Journal of Engineering Science and Technology Review
Special Issue on Telecommunications, Informatics, Energy and Management 2019

JOURNAL OF
Engineering Science and Technology Review

Conference Article

www.jestr.org

Neural Network Implementation for Detection of Denial of Service Attacks

Irina Topalova[1*], Pavlinka Radoyska[1] and Strahil Sokolov[1]

[1]*University of Telecommunications and Posts, Department IT, Sofia, Bulgaria*

Received 26 September 2019; Accepted 17 February 2020

Abstract

Denial of Service attacks are considered a major risk because they can easily interrupt a service, a business or educational process. These attacks are relatively simple to conduct, even by an unskilled attacker and cause significant loss. For this reason, it is particularly important that these attacks are detected, recognized and blocked in time. Most of the advanced methods and tools to protect against such attacks are based on monitoring and constant tracking in order to detect suspicious IP traffic. The application of these methods is associated with additional computational resources and expertise, which leads to subjectivity in the threat assessment. Therefore, it is necessary to propose methods for automated adaptive detection and recognition of Distributed Denial of Service attacks.
This study presents a method for automated detection and recognition of some of the Distributed Denial of Service attacks, by means of an automated adaptive system, based on a multilayer neural network. It is trained both normal and with signals reflecting different traffic conditions when Distributed Denial of Service attacks occur. The neural network is tested to recognize "baseline signals", representing different normal traffic conditions and to detect the abnormal traffic situations. The research is conducted for different kinds of internal Distributed Denial of Service attacks on a real Local Area Network. The obtained recognition accuracy results are represented and the achieved benefits are discussed.

Keywords: Denial of Service; Network security; Artificial intelligence; Neural networks.

1. Introduction

Distributed Denial of Service (DDoS) attacks are launched through remotely managed, well-organized and widely distributed zombie network computers. Many traffic or service requests are sent simultaneously or continuously to the target system. As a result, the targeting (victim) system becomes unusable, responds slow or completely collapses as a result of the attack [1].

Most of the advanced methods and tools to protect against such attacks are based on monitoring and constant tracking in order to detect suspicious IP traffic. The application of these methods is associated with additional computational resources and expertise, which leads to delay and also to subjectivity in the threat assessment. Therefore, it is necessary to propose methods for automated adaptive detection and recognition of DDoS attacks. A detailed analysis of the types of DDoS attacks, modern methods and defence tools were made by the authors of [2]. They conclude, that the artificial intelligence approach is currently one of the best performing ones.

The aim of this study is to propose and experiment the accuracy of detecting and classifying the most commonly used, even by unsophisticated intruders DDoS attacks, by training a Multi-Layer-Perceptron Neural Network (MLPNN). The neural network is tested to recognize "baseline signals", representing different normal traffic conditions and to detect the abnormal traffic situations. The research is conducted for different kinds of internal DDoS attacks on a real Local Area Network (LAN). The obtained recognition accuracy results are represented and the achieved benefits are discussed.

2. Related work

2.1 Classification of Distributed Denial of Service attacks

Security falls into three categories: confidentiality, availability, and integrity. It is obvious that Distributed Denial of Service (DDoS) attacks belong to the availability category DDoS attacks aim to exhaust some Internet resources in order to make the services unreachable for the legitimate or normal users by sending a large number of invalid requests - to target servers that provide the services. Legitimate traffic and the attack traffic differ in intention, not in the content.

There are different classifications of network attacks. In the [3] attacks are considered in the context of the TCP/IP model (table 1). Denial of Service (DDoS) attacks can be defined as Multiple Layers Attacks because they include mechanisms operating on Network and Data Link Layer.

A wider classification of DDoS attacks was made in [2], [4]. The authors apply four classification criteria: Degree of automation, Impact, Attack rate dynamics and Exploited Vulnerability.

* E-mail address: itopalova@abv.bg
ISSN: 1791-2377

1. The degree of automation reflects the extent of direct human participation in the attack.

    a. Manual attacks: Network investigation, malicious code installation, launching and guessing are performed manualy by the human.
    b. The semi-automatic attack includes two steps. The first step is made by the attacker which involves setting the type of the attack, selecting the address of the victim and organizing the attack timing/waives of the handlers' machines. The second step is deploying the handlers to automatically controlling the zombies and running the attack.
    c. Automatic: the human only initiates the attack by just a single command.

2. In terms of damage, the attacks are divided into Disruptive and Degrading. Disruptive leads to a complete denial of service. Degrading are more difficult to detect because the system continues to operate, although with limited capacity.
3. Attack rate dynamics is determined by the duration of the attack.

    a. Continuous rate attacks exhaust the resources of the attacked object by continually sending messages/requests. They cause a quick disrupting effect.
    b. Variable rate attacks are more flexible and more robust to detect. They are based on protocols' gaps. Low-Rate DDoS attacks are the example of Variable rate attack The effect does not appear very fast and can be expressed in both degrading and disrupting.

4. The most complex classification scheme is based on Exploited Vulnerability. Many authors classify DDoS attacks only in terms of Exploited Vulnerability [3] [4] [5] [6]. Attacks can be divided into four main groups under this criterion:

    a. Flooding attack – based on User Datagram Protocol(UDP) and Internet Control Message Protocol(ICMP).
    b. The aim of flooding attacks is to send a huge amount of traffic to the victim system to exhaust bandwidth and resources of the victim system. The attacks are based on ICMP and UDP protocols and the "replay" mechanism. There are three sub variants: direct flooding attacks, reflection-based flooding attacks and amplification flooding attacks.

        i. Direct flooding attacks use the victim system address as the destination address. ICMP flooding attack is executed by sending the huge amount of ICMP ECHO-REQUEST messages to the victim system. The victim system answers with the same amount of ICMP ECHO-REPLAY [2]. A UDP Flood attack is executed by sending the huge amount of UDP packet to a randomly generated ports on the victim system. The victim system releases that there not waiting application on the port and answers with ICMP "destination port unreachable".
        ii. Reflection-based flooding attacks are indirect attacks. The address of the victim system is used as the source address. The request packets with spoofing source address are sent to the intermediate system, called a reflector. The Reflector answers to the victim system with the normal packets. In this case, both the victim system and the reflector are overload.
        iii. Amplification flooding attacks are Reflection-based flooding attacks with amplification effect. The request packets with spoofing source address are sent to the broadcast IP target address. All receivers in the network segment send their replay packets to the victim system. Smurf and Fraggle attacks are amplification attacks, based respectively on ICMP and UDP protocols.

    c. Protocol exploit attacks are based on some specific feature or implementation bug of some protocol in order to consume excess amounts of victim resources. A classic example of these kinds of attacks is TCP-SYN. The incomplete 3-way handshake is performed. The TCP client (agent) sends TCP-SYN, the TCP server (victim) responds with TCP-SYN/ACK, but the third step does not happen. TCP-ACK is not received. Victim system waits a predefined time period and then rejects the session. When these unestablished TCP sessions are enough many, they block the system entrance and it cannot create real TCP sessions. Low-rate DDoS (LRDDoS) attack is another example [7] [8]. The TCP session is established successfully. The client activates the TCP slow-start mechanism, during the delivery of the requested data. TCP communication is in a run, the network traffic is low, but the TCP server is not accessible. Duplicate ACKs (or redundant packets) is another type of LRDDoS attack and is based on the fast-recovery TCP mechanism. Agent establishes a TCP session and sends a request to the victim system. Victim system responded with the first segment. Agent acknowledges the small part of the received segment. Victim system sends the unacknowledged portion of data. Agent acknowledges the next small part of the received segment and so on. The TCP session is in a run, communication is not interrupted, but the victim system is forced to generate a huge amount of extra traffic.
    d. Malformed packet attacks rely on incorrectly formed IP packets that are sent from agents to the victim in order to crash the victim system. Malformed packet attacks are divided into two groups: IP address attack and ill-formed packet attack. In the IP address attack, the IP address of the source and destination is the same and is the address of the victim system. As a result, the OS of the victim system is becoming confused and then crashes. In the IP ill-formed packet attack, the optional fields may jumble and the additional traffic analysis time is forcefully consumed.

The subject of our work is to detect flooding attacks. They are focused on disturbing genuine user's connectivity by exhausting victim network's bandwidth.

2.2 Defence mechanisms versus flooding DDoS attacks

Different publications offer different criteria and classification schemes for Defence mechanisms [2] [9]. The most complete classification is proposed in [2]. The authors

propose to divide the mechanisms of protection into two main groups: Statistical and Artificial Intelligence.

The statistical approach in flooding DDoS attack is quite spread applied. A statistical model for regular traffic is computed, and then, a statistical deduction test is used to determine whether a new traffic is an instance of this model. Traffic instances that do not meet the rules are classified as inconsistencies. Two trends are observed in the statistical approach: parametric and non-parametric. Parametric methods are a threshold-based model, statistical moment's parametric identification, spectral analysis. The most common non-parametric detection approaches are D-WARD, change aggregation trees (CATs), histogram-based detection, flow feature value (FFV), regression analysis, Markov method, statistical segregation, time series. The statistical approach shows a high degree of efficiency, but there is a serious problem - it needs preliminary traffic monitoring to build the statistical model and cannot react quickly to new types of attacks.

In recent years the Artificial Intelligence (AI) approaches are increasingly successful in Flooding DDoS attack defence. The main Machine learning technologies as the Bayesian theory of decision, multivariate techniques, clustering, multilayer perceptron, linear discrimination, local models, classification trees, reinforcement learning, and hidden Markov models are involved in some defence systems. The main advantage of AI defence methods is flexibility. Systems based on AI methods may alter the performance process based on recently collected data. The system can improve its performance on some test cases based on previous results.

Table 1. TCP/IP Layer based attack examples

Layer	Example of Attack
Application	Repudiation, Viruses, Malicious URLs
Transport	Session Hijacking
Network	Distributed Denial DDoS, Information Disclosure Spoofing Attack of Service, Packet Replication
Data Link	Flooding Attacks
Physical	Cable Cut, Jamming
Multiple	Denial of Service Attacks

3. The proposed method

The proposed method is based on the training and testing the MLPNN structure, shown in Figure 1. The neural network is trained with different signals representing normal (baseline) Internet Protocol (IP) traffic in LAN; six attacks with Ping flooding with different size of the ICMP packages; Slow rate attack and Replay attack.

The number of packets per second is measured through 50 seconds. Each amount of packets is connected to a separate input neuron. Thus the input MLPNN layer consists of 50 neurons, the second (hidden layer) has 25 neurons and the output layer has 8 neurons. The number of output neurons corresponds to the eight types of attacks mentioned above.

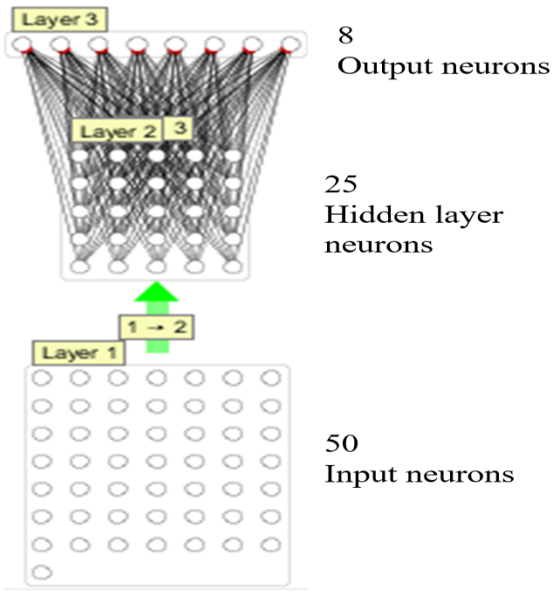

Fig 1. Applied 3-layered MLPNN structure with 50 neurons in the input layer; 25 neurons in the hidden layer and 8 neurons in the output layer

4.Experiments and results

The MLPNN was trained with six Ping flooding signals (Ping of Death) each having the correspondent packet size (PS) of 10000, 18000, 25000, 38000, 50000 and 65000 bytes. These train attack signal samples are shown in Figure 2. A software package was used to generate the other two types of attacks, namely *Slow rate* and *Replay attack*. The train and test samples of Slow rate and Replay signals are represented in Figure 3 and Figure 4 respectively.

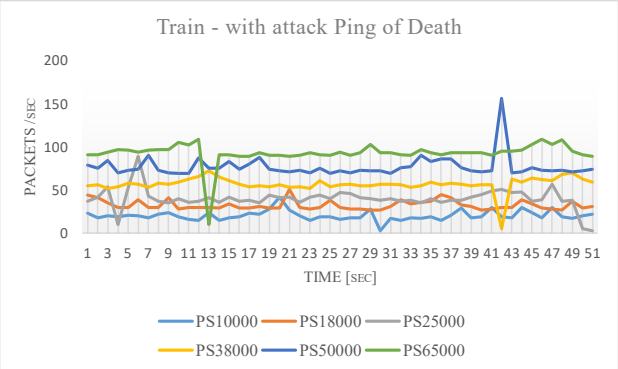

Fig 2. Train attack signal samples with different packet size

Fig 3. Train and test samples of Slow rate attack signals

The proposed MLPNN was trained also with "baseline" signals, shown in Figure 5, corresponding to the normal IP traffic, measured at different moments of the busy work period of the LAN.

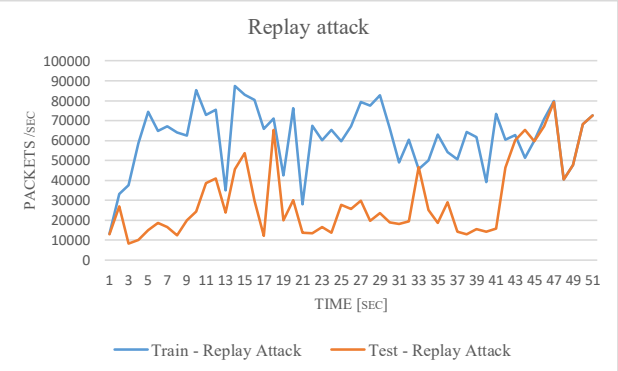

Fig 4. Train and test samples of Replay attack signals

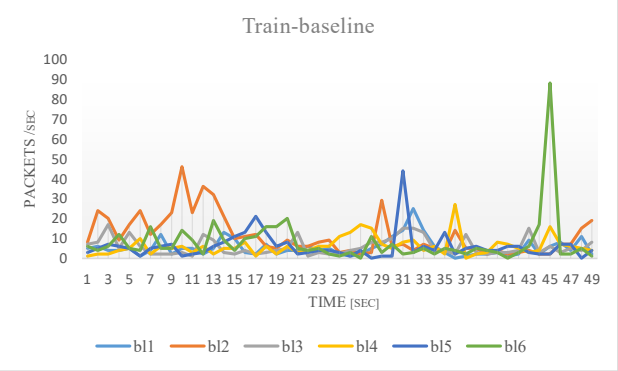

Fig 5. Train "baseline" signals, corresponding to the normal IP traffic

The MLPNN structure was trained with the input signals so described, applying tangens hyperbolicus as activating function of all neural network layers. The MLPNN stopped the iterations after reaching the preliminary fixed output error of 0.1, which error was achieved after 1267 iterations. Backpropagation algorithm was applied for training the MLP network. "SimBrain" software was used to implement the train and test stages of the method.

The next step is the implementation of the test phase, at which to the inputs of the already trained neural network, are transmitted signals that have not participated in the training samples, presenting various attacks and baseline traffic. Test "baseline" signals and test Ping flooding with PS15000, PS20000, PS30000, PS45000 and PS60000 are represented respectively in Figure 6 and 7.

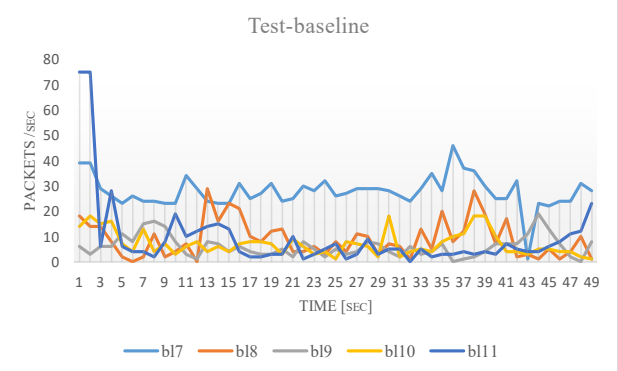

Fig 6. Test samples of "baseline" signals

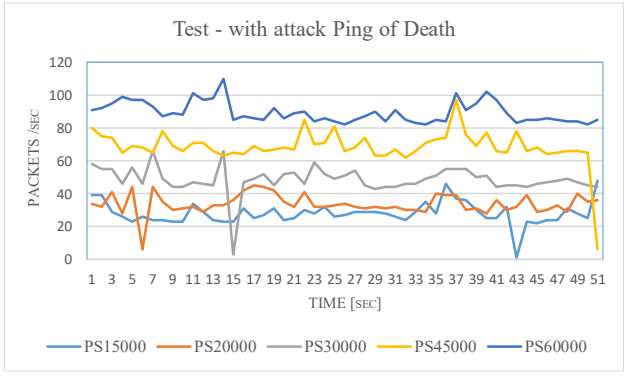

Fig 7. Test Ping attack signal samples with different packet size

The obtained recognition results of Ping attack with PS 20000 (output neuron number 3 has maximum potential and right recognized as PS25000) and PS 45000 (output neuron number 5 has maximum potential and right recognized as PS50000) are shown in Figure 8 and 9 respectively. They show the desired MLPNN output values and also the really achieved approximated output values in the test phase.

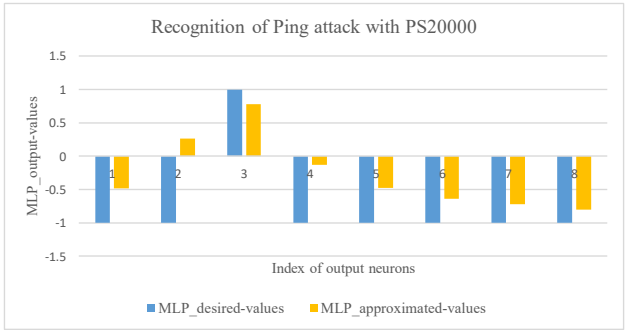

Fig 8. Recognition of Ping attack with PS 20000

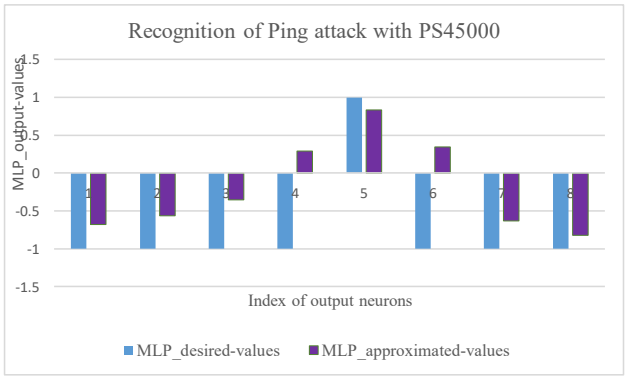

Fig 9. Recognition of Ping attack with PS 45000

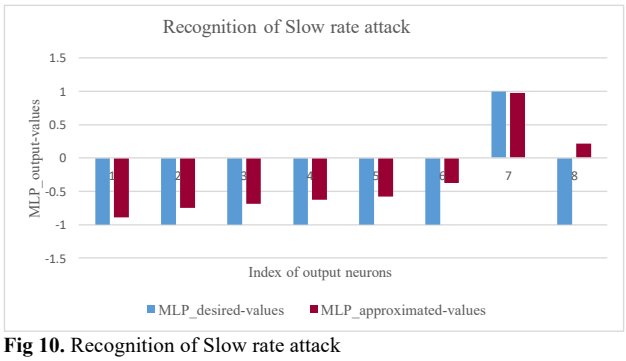

Fig 10. Recognition of Slow rate attack

Figure 10 represents the recognition of Slow rate attack – the maximum potential has achieved output neuron number 7, corresponding to the trained kind of attack. The

approximation error at the MPLNN outputs calculated after the test phase was defined according to equation (1):

$$Approx_{error} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(NN_{od} - NN_{or})^2} \qquad (1)$$

where *NNod* is the desired output neuron value, *NNor* is the real value at the test phase and *N* is the number of recognized classes i.e. N = 8. The recognition of "baseline" signals, reflects in no high output potential of any output NN neuron. Table 2 shows the calculated $Approx_{error}$ for the corresponding recognized attack class.

Table 2. $Approx_{error}$ for the corresponding recognized class.

Recognized class attack	**1: PS 10000**	**2: PS 18000**	**3: PS 25000**	**4: PS 38000**	**5: PS 50000**	**6: PS 65000**	**7: Slow rate**	**8: Replay**
MLP Approximation error	0.28	0.28	0.224	0.219	0.262	0.21	0.19	0.1837

5. Conclusions

The results of the recognition of the attack show high accuracy in approximating the values of the neurons in the output MLP layer in the test phase. This accuracy is the highest in recognizing Slow rate and Relay Attack. An advantage of the MLPNN learning method is the ability to detect attacks from close to those already trained in the network, as if correlated to one of the closest learned classes of attacks. An example of this is the correct classification of attack Ping PS20000 and PS 45000 to the nearest trained classes - PS25000 (class 3) and PS 50000 (class 5) respectively. Additionally, the proposed method does not require additional computing resources, subjectivity in the threat assessment or any preliminary knowledge.

References

1. S. Jamali and V. Shaker, ''Defence against SYN flooding attacks: A particle swarm optimization approach,'' Comput. Elect. Eng., vol. 40, no. 6, pp. 2013–2025, 2014.
2. Khalaf, B. A., S. A. Mostafa, A. Mustapha, M. A. Mohammed, W. M. Abduallah, Comprehensive Review of Artificial Intelligence and Statistical Approaches in Distributed Denial of Service Attack and Defence Methods, IEEE Access, Vol.7, pp. 51691-51713, Vol.7, 2019.
3. Bhavasar, D. "A survey on distributed denial of service attack and defence," in INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, pp. 41-47, 2015
4. Zargar, S. T., Towards Coordinated, Network-Wide Traffic Monitoring for Early Detection of DDDoS Flooding Attacks, Pittsburgh: University of Pittsburgh, 2014
5. AAMIR, M., M. A. ZAIDI, "A Survey on DDDoS Attack and Defence Strategies: From Traditional Schemes to Current Techniques," Interdisciplinary Information Sciences, vol. 19, no. 2, pp. 173-200, 2013
6. D.G. Kumar, C. V. Guru Rao, A. Gopal J, P. Mohan, "Distributed DDoS Attacks: Classification and Defence Mechanisms," International Journal of Information Engineering and Electronic Business(IJIEEB), pp. PP.703-708, 2013
7. Luo, J., X. Yang, J. Wang,J. Xu, J. Sun, K. Long, "On a Mathematical Model for Low-Rate Shrew DDDoS," IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, vol. 9, no. 7, pp. 1069-1083, 2014
8. Hubballi, N., J. Santini, "Detecting TCP ACK storm attack: a state transition modelling approach," IET Networks, vol. 7, no. 6, pp. 429-434, 2018
9. Douligeris, Chr., A. Mitrokotsa, "DDDoS attacks and defence mechanisms: classification and state-of-the-art," Computer Networks, vol. 44, p. 643–666, 2004