DDoS Attack Defense in Big Data-oriented Urban Rail Transit Safety Evaluation System

Qi Ding [1, a], Lijian Sun [1, b], and Xiaoyu Wang [1, c]

[1] Jinan Rail Transit Group Co., Ltd, Jinan, Shandong 250000, China;

[a] qdingjn@163.com, [b] ljliujn@163.com, [c] xywangjn@163.com

Abstract. Big data technology is playing an increasingly important role in the safety evaluation of urban rail transit, which can help improve the safety and reliability of the transportation system. However, such a big data system will also become the target of attackers, especially in the face of DDoS attacks, its stability and availability may be greatly affected. Various attacks against the Internet have become a major concern among the many flaws of the contemporary Internet. Unfortunately, attacks against the Internet continue to grow more sophisticated, unpredictable, and unstoppable. At present, many researchers have devoted themselves to the research of DDoS defense solutions, and it is necessary to summarize the existing research on DDoS defense work and the technologies used in it. In this article, we introduce the process of DDoS attacks, and summarize the characteristics, mechanism and weaknesses of various DDoS attacks according to the network layer where the attack is located. Finally, defending against DDoS attacks is crucial to the big data-based urban rail transit security evaluation system. This paper summarizes the content of this paper and discusses some open issues in the field of network intrusion detection.

Keywords: Rail transit; big data; DDoS; safety evaluation.

1. Introduction

The urban rail transit safety evaluation system relies on the real-time data collection, processing and analysis of big data, so its system stability is very important. DDoS attacks can cause system crashes or slow down by simulating a large number of fake requests, consuming network bandwidth and system resources. Studying DDos attacks can ensure system stability, data integrity and accuracy, improve user experience, ensure timely response, and protect data privacy. These measures will help improve the safety, make the urban transportation system safer and more reliable, and provide citizens with a better travel experience.

Urban rail transit safety evaluation requires real-time monitoring and rapid response to traffic incidents in order to take emergency measures and avoid potential dangers. DDoS attacks may cause system delay and slow response, thus affecting the timely handling of traffic incidents. Effective DDoS protection measures can guarantee the rapid response of the system and ensure that necessary security measures can be taken in time in emergency situations.

1.1 Rail Transit Data Security

The urban rail transit safety evaluation system may provide citizens with real-time traffic information and recommended safe travel routes. DDoS attacks may cause users to be unable to access these information and services normally, affecting citizens' travel experience. Defense against DDoS attacks can improve user experience and ensure that citizens can easily obtain safe and reliable traffic information.

Big data may contain some sensitive personal information, such as passenger travel trajectories. DDoS attacks may cause system vulnerabilities to be breached, and then leak these sensitive data. Defense against DDoS attacks helps to protect data privacy and prevent data from being illegally obtained and misused.

At present, some scholars have summarized the existing research results in the field of intrusion detection. Zargar et al. [1] first introduced DDoS flood attacks by categories according to the layers of the protocol (network layer/transport layer/application layer), then introduced Botnet, an important tool for launching DDoS flood attacks, and finally investigated DDoS Defense strategies

for flooding attacks. Jaafar et al. [2] surveyed 12 detected application-layer DDoS attacks (from January 2014 to December 2018) and reviewed the four defense stages of DDoS attacks and the latest DDoS attack detection methods. Jing et al. [3] first divided security-related data.

1.2 Data Security Protection Measures

Network data analysis methods include methods based on statistics, methods based on machine learning and methods based on knowledge. The statistics-based method first generates a portrait representing normal network behavior, and then uses statistical methods to calculate the difference between the traffic and the normal portrait. [4] The traffic whose difference exceeds the threshold is malicious traffic.

In contrast, the application of machine learning technology for intrusion detection is currently a popular research direction. Machine learning methods can build explicit or implicit models of the data by analyzing the characteristics of the traffic data. These methods have high detection rates and are able to be retrained and updated for new traffic. However, more computing resources are required during the training process.

Another approach is a knowledge-based approach, which matches network behavior with pre-defined rules or patterns to check for known malicious behavior. However, such methods have the disadvantage of not being able to detect zero-day attacks.

Researchers have summarized the existing research on the application of machine learning technology in the field of network intrusion defense. Berman et al. [5] discuss how each DL technique can be applied to cybersecurity, covering a broad range of attack types. Nisioti et al. [6] gave a comprehensive overview of unsupervised and hybrid (combined supervised and unsupervised) intrusion detection methods, related feature engineering techniques, and discussed the method of using IDS for attribution to identify attackers. [7] Based on the ability of different machine learning techniques to detect different attacks, the author compares and analyzes various machine learning techniques. Nguyen et al. [8] discussed some key requirements. In addition, Nguyen et al. [9] outlined an approach to applying deep reinforcement learning to the domain of cybersecurity.

The goal of this article is to select the appropriate data analysis method to detect the occurrence of network intrusion events, and set appropriate containment or mitigation rules to reduce the impact of attack traffic on victims. First, we'll cover popular DDoS attack categories. At the same time, according to the attack rate, it can also be divided into high-speed DDoS attack and low-speed DDoS attack.

2. Data Security Attack Process

The main characteristics of DDoS attacks are distribution and cooperation. There are roughly two ways to launch DDoS attacks, namely vulnerability attacks and flood attacks. Flooding attack is the most common DDoS attack, as shown in Fig.1.

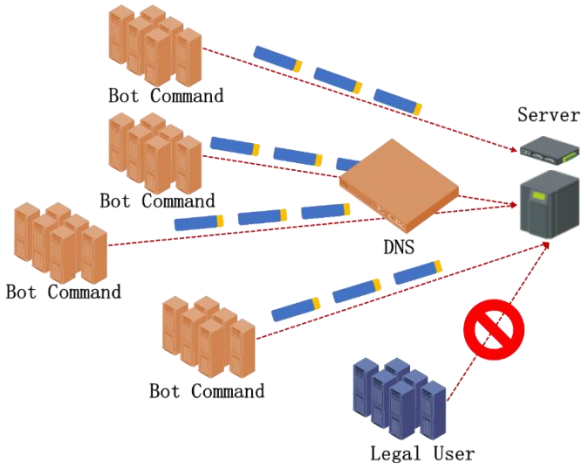

Fig. 1 DDos flooding attack

In order to realize the distribution and larger attack scale of DDoS flood attacks, as shown in Fig.2, attackers usually use a large number of nodes (Botnet) to attack victims. [10] Today's Botnet (such as P2P Botnet) can use the bot in the Botnet as a C&C server, which can effectively avoid the occurrence of a single point of failure. The bot administrator controls the host of the entire Botnet through the C&C server. The botmaster generates as little traffic as possible and avoids direct connections to hosts in the botnet in order to remain invisible and avoid detection by third parties (e.g. law enforcement agencies, researchers). Bots are infected hosts that execute Botnet programs, and they can perform tasks given by the C&C server [11].

Initial infection and secondary injection: Users will cause initial infection of the host in various situations, such as executing malware, opening emails containing malware, clicking unknown links, etc. [12] After successfully infecting the host, the next step is to download and run the Botnet program on the host, making the host a part of the Botnet (becoming a bot) and under the control of the botmaster.

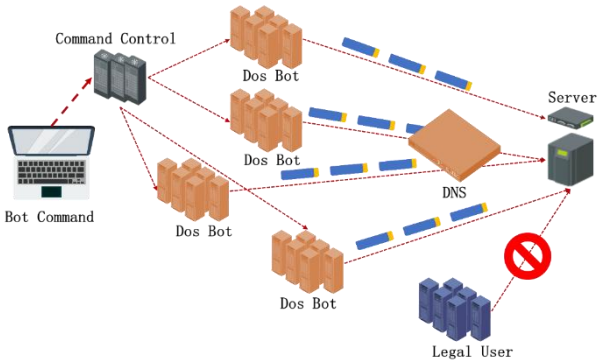

Fig. 2 Large scale DDos attacks

Status feedback and update: After becoming a bot, the infected host starts to log in to the C&C server, uses the protected session and parses and executes the commands in the session. [13] This process is performed periodically and requires authentication. Before the botmaster authorizes the attack, it usually sends an update command to the C&C server, and the C&C server contacts the bot and feeds back the status of the entire Botnet to the botmaster.

Such attacks are usually launched using UDP, TCP and ICMP protocol packets. A UDP flood attack sends a large number of UDP packets to a target server with the goal of overwhelming that device's processing and responsiveness.

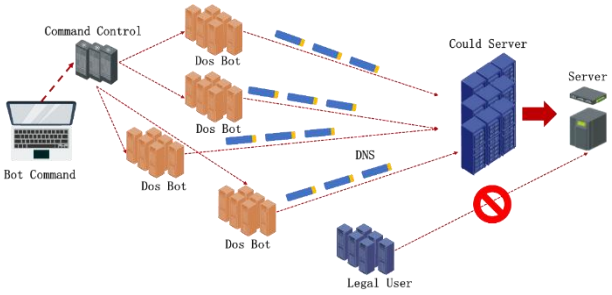

Fig. 3 Large scale DDos attacks on servers

These SYN packets will occupy a large amount of processing resources of the victim and generate a large number of half-open connections, [14] so that the victim does not have enough resources to process requests from legitimate users, as shown in Fig.3.

Ping flood attack is a simple DDoS attack using ICMP protocol. The attacker floods the target with ICMP "echo request" packets. Ping flood attacks can be made most effective by using the Ping flood option. Since the victim is constantly sending "echo replies" to the attacker, such an attack consumes both the victim's incoming and outgoing bandwidth, and may even consume its computing resources significantly.

3. Attack Detection Methods

This article introduces the life cycle of Botnet and the types, mechanisms and targeted weaknesses of DDoS flood attacks. In order to discover DDoS flood attacks, the characteristics of DDoS flood attacks can be used for direct detection, and the existence of Botnet can also be detected.

At present, the application of machine learning methods in the field of attack detection is a research hotspot. Before the popularity of machine learning algorithms, many scholars have conducted research on DDoS attack detection methods. We divide the existing DDoS attack detection methods into machine learning-based detection methods and traditional detection methods (such as statistical methods and knowledge-based methods) according to the types of data analysis methods, and introduce them respectively.

3.1 Machine Learning Methods

Subbulakshmi et al. [15] generated their own DDoS attack dataset on the experimental platform, which contains 10 types of network layer/transport layer DDoS attacks. After experimental comparison, it is concluded that the detection performance of EMCSVM is better than that of SVM.

Wang et al. [16] proposed a DDoS attack detection method based on the RDF-SVM algorithm. RDF-SVM is a machine learning learning algorithm proposed by the author that combines random forest and SVM. [17] After experiments In comparison, using this subset to train the algorithm can achieve a higher detection rate and recall rate.

Tan et al. [18] proposed a framework for detecting and defending against DDoS attacks in a Software defined network (SDN).

Qin et al. [19] first aggregate data packets into flows. The authors generate entropy vectors of five-dimensional features and utilize the K-means algorithm to generate portraits of normal traffic.

Rahman et al. [20] evaluated method in the SDN environment. The experimental, evaluation process includes the training of various models and their fusion with interception mechanisms.

3.2 Deep Learning Methods

Shaaban et al. [21] used convolutional neural network (CNN) to classify DDoS traffic and normal traffic. The author used two datasets for CNN model training, one is the public dataset NSL-KDD dataset, and the other is the dataset collected by the author himself using the Wireshark

traffic collection tool. [22]After feature generation and feature processing, the features of the two datasets are input into the neural network.

Li et al.[23] proposed a novel DDoS attack detection method, which combines long short-term memory artificial neural network and Bayesian method, called LSTM-BA. Using network traffic, LSTM-BA first learns DDoS attack patterns based on the LSTM method to obtain the predicted probability of DDoS attacks. Those with high predicted probability are identified as DDoS attacks, those with low probability are treated as ordinary traffic, and for those predicted values that are neither high nor low, the author further uses the Bayesian method to identify DDoS.

3.3 Real-time Detection Methods

In this section, we discuss traditional detection methods for network layer/transport layer DDoS attacks. Although the traditional methods basically do not have the ability to detect new types of attacks, the research on these detection methods has been relatively mature and has a high detection rate within the scope of application.

Bellaiche et al. used an entropy-based method to discover abnormal changes in the TCP handshake. Entropy is calculated based on the proportion of these suspicious packets that appear in normal traffic. When a SYN flood attack occurs, a large number of abnormal handshakes will be generated, causing the above four types of data packets to change drastically. By setting the entropy threshold, the SYN flood attack can be identified.

Sengar et al. first constructed the probability distribution of normal traffic using the normalized frequency of SYN, SYN-ACK, RST, and FIN packets. When detecting, the Hellinger distance between the probability distribution of the current traffic and the normal traffic distribution is used to compare with the set dynamic threshold, and when it exceeds the threshold, it is judged as a DDoS attack event.

David et al. detect DDoS attacks by observing connections. The authors construct dynamic thresholds to improve the flexibility of the detection system. [28] If the difference between the traffic entropy at the current time point and the entropy in the time interval is too large (more than the threshold), it is judged as a DDoS attack.

4. Rail Transit Data Security System

In the urban rail transit safety evaluation system based on big data, network security, as a key factor, needs to be given full attention. Urban rail transit system is a vital infrastructure in modern cities, and its stable and safe operation is crucial to the life and traffic flow of urban residents. However, with the widespread application of network technology, rail transit systems are also facing threats from network attacks.

DDoS flood attack is a common means of network attack. It floods the network bandwidth and resources of the target server, making the service unavailable and threatening the stable operation of the rail transit system. Therefore, in the urban rail transit safety evaluation system based on big data, as shown in Fig.4.

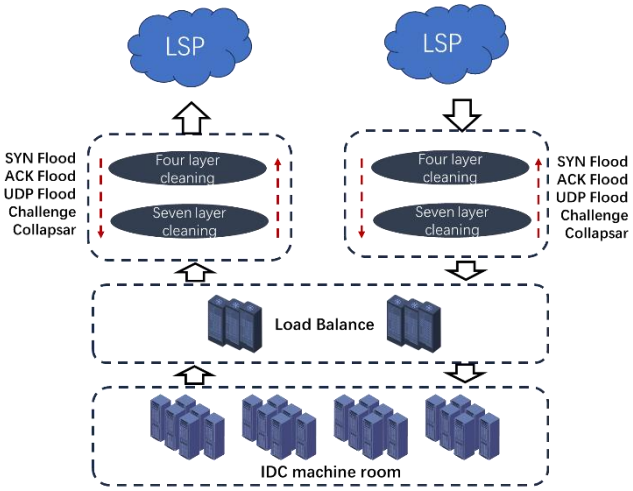

Fig. 4 The system architecture diagram

According to the network protocol layer used by DDoS flood attacks, DDoS attack detection methods can be divided into machine learning methods and traditional detection methods. Machine learning methods include traditional machine learning and deep learning. Traditional detection methods mainly rely on known rules or patterns for detection. Combining these methods, a multi-level and comprehensive rail transit network security monitoring system can be established for real-time monitoring and rapid response to DDoS attacks.

Secondly, the current open problems also have important implications for the construction and safety evaluation systems based on big data. Especially for the problems of attack detection and data sets, it is necessary to study and solve the characteristics and complexity of the urban rail transit system. By overcoming the shortcomings of traditional detection methods and improving the accuracy and response time of machine learning methods, effective defense against DDoS attacks and other network intrusions can be achieved. At the same time, the establishment of a comprehensive, credible, and high-quality network security data set, is conducive to the training and optimization, and improves the robustness and reliability of the entire security evaluation system.

5. Conclusion

For the urban rail transit security evaluation system based on big data, the research on DDoS attack detection methods and the open issues are of great significance, which can provide effective protection and protection for the network security of urban rail transit systems, and ensure that urban The stable operation of traffic and the safe travel of residents.

Acknowledgments

This work was financially supported by Jinan Rail Transit Group Co., Ltd. (Project Code: JGJS-QT-ZF-2021-075).

References

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

[2] Jing X, Yan Z, Pedrycz W. Security data collection and data analytics in the Internet: A survey[J]. IEEE Communications Surveys & Tutorials, 2018, 21(1): 586-618.

[3] Li P, Salour M, Su X. A survey of internet worm detection and containment[J]. IEEE Communications Surveys & Tutorials, 2008, 10(1): 20-35.

[4] Nisioti A, Mylonas A, Yoo P D, et al. From intrusion detection to attacker attribution: A comprehensive survey of unsupervised methods[J]. IEEE Communications Surveys & Tutorials, 2018, 20(4): 3369-3388.

[5] Mishra P, Varadharajan V, Tupakula U, et al. A detailed investigation and analysis of using machine learning techniques for intrusion detection[J]. IEEE Communications Surveys & Tutorials, 2018, 21(1): 686-728.

[6] Panimalar P, Rameshkumar K. A review on taxonomy of Botnet detection[C]//2014 International Conference on Advances in Engineering and Technology (ICAET). IEEE, 2014: 1-4.

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

[8] Zhijun W, Wenjing L, Liang L, et al. Low-Rate DoS Attacks, Detection, Defense, and Challenges: A Survey[J]. IEEE Access, 2020, 8: 43920-43943.

[9] Luo X, Chang R K C. On a new class of pulsing denial-of-service attacks and the defense[C]//NDSS. 2005.

[10] Subbulakshmi T, BalaKrishnan K, Shalinie S M, et al. Detection of DDoS attacks using Enhanced Support Vector Machines with real time generated dataset[C]//2011 Third International Conference on Advanced Computing. IEEE, 2011: 17-22.

[11] Wang C, Zheng J, Li X. Research on DDoS attacks detection based on RDF-SVM[C]//2017 10th International Conference on Intelligent Computation Technology and Automation (ICICTA). IEEE, 2017: 161-165.

[12] Vijayasarathy R, Raghavan S V, Ravindran B. A system approach to network modeling for DDoS detection using a Naive Bayesian classifier[C]//2011 Third International Conference on Communication Systems and Networks (COMSNETS 2011). IEEE, 2011: 1-10.

[13] Tan L, Pan Y, Wu J, et al. A New Framework for DDoS Attack Detection and Defense in SDN Environment[J]. IEEE Access, 2020, 8: 161908-161919.

[14] Qin X, Xu T, Wang C. DDoS attack detection using flow entropy and clustering technique[C]//2015 11th International Conference on Computational Intelligence and Security (CIS). IEEE, 2015: 412-415.

[15] Rahman O, Quraishi M A G, Lung C H. DDoS attacks detection and mitigation in SDN using machine learning[C]//2019 IEEE World Congress on Services (SERVICES). IEEE, 2019, 2642: 184-189.

[16] Saied A, Overill R E, Radzik T. Detection of known and unknown DDoS attacks using Artificial Neural Networks[J]. Neurocomputing, 2016, 172: 385-393.

[17] Kumar P A R, Selvakumar S. Distributed denial of service attack detection using an ensemble of neural classifier[J]. Computer Communications, 2011, 34(11): 1328-1341.

[18] Shaaban A R, Abd-Elwanis E, Hussein M. DDoS attack detection and classification via Convolutional Neural Network (CNN)[C]//2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS). IEEE, 2019: 233-238.

[19] Li Y, Lu Y. LSTM-BA: DDoS Detection Approach Combining LSTM and Bayes[C]//2019 Seventh International Conference on Advanced Cloud and Big Data (CBD). IEEE, 2019: 180-185.

[20] Bellaiche M, Gregoire J C. SYN flooding attack detection based on entropy computing[C]//GLOBECOM 2009-2009 IEEE Global Telecommunications Conference. IEEE, 2009: 1-6.

[21] Sengar H, Wang H, Wijesekera D, et al. Detecting VoIP floods using the Hellinger distance[J]. IEEE transactions on parallel and distributed systems, 2008, 19(6): 794-805.

[22] David J, Thomas C. DDoS attack detection using fast entropy approach on flow-based network traffic[J]. Procedia Computer Science, 2015, 50(4): 30-36.

[23] Sun C, Hu C, Liu B. SACK2: effective SYN flood detection against skillful spoofs[J]. IET information security, 2012, 6(3): 149-156.