International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Anomaly Detection in Distributed Denial of Service Attack using Map Reduce Improvised Counter Based Algorithm in Hadoop

Y.S Kalai Vani, P. Ranjana, M. Sankari

Abstract : A Distributed denial of Service attacks(DDoS) is one of the major threats in the cyber network and it attacks the computers flooded with the Users Data Gram packet. These types of attacks causes major problem in the network in the form of crashing the system with large volume of traffic to attack the victim and make the victim idle in which not responding the requests. To detect this DDOS attack traditional intrusion detection system is not suitable to handle huge volume of data. Hadoop is a frame work which handles huge volume of data and is used to process the data to find any malicious activity in the data. In this research paper anomaly detection technique is implemented in Map Reduce Algorithm which detects the unusual pattern of data in the network traffic. To design a proposed model, Map Reduce platform is used to hold the improvised algorithm which detects the (DDoS) attacks by filtering and sorting the network traffic and detects the unusual pattern from the network. Improvised Map reduce algorithm is implemented with Map Reduce functionalities at the stage of verifying the network IPS. This Proposed algorithm focuses on the UDP flooding attack using Anomaly based Intrusion detection system technique which detects kind of pattern and flow of packets in the node is more than the threshold and also identifies the source code causing UDP Flood Attack.

Keywords: Anomaly detection, Denial of service, Hadoop, Mapreduce.

I. INTRODUCTION

A Distributed Denial of Service (DDoS) attack is malicious attempt to make victim machine or web server and network resource unavailable to intended users by temporarily disrupting the services of a host machine or network resource. In earlier approach to hold the data flow is difficult because it has less capacity to store the data in memory to overcome the problem in this earlier approach Big data analytics is used to hold the huge amount of data.

In this approach is to weaken the victim's system by exhausting the resources such as input-output bandwidth, database bandwidth, CPU, memory.etc. DDOS attacks are classified into different types such as HTTP flood attack, UDP flood attack, ping of death attack

In this research paper we have taken a type of attack UDP flooding attack which comes under the classification of distributed of denial of service. UDP flooding attack can be detected by a improvised counter based algorithm in Hadoop. Threshold value is set for the UDP Packets, if UDP packets are exceeds the Threshold values then the UDP attack will be detected. In a UDP flood, the attackers send highly-hoaxed UDP(user datagram protocol) packet at a very high packet rate using a large Source IP range. Because of the UDP attack the victims network such as routers, firewalls, severs are exhausted by a large number of incoming UDP packets. This attack normally consumes network resources and available bandwidth, exhausting the network until it goes offline. UDP flooding attacks are very difficult to detect and blocking the network and it block the resources of the victim.

CLASSIFICATON OF DDOS ATTACK

A. Infrastructure attack

Infrastructure attack consists of Network bandwidth, routing equipment and computing resources. In this type of attack the attacker to overwhelm the resource capacity of node in a network by sending a large number of fake requests. Examples for this Infrastructure attacks are TCP/SYN flood,UDP flood, ICMP Flood etc.

TCP SYN flood:

This attack is caused by an attacker sends a lot of ordinary SYN segments to fill up resources causing a service to be denied for its connections.

UDP flood

In this attack huge amount of UDP packets are sent to random ports on the victims side . sometimes port are open without knowledge of administrator causing server to respond. A respond to a UDP Packet with an ICMP unreachable reply to the spoofed source IPaddress makes the situation worse by overwhelming network environment of the victimized IP address.

ICMP flood:

This attack is referred as a Smurf attack or ping of flood, is a ping based Dos attack that ends large numbers of ICMP packets to a server and attempts to crash TCP/IP Server.

B .Application level attack :

This type of attack which attacks the resources of the server by sending the HTTP request through the network. It has two categories which are given below

Manuscript published on November 30, 2019.
* Correspondence Author

Y.S.Kalai Vani MCA M,Phil, **(P**hD), Associate in Sindhi College, Banaglore and doing research in Hinsdustan Institute of Technology, Chennai, India

Dr. P.Ranjana, working **as a** Professor in Department of Computer Science in Hindustan Institute of Technology, Chennai.

M.Sankari ,MTech, (PhD), Research Scholar, CSE Hindustan Institute of Technology and Science Chennai,India.

Common application layer attack:

In this attack the attacker sends the normal request which consumes large amount of server resources or high work load requests across many TCP sessions are sent to the server,so it will cause common application layer attack.

HTTP flood attacks:

Some applications level DDoS are caused by HTTP GET flood HTTP GET flood attack caused by an attacker who sends the large amount of request in which it consumes a large amount of resources of a server because of HTTP GET packets are flooded and it consumes a resources of a server which server not able to process the requests because of HTTP Flooding attack.

II. RELATED WORK

The analysis of malicious network packets and processing of logs for threat detection has been a tedious task for years as attackers are changing their methods and tactics in launching DDoS attacks. In recent trends of Big Data Analytics Apache Hadoop and its ecosystem has attracted network security community because of the scalability, simplicity and fault tolerance features [06]

. The counter based method counts total traffic received or web page request from clients. Based on the threshold value specified, the server is alarmed. The process of access pattern method makes an assumption that the clients infected by the same protocol conduct similar behavior so that the attacker can be differentiated from normal client. But the proposed framework supports offline batch processing of traffic traces only.[22]

Sufian and Usman [24] proposed HADEC, a Hadoop based Live DDoS Detection framework to handle flooding attacks efficiently using MapReduce. They have implemented counter based DDoS detection algorithm to detect TCP SYN, HTTP GET, UDP and ICMP attacks only. A Good feature is important which can be derived using various methods for training the data and classification purposes.

Ganapathy et al. [26] specified a method wherein they used an intelligent rule-based attribute selection algorithm to determine the feature set. Seo et al. [27] propose another way of determining the feature set using a clustering-based method.

III. INTRUSION DETECTION APPROACHES:

An Intrusion detection system(IDS) is tool which is used to monitor the network for malicious activities in the network and it detects the attacks from the anomalous packets. Intrusion detection system is classified in to three types are:

Host based Intrusion detection System:

It is an IDS which monitors the internal part of the computer system and monitors the network packets in the network interface.

Hybrid Intrusion detection System:

It is an IDS which monitors the computer’s internal activities and network activities.

Network Intrusion Detection System:

It is an IDS that attempts to discover the unauthorized access to a computer network and detect the cyber threats in the network. It is classified into two types

i)Anomaly Detection system.

This type of approach analyses the network activities and looks for an unusual behavior in the network, if an anomaly found in the system then alarm is triggered.

ii) Misuse detection system

This approach is used to detect the known patterns of attacks. If the pattern is matched there Is possibility of an attack that alarm will be triggered. It is used to detect the attacks which are in the variations in known attacks.

To detect DDoS attack in the network sophisticated approach is needed which resolves the problem of finding the attack in huge volume of data. To handle this situation and solve the problem pervious model in DDoS attack is improvised. Hadoop is a platform which handles a huge volume of data and it is used to store the data in the form of zetabyte. Existing System not able to solve the problem of finding UDP flooding attack in a efficient manner. In this research paper the Map Reduce Improvised Counter Based algorithm is used to detect the UDP flooding attack.

IV. HADDOOP - PROPOSED SYSTEM FRAMEWORK

Hadoop is a proposed system frame work which handles huge volume of data with efficiency. Traditional approach is not suitable to handle the UDP packets because it can hold very less volume of data. Map Reduce is a programming model in Hadoop which is used to process distributed data. It consists of many cluster of machines in the distributed form.[2] It consists of two phases such as Map phase and Reduce phase. In map phase packets coming from different clusters and it are separated as a UDP Packets. It has the key as <source IP, UDP packets> in the map phase. In Reduce phase the filtered UDP packets are entered based on the Threshold value. These splits are then parallel processed by the mappers. In the Reduce Phase the intermediate results provided by the map phase are summarized and associated records are processed by single reducer.

The data set DARPA which is used for the experiment to detect the UDP flooding attack. The master distributes different sets of data among different mappers and the intermediate results are stored at the Mappers. The master then assigns the task of extracting information regarding the attack data to the reducers.

In Reduce phase has detection algorithm which is used to detect[4] the UDP Flooding attack based on the threshold value. If UDP packets are exceeding the Threshold value then the attack will be detected in the network and stores the IP address of the attackers also.

V. IMPLEMENTATION

The following algorithm which is improvised by Map Reduce in Hadoop. Hadoop is a platform which accepts huge number of data from the node and this algorithm is used to detect the unusual behavior of the packet in the network. We set the Hadoop environment which has Map[5] Reduce programming model which has two stages. In first stage this algorithm which separates the UDP packets from the network.

In second stage this algorithm detects the UDP flood attack which contains Threshold value as 1000000 for UDP packets and if the UDP exceeds the Threshold value then UDP flood is detected along with the source IP address.

Algorithm to Detect UDP Flood Attack(MRCIB)

```
STEP 1:   Start[map function]
STEP 2:   Identify nodes in network.
STEP 3:   If nodes in range(belongs to the same network)
STEP 4:   Set
THRESHOLD=1000000
THRESHOLD=UDP-PACKETS/SECOND
    MINIMUM=500
MAXIMUM=1000000
DEFAULT=10000
STEP 5:   CAPTURE packets.
STEP 6:   If packet not of standard type[Reduce function]
            Notify Malformed_ packet.
STEP 7:   Identify UDP_PACKET.
STEP 8:   If UDP_PACKET > THRESHOLD
            Notify UDP_FLOOD_ATTACK
STEP 9: Stop
```

Algorithm implementation in Java

The MRICB algorithm is implemented in java. Hadoop is Java based platform which needs to face the challenges like efficiency, accuracy, security. Since Java meets all the challenges because of its enhanced features. Java has a set of packages which will solve all requirements from the user side. The package Remote method invocation is used to make the establishment between the nodes in the network. There are different types of nodes are available in the network which want to communicate with server. It has the interface which has the method detection which has the implementation part of UDP flood detection. It accepts the inputs as UDP packet from the port 80 and checks the condition that it should not exceed the threshold 100000 as the maximum. If it exceeds the exception will throw as UDP Flood attack detection.

Code for MRICB algorithm in java

```
import java.rmi.*;
public interface DDOSservice extends Remote
{
//Interface of RMI service that will actually attack  on a target machine
public String attackdetection()
{
 throw RemoteException;
}
}
```

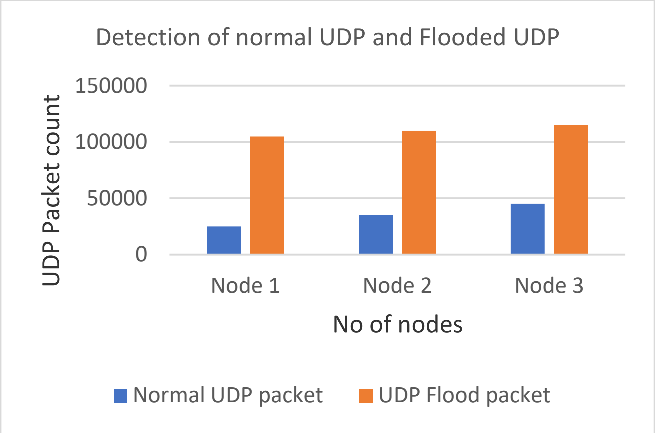
Server program

```
Public class DDOSServer extends UniCastRemoteObject implmenets  Runnable DDOSservice
{
   Final String Target = “xyz”;
 Static DDOSServiceServer_instance;
Public  DDOSServiceServer() throws RemoteException
{
 super();
}
Public void run()
{
//fix the Threshold value as 100000

/if it exceeds the Threshold identify the UDP flood exception
//Accepting the packets from the client
 For(int UDP=0;UDP<=1000000;i++)
{
try{
   Socket net = new Socket(Target,80);
sendRawLine(“GET/HTTP/1.1”,net);
sendRawline(“Host:”+Target,net);
System.out.println(“Attacking the”+Target”+” with connection”+UDP);
}catch(Exception e)
{
System.out.println(“UDP flood attack);
}
}
```

VI. RESULTS

Results of the MRICB(Map Reduce Improvised Counter Based) algorithm has as a graph which consists of two ranges of UDP packet such less than the Threshold value and more than the Threshold value. If it exceeds the threshold value UDP flood is detected otherwise the nodes will communicate normally. The map and reduce function has the methodology which has the two phases of separating the UDP packets and testing the threshold value.



VII. CONCLUSION & FUTURE WORK

This paper proposed a Hadoop based robust and efficient DDoS Detection model which has the mechanism to detect the UDP flooding attack by measuring the Threshold values for the nodes. MRICB algorithm implementation values are plotted in the graph shows the difference between normal UDP packet and flooded UDP. For the future work we plan to optimize the Map Reduce jobs, to enhance the model high end speed links and to enhance pattern matching algorithm for detecting the packets.

REFERENCES

1. Hall, London. Brown, S.L. and Eisenhardt, K.M. (1998) *Competing on the Edge: Strategy as Structured Chaos*, Harvard Business School Press, Boston.
2. Clutterbuck, D. and Crainer, S. (1990) *Makers of Management: Men and Women who Changed the Business World*, MacMillan, London.
3. Dolan, S.L., Garcia, S. and Auerbach, A. (2003) "Understanding and Managing Chaos in Organisations", *International Journal of Management*, Vol 20, No. 1, pp 23–35.
4. Evans, D. (1998) The arbitrary ape, *New Scientist*, Vol 159, No. 2148, 22 August, pp 32–35.
5. Farrell, W. (1998) *How Hits Happen: Forecasting Predictability in a Chaotic Marketplace*, Harper Business, New York. x
6. Fitzgerald, L.A. and van Eijnatten, F.M. (1998) "Letting Go For Control: The Art of Managing the Chaordic Enterprise", *The International Journal of Business Transformation*, Vol. 1, No. 4, April, pp 261-270. Goldberg, J. and Markoczy, L. (1998) "Complex Rhetoric and Simple Games", [online], Cranfield University, www.Cranfield.ac.za/public/cc/cc047/papers/complex/html /complex.htm.
7. McElwee, M. (1998) "Chaos Theory and Complexity as Fountainheads for Design of an Organization Theory Building Workshop", Paper read at XIVth World Congress of the International Sociological Association, Montreal, Canada, July.
8. Tom White, "Hadoop: The Definitive Guide", O‟Reilly Media, 2012
9. [22]Yeonhee Lee and Youngseok Lee, "Detecting DDoS attacks with Hadoop", In Proceedings of The ACM CoNEXT Student Workshop, pp. 1-2, 2011
10. [26]S. Ganapathy, K. Kulothungan, S. Muthuraj kumar, M. Vijayalakshmi, P. Yogesh, and A. Kannan, "Intelligent feature selection and classification techniques for intrusion detection in networks: a survey", EURASIP Journal on Wireless Communications and Networking, vol. 1, 2013. [27]J. Seo, J. Kim, J. Moon, B. J. Kang, and E. G. In, "Clustering based Feature Selection for Internet Attack Defense", International Journal of Future Generation Communication and Networking, vol. 1, pp. 91-98, 2008.

AUTHORS PROFILE

Y.S.Kalai Vani MCA M,Phil, (PhD), working as a Associate in Sindhi College, Banaglore and doing research in Hinsdustan Institute of Technology, Chennai, India Published more than eight research papers and membership in computer society of India.

Dr. P.Ranjana, working as a Professor in Department of Computer Science in Hindustan Institute of Technology, Chennai. Guiding more than ten PhD scholars and published more than twenty research papers in reputed journals,.

M.Sankari ,MTech, (PhD), Research Scholar, CSE Hindustan Institute of Technology and Science Chennai,India. Published more than eight research papers in reputed journal.