Study of Big Data Security and Privacy Protection

Guiling Fan

Xijing University, Xi'an, Shaanxi, 710123, China

Keywords: Big Data, Security and Privacy Problem, Technology and Skills

Abstract: With the rapid development of the global economy, data information has shown a surge, and the explosive information has increased the emergence of big data. Although it has brought great convenience to people's production and life to a certain extent, it also has corresponding zones. There are huge data security and privacy threats, information leakage has occurred, and it has brought great economic damage to people. In addition, the lack of big data security and privacy protection has caused many data leakage problems to be timely and powerful. The control has brought us great troubles and interventions. This paper intends to analyze the necessity of big data security and privacy protection from the concept of big data, and then analyze the challenges of big data security and privacy protection, propose targeted countermeasures, and improve the protection and level of big data security and privacy protection.

1. Introduction

Big data, in short, is a data set that is difficult to process and classify in the conventional sense of data storage and management tools. Big data, large scale, fast data transmission, and diversified features. Specifically, when China conducts activities through the Internet platform, it will generate some data information, such as pictures, texts, videos, etc., which are common data types. Most computer systems generate some data, such as files, databases, multimedia, etc., which are common data representations. The use of modern high technology to generate collected and acquired information, such as camera digital signals, is also a type of performance of data information. In short, big data is a data set generated by relying on the Internet platform. It is large-scale, high-speed, and diverse, and needs to be stored and processed by computer network technology.

2. Big data security issues

Due to the vulnerability of computers and external devices, they cannot withstand the damage of the natural environment and affect the security of big data. At present, the openness and service characteristics of computer networks are becoming more and more obvious, and the imperfections of their systems are becoming more and more obvious. The low security of the IP protocol adopted by the Internet is one of the security issues of big data. Cloud computing storage is a storage platform that is different from traditional data storage platforms. In the cloud storage platform, the data manager and the owner are separated from each other, and the cloud storage cannot guarantee the credibility of the cloud platform. Therefore, cloud storage platform data faces the risk of being sneaked or tampered with by third parties [5-7]. Traditional encryption methods are the primary method for solving such problems.

Because the operating system system is too large, there are inevitably security vulnerabilities, and its own security is difficult to guarantee. Most file protection mechanisms have a degree of security issues if they rely solely on the functionality of the operating system to implement an integrity verification mechanism. For example, the host-based file integrity protection method exposes itself to the guest operating system, and the isolation capability is poor. The malicious code can easily detect the detection system and try to bypass the detection and attack the system.

Big data relies on cloud computing. Data storage is stored in a cloud data cluster in a distributed manner. Once a virus or hacker invades a big data platform, data leakage and tampering will result in incalculable losses for individuals and businesses. A large number of facts show that big data is

Published by CSP © 2019 the Authors

not properly handled, which will greatly damage the privacy of users. According to different content to be protected, privacy protection can be further subdivided into location privacy protection, anonymous protection of identifiers, and anonymous protection of connection relationships. The threats people face are not limited to personal privacy leaks, but also based on big data predictions of people's status and behavior. A typical example is a retailer who knows the fact that his daughter is pregnant earlier than the parent through historical analysis and mails relevant advertising information to him. And social network analysis research also shows that the user's attributes can be discovered through the group characteristics. For example, by analyzing the user's Twitter information, you can discover the user's political inclinations, spending habits, and favorite teams. Current companies often believe that after anonymity, the information does not contain the user's identifier and can be publicly released. But in fact, privacy protection alone does not achieve privacy goals. For example, AOL has published a partial search history within 3 months after anonymous processing for people to analyze and use. Although personally relevant identification information has been carefully processed, some of the items can be accurately located to specific individuals. The New York Times immediately announced the 1 user it identified. The number 4417749 is a 62-year-old widowed woman who has 3 dogs in her family, has a certain disease, and so on. Another similar example is that the famous DVD renter Netflix has announced rental information for about 500,000 users and a reward of $1 million for the collection algorithm to improve the accuracy of the movie recommendation system. But when the above information is combined with other data sources, some users are still identified. Researchers have found that users in Netflix have a high probability of scoring non-top100, top500, top1000 movies, and de-anonymizing attacks based on the results of non-top videos.

One of the threats to big data credibility is to fake or deliberately created data, and erroneous data often leads to erroneous conclusions. If the data application scenario is clear, some people may deliberately create data, create some kind of "illusion", and induce analysts to draw conclusions that are favorable to them. Because false information is often hidden in a large amount of information, people can not identify the authenticity, and thus make a wrong judgment. For example, some of the fake comments on the review site are mixed in the real comments so that the user can't tell, and may mislead the user to choose some inferior goods or services. As the generation and dissemination of false information in the current online community becomes easier, the impact can not be underestimated. It is impossible to identify the authenticity of all sources by means of information security technology. The second threat of big data credibility is the gradual distortion of data in communication. One of the reasons is that the data acquisition process of manual intervention may introduce errors, and the data distortion and deviation due to mistakes will ultimately affect the accuracy of the data analysis results. In addition, data distortion has a factor in the version change of the data. In the process of communication, the reality has changed, and the data collected in the early days can not reflect the real situation. For example, the restaurant phone number has changed, but the early information has been included in other search engines or applications, so users may see conflicting information and influence their judgment. Therefore, users of big data should be able to understand the credibility of each data based on the authenticity of the data source, the data transmission path, the data processing process, etc., and prevent the analysis from producing meaningless or erroneous results.

3. Methods and measures to improve big data security and privacy protection

Big data protection technology is a direct carrier of big data security and privacy protection, which can ensure that data information is effectively stored and processed in the database field. Under the conditions of rapid development of modern science and technology, it is necessary to strengthen the protection of big data security and privacy, and to improve the innovation of big data protection technology. It can not only refine the source and record of data information, but also conform to the conformity of data and mark the data information in time. And verify the test to highly restore the true data information. With the advent of the era of big data, in many cases, users need to repeatedly confirm identity information and strengthen data protection technology to achieve the goal of maximizing user privacy protection, reducing unnecessary troubles and negative impacts, and avoiding bringing users Large economic losses and hidden dangers of personal and property safety.

With the advent of the data information era, many companies and many individuals have realized the importance of data information. As the source of innovation and the basis of execution, it is extremely important to optimize big data security and privacy protection. Many data tycoons use their own data and information resources to conduct bad operations and vicious violations, control the transmission and storage of many information, affect people's normal life and effective use of information, and hinder the healthy development of the era of big data. It is imperative to improve the correct use of data information and resist data monopoly. Only in this way can data information be used in a fair and reasonable competitive platform, make full use of the advantages of big data, and benefit the public, avoiding the illegal use of data information. IneviTable malignant effects.

With the advent of the era of big data and the rapid development of Internet technology, social networks as an important social product have become the link between people's communication and communication. Many people are active on social media. As part of social interaction, they all involve personal information. Partial disclosure, strengthening data supervision is extremely necessary. For anonymous social media information, it is necessary to protect the social network anonymously, make full use of the technology of modern technology to protect the security of personal information, and avoid the possibility of huge information loss caused by information leakage. Social networks have the function of convergence and communication. Inevitably, information exchange and transmission will occur between them. It is necessary to strengthen the comprehensive supervision of social content to fully protect the security of information and avoid the malicious application of information by unscrupulous people. Great personal and property damage to social subjects. Strengthening the overall supervision of data information in social networks requires the development of social realities and the popularity of big data security protection technologies to improve the comprehensiveness and pertinence of supervision, and thus improve the supervision effect.

The advent of the era of big data has provided a broader space for development and convenience for people's production and life. At the same time, the importance of big data security and privacy protection has gradually emerged, becoming the focus of widespread concern in society. However, due to the challenges faced by the big data era, China's efforts on big data security and protection measures are insufficient, making big data security and privacy protection technologies to be improved. In addition, it provides a good development space and legal protection for big data security and privacy protection to improve the comprehensiveness and pertinence of big data security and privacy protection.

4. Conclusion

Big data brings new security issues, but it is also an important tool for solving problems. This paper sorts out the key technologies related to big data security and privacy protection from the perspectives of privacy protection, trust, and access control of big data. But overall, the current research on big data security and privacy protection at home and abroad is not sufficient. Only through the combination of technical means and relevant policies and regulations can we better solve the problem of big data security and privacy protection.

References

[1] Feng Dengguo, Zhang Min, Li Wei. Big Data Security and Privacy Protection [J]. Chinese Journal of Computers, 2014, 37(1): 246-258.

[2] Chen Lili. Big Data Security and Privacy Protection [J]. Modern Industrial Economy and Informatization, 2017(04):92-93.

[3] Wang Yaohui. Big Data Security and Privacy Protection [J]. Communications, 2015(16):

224-224.

[4] Lü Xin. Research on Technology Architecture of Big Data Security and Privacy Protection [J]. Information Security Research, 2016, 2(3): 244-250.

[5] Ying Qin. Research on Big Data Security and Privacy Protection Technology [J]. Silicon Valley, 2014(10): 72-72.