Information Security on the Background of Big Data

Huijie Yang

Xi'an Peihua University, Xi'an, China

395436613@qq.com

Keywords: big data; information security; firewall technology

Abstract. The purpose of this paper is to expound the big data problem under the background of information security; First of all, from the perspective of the concept of big data, talked about in the context of "Internet +", the importance of information security. Secondly talked in big data networks, some problems such as information disclosure, and then what measures should be taken to for information security problem.

1 Introduction

In cloud computing, mobile Internet and the development of the Internet of things, people's life and production of data also show a geometric times growth trend; According to the international Internet Data Center, IDC (Internet Data Center) is expected, the existing within the Internet, more than 90% of the amount of Data is produced in recent years, and every two years, would double the speed of growth, by 2020, the global Internet Data quantity will increase 50 times. This suggests that, the Internet economy has already marched into the era of big data. In this case, the data is huge quantity. Have many characteristics, such as huge amount of information, information diversity present and update speed is very fast, etc. There are a lot of sensitive information and confidential information, it is easy to appear the information disclosure, virus attacks, data add, delete information theft and data tampering. At the same time, the network entities also are vulnerable to earthquakes, fires, floods, the effects of electromagnetic radiation; For example, in September 2014, the United States apple up "to" bump library "by hackers invasion, cause" the largest in the history of Hollywood yan zhao door "event. International hacking group anonymous attack north Korea many times website and cause the paralysis, and even cause the member account information. Thus, for individuals, businesses and even countries, how to do big data under the background of information security is a very important thing.

2 The Concept and Development

According to the Wikipedia definition, data that cannot be within a certain time with commonly used software tools to capture, management of data collection. Data what characteristics do you have?

2.1 Data Size Large (Volume). In today's world with index of the amount of data that are the size of the growth; According to the survey of the international data corporation (IDC), the data on the Internet will grow by about 50% a year, and now more than 90% of the world's data is created in recent years. Total data growth far more than the speed of computer hardware and software technology development, so that caused a crisis of data storage and processing.

2.2 Various Data Types (Appearance). Big data is the source of multi-source heterogeneous, their data types and formats increasingly rich and colorful, already broke through the previous limit of structured data category, include the semi-structured or unstructured data, through calculation, more than 95% of the data are semi-structured or unstructured data. In the past, the data can be stored in any other way and save, but now I have to do to save them, not only to analyze them, the purpose is to get more value.

2.3 Low Density Value (Value). Value of high and low density is inversely proportional to the total data size. How to use powerful machines algorithm to complete the value of a large amount of data \"purification\", is the problem urgently to be solved in the era of big data.

2.4 The Data Authenticity (Veracity). The content of the large data is related to happen in the real world. Research data is extracted from vast network data can explain and predict the process of real events.

2.5 Processing Speed Is Fast (Velocity). With the advent of the era of big data, in processing speed of data generated in the qualitative leap, this is the \"big data\" the most prominent feature distinguish from traditional data mining, it can get a faster way to meet the real-time requirements.

At present, various countries attached great importance to development of data. Have been issued on "big data" of national policies and strategies, and attaches great importance to the "big data" research and applications.2011 Japan launched a new comprehensive strategy after the earthquake "vibrant ICT Japan", focusing on "big data" the application of intelligent technology development. In 2012, the United States Government announced that it would invest $ 200 million to start the "big data research and development initiative", the United Nations has launched a "global pulse" project, hoping to promote the development of the global economy. In recent years, the large development of macro policy environment and constantly improve.2012, the national development and Reform Commission, Ministry of science and technology, Departments such as the Ministry of science and technology and supports a number of data-related projects in the areas of industrialization, promote technological research and development to achieve positive results.2014 clearly proposed in the Government work report, setting up new business innovation platform, in areas such as data catch up with the advanced, leading the future development.

3 Big Data Information Security Threats

As the big data technology in national defense, energy, aerospace, medical and other fields to present explosive growth, diversity, information has been penetrated into every corner of social life, closely combine with various fields, to the national information security and personal security brought serious threats and challenges.

3.1 The Threat Information Content. Under the background of big data, information content security is mainly two modes: information disclosure and information. With the development of e-commerce and mobile Internet behavior, the situation of the information received against more subtle than before; Therefore, to prevent the data are destroyed, tampered with, disclosure or steal the task is very difficult. data sources in the network space is very complex; On the one hand, a large amount of data collection, increases the risk of data leakage; Some sensitive data, on the other hand, the ownership and use right is no clear definition of the rights of the data boundary is fuzzy, many based on large data analysis for considering the personal privacy, which involved in unauthorized unwittingly once leakage or destruction of consequences.

3.2 Big Threats to Data Storage Carrier. Under the background of big data, the data mostly centrally stored together after, which makes it easier to "discover" in the Internet space, easy to become the first target of the hacker attacks. A large number of data files for storage and processing in the third party platform, its security is under great challenge. Although can through access to documents and authorization for more protection, but the protection mechanism itself is a problem that, they mostly depends on the security of the system itself, simple authentication, together with the deepening of social engineering in the field of invasion and the existence of security vulnerabilities and constantly emerging, vehicle safety are at stake.

3.3 Big Threats to Data Management. Big data is a dynamic process, involved in the role of the large number of more every day. In the communications industry, for example, the data are usually scattered in many systems, information source is very complex. When foreign cooperation, operator interaction and cooperation in research, and the development company exist a large number of data access. If not established in the process of data open to the public access management mechanism. How to effectively protect user privacy, prevent enterprise core data reveal that become meaningless, once a problem, not able to network attacks or internal personnel irregularities, such as real-time detection, monitoring, reporting and early warning. At the same time, when the accident happens,

can't provide the hacker attacks based on clues to track and solve the lack of control network and can review. This management oversight, resulting in loss of data can be searched without a trace, for data security management.

3.4 Intelligent Terminal Data Are Faced with The Threat. In the financial sector, for example, financial information, the network will necessarily promote the financial information system through the Internet connected to terminal intelligent devices; To participate in the financial information system of collection, storage, transmission and processing, information will be more and more; In the data exchange with the external terminal equipment, the already closed network opening to the outside world, will no doubt increased the rate of intrusion and attack. Intelligent terminal, therefore, data acquisition, storage, transmission, processing will increase financial information under the threat of attack on the other hand, has promoted the value will result in more data sensitivity analysis data transfer between mobile devices; Some malicious software or even have a certain data upload and monitoring function, can track the user's location, steal data or confidential information, a serious threat to information security of the individual, make the safety accident levels.

4 Problems of Information Security in the Context of Big Data

4.1 Data of Network Attack. Due to the large data size of the data in the Internet cloud is stored in a distributed storage form, have formed a unified view of data, storage, data protection is relatively simple, is easy for hackers to attack vulnerabilities, more in implementing high threat APT attacks by hackers, causing security problems. Due to large end users in the data environment is very large, and complex groups, so the system is difficult to fast real-time judgment on the legality of Internet users.

4.2 Public Non-rational Treatment. People love to play in fashion now, what "Sun" information, micro-micro-shopping, grab a red envelope. Life has become a "self-quantitative" model, these habitual actions, however, are likely to reveal your personal information, resulting in unnecessary losses. In a period of transition from traditional data and the age of big data, information leak also flooded.

4.3 Defects of the Traditional Security Technologies. Ta available within the Internet produced in recent years and growing doubling every two years, and by 2020, global Internet data will increase 50 times. This suggests that the Internet economy has entered the age of big data. Here refers to the large amount of data. With many features, such as pluralism presents enormous stockpile, information and updates very quickly and so on. There is a great deal of sensitive and confidential information, it is prone to leaks, virus attacks, by deleting the data added, information theft and data tampering. Meanwhile, the network entities are also very vulnerable to earthquakes, fires, floods, the influence of electromagnetic radiation. For example, in September 2014, United States Apple icloud hackers "hit" means invasion, triggering "Hollywood's biggest Zong Yan Zhao gate". International hacking group "anonymous" repeatedly attacked North Korean Web sites and leading to paralysis and even member account information. Therefore, for individuals, businesses and countries, how to do well the information security in the context of big data is a very important thing. At present, the network security threats emerge. Network hackers always attack were "innovations", edit the killing power powerful programming. So as to achieve the purpose of Internet server gets the data. Traditional firewalls and antivirus software only to deal with the invasion of the mobile device side means for addressing the hacker attacks on the cloud database.

4.4 Data Storage. Large data types and data structures are unmatched in traditional data, big data storage platform, is non-linear and even the amount of data growing exponentially, of every type and structure of data for data storage, it will trigger a variety of concurrent application process running and frequently out of order, vulnerable to dislocation data storage and data management, for large data storage and later pose a security risk. Current data storage management system that can meet the largest data of huge amounts of data in the context of data storage needs, remains to be tested.

4.5 Data Value of Risk. Due to the large value of the fundamental properties of low density data, hackers will make written attacks APT code, and place hidden in the data, making detection

protection software cannot be detected; Because of rapid identification code, and tend to ignore the virus code, will transmit the virus to big data in the cloud space Server; Through user information about small, to mine user information and profile.

5 Information Security Solutions in the Context of Big Data Analysis

5.1 Active Firewall Technology. Network provides us with fast, convenient and reliable service at the same time, also brings network security problems that cannot be ignored.

Not only to protect the safety of computer network equipment and computer network system of computer network security, and data security.at present, using a broader range of technical Firewall network security technology and intrusion detection technologies.

(1) What is a firewall? Firewall refers to isolation in local network defense system between the network and the outside world, this category of preventive measures in General.

Firewall products reflect a distributed architecture, a system of distributed firewall products designed to network nodes for the protection of objects, objects that can cover the maximum protection, greatly enhance the safety protection strength.

(2) The main technologies used by the firewall: packet filtering technologies packet filtering, proxy server technology, state inspection technology, network address translation, personal firewall technology.

Packet filtering technology is the core of the security policy, it can stop and check all incoming and outgoing data.

Proxy server technology: proxy server located between the client and server, completely blocking the exchange of data between them. Due to external systems with no direct data channel between internal servers, external malicious practices made it hard to compromise to the corporate network. Proxy Firewall usually supports some common application protocols are HTTP, SISSL, SMTP etc.

State inspection technology is based on the connection state detection mechanism is used, all living in the same connection the data flow of the package as a whole, constitute the connection status table, through the co-operation of rule tables and State tables and the various connecting factors to be identified in the table, table with the traditional static packet filtering firewall filtering rules, it has better flexibility and security.

5.2 Positive Education and Guidance to Improve Data Collection, Sharing and Use. Strengthen the education of general public to enhance its data collection, usage data, in order to enhance public confidence in big data.

5.3 Intrusion Detection, the Ability to Effectively Enhance the Security Management System. Intrusion detection is a type of active defense system can not only monitor operation, external and internal attacks, effective against network attacks, supplement the inadequacies of the firewall, it can combine with other network security products for network security for a reasonable amount of protection in order to protect the integrity of the information.

Common intrusion detection systems are of two types: IDS system based on network and host-based IDS systems.

Network-based IDS system is effective in scanning and DOS attacks by hackers;

However, switched and ATMA network, this model is not ideal; And the system is unable to monitor the other subnet across a router in attack; The system illegally tampered with login, Trojan attacks, account passwords, log file tampering attacks is not very effective.

Host based IDS system: is today one of the most important and dynamic security technologies, this technology through the research on intrusion behavior and characteristics, the security system for intrusion and invasion to make real-time responses.

5.4 Scientific Encryption, Strengthening the Safety and Security of Data Storage and Transmission. Encryption refers to encryption key and encryption algorithm is the process of turning data into ciphertext. It can enhance security and concealment of information data; Can also guarantee

the security of data transmission, even if the information is intercepted or stolen, criminals find it difficult to get the information.

5.5 Enhance the Safety Protection Technique to Guarantee Network and Information Security. One is to prevent APT attacks. With large data technology, according to the characteristics of APT attacks, designed with the capability of real time detection and subsequent backtracking full flow detection, application monitoring of hidden viruses.

Second, user access control. Depending on the data user identity to set different access levels permissions, and strictly control access.

Thirdly, data real-time analysis engine. Through the real-time engine data, you can quickly analyze the data of illegal operations, hacker attacks and other security events with the potential threat.

Data means greater opportunities, have enormous value, but also encountered challenges in the field of data security. Only these two issues while dealing with big data security can receive unexpected results.

5.6 Increase the Safety Awareness of Employees. The user's quality decides the efficiency of safety protection. Enterprises should strengthen the employee safety training, make its knowledge of the data is using value, fully understand their important role in enterprise data, improving the capacity of employees the recognition of large data security threat and completes the data security sense of responsibility, is that every employee can consciously install antivirus software. In time for the system patch, set a strong password, continuously reduce the safety risk.

References

[1] Xiong Haiqing. Large computer networks in the context of data security precautions [J]. Information security, 2015 (2): 160-162.

[2] Fang Shimin. Information security issues facing large data analysis [J]. Computer software and applications, 2013 (19): 20-22.

[3] Chen Peng. Study on the problem of information security in the age of big data [J]. Electronic production, 2015 (18) 22-32.

[4] Lai Jishun. Large computer networks in the context of data security and protection [J]. Information technology and information technology, 2014 (06): 40-42.

[5] Liu Zheng, Li Chunliang. Data information security and countermeasure study on [J]. Science and technology innovation and application, 2015 (11): 30-33.

[6] Zhang Bingjian. Huge amounts of data is upending traditional business thinking [n].Wen Wei Po, 2013 (3).

[7] Discussion of information security in the age of big data [J]. Online world, 2014 (03): 14-16.

[8] Chen Huoquan. Data governance in the context of big data network security policy [J]. Macroeconomic research, 2015 (08): 25-27.