ORIGINAL RESEARCH PAPER
Volume-6 | Issue-12 | December-2017 | ISSN No 2277 - 8179 | IF : 4.176 | IC Value : 93.98
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

A REVIEW OF MACHINE LEARNING TECHNIQUES EFFICIENCY IN DOS ATTACK DETECTION

Engineering

Zerina Masetic* International Burch University, Sarajevo, Bosnia and Herzegovina. *Corresponding Author

Nejdet Dogru International Burch University, Sarajevo, Bosnia and Herzegovina

ABSTRACT

Denial of Service (DoS) attacks are among most common attacks on the network. Therefore, the efficient attack detection mechanism should be developed to decrease the rate of these attacks. One approach that has been taken for DoS attack detection is the application of machine learning techniques. This paper aims to provide a review of recently published research papers that use machine learning techniques for DoS attack detection and verify their efficiency.

KEYWORDS

Denial of Service (DoS) attacks, machine learning methods, attack detection

INTRODUCTION

Although there are many active threats on the Internet, Denial of Service (DoS) attacks are among most common attacks on the network. DoS attacks prevent users from normal service access, due to the excessive consumption of network resources, memory, processor, etc. These attacks usually target web servers, CPUs, network or cloud storage (Top Threats Working Group, 2016).

The most common DoS attacks are when attackers "flood" the network with many requests at the same time, making the server unavailable to answer to that many requests. In this way, legitimate users are not able to communicate with the server. In a distributed denial-of-service (DDoS) attack, attacker uses more computers, which are geographically distributed, to launch the attack (Fig.1.) (McDowell, 2013).

In July 2009, the website of South Korean's largest daily newspaper, large-scale online auction house, a bank, and country's president website, White House, Pentagon, U.S. Forces Korea were under the DDoS attack (Sudworth, 2009), and in August 2013, the cloud hosting company, DigitalOcean was hit by DDoS attack (Kovacs, 2013). Furthermore, in the third quarter of 2017, resources from 98 different countries were victims of DoS attacks. The popular DoS attack recently is ransom DoS (RDoS) attack, in which attackers aim to gain the revenue from attacks (Khalimonenko, Kupreev, & Ibragimov, 2017).

It is important to deeply analyze these types of attacks, how they are caused, how to recognize and describe them, where they happen. After this information is obtained, the right detection mechanism should be developed. One approach that has been taken in research studies is application of machine learning methods.

The aim of this paper is to summarize the results in the recent works (from 2014 to 2017) in the area of DoS attack detection using machine learning techniques, considering the input features they used for analyzing these attacks and detection rate they obtained.

The paper is organized as following: Section II gives review of research works in mentioned area, with emphasis on efficiency of machine learning techniques. Section III presents the discussion and concludes the paper.

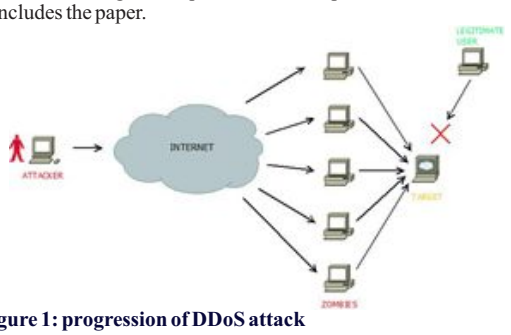

Figure 1: progression of DDoS attack

LITERATURE REVIEW

To detect DoS attacks, it is important to detect some unusual behavior and protect the system from possible attacks. These attacks can be described by two categories of input features: system performance data and network traffic data. System performance data, such as CPU, network, memory, disk, hypervisor, and OS performance, etc. could be seen at the end system while the attack is being executed. The other category, network traffic data, such as: source and destination IP address, protocol, packet length, port numbers, flags, etc. can give us more information about DoS attack before the system is attacked (Masetic, Hajdarevic, & Dogru, 2017).

These features are input to machine learning techniques. There are several studies that focuses on DoS attack detection using machine learning techniques. Barati et al. (Barati, Abdullah, Udzir, Mahmod, & Mustapha, 2014) proposed the approach that combines Genetic Algorithm (GA) for feature selection and Artificial Neural Networks (ANN) for DDoS attack detection. As the input features, they used freely available dataset to the research community, CAIDA dataset. The results are presented through several metrics: recall, precision, ROC curve, F-measure, true positive (TP) and false positive (FP) rates. The proposed approach achieved the rate 99.99 % correctly classified instances. Dhanabal & Shantharajah (Dhanabal & Shantharajah, 2015) applied Decision Trees (DT), Support Vector Machine (SVM) and Naive Bayes (NB) machine learning methods for attack detection on freely available NSL-KDD dataset, a refined version of KDD99 dataset. This dataset contains 41 input features to detect four types of attack, among which are DoS attacks. They obtained DoS attacks detection rates of 99.1 %, 98.7% and 75.2% for DT, SVM and NB, respectively. Gao et al. (Gao, Feng, Kawamoto, & Sakurai, 2016) proposed the approach for detection distributed reflection DoS (DRDoS) attacks that combines five features and machine learning algorithms for effective DRDoS detection. The classification process was done with the use of SVM algorithm, with normalized polynomial kernel. This algorithm was applied to five datasets obtained during the attack simulation. The results obtained vary from 89.74% until 100% detection rate. Kumar, Lal, and Sharma (Kumar, Lal, & Sharma, 2016) proposed a system consisting of a packet sniffer, a feature extractor, and a classifier, for detecting DoS attacks on VMs in the cloud. Specifically, they applied one-class SVM classifier on the network traffic dataset for classification of ICMP flood, Ping-of-Death, UDP flood, SYN flood, TCP land and DNS flood attacks and achieved classification accuracy of 100%, 94%, 97%, 96%, 98% and 99% respectively. Alkasassbeh et al. (Alkasassbeh, Hassanat, Al-Naymat, & Almseidin, 2016) proposed a detection mechanism that involves application of NB, Multilayer Perceptron (MLP) and Random Forest (RF) for "modern" DoS attacks, such as SIDDoS, HTTP flood, UDP flood and Smurf attack, considering some of the features: source and destination addresses, flags, number of packets, number of bytes in each packet, etc. They obtained following classification accuracies: 98.63%, 98.02% and 96.91% for MLP, RF and NB, respectively. Sofi, Mahajan & Mansotra (Sofi, Mahajan, & Mansotra, 2017) did a quite similar study to the previous authors. They proposed a detection model that incorporates four machine learning techniques: MLP, SVM, NB

and DT. These methods are applied to detect several types of DoS attacks, such as: SIDDoS, HTTP flood, UDP flood and Smurf. The overall accuracies were 96.89%, 98.89%, 98.91%, and 92.31% for NB, DT, MLP-ANN and SVM, respectively. He, Zhang & Lee (He, Zhang, & Lee, 2017) proposed a DoS attack detection system on the source side in the cloud that uses machine learning techniques for attack classification. Following DoS attack types are considered: SSH brute force attack, SYN flood, ICMP flood, and DNS reflection attacks. As detection techniques, supervised and unsupervised machine learning techniques were applied: Linear Regression (LR), SVM (with linear, RBF or polynomial kernels), DT, NB and RF algorithms, k-means and Gaussian Mixture Model for Expectation-Maximization (GMM-EM). The classification accuracy rates vary from 63.26% for GMM-EM to 94.96% for RF. Additionally, authors used joint data from multiple VMs, and obtained better results, from 66.53% to 99.73% for SVM with linear kernel.

The previous works is summarized in the Table 1.

TABLE – 1 PREVIOUS WORK ON DOS ATTACK DETECTION USING MACHINE LEARNING TECHNIQUES

Authors	Year	Input features	Machine learning methods	Classification accuracies
Barati et al.	2014	Total forward packets, maximum of backward inter arrival time, standard deviation of backward inter arrival time, mean of idle time, back urgent counter, etc.	Genetic Algorithm (GA) + Artificial Neural Networks (ANN)	99.99 %
Dhanabal & Shantharajah	2015	Duration, protocol type, flags, services, source and dest. bytes, content-, time- and host-based features	Decision Tree (DT) Support Vector Machine (SVM) Naïve Bayes (NB)	99.1 % 98.7 % 75.2 %
Gao et al.	2016	Number of packets with only IP header, UDP packet size, total number of packets, difference in the number of packets sent to and from target, maximum number of packets	Support Vector Machine (SVM) with normalized polynomial kernel	89.74 – 100 %
Kumar, Lal & Sharma	2016	Source and dest. IP addresses, bytes of data transferred, protocol, etc.	One-class Support Vector Machine (SVM)	94 – 100 %
Alkasassbeh et al.	2016	Source and destination IP addresses, flags, number of packets, number of bytes in each packet, etc.	Naïve Bayes (NB) Multilayer Perceptron (MLP) Random Forest (RF)	98.63 % 98.02 % 96.91 %
Sofi, Mahajan & Mansotra	2017	Source and destination IP addresses, flags, packet number, sequence number, packet size, packet type, etc.	Naïve Bayes (NB) Decision Tree (DT) Multilayer Perceptron (MLP) Support Vector Machine (SVM)	96.89% 98.89% 98.91% 92.31%
He, Zhang & Lee	2017	Number of key exchange packages, inbound/outbound DNS package ration, ICMP package rate, SYN/ACK ratio	Linear Regression (LR) SVM Linear Kernel SVM RBF Kernel SVM Poly Kernel Decision Tree (DT) Naïve Bayes (NB) Random Forest (RF) K-means Gaussian EM	97.77 % 99.73 % 98.15 % 99.13 % 99.07% 98.47 % 99.53 % 87.76 % 66.53 %

DISCUSSION AND CONCLUSION

Considering the results of research studies shown in Table 1, machine learning techniques could be the efficient tool for DoS/DDoS attack detection. Some authors of papers used freely available datasets, that contain features such as source and destination IP addresses, number of packets being sent, size of packets, flags, etc., and on the other side, some of them simulated attacks and obtained similar features. Furthermore, both supervised and unsupervised machine learning techniques were applied for attack detection. From the results shown above, we could notice that supervised machine learning techniques performed better than unsupervised. However, we cannot claim that these techniques usually perform better, because we presented only two results of unsupervised machine learning methods application in the table. The reason is that not many research articles could be found with these types of methods. Furthermore, it is not easy to determine what supervised learning technique perform the best, as results changed from study to study. Obviously, input features play a key role in detection technique performance. However, we can conclude that supervised machine learning techniques can be successfully applied for DoS/DDoS attack detection, considering network traffic features as input, as it can be seen in the Table 1.

REFERENCES

1. Alkasassbeh, M., Hassanat, A., Al-Naymat, G., & Almseidin, M. (2016). Detecting Distributed Denial of Service Attacks Using Data Mining Techniques. (IJACSA) International Journal of Advanced Computer Science and Applications, 436-445.
2. Barati, M., Abdullah, A., Udzir, N. I., Mahmod, R., & Mustapha, N. (2014). Distributed Denial of Service Detection Using Hybrid Machine Learning Technique . International Symposium on Biometrics and Security Technologies (ISBAST) , (pp. 268-273).
3. Dhanabal, L., & Shantharajah, S. (2015). A Study on NSL-KDD Dataset for Intrusion Detection System Based on Classification Algorithms. International Journal of Advanced Research in Computer and Communication Engineering, 446-452.
4. Gao, Y., Feng, Y., Kawamoto, J., & Sakurai, K. (2016). A Machine Learning Based Approach for Detecting DRDoS Attacks and Its Performance Evaluation. 11th Asia Joint Conference on Information Security, (pp. 81-86).
5. He, Z., Zhang, T., & Lee, R. (2017). Machine Learning Based DDoS Attack Detection From Source Side in Cloud. 4th International Conference on Cyber Security and Cloud Computing, (pp. 114-120).
6. Khalimonenko, A., Kupreev, O., & Ibragimov, T. (2017). DDoS attacks in Q2 2017. Moscow: Kaspersky Lab.
7. Kovacs, E. (2013, August 28). Cloud Hosting Company DigitalOcean Hit by DDOS Attack. (Softpedia News)
8. Kumar, R., Lal, S., & Sharma, A. (2016). Detecting Denial of Service Attacks in the Cloud. IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing.
9. Masetic, Z., Hajdarevic, K., & Dogru, N. (2017). Cloud computing threats classification model based on the detection feasibility of machine learning algorithms. 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), (pp. 1314-1318).
10. McDowell, M. (2013, February 6). Understanding denial-of-service attacks. (National Cyber Alert System, Cyber Security TIP ST04-015) Retrieved September 2017
11. Sofi, I., Mahajan, A., & Mansotra, V. (2017). Machine Learning Techniques used for the Detection and Analysis of Modern Types of DDoS Attacks. International Research Journal of Engineering and Technology (IRJET), 1085-1092.
12. Sudworth, J. (2009, July 9). New 'cyber attacks' hit S Korea. (BBC News)
13. Top Threats Working Group, C. (2016). The Treacherous 12 Cloud Computing Top Threats in 2016. Cloud Security Alliance.
14. The algorithm was applied to five sample datasets, combines with 5 test datasets.
15. Five categories of DoS attacks were classified.