International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 2153-2156 www.ijaem.net ISSN: 2395-5252

Detect DDoS (Distributed Denial of Service) Attack Using Machine learning

Rushikesh Vilas Chaudhari

Submitted: 15-06-2022 Revised: 25-06-2022 Accepted: 27-06-2022

ABSTRACT: Distributed Denial of Service (DDoS) attacks is a classification problem in machine learning. In relevance to Cloud Computing, the task of identification of DDoS attacks is a significantly challenging problem because of the computational complexity that has to be addressed. Fundamentally, a Denial of Service (DoS) attack is intentional attack attempted by attackers from a single source making an application unavailable to the target stakeholder. For this to be achieved, attackers usually stagger the network , halting system resources, a denial of for legitimate users. Contrary to DoS attacks, in DDoS attacks, the attacker makes use of multiple sources to initiate an attack. DDoS attacks are most common in a seven-layer OSI model's network, transportation, presentation, and application layers. In this paper, the research objective is to study the problem of DDoS attack environment by considering the most popular CICIDS 2017 benchmark dataset and applying multiple regression analysis for building a machine learning model to predict DDoS and Bot attacks by considering a Friday afternoon traffic logfile.

Keywords: DDoSattack and machine learning; multiple linear regression in DDoS; Cloud computing

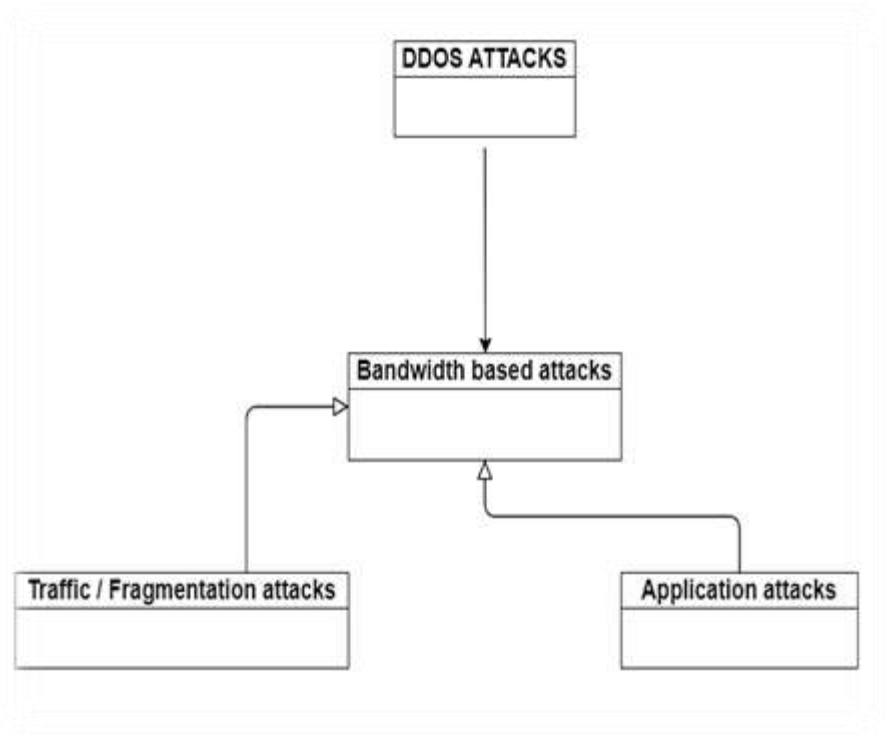
I. INTRODUCTION

Denial of Service attacks are intended primarily to disrupt computing systems in a network. Fundamentally, these attacks are initiated from a machine with the illegitimate intention of targeting a server system through an attack. A simple DoS attack could be a ping Flood attack in which the machine sends ICMP requests to the target server and a more complex DoS attack example could be a Ping of death attack. DDoS attacks are post cursor to DoS attacks, i.e., DoS attacks are a forerunner to DDoS attacks. DDoS attacks are attacks that are carried out in distributed environments. Fundamentally, a DDoS attack is an intentional attack type that is usually made in a distributed computing environment by targeting a website or a server to minimize their normal performance. To achieve this, an attacker uses multiple systems in a network. Now, using these systems, the attacker attacks the target website or server by making multiple requests to the target system or server. As these types of attacks are carried out in distributed environments, hence, these are also called distributed DDoS attacks.

The conventional way of DDoS attacks is the brute force attack that is triggered using Botnet wherein the devices of the network environment are infected with malware.

A Denial of Service attack, which is usually, is a purposeful attempt which is initiated so as to make an application or website unavailable to its legitimate users. This is achieved usually by flooding the website or application through network traffic. In order to achieve this, usually, one of the several choices of attackers is to apply diversified techniques that intentionally consume huge network bandwidth, thus causing inconvenience to legitimate users. Alternately, attackers also achieve this by handling system resources in an illegitimate manner. DoS attack is also called Non-distributed Directed attacker initiates DoS attack on the target system. The concept of DDoS attack is similar to DoS attack but the fundamental difference is that in DDoS attacks there are multiplethe attacker makes an attack by using multiple sources which may include routers, IoT devices, and computers in a distributed environment infected by malware. To make this possible, an attacker looks for availability of any compromised network. compromised networks, an attacker usually attacks through continuously packet floods or requests to the target system. The DDoS attacks Network layer, Transport layer, Presentation and Application layer of the 7-layer OSI reference model. Network layer and Transport layer attacks usually attacks the Presentation layer and as Application layer attacks.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 2153-2156 www.ijaem.net ISSN: 2395-5252



II. RELATED WORKS IN DDOS ATTACK

DDoS attack unexpected traffic on the highway so as to prevent the regular traffic flow arrival at its destination. DDoS attacks are usually performed through the use of a network of connected machines which are all connected via internet. The main problem with these types of attacks is the difficulty in discriminating between normal traffic and attack traffic as each Bot acts as if it is a legitimate one. DDoS attacks not only affect servers in distributed environments Cloud environments. DDoS attacks take the advantage of the services provided by Cloud environment such as (i) pay-as-you-go (ii) auto scaling and

(iii) multi-tenancy. In a typical Cloud infrastructure, Virtual machines (VMs) are run in large numbers by Cloud servers to provide uninterrupted services to legitimate users of the Cloud environment. Now, in an event of an attack by attackers, the server can consider this situation as an event of higher resource utilization. The result would be the server trying to utilize an auto scaling feature in Cloud computing. The result of auto scaling feature could be allocation of resources, and migration of resources so as to solve the server overloading problem. Now, assume that resource allocation and process migration continues as a result of an attack, then, the attacker eventually becomes successful in DDoS attack that was initiated and that such an attack affects either directly or indirectly the Cloud services and eventually implicates the financial revenues. Some of the DDoS attacks in Cloud environment are Buffer overflow, SYN flooding, Ping of death, IP spoofing and land attack. DDoS attack defence in (i) Preventing attacks in the first place, (ii) Detecting the attacks and (iii) Mitigation of attacks. Anomaly detection and Botnet detection techniques are used for preventing and detecting a DDoS attack.

Denial of Service and Distributed Denial of Service attacks are the important sources that make internet services vulnerable. Attack detection is not new in using machine learning techniques. it is learnt that signatures are used for signature-based Intrusion detection system, while detecting unknown attacks are the part of anomaly detection, whereas data flows generated by the unknown patterns gives the scope for studying DDoS. DDoS attacks can be prevented. Attacks are defined as violations of security policies of the network. Usually the attacks are classified into passive and active attacks. Passive attacks never affect the system but active attacks take control of system. The most frequent attacks in real-time network are Denial of Service (DoS) attacks. Distributed Denial of Service attack is multiple systems try to target one single system. By flooding the messages to the target system, the services in the systems are denied and considered as zombies Some of the types of DDoS attacks are Flooding, IP Spoofing,

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 2153-2156 www.ijaem.net ISSN: 2395-5252

TCP SYN Flood, PING Flood, UDP Flood, and Smurf attacks. Multiple machines are used to construct flooding in DDoS attacks.

• TCP SYN Flood attacks—The attack that spoofs the IP addresses is called TCP SYN Flood attack. This attack is more vulnerable as this is based on 3-way handshake protocol

• PING Flood attacks—PING attacks are based on packets of ICMP request. As the PING attack targets the system, the connection slows down and request packets cannot be communicated from the end users.

• UDP Flood attacks—Target system cannot handle authorized once the threshold limit is reached. As the servers reach the threshold limits, the other packet requests are discarded.

• SMURF attacks—This attack occurred because of spoofed PING messages. By pinging the IP address, huge ICMP requests are received, further, more bandwidth will be consumed which slows down the computer to work.

III. NEED FOR DDOS ATTACK DETECTION IN CLOUD:

The attack prevention, detection, and mitigation have received significant importance in relevance to the Cloud computing environment and the problem of DDoS attack detection has received primary importance from researchers. Researchers have been continuously working on proposing various methods and approaches to addressing the detection of DDoS attacks. Despite the availability of contributions that addressed methods and techniques to put a stop to DDoS attacks even today's deployment of the available methods could not resist the DDoS attacks affecting the Cloud.

MultipleLinearRegressionDistributed Denial-of-Service AttackDetectionFramework:

The fundamental research objective behind the proposed method is to design a machine learning model based on multiple linear regression analysis and perform data visualization by considering residual plots and fit charts. Study the possibility of applying multiple linear regression analysis to the CICIDS 2017 dataset which is used in some of the most significant recent research studies. The objective first applies the feature selection technique and determines the important attributes that are better deliverables for the prediction model. We have used the Information Gain approach method for carrying out feature selection. The information Gain approach is a widely applied model in several data mining-based applications. Multiple linear regression analysis and the behavior of chosen and retained important attributes dataset is studied by analyzing the fit charts and residual plots. The next subsection gives the experiment result from analysis using the proposed approach by the most popular CICIDS research dataset.

IV. RESULT ANALYSIS :

The dataset chosen for experimentation consisted of five-day log records from Monday to Friday in CSV format. For experiment analysis, we have considered the log file of Friday afternoon which also consisted of two class labels. The class labels are Benign and attack. The total number of traffic packets in the log file included 225,800 traffic packets.

The attributes in the Friday afternoon logfile are 78 with the last attribute being the class label there are 79 dimensions along with the class label. The modeling process started with the application of the feature algorithm which is based on the computation of information gained for each of the attributes in the dataset. The top 16 attributes have been considered for retention and other attributes in the attribute set are removed. The entire details of the ANOVA model and the list of the top 16 attributes with higher information gain are listed For mathematical modeling, we chose to perform multiple linear regression analysis by running regression analysis on the logfile with these 16 attributes. After initial analysis the attributes that are retained include the attributes at dimensions 1, 5, 6, 7, 9, 11, 13, 35, 36, 53, 54, 55, 56, 64,66 and 67. In performed using the reduced dimensionality log file with these 16 attributes. The mean absolute percentage error for the linear regression model is obtained as 0.2621. The percentage accuracy of the multiple linear regression model is obtained as equal to 80%

V. CONCLUSIONS:

DDOS attack has become more common in a distributed environment like Cloud, it is essential to detect the attacks which cause service unavailability in Cloud Computing. To identify such attacks, machine learning models can be used to train and test the attack detection datasets. Alternately, we can use the regression analysis technique by applying one of its important variants known as multiple linear regression analysis. The research objective behind this study is to build a machine learning model that is an ensemble of feature selection using information gain and regression analysis. The dataset considered was in the popularly known CICIDS 2017 dataset. It has been observed that through this ensemble model for the Friday morning dataset, prediction accuracy of

95% is achieved. Similarly, for the Friday afternoon log file, the prediction accuracy is obtained as 80% for 16 attributes obtained through information feature selection and regression analysis-based ML model. This paper the importance of regression analysis in building an ML model and also shows some of the important visualizations. In this work, we have limited our analysis to the one-day log file and in the future, this research may be extended to consider all traffic log files of five days and come out with a consensus-based machine learning model.

REFERENCES:

[1]. CIPD. 2020. Digital Learning | Factsheets | CIPD. online Available at: -factsheet
[2]. Accessed 21 August 2020
[3]. Radware'sDDoS Handbook: The Ultimate Guide to Everything You Need to Know About DDoS Attacks
[4]. DDoS A Complete Guide - 2021 Edition
[5]. Jia, B.; Huang, X.; Liu, R.; Ma, Y. A DDoS Attack Detection Method Based on Hybrid
[6]. Sharma, N.; Mahajan, A.; Mansotra, V. Machine Learning Techniques Used in Detection of DOS
[7]. Introduction to DDoS Attacks and Defense Mechanisms Gupta B
[8]. Detecting DDoS attacks against data centers with correlation analysis.