Asian Journal of Information Technology 15 (24): 5193-5196, 2016
ISSN: 1682-3915

Detection of DDoS Attack by Monitoring Network Traffic Using Z-Test Based Statistical Analysis

[1]K. Ganesh Kumar and [2]N. Rengarajan
[1]Department of IT, K. S. R. College of Engineering, Tiruchengode, India
[2]Nandha Engineering College, Erode, India

Abstract: Distributed Denial of Service (DDoS) describes the unavailability of service provided by the server in the network to its valuable clients in the network. Many researchers focus on the problems in network resources availability and its allocation to the clients without any time delay and in an efficient manner. The main task in the process includes the attack detection process which helps to identify the deviation of packet flow from the normal traffic. In this study, the detection of the DDoS attack is carried out with the help of the network capacity (called as router capacity) in processing and forwarding of the information packets. The packet flow is analyzed with the help of Z-test to check whether there is an abnormal traffic occurred in the network by focusing on the data arrival rate at different time intervals. Based on the result, the information packets are analyzed and check with the threshold value for forwarding processing by the router.

Key words: DDoS attack, attack detection, threshold value, router capacity, Z-test

INTRODUCTION

Denial of service attack describes the event that makes the user to pay their attention towards it. An event helps to prevent the authorized user from utilizing the available resources in the network is termed as denial of Service. Denying the permission for accessing the resources in any form comes under this attack. Many organizations face most of these kinds of problems. These types of attacks are performed by the system as an individual or in a group. Individual System can carry out the attack for the prevention of accessibility is called as Denial of Service attack (DoS) and group/many numbers of compromised systems involved is termed as Distributed Denial of Service attack (DDoS).

The need of study of the DdoS attack is most of the latest attacks are categorized under DDoS attack type (Braga *et al*., 2010). More number of systems present in the network utilizes the resources and due to the necessity of doorstep resource availability provision, every system tries to makes them to achieve its desired task by accessing the maximum bandwidth. As a result of this the accessing of the bandwidth resource of another system in the network may happen. It is the first step for the attack process. Attack happened in the system to be identified in order to provide efficient service to the intended user present in the network. To achieve the maximum efficiency and resource utilization, the network should be monitored periodically (Fig. 1).

Once the attack is identified and processed it is necessary to prevent the loss of the network resources and block its availability to the attackers. Detection of attack process (Leu and Lin, 2010) involves many phases such as analysis of network traffic in the router (Kumarasamy and Asokan, 2011), threshold value fixation, testing hypothesis, etc.

Literature review: The security measures available in the present system includes the processing of the incoming request packet, analyzing the sender information, checking the availability of sender name in the blacklist, etc. but there are some hackers in the internet likes to share the information stored by the others. It leads to the loss and modification of the information of the trusted user and it may block the necessary services provided by the system to its users. Loss and modification of the information should be blocked by increasing the security measures using the identity of the authorized users in the system.

Routing devices plays a main role in the security providing process. The packets arrived to the router are analyzed thoroughly in a manner such that the packet is free from untrusted user threats. Periodic analysis of the packet flow in the network is carried out to check the consistency of the network traffic. The detection of DDoS attack is the base for identifying the system performance.

Corresponding Author: K. Ganesh Kumar, Department of IT, K.S.R. College of Engineering, Tiruchengode, India

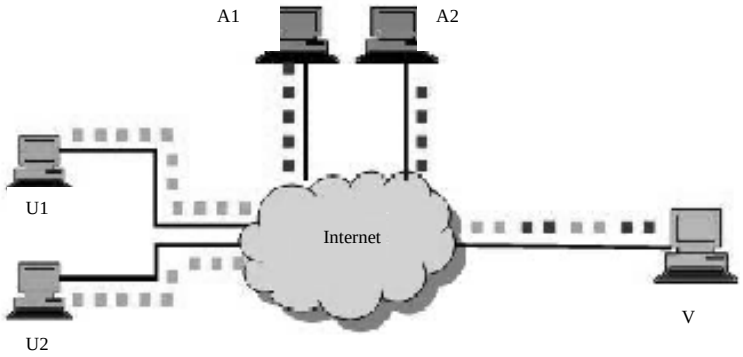

Fig. 1: Distributed denial of service attack (U1, U2-Legitimate User, A1, A2-Attacker, V-Victim)

MATERIALS AND METHODS

Attack detection methodologies: Attack can be defined as the process of making unavailable of a particular service provided by a computing machine. It can be identified by analyzing the performance of the system in terms of availability and efficiency. The process involves the generation of number of information packet more than that of the processing capacity of the routing machine. So that the system gets hang itself without any external source of attack. The following methods are used to detect the attack traffic (Chen, 2009) situations in the routing environment.

To identify the high network traffic, the threshold value of the network should be fixed for the maximum number of packets to be allowed. When there is a maximum number of packet arrival implies the chances of attack process in the network.

There are three various methods available in fixing the threshold value (Singh *et al*., 2014; Zhang *et al*., 2010) for router to forward the incoming packets. They are classified as:

- Type 1: based on the network capacity
- Type 2: based on the maximum number of packet arrived to the destination
- Type 3: based on the distribution for future prediction

Based on the network capacity: Network capacity is also called as router capacity which indicates the number of packets processed by an individual router at a time. The capacity is decided based on the model of the router. The situation in which the number of incoming information packets exceeds the capacity of the router leads to the jamming of the router and it leads to the decrease in the network efficiency.

Every router has its own processing capacity. According to the network packet processing, the router can be chosen. Routing capacity of the networking device indicates the processing capacity of the device in performing the routing operation for the incoming packets. This capacity is the limit of the maximum number of packets that a particular device can handle at any time. It can be denoted as the ratio of number of incoming packets received to the number of outgoing packets transmitted from the router.

Based on the maximum number of packets arrived to the destination: In case of non informative router, consider the number of packets arrived to router may vary from time to time and the maximum number of incoming packets is assigned as the capacity of the router.

Step 1: Let max be the maximum number of packets arrived and it is initialized to 0. Max = 0.

Step 2: If the maximum number of packets received at time T is greater than the packets received at time T+1, then the value of Max = number of packets received at time T+1. It can be denoted by using the code as follows:

Algorithim

```
Max=0;
for (int t=0; t<Rmax; t++)
{
If (Max (t+1)>Max (t))
Then
                    Max = Max (t+1);
Else
Max = Max (t);
}
```

Based on the testing of hypothesis: By considering the testing of hypothesis (Z-test), the packet arrival rate is classified and assigned the forwarding process. The various categories of packet analyzing are based on the following types:

- Test of significance for single sample proportio
- Test of significance for difference of two sample proportions
- Test of significance for single sample mean
- Test of significance for difference of two sample means

In this study, the test of significance for two sample proportions and two sample means are considered. Since the packets to be examined are drawn from two different timing intervals namely peak time and normal time, the value of incoming legitimate user packet and attack packet may vary.

The flow of incoming packets (Alkasassbeh *et al*., 2016) is less in normal time when compared to the peak time. In testing of hypothesis, assumption is made in the beginning and proceeds further to confirm based on the number of arrival of information packets. Hence, the testing of significance for two sample proportion and difference of two sample means are considered.

RESULTS AND DISCUSSION

Testing of significance for difference of two sample proportions: Assume the probability that the number of attack packets during the peak time is greater than number of attack packets in normal time. Then the population proportion for the difference of two samples is denoted by:

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$Q = 1 - P$$

Where:

P = Population proportion of difference of two samples
n_1 = Total number of packets arrived during heavy traffic
n_2 = Total number of packets arrived during normal traffic

The first process is to assign the null hypothesis (H_0) and alternate hypothesis (H_1) is as follows:

- H_0: assume that there is no significant difference of attack packets arrival in peak time and normal time, i.e., $P_1 = P_2$
- H_1: assume that the arrival of attack packets is greater in peak time when compared to normal time, i.e., $P_1 > P_2$

According to the testing of hypothesis (Z-test), the value of the Z-statistic is calculated as:

$$Z = \frac{P_1 - P_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Where:

P_1 = Proportion of attack packets in peak time
P_2 = Proportion of attack packets in normal time

The modulus of |Z |is calculated and compared with the table value of Z. The table value for Z at 5% level of significance for one tailed test is Z = 1.645.

If the calculated value of Z is less than the table value, then accept H_0, i.e., there is no significant difference of attack packets arrival in peak time and normal time (Fig. 2).

If the calculated value of Z is greater than the table value, then reject H_0 (accept H_1), i.e., the arrival of attack packets is greater in peak time when compared to normal time.

Difference of two means: The difference of mean is used to find out whether the possibility of arrival of attack packets during peak time is greater than the normal time.

The first process is to assign the null hypothesis (H_0) and alternate hypothesis (H_1) is as follows:

- H_0: Assume that there is no significant difference of attack packets arrival in peak time and normal time. i.e., $x_1 = x_2$
- H_1: Assume that the arrival of attack packets is greater in peak time when compared to normal time. i.e., $x_1 > x_2$

According to the testing of hypothesis (Z-test), the value of the Z-statistic is calculated as:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where:

n_1 = Total number of packets arrived during heavy traffic
n_2 = Total number of packets arrived during normal traffic
$\bar{x}_1$ = Mean of the attack packets at heavy traffic time
$\bar{x}_2$ = Mean of the attack packets at normal time
s_1 = Standard deviation during heavy traffic
s_2 = Standard deviation during normal time

The modulus of |Z |is calculated and compared with the table value of Z. The table value for Z at 5% level of significance for one tailed test is Z = 1.645. If the calculated value is less than the table value, then accept H_0, i.e., there is no significant difference between mean of attack packets in peak time and normal time.

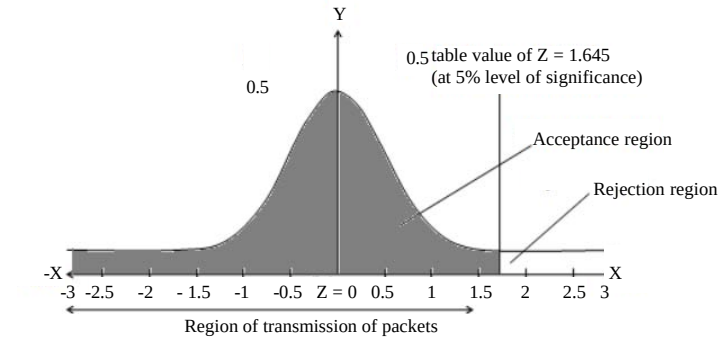

Fig. 2: Attack analysis using Z-test

If the calculated value of proportion is greater than the table value, then reject H_0 (accept H_1), i.e., there is significant difference between mean of attack packets in peak time and normal time.

CONCLUSION

Packet flow information in the router is utilized to identify the occurrence of attack packets in the network with the help of testing of hypothesis (Z-test). Once the test result shows that the value of Z lies within the transmission region it is under safe state. If the value gets increased over the threshold value, then the attack detection is achieved. Further remedy is to taken in order to avoid the attack in future and to find out the source of attack. Once the attack during the normal traffic time and heavy traffic time is categorized, the traceback approach is applied at the system by using the routing information stored in the packet. There are number of methods available to traceback the source of attack.

REFERENCES

Alkasassbeh, M., A.G. Naymat, A.B. Hassanat and M. Almseidin, 2016. Detecting distributed denial of service attacks using data mining techniques. Int. J. Adv. Comput. Sci. Appl., 1: 436-445.

Braga, R., E. Mota and A. Passito, 2010. Lightweight DDoS flooding attack detection using NOX open flow Proceedings of the 2010 IEEE 35th Conference on Local Computer Networks (LCN), October 10-14, 2010, IEEE, Brazil, ISBN: 978-1-4244-8387-7, pp: 408-415.

Chen, C.L., 2009. A new detection method for distributed denial-of-service attack traffic based on statistical test. J. Universal Comput. Sci., 15: 488-504.

Kumarasamy, S. and R. Asokan, 2011. Distributed denial of service (Ddos) attacks detection mechanism. Int. J. Comput. Sci. Eng. Inform. Technol., 1: 39-49.

Leu, F.Y. and I.L. Lin, 2010. A DoS DDoS attack detection system using chi-square statistic approach. J. Syst. Cybern. Inf., 8: 41-51.

Singh, B., D.S. Panda and D.G. Samra, 2014. Threshold based approach to detect DDoS attacks in cloud. Int. J. Innov. Res. Inf. Secur. IJIRIS., 2: 22-28.

Zhang, J., Z. Qin, L. Ou, P. Jiang and J. Liu *et al*., 2010. An advanced entropy-based DDOS detection scheme. Proceedings of the 2010 International Conference on Information, Networking and Automation (ICINA), October 18-19, 2010, IEEE, Changsha, China, ISBN: 978-1-4244-8104-0, pp: 67-71.