Research on the Security of Personal Information in the Era of Big Data

Cheng Chi[1(✉)], Tengyu Liu[2], Xiaochen Yu[1], Shuo Zhang[3], and Shuo Shi[3]

[1] People's Procuratorate of Heilongjiang Province, Harbin 150001, Heilongjiang, China
13904614111@139.com
[2] Higher Court of Heilongjiang Province, Harbin 150001, Heilongjiang, China
[3] Harbin Institute of Technology, Harbin 150001, Heilongjiang, China

Abstract. The advent of the era of big data has made people's lives more intelligent and convenient, and has brought enormous value to human beings. At the same time, it has brought about a lot of challenges. Personal information security is one of them. In the big data era, massive amounts of personal information are continuously input, simultaneously, the means of illegally acquiring, disseminating, and applying personal information are emerging in endlessly, Raising human thinking about information security. Therefore, the personal information security problem caused by big data should not be underestimated. Firstly, it introduces the related concepts of personal information, discusses the sources of risks faced by personal information in the process of implantation, dissemination and application. Secondly, it discusses the existing problems and solutions in China. Finally, the application difficulties and coping strategies of information security technology are analyzed, which provides theoretical support for personal information security in the era of big data.

Keywords: Personal information security · Source of risk · Legal protection · Information security technology

1 Introduction

With the vigorous development of network and information technology, the amount of data generated by human beings using the Internet is growing exponentially. In the era of information explosion, the concept of "big data" came into being. Nowadays, this concept is applied to almost all areas of human intelligence and development, such as advertising, finance, medical care, travel, artificial intelligence, etc., which not only promotes the digital transformation of government, enterprises, and social organizations, but also makes people's lives become more intelligent and convenient.

According to statistics, there are 2.9 million emails being sent every second in the world. If you read one article in a minute, it is enough for one person to read 5.5 years day and night. 28,800 h of video will be uploaded to YouTube every day, enough for one person to watch it day and night for 3.3 years. Amazon generates 6.3 million orders a day. Google processes 24 petabytes of data per day.

S. Han et al. (Eds.): AICON 2019, LNICST 287, pp. 107–114, 2019.
https://doi.org/10.1007/978-3-030-22971-9_9

The era of big data has come quietly. Whether people use “Internet +”, “Industry 4.0” or “Cloud Computing” to describe the current world, they are inseparable from the extensive use of “big data”. While the era of big data has brought about tremendous changes in human production and lifestyle, personal information security is facing serious threats and challenges, such as various fraudulent calls and harassing text messages. Our personal information is inadvertently being used illegally. The issue of personal information security brought about by the era of big data cannot be ignored.

As a product of the information age, big data aims to explore the potential value between data and serve the economic and social development through the storage and analysis of massive data. However, with the further development of the industry, the resulting problems have attracted more and more public attention. The world’s leading companies are caught up in the scandal of user information leakage, which poses a serious threat to the information security of citizens. In the era of big data, the market initiative lies in the hands of those who have the most data. Therefore, the demand for data including personal information is expanding, which further increases the importance of protecting personal information in the era of big data. In this context, this paper introduces the definition of personal information in the era of big data, the legal protection of personal information and Internet technology, and finally puts forward the protection countermeasures. It is hoped that the research in this paper can enlighten this subject.

2 Personal Information Security in the Era of Big Data

In a broad sense, personal information refers to the sum of all the content that can be transmitted related to a natural person. In the context of big data, personal information is often associated with personal information on the Internet. Therefore, personal information should be divided into two categories: the first category is the personal information that appears on the Internet with the birth of the Internet, such as e-mail addresses, online records, etc.; the second category is the personal information that existed before the birth of the Internet. Such as name, gender, age, ethnicity, occupational status, etc., after the birth of the Internet, this personal information becomes personal information on the Internet according to the needs of life or obeying social management.

Gemalto, a research company of digital security, released a report of Breach Level Index for the year of 2015. It pointed out that 2015 witnessed serious incidents of data breach. During the 12 months, security staff of the company collected and categorized 1673 cases of data breach, resulting in the breach of 707 million data records. The major forms include the attack of personal information and identity theft (Fig. 1). According to the report about the Protection of the Rights of Chinese Netizens 2015 released by Internet Society of China, the personal identity information of nearly 80% netizens had been breached. The information about personal online activities of over 60% netizens had been breached. In this case, the analysis of the means of obtaining information and the sources of risk is the first step to protect personal information.

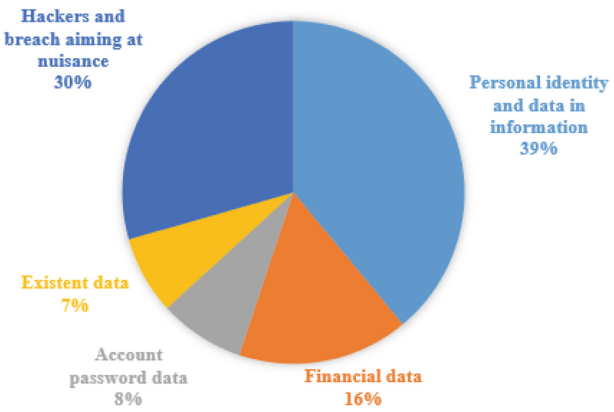

Fig. 1. The forms of data breach in 2015 (released by Gemalto)

With the implantation, dissemination and application of personal information, personal information faces three risks: access to information by network service providers, hackers and Trojans, and illegal circulation of information.

The first source of risk is the access of personal information to network service providers. Nowadays, the premise of people enjoying the services of various computers and mobile phone software and service platforms is registration, which inevitably involves personal information such as name, mobile phone number and e-mail address. In September 2017, Beijing police arrested illegal criminal suspects on 26 illegal websites and seized more than 1 million citizen information, involving information security of hundreds of thousands of people.

The second source of risk is that the means of obtaining personal information is itself an offence. Hackers and Trojans are products of the Internet era. In the era of big data, hackers and Trojans have become more rampant and embarrassing on the Internet. In September 2018, the information of 500 million users of China Lodging Group was leaked due to a hacker attack and turned to overseas websites for selling. In October 2018, online fashion retailer SHEIN was hit by a massive hack that resulted in a data breach for 6.5 million users.

The third source of risk is the subsequent disposal of the data after obtaining personal information. In the first case, the personal information is illegally transferred to the other people, and thus the personal data is "shared" between different subjects. The second case is that the Internet platform analyzes and infers the legally acquired personal information using big data technology to obtain the user's personal preferences. Such information mining may violate the conventions originally used for a specific purpose or within a specific scope. On December 28, 2017, the People's Court of Chang jiang Li Autonomous County of Hainan Province publicly pronounced a large- scale crime of infringing citizens' personal information. The defendant obtained a large amount of personal information from the hands of two people in the same village, deposited in Baidu's network disk, and send QQ information on the Internet sells more than 100 million citizens' personal information online and draws huge profits from it.

In the face of increasingly serious personal information security issues, although China has promulgated a number of related laws, there are still some loopholes in legislation, supervision, and legal remedies. It is necessary to protect personal information from multiple dimensions such as law and technology.

3 Research on Legal Protection of Personal Information

At present, the key to the protection of personal information in China can be expressed in two aspects, namely the legal protection and self-discipline protection of personal information. Among them, the legal protection of personal information is in the relevant laws and regulations, such as the "Personal Information Protection Law" and the "Internet Security Law of the People's Republic of China", through the formulation of personal information protection provisions to protect personal information, which can be mainly divided into legal Direct protection and indirect protection. The former is a law and regulation that explicitly proposes the protection of personal information, while the latter is a category that laws and regulations have proposed to protect personal dignity and personal privacy and personal information, thereby extending the protection of personal information. Compared with other countries, the protection of personal information in China is in a relatively backward position, and there are mainly three problems.

3.1 Lack of Clear Personal Information Protection Objects

In China's relevant laws and regulations, a big problem is the object of personal information protection, although the concept of "personal information" has been formed in relevant laws, its specific scope has not been clearly defined, such as personal information terms, personal data and personal information rights. If only according to the definition of personal information in law, the object of personal information protection is identified as all identifiable information related to individuals, then its extension will be very broad.

3.2 The Personal Information Protection Law Enforcement System Is Relatively Lagging Behind

There is a certain lag in the personal information protection law enforcement system. Although many departments in China currently supervise the network information security, there is no supervision responsibility for the actual processing of enterprise data. The current relevant laws and regulations have not yet reached a relatively perfect state, an effective regulatory system has not yet been formed. And law enforcement agencies lack a strong regulatory system. The law enforcement agencies did not construct the cyber security review mechanism and the threshold access system. Therefore, the information security threshold of the industry is very low, and all enterprises are mixed together, which creates great difficulties in management.

3.3 Improve the Existing Model Litigation System

In the era of big data, the infringement of personal information usually involves many groups, and the specific scope is widely dispersed. Therefore, compared with the past, the benefits of participating in litigation and efficiency cost considerations during this period are even more critical. We can study from foreign demonstration cases and parallel actions of group litigation to fully liberate the infringed information subject, and at the same time, the litigation is replaced by professional institutions and related personnel to reduce the individual's rights protection costs to a certain extent. In this way, the actual capacity of rights relief is effectively improved.

The protection of personal information is not only a simple legal issue, but also involves many factors such as the public's right to know and public opinion supervision. Therefore, both the criminal law and the judicial system need to be further improved. Of course, the protection of personal information is closely related to information technology, and technically improving the level of information security of big data helps to strengthen the level of information protection.

4 Research on Personal Information Security Technology

In order to implement the relevant requirements of the "Network Security Law" for the protection of personal information, the Central Network Information Office, the Ministry of Industry and Information Technology, the Ministry of Public Security, the National Standards Committee and other departments form an expert working group, and privacy provisions for 10 network products and services such as WeChat and Taobao. A review was conducted to regulate the collection, storage, use, and transfer of user personal information. In this special review of privacy protection, the transparency of privacy provisions and the increase of user choices have become highlights. However, efficient and streamlined data management is a higher requirement of enterprise practice. This includes establishing privacy management, planning data protection strategies, developing privacy policy procedures and guidelines. From the perspective of information technology, considering privacy protection in system and program design, and conducting privacy impact assessment, privacy protection can be placed at the "front end. "In the early stage of the product, the concept of privacy protection was added, and the whole life cycle management strategy was consolidated through continuous supervision and evaluation. But to achieve this goal, there are still some difficulties in the technical level.

4.1 Technical Level Application Difficulties

Big data is the product of the full development of information technology. The information form of big data is essentially different from the original information form. Therefore, it is difficult to analyze the existing information security technology from the information.

Mass and Diversification of Information. Smart devices record our personal information all the time, such as browsing records, online platform purchase records,

geographical location records, etc. Once we connect to the Internet, the service provider will collect and record the generated data, resulting in a huge amount of data. And the structure of these data is very confusing. When extracting data, it often adopts a non-differentiated way, which may expose some personal privacy information, or out of the actual information and incomplete information, thus affecting the confidentiality, authenticity and integrity of information security.

Information Can Be Digitized. Now, the information can be digitized, can be expressed by different combinations of binary numbers "0" and "1", which will make the data more easily obtained and transmitted, and data can make the stealing method more concealed, giving some information theft means provides convenience and makes information security more difficult.

Information Can Be Transmitted. After the information is digitized, it can be transmitted arbitrarily. There is no geographical or time limit on the transmission of information on the Internet, such features undoubtedly provide great convenience to those who anonymously collect information about netizens, distorted facts, and make a living. Due to the limitations of existing technologies, we cannot determine the information of collectors at present, so our information security protection has fallen into a very passive situation.

4.2 Technical Level Response Strategy

The threats to personal information in the era of big data come from the abuse of technology. Therefore, using advanced technology to prevent and control and counter the use of networks to invade personal information security is the most direct and efficient way.

Do a Good Job of Data Encryption Protection. The development of computer networks has greatly improved the efficiency of data transmission and circulation. In the era of big data, the transmission speed of data information is accelerated, and the effectiveness of data transmission is closely related to the security of computer networks. We must pay attention to the encryption protection of data, and use existing file encryption technology to protect the security of personal information. For enterprises, when conducting relevant enterprise information exchange, it is necessary to set relevant authority settings, and standardize the use of the internal network module information by the personnel of each work module to improve the effectiveness of personal information protection.

Data Technology for Real-Time Monitoring. In the protection of computer network security, it is of great significance to do relevant real-time monitoring. Doing relevant monitoring and intrusion detection is also an effective supplement to the existing personal information security protection measures. With the help of big data technology, the task of snooping information security threats can be completed, that is, a flexible predictive of the threat of attacks is completed. And a comprehensive test of the unknown danger, timely identify the phenomenon of violation of information security use, and do the relevant protection work. Through a series of data screening and other classification of abnormal phenomena, it is possible to perform alarm processing,

which can reduce the probability of information leakage in the shortest time and ensure the security of personal information.

Data Anonymization. The so-called anonymization is to erase the personal identity information from the database. These personally identifiable information includes name, address, credit card number, date of birth, and social security number. The remaining data can be used and shared. Although this measure does not completely eliminate the risk of personal information disclosure, it helps to maintain the security of personal information data, which is a feasible protective measure at this stage.

Strengthen Data Security Protection. At present, the more popular anti-virus software, security guards, computer butlers and other software are typical examples. Software engineers have fully developed Internet software to prevent and control viruses, Trojans, and malicious programs, and protect personal information from illegal attacks. Intelligent firewalls, access control technologies, data encryption technologies, and vulnerability scanning technologies are all common techniques used in big data Internet to prevent personal information from being attacked. In addition, using big data technology to address the information security challenges of the big data era is a useful attempt. For example, in response to harassing calls and text messages, some Internet technology companies have launched information interception service software based on large databases. This type of software effectively selects and identifies malicious calls and spam and intercepts them by creating a "call blacklist" database. Similar information technology tools can not only effectively respond to new data security threats, but more importantly, show us new ideas for dealing with information security threats.

5 Future Work

The development of computer network technology has brought about big data changes, but also brought the risk of information leakage. Computer network plays an important role in the development of modern society. It is related to people's daily life, business development and social harmony and stability. It is necessary to correctly recognize the development of computer network technology, attach importance to relevant information protection, and improve its own. Concept awareness, strengthen investment in relevant information protection, improve relevant legal and legal systems, do a good job in computer information security, and improve the use of computer networks. Only by continuously improving technology and improving the legal system, personal information will be protected.

6 Summary

In the context of the development of big data, data information has become an important asset of the state and enterprises, and information security has become a major strategic issue for the development of the country and society. However, due to improper management, hacker attacks and network system security vulnerabilities,

personal information always faces the possibility of being leaked. Strengthening the research on information technology has important practical value in the case that information leakage cannot be completely eliminated.

Big data has brought us great convenience, promoted the rapid development of the economy, the continuous improvement of the society, and it also brought us some unexpected troubles. We enjoyed the big data for us. Convenient and fast personalized life, while being troubled by these problems. There are many reasons for the security of personal information. To solve this problem, we need the joint efforts of the whole society. We also need to take corresponding measures at the legal level, technical level and personal level to coordinate the harmonious development of big data and information security based on the actual situation of today's society.

References

1. Huang, C.: Personal privacy in the era of big data. Huazhong Normal University, Wuhan (2015)
2. Zhang, M.: Risks and countermeasures of citizens' personal information data in the age of big Data. Inf. Stud. Theory Appl. **6**, 57–70 (2015)
3. Wang, Z., Yu, Q.: Privacy trust crisis of personal data in china in the era of big data: the survey and countermeasures. Comput. Law Secur. Rev. **31**(6), 782–792 (2015)
4. Chen, K.: The research of Open-sensitive data security. Zhejiang University, Hangzhou (2007)
5. Zou, H.: Protection of personal information security in the age of big data. In: International Conference on Computational Intelligence and Security 2016, CIS. IEEE, Wuxi (2016)
6. Ruixiang, W.: Research on the Legal dilemma and countermeasures of personal information protection in big data age. J. Inf. Secur. Res. **12**, 1097–1101 (2017)
7. Hao, X.: A preliminary study on the legal issues of citizen personal information security in the era of big data. Knowl. Base **12**, 9–10 (2017)
8. Li, F.: Analysis of personal information legal protection in the age of big data. Glob. Market Inf. Guide **25**, 30 (2017)
9. Dekker, M., Karsberg, C.: Technical guidance on the security measures in Paper 13a, Version 2.0. European Union Agency for Network and Information Security (ENISA) (2014)
10. Kshetri, N.: Big data's impact on privacy. Secur. Consum. Welf. **38**(11), 1134–1145 (2014)
11. Wang, L.: Privacy system in the United States and its impact on the development of China's Legislation (2007)