Personal Information Protection and Interest Balance Based on Rational Expectation in the Era of Big Data

A Case on the Sharing of Mobile Phone Signaling Big Data in Smart City Planning

Yong Lin¹,², Zhenjiang Shen¹,²* and Xiao Teng²

¹ Joint-lab SPSD, Fuzhou University-Kanazawa University
² School of Environmental Design, Kanazawa University
*Corresponding Author, Email: shenzhe@se.kanazawa-u.ac.jp

Received: April 7, 2021; Accepted: July 12, 2021

Keywords: Personal Information Protection and Utilization, Rational Expectation, Application Context, Data Sharing, Value Analysis, Risk Assessment, Interest Balance

Abstract: In the era of big data, personal information has been widely shared and used, which facilitates personal life, social production and public management but also brings the risk of personal information abuse. Personal information has multiple values involving with personality dignity and freedom, economic use, and public management. Meanwhile, the stakeholders relevant to personal information have become more and more diverse, leading to increasingly urgent demand for sharing and using personal information. With the great improvements in the processing efficiency and transmission rate of personal information, it has become much easier to share personal information, which makes the application of the principle of informed consent more difficult. In this circumstance, "Rational Expectation" rule becomes a new option of personal information protection in the era of big data. By assessing the risk of personal information sharing with matrix method in application contexts, it discusses the criteria of risk control under rational expectation rule. If the risk assessed is at the level of low risk, the sharing and use of personal information in this context complies with the rational expectation rule; If the risk assessed is at the level of medium risk, it is necessary to take measures timely and actively to reduce the risk and reassess the risk; If the risk assessed is at the level of high risk, the rational expectation rule is not applicable, the personal information controller should significantly inform the information subject and obtain consent before sharing the personal information. If there are multiple risk points in the application context, when and only when each risk level judged must be low risk, the rational expectation rule can be applied. Based on the rational expectation rule, we can achieve the balance of interests among personal information protection, digital economic development and public interest maintenance, so as to coordinate the promotion of digital innovation, economic development and social progress, and realize the unity of effective protection and rational use of personal information.
1. INTRODUCTION

In the era of big data, large amount of diverse dynamic data is generated in the context of "digitalization". Big data greatly facilitates people's daily life. For example, selection of daily travelling route can be made with the assistance of big data. In the field of market, big data is used for providing personalized services, which enormously improves the operators' profitability and also brings people a higher quality of life. Governments also make use of big data to promote e-government, intelligent management, and social governance, make scientific and rational decisions, and better provide public services.

However, the above big data is often related to personal information. Personal information is generally defined by "identification theory" (Qi & Zhang, 2018). Namely whether a specific natural person can be identified alone or in combination with other information is used as the core criterion for determining personal information(1). Liu (2017) pointed out that all rational means possible to be taken by controller or third-party user of personal information under the conditions at that time should be taken into account to judge whether the information was identifiable; Hon, Millard, and Walden (2011) believed that when processing information for special purposes, non-personal information might also be transformed into personal information. From the perspective of the attribute of right for personal information, people have different understandings of the core interests under the information protection law due to the different legal cultures and social backgrounds of different countries (Boshe, 2015). Yang (2016) reckoned that personal information had connotations of both spiritual personality interests and property interests, but the attribute of personality right was the essential attribute of personal information; Long (2017) thought that a new type of data property right should be constructed on the basis of distinguishing personal information and data assets; Xiang (2018) proposed that personal information property right should be independently confirmed, exercised, and protected, and efficiently configured by market; While, Gao (2018) brought forward the "social cybernetics" of personal information which argued that personal information right no longer only appears as an absolute property right or personality right. For the research of personal information protection and utilization mode, Rice and D'Arcy (2007) raised a differentiated privacy protection strategy which provided customized privacy levels for users’ personal information and measured the value of privacy; Yin and Wang (2016) stated that it is necessary to establish a comprehensive personal data traceability management system to protect personal information; Sun (2016) suggested to establish a tort law system with the end user of information as the responsible person to guide secondary dissemination and utilization of information.

Data Sharing is a key way of utilization of personal information. Data Sharing is the way that data controller shares the collected information with a third party with or without charge, forming a civil and commercial legal relation based on data right distribution between the data controller and sharer. In the era of big data, data resource has become an important production factor. And high-level development and use of data resources are the premise of both information-based improvements in traditional industry and rapid development of modern information service industry. By data

sharing, it is both available to reduce the cost on data acquisition and make full use of resources. The essence of data sharing is the acquisition, transmission, and reuse of personal information. In principle, the sharing of personal information should obtain the informed consent of the information subject\(^2\). According to its sensitivity, personal information can be divided into sensitive personal information and general personal information. Once sensitive personal information is unlimitedly shared, the personal privacy and personal property security of information subject will face serious threats\(^3\). In practice, after sharing personal information, the information subject is possible to lose the right of making decision on the personal information. If the personal information is improperly used, the privacy of the information subject may be infringed and the subject may be discriminated against. Furthermore, it is common that the information subject may be defrauded and subject to personal and property damages in case that the personal information is disclosed. If business managers do not process the shared data as provided by law, it is easy to violate the law and even constitute a criminal offence, such as the crime of infringing citizen's personal information\(^4\). In any case personal information should be used rationally under the guidance of legal rules. The rational utilization of personal information shall be based on lawfulness, justification and necessity, and shall not be excessively processed\(^5\). Legitimacy is the premise and bottom line of the rational utilization of personal information. Rationality is mainly related to the balance of interests of the relevant stakeholders and the maximization of the overall interests under the premise of protecting the core interests of information subjects.

Informed consent is the basic principle of personal information protection, however, it faces the challenges of information overload, status asymmetry, data explosion and rapid transmission in the era of big data (Lv, 2021). As a coping strategy, the protection and utilization of personal information can also take into account whether the processing of personal information in a specific context conforms to the "rational expectation" of the public (Wang, L., 2019). In other words, when personal information is used in a specific context, if the information processing behavior can be expected by the relevant parties, especially the information subject, and this processing behavior can be recognized at the level of general social cognition, then the processing behavior is reasonable, there is no need to obtain the informed consent of the information subject. Rational expectation rule is based on the integration of Contextual Integrity theory and Risk Management theory. Simitis (1999) first proposed Context-oriented Rules in the investigation report on the implementation of Convention for the protection of individuals with region to automatic processing of personal data. And insisted that it is necessary to consider the sensitivity of personal information based on the specific context, and take corresponding protection measures accordingly\(^6\). Nissenbaum (2005) further put forward the theory of "Contextual Integrity", which holds that the specific context of the original collection of personal information should be respected, and its subsequent

\(^2\) Civil Code of the People's Republic of China (2021), Article 1038.
\(^3\) Investigation report on the protection of Chinese Netizens' Rights and Interests (2016).
\(^4\) Criminal Law of the People's Republic of China, Article 253 (1).
\(^5\) Civil Code of the People's Republic of China, Article 1035; Amended Act on the Protection of Personal Information (Japan, 2016), Article 15-19.
\(^6\) Spiros Simitis, Review of the answers to the Questionnaire of the Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS 108), 1999.
dissemination and use should not exceed the situation at that time. That is, in a specific context, the information processing should meet the expectations of the information subject, and the specific information processing should match the specific context. Personal information collected in a specific context shall not be processed beyond that context. She believes that the key to protecting personal information is to ensure the "Contextual Integrity" of the flow and utilization of personal information. The context is composed of many factors, and the influence of different factors is different, and that the risk may be different after the combination. Fan (2016) put forward a idea based on big data context and risk management; Pearce (2017) explored the possibility and feasibility of making interdisciplinary research on personal information protection and risk management; Tian (2018) held that prior informed consent should be transformed into hierarchical consent on the basis of information classification and context-based risk assessment, and one-time consent should be changed into continuous information disclosure and dynamic consent; Ding (2018) presented that the rationality of personal information circulation should be judged in specific context and community.

In the era of big data, personal information has become an important resource for economic production and social management. The interest subjects of personal information are diversified, and the interest demands of each stakeholder are different, or even conflict with each other. Driven by interests, a large amount of personal information is shared and used. But In the above researches, there is still no answer to what kind of context is rational to share and use personal information. In other words, what are the criteria for the rational utilization of personal information through risk assessment? And who will benefit from using personal information in this context?

This paper discusses the protection and rational utilization of personal information based on rational expectation rule and puts forward the criterion of rational expectation by assessing the risk of sharing and using personal information in specific application context, in order to achieve the balance of interests among the personal information subjects, business entities and public managers.

2. RESEARCH APPROACH

Firstly, through the analysis of the value of personal information and the measurement of interests, this chapter defines the core interests of personal information subjects and the relationships among the protection of personal information, the development of digital economy and the maintenance of public interests. For the purpose of protecting personal core interests and balancing the interests among stakeholders, the risk assessment model is constructed:

\[
R (I, T, V) = R (P(T, V), S (T, I)) \tag{1}
\]

- Risk
- Possibility
- Severity
- Threat
- Vulnerability

Then, the criteria for determining rational expectations were discussed through the risk assessment on personal information by using matrix method in specific context. And taking the application of mobile phone signaling big data in smart city planning as an example, the combination of quantitative analysis and qualitative analysis to evaluate the risk of personal information sharing in the specific context of smart city planning and management.

Finally, some proposals would be put forward to realize the balance of interests among personal information subjects, enterprise entities and public...
managers based on rational expectation, so as to protect and make rational utilization of personal information in the era of big data.

3. DIVERSIFIED INTEREST SUBJECTS OF PERSONAL INFORMATION AND THEIR INTEREST DEMANDS IN THE ERA OF BIG DATA

The innovative application of big data stimulates huge economic and social values but also brings legal problems about rational use of personal information (Zhou, 2019). In the process of data sharing, diversified stakeholders are involved and the interest conflict between personal information subjects, business entities and public managers is becoming prominent. There is an urgent need to balance the stakeholders' interests in order to ensure personal information protection and rational utilization.

3.1 Analysis of the Values of Personal Information

Personal information has multiple values, which is not only an important factor affecting the risk assessment of information sharing, but also the basis for stakeholders to achieve interest balance.

3.1.1 Personality Dignity and Freedom Value of Personal Information

Sensitive information in personal information often involves personal privacy and thus is protected by personality right. While personality taking personality dignity and freedom as the value basis is the specific embodiment of personality dignity value and makes it possible to realize independence and free development of personality (Whitman, 2003). The term of "personality dignity" was mentioned in the Charter of the United Nations (1945) for the first time and initially confirmed as a basic human right in form of legal document in the Universal Declaration of Human Rights (1948). Personality dignity itself has specificity and thus that of personal information subject is maintained by law7), which prevents others from infringing sensitive personal information and provides precondition for personal information subject to exercise his/her personality right. A person's

7) Civil Code of the People's Republic of China (2021), Article 109
personality dignity is equally protected by law regardless of his/her family background, wealth and status, which precisely reflects the core value of personality dignity. With the development of society, various new personality interests appear especially in the era of big data. Whether to incorporate the interests in the protection range of personality right or not should be determined on the basis of personality dignity. So, personality dignity and freedom is the core value of personal information.

3.1.2 Economic Use Value of Personal Information

The economic use value of personal information is closely relevant to the substantial improvements in data acquisition capacity, storage capability and computing power based on big data and the resulting changes in industrial production mode and commercial marketing model. With the development of IT and AI technologies, data resource has become one of the five core production factors, and the deep integration of informatization into economic and social development has greatly affected industrial division and the adjustment of business conditions and reshaped the competition pattern of world economy. Supported by big data technology, industrial production can be carried forward on the basis of accurate understanding of consumer demand. Commercial marketing has also been transformed from the early massive advertising to oriented marketing and data pushing in due time, resulting in significant increase in economic efficiency. And the integration, sharing and use of information have become a new profit source for enterprises. Meanwhile, operators have elaborately improved their products and services taking consumer demand as the orientation, which greatly meets consumer demand and improves consumer experience. Overall improvement in the use level and economic value of personal information resources is the objective needs of economic development in the era of big data.

3.1.3 Public Management Value of Personal Information

Personal information contains public management value. Since ancient times, rulers of various countries in the world have always been collecting and using personal information to govern the country. The census conducted by government body is a traditional way of collecting and using personal information. In 4500 BC, the ancient Kingdom of Babylon had conducted national census; and the world first modern census was conducted in America in 1790. By virtue of advanced information technology in the era of big data, governments of various countries can collect and use personal information more quickly at low cost, determine social conditions and public opinions by making extensive specimen analysis, make scientific and rational decision and better boost public management and public service. Personal information of criminal suspect can be used by relevant government sector to effectively trace the clues and make information analysis, so as to hit and prevent crime and fully guarantee public safety. Meanwhile, personal information of epidemic patient can be applied in public health field to timely trace and cut off the transmission route of the epidemic and effectively prevent the transmission of pandemic such as COVID-19 (Benreguia, Moumen, & Merzoug, 2020).

8) Civil Code of the People's Republic of China (2021), Article 990
3.2 Diversified Interest Subjects of Personal Information and their Interlaced Interest Demands

Due to the multiple values of personal information, diversified interest subjects have different and even conflicting interest demands for personal information. According to the analysis on the values of personal information, the interest subjects of personal information can be classified into personal information subjects, business entities and public managers. Among them, business entities can be further divided into the traditional industrial operator transforming to information-based operation model and the service provider engaging in supply of data resources, while public managers also include the central government, local government and third-party institutions, such as industry associations authorized by government.

3.2.1 Interest Demands of Personal Information Subject

First, personal information subject not only treat personal information as a kind of "data resource" but also has the demand for protecting the sensitive information involving personal privacy and the personality interest contained in his/her personal information by law. In the context of big data, the terminals of Internet of Things can collect personal information in almost the entire time and space. Consequently, the virtual nature and anonymity of cyberspace and the convenience in information transmission stimulate more intensified demand for protecting personality dignity. And this demand has evolved into general demand of the whole society (Wang, L., 2021a). Second, with the economic use value and public management value as discovered by big data mining technology, non-sensitive information in personal information has become a "data asset" that can bring huge interest and the personal information subject should have the right to enjoy the interest and process it. Third, in the era of big data, the entire society has intensified dependence on information resource. Even the persons involved in the informatization progress may use relevant information service provided by public and private institutions to meet their daily living needs. Hence, individual is required to alien certain personal information to meet the entire society's demands for information product and service supply.

3.2.2 Interest Demands of Business Entities

With the development of digital economy, business entities have increasingly high demand for using personal information. Traditional business operators can collect and process customers' personal information to know about the market demand to make targeted production and sales plan and thus improve the profitability. Due to the extensive demands for personal information, personal information service providers specialized in collection, storage, transmission and supply of information have also emerged. By collecting large amount of personal information and analyzing and digging the information by certain rules, they have constructed various databases to provide information service with or without charge and meet the special and individualized demand of economic production for data resources. Meanwhile, large e-commerce transaction platform and social networking services collect massive personal information from the extensive transaction data and social data by virtue of their channel advantages so that

10) General Data Protection Regulation (EU, 2018)
they can provide extended information service and serve as specialized information providers to obtain excess earnings beyond the main business, while improving their service quality and economic efficiency.

3.2.3 Interest Demands of Public Managers

It has become a common practice in countries all over the world to collect and use necessary personal information under legal framework to better provide public service and administrate the society (Zhang, J., 2021). As government management and service is oriented to people, various personal information is surely necessary to be obtained and positively used in order to guarantee public safety, provide public service and realize effective social governance, which is also the intrinsic demand of "people-oriented government". The public managers represented by the government are always the greatest collectors and users of personal information and the protectors of such information (Zhang, X., 2015). Public power cannot be exercised without written regulations in law. As the authority of public power, government is responsible for protecting citizens' personality rights and property rights and interests and cannot collect and use personal information at will, and its way and extent of using personal information should be restricted by law. Meanwhile, the collection and disposal of personal information is restricted legally by government bodies, which is not only a reflection of respecting human rights but also is needed for protecting equal competition in the market and more for maintaining social stability and the validity of the government's political power. The central government should formulate guidelines to ensure that business entities act appropriately and effectively, and support the personal information protection measures formulated or implemented by local governments to seek proper processing of personal information. The Local government can use large amount of personal information to improve public management and service quality, and also promote economic development by regulating business entity's rational use of personal information while fully protecting sensitive personal information to maintain social stability. The government can also authorize professional third-party organizations, such as industry associations, to formulate the implementation norms of personal information protection and reasonable utilization in relevant fields, so as to make the protection and utilization of personal information more practical.

In summary, personal information bears personal interests, economic interests and public interests. Respecting and protecting personality dignity is the core interest demand of personal information subject, and also the premise of the protection and rational use of personal information.

4. RATIONAL EXPECTATION AND THE RISK ASSESSMENT ON PERSONAL INFORMATION IN SPECIFIC APPLICATION CONTEXT

Informed consent is the basic principle for personal information protection[11]. In the era of big data, with the breakthrough in computing power and the emergence of Internet of Things and 5G communication technologies, the processing efficiency and transmission rate of personal

---

information have been greatly improved. Followed by, data is shared for many times and used for uncertain purposes. This case increases the difficulty in applying the principle of informed consent so that the traditional protection model oriented to informed consent is confronting with great challenge (Fan, 2016). To evade legal risk and meet legal requirement for informed consent, business entities often list tedious and obscure privacy policies that are difficult to be read and understood by users (McDonald & Cranor, 2008). In order to use the product or service, user has to passively accept the privacy policies so that the policies become non-sense and cannot provide personal information subject substantial guarantee (Choi, Park, & Jung, 2018).

In need of using massive dynamic data, for instance using mobile phone signaling related big data to make planning for intelligent city (Manfredini, Pucci, & Tagliolato, 2014), obtaining informed consent of the information subject is less feasible in practice. Hence, only on the principle of informed consent, it has already been difficult to adapt to the current economic and social development trends and not available to effectively balance the interests between personal information subjects, business entities and public managers.

4.1 Rational Expectation Rule: New Option of Personal Information Protection Model in the era of Big Data

In the era of big data, the rule of rational expectation has become an innovative option of personal information protection (General Data Protection Regulation (EU, 2018); Consumer Privacy Bill of Rights Act (Draft) (USA, 2015)). Whether a personal information processing behavior gets in breach of the rights and interests of the information subject can be judged by measuring the information subject's rational expectation for protection and use of personal information in specific application contexts (Nissenbaum, 2009). In other words, in specific context, a behavior of collection and sharing of personal information is rational even without informed consent of the information subject, provided that the behavior is expectable for relevant parties and can be generally accepted by the society and the risk of sharing or collecting information is under control. From the perspective of interest balance, the target of personal information protection is to make rational use of personal information on the premise of strictly protecting personality dignity of the information subject. Definition of the circumstances of rational use constitutes the "boundary" of personal information protection. What rational expectation is on earth and whether the expectation complies with common social cognition or not should be judged by making risk assessment on the shared personal information in specific context, namely admitting the necessary existence of risk and getting the risk controlled within acceptable range.

4.2 Risk Assessment on Personal Information in Specific Context

Risk assessment on personal information is a process to verify the legality and compliance of personal information processing behavior, judge the risk level of this behavior to harm legal rights and interests of the personal information subject, and assess the effectiveness of various
measures taken for protecting the subject\textsuperscript{13}. This risk assessment should obey the laws or the industrial standard formulated with legal authorization, integrate many subjective and objective factors, make specific survey on the protection and use of personal information in specific application contexts and analyze and judge the possibility of the behavior to cause risk and the severity of the harm caused by the risk. European Data Protection Board suggests that data protection impact assessment (DPIA) should be carried out before processing any data involving innovative technologies (e.g. AI and machine learning) and sensitive personal information (e.g. biometric data, health data and credit data) and tracing a person's location or behavior\textsuperscript{14}.

4.2.1 General Flow of Risk Assessment on the Sharing of Personal Information

At present, European and American countries have established their specific information risk assessment standards which provide risk prevention and control tools for protecting personal information\textsuperscript{15}. In specific context, general flow of the risk assessment is as shown in Figure 2

The first is to comprehensively sort out the personal information to be assessed and form a complete list of shared data and data flow charts, and focus on describing the type, amount, sensitivity, anonymization, and cross-border transfer of personal information.

The second is to confirm with the data sharer the application contexts of the personal information to be assessed and focus on describing the identity of third party of the sharing, the purpose of processing personal information, detailed processing method, safety measures to be taken in the processing process, the persons having access to personal information, the status of third party's access to information systems, the list of interfaces for external transmission of personal information, data storage and deletion plans, and the disposal of storage media, and so on.

The third is to identify possible threats (including threat sources, affected objects, occurrence probability and frequency) in personal information processing in this context, generally from the aspects of network environment and technical measures, personal information processing flow, third parties and participants, business characteristics, scale, and security situation\textsuperscript{16}.

The fourth is to analyze possible threats and assess the existing vulnerabilities in this context, judge the efficiency of current security measure for guaranteeing data sharing (security management guarantees and security technical safeguards for special data recipients) and assess whether the rights and interests of personal information subjects are possible to be harmed or not.

The fifth is to assess the values of personal information and the interests of personal information subject, and analyze the possible impact of the

\textsuperscript{13} Information security technology—Guidance for personal information security impact assessment (China, GB/T 39335-2020)
\textsuperscript{14} Guidelines on Data Protection Impact Assessment (DPIA) (European Data Protection Board, 2017)
\textsuperscript{15} Information technology-Security techniques-Privacy impact assessment (ISO/IEC FDIS, 2017); Handbook on Security of Personal Data Processing (ENISA, 2017); Guidelines on Data Protection Impact Assessment (ICO, 2017)
\textsuperscript{16} Information security technology—Guidance for personal information security impact assessment (GB/T 39335-2020)
threats identified in this context on the rights and interests of personal information subject and the impact extent. This impact can generally be measured from the following dimensions: the restriction in information subject's right to make decisions independently, in fringe personality dignity (e.g. the possibilities to trigger discriminatory treatment), and harm personal and property safety.

The sixth is to generate a two-dimensional risk judgment matrix based on two results to judge the risk level of the personal information processing and judge whether the big data is used within safe range of rational expectation or not.

The seventh is to put forward applicable improvement suggestions to complete the security guarantee measures and finally form the assessment conclusion.

4.2.2 Calculation model of the risk of personal information sharing

To calculate the risk, it is necessary to determine the factors that affect the risk, the interaction between the factors and the specific calculation method. In the risk assessment of personal information sharing, the impact factors involved in the calculation of risk value are generally information subject interest, threat and vulnerability, and the interaction between the factors is shown in Figure 2. Firstly, the interest, threat and vulnerability of personal information are identified (according to the existing security measures). Then, through the impact of possible threats and the vulnerability of existing security measures, the possibility of damage to the rights and interests of personal information subject is determined. And through the impact of threats and the value of information subject interests to determine the extent of damage to the rights and interests of personal information subject. The damage to the rights and interests of the personal information subject generally includes the restriction of the information subject's independent decision-making power, the violation of humanity dignity (such as causing discriminatory treatment), and the injury of personal property.

The basic model of risk calculation is as follows:

\[ R (I, T, V) = R (P(T, V), S(T, I)) \]  

Among them,

- \( R \) - Risk of damage to the rights and interests of personal information subject in a specific context
- \( P \) - Possibility of damage to the rights and interests of personal information subject caused by threat exploiting vulnerability
- \( S \) - Severity of damage to the rights and interests of personal information subject caused by the threat

\[ \text{Figure 2. General Flow of Risk Assessment on the Sharing of Personal Information} \]
I - Interest of personal information subject
T - Threat considering the probability of threat occurrence and its harmfulness
V - Vulnerability according to existing security measures

The values of the three factors I, T, and V can be determined through statistical analysis or professional experience.

The calculation of the three functions of R, L, F can adopt qualitative analysis method or quantitative analysis method, or combine qualitative and quantitative.

The Matrix Method is a combined method of qualitative and quantitative that is commonly used in the current risk calculation. The Matrix Method is mainly applicable to the case that the target element value is determined by two known element values. Firstly, a two-dimensional calculation matrix is constructed. The value of each element in the matrix can be determined according to the specific situation, and it does not necessarily follow a unified calculation formula, but it must have a unified increasing and decreasing trend. Then, the values of the two elements determined are substituted into the matrix for comparison, and the intersection of rows and columns is the target element, the calculation result of prime value. The characteristic of Matrix Method is that it can clearly list the change trend of elements by constructing the calculation matrix of pairwise elements. It has good flexibility and been widely used in risk analysis.

4.2.3 Risk Levels of Sharing Personal Information

By identifying the risk source of data sharing in specific context and taking current personal information guarantee measures, the possibility of the sharing to harm the information subject's rights and interests was assessed and generally divided into three levels: considerably possible, rationally possible, and less possible. By analyzing the types of harm that data sharing in the specific context might cause to the information subject's rights and interests on the basis of the possible impact of the identified risk on the information subject's rights and interests, the harm degree was comprehensively assessed and generally divided into three degrees: severe harm, moderate harm, and small impact. Further, by analyzing the possibility to harm the personal information subject's rights and interests and the harm degree, a two-dimensional matrix was constructed as shown in Figure 3 to judge the risk level of sharing data in this context.
### Figure 3. Matrix for Accessing the Risk Level of Sharing Personal Information

<table>
<thead>
<tr>
<th>Severity of damage</th>
<th>Possibility of damage</th>
<th>Less possible</th>
<th>Rationally possible</th>
<th>Considerably possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe harm</td>
<td>Low risk</td>
<td>High risk</td>
<td>High risk</td>
<td></td>
</tr>
<tr>
<td>Moderate harm</td>
<td>Low risk</td>
<td>Medium risk</td>
<td>High risk</td>
<td></td>
</tr>
<tr>
<td>Minimal impact</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td></td>
</tr>
</tbody>
</table>

#### 4.3 Judgment of Rational Expectation

If the information sharing risk assessed (as Figure 3) is at the level of low risk, the sharing and use of personal information in this context complies with the rational expectation. If the risk assessed is at the level of medium risk, it is necessary to take measures timely and actively to reduce the risk and reassess the risk. If the risk assessed is at the level of high risk, the rule of rational expectation doesn't apply and the personal information controller should significantly inform the information subject before sharing the personal information. In this case, data sharing can be carried out only after obtaining user's express consent. If there are multiple risk points in the application context, when and only when each risk level judged must be low risk, the rational expectation rule can be applied. It must be noted that the risk assessment and judgement of rational expectation must be made by an independent third-party assessment agency or industry association authorized by law17).

#### 5. INTERESTS BALANCE OF DIVERSIFIED STAKEHOLDERS BASED ON RATIONAL EXPECTATION

In the era of big data, personal information not only has the attribute of personality interest, but also has the attribute of economic interest and public interest. General personal information has the characteristics of sharing and non-exclusive, and has become a data asset that can significantly create economic value and social value. On the basis of protecting the core interests of personality dignity and freedom, the personal information subject will benefit from the rational utilization of personal information, such as obtaining the economic interest and/or convenient service. From the view of economic interest, business entities make full use of personal information to maximize the economic interest from business activities and become an independent stakeholder. To better provide public management and public service, governments substantially take part in the collection and use of personal information, leading to increasingly weakened private right and

---

17) Consumer Privacy Bill of Rights Act (Draft) (USA, 2015)
gradually enhanced public interest of personal information. Governments contemporarily are both user and protector of personal information.

How to balance the interest demands among personal information subjects, business entities and public managers has become a practical problem to be solved urgently. Among the said types of interests, the most important and fundamental interest is personality dignity interest. Strictly protecting personality dignity and freedom is a common value\(^{18}\). And only in this way can it be possible to balance the interests of the stakeholders and make rational use of personal information. Otherwise, other interests will be like water without source. Provided that the core interest of personal information subject is fully guaranteed, it is crucial to construct rules for balancing the interests of stakeholders based on rational expectations so as to protect and make rational use of personal information. (As shown in Figure 4)

![Figure 4. Interest Balance between Personal Information Subjects, Business Entities and Public Managers](image)

For public managers, complete personal information protection laws and regulations or industrial standards should be formulated to strictly protect sensitive information involving personality dignity in personal information. Operable laws and regulations and industrial standards should be formulated to ensure that business entities use big data legally and rationally, specify business entities' behaviors of collecting, using and sharing personal information, restrict personal information from being used for unfair competition in breach of rules, and strictly hit the criminal behavior that seriously infringes personal information\(^{19}\). In need of collecting, using and sharing personal information for the purpose of public management and service, the public use related law should be formulated in advance and authorized legally, but the capacity of exercising public power should be self-restraint and not over intervene in private rights and interests. To improve the efficiency of public management and public service level, government sectors should authorize partial function of public management to specialized third-party agency such as industrial association to participate in and even lead the collection, use and sharing of personal information in relevant field.

For business entities, it is necessary to obey the laws and regulations and industrial standards for personal information protection, not abuse big data resources to monopolize or use asymmetric information to make

\(^{18}\) Universal Declaration of Human Rights (UN, 1948)

\(^{19}\) Amendment (IX) to the Criminal Law of the People’s Republic of China (2015)
discriminatory transaction, so as to maintain a good business environment. Business entities should respect the private rights and interests of information subjects and improve the level of guaranteeing personal information safety from the perspectives of operation concept, safety regulations and technological means. Personal information should be obtained or shared legally and used rationally and appropriately\textsuperscript{20)}. The personal information subjects should have necessary rights to control the personal information, such as the right to know, the right to correct, the right of being forgotten, restricted processing right, the right to carry and the right to reject\textsuperscript{21)}. Business entities should comply with general social expectation for protecting and using personal information and give the information subject rational price (economic compensation or convenient service). Business entities should not blindly pursue maximum use of personal information. Only if personal information is protected legally can it be possible to get consumers' trust, maintain fair competition between information service providers and maintain the interests of the business entity itself.

Personal information subjects should maintain "rational expectation" for use of personal information in the big data information-based society\textsuperscript{22)}, understand that only if being used collectively can personal information give play to its economic value and social value to the maximum extent. For general personal information not involving personal privacy, the information subject can alien it to business operator for rational sharing and use, provided that there are credible laws and technologies to guarantee the security, in order to obtain convenient life and corresponding economic price. Based on public interest and on the premise of getting legal endorsement, public sectors are allowed to collect personal information to obtain better public service and public security guarantee. Sensitive information involving personal privacy is relevant to freedom and dignity of the information subject and thus should be strictly protected by law and should not be processed at will even getting consent of the right holder (Wang, L., 2021b). It is neither allowed to infringe the personality dignity of the information subject for any reason or in any way nor allowed to break the baseline of law and ethic\textsuperscript{23)}.

6. **CASE STUDY: RISK ASSESSMENT AND CONTEXT SIMULATION**

In this paper, we take mobile signaling big data as an example to simulate personal information sharing in specific context of smart city planning, and discuss the application of rational expectation rule based on risk assessment in specific context.

6.1 **Description of the Context**

Smart city is an advanced concept of modern social development, is the product of the organic combination of people-oriented city and information

\textsuperscript{20)} Civil Code of the People's Republic of China (2021), Article 111, Article 1035.
\textsuperscript{21)} General Data Protection Regulation (EU, 2018)
\textsuperscript{22)} Consumer Privacy Bill of Rights Act (Draft) (USA, 2015)
\textsuperscript{23)} Civil Code of the People's Republic of China (2021), Article 991/992
\textsuperscript{24)} The "13th Five-Year" National Informatization Plan, China, 2016
city, which is to use big data, internet of things, artificial intelligence and spatial information integration technologies to achieve the humanization, intellectualization, systematization and sustainability of urban system. The core of smart city planning is people-centered. It is based on the analysis framework of space-time-behavior, and makes "dynamic" analysis of urban system with the support of advanced information technology, by mining, relating, identifying, and integrating all kinds of big data (especially including personal information). It combines human elements to improve the intelligence and sustainability of urban system spatial structure analysis (Niu, Wang, & Ding, 2017), urban land functional area identification (Jin, Chen, & Sun, 2018), urban construction environmental assessment (Wang, D. et al., 2015), urban activities and community differentiation (Yan et al., 2018), and so on. As the control instruction in the mobile communication system, the initial technical purpose of mobile phone signaling is to control the connection of the channel and transfer the network management information to maintain the normal operation of the communication system. With the characteristics of full sample, low cost, massive, dynamic, fast and continuous, mobile phone signaling big data is widely used in smart city planning. Mobile phone signaling big data not only has time and space dimensions, but also has significant human behavior attributes. Even if the mobile phone signaling big data has been processed anonymously, it will inevitably show some specific location attribute information of users. Once leaked or abused, it is easy to infringe personal privacy of information subject. Only rely on anonymization rule to share mobile phone signaling big data is not enough to protect personal information security in smart city planning (Lin, Shen, & Teng, 2021). Therefore, taking mobile phone signaling big data as an example, the following simulates and discusses the application of rational expectation rule in smart city planning sharing and using mobile phone signaling big data.

Application context 1: Measure the hierarchical structure of urban system
Big Data: mobile phone signaling big data
Data size: exceed 100 million
Number of information subjects involved: 1.39 million
Sharing mode: operation in the physically isolated data system of Telecom Operator
Grid cell size: take town as a cell
Study area: Chang-jiu urban agglomeration, China
Sources: China Unicom

Application context 2: Study the spatial characteristics of urban activities and community differentiation
Big Data: mobile phone signaling big data
Data size: exceed 100 million
Number of information subjects involved: 1.65 million
Sharing mode: transmission to the data system controlled by researcher
Grid cell size: 200m×200m
Study area: Changchun, China
Sources: China Mobile

6.2 Identification of Factors Affecting Risk

6.2.1 Identification of Personal Information Subject Interest

The interests of personal information subject generally include the interest of personality interests (personality dignity and freedom), economic benefits
and convenience services. Mobile phone signaling big data can generate customer location information and activity track information, which involves personal privacy, has high sensitivity and significant value of personality interests; however, in smart city planning, sharing mobile phone signaling big data has no direct significance for information subjects to obtain economic benefits and convenient services. Therefore, the important interest of personal information subject identified in this context is the value of personality dignity and freedom (I1).

6.2.2 Identification of Threats

Generally from the network environment, personal information processing flow, way of data sharing with third parties and other aspects to identify the threat of data sharing.

The network environment of the information system dealing with personal information includes internal network or Internet. Different network environments face different threat sources and the threat of the information system connected to the Internet is higher. In view of the processing big data of mobile phone signaling is generally carried out in the private network, this threat may not be considered.

In specific application context, the most important threat (T1) is the information recognition of mobile phone signaling data in the process of personal information processing. The threats to the big data recognition of mobile phone signaling include trajectory identification, workplace-residence identification, accurate positioning to specific individuals, user portraits, etc., as well as high-frequency tracking or long-time monitoring of the whereabouts of information subjects.

The way of sharing data with third parties is also an important aspect that poses a threat (T2). Confirm the data interaction mode between the mobile phone signaling data system of telecom operators and the third party data processing system. If the telecom operators only open part of the platform to the third party, and the third party processes the data on the physically isolated data system designated by the telecom operators and obtains the calculation results, the threat is mainly reflected in the code and plug-in of the third party, and the threat is low; if the telecom operators share the data by transmitting mobile phone signaling big data to the third party, the threat increases significantly It involves whether the third party receiving personal information strictly implements the contract agreement, whether the purpose of use will be changed, whether the storage time of personal information is minimized, whether it is timely deleted beyond the time limit, and whether the necessary security management measures are formulated and implemented according to the business security requirements.

6.2.3 Identification of Vulnerability

It mainly focuses on the vulnerability brought by the spatio-temporal attribute and data scale of mobile phone signaling big data in the specific context of smart city planning using mobile phone signaling big data. Mobile phone signaling big data belongs to spatio-temporal big data, which has time and space dimensions, can reflect the spatio-temporal position, and is easy to be used by attackers. It is the first vulnerability (V1). Using mobile phone signaling big data for smart city planning, the number of data is often more than 10 million or even hundreds of millions, and the scale of information
subject is more than 1 million. Once leaked, the consequences will be serious. It is the second vulnerability (V2).

Based on the generality of the context risk assessment (rather than individual case), and considering the technical strength and management ability of telecom operators, this paper assumes that the mobile phone signaling big data has been De-identified before sharing; The border protection equipment has been deployed at the network boundary, the border protection strategy has been configured, and the data leakage prevention and intrusion prevention technical measures have been implemented; A complete network security incident warning, emergency response, notification and reporting mechanism has been established; Carry out regular security inspection, evaluation and penetration test for information system, and timely update patches and reinforce security; Have signed confidentiality agreements with relevant personnel engaged in personal information processing positions, and conducted background checks on a large number of personnel who have contact with sensitive personal information; Have carried out professional training and assessment of personal information security for relevant personnel in personal information processing positions; Have signed a binding contract and other documents with a third party, stipulating the purpose and method of processing personal information after it is transferred to the third party, as well as the data retention period and the processing method after the data exceeds the period. That is, the risks that may be caused by the above factors are not considered.

In summary, in the above application contexts, personal information subject interest I1 faces two main threats: information identification T1 and data transmission T2. The threat T1 exploits the vulnerabilities are data spatio-temporal attributes V1 and data scale V2, and the threat T2 exploits the vulnerability is data scale V2. So there are three risks in personal information sharing.

6.3 Assessment of Factors Affecting Risk

All kinds of factors are divided into five levels according to their degree, which are represented by 1-5 values from small to large. If the influence of factors is not significant, it will not participate in the risk calculation. (As shown in Table 1)

Personal information subject interest I1:
The value of personality dignity and freedom is the core value of personal information. Let I1 = 5.

Threat T1:
In application context 1, taking the town as a unit for trajectory identification, the granularity is large and the accuracy is low. It is difficult to identify the identity of information subject. Let T1=1.

In application context 2, taking the scale of 200m×200m as a unit to identify workplace and residence, as well as to identify the place of consumption and leisure. It can accurately locate the residential area, work place, daily consumption and leisure place, and it is very easy to identify the identity of the information subject, or even make a portrait of the information subject. Let T1=5.

Threat T2:
The above two context have not specified the specific sharing way of mobile phone signaling big data. For case comparison and analysis, it is assumed that the first application context is that the third party processes data on the physically isolated data system designated by the telecom
operator and obtains the calculation results. Let $T_2=1$. And it is assumed that the second application context is that telecommunication operators share data by transmitting mobile phone signaling big data to the third party. Let $T_2=3$.

**Vulnerability $V_1$:**

Spatio-temporal attribute is the inherent attribute of mobile phone signaling big data. The positioning accuracy of mobile phone signaling big data depends on the base station density. Generally, the accuracy is about 200m, and the vulnerability is moderately controllable. Therefore, $V_1=3$ is set for both application contexts.

**Vulnerability $V_2$:**

The characteristics of mobile phone signaling big data are full sample, dynamic and continuous, which will inevitably lead to a large number of data to be used. The number of data used in application context 1 is more than 100 million, and the number of active users recorded is 1.39 million. Although the number of data is not explained in application context 2, 1.65 million people are identified from the research results and the data time span is 14 consecutive days, so the number of data should also exceed 100 million. Referring to the *Guidelines for Personal Information Security Impact Assessment*, the scale of processing personal information exceeds 1 million people, its vulnerability is moderate. And the data has been anonymized, its vulnerability can be reduced as appropriate. Therefore, $V_2=2$ is set for both application contexts.

<table>
<thead>
<tr>
<th>Table 1. List of assignment of risk factors under specific application contexts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I</strong></td>
</tr>
<tr>
<td>Application Context 1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>$I_1=5$</td>
</tr>
<tr>
<td>$T_2=1$</td>
</tr>
<tr>
<td>Application Context 2</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>$I_1=5$</td>
</tr>
<tr>
<td>$T_2=3$</td>
</tr>
</tbody>
</table>

### 6.4 Risk Calculation of Personal Information Sharing

The calculation process of the three risk values is similar. Now take the threat $T_1$, the interest $I_1$ and the vulnerability $V_1$ of application context 1 as an example to calculate and demonstrate.

Calculating the possibility of the harm of the rights and interests of the subject of personal information

Firstly, the possibility matrix of harm occurrence is constructed according to the empirical function $z_{ij}=x_i\times y_j$, as shown in Table 2. Substituting $T_1=1$ and $V_1=3$ into the matrix, the probability of harm was determined as 3.

$$P_1(T_1, V_1)=(1, 3)=3$$

<table>
<thead>
<tr>
<th>Table 2. Possibility matrix of harm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vulnerability Severity (V)</strong></td>
</tr>
<tr>
<td>Threat Impact level (T)</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>
Secondly, the possibility of harm is graded. As shown in Table 3, the possibility of harm is graded as "Less possible".

<table>
<thead>
<tr>
<th>Possibility level</th>
<th>Less possible</th>
<th>Rationally possible</th>
<th>Considerably possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possibility value</td>
<td>1-5</td>
<td>6-15</td>
<td>16-25</td>
</tr>
</tbody>
</table>

Calculating the severity of the harm to the rights and interests of the subject of personal information

Firstly, the severity matrix of harm is constructed according to the empirical function \( z_i = x_i + y_j \), as shown in Table 4. Substituting \( T_1 = 1 \) and \( I_1 = 5 \) into the matrix, the severity of harm was determined as 6.

\[
S_1 (T_1, I_1) = S_1 (1, 5) = 6
\]  

<table>
<thead>
<tr>
<th>Personal Information Subject Interest (I)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threat Impact level (T)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

Secondly, the severity of the harm is classified. As shown in Table 5, the severity of the harm is "Moderate harm".

<table>
<thead>
<tr>
<th>Severity level</th>
<th>Minimal impact</th>
<th>Moderate harm</th>
<th>Severe harm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity value</td>
<td>1-4</td>
<td>5-7</td>
<td>8-10</td>
</tr>
</tbody>
</table>

6.5 Risk Level Judgment

Application context 1: Based on the comprehensive analysis of the possibility and severity of the harm to the rights and interests of the personal information subject, according to the risk calculation model \( R (P(T, V), S(T, I)) \), and comparing with the risk judgment matrix (Figure 3), it is judged that the level of risk \( R_1 \) of personal information sharing is low risk.

Similarly, according to the above risk calculation process, it is judged that risk \( R_2 \) is low risk and risk \( R_3 \) is low risk.

To sum up, in the application context 1 "Using mobile phone signaling data to measure the hierarchical structure of urban system", the risk level of personal information sharing is low. Using mobile phone signaling data to measure the hierarchical structure of urban system, personal information can be used based on rational expectation rules.

Application context 2: Using the same method, it is judged that risk \( R_1 \) is high risk, risk \( R_2 \) is high risk, and risk \( R_3 \) is high risk. In the application context 2 "Using mobile phone signaling big data to study the spatial characteristics of urban activities and community differentiation", the risk level of personal information sharing is high risk. Using mobile phone
signaling big data to study the spatial characteristics of urban activities and community differentiation is not applicable to rational expectation rule.

6.6 Discussion and Suggestion

The rational expectation rule can only be applied when and only when all the risks judged are low risks. Obviously, in application context 2, it does not meet the conditions of reasonable expectation rule. The mobile phone signaling big data controllers should significantly inform the information subject before sharing it in this context. Through the above case simulation, we can also find that sharing mode and grid cell size are the key points of the rational utilization of mobile phone signaling big data in smart city planning. In application context 2, if the technical conditions and research accuracy permit, the risk can be reduced by optimizing the sharing mode and increasing the grid size to meet the applicable conditions of rational expected rule.

7. CONCLUSION

Personal information includes the values of personality dignity and freedom, economic use, and public management. In the era of big data, personal information has been extensively used in various fields of social life. And stakeholders of personal information have become increasingly diverse. Personality interest is the core interest of the personal information subject. Sharing and making rational utilization of personal information on the premise of guaranteeing core interest of personal information subject and balancing the relation among personal information protection, digital economic development and public interest maintenance has become the intrinsic demand of the rapid development of economy and social. In the era of big data, the informed consent oriented traditional protection model is facing difficulties. While, the rule of rational expectation becomes a key option of personal information protection. The boundary of rational expectation can be judged by assessing the risk level of sharing personal information in specific context. If the personal information sharing risk assessed is at the level of low risk, the sharing and use of personal information in this context complies with the rational expectation. If the risk assessed is at the level of medium risk, it is necessary to take measures timely and actively to reduce the risk and reassess the risk. If the risk assessed is at the level of high risk, the rule of rational expectation doesn't apply and the personal information controller should significantly inform the information subject before sharing the personal information. In this case, data sharing can be carried out only after obtaining user's consent. If there are multiple risk points in the application context, when and only when each risk level judged must be low risk, the rational expectation rule can be applied. Namely, risk is tolerable but should be controlled in acceptable range of low risk, so as to achieve the balance of interests among the stakeholders of personal information.
REFERENCES


