Review on Data Sharing in Smart City Planning Based on Mobile Phone Signaling Big Data From the Perspective of China Experience: Anonymization VS De-anonymization

メタデータ	言語: English 出版者: 公開日: 2021-06-02 キーワード (Ja): キーワード (En): 作成者: Lin, Yong, Shen, Zhenjiang, Teng, Xiao メールアドレス: 所属:
URL	https://doi.org/10.24517/00062380

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 International License.

International review for spatial planning and sustainable development, Vol.9 No.2 (2021), 76-93
ISSN: 2187-3666 (online)
DOI: https://doi.org/10.14246/irspsd.9.2_76

Copyright@SPSD Press from 2010, SPSD Press, Kanazawa

Review on Data Sharing in Smart City Planning Based on Mobile Phone Signaling Big Data

From the Perspective of China Experience: Anonymization VS De-anonymization

Yong Lin[1,2], Zhenjiang Shen[1,2*] and Xiao Teng[2]
1 Joint-lab SPSD, Fuzhou University-Kanazawa University
2 School of Environmental Design, Kanazawa University
** Email: shenzhe@se.kanazawa-u.ac.jp*

Received: *July 20, 2020* Accepted: *Jan 2, 2021*

Keywords: Mobile Phone Signalling Big Data, De-Anonymization, Data Sharing, Sensitive Information, Personal Information Protection, Smart City Planning

Abstract: In the smart city planning based on spatiotemporal big data, the mobile phone signaling big data is the most commonly used data source at the moment. This kind of big data has time and space dimensions and also significant human behavior attributes. According to the relevant Chinese law, the data has been anonymized before sharing, i.e. cannot be identified as a specific individual and cannot be restored again, thus is no longer regarded as personal information. In smart city planning, the mobile phone signaling big data is used to construct the basic dynamic analysis framework of "space-time-behavior". Even if the mobile phone signaling big data has been processed anonymously, it will inevitably show some specific location attribute information of mobile phone users. The anonymous track information can be matched to the corresponding geographical space, so as to mark the active location information of the information subject in a specific period of time. It can easily identify the specific location information such as the job and residence of mobile phone user, and even give user portrait. Existing technology shows that the mobile phone signaling big data is easy to be de-anonymized, and Anonymity rule are not applicable to the sharing of mobile phone signaling big data in the smart city planning. Mobile phone signaling big data belongs to personal sensitive information. Once leaked or abused, it is easy to infringe personal privacy of information subject. Therefore, only using current anonymization means to share the mobile phone signaling big data are not enough to protect the security of personal information in smart city planning, and sharing the mobile phone signaling big data should follow the basic principle of explicit informed consent. In special circumstances or scenarios, breaking through the basic principle of the mobile phone signaling big data sharing should have clear legal provisions and comply with legal procedures.

1. INTRODUCTION

Big data is widely used in various fields of society and in a myriad of ways. Particularly speaking, spatiotemporal big data is the basic support of smart city, and mobile phone signaling big data is the most commonly used spatiotemporal big data when it comes to smart city planning. Smart city planning is a new concept of urban development, which is to use big data, cloud computing, Internet of things, 5G communications and spatial

geographic information integration technology for promoting the level of intelligence in urban planning and management. The core of a smart city is people-oriented. Under the support of the new generation of information communication technology (ICT), it integrates human behavior and activity factors, collects and uses massive data including personal information, and carries out a more dynamic analysis of the urban space, so as to promote the intelligent levels of national spatial planning, natural resource development, municipal facilities improvement, public administration management and a positive transformation on the community livelihood services. All of that can promote the scientific, efficient and sustainable development of the city, and constantly improve the modernization level of the urban governance system and capacity.

In recent years, from the perspective of people-oriented, scholars have performed dynamic analysis of space-time big data based on mobile phone signaling, mining the characteristic information of urban space and population distribution and activity (Shen & Li, 2018), researching in urban system planning (Niu, Wang, & Ding, 2017), urban agglomeration spatial identification (Zhao, P. et al., 2019), metropolitan area boundary division (Wang, D., Gu, & Yan, 2018), urban spatial structure analysis (Niu, Ding, & Song, 2014), urban land functional area identification (Jin, Chen, & Sun, 2018), and urban construction Environmental Assessment (Wang, D. et al., 2015), intensity change of urban activities (Manfredini, Pucci, & Tagliolato, 2014; Ratti et al., 2006), street vibrancy (Long & Zhou, 2016), distribution of residential population (Becker et al., 2011), workplace-residence relationship (Yang, Zhou, & Zhang, 2019; Zhang, T., 2016), travel characteristics (Yuan, Raubal, & Liu, 2012; Lu, Long, & Yu, 2019). By building up the basic dynamic analysis framework of "space-time-behavior" and using mobile phone signaling big data to mine the temporal and spatial characteristics of behavioral trajectories, it is possible to identify the dynamic characteristics of individual behavior activities such as living and working habits and recreation (Zhong, W. et al., 2017).

In fact, the number of institutions that are collecting big data with real-time positioning information to provide personalized services have been increasing. This phenomenon indeed brings more quality and convenience to people, but also shines a light on issues such as illegal collection and use, and large-scale leakage of personal information, leading to serious security threats, discrimination and reputation damage. Discussion on how to legally use big data, not only to meet the needs of economic development and social public interest, but also to protect personal information security, have been gaining a lot of attention recently. Smart city planning base on mobile phone signaling big data inevitably involves data collection, sharing and use. At present, the mobile phone signaling big data used in the field of smart city planning mostly comes from the data shared from mobile communication operators, but it is not ruled out that some data are obtained through a third party, and there is data shared for many times. It is important to considered that mobile phone signaling big data can generate user's trace, which belongs to personal sensitive information. Currently, the sharing and use of such data is mostly in an anonymous way. In the application practice of smart city planning, in most cases, it is necessary to use mobile phone signaling big data to mine the spatiotemporal characteristics of human behavior trajectory, and then match it with geographic information space to construct the basic dynamic analysis framework of "space-time-behavior". Therefore, even if the mobile phone

signaling big data is processed anonymously, it will inevitably show some specific location attribute information of users. In other words, the mobile phone signaling big data can be de-anonymized in smart city planning. Therefore, sharing and using this type of data has the legal risk of violating personal privacy.

At present, few studies have explored the legal and reasonable use of mobile phone signaling big data in smart city planning from the perspective of data sharing and personal information protection. Through the analysis of de-anonymization technology of mobile phone signaling big data, this paper discusses the issues that might arise over the sharing and protection of personal sensitive information, with a focus on mobile phone signaling big data in smart city planning.

2. RESEARCH APPROACH

In smart city planning, the mobile phone signaling big data is used to construct the basic dynamic analysis framework of "space-time-behavior". However, it can easily identify the specific location information such as the job and residence of mobile phone user, and even give user portrait. In present work, we discussed anonymization VS de-anonymization of the data sharing from a view of protection of personal sensitive information.

First of all, previous studies have already proved that mobile phone signaling big data is widely used in the field of smart city planning in China and applied on studies about the relationship between human activities and urban space, it is also used to construct the basic dynamic analysis framework of "space-time-behavior". Then the reason why mobile phone signaling big data becomes the main data source of smart city planning have also been previously analyzed. As a kind of spatiotemporal big data, mobile phone signaling big data not only has time dimension and spatial dimension, but also can reflect user location and behavior attributes, and generate human activity track in urban space-time. Such characteristics are in accordance with the definition of personal information according to the law, so mobile phone signaling big data belongs to this specific category. Anonymization technology is commonly used in the sharing and use of mobile phone signaling big data in China. However, through inquiries on patent websites, it was found that there are already some de-anonymization technologies in the field of smart city planning, which are used for planning practices. Then a simple identification method is proposed, based on anonymous track information that can be matched to a corresponding geographical space, so as to mark the active location information a subject within a specific period of time. This method can easily identify the specific location information (e.g. job and place of residence of the user) and create user portraits. However, it is important to consider the possibility of such features infringing policies of personal privacy. Finally, legal research has clarified that explicit informed consent is the basic principle of data sharing. So, as a kind of sensitive information, mobile phone signaling big data should follow the basic principle of explicit informed consent. It is also important to highlight that breaking consent principles under special circumstances or scenarios should results in clear legal provisions and follow rules that are created based on legal procedures.

3. MOBILE PHONE SIGNALING BIG DATA AS THE MAIN DATA SOURCE OF SMART CITY PLANNING

In this work, mobile phone signaling big data is recognized as a typical spatiotemporal big data, and it is also one of the main data sources for smart city planning in China.

3.1 Overview of spatiotemporal big data

Spatiotemporal big data refers data based on a unified time-space reference (spatial reference system, time reference system), that moves and changes according to a time and space, related to an specific location directly (fixed position) or indirectly (spatial distribution) (Wang, J. et al., 2017). In addition to time and spatial dimension, spatiotemporal big data also has multi-dimensional information such as thematic attribute dimension.

(i) Spatial dimension (Si: X_i, Y_i, Z_i) - refers to data has three-dimensional spatial location or distribution information.

(ii) Time dimension (T_i) - refers to the dynamic changes of data over time.

(iii) Attribute dimension (D_i) - refers to various thematic attribute information loaded on spatial dimension with elements (phenomena) that change with time.

Spatiotemporal big data includes spatiotemporal benchmark data, position trajectory data, geodetic data, remote sensing image data, and spatial media data associated with position (Wang, J. et al., 2017). Compared with other data, spatiotemporal big data is based on unified spatial reference and temporal reference, so it is more complete and structured. Because of the organized structure of this kind of data, its value is high and it can be applied in a myriad of ways. Spatiotemporal big data can be used to describe the information of ground features and human activities, which is an important feature to support of smart city planning and management. It integrated with other types of data to provide a four-dimensional environment composed of spatiotemporal interweaving, and make planning, layout, analysis and decision-making based on unified space-time. From the perspective of visualization, the virtual display capability of spatiotemporal big data best meets the needs of human perception and can intuitively provide the spatial distribution and time identification of data for people to better represent the construction achievements of smart city under multivariate data.

3.2 Why Mobile phone signaling big data

Mobile phone signaling is the control instruction in mobile communication system. It establishes a temporary communication channel between designated terminals and controls the connection of channel and transfers network management information, to maintain the normal operation of the communication system. The basic format of mobile phone signaling raw data includes mobile phone IMSI number, time stamp, cell global identity (CGI), event type and other fields. When a mobile phone user is starting, talking, texting or surfing the Internet, the mobile terminal will then communicate with the transmitting base station to generate signaling data and record the change of a user's behavior. At the same time, the mobile phone signaling can record the user's location in the coverage area of the base station, update the user's

location periodically according to the fixed time interval, and track the user from one base station coverage area to another. Therefore, the mobile phone signaling big data is a type of spatiotemporal big data, moving and changing in time and space and able to reflect the user's location and attributes under a unified spatiotemporal benchmark.

Through the mining of mobile phone signaling big data and by using its continuous changes in time and space, it is possible to generate the user's (information subject) activity track. After that, we can then analyze the function layout and connection of urban space, as well as the relationship among residents' work, residence, rest and space-time, which provides a new research perspective and technical means for smart city planning and public participation in urban management. The Civil Code of the People's Republic of China, personal information refers to various information recorded electronically or in other forms that can identify a specific natural person separately or in combination with other information, including a person's name, date of birth, identity card number, biological recognition information, address, telephone number, e-mail address, health information, and whereabouts information, among others, for instance. Needless to say, the characteristics of mobile phone signaling big data conform to the legal definition of personal information, and obviously belong to the category of personal information.

According to the data released by the Ministry of Industry and Information Technology of China, by the end of 2018, the total number of mobile phone users in China had reached 1.57 billion, and the nationwide mobile phone penetration rate was 112.2 units per 100 people. According to the calculation of 100 signaling data generated by each mobile phone every day, a city with a population of 1 million people has the capacity to generate hundreds of millions of mobile phone signaling per day. Mobile phone signaling big data covers a wide range of time-space. It is almost full sample, and can be obtained anytime and anywhere. It also features important characteristics of being massive, dynamic, fast and continuous. Moreover, compared with the traditional data, the cost of mobile phone signaling big data acquisition is lower. In 2018, the *"Top 10 Big Data Application Institutions of Planning Industry in China"* selection activity sponsored by Urban Data Party of China studied about 230 planning projects based on big data, and evaluated each one from 12 dimensions. The result showed that mobile phone signaling big data has become the most important big data source for urban planning, accounting for more than 30%.

As discussed above, we take Mobile phone signaling big data as case, study for rules of data sharing in application for smart city planning by analyzing anonymization vs de-anonymization from the Perspective of China Experience.

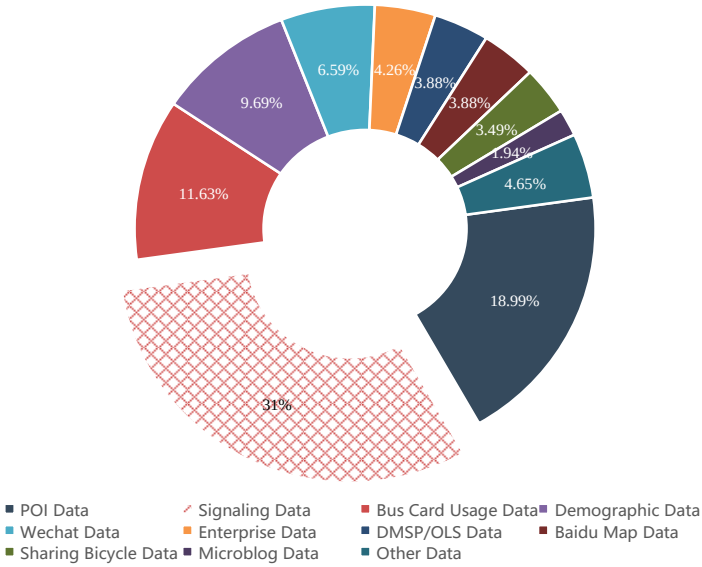

Figure 1. Proportion of major big data sources in smart city planning

4. ANONYMIZATION VS DE-ANONYMIZATION

The advantages of using mobile phone signaling big data to support smart city planning are evident, but sharing and using such information may also causes issues regarding informational leakage. *The Civil Code of the People's Republic of China* stipulates that, without consent of a natural person, no personal information shall be illegally provided for any other person, excluding the information through which the specific individual cannot be identified after processing and which cannot be restored. In order to assure personal information security, anonymous technology is widely used in this realm.

4.1 Anonymization of mobile phone signaling big data

Considering the potential aforementioned issues with leaked information it is necessary that China's network communication operators (personal information controllers) apply appropriate privacy protection technology to preprocess the data before sharing and using mobile phone signaling big data. At present, the commonly used methods can be divided into two categories. The first one modifies the original signaling data as necessary, that is, disturbance, swapping or differential processing, to reduce the accuracy of trajectory in time and space and achieve the purpose of privacy protection. However, this kind of processing technology may cause spatiotemporal data distortion, lowering the confidence of the derived conclusions which does not meet the requirements of smart city planning for data accuracy (Zhou et al., 2009). The second possible method is to anonymize the original signaling data. The so-called anonymization of information refers to a process that makes the information subject's identity impossible to be identified so the processed information can't be restored under the existing technical conditions by blocking the association between the personal information and the

information subject. K-anonymity[①] is a common information protection technology and it is also used to make track information anonymous. For any track, at least k-1 other tracks are transformed into identical anonymous ones to form an anonymous trajectory set. In this case, there are at least k indistinguishable trace records in quasi-identifier, so potential attacker cannot identify the specific subject of a track. K-anonymity uses the parameter K to set the maximum risk of information disclosure. Without background knowledge, the attacker has only 1/K probability to guess the real trajectory of a specific subject, guarantying more possibilities of data protection.

4.2 De-anonymization of mobile phone signaling big data

Everything in the world always mutually reinforces and neutralizes each other, which is like a spear and a shield. The same is true for personal information protection technology. With anonymization technology, de-anonymization technology will inevitably follow. The two compete and promote each other.

4.2.1 De-anonymization technology

De-anonymization is essentially a data mining strategy and a technology that is used to re-identify information subjects from an anonymous data set. For matrix data sets, data sets from different sources are generally used for matching and corresponding relationships are established to achieve de-anonymization. For network data sets, seeds are generally implanted or identified, and node mapping is established to achieve de-anonymization. Taking k-anonymity as an example, if someone wants to identify the information of a specific individual from the data set that has been anonymized, the most common method is to use other source data sets that are known to have overlapping attributes for matching operations. Even if only the limited information records of an information subject are mastered, and the sensitive attributes of the records are not homogeneous or similar, it is still possible to find the corresponding information records from multiple ones by chain attack based on the background knowledge mastered. It is also possible to identify specific individuals, as a way to obtain the privacy information of the subject (Zang & Bolot, 2011).

With the development of machine learning and artificial intelligence technology, the methods for de-anonymization are getting more and more advanced. Recently, Rocher, Hendrickx, and de Montjoye (2019) discovered a method that uses generation model to evaluate whether a person's identity can be re identified from an incomplete anonymous database. Based on these findings, a new AI program was developed to accurately estimate the likelihood of re-identifying individuals through anonymous data sets. The authors claimed that only a few attributes are required to generally re-identify individuals with high confidence even if the data set is incomplete.

① https://en.wikipedia.org/wiki/K-anonymity

4.2.2 De-anonymization technology of mobile phone signaling big data in smart city planning

By consulting the patent announcement website of the State Intellectual Property Office of China②, enter "mobile phone signaling and identification" , as of December 31, 2019, in various scenarios of smart city planning, there are 24 patents (including patent applications) related to re-identification of anonymous mobile phone signaling big data, including 4 authorized patents and 19 effective substantive examinations. The detailed of all the patents are shown below in *Table 1*.

Table 1. List of invention patents for mobile phone signaling big data de-anonymization technology in China

No.	Application Number	Effective Filing Date	Patent Name	Legal Status
1	2015101597226	2015.04.03	Method for recognizing expressway traffic state based on the quality perception of mobile phone signaling data	Authorized
2	2015104524034	2015.07.29	Method for identifying and portraying travel chain of travelers based on mobile phone signaling data	Authorized
3	2015106651759	2015.10.14	Method for road state recognition based on mobile phone signaling	It's deemed to be withdrawn after the disclosure of the invention patent application
4	201510970023X	2015.12.22	Method for resident rail transit travel mode recognition based on mobile phone signaling data	Authorized
5	2016106536280	2016.08.10	Method for mobile phone user travel mode recognition based on mobile phone signaling data and navigation route data	Authorized
6	2017101016886	2017.02.24	Method for user traveling and staying behavior recognition based on mobile phone signaling data	Invention disclosure and effective substantive examination
7	2017101901808	2017.03.27	Method for passenger travel route recognition based on track IC card and mobile phone signaling data	Invention disclosure and effective substantive examination
8	2017103051831	2017.05.03	Method for population recognition based on mobile phone signaling data	Invention disclosure and effective substantive examination

② http://epub.sipo.gov.cn

9	2017109283724	2017.09.22	Mobile high-speed rail user identification terminal based on mobile phone signaling	Invention disclosure and effective substantive examination
10	2017113930854	2017.12.21	Method for bus line recognition based on user mobile phone signaling	Invention disclosure and effective substantive examination
11	2018100275996	2018.01.11	Method for regional congestion recognition based on user mobile phone signaling data	Invention disclosure and effective substantive examination
12	2018100536407	2018.01.19	Method and system for path recognition based on mobile phone signaling data	Invention disclosure and effective substantive examination
13	2018110302330	2018.09.05	Method for airport rail passenger identification based on mobile phone signaling data	Invention disclosure and effective substantive examination
14	2018110303865	2018.09.05	Method for airport passenger travel OD recognition based on mobile phone signaling data	Invention disclosure and effective substantive examination
15	2019100777094	2019.01.28	Method for mobile user payment identification based on mobile phone signaling big data	Invention disclosure and effective substantive examination
16	2019101928922	2019.03.14	Method and system for population identification based on mobile phone signaling data	Invention disclosure and effective substantive examination
17	2019101976184	2019.03.15	Method for identifying source of urban traffic congestion in morning peak hours based on mobile phone signaling	Invention disclosure and effective substantive examination
18	2019102757339	2019.04.08	Method for airport arrival and departure passenger identification and passenger condition analysis based on mobile phone signaling data	Invention disclosure and effective substantive examination
19	2019104674222	2019.05.31	Method for identifying population type of railway transportation hub based on mobile phone signaling data	Invention disclosure and effective substantive examination
20	2019104671205	2019.05.31	Method for crowd type identification based on mobile phone signaling data	Invention disclosure and effective substantive examination

21	2019105298062	2019.06.19	Method for improved mobile phone signaling data travel identification based on dynamic space threshold	Invention disclosure and effective substantive examination
22	2019105298005	2019.06.19	Method for identifying and correcting spatial deviation of mobile phone signaling data	Invention disclosure and effective substantive examination
23	2019106291465	2019.07.12	Method for user activity space identification based on mobile phone signaling	Invention disclosure and effective substantive examination
24	2019106288566	2019.07.12	Method for identifying population migration based on mobile phone signaling	Invention disclosure and effective substantive examination

Recognition technology of anonymous mobile phone signaling big data can be used for de-anonymization in all the different states or applications exemplified above. As recorded in Article 23 of *Table 1 - Method for User Activity Space Identification Based on Mobile Phone Signaling* - the invention "discloses a mobile phone signaling based user activity space recognition method, which realizes the automatic identification of user's activity points by using the spatiotemporal clustering model", "through this identification method, the activity rules of users are identified, and then the user's residence, work and other activity destinations are identified", "improves the authenticity and accuracy of the rules of user activities, and also helps to grasp the activity rules and objectives of users dynamically", "and" it's conducive to the improvement of the scientificity of the planning layout construction of residential land, working land and public facilities within urban space planning " (Zhao, B. et al., 2019).

5. A SIMPLE IDENTIFICATION METHOD IS ENOUGH TO CAUSE PRIVACY RISKS

In the application practice of smart city planning, mobile phone signaling big data is used to mine the spatiotemporal characteristics of human behavior trajectory. The data is then matched with geographic information space to construct the basic dynamic analysis framework of "space-time-behavior". However, there is a simple method to identify work place and residence, as well as other specific location information.

5.1 Workplace-residence identification

Regarding a specific information subject, its activity track generally has relatively fixed characteristics. The anonymous track information can be matched to the corresponding geographical space through space-time analysis and can label the active location data of the information subject. For example, it is possible to identify the residence and work place of mobile phone users (information subjects) by tracking the anonymous trajectory information

regularly. Generally, it is assumed that from 9:00 to 12:00 and from 14:00 to 17:00 in weekdays are the most likely working hours of residents; and from midnight to 6:00 is the most likely the time gap that the users are home. Based on this hypothesis, the stable point with the most active times was observed during work and home hours. Evidently, the point with the most active times within the working period can be basically determined as the user's work place, and the one that is more active within the gap at home time can be basically determined as the user's residence. Then, the user's identity and activity trajectory can be identified with high probability by using the data sets from other sources that contain known identity information and details of one's job/residence location information to attack the identified sensitive location information such as job and residence. Furthermore, the side channels of specific information subjects obtained from public networks and social media can be used to determine more location information. Then, by comparing the trajectory features of anonymous trajectory information of corresponding time points with spatiotemporal analysis technology, the complete historical activity trajectory of users is finally completely tracked, and then identify the user's job and residence information (Zhong, J. et al., 2016). In addition, it is possible to also identify the frequent places of the information subject by localizing the stable points with more active times in each time interval of the track information.

5.2 User portrait and privacy risks

The trajectory information can be used not only identify the user's identity, location and behavior, but also to make a user portrait. For instance, if a mobile phone signaling big data generates a user's daily activity track (as shown in *Figure 2*), the stable point with more active times of the track is investigated. It can be inferred that someone lives in a high-end community by looking at the details of "home time", and that the same person works in a university by looking at the details of "working hours". It is possible to identify that the user has a preschool child since the signal tracks the person going to a kindergarten every day. It is also possible to identify that the person in this case is a female because it frequent visits to a beauty salon. Even the user's consumption preferences can be inferred by looking at the frequent visits to a high-end department store. Thus, we can sketch a detailed user portrait of a white-collar young woman (as shown in *Figure 3*).

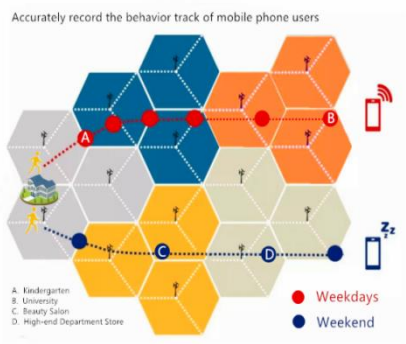

Figure 2. User activity trajectory

Figure 3. User portrait③

③Image Source: Docer Picture Material Library provided by WPS Office Software.

Mobile phone signaling big data can generate activity tracks that have strong identification. Even if the tracks are anonymous, it is easier to be identified. The development of technology for identifying mobile phone signaling big data increases concerns that people might having regarding personal information leakage. Track recognition technology itself is a double-edged sword; the key is to look at what the technology is used to do. When it is used in smart city planning and management, especially in urban public emergency management, it can be very beneficial. During the battle against covid-19 epidemic, for instance, some countries use activity trajectory data to track close contacts,④ effectively blocking the transmission chain, which is a very positive application of such technology (Benreguia, Moumen, & Merzoug, 2020). However, when something is leaked and somehow ends up with online hackers that can infringe one's personal privacy, then we are dealing with huge negative effect of the same tool. According to the *"Investigation Report on the Protection of Chinese Netizens' Rights and Interests (2016)"* issued by the Internet Society of China, 54% of the people in China think that personal information leakage is a serious issue, and 21% of them think it is extremely serious. 84% of the netizens have declared that they personally felt the adverse effects of personal information leakage. Due to the leakage of personal information, junk information and fraud become a rampant issue. The report estimated that the overall economic loss suffered by Chinese Internet users is about 91.5 billion Yuan a year. It also highlighted the possibility that after personal information is leaked, it may be even reselled and transferred many times, causing further harassment and infringement for the information owners. The consequences of this cycle are challenging and the losses that it might cause might not be reversible. With the development of artificial intelligence, cloud computing and other technologies, trajectory identification technology tend to become even more advanced, so it becomes unrealistic to make trajectory data completely anonymous. Moreover, in view of the huge demand for trajectory data in the field of smart city planning, it is also unrealistic to prevent the sharing and use of such information. Taking into consideration the current smart city planning practice in China, only relying on anonymization to share and use mobile phone signaling big data is far from sufficient to protect personal information security. The key to solve this issue is figuring out how to make effective data sharing rules to ensure the security of personal information to the maximum extent.

6. MOBILE PHONE SIGNALING BIG DATA SHARING AND PERSONAL SENSITIVE INFORMATION PROTECTION

Data sharing is the act of sharing the information collected by a data controller with a third party, forming a civil and commercial legal relationship based on the distribution of data rights and interests between the data controller and the sharer. The legal essence of data sharing is the collection, transmission and reuse of information. In the era of big data, relying on technologies such as the Internet of things, artificial intelligence and other related, data users can share data resources more conveniently, reduce the cost

④https://blogs.worldbank.org/eastasiapacific/koreas-response-covid-19-early-lessons-tackling-pandemic

of data collection, realize multiple utilization of data, and extract "Data Gold Mines" to the maximum extent (Li, 2014).

6.1 Sharing of mobile phone signaling big data

As the control instruction of mobile communication system, the initial technical purpose of mobile signaling is to control channel connection and transfer network management information, and also to maintain the normal operation of the communication system. Now the big data technologies are constantly being proved, mobile phone signaling big data, with its full sample, low-cost, massive, dynamic, fast and continuous characteristics, intuitively provides the spatial location, spatial distribution and time identification of data for people. It is widely used in the field of smart city planning, and its technical and economic value has been greatly expanded. In other words, mobile communication operators (data collectors and controllers) share mobile phone signaling big data, while urban planners (data sharers) make secondary development and use of this data. Considering that mobile phone signaling big data can accurately identify personal activity trajectory, and personal activity trajectory belongs to sensitive information⑤, mobile phone signaling big data should be included in the category of personal sensitive information. Once abused, it is easy to infringe personal privacy and cause personal and property interests to be damaged. So, it remains necessary to create more stringent sharing rules of sensitive information to regulate the sharing and the secondary utilization of mobile phone signaling big data.

6.2 General principle of data sharing: informed consent

The principle of informed consent is a general principle for data collection and utilization, and it is considered the most basic applicable rule for personal information protection.⑥ The personality rights and interests of information subject are exclusive. This principle puts forward the idea that the collection and use of personal information shall only happen after the informed consent is obtained from the subject. For information subjects, there is no essential difference between the data sharing behavior and the recollection of personal information (Wang, L., 2019). In fact, the information sharing person is, at the same time, the personal information collector. In this case, the sharing behavior should also be adjusted according to the principle of informed consent. Under normal circumstances, data sharing should have a specific and proper purpose. With the informed consent of the information subject and within a scope of the limited purpose, only the valid information that is necessary to achieve a pre-established purpose should be shared. Without the informed consent of the information subject, data should not be shared and used beyond the scope of the purpose established at the time of personal information collection.⑦

Practices regarding the determination of informed consent for personal information sharing in the US and EU practices are not consistent. The United States has adopted the practice of "implied consent", that is, when data is shared, the information subject should be properly informed of the shared

⑤ *Civil Code of the People's Republic of China*, Article 1034.
⑥ *The Privacy Act(USA)*；*General Data Protection Regulation (EU)*；*Civil Code of the People's Republic of China.*
⑦ *Civil Code of the People's Republic of China*, Article 1035.

object, information category, purpose of use and other related content.⑧ If the information subject does not make explicit objections, it is implicitly determined that the information subject agrees to data sharing. *The General Data Protection Regulation* stipulates that the information subject's explicit consent and authorization must be obtained when data is shared.⑨ No matter what the data source is – whether it was obtained through automatic data processing or not - as long as it is considered personal information, it should be protected. The authors believe that for personal sensitive information, from the perspective of protecting personal information security and reliance interests, when information controllers share personal information, they should use clear and popular language to inform information subjects of the purpose, method and scope of sharing and use it in a comprehensive, accurate and timely manner, It has also been stipulated that the informed consent of the information subject shall be made with a clear intention or active behavior, and that if the subjects remains quiet without refusal, it shall not be regarded as consent, unless otherwise specifically provided by law. Another possibility that would be considered ethical is that if there is a clear agreement between the information subject and the information controller.

6.3 Special rules for data sharing

In some special cases, data-sharing behavior is difficult to fully or timely achieve individual informed consent (Zhang, Y., 2020), so it is necessary to regulate the protection of information subject through some special rules to make up for the flaws of the principle of informed consent application.

6.3.1 First exception to the principle of informed consent: anonymity

In general, personal information is no longer regarded as personal information after being anonymized, which is stipulated in relevant laws on personal information protection in both China and Japan.⑩ After anonymous processing of personal information, if the information subject cannot be identified according to the existing technical means, it can be considered that the association between the information and the individual has been blocked, and it is no longer regarded as personal information.[11] In this case, the informed consent of the information subject is usually not required for data sharing.

6.3.2 Second exception to the principle of informed consent: reasonable expectation

In some cases, when determining if any personal information rights and interests were infringed, it is necessary to consider the reasonable expectations of relevant actors in specific situations (Wang, L., 2019). That is to say, in a specific situation, even without the explicit informed consent of the information subject, a behavior can be considered reasonable as long as the

⑧ *California Consumer Privacy Act (USA)*, Article 1798.
⑨ *The General Data Protection Regulation (EU)*, Article 6.
⑩ *Civil Code of the People's Republic of China*; *Correction of Protection Act on Personal Information*（*Japan*）.
11 *Civil Code of the People's Republic of China*, Article 1038.

information collection can be expected by all parties involved, and this expectation is accepted by general social cognition. For instance, when it comes to smart city planning, in order to optimize urban traffic management and provide timely feedback of road conditions, traffic management departments need to arrange intelligent sensing equipment to collect data of people and vehicles, and hand them over to a professional third party for real-time dynamic analysis. However, in such situations, it can be challenging to obtain the consent and authorization of the information subject. At this time, identifying if the behavior violates personal privacy depends on whether the data collector has conducted the action careful and reasonable, and whether it goes in accordance with the general social cognitive expectations. Therefore, in addition to explicit consent of the information subject, the information controller can also claim the reasonable expectation rule applicable in specific scenarios.

6.3.3 Third exception to the principle of informed consent: public interest needs

If relevant data collection and shared behavior are both on the side of public interest, specific public power departments might have the right to obtain and share information without informed consent of the information subject to a certain extent. However, the behavior must be clearly stipulated by laws and follow the legal procedures, and it also should be conducted under proper supervision and any necessary restrictions. The definition of public interest must also be clearly defined by the law, to avoid the power abuse.

6.3.4 Forth exception to the principle of informed consent: emergency needs

Another situation where it might be ethical to break the informed consent rules happens when relevant data collection and sharing behavior is considered an emergency, in order to better protect the personal interests of the information subject, or to protect the life and health of the third person and other major interests. However, all should still be in accordance with the law and inform to the information subject in an appropriate way afterwards.

6.4 Sharing rules of mobile phone signaling big data

The mobile phone signaling big data involves personal trace, which belongs to personal sensitive information. There are a lot of possible de-anonymization technologies for this type data, so it is impossible to completely block the relationship between information and the subject. Therefore, in the smart city planning, sharing of mobile phone signaling big data should not apply to the anonymity rule. However, considering the current smart city planning practice in China, a large number of “anonymous” mobile phone signaling big data is used without the subject’s consent, evidently has a very big risk of infringing personal privacy in law.

The principle of reasonable expectation involves general cognition and expectation of personal privacy protection recognized by the society. Chinese mobile phone users have a relatively tolerant attitude towards mobile phone signaling sharing and are more likely to exchange part of their privacy as a way to get access to convenient services. However, in Europe and in the United States, users have different perceptions and expectations of the rules

of privacy protection, leading them to act less tolerant (Zeng, 2007). Some people might even prefer to sacrifice convenience so they don't need to make a commitment that is attached to a great amount of personal information. In this case, if the shared information is considered sensitive, the "reasonable expectation" of the subject is quite limited (Wang, L., 2019).

In a sense, smart city planning with spatiotemporal big data such as mobile phone signaling big data is, somehow, a need of public interest. However, there are different definitions of public interest in different nations and society structures. Whether smart city planning belongs to the public interest remains a controversial question according to the law. Unless it is clearly stipulated by national laws and regulations and through legal procedures, the sharing and the usage of mobile phone signaling big data in the field of smart city planning should not exceed the basic principles of personal information protection.

In some countries such as China and South Korea, if the mobile phone signaling sharing behavior is meant to meet emergency need (e.g. tracking activities and close contacts for new coronavirus infected persons, controlling the spread of the epidemic, and protecting the people's basic right to life and health), the regulation of the informed consent principle can be broken through. On the contrary, European and American countries are more inclined to try other technical solution and carry out digital contact tracking with the help of mobile application software and Bluetooth technology (Dai, 2020).

In summary, the sensitive information such as mobile phone signaling big data can generate trace, it is easy to infringe on the personal privacy of the information subject and cause damage to personal and property rights and interests. Sharing mobile phone signaling big data, the information controllers should comprehensively, accurately and timely inform the information subject of the purpose, method and scope of sharing and use in clear and popular terms, obtain the explicit consent and explicit authorization of the information subject in advance, and ensure that the express consent of the personal information subject is an independent and clear expression of will on the basis of full knowledge, and cannot Share the sensitive information again at will or allow others to use it again. In special circumstances or scenarios, breaking through the express informed consent principle of mobile phone signaling big data sharing should have clear legal provisions and comply with legal procedures. It should be recognized that different countries and societies have different perceptions of the sensitivity of specific types of information. In order to protect personal privacy and personal information security, countries should follow their own political ideas, religious beliefs, value orientation, cultural customs to formulate specific rules in the form of legislation to protect personal sensitive information such as mobile phone signaling big data.

7. CONCLUSION

We are now experiencing the fourth technological revolution, and big data has been widely applied in all fields of social life. The sharing and utilization of spatiotemporal big data provides new ideas, new technologies and new models for smart city planning, and also brings a new discussion over the issue of personal information protection. Mobile phone signaling big data is relatively simple to be de-anonymized and belongs to the category of sensitive information. The principle of informed consent is the most basic principle of data sharing. Anonymity rule are not applicable to the sharing of mobile phone signaling big data in the smart city planning. At present, in China's smart city planning practice, if the principle of informed consent is not applied and only

relying on "anonymous" means to share and use mobile phone signaling big data, there is a huge risk of violating the relevant laws of personal information protection. The sharing and using of mobile phone signaling big data and other personal sensitive information in smart city planning should obtain the explicit consent and clear authorization of the information subject. Breaking through the principle of explicit informed consent in specific scenarios or special circumstances, such as smart city planning or urban emergency management, using mobile phone signaling big data should have clear provisions in laws and regulations, as well as to comply with legal procedures.

REFERENCES

Becker, R. A., Caceres, R., Hanson, K., Loh, J. M., Urbanek, S., Varshavsky, A., & Volinsky, C. (2011). "A Tale of One City: Using Cellular Network Data for Urban Planning". *IEEE Pervasive Computing*, *10*(4), 18-26. doi: https://doi.org/10.1109/MPRV.2011.44.

Benreguia, B., Moumen, H., & Merzoug, M. A. (2020). "Tracking Covid-19 by Tracking Infectious Trajectories". *IEEE Access*, *8*, 145242-145255. doi: https://doi.org/10.1109/ACCESS.2020.3015002.

Dai, X. (2020). "Information Governance in 'Pandemic Control State': Practices and Ideas". *Beijing Cultural Review*, (5), 86-94. doi: http://d.wanfangdata.com.cn/periodical/whzh202005010.

Jin, P., Chen, M., & Sun, Z. (2018). "Urban Land Use Functional Area Identification Method Based on Mobile Phone Signaling Data". *Information & Communications*, (1), 268-270. doi: http://dx.chinadoi.cn/10.3969/j.issn.1673-1131.2018.01.133.

Li, J. (2014). "How Do Telcos Mine the 'Gold Mine' of Big Data". *China Telecommunications Trade*, (3), 82-83. doi: http://dx.chinadoi.cn/10.3969/j.issn.1671-3060.2014.03.032.

Long, Y., & Zhou, Y. (2016). "Quantitative Evaluation on Street Vibrancy and Its Impact Factors : a Case Study of Chengdu". *New Architecture*, (1), 52-57. doi: http://dx.chinadoi.cn/10.3969/j.issn.1000-3959.2016.01.009.

Lu, Z., Long, Z., & Yu, Q. (2019). "Analysis on the Job-Housing Spatial Distribution and Commuting Characteristics of Kunshan City Based on Cellular Signaling Data". *Modern Urban Research*, (3), 50-55. doi: http://dx.chinadoi.cn/10.3969/j.issn.1009-6000.2019.03.007.

Manfredini, F., Pucci, P., & Tagliolato, P. (2014). "Toward a Systemic Use of Manifold Cell Phone Network Data for Urban Analysis and Planning". *Journal of Urban Technology*, *21*(2), 39-59. doi: https://doi.org/10.1080/10630732.2014.888217.

Niu, X., Ding, L., & Song, X. (2014). "Understanding Urban Spatial Structure of Shanghai Central City Based on Mobile Phone Data". *China City Planning Review*, (6), 61-67. doi: http://dx.chinadoi.cn/10.3969/j.issn.1000-3363.2014.06.009.

Niu, X., Wang, Y., & Ding, L. (2017). "Measuring Urban System Hierarchy with Cellphone Signaling". *Planners*, *33*(1), 50-56. doi: https://doi.org/10.3969/j.issn.1006-0022.2017.01.008.

Ratti, C., Frenchman, D., Pulselli, R. M., & Williams, S. (2006). "Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis". *Environment and Planning B: Planning and Design*, *33*(5), 727-748. doi: https://doi.org/10.1068/b32047.

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). "Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models". *Nature Communications*, *10*(1), 3069. doi: https://doi.org/10.1038/s41467-019-10933-3.

Shen, Z., & Li, M. (Eds.). (2018). *Big Data Support of Urban Planning and Management: The Experience in China*. Berlin: Springer.

Wang, D., Gu, J., & Yan, L. (2018). "Delimiting the Shanghai Metropolitan Area Using Mobile Phone Data". *Acta Geographica Sinica*, *73*(10), 1896-1909. doi: http://dx.chinadoi.cn/10.11821/dlxb201810006.

Wang, D., Zhong, W., Xie, D., & Ye, H. (2015). "The Application of Cell Phone Signaling Data in the Assessment of Urban Built Environment: A Case Study of Baoshan District in Shanghai". *Urban Planning Forum*, (1), 82-90. doi: http://dx.chinadoi.cn/10.16361/j.upf.201505010.

Wang, J., Wu, F., Guo, J., Cheng, Y., & Chen, K. (2017). "Challenges and Opportunities of Spatio-Temporal Big Data". *Science of Surveying and Mapping*, *42*(7), 1-7. doi: http://dx.chinadoi.cn/10.16251/j.cnki.1009-2307.2017.07.001.

Wang, L. (2019). "Data Sharing and Personal Information Protection". *Modern Law Science*, *41*(1), 45-57. doi: http://dx.chinadoi.cn/10.3969/j.issn.1001-2397.2019.01.04.

Yang, L., Zhou, L., & Zhang, X. (2019). "Research and Evaluation of Jobs-Housing Space Characteristics Based on Mobile Phone Signaling Data: A Case Study of Guangzhou". *Urban Insight*, (3), 87-96. doi: http://dx.chinadoi.cn/10.3969/j.issn.1674-7178.2019.03.008.

Yuan, Y., Raubal, M., & Liu, Y. (2012). "Correlating Mobile Phone Usage and Travel Behavior – a Case Study of Harbin, China". *Computers, Environment and Urban Systems*, *36*(2), 118-130. doi: https://doi.org/10.1016/j.compenvurbsys.2011.07.003.

Zang, H., & Bolot, J. (2011). "Anonymization of Location Data Does Not Work: A Large-Scale Measurement Study". Proceedings of the 17th annual international conference on Mobile computing and networking, pp. 145-156. doi: https://doi.org/10.1145/2030613.2030630.

Zeng, L. (2007). "A Contrastive Analysis of the Chinese and Western Privacy Right". *Journal of Hubei University (Philosophy and Social Science)*, *34*(4), 35-39. doi: http://dx.chinadoi.cn/10.3969/j.issn.1001-4799.2007.04.013.

Zhang, T. (2016). "Job-Housing Spatial Distribution Analysis in Shanghai Metropolitan Area Based on Cellular Signaling Data". *Urban Transport of China*, *14*(1), 15-23. doi: http://dx.chinadoi.cn/10.13813/j.cn11-5141/u.2016.0103.

Zhang, Y. (2020). "Legal Protection of Personal Information Related to the Epidemic in the Context of Big Data". *Henan Social Sciences*, *28*(4), 56-65.

Zhao, B., Tang, X., Gao, Z., & Zhang, J. (2019). "Method for User Activity Space Identification Based on Mobile Phone Signaling: China". (2019106291465),

Zhao, P., Hu, H., Hai, X., Huang, S., & Lyu, D. (2019). "Identifying Metropolitan Edge in City Clusters Region Using Mobile Phone Data: A Case Study of Jing-Jin-Ji". *Urban Development Studies*, *26*(9), 69-79. doi: http://dx.chinadoi.cn/10.3969/j.issn.1006-3862.2019.09.014.

Zhong, J., Chang, S., Liu, X., & Song, H. (2016). "De-Anonymization Attack Method for Mobile Trace Data". *Computer Engineering*, *42*(12), 133-138. doi: http://dx.chinadoi.cn/10.3969/j.issn.1000-3428.2016.12.024.

Zhong, W., Wang, D., Xie, D., & Yan, L. (2017). "Dynamic Characteristics of Shanghai's Population Distribution Using Cell Phone Signaling Data". *Geographical Research*, *36*(5), 972-984. doi: http://dx.chinadoi.cn/10.11821/dlyj201705013.

Zhou, S., Li, F., Tao, Y., & Xiao, X. (2009). "Privacy Preservation in Database Applications: A Survey". *Chinese Journal of Computers*, *32*(5), 847-861. doi: http://dx.chinadoi.cn/10.3724/SP.J.1016.2009.00847.