Discussion on Network Intrusion Detection Technology

Jinning Guo[1], Jiayi Zhao[1], Yuxuan Peng[2], Hao Wu[1]

[1]Nanjing University of Information Science & Technology, Nanjing, 210044, China

[2]Nanjing Forestry University, Nanjing, 210037, China

Keywords: Intrusion detection, Machine learning, Deep learning, Internet of things

Abstract: With the rapid development of modern network technology, serious network intrusion problems have arisen and the development of network intrusion detection technology has been promoted indirectly. In recent years, IDS based detection system has been gradually applied to all walks of life, playing an important role in monitoring and prevention. The core technology of intrusion detection also changes with the development of related disciplines. Machine learning and deep learning gradually replace the traditional detection technology based on state transition and monitoring with its solid theoretical foundation and strong adaptability, and begin to provide network security services for the new network represented by the Internet of things, making the network intrusion detection intelligent and adaptive. However, the current network intrusion detection still has low accuracy. This paper introduces the principle and working mode of mainstream intrusion detection technology based on the status of network intrusion phenomenon and the current commonly used models, points out the key problems existing in the industry at present, and forecasts the solution of the problems and the future development of network intrusion detection technology.

1. Introduction

As early as the 1980s, network intrusion behavior attracted the attention of the computer network industry. Anderson defines intrusion as: unauthorized intentional attempt to access information, tampering with information, making the system unreliable or unusable [1]. Common network intrusions are mainly divided into two categories: passive attack and active attack. Passive attack refers to the intruder stealing user information or key data without the user's consent. The main forms include eavesdropping, stealing traffic data and cracking data stream. Active attack means that intruders can't use terminal services normally by maliciously changing the data stream. The main forms include tampering, forgery and denial of service (DoS).

The existence of network intrusion will affect users' normal use, and may cause serious property loss and privacy leakage. Therefore, combating network intrusion has always been a major research focus of network security technology. Traditionally, the tools to deal with network intrusion are firewalls and proxy servers. By setting up a communication monitoring system on the network boundary, the firewall isolates the internal and external networks and controls the entry and exit of boundary data, thus preventing intruders from invading from the external network. The proxy server can act as an intermediary for the external network to apply for access to the internal network. Through the proxy server, you can hide your IP address from the intruder, so that the intruder can't achieve the purpose. However, with the continuous evolution of network intrusion behavior, the traditional defense methods can not fully meet the security needs of users. Therefore, Intrusion Detection System (“IDS”) has gradually replaced firewall and become the mainstream network security protection technology. IDS monitors the computer system, and alerts the user when it finds any abnormality. Later, IDS gradually developed into Intrusion Detection Expert System (IDES), and on this basis, many branches were derived and applied to different detection environments, among which the industrial control intrusion detection system (ICSIDS for short) was the most influential one [3]. As show in fig.1.

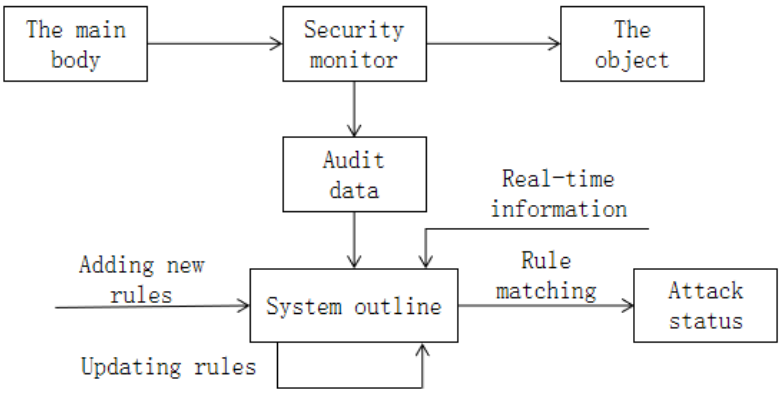

Fig.1 Basic Working Principle of Ides

With the rapid development of modern network technology, especially the Internet of Things technology, the number of users and devices accessing the Internet in the world has increased dramatically, and the forms and coping strategies of mainstream network intrusion have also changed significantly. Invaders are no longer limited to kidnapping a single user or device, but instead use an intrusion mode with a wider scope of damage, among which the most famous and representative one is Distributed Denial of Service (DDoS). On October 21st, 2016, Dyn DNS Company, which provides dynamic DNS services, suffered a large-scale DDoS attack. This attack made many websites in the Eastern District of the United States unable to access normally. Twitter, Spotify, Reddit, SoundCloud and other websites were severely attacked, and Twitter even had 24-hour zero access. Traditional anomaly intrusion detection has been unable to resist various invasion. However, the large-scale intrusion behavior also provides sufficient samples for researchers. This condition gives machine learning and deep learning technology the opportunity to make great achievements in the field of network information security. Since then, researches on the application of machine learning and deep learning in the Internet of Things have emerged in an endless stream, and the judgment accuracy of related technologies for network intrusion has been increasing, showing the trend of replacing traditional algorithms.

2. Application of Network Intrusion Detection Technology

2.1 Traditional Intrusion Detection Technology

Traditional intrusion detection technologies are divided into two types: Anomaly intrusion detection and Misuse intrusion detection. Abnormal intrusion detection refers to establishing the normal mode profile of the system. If the difference between the profile value of the system or user obtained in real time and the normal value exceeds the specified threshold, an intrusion alarm will be given. Misuse intrusion detection refers to detecting intrusion according to known attack characteristics, which can directly detect intrusion behavior [2]. Anomaly intrusion detection is generally based on feature selection, Bayesian inference, Bayesian network, Bayesian clustering or data mining. Its advantage is that it does not depend on the characteristics of intrusion behavior, but finds anomalies from the perspective of users and terminals. Misuse intrusion detection is generally based on state transition analysis, keyboard monitoring or self-defined rules. Besides, there is a hybrid intrusion detection system which combines the characteristics of two intrusion detection technologies in practical application. Traditional intrusion detection technology is a cornerstone in the early development of IDS, and it has played an important role in the field of network information security from last century to the beginning of this century. However, with the development of new network technology and the increasing demand of information security, these traditional detection methods and classical algorithms generally show the problems of low efficiency and low success rate in the face of today's complex intrusion methods, so they are no longer applicable to the current intrusion detection system.

2.2 Application of Machine Learning in Intrusion Detection Technology

Machine learning, with its strong adaptability and self-learning ability, provides an effective analysis and decision-making tool for information security. It solves the problems of low efficiency and poor adaptability of traditional methods. The growing Internet system provides a large number of data samples, which makes it possible for machine learning to be used in intrusion detection. Zhang Lei et al. [4] elaborated a series of ideas and solutions for the application of machine learning in network security. Song Yong et al. [5] Based on information theory and big data technology, provided a method to extract effective intrusion detection data from data samples and normalize them. Inspired by large-scale DDoS attacks, Doshi et al. [6] and Bhatt et al. [7] collected intrusion data by simulating the Internet of things environment and simulating DDoS attacks, and used a variety of machine learning classification algorithms (decision tree, support vector machine, random forest, etc.) for analysis, and finally achieved a detection success rate of 94-99%. He Xiang et al. [8] analyzed the performance of different machine learning algorithms from the perspectives of training time, false alarm rate, detection ability of unknown intrusion information, and memory occupation of modules, and elaborated the application fields of various algorithms.

The role of machine learning in information security has been widely recognized and put into practical production, which is the main force of intrusion detection technology.

2.3 Application of Deep Learning in Intrusion Detection Technology

Although machine learning realizes the intelligence of intrusion detection technology, its high success rate actually depends on the known knowledge,the training of a large number of intrusion data does not have the ability of continuous learning. However, in the background of network intrusion behavior diversification and strong suddenness, the single machine learning technology may not be able to adapt to the improved intrusion.

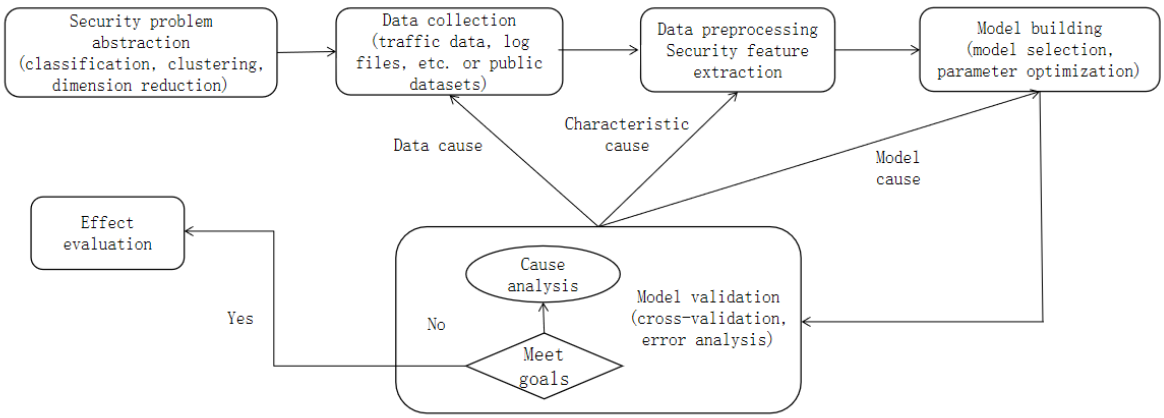

Fig.2 Principle of Intrusion Detection Using Machine Learning

In fig.2, Deep learning is a feasible solution to this problem. The purpose of deep learning is to achieve the goal of machine learning and the development of machine learning system. In terms of network security, deep learning can be used to continuously learn new intrusions, so that users and terminals can get long-term and reliable protection. Zhang Yuqing et al. [9] analyzed the research status and progress of deep learning applied in Cyberspace Security from the aspects of classification algorithm, feature extraction and learning effect, and summarized the problems and opportunities of deep learning in related fields. Vinayakumar et al. [10] discussed a deep learning model (DNN), and comprehensively evaluated the experiments of DNNS and other classical machine learning classifiers on various open intrusion data sets. Through strict experimental tests, it is proved that DNNS has better performance than classical machine learning classifiers.

Compared with machine learning, deep learning has the ability of continuous learning, but it has disadvantages in classification absolute accuracy, training speed, module occupied memory and

other performance indicators, and the cost of actual production is high. Therefore, it has not completely replaced the status of machine learning in intrusion detection technology. It can be predicted that in the future network security technology, the application of deep learning will continue to develop and play a greater role.

3. Problems in Current Intrusion Detection Technology

3.1 It is Difficult to Maintain a High Level of Detection Accuracy

Although in Doshi's research, the detection accuracy of several classical machine learning classifiers has reached more than 99%, but the detection accuracy of machine learning is still not guaranteed. On the one hand, Doshi's machine learning data set design is not reasonable. Among the nearly 500000 data packets used for training model, less than 7% of them are normal, and the rest are all malicious packets for intrusion. In such a large proportion, the model is prone to over fitting problems, and the actual detection accuracy will be discounted. On the other hand, Doshi's data collection work is completed in a short time, and the training model can't cope with other versions of intrusion packets, which may not be applicable now. While deep learning also has the problem of accuracy. In Vinayakumar's deep learning model, the detection accuracy is lower than that of Doshi's machine learning model. However, in actual large-scale DDoS attacks, the number of malicious packets is extremely large, so the sacrifice of the absolute accuracy of deep learning may make its overall reliability questionable.

3.2 Users Lack of Coping Means

Even if the existing intrusion detection technology can detect the intrusion correctly, how to deal with it is a thorny problem. The detection of machine learning and deep learning can only detect intrusion in essence, and can not accurately describe the possible consequences of intrusion. At the same time, users are often lack of network security related knowledge reserves and processing experience, unable to make correct decisions for system prompts. This disadvantage is more prominent in the development trend of “Internet of things”: in the future, most terminals will be connected to the existing network, including many precision instruments and key equipment, such as the transfer device of large power station, and the instrument for monitoring the patient's condition. The operation of these instruments will lead to a series of serious consequences. At the same time, the focus of this research is to provide users with more efficient intrusion detection and in-depth learning.

4. Future Prospects

In the next few years, the Internet of things technology will continue to develop at a high speed, and the scale of network and data packets used by the whole society will continue to grow. Artificial intelligence technology represented by machine learning and deep learning will still be the main research direction of intrusion detection technology. Considering the development of the industry, I think the most important problems to be solved are:

How to improve the detection accuracy?

How to reduce the loss caused by network intrusion as much as possible?

In my opinion, under the premise of machine learning, in order to improve the detection accuracy, it is necessary to ensure that researchers can obtain enough data samples. In previous studies, researchers generally use temporary network models, and the data packet samples collected can not be compared with the actual situation in terms of quantity and type. Therefore, professionals in the industry should establish an open source, standardized database as soon as possible to meet the needs of research. Secondly, optimizing the index selection and parameter setting of machine learning may also make the detection accuracy rise to a higher level. In general, the prospect of improving the detection accuracy by optimizing the current technology is relatively clear.

As for how to reduce the loss, I think a potential solution is to classify the harmfulness of intrusion behavior according to its characteristics, and display the classification results to users, so as to improve the probability of users making correct decisions. However, it is difficult to implement

this scheme, because the essence of classification is to classify intrusion packets again, and the number of classification is more, the performance of the classifier needed is naturally higher. However, the accuracy of the existing classifier model in dealing with the simple binary classification of "whether it is a malignant packet" is still not satisfactory. Obviously, the current technology can not complete the task of hazard classification. Therefore, the most powerful means of prevention is to improve the relevant legal system and improve the safety awareness of users.

References

[1] ANDERSON J P. Computer Security Threat Monitoring and Surveillance[R]. James P Anderson Co, Fort Washington, Pennsylvania, 1980.

[2] Qing Sihan, Jiang Jianchun, Ma Hengtai, Wen Weiping, Liu Xuefei. A review of intrusion detection technology [J]. Acta communica Sinica, 2004 (07): 19-29

[3] Yang An, Sun Limin, Wang Xiaoshan, Shi Zhiqiang. A summary of intrusion detection and measurement technology in industrial control system [J]. Computer Research and Development, 2016,53(09):2039-2054.

[4] Zhang Lei, Cui Yong, Liu Jing, Jiang Yong, Wu Jianping. Application of machine learning in the research of network space security [J]. Journal of Computer Science, 2018,41(09):1943-1975.

[5] Song Yong, Cai Zhiping. Data normalization method of intrusion detection based on information theory in big data environment [J]. Journal of Wuhan University (Science Edition), 2018,64(02):121-126.

[6] R. Doshi, N. Apthorpe and N. Feamster, "Machine Learning DDoS Detection for Consumer Internet of Things Devices," 2018 IEEE Security and Privacy Workshops (SPW), 2018, pp. 29-35, doi: 10. 1109/SPW. 2018. 00013.

[7] P. Bhatt and A. Morais, "HADS: Hybrid Anomaly Detection System for IoT Environments," 2018 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC), 2018, pp. 191-196, doi: 10. 1109/IINTEC. 2018. 8695303.

[8] He Xiang, Zhan Liu, Jiang Jiguo. Comparative study of intrusion detection methods based on machine learning [J]. Information Network Security, 2018(05):1-11.

[9] Zhang Yuqing, Dong Ying, Liu Caiyun, Lei Kenan, Sun Hongyu. Current situation, trend and prospect of deep learning applied to cyberspace security [J]. Computer Research and Development, 2018,55 (06): 117-1142.

[10] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," in IEEE Access, vol. 7, pp. 41525-41550, 2019, doi: 10. 1109/ACCESS. 2019. 2895334.