The Detection Method of DDoS Attack based on Information Entropy

Guojun Chen

College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract

Distributed Denial of Service (DDoS) attack is one of the malicious attacks that affect network security. In recent years, people have made a lot of detection methods in DDoS detection, but these algorithms still need to be improved. In this paper, we improved the traditional entropy algorithm for the purpose of reducing the false alarm rate and improving the detection rate, and proposed a DDoS attack detection algorithm based on information entropy. Based on this algorithm, we constructed an online experimental network environment. The experimental results show that the application of DDoS attack detection algorithm based on information entropy has better detection and corresponding effect.

Keywords

Information Entropy, Ddos, Attack Detection.

1. Introduction

With the rapid development of computer network technology, Internet closely linked to people's learning, work and life, and has been related to national government, military, cultural and educational fields, which becomes an important guarantee for the development of the whole society. Because the Internet is open and resource-sharing features, any group or individual can easily download or upload a variety of information online. This is the value of the Internet, but also determines the information on the Internet is difficult to monitor, and network violations are difficult to leave the relevant evidence, so it is difficult to find and confirm the initiator of network attacks. Due to the lack of effective detection methods, network attacks are becoming more and more frequent, the attack methods used by attackers are more and more advanced and mature, resulting in the loss is also growing [1, 2, 3]. Therefore, network security is facing more and more severe challenges. Specific cyber security threats include malicious attacks, security flaws, structural dangers and software vulnerabilities, where malicious attacks are the biggest threat. Malicious attack refers to the intentional destruction of the enemy in retaliation, personal interest or other purposes. Malicious attacks have the following characteristics: Malicious attacks are generally professional and technical personnel, they have a high professional skill. The concealment of malicious attacks is very strong, which is difficult to cause suspicion and tracking, and the crime technology is difficult. Malicious attacks on financial institutions' network systems can cause huge losses to financial institutions and even lead to their bankruptcy. The range of malicious attacks is varied.

From the perspective of the overall performance of the network, distributed denial of service (DDoS) attacks are one of the most serious threats to malicious attacks, and it poses a serious threat to the stable operation of the Internet. DDoS is a new type of denial of service attack method which has been developed from denial of service attacks [4, 5]. DDoS attacks mean that multiple attackers at different locations simultaneously attack one or more targets, or one or more attackers control multiple machines located in different locations to attack the victim at the same time. As the attack point is distributed in different places, so this type of attack is called a distributed denial of attack.

DDoS attacks are easy to implement, even if one do not understand the network security, and even just learn to use the computer, they can use the online attack software to attack, and the emergence of a more powerful attack tool in the future will bring more development possible to DDoS attacks. Therefore, how to detect and prevent DDoS becomes an important issue.

2. Methods

2.1 Analysis of DDoS features

2.1.1 Information entropy

Information entropy is a method of probabilistic testing and mathematical statistics to study the basic problems of communication[6,7]. It was first used in thermodynamics, the reaction reflects the degree of uncertainty, that is, the more ordered the system, the higher the entropy. The information entropy is formed by extending the thermodynamic probability to the probability that each information source signal appears[8]. From the communication point of view, the random interference of the source is unavoidable, the communication system has statistical characteristics, and the information source can be regarded as a set of random events. The randomness uncertainty of the set is similar to that of microcosm in the thermodynamics. The information entropy can effectively reflect the amount of information in the message.

In recent years, information entropy is used for network intrusion detection, we can regard the measure data as a discrete information source, and regard the measurement data in the various attributes as a set of random events, then we can analyze its information entropy. It can be used to detect large-scale network traffic DDoS attacks. The information entropy is defined as follows:

$$H(X) = -\sum_{i=1}^{N} P_i Log_2 P_i \tag{1}$$

Where X is the state of the source space, and it has a total of N states, that is, $X = (X_1, X_2, \cdots, X_N)$.

In this paper, N is a time period of different IP addresses or the total number of ports, X_i represents one of the IP addresses or ports, the probability of X_i occurrence is P_i, and $\sum_{i=1}^{N} P_i = 1$.

When DDoS attacks, a large number of puppet machines are controlled to attack at the same time, and the attacker will randomly forge the attack packet source IP address, or use the reflection DDOS attack mode, so the source IP address is more dispersed, and the number will surge. In the case of suffering DDoS attacks, there are plenty of packets which are unable to establish communication with servers in the network, as is called one-way connection packet, whose capacity can be reflected by the density of one-way connection. The network packets are flooded into the victim, and some of the ports that provide the service are attacked. Therefore, the destination IP address and destination port address of the network packets are quite centralized. Information entropy can effectively represent the same attribute on the corresponding data of the concentration and dispersion.

2.1.2 Unilateral connection density

When DDoS attacks, the attacker will forge the source address, or use the reflection attack mode, which cannot establish a normal connection, and the data will not be able to establish TCP / IP protocol 3 times the handshake, so it cannot establish a normal connection. OWCD [8] can be a good response to this phenomenon that it cannot establish a normal connection. The process can now be reformulated with more detail as follows: suppose $R = \{p_1, \cdots, p_i, \cdots, p_j\}$ is a set of IP packets, among them, the packet elements are in triad form, that is, $p_i = (SrcIP_i, Src\,\mathrm{Prt}_i, DstIP_i, Dst\,\mathrm{Prt}_i, c_i)$, in which, $SrcIP_i$ represents the source IP address of packet i, $Src\,\mathrm{Prt}_i$ indicates the source port number, $DstIP_i$ indicates the destination IP address, $Dst\,\mathrm{Prt}_i$ indicates the destination port number, c_i is the TCP control bit, or ICMP request. If there is a certain two IP meet $SrcIP_i = DstIP_j$, $Dst\,\mathrm{Prt}_i = Src\,\mathrm{Prt}_j$,

$DstIP_i = SrcIP_j$, $Dst\Pr t_i = Src\Pr t_j$, then constitute a two-way connection (TWC), on the contrary, constitute a one-way connection, connection (OWC). The OWCD formula is as follows:

$$OWCD = \sum \frac{OWCPackets}{IPPackets} \times 100\% \quad (2)$$

Among them, OWC Packtes refers to the number of pairs of IP addresses that are normally connected, and IP Packets refers to the total number of addresses. When an attack occurs, the data flow will increase that those cannot establish a normal connection in the network, and it will make OWCD significantly increased. The experiments show that OWCD value is generally below 39 in the normal flow, but in the DDoS attack, OWCD will tend to 100. In this paper, we only consider IP bidirectional flow of three kinds protocol, which are TCP, UDP and ICMP, and count the OWCD values of the three protocols respectively.

2.2 DDoS detection algorithm

This algorithm is divided into three stages: pretreatment, training and detection. The process is shown in Figure 1, By extracting the IP information of the network packet and calculating the entropy value and OWCD value of the packet, the preprocessing stage obtains the DDoS attack characteristic representation. In the training phase, the data entropy is clustered by using the weighted Euclidean distance formula and the thresholds are established. The training set of normal behavioral characteristics is established. The detection phase is also subjected to a clustering process, and then the threshold is analyzed. If the data belongs to the normal range, the normal data is classified into the training set and classified, and then the attack response is carried out.

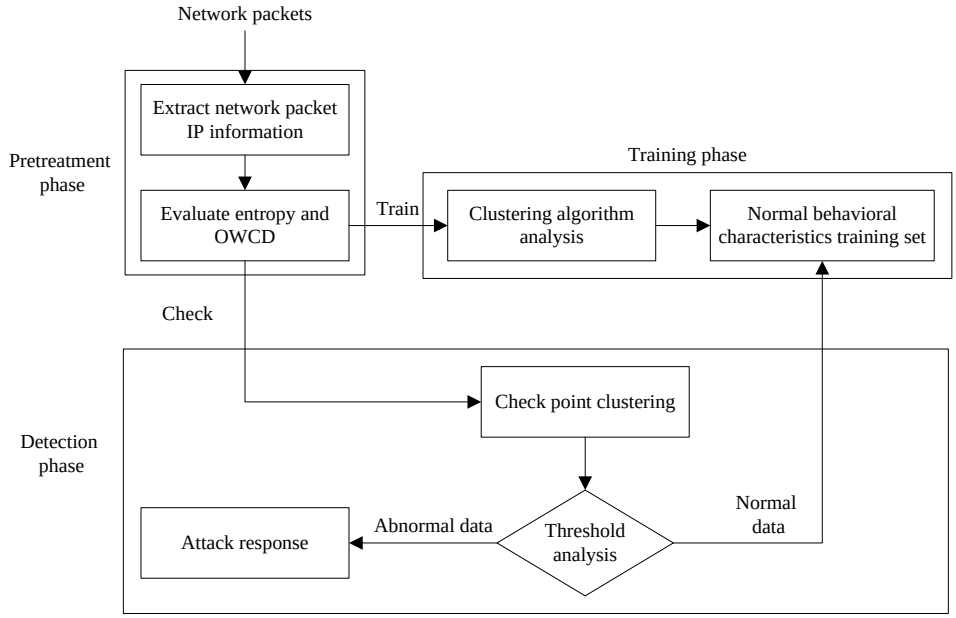

Fig. 1 DDoS detection algorithm flow

Because this article is not only to detect DDoS attacks, but also to detect whether it belongs to TCP SYN flood, UDP flood and ICMP flood attack, so we have to establish a total 4 training set of DDoS, TCP SYN flood, UDP flood and ICMP flood. In the training process, there is no need to establish a MIX flood training set, because the proportion of TCP, UDP, and ICMP packets is flexible during DDoS attacks. If the training set of MIX flood is set up, it is necessary to add different proportion of data packets for training, and establish different training sets. In this paper, if two or more than two attacks of TCP flood, UDP flood and ICMP flood occurred at the same time, it is considered to be a MIX flood attack.

3. Results

3.1 Experimental network environment

In order to verify the effectiveness of the DDoS detection system, this paper constructs an online experimental network environment, whose topology is shown in Figure 2. Host D is the attacker, host A is used to generate background traffic, and hosts B and C are used to generate DDoS attack traffic. Because of this detection system is not only to detect attacks, but also take the initiative to attack response, so the PC machine is installed on the double card and Linux operating system simulation and router, while the detection system deployed on the machine, which can achieve filtering attack. Through the configuration in the Linux operating system under the dual network card routing table, it can achieve the function of the router.

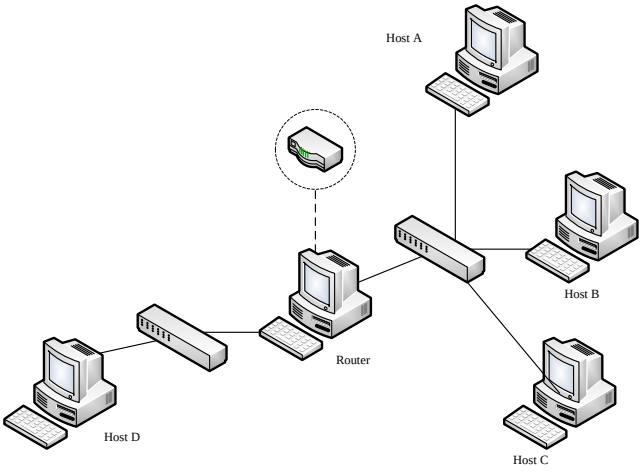

Fig. 2 Experimental network topology

In order to achieve the simulation of the actual network environment effect, we use the traffic generation tool npag to generate different rates of TCP, UDP and ICMP background traffic. In the host A, B running TFN2K [9] to generate attacks, TFNZK is able to run in the Linux operating system to attack tool, and it can simultaneously open multiple threads to initiate a pseudo-source IP address specified protocol attacks. In this experiment, the number of open threads is 15, respectively, launched TCP SYN flood, UDP flood and ICMP flood attacks. However, the MIX flood attack uses TCP, UDP and ICMP ratio of 1: 1: 1 alternately initiated formation. Attack strength is difficult to determine in the experiment, the current study does not have the exact conclusion in this regard. In this experiment, two attackers form TCP, UDP and ICMP attack packets that can reach an average of 7368 / s, 14958 / S, 5793 / S, this rate is enough to cause the victim machine crash.

3.2 Result analysis

According to the relationship between the sampling specimen and the detection rate, the training is to collect 800 training samples, and the 3 class has a total of training samples of 2400, and carry out the clustering training respectively, while establish the source IP address database. And then send the test sample to the training set for testing, which can give the detection rate of each class.

Table 1 lists the detection system for each type of DDoS attack detection rate and background traffic false alarm rate. It can be seen that the detection rate of the on-line attack is slightly lower than that of the corresponding off-line attack, but still shows a better detection effect than the false positive rate in the off-line detection.

When an attack is detected, the attack response is started and the attack packet is filtered. TCP SYN flood is as shown in Figure 3 and Figure 4, they represent the corresponding packet changes that did not initiate the attack response and start the attack response, respectively. It attacked at 755, 255 and 655, attack time was 85, at the time of fifty-fifth, the attack stream was filtered out, the data flow soon returned to normal.

Table 1. Online detection results of different types of DDoS attacks

Attack type	Detection rate /%	False rate /%
TCP SYN flood	91.1	13.5
UDP flood	89.9	15.7
ICMP flood	85.1	14.1
MIX flood	91.7	9.5

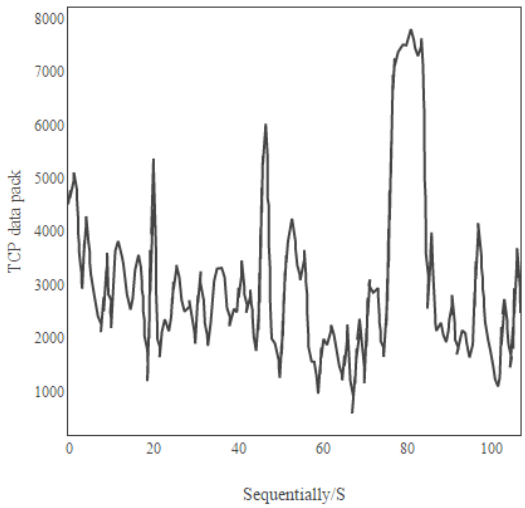

Fig. 3 The number of packets without attack response (TCP SYN flood)

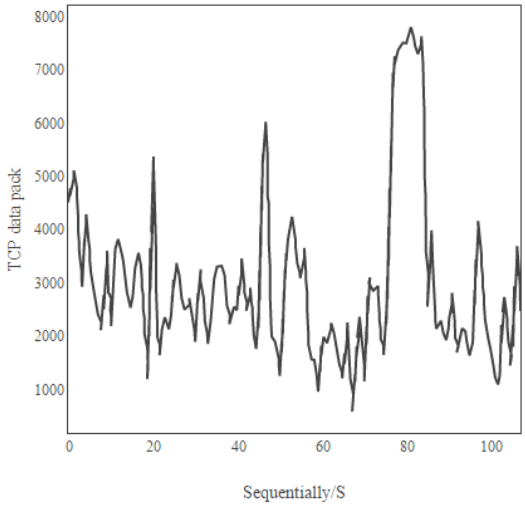

Fig. 4 The number of attack packets that open the attack response (TCP SYN flood)

As can be seen from the comparison of Figure 3 and Figure 4, in Figure 4, the data flow in the 50S-56S has declined, but no attack has occurred at this time. This is due to the detection algorithm has a certain false alarm rate, resulting in regards the normal flow of the burst as the attack traffic, and the normal flow has been filtered.

4. Conclusion

In this paper, we first proposed a DDoS attack detection algorithm based on information entropy, and then constructed an online experimental network environment, and carried out the experimental research in this experimental environment. The on-line test shows that the detection rate of the model

for different attack types and the detection rate in the offline state were basically similar, the detection rate had reached more than 85%, and the false alarm rate of the normal data was within 15%, which was also at a level that can be received. At the same time, the experiment proves that the detection system has the defense ability to DDoS attack, and can filter the high intensity DDoS attack packet in time, so that the legitimate traffic can be protected better.

References

[1]. Ingols, K., Chu, M., Lippmann, R., Webster, S., & Boyer, S. (2009, December). Modeling modern network attacks and countermeasures using attack graphs. In *Computer Security Applications Conference, 2009. ACSAC'09. Annual* (pp. 117-126). IEEE.

[2]. Ourston, D., Matzner, S., Stump, W., & Hopkins, B. (2003, January). Applications of hidden markov models to detecting multi-stage network attacks. In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on* (pp. 10-pp). IEEE.

[3]. Sunny Behal, Krishan Kumar. Trends in Validation of DDoS Research [J]. Procedia Computer Science, 2016, 85.

[4]. Zargar, S. T., Joshi, J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. *IEEE communications surveys & tutorials*, *15*(4), 2046-2069.

[5]. Compagno, A., Conti, M., Gasti, P., & Tsudik, G. (2013, October). Poseidon: Mitigating interest flooding DDoS attacks in named data networking. In *Local Computer Networks (LCN), 2013 IEEE 38th Conference on* (pp. 630-638). IEEE.

[6]. Han C K, Choi H K. Effective discovery of attacks using entropy of packet dynamics. Network, IEEE, 2009, 23(5): 4 -12.

[7]. Cover T M, Thomas J A. Elements of information theory. John Wiley & Sons,2012.

[8]. Yong Hao Gu,Wei Ming Wu. DDoS Detection and Prevention Based on Joint Entropy and Conditional Entropy [J]. Key Engineering Materials, 2011, 1244(474).

[9]. Nugraha, M., Paramita, I., Musa, A., Choi, D., & Cho, B. (2014). Utilizing Open Flow and sFlow to Detect and Mitigate SYN Flooding Attack. *Journal of Korea Multimedia Society*, *17*(8), 988-994.

[10]. Wang, Y., Jiang, H., Liu, Z., & Chen, S. (2015). A CRF-Based Method for DDoS Attack Detection. In *Proceedings of the 4th International Conference on Computer Engineering and Networks* (pp. 81-87). Springer International Publishing.