# Diagnosis of Distributed Denial of Service Attacks using the Combination Method of Fuzzy Neural Network and Evolutionary Algorithm

#### Saeid Mahmoudpour<sup>1\*</sup> and Seyed Javad Mirabedini<sup>2</sup>

<sup>1</sup>Department of Computer Engineering, Damavand Science and Research Branch, Islamic Azad University, Damavand, Iran; Sdmpour@gmail.com <sup>2</sup>Department of Computer Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran; jvd2205@yahoo.com

### Abstract

Availability, integrity and confidentiality are the key concepts of cyber security. Distributed denial of service attacks affecting the availability of information sources. This type of attack when be successful which led to the non-availability to the information sources. The success and impact of service denial attacks will be identified based on the victims and the risk level, threat and consequences of this attack, according to each case is different. The aim of this study is DDOS attack detection with differential evolutionary algorithm combination method and the fuzzy neural network ANFIS. For this purpose will pay to simulate and evaluate the proposed approach. In order to simulate in this research which is used MATLAB 2013b software which is a programming language and environment for scientific computing. The result of comparison showed that the ANFIS combination algorithm and differential evolution than to the ANFIS algorithm has more accuracy.

Keywords: ANFIS, Cyber Security, DDOS Attacks, Differential Evolution

## 1. Introduction

Distributed Denial of Service<sup>1</sup> attacks (DDoS) are the one of the most important attacks to avoid the uninterrupted performance of Internet service considered<sup>1</sup>. DDoS attack simply means pouring the more demands on a server (victim computer or target) and overuse of resources (CPU, database, bandwidth, memory and etc. as normal serving to the users due to the high volume of processing or the so-called overload operation of server, impaired or unavailable<sup>2</sup>. In the attacks with nature of service denial numbers of packages through (DoS) or more (DDoS) car to disable the computing power and network resources or enable target car set. DDoS attacks are more powerful and diagnosis and deal with them is more difficult from DoS attacks. Because in these attacks, several machines can be integrated in order to set the small flow of traffic into the target machine which manage of the total traffics for the target machine is difficult<sup>3</sup>. In an attack DDoS, a large number of compromised nodes attack to a similar target that will be led to denial of target system service. In fact, this kind of attacks founded by using botnet<sup>2</sup> of infected machines that unfortunately by increasing the number of hosts at risk, do not need to detect attack traffic from any source of DDoS attack as individually to create a powerful attack<sup>4,5</sup>.

Since that surely cannot be say any high volume application did not consider a DDoS attack, in this paper based on fuzzy neural network a method to diagnosis deal with these attacks will be detected. To this end, we plan to educate and to determine the optimal parameters use the differential evolution algorithm.

Differential evolution is a possible search methods based on the population in 1995 year by Setoren and Price

was invented. Evolutionary difference while similarities with other evolutionary algorithms but use of the distance information and direction from the current population to carry out search operations, it is distinct from other evolutionary algorithms. Adaptive neural fuzzy inference systems approach (fuzzy neural network ANFIS)<sup>6,7</sup> is a method that uses a multi-layer recursive network and neural network learning algorithms and fuzzy logic to design nonlinear mapping between the input space and output deals.

# 2. Research Empirical Background

The first investigation was carried out in order to detect DDoS attacks on the basis of some characteristics of specific disorders such as traffic sudden changes or Time To Live (TTL) was weird<sup>4</sup>. These methods usually have two problems: First, the diversion of traffic characteristics might see too small especially in monitors that are close to sources of attack.

Second, the accuracy and correctness of diagnosis limited to a variety of DDoS attacks because a property cannot cover all types of DDoS attacks. For example, the rate of packets SYN, to the packages RST cannot be used to detect SYN. To fix the problem, some of the works<sup>8,9</sup> the use of multiple traffic characteristics offered in the DDoS detection.

In 2013 year Raj Kumar and Selva Kumar in<sup>10</sup> used classifiers combination for detection of DDoS attacks. They proposed algorithm called NFBoost in terms of weight update distribution strategy, minimize the cost of errors and results combination method, the output of the has differ with other available methods.

Vysrz and others in<sup>11</sup> in response to a new type of vulnerability emerged called the application layer denial of service attacks that target Web services, attack productive tool to test and confirm the reported vulnerability introduced. Their results showed that, attacks devastating impact on the availability of Web services even when the absolute minimum of attack resources used. Because this type of attack is very simple to set up, it is clear that given the growth of cloud and web services, it is necessary to defend against them. This article smart system, fast and adaptive to layer attacks detect, the program has proposed XML and HTTP. The system they designed can be added in a cloud environment. Aspydvrvpvlvs and others<sup>12</sup>, the method of game theory which already in the field of network security in order to explore the interaction between attacker and defender was used during the scenario of distributed denial of service attack, (DDoS) used. In this paper, combined with previous work with this model create a richer set of available options to attackers. Writers multiple permutations in term of the cost to carry out an attack, the number of nodes being attacked, probability distributions of malicious traffic and their parameters discussed and they showed that a unique optimal strategy available to the defendant. By adopting this strategy, the upper border defense, chooses to deal with a real attack.

Zhou and his colleagues<sup>13</sup>, a new method suggested to detect DDoS attacks of application layer or AL-DDoS. Their work differentiated itself with previous methods according to attack detection AL-DDoS in backbone heavy traffic. Authors provided a modular architecture for defense that consists of manual end sensor, a detection module and traffic filter. Their proposed method to build a Realtime Frequency Vector (RFV) and traffic real-time specification as a set of models developed. This model can be used to diagnose AL-DDoS attacks.

## 3. Hypotheses

- Detecting DDoS attacks needs to adaptive learning and increase classifier.
- Detecting DDoS attacks needs to lower computational complexity and make smarter decisions from uncertain information.

# 4. Methodology

The aim of this study was to detect DDOS attacks with the combination method of differential evolution algorithm and the fuzzy neural network ANFIS. Here we will pay to the simulations and evaluate the proposed approach. In order to simulation was used MATLAB 2013b software.

## 4.1 Used Algorithms

# 4.1.1 Adaptive Neural-Fuzzy Inference System (ANFIS)

Adaptive Neural-Fuzzy Inference System method is a method that uses by a multi-layered recursive network and neural network learning algorithms and fuzzy logic to design nonlinear mapping between the input space and output deals. Figure 1, the structure of a simple ANFIS network with two inputs variable x and y shows. The network is composed by 5 layers that each inputs variable have two fuzzy subset. As Figure 1 shows proceed; and are subset of x and are subset of y.



**Figure 1.** The structure of adaptive network based on fuzzy inference systems<sup>12</sup>.

The structure of this network including nodes and orientation arc communicate the relationship between the nodes. As the name of network inferred and part of nodes or all of them have capable of adaptation. It means that the node output dependent on the parameters related with node. How and rate of change in the parameters, in order to reduce error defined by the rules of learning.

Fuzzy inference system of this network is type of Takagi-Sogono and its structure has 5-layer. The first hidden layer, input variables such as relative to the membership functions. The nodes of this layer are adaptive and output of this layer based on a formula (1) is calculated.

$$\boldsymbol{o}_{i}^{l} = \boldsymbol{\mu}_{Ai}(\mathbf{x}) \tag{1}$$

x input of node i, Ai is fuzzy subset of the input variable x,  $O_i^1$  output from the i-th node from the first layer and  $\mu_{Ai}$  (x) Membership function related to fuzzy subset of input variable x with value maximum of 1 and the value minimum of 0.

Fixed nodes in the second layer, the threshold rate for each rule calculated by multiplying the input values and consider as output. The function of each node in this layer, finding the weight (w) for the fuzzy rules by using membership functions are. To achieve this amount through formula (2) may be calculated.

$$w_i = \mu_{\rm Ai}(\mathbf{x}) \times \mu_{\rm Bi}(\mathbf{y}), i = 1, 2, ...$$
 (2)

Obtained weights from the second layer by nodes of third layer and through formula (3) are normal.

$$\overline{w_i} = \frac{w_i}{w_1 + w_2} \tag{3}$$

Node i-th in the fourth layer, the share of the i-th law in the final output through the node function (4) calculated. Wi is output of the third layer, {pi, qi, ri} parameters collection called parameters of consequences. This layer shows part of law result in the ANFIS model.

$$O_{i}^{4} = \overline{w_{i}}f_{i} = \overline{w_{i}}(p_{i}x + q_{i}y + r_{i})$$

$$\tag{4}$$

Only node of the fifth layer, the final output through the calculation of input signals outcome and acquired by using the formula (5),<sup>12</sup>.

$$O_i^4 = \text{overall output} = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$
(5)

#### 4.1.2 Differential Evolution Algorithm

This algorithm is a method of calculating the real functions (Real value) using evolutionary strategies.

Evolution process in this algorithm based on a gradual and continuous improvement in the initial guess (candidate response) and accordance the principles of all evolutionary class algorithms need to a fitness function (Fitness function) for compare the response.

The strength point of the algorithm DE compared with methods of solving other real equations (such as Newton's method), it does not need the gradient or function slope. As a result using this algorithm without any information about the type of the function can compute a relatively optimal response for a variety of continuous/noncontinuous multi-dimensional functions, time-variable and irregular hoped.

Differential Evolution algorithm steps are as follows<sup>14,15</sup>:

- Production of the initial population from candidate responses. Each candidate response is a vector of real numbers to the number of issue dimensions (unknown parameters) is.
- For each candidate response X, distinguished three responses a, b, c of the population have selected.
- Determine the random parameter R in the range of 1 and the dimensions of the issue.
- Calculate improved response Y so that for each dimension X (i) from X, if R equality with i and or estimate the combined probability p (i), the estimate formula Y (i) = a (i) + F \* (b (i) c (i)), to calculate the y (i). And otherwise, the X (i) intoY (i) is assigned.
- Accept the new response Y, if the fitness is greater than X.

• Repeat steps 2 to 5 until the realization of the termination condition.

The evaluation procedure in this paper presented is in Table 1.

Table 1.The evaluation criteria used in this study					
MSE (mean squared of predic- tion error)	$MSE = \frac{\sum_{i=1}^{n} (e_i)^2}{n}$				
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (e_i^2)}{n}}$				
MEAN (average of prediction error)	$Mean = \frac{1}{n} \sum_{i} (\hat{y}_i - \hat{y}_i)$				
STD (standard deviation of prediction error)	$STD = \sqrt{\frac{1}{n} \sum_{i} \left( \hat{y}_{i} - \overline{\hat{y}}_{i} \right)^{2}}$				

y and ŷ indicate actual value and predicted variables are at time t.

CR parameter is composition probability (such as genetics) and p (i) the chance of combination realize for each dimension of the response. F value is also an integer value and constant which is selected according to the type of problem. (The appropriate value selection method for F has story)

## 4.2 The Proposed Approach

In the proposed method in this research, to train ANFIS fuzzy neural network and optimizing the parameters of differential evolution algorithm is used and in the related data to distributed denial attack use.

The proposed algorithm in this study as follows:

- Pre-Processing.
- Get training data.
- Create a base phase system.
- Set the parameters of base phase system according to modeling error by differential evolution algorithm.
- Restore of phase system have best parameter values as the final result (the best parameters, those have the least amount of errors).

### 4.3 Pre-Processing Data

In pre-processing step in order to achieve better results values for each feature for all patients normalized between 0 to 1, then the matrix rows of data randomly displacement until the data arrangement out of from the collected initial state. Normalization, for achieve higher accuracy. If A is the amount of features into a column  $A_{max}$  the value maximum of the property and  $A_{min}$  value minimum of the property and  $\tilde{A}$  consider normal value, the following formula is used to normalize the data:

$$\overline{A} = \frac{A - A_{\min}}{A_{\max} - A_{\min}} \tag{6}$$

## 4.4 Fuzzy Neural Network Combination of ANFIS and Differential Evolution Algorithm

The combination of differential evolution algorithm and fuzzy neural network of ANFIS do in this case that training fuzzy neural system ANFIS and determining its parameters optimized value to be done using differential evolution algorithm. To do this, the training of fuzzy nervous system has turned into an optimization problem and we solve it using differential evolution algorithm. First, we must have a fuzzy system. Here, fuzzy systems used the type of Sugeno fuzzy. Using ANFIS, generated fuzzy systems teach. In ANFIS for training two methods of back propagation and hybrid used. In this combination, we intend used differential evolution algorithm for training and we achieved better results, used for optimization of the system error criteria. By differential evolution algorithm by changing the parameters this error is less and more and can be achieved optimal values of the parameters.

By using optimal amounts which by differential evolution algorithm is proposed, better results can be achieved and we have better diagnosis.

# 5. Finding

### 5.1 Evaluation Criteria

To assess compliance of a forecast with a data pattern was used measure criteria of forecast error. In this study, the assessment indicators MSE, RMSE, Mean and STD have used.

### 5.2 Data Set

Group ISI from the Laboratory of MIT L lincoln under the DARPA and AFRL/SNHS the first standard data for review and evaluation of intrusion detection systems, were collected. This information, during a few weeks in a simulation used for DARPA intrusion detection systems testing. This data set, based on years of information collecting (1998-1999) classified.

Data collection in 1999 year that to try and by monitoring Lee and collected during his doctoral project in the third international contest of knowledge discovery and data mining, KDD CPU 99<sup>15</sup> and at the Fifth Conference it was placed in use. This database includes standard connection records that series of attacks and intrusions simulated in a military network is included.

There are 4 categories of attack in this data set.

n DOS: In this attack, system resources excessive use is embedded and causes normal request, to provide the resources, be rejected.

n R2L: in the attack type of R2L, attacker with intrusion from far away to the victim machine, began to abuse the legal user of the account and attempts to send packets on the network.

n U2R: This type of attack run successfully at the victim machine and roots provided.

n Probing: In this type of computer attacks to collected information or to find known vulnerability potential to be scanned.

Four categories in the data sets KDD CPU 99 and its subsets in Table 2 are displayed.

Table 2.Existing categories in the data collection KDDCPU 99

DOS (Denial	R2L (User-	U2R	Probing
Of Service)	to-Root)	(Remote-to-	
		Local)	
Pingflood	Dictionary	Perl	IPsweep
SYNflood	FTP write	Fdformat_	Saint
Mailbomb	Sendmail	Loadmodule	Satan
DDoS		Eject	

#### 5.3 Analysis of Data

Here we intend to set parameter values and with using differential evolution algorithm, to obtain optimal response. To this end, consider the following relationship:

$$P_i^* = x_i \times P_i^0$$

Where  $P_i^*$  optimum value of parameter i and  $Pi^0$  is the corresponding parameter value in the initial basis fuzzy system.  $x_i$  Value will determine the differential evolution algorithm. For  $x_i$  can be determined number within the range of positive or negative of a specified amount.

The used cost function for the differential evolution algorithm is the error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (e_i^2)}{n}}$$

 $e_i = t_i - y_i$ 

n is the number of data total, t is target value and y is output value of model.

In this study, the number of generation is 100 and number of early population considered 20.

In this study, the criteria of mean square error (RMSE), Root of Mean Square Error (MSE), mean error (Mean) and standard deviation of errors (STD) we use.

First evaluate the predictive power of fuzzy neural network than neural network in identify.

In this test, we consider the MSE criterion to compare the results of ANFIS and neural network.

First, the data sets are normalized to fall between 0 and 1.

If A is the amount of features into a column  $A_{max}$  the value maximum of the property and  $A_{min}$  value minimum of the property A Ra value is considered normal and  $\tilde{A}$  consider normalize value, following formula is used to the data normalize:

$$\overline{A} = \frac{A - A_{min}}{A_{max} - A_{min}} \tag{7}$$

Then the data set split into two parts of training and testing. To this end, 70% of the data to training, 30% assign to the testing.

In the next step applied the ANFIS algorithm on the pre-processed data.

For comparison with standard neural will consider network MSE.

Here, for the training data set, the MSE value obtained 0.006.

For running the data on the neural network use neural network nf tool tools available in MATLAB. 70 percent of data as a training set and from between thirty percent remaining 15 percent as a validation dataset and 15% consider as a testing data set.

Results are as Figure 2.

Results		
	뤓 Samples	🔄 MSE
🗊 Training:	345815	5.38018e-2
🕡 Validation:	74103	4.68620e-3
🔰 Testing:	74103	6.22254e-2

**Figure 2.** The results of the implementation of the neural network on the study data collection.

As can be seen, MSE values in ANFIS fuzzy neural network has been lower than the neural network

Table 5. Compares the algorithm used in this study and Artific algorithm							
Method	Data set	MSE	RMSE	Mean	STD		
The combination method of differential evolution algorithm	Training	0.0015597	0.039493	0.0054167	0.0393395		
and ANFIS	test	0.0029806	0.054595	0.017419	0.052597		
ANFIS fuzzy neural network	Training	0.0066653	0.081641	0.023472	0.078743		
	test	0.082323	0.090732	0.047742	0.078431		

Table 3. Compares the algorithm used in this study and ANFIS algorithm

In the case of ANFIS fuzzy neural network for training data set, the error average value equal to 0.023472, the standard deviation value equal to 0.078743, the error value of MSE equal to 0.0066653 and the error value of RMSE obtained are equal to 0.081641. For the training data set, the average value of error equal to 0.047742, the standard deviation value equal to 0.078431, the error value of MSE equal to 0.0082323 and error value of RMSE obtained 0.090732.

In the next step, the differential evolution combination algorithm and ANFIS applied on the pre-processing data. The following results were obtained.

For training data collection, the errors mean value equal to 0.0054167, the standard deviation value equal to 0.039395, the MSE error value equal to 0.0015597a nd the RMSE error value, equal to 0.039493 obtained.

Result of running the differential evolution combination algorithm and ANFIS on the testing data set this means that the mean value of errors equal to 0.017419, the standard deviation value of error equal to 0.052597. As for the test data set with the implementation of differential evolution algorithm and ANFIS, the MSE error value equal to 0.0029806 and the error value of RMSE equal to 0.054595.

Table 3 display compares between the usage combination algorithm in this study and ANFIS.

In this study, using the program at first the data normalizing data, the order of records commix and then training dataset equal to 70% of the data set means the 73 record and the test data collection equal to the remaining 30% consider and implement the programs. As in Table 3 can be seen, from the point of view of search criteria, the performance of differential evolution combination method and ANFIS is better than ANFIS.

# 6. Conclusion

Nowadays, security and its branch in the web space have become critical and epidemics especially for owners of sites and more importantly for Web server administrators because vulnerabilities and security weaknesses as an inhibitors factor in the path of progress and development goals on the web. In this study pay to the diagnosis of distributed denial attacks to providing the combination method of fuzzy neural network ANFIS and differential evolution. To evaluate the results, at first fuzzy neural network compared with the neural network. To compare these two algorithms, the Mean Square Error criteria (MSE) were used. The results show the MSE error of less were in the ANFIS and this indicates that, fuzzy neural network than the neural network has higher efficiency. Then, with the goal of increasing the accuracy of diagnosis to train ANFIS fuzzy neural network use the differential evolution algorithm and the obtained results of this method were compared with ANFIS fuzzy neural network. The criteria used in this study to compare ANFIS and ANFIS combination algorithm and Genetics, MSE, RMSE and standard deviation of error and the mean errors. Results of comparison according to these criteria showed that a combination algorithm ANFIS and differential evolution compared to the ANFIS algorithm has higher accuracy. One of the limitations of this study, the volume of related data collection and lack of single data to apply for research students and researchers.

## 7. References

- Bhuyan MH, Bhattacharyya DK, Kalita JK. An empirical evaluation of information metrics for low-rate and highrate DDoS attack detection, Pattern Recognition Letters. 2015 Jan; 51:1–7.
- 2. Wang F, Wang H, Wang X, Su J. A new multistage approach to detect subtle DDoS attacks. Mathematical and Computer Modelling. 2012 Jan; 55(1-2):198–213.
- 3. Thing Ling V. Adaptive response system for distributed denial-of-service attacks. [Ph.D. Thesis]. College London; 2008.
- 4. Yu J, Lee H, Kim MS, Park D. Traffic flooding attack detection with SNMP MIB using SVM. Computer Communications. 2008 Nov; 31(17):4212–9.
- 5. Mansfield-Devine S. Computer Fraud & Security. 2014; 10:15–20.
- 6. Santhanam T, Ephzibah EP. Heart disease prediction using hybrid genetic fuzzy model. Indian Journal of Science and Technology. 2015 May; 8(9):797–803.

- 7. Okolobah V, Ismail Z. New approach to peak load forecasting based on EMD and ANFIS. 2013 Dec; 6(12):5600–6.
- 8. Zaroo P. A survey of DDoS attacks and some DDoS defense mechanisms. A part of course textbook Advanced Information Assurance (CS 626) in Purdue University. 2002.
- 9. Shiaeles SN, Katos V, Karakos AS, Papadopoulos BK. Real time DDoS detection using fuzzy estimators. Computers and Security. 2012 Sep; 31(6):782–90.
- Raj Kumar PA, Selvakumar S. Detection of distributed denial of service attacks using an ensemble of adaptive and hybrid neuro-fuzzy systems. Computer Communications. 2013 Feb; 36(3):303–19.
- 11. Vissers T, Somasundaram TS, Pieters L, Govindarajan K, Hellinckx P. DDoS defense system for web services in a cloud environment. Future Generation Computer Systems. 2014 Jul; 37:37–45.

- 12. Spyridopoulos T, Karanikas G, Tryfonas T, Oikonomou G. A game theoretic defence framework against DoS/DDoScyber-attacks. Computers and Security. 2013 Oct; 38:39–50.
- Zhou W, Jia W, Wen S, Xiang Y, Zhou W. Detection and defense of application-layer DDoS attacks in backbone web traffic. Future Generation Computer Systems. 2014 Sep; 38:36–46.
- Qi Y, Tian J, Dai RW. Fingerprint classification system with feedback mechanism based on genetic algorithm. in Proc Int Conf Pattern Recognition. Brisbane QD 1998 Aug 16-20; 1. p. 163–5.
- 15. Available from: http://www.mshams.ir/blogs/2010/11/ Introduction-evolution-algorithm-differential-differential-evolution