A Defense System for DDoS Application Layer Attack Based on User Rating

Gaojun Jiang[1, a], Zhengping Jin[2, b]

[1] Beijing University of Posts and Telecommunications, Beijing, 100876, China

[a]buptjgj@126.com

Keywords: Application Layer DDoS, User Scoring, Flow characteristics, DDoS Defense.

Abstract. Based on the DDoS attack on application layer and the difference of normal users' access habits and the traffic features, we proposed a user rating based defense system for DDoS application layer attack. This system can significantly improve the detection rate through analyzing and scoring the access record of source IP address as well as the traffic features. Besides, the new system supposes to send verification code to users whose scores are below the threshold, which can efficiently promote detection rate of DDoS attack, reduce false alarm rate and significantly enhance normal users' experience.

1.Introduction

It has been very difficult to achieve the goal of attack for the bottom layers DDoS attacks, as the detection technology against DDoS attacks in network and transport layers become more and more mature and the correspondent defense continued to be strengthened. Therefore the attackers turn to transfer the target to the application layer. Application-layer DDoS attacks usually consume server a lot of computing resources to prevent normal accesses to legitimate users. Unlike the traditional network-layer DDoS attacks, the application-layer attack happens in the application layer. Since the application layer attacks using high-level protocol, the service request submitted to the target server is as same as the request submitted by legitimate users. So these service requests do not need to utilize network protocol vulnerabilities, and they are all legitimate. More importantly they are very similar to bursty traffic FC (Flash Crowd). Thus, the conventional method for detecting DDoS attacks at the network layer and the transport layer can not be directly used for application-layer DDoS attack detection.

In recent years, researches on application-layer DDoS attack detection methods have become a new popular focus in the field of network security.Xie et al.[1] presented a method that will make statistic fort users' browsing behaviors. According to features such as websites accessed by users and the responses from server or proxy server, they apply hidden semi-Markov model to simulate the legitimate users' access behavior. If the monitored data stream has a considerable bias from the simulated legitimate users' access behavior, the monitored access behavior then is viewed as a source of attack. Zhao et al.[2] proposed a detection method based on analysis of users' behavior characteristics. This method built a legitimate user model from aspects such as legitimate users' browsing webpages, staying time on the page, etc. If a user's behavior characteristics did not match this model well, the system affirmed that an application layer attacks occurred. Ahn et al. [3] proposed PUZZLE detection method. When the server resource was excessively consumed, the system can suspect itself under attack. It then requires the user to answer a few simple questions to determine the legitimacy of the user, since the puppet program cannot correctly recognize verification code pictures.

In this paper, we propose a novel detection model for application-layer DDoS attack. Firstly, We make feature extraction and analyze the IP address access features. Then the weight of each feature is determined by using AHP (analytic hierarchy process) and a score is given respectively to each feature. Through scoring and evaluating users' access information, our method can efficiently detect application-layer DDoS attack. The simulation and experiment show that comparing to traditional detection methods this approach can be more easily implemented, and get a higher detection rate and

a lower rate of false alarm . Besides, our method can dramatically enhance legitimate users' access experience. More details are shown in the Table 1 below.

Table 1. Comparison of detection methods

Detection methods	Detection rate	False alarm rate	Advantage	Disadvantage
HSMM	About 95%	About 3%	The calculated results are more reliable and accurate	Malicious attack software can simulate the normal user behavior, and for each user to build a model of the workload is more complex
Based on analysis of users' behavior characteristics	About 96%	About 3%	Ditto	Ditto
PUZZLE	Depending on the verification code	Depending on the verification code	Simple and effective, can quickly identify whether the normal user	Will seriously affect the user experience
Based on user rating	About 97%	About 2.88%	Simple and effective, with high detection efficiency, and has a good user experience	The evaluation system has a certain lag

2.Principle definition

The closer the user access data distribution is to the normal data distribution, the higher score it is assigned and vice versa. Evaluation indicators are as follows:：

The probability that the server was attacked in the last week A1: The marked number when accessed server being under attack in the latest week or the total number of visits can intuitively reflect the direct relevance of this IP address with DDoS attacks.

The last time the verification code is through A2： The records whether the last access triggers verification code and passes through the verification can reflect that in the recent time accesses are normal or not.

Last month's visit days A3: the number of days accessing the specific website in last month can reflect whether this IP address is an old user that access frequently.

Average download traffic per second A4: This may reflect the degree of consumption for IP network traffic. In a certain degree, it can be referred as indicators of judging application-layer DDoS attacks.

Packet density A5: the number of packets sent / data packet transmission time. The small density often means that there is a slow connection attack; if the density is particularly large, it means that there are often possible flooding attacks. Therefore, the normal value should be within a range. Being larger or smaller than this range means that threats are possibly aroused.

The average number of requests per second A6: the number of requests which normal browsing sends is supposed to be lower than a threshold value. If the request is sent too frequently, DDoS attacks possibly occurred.

IP addresses are domestic or not A7: the possibility of domestic IP address launching attack is smaller, while the foreign IP address is more likely to be puppet machine.

3. System Design

In this paper we focus on application-layer DDoS attack type and set the access habit and traffic as our research object. The defense system model based on users' rating for application-layer DDoS attacks is shown in Figure 1. It is mainly divided into network data capture module, scoring and computing module, the verification code module and limited-function modules. The network data

capture module accomplish network data acquisition, the scoring and computing module accomplish the calculation and storage of various evaluation and scoring; verification code module is to complete the specified user authentication code verification; limited-function modules accomplish access limit to suspicious users who obtain very low scores.

Traditional detection methods just simply distinguish users' IP from being normal and abnormal, which will cause a great randomness. To solve this problem, in our system, we evaluate and rate all the access IP according to the normality of daily operations and the use frequency, and store the result into the database. When significantly increased data is detected, our system will use the amount of new IP and IP of low scores to determine whether it is under attack. If the system detects attack, it will automatically launch the verification code module for the requests from IP of which the score is below threshold, and disconnect those connections, which the client traffic is beyond threshold and have not got verified by verification code. For users with high scores, the system does not put any limitation to them, however at the same time greatly reducing their rates. Thus the system can avoid the complex process for each user's operation modeling, greatly saving server resources. Besides the efficiency and accuracy will be greatly improved. Moreover, it greatly enhances the legitimate user experience.

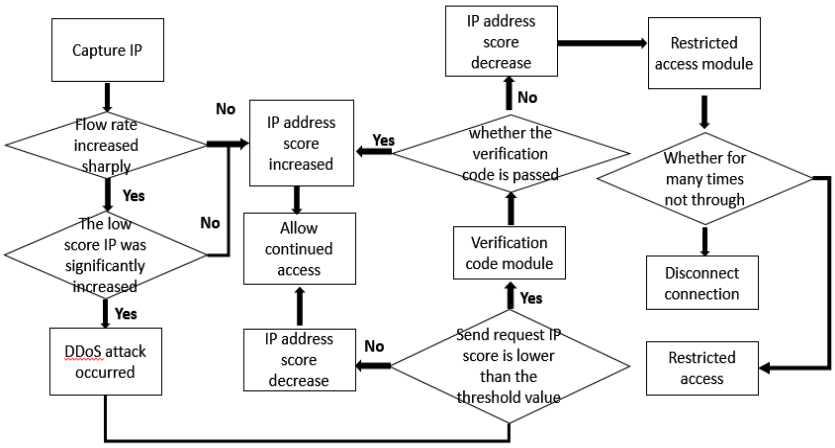

Figure 1. A Defense System for DDoS Application Layer Attack Based on User Rating Model

4. System Implementation

4.1 Determing rating index weight

We apply AHP (analytic hierarchy process) to determine the index weight [4]. By comparing the previous importance of each index, as shown in Table 2, we apply the 1-9 scale method to get the judgment matrix B_{7*7}. The relevant importance of Judgment Matrix Index is derived from previous research experience as well as consulting relevant archives. As in the judgment matrices A1 is slightly more important than A2, $b_{1,2} = 3$, $b_{2,1} = 1/3$. While A1 is extremely more important than A7, so $b_{1,7} = 9$, $b_{7,1} = 1/9$. In this way we can get the value of other elements in the matrix. The judgment matrix was shown in Table 3.

Table 2. Meaning of scale

A Indicator than the B Indicator	A Indicator evaluation value	Remarks
Extremely important	9	Take 8,6,4,2,1/2,1/4,1/6, 1/8 as the intermediate value of the above evaluation value
Very important	7	
Important	5	
Slightly important	3	
Equally important	1	
Slightly secondary	1/3	
Secondary	1/5	
Very secondary	1/7	
Extremely secondary	1/9	

Table 3. Evaluation matrix B_{7*7}

Indicator	A1	A2	A3	A4	A5	A6	A7	Multiplicative	Square root	Score
A1	1	3	5	6	6	7	9	34020	4.44008	39.9635
A2	1/3	1	3	5	5	6	8	1200	2.75349	24.7831
A3	1/5	1/3	1	5	5	6	7	70	1.83479	16.5142
A4	1/6	1/5	1/5	1	1	3	7	0.14	0.75512	6.7966
A5	1/6	1/5	1/5	1	1	3	5	0.1	0.71969	6.4777
A6	1/7	1/6	1/6	1/3	1/3	1	4	0.00176	0.40423	3.6383
A7	1/9	1/8	1/7	1/7	1/5	1/4	1	0.0000141	0.20293	1.8265

4.2 Setting the scoring sheet according to the weights

Let individual score out of 100 points. We apply AHP to get the weight of each index to the whole target. By multiplying the weight with the score, we can get the overall score of each index. Then according to every possibility of each index, we set the corresponding alternative score value of each index [5]. The value of alternative scores can be derived from statistic of test data. And those alternative scores that perfectly match the use features of legitimate users will be given higher value. The scoring table we obtained is as follows:

Table 4. Score table

Score indicator	Full marks	Option	Score
A1:The probability that the server was attacked in the last 1 weeks	40	0%-5% 5%-10% 10%-20% 20%-100%	40 20 10 0
A2:The last time the verification code is through	24	Yes No	24 0
A3: Last month's visit days	17	15-31 8-15 3-8 0-3	17 12 8 0
A4: Average download traffic per second	7	0-10KB 10-100KB >100KB	7 5 0
A5: Packet density	6	10-500 0-10 >500	6 0 0
A6: Average number of requests per second	4	0-20 20-50 >50	4 2 0
A7: IP addresses are domestic or not	2	Yes No	2 0

After getting the above score table, the next step is to correspond the statistical data of each IP to the score table, and sum up each individual score to get the total score of the IP respectively. Provisions for the score is not higher than the threshold for the occurrence of attacks need to enter the verification code of the user, higher than the threshold value can be accesse. The higher the score, the less possible of being suspected when DDoS attack occurs.

4.3 Scoring system simulation

In this simulation test, the training data set is consist of 5000 groups of data collected when legitimate users accessing the website. In addition, one server is used to simulate web server and database server. And we run the library management information system on the server to manage information system, while using MySQL to achieve database service. 20 hosts simultaneously launch App-DDoS attack. Each machine runs 10 attack-programs. One type of attack sends GET request to server within nearly zero intervals to simulate flooding attack. While the other type of attack constantly sends heavy load requests to generate some real web application-layer DDoS attack data. Those data are considered as abnormal test data, which will be used to test the detection rate and false alarm rate of the model [6]. The detection rates and false alarm rates obtained in using different scoring threshold are shown in Fig.2:

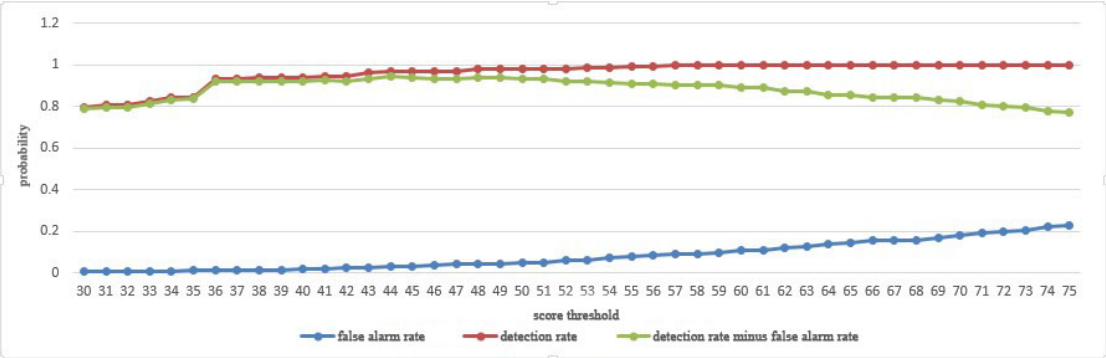

Figure 2. Simulation result

As the above figure shows, only when a legitimate user's access habit is very abnormal (such as download data flow is huge or send request too frequently and code verification failure), will our system mistakenly detects the legitimate user as an illegal one.

Through simulating the data, we derive that when the scoring threshold is 44, the detection rate is 97% and the false alarm rate is 2.88%. Under this condition, we achieve the maximum value of detection rate minus false alarm rate, which is 94.12%. Thus, our method can efficiently detect DDoS attacker, and meanwhile it can relieve the pressure of server. Besides it can enhance the legitimate users' access experience. Overall, our method proposes a new solution to DDoS attack problem.

5. Conclusion

In this paper, the normal users' access habits and traffic characteristics are analyzed, using the analytic hierarchy process to establish the IP credit evaluation model, and in the simulation of the network environment for the empirical test and comparative analysis. Results show that the scoring system relative to the traditional method can more effectively detect the application layer DDoS attacks and simultaneously, in the greatest degree, enhance the user experience of access, which provides a new idea for the application layer DDoS attack detection.

Acknowledgments

This work is supported by NSFC (Grant Nos. 61300181, 61502044), the Fundamental Research Funds for the Central Universities (Grant No. 2015RC23).

References

[1] Y.Xie and S.Yu.A detection approach of user5 behaviors based on HsMM[C].Proc.19th Int.Teletraffic Congress(ITC 19),Beijing,China,2005 Aug.29-Sep.2.451-460

[2] Guofeng Zhao,Shoucheng Yu,Sheng Wen.Detecting application-layer DDoS attack based on analysis of users' behaviors[J].Application Research of Computers[D],2011,28(2):717-719.

[3] Luis von Ahn,Manuel Blum,Nicholas J.Hopper and John Langford.The CAPTCHA.http://www.captcha.net.2000.

[4] Hongcai Sun,Ping Tian,Lianfen Wang.Network analytic hierarchy process and Decision Science[M]. National Defence Industry Press,2011.

[5] Xiaoming Zhu,Zhiguo Liu.Summary of credit scoring model[J]. Statistics and decision making,2007(01):103-105.

[6] Jian Cheng,Yujun Lian. Research on the modeling and verification of credit scoring system[J]. International Financial Research,2007(6):50-59.