

A network graphic consisting of interconnected nodes and lines, forming a complex web structure. The nodes are represented by small circles, and the lines are thin, light-colored lines connecting them. The overall shape is somewhat triangular, with the top part being wider and tapering towards the bottom.

**Annual Meeting of the New
Champions 2018**
2018新领军者年会

The 4th industrial revolution and the innovative society
第四次工业革命与创新型社会



Big Data Application and
Urban Public Security

**大数据运用
与城市公共安全**

Annual Meeting of the New Champions 2018

2018 新领军者年会

Big Data Application and Urban Public Security

大数据运用与城市公共安全

Big Data Application and Urban Public Security

Wang Ye Wen Zhiqiang

Tianjin Normal University

Abstract: In face of serious problems in urban public security and growing security governance needs, the application of new technologies represented by big data plays a more and more important role. The existing problems regarding data application in the urban public security includes decentralized data collection, poor data quality, overdependence on artificial intelligence, limited data mining depth and width, low sharing degree and lack of security guarantee for public data. We should adopt big data thinking to solve these problems by making full use of technologies such as internet big data and artificial intelligence to perform intelligent data filtering, classification, study and judgment, smart forwarding. We should also skillfully use “data exhaust” to deeply explore the value of information resources and spare no effort to build a “smart and safe city”.

Through building big data bases, enterprises and disaster prevention and early warning centers, we can actively digitalize urban resources. At the technical and institutional levels, we can

integrate and open public security resources in the process of acquisition, storage, analysis and practical decision-making, so as to realize the cross-regional cross-departmental information sharing; and we also should improve the data service and resources exchange system to gradually build the big data system focusing on urban public security. Meanwhile, this paper also forecasts the application prospects of the latest block chain technology in personal digital credit remodeling, privacy protection, public security information sourcing and public data storage and access management.

Catalog

1	Big Data application and urban public security: background and trend	1
1.1	Increase in urban public security governance needs	1
1.2	Increase in Big Data application in urban public security	1
1.3	Better Big Data security technologies	4
2	Data application in urban public security: problems and barriers	5
2.1	Data acquisition	5
2.2	Data processing	6
2.3	Data use	8
2.4	Data management	10
3	Opportunities for urban public security management in the Age of Big Data	11
3.1	Intelligent filtering and classification of data intelligence ..	12
3.2	Intelligent study and judgment on data	13
3.3	Intelligent forwarding of effective information	16
3.4	Intelligent change of study and judgment model through full data fusion	17
3.5	Deep mining of information value with the help of “Data Exhaust”	18
4	Urban public security-oriented Big Data system construction ..	19
4.1	Strengthen the Big Data base development and digitalize urban	

resources	19
4.2 Open and share data	31
4.3 Application of block chain technology	40

1 Big Data application and urban public security: background and trend

1.1 Increase in urban public security governance needs

General Secretary Xi Jinping said, “Public security is an important manifestation of social stability and social order, and an important guarantee of people’s happy life.” We should firmly establish the concept of safe development, maintain public security to safeguard the fundamental interests of the people, do a solid job in building an all-around and three-dimensional public security net for people’s happy life, social stability and order and for the nation’s long-term peace and stability.

Despite of sound overall situations in urban public security, rapid urbanization and modernization increases the number of incidents, such as Hajj stampede in Mecca, attack on a mosque in Canada, car bombing in Barcelona, Takada’s airbag problem and e quick loan in Huizhou. These problems have become increasingly prominent, demanding better urban public security governance. In face of people’s public security needs, the new technologies play a more and more important role.

1.2 Increase in Big Data application in urban public security

1.2.1 Adoption of Big Data thinking

In the Age of Big Data, whole sample replaced sampling, correlation replaced causality and the analysis of overall

characteristics with overall data takes place of that with sample data, and the passive analysis and initiative transformation are actually applied in the quick judgment on a correlation or a trend. The big data thinking provides a new idea for the public urban security governance. For example, New York City developed a new fire prevention program. In this program, data collected by 19 municipal departments were added to the database of 900,000 buildings in the whole city, including records of detention because of tax debt, abnormal water and electricity use, overdue payment and rat complaints. Besides, through the calculation and analysis of fire records in the past five years, it was found that the building type and construction year are correlated with fire probability and that the buildings with illegal partitions inside have a higher probability. Based on the above findings, a new fire prevention program was made. It is thus clear that in the urban public security management, such all data thinking model and methodology should be deeply rooted in the specific measures, so as to better govern the urban public security.

1.2.2 Application of Big Data technologies

At present, many big data technologies are applied to solve relevant urban public security problems, such as that in the intelligent transportation system. For example, Huawei has expanded itself to the transportation field in recent two years and helped to solve the traffic jam and parking difficulty in China. In China, the

traffic jam cost accounts for nearly 2% of GDP, about RMB 1 trillion Yuan, every year. Such a high traffic jam cost, coupled with the huge traffic flow, severely challenges the urban traffic management. In order to solve this problem, based on cloud computing and big data, Huawei and Shenzhen Traffic Police together raised data efficiency. Huawei put forward NB-LOT intelligent parking, which has three advantages - low cost, convenient deployment and good immediacy. Based on the three advantages, the vehicle detector can directly report information to the operator's wireless network without aggregated gateway, which reduces the total cost. Besides, the vehicle detector is a plug-and-play device, so it can be installed easily without the support of network technicians, and the subsequent intelligent parking capacity expansion is easy. At the same time, mobile APP and parking management data platform are adopted to provide uniform intelligent parking services for the urban residents, which can facilitate the residents and save the management cost. Meanwhile, the roadside parking system can use geomagnetic parking space (built-in NB-LOT communicating module) detector to collect the parking space occupancy state; the parking management platform can automatically generate billing data according to the parking time and charging rule. Based on the above, the users can view information on their mobile phones, such as the system's equipment operational condition, parking space occupancy state and charge details, and are immediately informed of the number of

vacant parking place in the surrounding areas.

1.3 Better Big Data security technologies

In recent years, big data and relevant technologies have improved, becoming a prime power of disruptive innovations in various walks of life. So far, four big data technologies have been widely used in the security. Firstly, big data fusion technology. Cloud storage system can not only realize the effective data fusion and sharing, but also solve the problem of normal storage and recover video data in case of hardware failure. Secondly, big data processing technology. Cloud computing system can be used to structure the unstructured data, and then to realize the analysis and mining of security data. Thirdly, monitoring technology. Computer image vision analysis technology can be used to separate the background from the target in scenes and then analyze and track the target in the camera scenes, including identification of human face, license plate number and traffic lights. Fourthly, big data analysis and mining technology. Based on cloud storage system and cloud computing system, quick analysis on “big data” can be realized, such as cloud-based license plate identification. Through the analysis on massive videos, it can quickly fetch license plate information and form effective “files”, and finally realize the prevention and alarming of incidents and build a security system.

2 Data application in urban public security: problems and barriers

Although the big data application in the urban public security is constantly enhanced and the big data security technology improves, the data application in the urban public security is far from perfect on the whole.

2.1 Data acquisition

2.1.1 Decentralized data collection

In the context of big data, the demands for real-time data processing, security control operation and rapid information synchronization raise a high requirement for data acquisition. However, the urban data collection system is a large-scale decentralized system, and the data are mainly from social networks, mobile devices, monitoring equipment and sensors, so the required data are decentralized and take diversified methods to collect. As a result, fragmentation makes it difficult to collect and analyze the data. In addition, in the urban public security management, various departments in different administrative regions are mutually independent in basic data acquisition with diversified focus, resulting in overlapping information. This not only increases the urban work cost, but also reduces the working efficiency, which is, to some extent, a waste of resources.

2.1.2 Poor data quality

Data quality is the first issue to be considered after data acquisition. As the data become more extensive and complex, the risk of low-quality data generation is intensified. Low-quality data not only undermines the availability of information system, but also seriously affect the sustainable use of data. The quality problems of raw data mainly include noise, abnormal value, numeric value missing and numeric value repetition. In the data acquisition process, manual data entry can easily lead to such errors as input value error and subjectively assumed value. “Dirty” and “ill-conditional” data can be found everywhere, and there is not a universal and effective method to process these data. In many researches about data analysis, only simple manual data cleaning is often adopted. Consequently, the analysis results based on these data are usually incorrect or one-sided. Therefore, much attention must be paid to data quality when big data are used to govern the urban public security.

2.2 Data processing

2.2.1 Complex data cleaning process

Data cleaning is the first step of data processing. It uses computer programs to automatically process the data or uses Microsoft tools, like Excel, to manually process the data. Data cleaning mainly covers data source arrangement, merger and data cleaning. At present, data cleaning cannot be completely done by computers, because they are

not intelligent enough to make decisions about data conversion as the experts in this field. Although there are many mature data cleaning technologies in the database research field, most of them are just for some specific data quality problems. For example, the interactive visualization tool can be used for pattern matching, entity analysis and data cleaning, but the interaction function of these systems is limited. When the data set changes or used by other users, a large amount of time will be consumed in the cyclical and repeated data cleaning. Meanwhile, much attention should be paid to excessive cleaning, which may lead to “dirty” data set. Therefore, we should balance the relationship between data cleaning and data analysis and develop an efficient data cleaning tool.

2.2.2 Overdependence on artificial study and judgment in data processing

In the data processing, because the technologies are not perfect, overdependence on artificial study and judgment reduces the information validity. At present, data analysts subjectively judge the availability of data for the research question, identify data quality problems, and then clean the required data to meet requirements. In addition, overdependence on business personnels’ data information processing may lead ton inconsistency between results and common sense. Lastly, in the specific data mining, a result evaluation is required. To make full use of the data mining results, professional decision-makers are needed to interpret the data and find the

available information. But the data mining results are mainly correlativity, rather than causality, and they may be greatly instable and sporadic, which will lead to information misrepresentation and distortion and thus mislead the decision. If this problem cannot be effectively solved and avoided, no correct result can be obtained even though a scientific data mining is conducted.

2.3 Data use

2.3.1 Insufficient data mining depth and width

In the urban public security, many problems in the data mining process lead to lack of depth and width and use and failure to identify the core value. The problems are mainly as follows. For one thing, in face of massive and complex data, the departments have no awareness of deep data mining. They just simply collect and store the data, but seldom use them, leading to the waste of mass data. For another, people are still learning about latest technologies in big data storage and processing, so they prefer to process data with the traditional methods, which will lead to a deviation from data value. In addition, how to use the accurate intelligent model and machine to mine the essential characteristics and inherent law of data in the massive and miscellaneous data is what the whole industry is studying. Meanwhile, the expansion of boundaries of the whole database, including the association mining of “junk data” is not well understood. Moreover, the business staff and data mining engineers

often have no good communication or cooperation, so they cannot effectively dock a department's business needs and data mining, which is an important cause for insufficient data mining depth and width.

2.3.2 Low degree of data sharing

Among the departments of various administrative areas in a city, basic information resources are seldomly shared. There is no institutionalized process management. Relevant laws and regulations are outdated and blocks of data not coordinated with virtually no horizontal data sharing. The reasons are mainly as follows. Firstly, the departments adopt different data indicators, so it is hard to integrate the data. Secondly, the data formats of different departments are incompatible, so a large amount of basic information cannot be shared and integrated. Thirdly, every department separately collect and filter the data it needs and may lay aside or eliminate other possibly valid data, which leads to low data utilization rate and increase in data use cost. Fourthly, data opening and security cannot be guaranteed at the same time. On the one hand, the public institutions master rich data resources, but there is not a clear opening standard for data among departments, so the data cannot be effectively shared and used. On the other hand, the data security problem becomes more and more prominent and there is a risk of leaking such important data as citizens' privacy.

2.4 Data management

2.4.1 Defects in data management

Big data bring a fundamental change of data and information processing mode and also promotes the development of all walks of life. However, the public security incidents frequently occurring in recent years reflect the imperfect emergency management system in China and also reflect the inadequacy in urban public security data management. As the urban information and digital infrastructures are improved and new intelligent devices emerge constantly, the urban data experience an explosive growth. In face of massive data, our data management is still in the start-up stage and has the following defects: decentralized data, delayed data recovery, hidden dangers in storage safety, high centralization degree, lack of technicians, imperfect institution setting and poor emergency management ability, etc. Therefore, all links of data management should be strictly controlled when big data are used to govern the urban public security.

2.4.2 Lack of privacy and access protection

The advent of the age of big data accelerates the transformation and upgrading of all areas of society, but the data leakage problems, such as spam mail, crank call, user information crime, become increasingly prominent, exerting a bad influence on the urban public security. For example, the data of 87 million users of Facebook

leaked; Cambridge Analytica LLC took advantage of the user information to send messages purposely and manipulated the election result. In addition, the hidden dangers of data security, such as unauthorized access to data and malicious modification of data are also serious.

2.4.3 Lack of security guarantee for urban public data

In February, 2017, a Russian black hat hacker “Rasputin” obtained the system’s access authority by taking advantage of SQL injection vulnerabilities, breaking down the systems of over 60 universities and USA’s government organizations break down and acquiring a large amount of sensitive information. It is thus clear that there are still enormous problems in the security guarantee for urban public data. For example, because the cryptosystem of hardware facilities is imperfect, the sensitive data in the database cannot be protected and the low firewall safety factor leads to invasion of illegal users. Besides, there is not a uniform data platform to comprehensively manage the data. The corresponding rules and regulations are incomplete, and there is no regular information security education and training for the workers. The above factors also undermines the security guarantee for urban public data.

3 Opportunities for urban public security management in the Age of Big Data

The data application problems in the urban public security also

bring opportunities for urban public security management in the age of big data. Here, intelligence information study and judgment covers analysis, study and judgment, processing, forecast and notification, and practical application of intelligence information; and it plays a very important role in the practices of intelligence information providing effective services for urban public security.

3.1 Intelligent filtering and classification of data intelligence

Intelligence filtering and classification is the first link of intelligence information utilization and the basis of intelligence analysis and processing. Take the urban video surveillance cameras for example: Every camera collects about 24~48 GB HD videos every day, and the data covers various data formats, such as multimedia, image and other unstructured data. Besides, valuable data are only contained in several frames of massive video data that are called “key frames”. If the massive video data are idle for a long time and no valid information is extracted, the data resources will become “zombified”. Video intelligent filtering and classification (object detection, target tracking, behavior analysis and incident analysis) and video structural description are used to process the original videos; the processing results are some data frames containing people and vehicles and structural description of these data frames in standard format, which are used to discover the abnormal behavior. For example, the UK’s Advisor (Annotated

Digital Video for Intelligent Surveillance and Optimised Retrieval) program can estimate the crowd density on the subway and around it, analyze people's behavior and forecast the potential dangers. "Intelligent Vision Kit" based on automatic incident detection and distributed technology which was developed by IBM is applied in the deployment of large systems and internet businesses.

Text-type intelligence information can be processed in four steps - text pre-processing, redundant word removal, feature identification and extraction, and text classification. For example, after neural network approach is introduced to the classification of security topic texts in social media, the text retrieval efficiency and quality can be enhanced effectively, and the social media data can, to a larger extent, reflect the users' attention to relevant events of public opinions. Based on this, it can help the relevant government departments to obtain the situation of public opinions, learn the evolution dynamics of "data security" topic, guide the trend of public opinions in time and provide decision-making supports for the government departments to supervise and cope with relevant internet events of public opinions.

The text-type unstructured data and image investigative information can be structured through clustering model or combination model.

3.2 Intelligent study and judgment on data

In the age of intelligence 3.0, the massive data resources supported

by the internet bring the serious problem of information overload when facilitating the intelligence agents. Intelligence study and judgment stresses analysis on massive data, and the intelligence analysis target cannot be accurately pinpointed by manpower only. In order to make up for the limited analytical ability of human brains, we train computers to simulate human's thinking mode and adapt them to different data conditions by strengthening the computation methods or analysis techniques. When a large intensive activity is monitored, we can quickly study and judge and give an early warning of emergencies by monitoring such real-time data as people flow, traffic flow, passenger flow and internet public opinion in combination with relevant data mining algorithms. In the specific implementation, we can first define the relevant data about areas of interest that should be extracted according to the nature, type, scale, place and security emergency requirements of the large gathering activity, and establish special data warehouses through data filtering and integration, such as passenger flow data warehouse, traffic flow data warehouse, and emergency treatment expert knowledge warehouse. Then, we can make use of data mining algorithms to establish corresponding mining models to make intelligent study, judgment and analysis on the data, such as attribute classification, association analysis, semantic reasoning, tracing and forecasting, early warning and decision tree. At last, we can visually present the analysis results of various types of data about the area of the large

gathering activity in real time, monitor the occurrence probability, type, scale, impact scope, harm extent, development and evolution trend of emergencies, provide intelligence support for the large gathering activity and supporting decisions and contingency plans for the deployment of the emergency command department.

With the development of handheld devices and sensor technology, the temporal and spatial trajectory data of a large quantity of users are recorded. On the premise of protecting the users' privacy, intelligent study and judgment based on temporal and spatial data can help us discover the behavior patterns of many individuals and groups and give early warnings to relevant departments. In the practical application, it can identify a large activity with abnormal individual movement mode. Namely, through the trajectory of a known suspect, it can find people with similar sub-trajectory and then find the gang. In the aspect of urban rail traffic network short-term forecasting, it can capture the traffic flow data and forecast the relationship between the traffic state in the future and that in the past based on the time sequence forecast model. And it can also use the multi-mode characteristic of traffic flow data to carry out forecasting, and the effective use of periodicity can effectively enhance the forecasting accuracy. Through the bus card exchange data, GPS trajectory data and video monitoring data, it can complete short-time forecasting of regional people flow in the city and traffic flow on highways so as to play the role of early warning,

relieve traffic jam and reduce stampede, and make the relevant urban administrators take precautions.

3.3 Intelligent forwarding of effective information

Through intelligent forwarding, the data analysis results and other information, messages can be sent effectively to receivers, so as to increase the data circulation efficiency and give full play to the data value. People may easily misunderstand “personalized forwarding”. In fact, the system contains both personalized and non-personalized messages. Especially when the security information is forwarded, non-personalized information is required.

Intelligent forwarding covers two meanings - active acquisition and passive acquisition. For example, when the police officers keep watch on a suspect, they may encounter a lot of doubtful information. If a vehicle enters the residential community of the suspect, the traditional way may be posing request to the command center by telephone or other ways. The command center then dispatches relevant personnel to collect necessary information, which is inefficient. In addition, the acquired information is mainly unprocessed raw data, so it is difficult to learn the potential association between the vehicle and the suspect based on these data. But in the context of big data, before the police officers send a request, the urban camera system can actively follow this vehicle; the

backstage can make a quick analysis on the relevant data of this vehicle and filter important information which is then sent to the forward scouts.

3.4 Intelligent change of study and judgment model through full data fusion

The premise of intelligent study and judgment is data fusion, which requires combining the data of multiple platforms, including the internal and external data of the enterprise. By studying and judging the vertical and horizontal data association, the coordination between business and technology, city departments and application systems can be realized. Through the online and offline interaction, the news media can analyze disaggregate behavior model in a crisis, build an intelligent participation foundation of urban public security emergency management field and break through the bottleneck of complete separation of emergency management from informatization. At the same time, in the age of big data, digital subsistence and network survival become the mainstream survival and life style of the young, and the mobile social media, such as microblog, WeChat or QQ group, have become the preferred channels to cope with emergencies and important platforms of political information release, and they play an important role in solving urban emergencies, setting public issues, gathering social resources and generating public opinions.

3.5 Deep mining of information value with the help of “Data Exhaust”

Data exhaust refers to all relevant data generated in digital activities. In the urban public security field, through the large-scale complex relationship network dynamic model, we can make full use of “data exhaust”, help the analyst to quickly identify the figures and events related to the suspect, construct the suspect’s social graph in combination with the temporal and spatial association analysis and multidimensional statistical filtering. With these information, we can analyze, study and make judgment, master the organizational framework of the criminal gang, detect the behavior characteristics, activities and abnormal behavior of the gang, and then enhance the forecasting, early warning and prevention ability of serious social events.

Through multi-level fund flow analysis, we can collect information about the bank account, fund transfer, ATM withdrawal and multi-media images and videos. Through the analysis on fund flow in multi-level exchanges and spatio-temporal analysis of a large quantity of account, we can identify the abnormal exchanges in the complex fund exchange network, doubtful accounts, and the suspect when we consider the location of ATM withdrawal and the image and video information.

In terms of social media analysis, we can pool multiple IDs of virtual social media accounts together and match them, capture and

integrate all the associated articles, posts, comments, messages and forward information of relevant social media accounts, support the analysts to adopt such social network analysis method as “spreading path, guiding opinions ” to quickly identify the information source, spreading path and core spreaders, which can thus help effectively identify and stop the spread of rumors, and crack down on crimes.

4 Urban public security-oriented Big Data system construction

4.1 Strengthen the Big Data base development and digitalize urban resources

4.1.1 Build Big Data bases

(1) Significance and value of building Big Data bases

The development of big data bases for urban public security follows the trends of the times, and it involves the storage and use of data, contingency plans, models, knowledge and decision-making techniques as the core. The public security data base is a comprehensive integration of storage device, network device, application data and programs, and it has such functions as storage, inquiry, management and maintenance. The base is helpful to quick collection of data, basic and business, including the sustainable acquisition, storage, maintenance, sharing and application of big data in urban public security, and comprehensive situation of political data resources in the urban public security field. In terms of a specific business, data support can be provided for urban public

security management. The big data base for urban public security should mainly be operated in the way of providing services so as to provide various data and application services for the city government and various competent public security departments, including the data acquisition, processing, release and analysis service, model calculation and application system hosting in the urban public security.

The urban public security management is featured by multiple targets, large quantities of information, complex and diversified data, wide scope of coverage and complex types and so on. It requires super-size, multi-source heterogeneous, multi-type and rapidly changing data support. Therefore, the government can build big data bases to acquire data through system log, network data and other methods, apply the cloud computing technology to establish cloud storage platform so as to store the heterogeneous data, grade and classify information, store, share and analyze urban public security data. As a result, it can intelligently detect various risks in the city operation in time, effectively evaluate urban public security, master the main risks in the city management and operation, and help decision-making to solve the urban public security problems, in order to build a smart and safe city.

(2) Structure of Big Data base for urban public security

The urban public security data are similar to scientific data, and have the full life cycle of “acquisition - storage - analysis - release -

filing and long-term preservation”. Specially, it can be divided into six stages - data acquisition, data processing, data storage, data mining and analysis, knowledge visualization, long-term preservation and utilization of data.

Firstly, data acquisition stage. Data acquisition is the first link of supervising the urban public security data. As the urban internet of things system becomes better and better, the sensors in all areas of the city send a large quantity of data all the time. The data of the urban security data supervision platform should be at least from two sources. One is the urban internet of things system, such as video monitoring system, image collectors, bus card system and GPS. Besides, the user data recorded in the personal handheld devices and smart mobile phones are also a component of the urban internet of things system. The other is the social media data. The social networks, to some extent, can reflect the trends of public opinions and the residents’ travel trend. Through the real-time data collection, the city dynamics can be realized in time, and then precautions can be taken.

Secondly, data processing stage. At the data processing stage, pre-processing of raw data, which covers data extraction, conversion and loading, is required. The data analysts in the supervision center extract the data from the scattered and heterogeneous sources to the temporary intermediate layer, then carry out data cleaning, conversion and integration, and finally load the data to the data

warehouse according to the pre-defined data model. Pre-processing can provide data support for further analysis and data mining.

Thirdly, data storage stage. A city can generate a large quantity of data with diversified data structures, so large capacity that can store both unstructured and structured data and enhance the analysis performance is required. In this system model, distributed cloud storage architecture is selected as the data storage scheme, where the urban public security data are stored on the cloud distributed server. In terms of storage mode, the traditional relational database cannot satisfy the needs of massive urban data storage in the aspects of expansibility, read-write rate and capacity support. In order to satisfy the needs of semi-structured and non-structured data storage, this data supervision platform adopts NoSQL storage system, which supports the storage and flexible management of massive data in the urban public security field.

Fourthly, data mining and analysis stage. Data mining and analysis is an important link in the data life cycle. In this stage, various data mining algorithms are used to arrange the massive unordered data, search for information hidden in them and covert them into understandable knowledge. Many existing statistical analysis methods, machine learning algorithms and neural network approaches can forecast the urban emergencies.

Fifthly, knowledge visualization stage. Many data have non-spatial feature, which makes it hard for the users to understand

the data. In the knowledge visualization stage, the large-scale abstract data mining results are presented in the form of images as a good assistance and supplement, so that the data supervisors or decision makers can intuitively see the information conveyed by the data and understand it easily. By integrating technologies in various fields such as image processing, computer-aided design and man-machine interaction, this stage, with the help of visualization tools, can realize the visualization of data products, and provide such services as intelligent retrieval and information modeling. The existing knowledge visualization tools include concept map, semantic network and thinking map, which can help the supervisors to carry out real-time inquiry, modeling and analysis and also can help the decision-makers to make a decision quickly.

Sixthly, stage of long-term preservation and utilization of data. The long-term preservation of urban public security data is also very important. The long-term preservation and sharing of data is helpful to data verification and reuse. Only after a reasonable storage method is used, can the cost be reduced. Therefore, Hadoop distributed file system (HDFS) may be used to store the data in the form of file; and the triangle model can be used to process the data. As time goes by, a large quantity of small files are gradually merged. The older the data are, the higher the merger efficiency is. Gradually, a balance between the real-time response and long-term preservation of massive data can be reached and the efficient data management can be met. In

addition, only in the constant use, can the data generate more value, so the data collected by the public security data supervision platform should be, to some extent, open to the visitors, so as to maximally reuse the data, and then support the cross-field and interdisciplinary merger and reuse of data.

With the data life cycle as the clue, the data monitoring work that should be completed in various stages with cloud platform as the solution is narrated in detail.

4.1.2 Strengthen efforts to attract and develop Big Data enterprises

With the rapid growth of urban public data, the data management and application needs of the public security industry have exceeded the traditional one for historical data backup. How to explore the association and value in the massive public data and enhance the ability and level of the public security institution to maintain the social harmony and stability have become a core problem of the public security governance. In the traditional public security response, the government departments are almost the only governance subject. But nowadays, enterprises gradually become an important participant of the public security governance, especially some high-quality enterprises in the internet industry and information industry. With the help of their big data processing technologies, they can help the government administrators to mine beneficial information in the massive data. Therefore, the big data enterprises

should not only be able to promote the big data integrated solution for the public security based on the latest big data technologies, but also be able to comprehensively process, analyze and mine the floating population data and data on the social acquisition platform, and realize discovery of key personnel, early warning of key events, population mobility trend analysis, people's relation discovery, specific event analysis and early warning of regional overrun. At the same time, they should also be able to provide interactive web query and analysis interfaces to provide information support for stability maintenance, criminal investigation, early warning, study and judgment, and command decision making. Take Google for example. Several weeks before H1N1 broke out, the engineers of this company made a judgment on the flu source and spread situation by summarizing and comparing a large quantity of internet information. This judgment is basically consistent with the official conclusion given later, and becomes a more effective and timelier indicator.

When paying full attention to the non-government subjective role, we should also have a clear understanding of the government departments' leading role. Many public security fields involve the national security and political and social stability, so there must be an authoritative public subject to dominate and coordinate the work, and this subject must be the government departments. The current problem is that the failure to coordinate the departments makes "the department with data have no ability to analyze the data and the

department with analysis ability have no data”. To sum up, the public security “smart governance” in the age of big data has the requirement for its subject that the government play its leading role in the public security governance and build a “government-society-market” co-governance mechanism while effectively coordinating its different departments.

We can focus on extending the big data industry chain by attracting the leading enterprises and fostering local enterprises.

Firstly, it is urgent to attract some domestically and internationally renowned enterprises with a leading role in cloud computing and big data. We also need a lot of small and medium-sized enterprises specialized in big data acquisition, storage, analysis, processing and application, so as to form an innovative R&D platform, break through the big data key technologies, and foster big data-based emerging modes in such fields as information consumption, cultural creative and advanced manufacturing.

Secondly, we should seize the opportunity of data resources gathering and application to drive the development of data industry, take investment and knowledge as the breakthrough and policy support as the opportunity, in order to optimize the industry development pattern. We also need to mobilize relevant government departments and social organizations, and attract large domestic internet information enterprises, professional data centers and enterprises that can generate data themselves to the city to build the

basic data center and backup data center. Specifically, the following two projects can be carried out. Firstly, big data strategic action planning project. We should introduce and foster big data enterprises, accelerate the development of big data industry, accelerate the gathering, sharing and opening of government and public data, comprehensively carry out innovative applications of big data, build a big data development support system and promote the concentrated development of big data industry, and provide financial support for the core big data industry projects. Secondly, big data application demonstration project. We should support the application of big data in the full life cycle, such as R&D, design, management and after-sales service of manufacturing industry, foster some pilot demonstration projects of new technologies, new products and new models of big data, and give rewards to the enterprises receiving approval of pilot demonstration. For the big data application service enterprises settling in the standard workshops and offices constructed with the government investment, the government of the area where the enterprises are should provide proper premises rent allowance. In addition, big data industry chain fostering action cannot be implemented. We should build a big data industrial park and support the existing data center to expand the industry application foundation, guide the small, medium and large enterprises to cooperatively build the big data industry ecosphere, actively promote the development of big data industry alliance, build and improve the big data industry

chain and build a perfect government-industry-education-research-utilization linkage mechanism.

Lastly, on the basis of promoting the cluster construction of data centers and then preliminarily forming the data resource depression, we should support information providers with good integration capability to build several big data service platforms to provide integrated sharing services, such as big data analysis public support and application in key fields. At the same time, in the process of integrating the software and hardware enterprises and service enterprises, to extend the development of relevant industries, we should provide software and hardware integrated big data solutions, and finally form an all-round product and service supply system.

4.1.3 Establish disaster prevention and early warning data center

At present, natural disasters and extreme weather-climate events occur frequently with complex causes, which seriously threatens the urban public security, so accurately forecasting the unexpected disaster events in time becomes the core goal of establishing an urban public security early warning mechanism and conducting disaster prevention and mitigation.

The disaster prevention and early warning data center is a whole set of complex system that conveys, accelerates, presents, calculates and stores disaster prevention and early warning data information. It not only contains the computer system and other matching systems

(such as communication and storage systems), but also contains redundant data communication connections, environment control devices, monitoring equipment and various security systems. It can accurately forecast the unexpected disaster events in time and play a role in safeguarding the urban public security.

After years' exploration and practice, China has accumulated rich experience in data center construction and management and gradually formed a complete and scientific system, providing a good theoretical and practical foundation for the establishment of disaster prevention and early warning data center. The system architecture of the disaster prevention and early warning data center mainly contains the following two systems: firstly, disaster prevention and early warning data center storage system. This system gives full play to the performance of distributed storage system, uses the flash memory disk and storage virtualization technology to integrate the storage hardware in a virtual pool, layers the "cold, warm and hot" data according to their access, read and write frequency characteristics. It also gives full play to the high-speed read-write ability of flash memory disk, and establishes independent storage system's metadata layer to enhance the data query speed and efficiency. Thus, it effectively guarantees the service efficiency and safety of the core storage system. Secondly, disaster prevention and early warning data center backup system. This system backs up the released data according to the main needs of the internet areas, and backs up the

virtual servers in the private cloud in combination with relevant software. In terms of released data, it carries out complete backup every month and differentiated backup every week; in terms of virtual servers, it carries out complete backup every month and differentiated backup at any time according to the system adjustment.

In addition, the establishment of disaster prevention and early warning data center also covers the construction of security system, and it needs to cope with the problems in physical security, network security, data security, system security and application security. Only after the safe operation of infrastructure is guaranteed, can the safety and reliability of the data center be effectively guaranteed.

We should formulate the standards for the disaster prevention and early warning data center at the beginning of its planning and establishment. When the overall architecture planning is made, we should formulate the detailed infrastructure construction standards and system support standards, so as to effectively realize an infrastructure environment platform with uniform service interface standards and provide a good basic environment support for the service application platform. At the same time, we should formulate the optimal standard for the data format and storage structure so as to facilitate the later data mining and standardized interface service. In the utilization and maintenance process, we should carry out standardization according to the standards so as to effectively enhance the full-open support of the infrastructure for the application

system standardization and more effectively reduce the islanding effect of information system. At the same time, we should promote data sharing and enhance the data mining capability and service capability.

Meanwhile, the urban disaster prevention and early warning data center should have both the functional coverage of national disaster prevention and early warning data center and pertinence of urban disaster prevention and early warning data center so as to have a quick transmission and early warning response and take defensive measures quickly to carry out emergency treatment and mitigate the disaster. We should overcome the data crossing of such urban data centers as earthquake data center, weather data center and traffic data center, realize the location aggregation and data integration and sharing, enhance the utilization rate of data resources and reduce the cost.

4.2 Open and share data

4.2.1 Promote the fusion and opening of public data resources

The convergence and fusion of information technology and economic society caused the rapid growth in data, and data have become a national basic strategic resource. In the urban public security field, big data have a broad application space. To deepen the big data application, promoting the fusion and opening of public data resources is the only way.

The *Platform for Action of Promoting the Big Data Development* released by the State Council in 2015 put forward that we should strengthen the top design and overall coordination, greatly promote the connectivity, opening and sharing of the government information system and public data, accelerate the integration of government information platforms, eliminate the information island, promote the data resources to open to the public, strengthen the administrative accountability, guide the social development and serve the public enterprises. We should support scientific macro-control, promote precise government governance, advances convenient commercial services, promote efficient security guarantee and accelerate universal service for the people's livelihood. Beijing has released *Action Plan of Beijing for Big Data and Cloud Computing Development (2016-2020)* (hereinafter referred to as *Action Plan*). The *Action Plan* put forward that Beijing will build internationally first-class big data and cloud computing infrastructure, make over 90% of the units opening the public data with a data opening rate of over 60%. Fushun City in Guangdong Province, with "Internet +" technology as the support, breaks through the department and level boundary, political barriers and information island, integrates several doors into one and several windows into one, and realizes the "one-door one-website" new political service mode where one can handle various affairs in one door and enjoy the whole-course process on one website. Based on the principle of cost minimization

and benefit maximization, it spares no effort to cut through the data conversion channels between different systems and makes the departments handle affairs cooperatively rather than people repeatedly travel between different departments.

The fusion and opening of public data resources can be realized mainly at the technical level and institutional level:

At the technical level, there are the following four aspects. Firstly, to realize the diversification of data acquisition and consolidate the data foundation. By improving such urban infrastructure construction as PFID and internet of things, we can strengthen the perception and acquisition of various social activities and city operation data and provide sufficient data foundation for the public security monitoring, early warning and disposal. Secondly, to realize the fusion of data storage and strengthen the data association. We should associate information acquired through various channels, make the original information really become fine and available data from extensive and unordered data and finally realize the efficient application of data. Thirdly, to realize the real-time data analysis and guarantee the efficient processing. The public security guarantee system should be able to quickly analyze and process the massive, heterogeneous and dynamic data, detect the possible hidden dangers and find the best solution in proper order after an emergency happens. Fourthly, to realize intelligent decision making about public security practices and promote the scientific decision making. We should strengthen

the development of public security data analysis tools and establish a data analysis model so as to analyze and process the massive security data and make a scientific judgment and take countermeasures in a short time.

At the institutional level, the following three parts are covered. Firstly, to improve the urban security data management network. We should strengthen the network management over the areas with a high density of population or facilities and high hidden safety dangers, improve the data acquisition system of various network areas, make the urban security monitoring network extend to every risk point and avoid the dead zone. Secondly, to promote the data integration. It is suggested that the governments should make full use of the cloud computing technology to establish a big data management cloud platform for urban public security so as to manage various data related to the public security in a centralized manner, effectively drive the exchange and sharing of public security information between the departments, governments and among the public according to the functions and needs. Thirdly, to strengthen the big data security guarantee. We should not only promote the development and utilization of public big data resources, but also protect the citizens' privacy and business secrets. Therefore, after acquiring the public security data, the governments should ensure that the data are neither leaked or misused, and only use these data for public security management and other legal public management

functions, but not for other purposes.

4.2.2 Promote the cross-departmental data resources sharing

Urban security is complex system engineering and it requires the cooperation of all walks of life and various departments. In recent years, with the frequent occurrence of highly-associated and highly derivative safety accidents in Chinese cities, the co-building and sharing of data resources between each other is especially required; and the successful experience of foreign countries also should be learned. Take New York for example. Since 1990, New York City Management Department has proactively deployed “data-driven” management model. After Bloomberg assumed office in 2002, he began to use the data technology to promote such reform as government behavior transparency. In 2013, Mayor’s Office of Data Analytics (MODA) based on Data Bridge was formally established. Different from the traditional offices, this is a data analysis team gathering and integrating the data resources of all the public sectors of the whole New York and using the big data analysis technologies to forecast and prevent the urban hidden dangers. In the aspect of urban fire control, MODA and Fire Department of New York (FDNY) cooperates whole heartedly: the former provides data support and analysis capability while the latter provides guidance with rich fire control practices. They together developed the big data fire analysis and forecast algorithm model - Fire Risk Monitoring

System (RBIS).

In China, over 80% data resources are controlled by the government and most of them are not open to the public; and the connectivity and sharing between departments is also few. But as the urban public security problems become more and more complex, since massive data scatter in different industry departments, a single department has no sufficient data resources to solve the increasingly complex urban security problems. The “13th Five-year” Development Plan pointed out that we should “reform the urban management and law enforcement system, promote the fine, full-cycle and cooperative urban management”. Promoting the cross-departmental data resources sharing is helpful to strengthen the construction of big data system for urban public security.

When promoting the cross-departmental data resources sharing, we should do the following things: firstly, we should set up the idea of overall urban public security governance, get rid of the disadvantages from “fragmentation”, realize the cross-industry overall governance, and strengthen the intra-department data resources connectivity. Secondly, we should integrate the management departments and organizations of urban security industries, and strengthen the overall data resources processing capability. Thirdly, we should improve the data information sharing catalogue and exchange system and eliminate the poor connectivity when the data resources are shared. Fourthly, we should improve the

inter-departmental information resources sharing incentive and restrictive mechanism. The cross-departmental data resources sharing cannot be realized automatically and it must be guided through appropriate incentive and restrictive mechanism.

Based on the national government data uniform opening platform, we should unify the internal data standards, establish the security emergency database at all levels, set up intra-department data resources platform, establish the department sharing mechanism of data resources, and on the basis the data association mining, we should set up multi-level information system and information management platform based on the technologies such as cloud computing, big data and internet of things, enhance the industry departments' emergency quick response capability, real-time linkage capability and resource allocation ability. Then, when ensuring the data resources are safe and permitted by laws, we should make full use of the opportunity of department data opening and sharing to guide the enterprises and institutions, industry associations, scientific research institutions to consciously acquire the valid data and open them to the public so as to provide safe and sufficient information resource support for the operation of urban security big data system platform.

4.2.3 Improve the public data service and resources exchange system

Urban public security has a bearing on people's immediate

interests, and ordinary people also have the right to know public security information and the right to participate. In 2009, the American Obama administration provided various government data for the public by launching Data.gov. As of October, 2014, this website had collected over 130,000 data sets, and the content covers over 20 categories, including agriculture and education. Besides, this website also provides such new functions as data grading evaluation, advanced query and social network site, and also motivates many government departments, including the UK, to promote similar actions.

The construction of big data system for urban public security should be oriented by providing services for the public, and we should set up a new public security information service idea targeted to serve and benefit people with information. We should guide the industry departments, enterprises and public institutions to carry out big data application demonstration in such fields as urban-rural service, disaster mitigation and relief, traffic tour and social security, accelerate the deep fusion of traditional public security service and internet or traditional public security service and mobile internet, and build a public-oriented integrated online public service system. We should integrate the public security functions of various industry departments, construct a comprehensive urban public security service platform and provide one-stop information consumption service for the public. Besides, we should strengthen the evaluation and

inspection on the government public security industry supervisory departments' portal websites and assure the public security information services for the public, including information inquiry, online reception and online consultation. We should accelerate the deep fusion of internet and urban public security service system, promote the public security data resources to be opened to the public, realize the public-oriented online query and customized forwarding and provide personalized and intelligent services for the public.

In order to further eliminate information island, we still need to improve the resources exchange system between different industries. Firstly, we should improve the regulation and policy system, establish the data resources exchange platform with uniform rules and consistent standard and exchange regulations and policy system, so as to provide a complete system support for the data resources exchange platform. Secondly, we should establish a data resources exchange information system to provide reliable technology support for the efficient services of data resources exchange platform and actively give play to the advantages of information sharing. Thirdly, through the platform exchange information database, we should deeply analyze the past similar data exchange and evaluation situations, try to move forward the overall data processing and pre-process the original exchange data information. Fourthly, we should make constant innovations on the platform exchange form, and try to establish the public data resources exchange “handheld

platform” integrating WeChat and mobile phone APP as a whole. Fifthly, we should build an information-based monitoring system, make use of the functions such as dynamic records and track tracing of the public resources exchange platform informatization system to prevent the illegal behaviors, so as to provide a powerful supervisory support for regulating the operation of the data resources exchange platform. Sixthly, we should give play to the fundamental rule of the market and leading role of the government, optimize the market environment, realize the opening, transparency and resource sharing in the whole exchange process, guarantee the exchange fairness, enhance the service quality and efficiency and promote the construction of big data system for urban public security.

4.3 Application of block chain technology

4.3.1 Personal digital credit remodeling

For many application programs, identity management is a key link. Establishing a personal credit system with block chain technology means that all the personal records can be stored in the same system. And that every person has a comprehensive digital identity, in which the authoritative information cannot be changed arbitrarily. The block chain consensus mechanism greatly enhances the information reliability and ensures that information can be changed only after all the parties agree.

Therefore, with block chain distributed ledger as the security

guarantee, when a user is decided to be anonymous or have an identity through design selection, and the block chain technology is used to store all the system records of change or exchange on the cloud system after an individual gets permission, he can inquire the source and authenticity of data in the transmission process. Here, digital identity covers personal identity authentication and digital signature function, and it will provide a reliable personal identification mechanism as a basic component of one-stop service.

4.3.2 Personal privacy protection

In the information age, the public are deeply troubled by being “transparent”, and data exchange and personal information leakage incidents happen frequently. Therefore, an ideal personal information management scheme should have several guarantee mechanisms. If the block chain technology is adopted, in order to maintain data synchronization between scattered nodes and reach a consensus on the exchange and information content, some information must be appropriately opened. But on the other hand, as the block chain technology is applied in the financial field, such as banks, such information as transaction record is important and sensitive data, so extra measures should be taken to restrict the use of the unauthorized users. In addition, the transaction records usually can reflect some sensitive information and may leak the user’s privacy. For example, a user’s shopping transaction records can reflect his consumption

level and life status. In order to protect the user's privacy, some sensitive data must be processed to reduce the risk of privacy leakage.

Therefore, we can use block chain distributed P2P network to convey and process the user's privacy information and adopt relay forward mode to communicate between nodes. In this way, it will be difficult for the attackers to steal and intercept the information communicated in the network, and they cannot learn the true source and direction of the information, either. All the additions, deletions, modifications, inquiries, retrievals and accesses will be recorded on the block chain. Once a suspected criminal act happens, it will be exposed to the public and cannot be changed or denied. The application procedures based on the block chain technology usually have decentralized architectures, so it needn't store such sensitive information as account and password on the central server. Therefore, we can avoid the traditional risk of privacy data leakage caused by server being hacked and then effectively cope with network attack.

4.3.3 Public security information sourcing

In order to build an open public sector and make the public play a role in supervising such relevant information as economic and social development, market supervision and public services, block chain can create a new mechanism to confirm whether the information or public opinion is trustworthy. This mechanism covers information

source evaluation, content tamper-resistance and multi-node content verification, which supplement each other and constitute a complete constraint mechanism. The whole network information will become opener and more transparent; and “information disintermediation” will be realized so that we can acquire information without cost or with a very low cost.

Therefore, from the perspective of information summary and statistics, we can make public opinion analysis on the activity rules of an individual or organization according to the records of such details as data values, occurrence time, transaction process and source certificate. We can also quickly re-check the information and make countermeasures, shorten the period of useful effect on the target groups and realize the clear, quick and accurate judgment on and response to the information. By uploading and disclosing the complete test data of a product through the block chain, the information including sourcing the place of origin of the product, transportation and moving trajectory will become more convincing, which will ensure the security of all links on the processing chain. In addition, we can also effectively control the spread of public opinion and prevent various panics caused by such incidents as “plastic laver” incident and “Japan Tsunami”. Therefore, by integrating the block chain’s information sourcing characteristics and the government’s public chain, we can enhance the government credibility and prestige. When a rumor event happens, the government can integrate

information in time, trace the information publisher and propagating source through the block chain, analyze the information authenticity and control the public opinion to prevent the public panic and social instability caused by its further deterioration, and strengthen the governance effectiveness of public sectors.

4.3.4 Public data storage and access management

Nowadays, the massive data generated by the urban public sectors in the world are stored in the centralized database which can be easily attacked, posing huge risks. Block chain, as a tamper-resistant distributed ledger, can make a detailed record of all the processes and stages of a thing, including its generation, development, exchange and storage, and it cannot be easily changed. Through the consensus protocol, we can check whether the new activity data record is valid and judge whether the condition being added to the block chain is satisfied. And we can also track the activities of these entities through sharing records and avoid the influence from hacker attack and unauthorized change. Once a sharing and consensus mechanism is established through point-to-point network, several nodes will jointly guarantee the data integrity, which will thus guarantees the three key factors - data authority, data accuracy and data access control.

These three key factors are specified as follows. Firstly, data authority. Such authority not only shows the data origin, but also

stipulates the data accuracy, operation authority, life cycle management and storage location of final authoritative version of data. Secondly, data accuracy. Accuracy is a key characteristic of data. It means the value data record of any object is accurate; it can represent the correct value and official viewpoint; the form and content are consistent with the described object. Thirdly, data access control. Block chain can preserve the records of access to public information and individual information, including the detailed information of the data, data exchange process and information of the person with the data updating authority. For example, in the aspect of land registration, recording with block chain can guarantee the complete land circulation information, including the land location, size, ownership and transaction records. In the aspect of vehicle traffic management, once the violation information is recorded in the block chain system, the license plate number, violation time and place and penalty record will be always preserved on the block chain and cannot be intervened or manually deleted, which thus guarantees the deterrent force of traffic management rules on all people. In the aspect of financial securities exchange and government internal affairs, by forming a “digital ID card”, all the certificates, physical assets or intangible assets, private records and documentary evidence will be transferred to the block chain, and the whole-network multi-verification forms a data information “self-certification” mode, so the government and public sectors needn’t rely on the data

information managed and provided by the third-party organization and then enhance their working efficiency and quality. In the future, block chain technology will also be applied in the governmental personnel's welfare fund distribution supervision, tax supervision, charity supervision, passport issuance and food supply chain safety, etc.

Research Team

Wang Ye, Professor, Tianjin Normal University

Wen Zhiqiang, Professor, Tianjin Normal University

Wang Shiwen, Associate Professor, Tianjin Normal University

Wang Shuyi, Lecturer, Tianjin Normal University

WangLin, Professor, Tianjin Normal University

Li Huiyu, Graduate student, Tianjin Normal University

Yan Jiahong, Graduate student, Tianjin Normal University

Zhang Junya, Graduate student, Tianjin Normal University



Annual Meeting of the New Champions 2018

2018新领军者年会

Big Data Application and Urban Public Security

大数据运用与城市公共安全