Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 15 Issue 6 Version 1.0 Year 2015
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Effective Detection and Prevention of Ddos based on Big Data-Mapreduce

By Sumathi Rani Manukonda & Dr. Koppula Srinivas Rao

CMR College of Engineering & Technology, India

Abstract- Distributed Denial of Service (DDoS) attacks is large-scale cooperative attacks launched from a large number of compromised hosts called Zombies are a major threat to Internet services. As the serious damage caused by DDoS attacks increases, the rapid detection and the proper response mechanisms are urgent. However, existing security methodologies do not provide effective defense against these attacks, or the defense capability of some mechanisms is only limited to specific DDoS attacks. Therefore, keeping this problem in view author presents various significant areas where data mining techniques seem to be a strong candidate for detecting and preventing DDoS attack. The new proposed methodology can perform detecting and preventing DDoS attack using MapReduce concepts in Big Data.Thus the methodology can implement for both detecting and preventing methodologies.

Keywords: *ddos attacks, data mining-big data, mapreduce.*

GJCST-C Classification : *C.2.4 E.2*

Strictly as per the compliance and regulations of:

Effective Detection and Prevention of DDoS based on Big Data-Mapreduce

Sumathi Rani Manukonda[α] & Dr. Koppula Srinivas Rao[σ]

Abstract- Distributed Denial of Service (DDoS) attacks is large-scale cooperative attacks launched from a large number of compromised hosts called Zombies are a major threat to Internet services. As the serious damage caused by DDoS attacks increases, the rapid detection and the proper response mechanisms are urgent. However, existing security methodologies do not provide effective defense against these attacks, or the defense capability of some mechanisms is only limited to specific DDoS attacks. Therefore, keeping this problem in view author presents various significant areas where data mining techniques seem to be a strong candidate for detecting and preventing DDoS attack. The new proposed methodology can perform detecting and preventing DDoS attack using MapReduce concepts in Big Data.Thus the methodology can implement for both detecting and preventing methodologies.

Keywords: DDoS attacks, data mining-big data, mapreduce.

I. Introduction

Today, the number of attacks against large computer systems or networks is growing at a rapid pace. One of the major threats to cyber security is Distributed Denial-of-Service (DDoS) attack .In which the victim network element(s) are bombarded with high volume of fictitious attacking packets originated from a large number of Zombies. In the modern computer world, maintaining the information is very difficult. Some interrupts may occur on the local system (attack) or network based systems (network attack). The aim of the attack is to overload the victim and render it incapable of performing normal transactions.In this paper a new cracking algorithm is implemented to stop that DDOS attacks. This is most critical attack for a network called distributed denial of service attack.To protect network servers, network routers and client hosts from becoming the handlers, Zombies and victims of distributed denial-of-service (DDoS) attacks data mining approach can be adopted as a sure shot weapon to these attacks. Without security measures and controls in place, our data might be subjected to an attack. In our algorithmic design a practical DDOS defense system that can protect the availability of web services during severe DDOS attacks. One common method of attack involves sending an enormous amount of request to the server or site and server will be unable to handle the requests and site will be offline for some days or some years depends upon the attack.

The proposed system identifies whether the number of entries of client exceeds more than five times to the same sever, then the client will be saved as a attacker in blocked list and the service could not be provided. So our algorithm protects legitimate traffic from a huge volume of DDOS traffic when an attack occurs.This approach uses the automatic feature selection mechanism for selecting the important attributes using the Map Reduce methodology. And the classifier is built with the theoretically selected attribute through the neural network. And then, our experimental results show that our approach can provide the best performance on the real network, in comparison with that by heuristic feature selection and any other single data mining approaches.

II. Overview

a) DDoS ATTACK

Distributed Denial-of-Service (DDoS) attack is the one in which the victim's network elements are bombarded with high volume of fictitious attacking packets that originate from a large number of machines.

In Recruiting phase attacker initiates the attack from the master computer and tries to find some slave computers to be involved in the attack. The number of DDoS attacks grew 20 % last year - a major decrease in the rate of attacks from 2007 to 2008, when these devastating attacks increased 67 percent, according to a report.1 According to a report Internet Service Providers (ISPs) are most worried about botnet-driven distributed denial-of-service (DDoS) attacks2 .A small piece of software is installed on the Zombies to run the attacker commands. The Action phase continued through a command issued from the attacker resides on the master computer toward the Zombies computers to run their pieces of software. A successful attack allows the attacker to gain access to the victim's machine, allowing stealing of sensitive internal data and possibly cause disruption and denial of service (DoS) in some cases. The mission of the piece of software is to send dummy traffic designated toward the victim. Therefore, a mechanism that is strong and reliable is desired. Hence

Author α : Asst.Prof, Department of CSE, CMR College of Engineering & Technology Hyderabad, Telangana State, INDIA.
e-mail: ksreenu2k@yahoo.com
Author σ: Professor & Head of Dept. CSE, CMR Collge of Engineering & Technology, Hyderabad, Telangana State, India.
e-mail: sumathibabum@gmail.com

the key idea is to use data mining techniques to discover consistent and useful patterns of system features that describe program and user behavior of attack.

III. Data Mining Data

Data mining is becoming a persistent technology in activities as diverse as using historical data to predict the success of a marketing campaign, looking for patterns in network traffic to discover illegal activities or analyzing sequences. A data mining application is typically a software interface which interacts with a large database containing Network traffic parameters or other important data. From this outlook, the approach is gaining importance in the field of DDoS attacks. Data mining applications are computer software programs or packages that enable the extraction and identification of patterns from stored data.

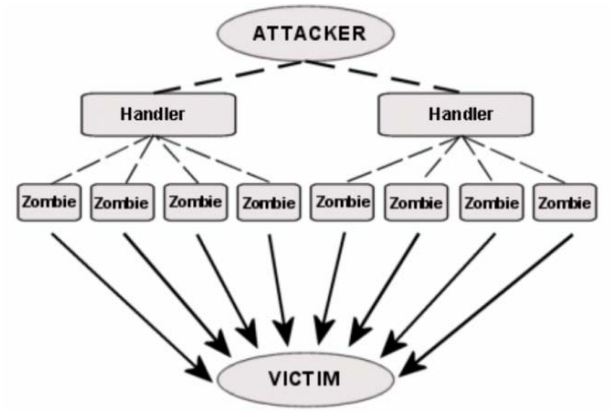

Figure : Architecture of DDoS attack

Data mining is, at its core, pattern finding. Data miners are proficient at using specialized software to find regularity (and irregularities) in large & complex data sets. Data mining is widely used by companies and public bodies for marketing, detection of fraudulent activities such as DDoS attacks. Huawei 2013 Security Research Report 29.81% more DDoS attacks occurred than last year, More than 72.91% attacks larger than 1Gbps. Http application protocols Attacked up to 87.74% and Longest DDoS attacks last 349 hours 36 minutes 42 seconds.

IV. Various Application Areas of Data Mining in Ddos Attacks

Recently, data mining has become an important component for DDoS attack prevention. An overview of real time data mining-based intrusion detection systems (IDSs) is presented by researcher that focused on problems related to deploying a data mining-based IDS in a real time environment also discussed a distributed architecture for estimating cost-sensitive models in real time.

Different data mining approaches like classification, association rule, clustering, and outlier detection are the few techniques frequently used to analyze network traffic or data to gain knowledge that helps in controlling intrusion Various applications where data mining approach can be used in prevention and detection of DDoS attacks are discuss below:

V. Implementation

a) Big Data-Map Reduce

i. *Map step*

The average output of the map will be recorded ID as the key and retired as the value. Every mapper maintains a collection bearing the canopy center candidates it has learned thus far. During every map the mapper determines if each successive record is within the distance threshold of any already determined canopy center candidate. The intermediate output sent to the reducer has the record ID as the cay and the list of retired-rating pairs as the value.

ii. *Reduce Step*

The yield of the reduce step will simply output record ID as the key and concatenate the rater IDs for that record into a comma separated list. The reducer repeats the same procedure as the mappers. It meets the candidate canopy center record IDs, but takes out those which are inside the same threshold limit . In other words, it removes duplicate candidates for the same canopy center. In order for this to operate correctly the number of reducers is set to one.

iii. *DDoS attack impact on Big Data*

As the Internet continues to grow and prosper, hacker attacks continue to increase in severity and frequency. As Internet bandwidth has expanded, so too has the scale and frequency of DDoS attacks. For example, in March 2013, European anti-spam company Spamhaus experienced multiple 300 Gbit/s DDoS attacks, the largest such attacks in history. The real-time example, Huawei is the first anti-DDoS solutions provider to apply Big Data technology to DDoS detection and prevention. Huawei leads the industry in eliminating covert DDoS attacks disguised as normal access requests.

iv. *DDoS Trends Challenging Attack, Defense Technologies*

Typically, DDoS attacks originate from mock sources, such as Synchronize (SYN) flood, User Datagram Protocol (UDP) flood, and Domain Name Service (DNS) flood, and are carried out by zombie hosts. When DNS servers are paralyzed a wide range of network services will be blocked or broken.The more bandwidth the attack consumes, the bigger the threat to network infrastructure. Such attacks that target specific applications, such as HTTP Flood attacks against e-

commerce websites and web games, require a TCP connection between the zombie host and servers targeted for attack. To avoid detection, hackers reduce the attack traffic rate so that the attack footprint resembles a legitimate request.

VI. Framework & Algorithm

These heavy-traffic DDoS attacks are the easiest to detect, but require the highest processing performance to affect the necessary rapid response; otherwise, the network links will become jammed, completely flooded, while security devices deployed on the access side are failing. Until the recent arrival of cost-effective flow analysis technology, these super-large-bandwidth DDoS attacks were best handled by commercial anti-DDoS SPs. Blocking such attacks requires the deployment of super-large capacity prevention systems on the upstream side of the network. Effective enterprise anti-DDoS systems must be based on high-performance hardware platforms with a minimum 100-Gbit/s defense capacity, or the defense device itself will likely become the network bottleneck. We have now entered the era where these high performance tools are now available for enterprises.

Because DDoS attack detection systems rely on traffic models for attack detection, the better the traffic model the higher the probability of detecting attacks. The difficulty in detecting light-traffic attacks is that the small numbers of attack packets are concealed in massive volume of legitimate network access packets. A further challenge to detecting low volume DDoS attacks is that application layer attacks strongly resemble legitimate access requests, and that even with increased sampling rates, flow analysis is unsuitable for detecting application layer attacks because QPS analytics are not included in the access traffic model. Mitigating this type of attack using traditional prevention systems can only limit the connections of legitimate access sources.

a) DDoS Defense Technology Based on Big Data -Map Reduce

As an industry-leading anti-DDoS solutions provider, Huawei is the first vendor to apply Big Data technology to the detection and prevention of covert DDoS attacks disguised as normal access requests.

These heavy-traffic DDoS attacks are the easiest to detect, but require the highest processing performance to affect the necessary rapid response; otherwise, the network links will become jammed, completely flooded, while security devices deployed on the access side are failing. Until the recent arrival of cost-effective flow analysis technology, these super-large-bandwidth DDoS attacks were best handled by commercial anti-DDoS SPs. Blocking such attacks requires the deployment of super-large capacity prevention systems on the upstream side of the network. Effective enterprise anti-DDoS systems must be based on high-performance hardware platforms with a minimum 100-Gbit/s defense capacity, or the defense device itself will likely become the network bottleneck. We have now entered the era where these high performance tools are now available for enterprises.

Because DDoS attack detection systems rely on traffic models for attack detection, the better the traffic model the higher the probability of detecting attacks. The difficulty in detecting light-traffic attacks is that the small numbers of attack packets are concealed in massive volume of legitimate network access packets. A further challenge to detecting low volume DDoS attacks is that application layer attacks strongly resemble legitimate access requests, and that even with increased sampling rates, flow analysis is unsuitable for detecting application layer attacks because QPS analytics are not included in the access traffic model. Mitigating this type of attack using traditional prevention systems can only limit the connections of legitimate access sources.

b) DDoS Defense Technology Based on Big Data -Map Reduce

As an industry-leading anti-DDoS solutions provider, Huawei is the first vendor to apply Big Data technology to the detection and prevention of covert DDoS attacks disguised as normal access requests.

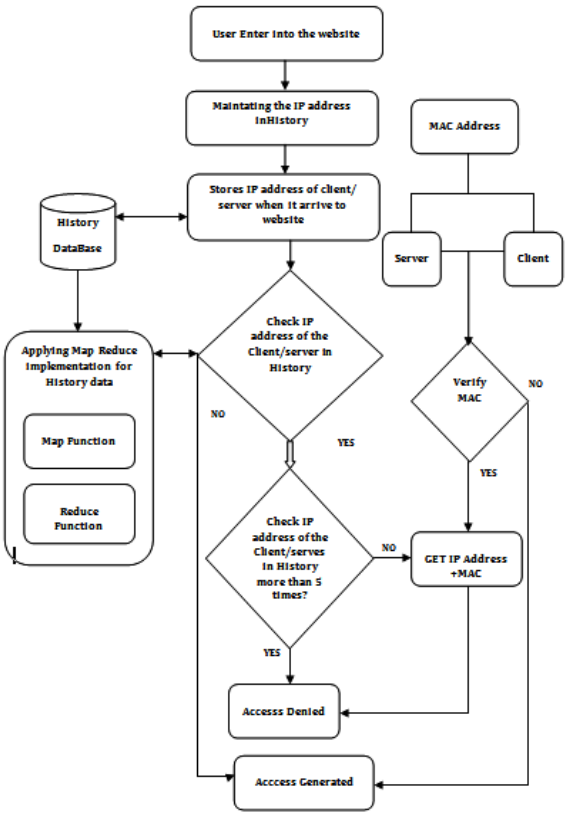

Figure : Framework

These heavy-traffic DDoS attacks are the easiest to detect, but require the highest processing performance to affect the necessary rapid response; otherwise, the network links will become jammed, completely flooded, while security devices deployed on the access side are failing. Until the recent arrival of cost-effective flow analysis technology, these super-large-bandwidth DDoS attacks were best handled by commercial anti-DDoS SPs. We have now entered the era where these high performance tools are now available for enterprises.

Because DDoS attack detection systems rely on traffic models for attack detection, the better the traffic model the higher the probability of detecting attacks. The difficulty in detecting light-traffic attacks is that the small numbers of attack packets are concealed in massive volume of legitimate network access packets. A further challenge to detecting low volume DDoS attacks is that application layer attacks strongly resemble legitimate access requests, and that even with increased sampling rates, flow analysis is unsuitable for detecting application layer attacks because QPS analytics are not included in the access traffic model.

c) DDoS Defense Technology Based on Big Data -Map Reduce

As an industry-leading anti-DDoS solutions provider, Huawei is the first vendor to apply Big Data technology to the detection and prevention of covert DDoS attacks disguised as normal access requests.

New algorithm using the Map Reduce Methodology
H=Maintaining IP address as History;
U=User enter into website;
I=Storing Each Client IP address;
N=New IP address;
MACM =Mapped MAC/IP Address
MACR =Reduced MAC/IP Address
MACN =New MAC/IP Address

Start the Process
Check each time U in server, If ((I==H)==N) { Else If (I<5)
{
IP=Get the IP address;
MaP((I value & H)U), /(Applying MaP FUNCTION to I&H*
*with User web address / IP Adress) */*
{
FOR each{ U,I value}in {(I,H value) },//mapping (split) key values/ pairs
do
MACM=(U,I value) ∪H
Return MAC;
}
REDUCE((I value & H)U,MACM), /(Applying REDUCE FUNCTION to I*
*value of User in H) */*
{
FOR each {(I value ,U) €H ,MACM},//reducing (merg) key values/ pairs
do
MACR=(U,I value) ∩H
}
MACN 1=IP+MACR // Read Previous MAC
Algorithm Server=MAC1;
Client=MAC1;
If (Server=Client)
{
Accept the request from the client Send the response for the request.
}
Else
{
Add the User.IP to the Attacker List,
Print: "Access Denied"
} } }
Else
{
Accept the request from the IP Send the response for the request.
}
End

d) Challenges

- *High security:* The security solution must be able to defend against DDoS attacks of various types, regardless of the traffic attacks or application-layer attacks, to protect all online services from attacks.
- *High performance:* To avoid being the bottleneck of the whole system, the security solution must feature high-performance defense capabilities so that it can deal with the traffic flooding attacks on Tencent's large-scale services.
- *High scalability:* The security solution must support flexible performance expansion to vary with service requirement changes, catch up with service mode innovation, and form an architecture required for long-term service development, in order to protect previous investment and reduce total investment cost..
- *High availability:* The security solution must ensure reliable service connections, precisely differentiate attack traffic from normal traffic, and accurately identify attacks
- *Low O&M cost:* Considering that O&M cost significantly affects Tencent, the security solution must be small-sized, consume low power, minimize occupied equipment room space and consumption with improved performance, and greatly reduce the TCO for deploying multiple nodes in batches.

VII. Conclusion

DDoS attacks are quite complex methods of attacking an ISP in large data. These attacks are an aggravation at a minimum, and if they are against a particular system, they can be brutally destroyed. This paper discussed various detection algorithms which are using data mining concepts &algorithms for DDoS detection & prevention. The proposed methodology we can detect the attack easily by implementing Map Reduce functionalities at the stage of verifying the Network IPs. Loss of network resources, costs money, delays work, and interrupts communication between various legal network users , thus it can prevent all these complexities effectively. Detecting, preventing, and mitigating DDoS attacks is important for national and individual security. But with the improvement in technology new areas are emerging where data mining techniques can be utilized for handling DDoS attacks that are to be discussed in future.

VIII. Acknowledgment

We would like to thank everyone who has motivated and supported us for preparing this manuscript.

References Références Referencias

1. J. Mirkovic and P. Reiher, "A Taxonomy of DDoS Attack and DDoS Defense Mechanisms," ACM SIGCOMM Computer Communications Review, Volume 34, Number 2, April 2004, pp. 39-53.
2. Rui Zhong, and Guangxue Yue "DDoS Detection System Based on Data Mining" ISBN 978-952-5726-09-1 (Print) Proceedings of the Second International Symposium on Networking and Network Security (ISNNS '10)Jinggangshan, P. R. China, 2-4, April.2010, pp. 062-065
3. Bremler-Barr and H. Levy. Spooling prevention method. In Proc. IEEE INFOCOM, Miami, FL, March 2005. BRUTLAG, J. D.
4. Aberrant Behavior Detection in Time Series for Network Monitoring. In Proc. of USENIX LISA (2000).
5. C.SIRIS and F. PAPAGALOU. Application of anomaly detection algorithms for detecting syn flooding attacks. In Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '04), volume 4, pages 2050–2054, Dallas, USA, 2004
6. Douligeris and A. Mitrokotsa. DDoS attacks and defense mechanisms: classification and state of-the-art. Computer Networks, 44(5): 643-666, 2004
7. Huang Kai, Qi Zhengwei, Liu Bo" Network Anomaly Detection Based on Statistical Approach and Time Series Analysis" 2009 International Conference on Advanced Information Networking and Applications Workshops.
8. O. Salem, S. Vaton, and A. Gravey. An efficient online anomalies detection mechanism for high-speed networks. In IEEE Workshop on Monitoring, Attack Detection and Migitation (MonAM 2007), November 2007.
9. Dietrich, S., Long, N., and Dittrich, D. 2000. Analyzing distributed denial of service attack tools: The shaft case. In Proceedings of 14th Systems Administration Conference. New Orleans, Louisiana, USA, 329-339.
10. Wenke Lee, Salvatore J. Stolfo , Philip K. Chan Eleazar Eskin , Wei Fan , MatthewMiller , Shlomo Hershkop and Junxin Zhang "Real Time Data Mining-based Intrusion Detection "2002, ieexlore. ieee.org.

This page is intentionally left blank