parameter. Ordinarily one epoch of preparing is characterized as a solitary introduction of all input vectors to the network. The network is then refreshed by the consequences of every one of those introductions. Preparing happens until a most extreme number of epochs happen, the presentation objective is met, or some other training function of the preparation work happens.

After executing the neural network we showed signs of improvement location exactness as 55.86 and the number of feature set matrix is depicted in figure 14.



	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
G1	1	0	0	2	2	4	5	19	11	5
G2	15	0	0	1	6	3	0	16	2	7
G3	8	0	1	0	2	15	0	1	1	1
G4	0	4	3	0	1	0	0	4	0	17
G5	0	8	15	0	0	0	16	6	0	6
G6	1	17	7	3	0	0	7	1	0	20
G7	0	2	0	27	0	9	1	2	1	0
G8	0	1	2	5	18	1	1	2	26	0
G9	1	1	0	9	0	8	4	0	3	0
G10	7	0	0	10	5	19	18	0	5	0
Tot	4	0	21	31	12	25	21	0	7	18
CCR	80%	81%	90%	91%	89%	98%	90%	98%	96%	83%

No of factors
Network Output for 10 factors

Figure 9: Feature vector matrix for the proposed network

The error convergence for the proposed work is portrayed in figure 15 and it could be seen that the proposed technique can accomplish a quicker convergence when contrasted with existing ordinary ANN models. Moreover, the data dimension has been greatly reduced due to integration PCA with the ANN optimization.

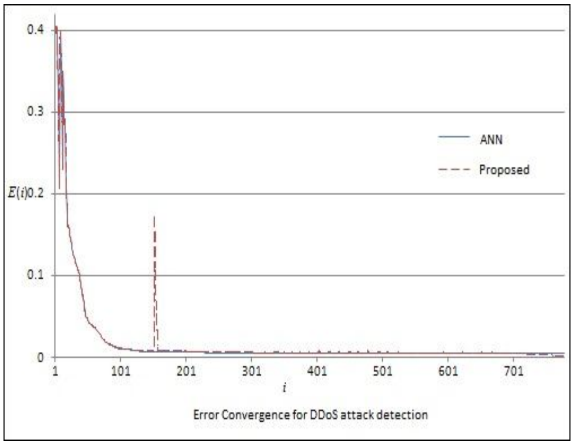
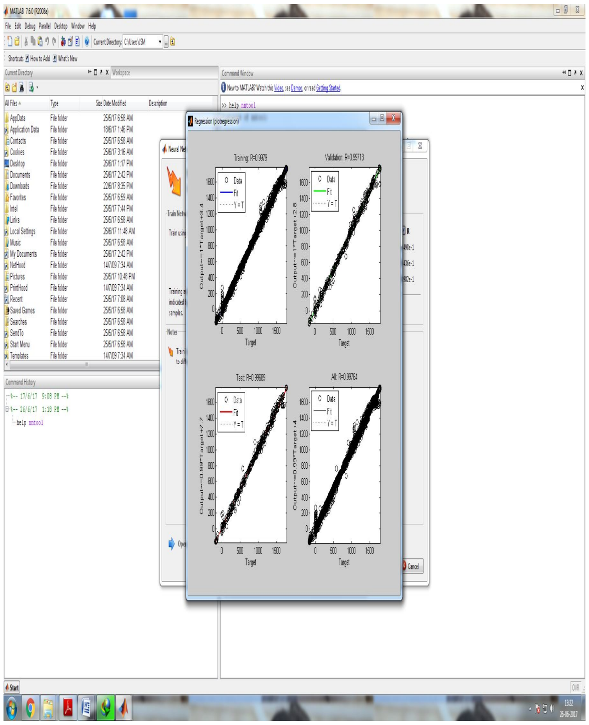


Figure 10: Error Convergence plot of proposed model

Figure 16 depicts the regression plot for the proposed ANN model.



New to MATLAB? Watch this Video, see Demos, or read Getting Started.

```
>> help nntool
```

Figure 11: Regression Plot for the proposed network

The relapse plot has been gotten for test, approval and in general periods of the ANN execution. The ideal mistake estimation of 0.9975E-05 is acquired which gives a decent rate of classification. An analysis of the receiver output characteristics has been tabulated in table 1

Table 1: Receiver operating characteristics

Parameters	ANN	Proposed ANN
Transmission throughput	790kbps	646kbps
Reciever throughput	812kbps	770kbps
Packet delivery ratio	0.74	0.51
Elapsed Time	0.92s	0.71s

An analysis of the predictions of the proposed ANN which are correct/incorrect are given below in figure 12.

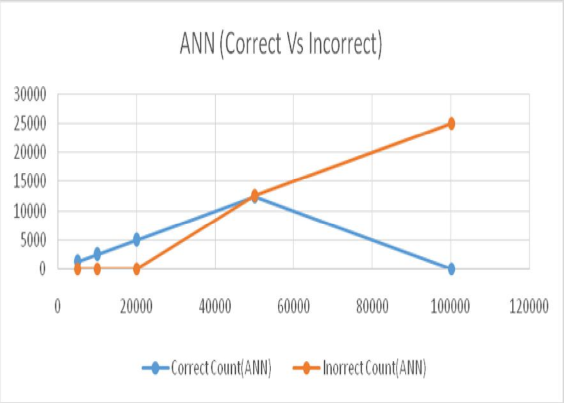


Figure 12: Prediction performance of proposed ANN

5. CONCLUSION

This paper has effectively dealt with implementation of an improved ANN model with data dimension reduction feature for implementing an attack detection system especially DDoS types of attacks which prove to be a menace in recent times. The proposed model has been actualized utilizing requirements and mathematical formulations widely managed with and elaborated in the earlier sections of this chapter and tested against a wide range of network attacks. Experimental perceptions have been classified and visually spoken to and the discoveries demonstrate a predominant exhibition when contrasted and existing ANN systems for DDoS attack. The proposed strategy is prepared by LM algorithm which observes the irregular patterns of the incoming patterns and assigns them into infected packets which are isolated from creating additional infections as they advance along with the network. When these tainted agents are cleared, the network bandwidth capacity is cleared and the original speed of web utilization is reestablished to the client, therefore, meeting the problem objective.

REFERENCES

1. Ali E.Taki, El_Deen, El-Sayed, A.El-Badawy and Sameh N.Gobarn **"Digital Image Encryption Based on RSA Algorithm",** *International Journal of Electronics and Communication Engineering*, Vol.9, issue-1, pp. 69-73.2013.
2. Arumugam N and C. Venkatesh (2013), **"Ant system algorithm based IP traceback method to detect denial of service attack on data network",** *Australian Journal of Basic and Applied Sciences*, Vol. 7, 2013, pp. 732-741,2013.
3. Audrey A. Gendreau, Michael Moorman (2016), "**Survey of intrusion detection systems towards an end to end secure internet of things**", *Proceedings of international conference on future internet of things and cloud,2016*

4. Bass T ,"**Intrusion detection systems and multi-sensor data fusion**", *Communications of the*

ACM, Vol. 43, No. 4, pp. 99-105,2008.

5. Batalla J M, P. Krawiec **“Conception of ID layer performance at the network level for Internet of Things”**, *Springer: Personal and Ubiquitous Computing*, Vol. 18, pp. 465-480,2014.

6. Bharathi and Sukanesh (2012), “**A PCA framework for detection of application layer DDoS attacks”,** *Transactions on information science and applications*, Vol. 12, No. 9, pp. 389 – 398,2012.

7. Chen Y and K. **Hwang “Collaborative detection and filtering of shrew DDoS attacks using spectral analysis”,** *Journal of Parallel and Distributed Computing*, Vol. 66, no. 9, pp. 1137-1151,2006.

8. Chen, Min, Wan, Jiafu, Li, Fang **“Machine-to-machine communications: Architecture, standards and applications**”, *KSII Transactions on Internet & Information Systems,* Vol. 6, No. 2, pp.480 – 497,2012

9. Choi J, C. Choi, B. Ko, D. Choi, and P. Kim, “**Detecting web based DDoS attack using Map Reduce operations in cloud computing environment**”, *Journal of Internet Service Information Security*, Vol. 3, No. 3/4, pp. 28–37,2014.

10. Dewan Md. Farid, Mohammad Zahidur Rahman **“Anomaly Network Intrusion Detection Based on Improved Self Adaptive Bayesian Algorithm”**, *Journal of Computers,* Vol. 5, No. 1, January 2010.

11. Jerbi, M. **“Towards efficient geographic routing in urban vehicular networks”,** *Vehicular Technology, IEEE Transactions on*, Vol. 58, No. 9, pp. 5048–5059,2009.

12. Joo P. Amaral, Lus M. Oliveira, Joel J. P. C. Rodrigues, Guangjie Han, Lei Shu **, “Policy and Network-based Intrusion Detection System for IPv6-enabled Wireless Sensor Networks”**, *IEEE Conference on Communications Software*, Services and Multimedia Applications,2014.

13. Juneja D and N. Arora **“An ant based frameworks for preventing DDoS attack in wireless sensor networks***”*, *International Journal of Advancements in technology,* Vol. 1, 2010, pp. 34-44.

14. Latifur Khan, Mamoun Awad, Bhavani Thuraisingham **“A new intrusion detection system using support vector machines and hierarchical clustering”, Journal of VLDB Journal**, Vol.16, pp.507-521,2007.