ISSN: 2664-9640

Research on Internet Anomaly Monitoring Model Based on Big Data Environment

Jie Fang [a], Qingchun Fan

School of Computer Science and Technology, Hefei Normal University, Hefei, 230601, China

[a]fangjie@hfnu.edu.cn

Abstract

In the complex and ever-changing Internet situation, the base number of monitoring and management data required for network security has increased geometrically. Network attacks have new trends and new characteristics, such as more covert attacks, more changeable behaviors, long attack latency cycle and the emergence of new attack methods. The traditional network anomaly detection technology is facing the dilemma that it can not effectively deal with, and it is urgent to solve the low efficiency of existing anomaly detection, New attacks cannot be detected. In order to quickly and accurately respond to the situation awareness needs and network security threats in the big data environment, relying on the functional advantages of big data technology and machine learning, an Internet anomaly monitoring model is proposed to realize the needs of rapid collection, intelligent perception, accurate analysis, convenient storage and rapid query.

Keywords

Big data; internet; Anomaly monitoring technology; Multi source heterogeneous data environment.

1. Introduction

With the in-depth promotion of Internet application technology and the opening and easy availability of the underlying technology, the Cyberspace Security Situation and network security are becoming more and more serious. Based on strategic considerations, the country has built firewalls at all levels through the "moat", but in the face of new technologies and new means, the national cyberspace security situational awareness and prevention and control power have been gradually weakened. At present, although Internet-based applications have touched and extended to all levels of social and economic life, we should see that the data structure and level of a country with a population of 1.4 billion are extremely complex, and the underlying cyberspace data are extremely complex. The current network security situation awareness and anomaly monitoring system, or the macro Cyberspace Security Protection System Based on it, has no response strategy or better solution to deal with the scattered and complex security problems of the underlying network. With the passage of time, the foundation of network security will become increasingly unstable, and the foundation of the overall interests of the country will be affected. It is urgent to consolidate the foundation of network security, resolutely prevent and resolve the physical risks of network security, and actively resist the national cyberspace security crisis.

With the advent of 5g and big data era, problems such as network and information security and privacy of public infrastructure telecommunications facilities and individuals are becoming more and more prominent. Changeable and diverse attack behaviors are overwhelmed. New attack methods emerge in endlessly, such as apt attack and Zero Day attack. The information security events caused by such attacks are becoming more and more prominent and severe. The emergence of such problems reflects that there are urgent problems to be solved in the

ISSN: 2664-9640

current security situation awareness and security technical means. Firstly, the information infrastructure, security standards and architecture of Telecommunications enterprises do not match the development of existing attack technologies, resulting in the complexity and high coupling of information data. Secondly, the existing network security situational awareness and data analysis capabilities can not effectively match the security monitoring requirements of multi-source heterogeneous data in the big data environment; In addition, for the constantly updated new attack means and technologies, the traditional network security monitoring, analysis and early warning model can not make accurate judgment and timely response, resulting in high delay in finding problems.

The problems in the above aspects are becoming more and more obvious, which urgently needs to be paid attention to and solved by relevant industries. By optimizing and transforming the existing data infrastructure and security architecture, we can improve the security application value of data. This paper takes the abnormal data monitoring based on big data as the research goal, draws lessons from and considers the characteristics of the current distributed data architecture in the Internet industry, and puts forward the Internet abnormal data monitoring model to improve the efficiency of Internet abnormal data monitoring, perception and analysis, so as to save the cost of hardware resources and build a distributed data cloud architecture. Based on the big data analysis and processing technology, by introducing measurable data monitoring and traceable analysis algorithms, Support the actual needs of refined security analysis, identify hidden and complex security vulnerabilities from the detection stage, so as to provide accurate judgment and analysis of subsequent attacks and abnormal behavior data, and continuously improve the level of network security situational awareness and network security monitoring.

2. Research objectives

Research on the collection and analysis of massive multi-source heterogeneous big data on the Internet. Based on various technical means such as security attack and abnormal behavior monitoring, big data collection and preprocessing, machine learning and so on, explore and implement the research on security data monitoring and mining in information security situational awareness in a forward-looking and scientific way. Through the integration of security integration event analysis and event linkage disposal, Build an abnormal security behavior management index system, scene and quantitative evaluation system model with information security specific indicators, conduct accurate event correlation analysis, real-time scene perception and tracking of security events, and effectively trace the source, so as to realize the situation awareness, visual management and panoramic display of abnormal behavior and risk events. Further improve the security and stability of situation awareness system and information system, realize the comprehensive control of network security risks and improve the level of linkage disposal, and provide power guarantee for the development and innovation of information security and situation awareness.

The key problems to be solved in the research process of anomaly monitoring based on big data include: (1) implement cross system security data collection, mining and event analysis based on internal gateway. (2) Linkage technology analysis of external attack or abnormal security behavior and internal system security vulnerability. (3) Index system standards based on abnormal security behavior situation analysis, as well as design matching implementation scenarios and rehearsal models. (4) Cross domain and cross department strategic coordination of multi type and multi-dimensional security technology, and qualitative and quantitative evaluation criteria for abnormal security behavior. (6) Dynamic monitoring display of abnormal safety behavior based on multi-dimensional perspective.

ISSN: 2664-9640

3. Anomaly detection technology

As a widely used technology in the field of information security, anomaly detection is usually used to detect abnormal behavior or abnormal data that violate statistical rules. With the advent of the era of "Internet of things", wireless sensor networks are widely used in Internet of things applications because of their flexibility [1]. A large number of tens of thousands of sensors may be deployed in the core backbone equipment. At present, more than 100 sensors of the new Tesla car are installed and used for real-time remote communication with the core controller of the car to assist the needs of automatic driving. The sensors capture the detection data from the internal and external observation of the car in real time, which are used to perceive and judge the abnormality, so as to predict and analyze the automatic assisted driving. Some sensors transmit data through the Internet of things channel, which has the risk of being intercepted and tampered with [2].

Since most traditional anomaly detection methods are highly dependent on the category of known patterns, whether normal patterns or abnormal patterns. When the detected behavior matches the known abnormal pattern, it is determined and recognized as abnormal behavior. When a new stealth attack abnormal behavior occurs, its existence is likely to be incorrectly detected as normal[3]. In view of this, some scholars propose to introduce the method of deep reinforcement learning in some anomaly detection strategies to update the existing anomaly patterns by learning from emerging behaviors.

However, the data changes from sensors and the coupling correlation measurement of data changes are often ignored[4]. Suppose a humidity sensor deployed in a smart agriculture sends back an abnormal humidity value alarm to the management node of the data monitoring center at a certain time, but then it will send back a series of normal reference values. This anomaly may be caused by raindrops or dew acting on the humidity sensor. When raindrops or dew drops evaporate, the humidity value will naturally return to normal. Obviously, it is not necessary for this situation to trigger an alarm. However, if the humidity sensor constantly sends abnormal values back to the management node of the data monitoring center, this may be a real alarm. Therefore, it is urgent to propose heuristic anomaly monitoring models and monitoring methods to accurately and scientifically detect abnormal behaviors and events [5].

4. Network anomaly analysis model based on big data

The core of the network anomaly analysis model based on big data is to meet the needs of network anomaly analysis models such as big data collection, data preprocessing, situation analysis, data streaming processing and cloud storage by providing a large number of multi-source heterogeneous security data collection and storage solutions, based on distributed message queue, distributed offline analysis module and streaming processing module, and relying on machine learning algorithms. The network abnormal behavior analysis architecture based on big data can be composed of five layers: big data acquisition layer, data preprocessing layer, machine learning and big data analysis layer, data storage layer and situation awareness and visualization layer[6]. The low coupling between each layer is independent of each other. Even if there is a problem with the components of one layer, the impact on the components of other layers will be minimized. The standardized format design makes the data format of each level unified, and the standardized interface is used for calling between each layer. The user configures the cluster development of each component, monitors and collects the network data flow, through a series of feature extraction and preprocessing processes, through classification and screening based on the network abnormal behavior detection model trained by intensive learning, distinguishes between normal behavior and different attack behavior, and displays the detection results and query results to the user.

ISSN: 2664-9640

5. Design and implementation of behavior precise perception module

With the continuous integration and development of the interconnection of all things, the complexity of Telecommunications infrastructure information system is becoming more and more prominent. The accumulation of factors such as different quality of employees, outsourcing introduction, loopholes and defects in system design makes the threat of abnormal behavior to the information ecosystem more and more serious[7]. Based on the goal of controlling abnormal behavior faster and more accurately, this paper takes abnormal behavior monitoring and situation awareness management as the main line, through the construction of information system in the whole life cycle, from the demand design and implementation of business system to the completion of system model construction, and then to the whole process of network access test and acceptance, operation and maintenance management of monitoring system, and the implementation of situation awareness and monitoring of network monitoring and core business system, Organically integrate intelligent sensing strategy, big data acquisition, preprocessing and analysis, machine learning abnormal behavior quantitative evaluation model, data visualization, situational awareness and other technologies to integrate the risks in the process of information construction. At the same time, it is reconstructed from multiple perspectives such as management, monitoring and operation and maintenance, and uses multi-dimensional methods to carry out all-round situational awareness and view display of network security risks, so as to realize the transformation from passive to active, Switch from passive information security management to active information security management, and continuously optimize and improve the ability of accurate control, dynamic perception, decision optimization and continuous improvement of abnormal behavior. The contents to be solved in this module include:

(1) Starting from the goal of global risk health algorithm analysis, the model of system integrated operation and maintenance is established.

(2) From multiple perspectives and dimensions such as region, business system, data type and time cycle, the whole process security situational awareness is displayed in an all-round combination.

(3) Based on risk visualization, implement the whole life cycle control and management of data from monitoring and display to disposal tracking.

(4) Realize the distribution display and detailed description of security problems, and implement them from the whole bank's internal and external abnormal behavior, security events, vulnerability, compliance problems, alarms, terminals (including ATM) and other elements.

(5) Online path restoration and traceability display are implemented through internal and external abnormal behavior attacks.

6. Design and implementation of abnormal behavior detection and analysis module

Hacker attack methods have quietly changed from "short-term publicity", "extensive attack", "using existing tools" and "brute force cracking" to "long latency period", "increasingly accurate directional attack", "customized attack tools" and "increasingly refined DDoS", among which apt (advanced persistent thread) attack is the most representative. Its attack harm is huge and destructive. If we can effectively identify and respond in the early attack link, it will help to suppress the attack and reduce the loss in the later stage. Lockheed Martin Space Systems Company has a strict definition of the link of security attack. The killing chain model is described as follows: the link of security attack has seven steps, namely (1) deception and reconnaissance (2) weaponized (3) delivery (4) using attack (5) installing (6) command and

ISSN: 2664-9640

control (7) leakage. At present, most technical means are for the discovery attack in step 6 and 7. For the first step of the model, there is no mature theory and prototype.

The researchers found that by advancing the steps of discovering attacks to steps 4 and 5, we can shield the attacks of hackers and minimize the free attack time of hackers, so as to quickly improve the protection system to the optimal state and carry out targeted and effective defense against abnormal behaviors and attack sources. Abnormal behavior detection and analysis is mainly divided into the following steps: internal and external abnormal behavior information collection, search and sharing, early warning and alarm based on analysis and situation awareness, automatic upgrading of security capability and unified output of abnormal behavior[8].

7. Design and implementation of abnormal behavior attack path traceability module

In a broad sense, attack path tracing refers to the process of restoring attack methods through the characteristics of security events in the network and finally tracing the attacker[9]. Tracing back the attacker and thoroughly investigating the weak points are also the core purpose of tracing the source of the attack path. The core of attack traceability includes process traceability and network traceability. The final output of process traceability is the overall operation steps of the attack, which is used to show the attacker object, attack purpose and attack process, while the final output of network traceability is the network topology, which realizes the perception and display of attack means and attack tools. Security event warning is the core dependence of attack traceability. The attacker's social attributes and geographical location depend on the perception and monitoring of abnormal behavior information.

8. Design and implementation of abnormal behavior perception computing module

In the Telecommunications enterprise information system, the data is multi-source and heterogeneous, and the scale is large, but there are few effective abnormal information data that can be monitored smoothly. How to mine the valuable data in the massive big data, and then find the hidden abnormal behavior and attack behavior is the focus of this paper. Abnormal behavior perception calculation depends on business operation behavior. Through the mining operation of various log data, starting from different dimensions, the operation log data is deeply mined and statistically analyzed based on statistical analysis, association analysis, cluster analysis, time data mining and model analysis, and finally the business operation behavior is displayed to form a business operation baseline[10]. Then, by comparing with the operation baseline, we can actively find and identify suspected abnormal behaviors, prevent attacks, and give timely warnings to suspected abnormal behaviors, so that business personnel can deal with abnormal behaviors in time.

9. Conclusion

This paper collects and analyzes the massive multi-source heterogeneous data of the whole network, comprehensively uses various technologies such as security threat early warning calculation, big data collection and preprocessing, machine learning and so on, and prospectively and scientifically explores the technologies and methods of security data mining, security monitoring and application integration, security event analysis and situation awareness, and security event linkage handling in information security situational awareness, Build an abnormal security behavior management index database, scene database and quantitative evaluation model with the characteristics of the telecommunications industry, so

ISSN: 2664-9640

as to carry out intelligent correlation analysis of security events, achieve visual management of abnormal behavior and panoramic display of risk situation with the implementation of scene perception and tracking traceability, further improve the comprehensive control and linkage disposal level of security operation risk of information system, Escort for the development of telecommunications business, Provide power guarantee for innovation leading.

References

[1] Yan Zhi, Zhan Jing. Big data application mode and abnormal security behavior analysis. Computer and Modernization, 2014: 2-3
[2] Yin Chuanlong. Research on network anomaly detection technology based on deep learning [D]. Strategic Support Force Information Engineering University, 2018.
[3] Chen Xingshu, Zeng Xuemei, Wang Wenxian, Shao Guolin. Network security and intelligence analysis based on big data [J]. Engineering Science and Technology, 2017, 49(03): 1-12.
[4] Dong Guishan, Wang Zheng, Liu Zhenjun. Analysis of Digital Ocean System and Security Requirements Based on Big Data. Communication Technology, 2015: 2-3
[5] Zhang Bin. Application of Big Data Analysis Technology in Security Field. Telecommunications Engineering Technology and Standardization, 2015: 4
[6] Li Ruopeng. Design and implementation of network abnormal behavior detection platform based on big data [D]. South China University of Technology, 2018.
[7] He Bin, Yang Yuanjin. Research Review of Big Data Security Analysis and Management Platform. Network Security Technology and Application. 2016: 3
[8] Liao Minghui. Data analysis and information security protection in the era of big data. Electronic Technology and Software Engineering. 2016: 1
[9] Lin Ning, Zhang Liang. User information security and privacy protection in the era of big data. Information Technology. 2016: 1
[10] Dong Na, Liu Weina, Hou Botao. Modeling method of abnormal network behavior based on big data [J]. Electric Power Information and Communication Technology, 2018, 16(01): 6-10.
[11] Yang Xi, GUL Jabeen, Luo Ping. Big Data Security Technology in the Cloud Era. ZTE Technology. 2016: 2-3
[12] Lu Xin, Han Xiaolu. Research on Big Data Security and Privacy Protection Technology Architecture. Information Security Research. 2016: 7
[13] Belouch M, El Hadaj S, Idhammad M. Performance evaluation of intrusion detection based on machine learning using Apache Spark[J]. Procedia Computer Science, 2018, 127: 1-6.