Advances in Computer Science Research, volume 86
International Symposium on Communication Engineering & Computer Science (CECS 2018)

Analysis and Research of Information Security Based on Big Data

Yan Hou
Qilu Normal University, Jinan, Shandong, 250014, China
122365871@qq.com

Abstract. big data is an emerging product in today's era. Under the premise of the increasing development of this emerging product, information security issues come along with it. It believes that the security of personal information in the context of big data is not just a private issue, it is also a public management issue that is highly valued by the society. Based on this understanding, from the perspective of individuals, enterprises and governments, the issue of personal information security under the conditions of big data is examined. By reading the literature and questionnaires, statistical software is used to statistically analyze the data. From the perspective of individuals, enterprises and governments, the problems of personal information security are analyzed, such as serious disclosure of personal information, weak technical protection ability of enterprises and insufficient government crackdown. Therefore, it is necessary to strengthen information security protection.

Keywords: big data; information; information security.

1. Introduction

Big data primarily refers to a collection of data that can't be captured, managed, and processed with conventional software tools over an affordable time frame. After 2012, big data is being recognized by more and more people, and people use it to describe and define the massive data generated in the information age. The arrival of the era of big data has made people's communication more close and life more convenient. According to a column in the February 2012 "New York Times", the era of big data has arrived, and in business, economics, and other areas, decision-making will increasingly be based on data and analysis. In general, big data analytics is tied to cloud computing because real-time large dataset analysis requires a framework like Map Reduce to distribute work to hundreds or thousands of computers. The arrival of the era of big data has made people's communication more close and life more convenient. However, in this process, the era of big data faces serious information security challenges. The information of a large number of users has been stolen and leaked, which has a very big impact on people's lives. Therefore, people need to accurately grasp the opportunities of information security development in the era of big data, actively respond to the challenges in the development of information security and take a dialectical view of the information security opportunities and challenges faced by the big data era, so that big data technology can be better. Serve people's production and life. people need to use a dialectical perspective to view the information security opportunities and challenges faced by the big data era, so that big data technology can better serve people's production and life.

2. Literature Review

In 1890, Americans Brandys and Warren first mentioned the concept of "personal information protection" in the article "On privacy". The statement at the time was "privacy", which was the earliest time when the concept of personal information protection could be queried [1].

On the issue of changing the traditional model, David believes that different groups are superficially related, but the essence is really different [2]. Each country has different processing power for cyber security, which determines that each country's level of information security and social stability is different [3]. Meyer and Schonberg believe that our current priority is the degree of responsibility of users who need to solve information about their personal behavior, with a focus on the right to know [4].

In the risk brought by the era of big data, Colin believes that merchants have a large amount of personal information as a collection of personal information collected by individual consumers [5]. The hacker takes this point, the merchant is attacked by the hacker, and a large number of spam messages and harassing calls are aggregated on the consumer's mobile phone [6].

Domestic scholars' research on personal information protection is still in its infancy, and many questions have not yet been answered in common [7]. In terms of personal information protection: Li proposed to prevent personal information from being leaked in terms of objects of protection and methods of protection [8]. Firstly, ensure data security, and strengthen the protection of computer hard drives. The second is to take measures to combine the use of data and control, the third is to encrypt the information to avoid theft of information, and the fourth is to regulate and protect personal information through legislation [9]. Feng believes that the disclosure of personal privacy occurred after the collection of big data, so he proposed encryption technology such as anonymous information collection, anonymous technology protection, watermark theft, data traceability [10].

From a legal perspective: Liu and Hu analyzed the network privacy information in the context of big data and proposed that efforts should be made to strengthen network supervision and improve legal protection to create a pure network atmosphere [11]. Li believes that in the current era, strengthening personal privacy is a top priority [12]. To promote the protection of personal information to the height of national strategic resources, it is necessary to speed up legislation, protect individuals from the legal level, and increase the supervision of government departments on the basis of implementing laws [13]. After studying and researching some basic models, such as industry self-discipline mode and personal privacy processing mode, Wang and Tian proposed that strengthening industry self-discipline is an effective way to protect personal privacy and pointed out that government supervision is a macroscopic guiding means. Enterprises have supervised China's information security industry through industry self-discipline and strengthening their moral cultivation [14].

From the perspective of legislation: many scholars have discussed and analyzed personal information protection from the perspective of legislation, including criminal law, civil law, administrative law, etc. [15]. In the "Criminal Law Amendment (A)" adopted by the National People's Congress in 2009, it added "crime for illegally providing and illegally selling citizens' personal information" [16].

3. Methodology

3.1 Research Method

It is proposed to analyze the current status of personal information security, and the problems that may be encountered, and propose what difficulties we face in such an environment and intend to propose solutions to these problems through questionnaires and other means.

The main purpose of this paper is to analyze and study the personal information security problem in the era of big data. The research methods mainly include literature analysis method and questionnaire survey method.

Literature analysis

By consulting the relevant literature of China National Knowledge Infrastructure, input keywords such as “Big Data” and “Personal Information Security” and searched for tens of thousands of related articles. In terms of quantity, it can be concluded that this topic is very popular in the current era, and many scholars are conducting various researches. Therefore, before reading this article, it is necessary to read and understand the relevant literature. Only by reading through the relevant literature can we have a unique understanding of the topic of this paper.

Questionnaire method

On the basis of comprehensive reading of related literatures, a questionnaire is designed. The contents of the questionnaire cover everyone's understanding of big data, their views on personal privacy, and suggestions for personal information security. The recovered questionnaires will be classified according to relevant aspects, and analyzed and summarized by EXCEL.

3.2 Questionnaire Analysis.

In this survey, 300 questionnaires were distributed through various channels, and 278 questionnaires were returned, of which 265 were valid questionnaires. The analysis of the questionnaire was mainly based on the channels of personal information, the recognition of information security and the use of security measures, and the perception and practice of information disclosure.

Channels for personal information provided by respondents

By providing a survey of respondents with personal real information, the results are shown in Table 1:

Table 1. Channel analysis table for respondents to provide personal real information

Channels/number of people	Sign up for email, social events, etc	To apply for a job	Member or questionnaire survey
number of people	218	150	178
The percentage	83%	57%	68%

From the above table, it can be concluded that the real information about individuals involved in three aspects of the questionnaire, and it exists in many people. 218 people filled in their real information when they registered their emails and conducted social activities. These people accounted for 83% of the valid questionnaires. There were 150 people who filled in the personal information in the job search. These people were mainly graduates of universities, secondary schools, and job-hopping families, accounting for 57% of the valid questionnaires. There were 178 people who filled in their personal information when they applied for membership cards of major stores and websites, accounting for 68% of the effective questionnaires. These people were mainly housewives and the elderly.

Whether the respondent has lost personal items

Statistics were made on whether the respondents had lost personal belongings. The statistical results are shown in figure 1:

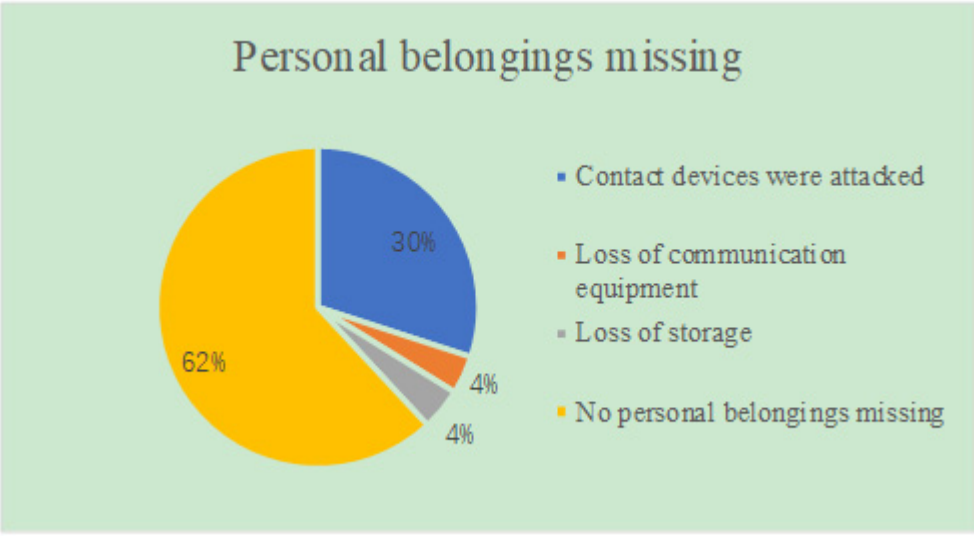

Figure 1. Analysis of the loss of personal belongings.

Through the statistics in the above table, it can be concluded that the contact device has been attacked, including mailbox theft, WeChat remote login, QQ stolen, etc. In the valid questionnaire, 78 people have lost their mailboxes, and they have encountered QQ or WeChat remote logins. Their personal information was lost in these situations, and the same threats were also the personal

information of friends in their address book. A small number of people had lost their information and friends' information after their communication equipment was lost.

Respondents' personal information technology protection

The respondent's personal information technology protection situation was understood, and the results are shown in Figure 2:

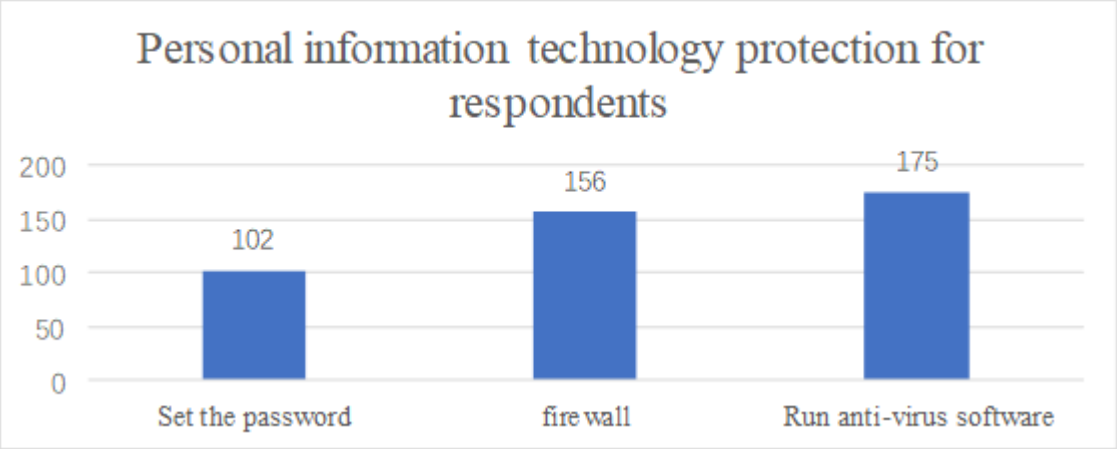

Figure 2. Analysis of the personal information technology protection of respondents.

Personal communication equipment and storage equipment are often used by us. The memory inside is occupied by many things in our lives. The information inside is the epitome of each of us. If the information is not encrypted, they will become transparent and appear in front of each other without any omissions. However, through the questionnaire, the results obtained are not optimistic. On the basic defense issues of setting passwords, installing firewalls, running anti-virus software, etc., only half of the people use it, while others don't.

4. Results and Discussion

This questionnaire starts with the basic situation of the respondents and provides an in-depth understanding of the respondents' usual online habits and awareness of personal information protection. From the technical protection degree of electronic equipment, the attitude towards information leakage, the methods adopted, and the countermeasures when the information leakage caused losses, the problem design and statistics are presented, and the following problems are presented:

The disclosure and abuse of personal information is serious. In terms of e-mail, record archiving, social networking, etc. on the web, there is a large amount of data generated every day, which contains a large amount of user privacy. Therefore, with the development of the era of big data, a large amount of data is generated, which increases the risk of leakage of users' personal privacy. The data of these users includes not only the data operated by the enterprise, but also the privacy of the user and the behavior of the user's web browsing. Once the user's personal information is leaked and used by criminals, it may cause damage to the user's property and even threaten the user's personal safety.

Storage devices used by enterprises are often compromised. The leakage of personal information by enterprises is manifested in the technical settlement of firewalls and servers, and the related technical personnel are driven by interests and lost their way. The main reason for corporate information leakage is that software and hardware equipment are backward, technology can’t keep up, and investment in security is insufficient.

The illegal use of information causes property damage and difficulty in rights protection. The problem of the difficulty of rights protection reflected in the questionnaire is that China’s crackdown is insufficient. When citizens receive harassing calls and spam messages, many people choose to be silent. Only a small number of people choose to protect their rights. It is precisely because China's

efforts to deal with illegal acts are insufficient, which makes the road to safeguarding rights difficult and long.

Individuals are not sufficiently aware of their information protection. Based on the serious situation of personal information leakage, the main reason for the information leakage from the personal perspective is that individuals have insufficient awareness of their own information protection.

5. Conclusion

Big data is an emerging product of the modern era. Under the premise of the development of this emerging product, personal information security problems have emerged. It is believed that the security of personal information in the context of big data is not only a private issue, it is also a public management issue that is highly valued by the society. Based on this understanding, the issue of personal information security under the conditions of big data is examined from the perspective of individuals, enterprises and government. The main manifestations and causes of this problem are analyzed, and targeted countermeasures are proposed.

Individuals are not sufficiently aware of their own information protection, resulting in serious disclosure and abuse of personal information. In response to this problem, individuals are required to take the initiative to receive safety education, develop good online habits and master basic information protection skills.

In view of the problem that the system is attacked due to the relatively backward software and hardware technologies for personal information protection, it is proposed that enterprises should strengthen their sense of responsibility for information security, strengthen the ethical quality of the merchant, strengthen the self-discipline of the industry, and strengthen the feasibility suggestions for pre-existing prevention, in-process control, and active handling afterwards.

In response to the problem of personal property loss and difficulty in protecting rights caused by information leakage, it is proposed to speed up the promotion and popularization of VIEID, strengthen the security construction of information systems, increase the training of talents related to big data, strengthen the supervision of relevant departments, and improve relevant legal.

References

[1]. David Laser. "Computational Social Science". Science. 2009, 323(6), pp.721-723.

[2]. Colin Tankard. Big data security. Feature, 2012(7):5-8., pp.101-103.

[3]. Xu L, Jiang C, Wang J, et al. Information Security in Big Data: Privacy and Data Mining. IEEE Access, 2017, 2(2), pp.1149-1176.

[4]. Hu J, Vasilakos A V. Energy Big Data Analytics and Security: Challenges and Opportunities. IEEE Transactions on Smart Grid, 2016, 7(5) pp.2423-2436.

[5]. Kaur P D, Monga A A. Managing Big Data: A Step towards Huge Data Security. International Journal of Microwave & Wireless Technologies, 2016, 6(2): pp.10-20.

[6]. Moura J, Serrao C. Security and Privacy Issues of Big Data. Handbook of Research on Trends and Future Directions in Big Data and Web Intelligence. 2016: pp.182-187.

[7]. Strang K D, Sun Z. Big Data Paradigm: What is the Status of Privacy and Security. Annals of Data Science, 2017, 4(1): pp.1-17.

[8]. Gholami A, Laure E. Big Data Security and Privacy Issues in the CLOUD. International Journal of Network Security & Its Applications, 2016, 8(1): pp.59-79.

[9]. Ye H, Cheng X, Yuan M, et al. A survey of security and privacy in big data. International Symposium on Communications and Information Technologies. IEEE, 2016: pp.268-272.

[10]. Klein J, Buglak R, Blockow D, et al. A reference architecture for big data systems in the national security domain. 2016, 2: pp.51-57.

[11]. Gahi Y, Guennoun M, Mouftah H T. Big Data Analytics: Security and privacy challenges. Computers and Communication. IEEE, 2016: pp.952-957.

[12]. Chin W L, Li W, Chen H H. Energy Big Data Security Threats in IoT-Based Smart Grid Communications. IEEE Communications Magazine, 2017, 55(10): pp.70-75.

[13]. Dan W, Zhao W, Ding Z. Review of Big Data Security Critical Technologies. Journal of Beijing University of Technology, 2017, 43(3): pp.335-349.

[14]. Yang M, Zhou X, Zeng J, et al. Challenges and Solutions of Information Security Issues in the Age of Big Data. China Communications, 2016, 13(3): pp.193-202.

[15]. Blobel B, Lopez D M, Gonzalez C. Patient privacy and security concerns on big data for personalized medicine. Health & Technology, 2016, 6(1): pp.75-81.

[16]. Hong X, Yu Z, Haibo H U. Big data and security visualization. Journal of Chongqing University, 2016, 39(2): pp.71-81.