A Preliminary Study on Network Privacy Protection in the Context of Big Data

Chunying Wu

1Department of Military Management, China People's Police University, Lang fang City, Hebei Province, 065099, China

*Corresponding author’s e-mail: 775251252@qq.com

Abstract. Privacy protection on the Internet is a key concern of people in the context of current big data, directly affecting people's information security and personal safety. Based on this, this paper first analyzes the current status of network privacy in the era of big data, and summarizes the common theories and data analysis methods and models of existing network privacy protection. Further, the paper makes an analysis of the current status of relevant network privacy laws in China, and finally expects the outlook of future development.

1. Introduction
The global data explosively increases, with the rapid development of high technology such as the Internet, the Internet of Things and cloud computing, which has marked the entry of human society into the era of big data. However, it also brings serious negative effects to people, along with technological development. In March 2018, the British media revealed that the profiles of more than 50 million Facebook users were illegally used to send political advertisements. Google acknowledged an API vulnerability in Google+, bringing consequences that the name, email, occupation, and age of 52.5 million users were accessed in the six days of November 2018[1]. The collection of massive data increases the possibility of exposing sensitive data, and the abuse and misuse of big data also increase the risk of privacy leakage. Once exposed or leaked, it will bring serious losses and great troubles to network users.

2. Current status of network privacy in the context of big data
According to survey conducted by China Consumer News in 2015, 76% of respondents believe that personal information in the era of big data is more likely to be leaked, among which personal communication information is the most vulnerable [2]. Privacy is a basic right of every citizen. In June 2017, China officially implemented the "Network Security Law of the People's Republic of China", the provisions on data security and privacy protection have become the focus of attention. The Network Security and Security Bureau of the Ministry of Public Security issued "the Guidelines for the Protection of Internet Personal Information Security (Draft for Comment)", which provides comprehensive guidance on the protection of personal information of enterprises from the aspects of security management mechanism, security technology measures and business processes of personal information security protection[1].
Big data is characterized by massive data scale, rapid data flow, diverse data types and low value density [3]. In the context of big data, big data privacy protection faces many new challenges [4]. Big data is a data set based on the Internet platform. As an internet-based big data, it will face various security risks brought by the network [5].

The amount of data in big data is huge and concentrated. It is easier to be found, and becomes the primary target of hackers, because they can get more profits. With the development and application of data mining, machine learning and other technologies, big data analysis becomes more and more powerful. Massive data contains value, comprehensive analysis of which can obtain more personal information, increasing the risk of personal privacy disclosure. Data monitoring technology is widely used, such as the purchase of tickets, outpatient visits, network shopping, driving location and other information, so people's living habits are often snooped, and personal privacy will be violated.

In addition, it is difficult to determine the scope of violations of personal privacy and the protection of personal privacy [6]. General researcher is difficult to obtain massive original data related to the major required by big data research. What’s more, some unscrupulous people by policy loopholes carry out illegal activities and affect people's normal life. Therefore, a sound policy system is the premise basis for ensuring the optimization of the current network environment.

3. Common theory of existing network privacy protection

3.1. Evolutionary Game Theory.
In recent years, many scholars have begun to apply the evolutionary game approach to the privacy protection of users and establish game models. Evolutionary game theory is the combination of evolutionary theory and game theory in evolutionary economics. It is assumed that the game players are randomly selected from the largest population, and the players play the game repeatedly according to the social pattern [7]. In 2017, Feng Mining and other scholars found that user trust, social platform influence and reward, and third-party supervision are the key factors influencing the evolution of privacy policies of users and social platforms [8].

3.2. Protection Motivation Theory.
The theory suggests that whether users have protective behavior motivation depends on the results of threat assessment (perceived susceptibility, perceived severity) and response assessment (response efficiency, self-efficacy) in the cognitive process [9]. The study of Xie Weihong and Qu Jingjing in 2018 shows that threat assessment positively stimulates users' online privacy concerns and prompts them to take corresponding privacy protection actions [10].

3.3. Privacy Calculus Theory.
This theory suggests users decide whether to upload their personal information after calculating the revenue and risk, and privacy disclosure behavior is decided whether to implement after weighing perceived revenue and privacy concerns. When the concern cost caused by privacy disclosure is greater than the benefit gained from privacy disclosure, this behavior will not be adopted [11]. In 2016, Li Haidan and other learners concluded that perceived risk, perceived benefit and procedural justice are strong influencing factors of privacy disclosure intention, which further affects disclosure behavior [12]. In 2018, Zhao Dongxiang and other researchers showed that perceived privacy control, as an important psychological element, can positively affect perceived benefits and negatively affect privacy concerns [13].

3.4. Theory of Planned Behavior
All factors that may affect behavior are indirectly affecting behavior through behavioral intentions. Behavioral intentions are influenced by three factors: the attitude derived from the individual himself, the subjective norm derived from the external, and the perceptual behavior control [14-15]. Based on this theory, scholars Hu Xiaomei in 2016, the study found that privacy protection awareness and social
network experience are the factors that adolescents adopt privacy protection behavior in social networks. Parental education level and parental control of their children's online behavior are other influencing factors. This two jointly determine the privacy protection behavior of teenagers in social networks [16].

3.5. Uses and Gratifications Theory
From the standpoint of the audience, the theory examines the psychological and behavioral effects of mass communication on humans by analyzing the audience's motivation for using the media and obtaining satisfaction with the needs [17]. It emphasizes the audience's initiative, highlights the status of the audience, combined with the relevant knowledge of psychology and sociology, explains people's behavior of using the media to get satisfaction, and proposes the social reasons and psychological motives of the audience to accept the media. In 2018, scholar Vishwanath and others published a paper and concluded the influence of social media users' information needs, entertainment needs and social needs on their privacy protection behavior [18].

4. Common data analysis methods and models for current network privacy protection

4.1. Descriptive Statistical Analysis Method
Through frequency analysis, central tendency analysis, discrete degree analysis and other analysis methods, in the form of tabulation, statistical graph, etc., this method gives a summary and statistical description of the relevant data of all variables in the overall survey.

4.2. Structural Equation Modeling
Structural equation modeling is a method for establishing, estimating and testing causal models. The model contains both observable explicit variables and potential variables that cannot be directly observed. Structural equation modeling can replace multiple regression, path analysis, factor analysis, covariance analysis and other methods, and clearly analyze the effect of single index on the whole and the relationship between single index[19].

4.3. Linear Regression Model
Linear regression model refers to modeling by building the relationship between independent variables and dependent variables, judging the relationship between variables and making predictions.

5. Analysis of the Current Situation of Relevant Network Privacy Laws in China
China's provisions on privacy protection are scattered in the constitution, criminal law and other laws, but in terms of interpretation, there is no clear and complete definition of the boundary of privacy. With the development of big data and artificial intelligence, the value of privacy is increasing gradually. In the network media, especially in the social media environment, the boundary of privacy is more ambiguous.

In 2001, China issued "the interpretation on several issues concerning the determination of liability for damages for tort spirit", one of which stipulates: "Where a person violates the public interest, social morality, or infringes upon the privacy of others, and the victim brings a suit in a people's court for compensation for the mental damage on the grounds of infringement, the people's court shall accept the case according to law." This is the first time that China has positively protected privacy and privacy rights in the form of legal documents [20]. In order to comply with the rapid development of the Internet and the requirements of The Times, China issued and implemented the "network security law" in June 2017. It has better guaranteed and safeguarded China's cyberspace sovereignty, national security and social and public interests, protected the legitimate rights and interests of citizens, legal persons and other organizations, and promoted the healthy development of economic and social informatization. There is a principle in the network security law that personal information should be collected and used in a lawful, legitimate and necessary manner. In May 2018,
China implemented “the personal information security specifications for information security technology”. In the form of national standards, the specification clarifies the compliance requirements for the collection, storage, use and sharing of personal information, and provides guidance for network operators to formulate privacy policies and improve internal control [1]. The Network Security and Security Bureau of the Ministry of Public Security issued "the Guidelines for the Protection of Internet Personal Information Security (Draft for Comment)" In November 2018. In May 2019, the cyberspace administration of China released “data security management measures (draft for comments)”, which stipulated the collection methods of personal sensitive information and other contents. In recent years, China has issued and implemented a number of laws and regulations, and relevant personal information and privacy protection policies will be gradually improved in the future.

6. Future Prospects and Development Direction

Big data is closely connected with the Internet. While big data brings convenience, it also poses a threat to the security of big data and the privacy of users.

Firstly, effective assessment and control should be actively carried out for the existing security vulnerabilities and network risks in the current network. The new technology should be applied flexibly to enhance the ability to solve network security problems and build a perfect and harmonious network environment for users.

Secondly, in view of the security risks of big data, the application of big data security technology should be enhanced to ensure the security of big data information in all links (collection, storage, processing and transmission) relying on technology. Differential privacy and k-anonymity privacy protection technologies should be further studied and optimized, so as to further protect users' privacy, avoid the loss of users' interests, and make big data play a better role for the society.

In addition, common data analysis methods for online privacy protection should be further improved. Web crawlers should be legally used to collect relevant information of research targets, and multiple linear regression, clustering and other machine learning algorithms should be used for analysis, so as to further improve the intelligence and accuracy of data analysis.

Finally, the policies, regulations, laws and regulations of enterprises related to network privacy should be better improved to promote the popularization and scientificity of network security legislation. Reasonable third-party supervision strategies can be studied to supervise the use of big data by relevant enterprises and industries on the premise of consciously protecting citizens' privacy [6].

7. Conclusion

Personal privacy protection in the big data environment is a new hot issue. It needs comprehensive scientific research at all levels of the country and society to be completely solved and big data can better meet the development of the times.

Acknowledgements

Supported by Key R&D Projects in Hebei Province (Grant No. 18210706), National Key Projects of Statistical Science Research (Grant No. 2017LZ37), Research and Practice Project on Higher Education Reform in Hebei Province (Grant No.2018GJG450), the Humanities and Social Sciences Fund of the Education Ministry (Grant No.17YJC630214), China People's Police University Project (Grant No. ZQN2018056).

References

[1] Li Shaopeng. Cybersecurity Highlights for 2018, 2019-01-02, https://mp.weixin.qq.com/s/YvUX8Zip9gfAtj6YY27BA.

[2] Xie Weihong, Qu Jingjing. Impact of Threat Assessment on Internet Privacy Concerns in the Context of Big Data: Regulation of Organizational Privacy Policy[J], Science and Technology Management Research, 2018 No. 19, 188-194.
[3] Big Data: The Next Frontier for Innovation, Competition, and Productivity, Mc Kinsey Global Institute Report, (2011-05) [2015-10-11]. http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

[4] Xue Fu, Chen Hongbing. Research on Big Data Privacy Ethics [J]. Research on natural dialectics , 2015,31(2),44-48.

[5] Gao Xiang. Necessity and Measures of Big Data Security and Privacy Protection, Electronic Technology & Software Engineering[J],2016(10),208.

[6] Yang Bo. Personal Privacy Protection Mechanism in Big Data Environment, Software Guide [J], 2016(8), 168-170.

[7] Li Junwei. The Study of Social Networks Users Privacy Based on Evolutionary Game [D], 2013.12.

[8] Feng Mining, Zhu Guang, Yang Jiayun. Research on Social Network Privacy Protection Based on Evolutionary Game [J]. Journal of Intelligence, 2017(9): 127-132,85.

[9] Rogers R W, Cacioppo J T, Petty R. Cognitive and Physiological Processes in Fear Appeals and Attitude Change: A Revised Theory of Protection Motivation [M] . Social Psychophysiology: A Sourcebook, 1983.

[10] XieWeihong, Qu Jingjing. Impact of Threat Assessment on Internet Privacy Concerns in the Context of Big Data: Regulation of Organizational Privacy Policy, Science and Technology Management Research, 2018(19):188-194.

[11] Shi Shuo. Factors Affect Privacy Disclosure on Social Network Sites: Integrity of Privacy Calculate Theory and TPB Model [D], Nanjing University, 2011.

[12] Li Haidan, Hong Ziyi, Zhu Hou. Influencing Factors of User’s Privacy Disclosure from PrivacyCalculus and Justice Perspectives [J], Library Information Knowledge, 2016(6):114-124.

[13] Zhao Dongxiang, HuoZhaoguang, Fan hao. Examining Individuals’ Behavior Intention of Mobile Health Management Services: An Empirical Study from Information Privacy Perspective[J], Journal of Modern Information ,2018(5):74-81.

[14] Ajzen I. From Intentions to Actions: A Theory of Planned Behavior[M]// Action Control. Springer Berlin Heidelberg, 1985:11-39.

[15] Planned Behavior Theory, MBA think tank encyclopedia, https://wiki.mbalib.com/wiki/ planned behavior theory.

[16] Hu Xiaomei. Research on Influencing Factors of SNS Users’ Privacy Protection Behavior -- Based on the Analysis of Pew Survey Data of Teenagers [J]. Southeast Communication, 2016(10):61-65.

[17] Use and Satisfaction Theory, 360 Encyclopedia, https://baike.so.com/doc/5614999-5827609.html.

[18] Vishwanath A, Xu W, Ngoh Z. How people protect their privacy on facebook: A cost-benefit view[J]. Journal of the Association for Information Science & Technology, 2018, 69(5):700-709.

[19] Structural Equation Model, 360 Encyclopedia, https://baike.so.com/doc/892705-943678.html.

[20] Wang Siwen. The Establishment of Individual Privacy Classification System and Property Right System in the Field of Public Opinion [J]. Journalistic Knowledge,2019(2):54-59.