The Design of Police Big Data Mining Model

Qingguo Geng[1, a, *]

[1] Department of criminal investigation, Jilin Police College, Changchun, Jilin, China

[a] 51982668@qq.com

[*]Corresponding author

Keywords: Internet of Things technology; big data analysis; police information; model design

Abstract: Big data, as a technological product of the information explosion era, has had a significant impact in all walks of life. The emergence and application of the Internet of Things technology has revolutionized the multi-angle, comprehensive, and in-depth collection of information, provided new channels for obtaining police intelligence, and resolved some blind spots in the construction of police intelligence sources. Effectively extend and expand the time and space for obtaining police intelligence. How to use the Internet of Things technology and big data analysis algorithms to improve the efficiency of case handling is one of the important topics of police analysis and research in various countries. This article conceives a set of police big data analysis models based on the Internet of Things technology, including data perception layer, data transmission layer. It is expected to be able to mine police intelligence from the massive, dynamic, and heterogeneous data of the Internet of Things in real time and efficiently, improve the correctness of police decisions, and better meet the needs of dynamic management and combating crime.

1. Introduction

Facing the increasingly dynamic, intelligent, violent and organized crime situation, if the police department wants to master the initiative and improve the ability to attack, prevent and control illegal crime, it must improve the ability of police information acquisition, research and decision-making in a wider range and deeper level, and give full play to the advantage of intelligence leading police service. It can't meet the needs of the society to solve the case by the police. Therefore, the state puts forward the development strategy of "strengthening the police with science and technology", and helps the police to crack down on illegal crimes by means of high-tech equipment and information technology. The "Golden Shield Project" has pushed the construction of police information to a climax. However, with the advent of the information age, the means of committing crimes become more and more changeable and hidden, and the traditional means of information technology has been difficult to help the police to quickly analyze and judge. The emergence and application of Internet of things technology has played a revolutionary role in the multi angle, all-round and deep-seated acquisition of information, provided a new platform for the acquisition of police information, solved some blind areas in the construction of police information sources, and effectively extended and expanded the time and space for the acquisition of police information. In order to provide high-quality information for the police department, it is necessary to analyze the police information based on the massive dynamic data of the Internet of things from the perspective of high-quality information acquisition and application according to the needs of police information construction. This paper attempts to build a set of police big data mining model based on Internet of things technology, in order to improve the accuracy and efficiency of police clues collection, analysis and judgment process.

1.1. Overview of Internet of things technology

Internet of things is the product of the integration of sensor network, communication network, Internet, social network and other networks. It mainly uses RFID technology, sensor technology,

  DOI: 10.25236/icme.2019.040

barcode and QR code, infrared sensor, laser scanning technology, and in accordance with the corresponding protocol rules, to realize the connection of "things to things", "people to things", "people to people" at "any time", "any place", "any network", so as to realize the intelligent network system of information exchange and communication, so as to realize the intelligent An information network for identification, location, tracking, monitoring, and management. At present, the Internet of things and its related technologies are gradually extending to all aspects of society, which also brings opportunities and challenges to the police work, and provides a new platform for all-round access to police intelligence. Generally, the "things" of the Internet of things should meet the following conditions: there is a corresponding information transmission / receiver; there is a data transmission path; there is a certain storage function; there is a microprocessor; there is an operating system; there is a special application program; follow the communication protocol of the Internet of things; there is a unique number that can be identified in the network world. [1] It can be seen that, as the "Internet of things connected", the core and foundation of the Internet of things is still the Internet, which is an extended and expanded network based on the Internet; at the same time, its user end extends and expands to the information exchange and communication between any goods.

1.2. Characteristics of data in the Internet of things

There are six main characteristics of data in the Internet of things: first, the amount of data is huge. The Internet of things has a large number of sensor devices, they continuously transmit the collected information to the data center, which needs to process a huge amount of data. Second, the data type is complex. There are many kinds of entities in the Internet of things, and different types of information collected by different terminal sensors lead to data collected by Internet of things including numbers, data, texts, videos, and simultaneous interpreting. Third, the data is heterogeneous. The format and semantics of the data collected by the Internet of things sensor terminals are different, and the data heterogeneity is obvious, which increases the difficulty of data transmission, storage and processing. Fourth, it is highly dynamic. The Internet of things contains a large number of sensor nodes, and their frequent changes lead to high dynamic data of the Internet of things. Fifth, the complex characteristics of time and space. This is a significant feature of the Internet of things data. Because the sensor terminals are widely distributed, the collected data reflect the state of an object at that time and other information, so these data are meaningful in a specific time and space. Sixth, the data is incomplete. When the sensor terminal of the Internet of things works, it may be affected by natural factors or human factors, and the sensor terminal is difficult to cover every corner, resulting in incomplete or even unsafe data collected. [2]

1.3. Data mining in the Internet of things

Data mining refers to the high-level process of extracting people's unknown but potentially useful information knowledge from data and expressing it into a pattern that can be understood ultimately. It integrates database, artificial intelligence, machine learning, statistics, visualization technology and other disciplines, uses pattern matching and a variety of algorithms to find the relationship between data, and carries out concept description, association analysis, classification and prediction, cluster analysis, and can use intuitive graphics to express information patterns, data association trends.[3]

Police information data mining is to extract potentially useful sensitive information and rules from massive, fuzzy and incomplete case information. At present, the problem facing the police department is how to analyze and mine the rapidly growing mass of information, such as personnel information, goods information, hotel industry information, flight information, vehicle information, etc., to quickly find out the information needed. Applying data mining technology to the analysis of police security information, mining out the association rules of case attributes and criminal attributes, and then finding out the laws and characteristics of crime, which is of great theoretical and practical significance to prevent and combat crime.

1.4. Overview of police Internet of things

In view of the increasingly information, intelligent and organized crime situation, the public

security department can improve the ability of information acquisition, transmission and processing with the help of Internet of things technology. Police Internet of things technology is to apply a variety of intelligent communication and network technology to the traditional police work, forming a platform, which can record, store and process the information of various public security elements, and connect all the object information through the Internet to realize resource sharing and police management intelligence.[4]

The Internet of things is the further embodiment of police science and technology. In the future, the police Internet of things can realize the collection, intelligent identification, positioning, tracking, monitoring and management of information needed in public security prevention and control, investigation and case handling, and can innovate the police coordination mode to maximize the overall power of police resources.

With the advancement of police information work, the combination of Internet of things key technologies (information perception, information transmission, information processing, etc.) and police business work can improve the level of police information and intelligent management, and enhance the ability of police organs to serve the society and protect people's safety.

1.5. Big data and big data analysis and mining

Big data refers to the data collection that can't be captured, managed and processed by conventional software tools in an affordable time range. It is a massive, high growth rate and diversified information asset that needs new processing mode to have stronger decision-making power, insight and discovery power and process optimization ability. Big data has 5V characteristics (proposed by IBM):[5] volume, velocity, variety, value, veracity. With the advent of the era of big data, big data analysis has emerged. Big data analysis refers to the analysis of large-scale data. It refers to the process of using appropriate statistical analysis methods to analyze a large number of collected data, extract useful information and form conclusions, and then make a detailed study and summary of the data.

The key to police analysis and research is to discover the seven elements of time, space, personnel, organization, case (event), position, and article. If the seven elements of information can be obtained with the help of the Internet of Things technology, then the police will greatly improve the clues，the efficiency of information, while avoiding human error. In the face of complicated clue information, big data analysis technology can be used to achieve in-depth and efficient mining analysis, and then quickly find the correlation between the seven elements.

2. Manuscript Preparation

2.1. Technology Architecture

The police big data analysis model based on the Internet of Things technology is a system model that uses big data analysis algorithms to deeply analyze and analyze the massive police Internet of Things clue information. It includes the IOT perception layer, data transmission layer, data analysis layer, and data display layer. At the same time, it also needs to effectively integrate with the existing public security information network to achieve information sharing collisions.

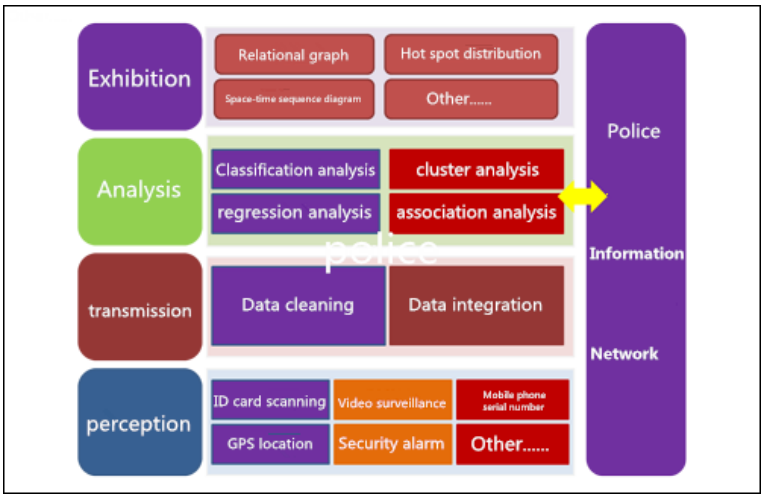

Figure 1 Technical architecture diagram.

2.2. IOT perception layer

Police Internet of Things refers to the use of sensing technology and intelligent devices to automatically perceive the objects of interest in police work. Through the network, technical processing and intelligent analysis, the real-time grasp of the status and situation information of the objects of interest is achieved, and the dynamics of the objects of interest are achieved, the purpose of monitoring, precise management and scientific command. Currently applied by the police are passenger ID verification, police car police GPS positioning, urban video surveillance, road checkpoint vehicle monitoring, key area alarm prevention, key personnel mobile phone positioning, gun RFID tag positioning, etc, which can collect personnel and vehicles As well as the urban spatio-temporal data of items, it provides a lot of true and reliable raw data for the police to establish basic elements for investigation.[6]

2.3. Data Transfer Layer

The IOT data transmitted to the system model via media such as mobile Internet and wired networks has a large amount of data and is distributed in different time domains and airspaces. It cannot be used directly. To this end, they need to be processed according to certain standards Preliminary screening and classification integration, which will be converted into compatible system models, with "metadata" with unified norms and standards.

2.4. Data Analysis Layer

After preliminary classification and integration, the data is still massive and lacks intuitive connections, which cannot provide the police with a basis for research and judgment. To this end, professional data mining algorithms are needed to further analyze and integrate these data. At the same time, it must collide with clues and intelligence in the public security information network to reduce problems such as low information reliability caused by a single data source. Commonly used analysis algorithms are classification analysis, regression analysis, cluster analysis, and association analysis.

2.4.1. Classification Analysis

The process of dividing data with different characteristics into different categories according to certain classification criteria. Taking gang crime in an area as an example, the area has a high incidence of gang crime in the recent period. The police conducted a centralized investigation of personnel in entertainment venues and rental houses in the area and nearby areas, and collected a large number of ID information and mobile phone serial number information. By colliding with the ID number and mobile phone serial number during the midnight period, the police can screen out a list of suspicious people who frequently come and go in the high crime area, thereby helping the police narrow down the scope of investigation.

2.4.2. Regression Analysis

By making a certain correlation analysis between the independent and dependent variables, a regression equation is established to predict the dependence of the variables. California police have used fire warning systems to predict building fires and analyze arson cases.

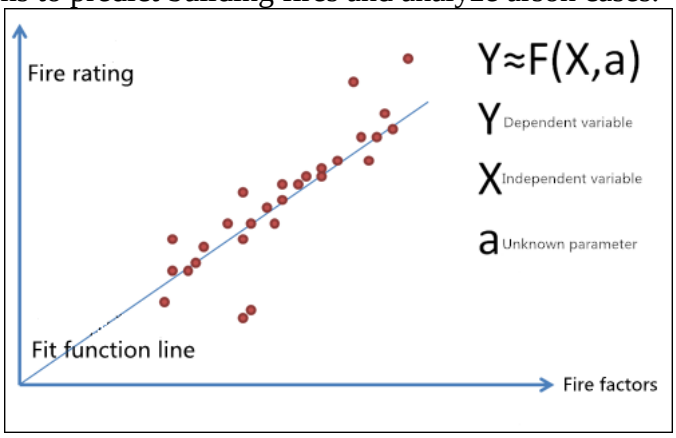

Figure 2 Regression analysis graph.

The California Police formed a set of fitting functions for the fire level and fire factors by digitizing the fire case within a year, the weather of the day, and the factors of the building itself. The more detailed and smooth the combined function curve is, the higher the accuracy will be, and then the empirical data will be formed to effectively improve the fire warning ability. At the same time, the police did not let go of the abnormal points, because the abnormal points often represented the suspicion of "artificial arson". The police then analyzed the abnormal points to find out the circumstances behind the fire.

2.4.3. Cluster Analysis

Unlike classification analysis, cluster analysis does not have prior knowledge. Generally, a group of seemingly irregular data is divided according to certain characteristics, and data with different attributes are divided into different groups. The police can group the collected ID number, alarm information, mobile phone serial number, etc. according to the time or space as the benchmark attribute, and then find suspicious clues.

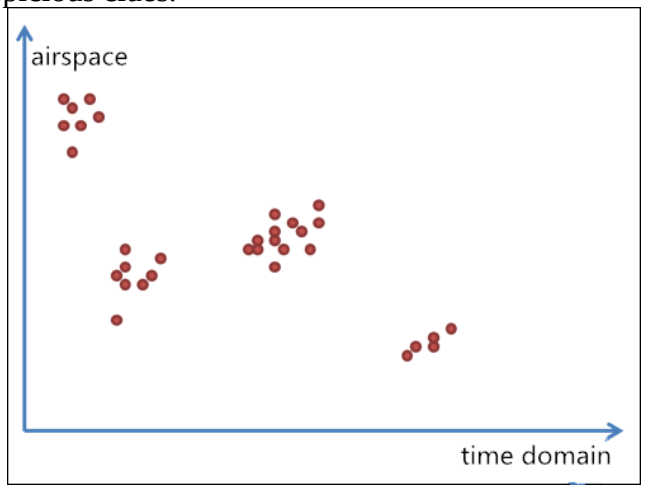

Figure 3 Cluster analysis diagram.

For example, a cultural relics theft case occurred at midnight at a cultural relics protection unit. By sampling the unit's surveillance video and mobile phone serial number in the past week, the police first locked suspicious targets on the four mobile phone serial numbers that appeared at midnight. Because this period cannot be a normal tourist visit period, it is suspicious, but because the perpetrator brought a hood and gloves, he could not obtain a valuable surveillance video of the crime. The police carried out investigations and collisions for the rest of the time and found that

three days before the incident, the suspicious target mobile phone serial number appeared in the cultural relics exhibition period. By examining and positioning the surveillance video and mobile phone serial number during this period, the investigation scope could be effectively narrowed.

2.4.4. Correlation Analysis

It is used to find the correlation between valuable data in a large amount of cluttered data. By analyzing the basic information of the suspect, relatives and friends, transportation, bank accounts, and travel records, etc., a relationship network of the suspect can be drawn, and then the police can provide a powerful clue to quickly grasp the suspect's movements.

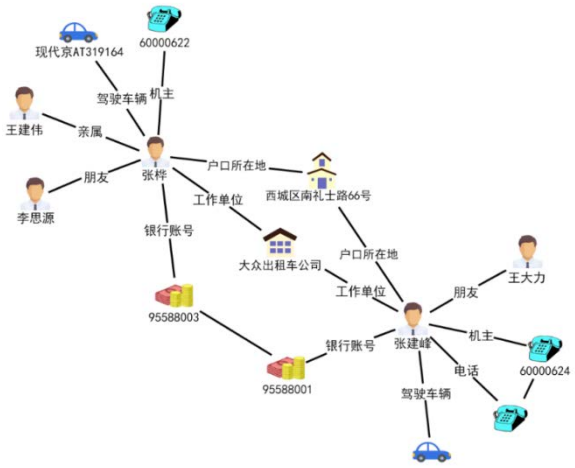

Figure 4 Association diagram.

2.5. Data Display Layer

The technology that comes along with the application of big data and visual display technology can intuitively display the obscure data generated by big data analysis in the form of charts, helping police to quickly judge based on the characteristics of the charts. According to the characteristics of police case research and judgment, this model selects several common chart models for display.

2.5.1. Association Diagram

Association diagram, also known as association diagram, is a diagram used to analyze complex relationships such as "cause and effect", "purpose and means", etc. It can help people to find logical relationships Solutions to the problem. The association diagram is used in conjunction with the association analysis algorithm to show the relationship between the elements of the research analysis result. The figure divides each element into a central element and an associated element, and the associated element expands around the central element. It is usually used for the relationship between the suspect and the subject. [Note: The element types of the center element and the associated elements can be the same or different]

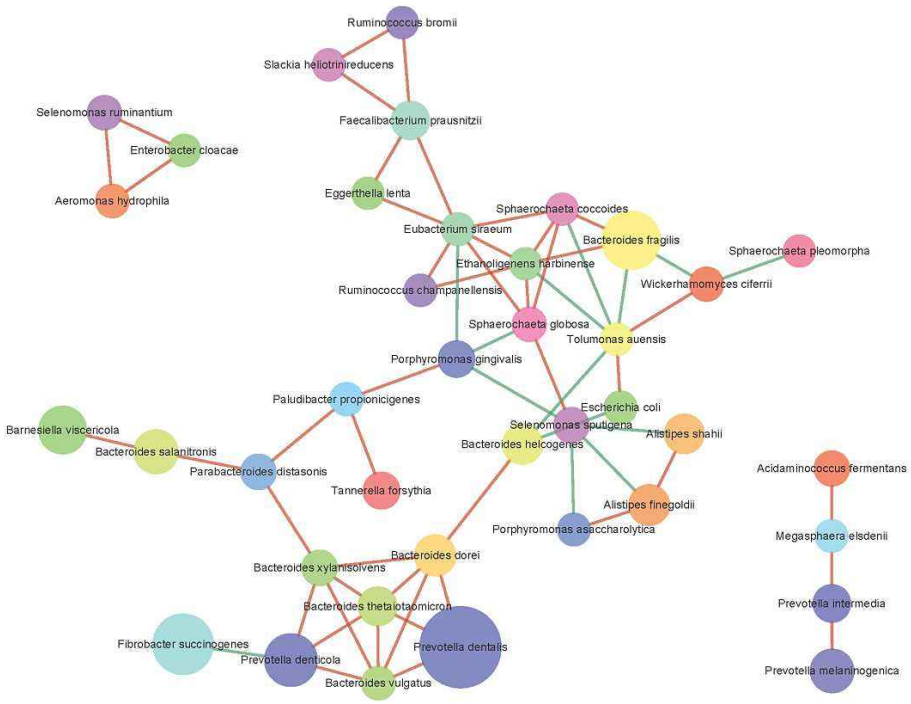

Figure 5 Association diagram.

2.5.2. Spatiotemporal Timing Diagram

The spatiotemporal sequence diagram is an interaction diagram that describes the time and space sequence of passing messages between objects and between objects and participant instances. It is usually used in conjunction with classification analysis algorithms and cluster analysis algorithms to display the behavior trajectory of the subject within a certain time and space according to the spatiotemporal attributes. It is usually used to display the behavior trajectory of a suspect or a suspected vehicle.[7]

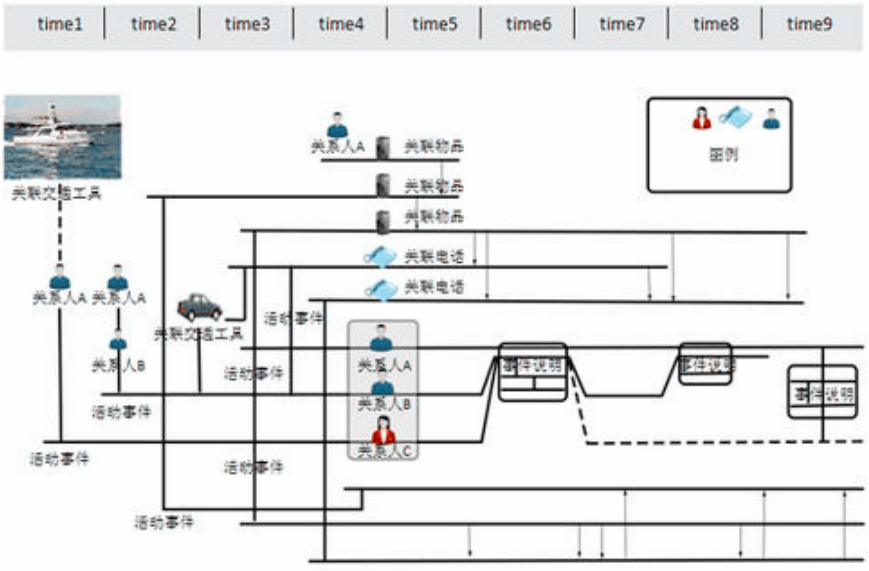

Figure 6 Spatiotemporal timing diagram.

2.5.3. Hot Map

The hotspot issue map is an analysis method that uses different signs to mark and display areas on the map or page according to different levels of attention. The method of labeling generally uses the form of color depth, density of points, and proportion No matter which method is used, the end result is the same, that is, the eyes are suddenly bright. Usually used in conjunction with cluster

analysis algorithms, it is used for collision digger behavior habits, associations, and "commonality" between abnormal behaviors.

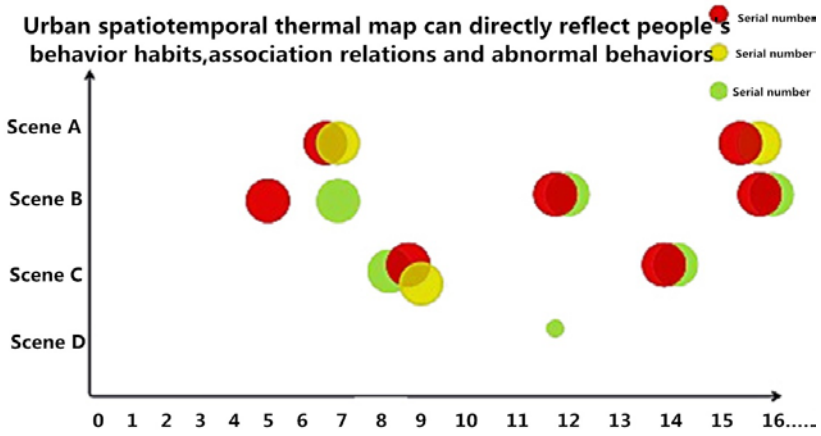

Figure 7 Hot map.

2.6. Application Architecture

The police big data analysis model based on the Internet of things technology itself cannot solve the case directly, but it can help the police narrow the investigation scope and shorten the investigation time. Because it needs to integrate multi-party information resources, it is a set of cross-department, cross-police analysis and application models.

2.6.1. Horizontal Dimension

In the process of handling a case, the police often encounter the cooperation of multiple police types, and each police type may have different channels for obtaining clue information, resulting in different data for the same target element. The information is verified, and the information of one or more people is used as the final conclusion data. (Note: The data obtained by relying on the Internet of Things technology will reduce the probability of data differentiation, but the case clue is not only the Internet of Things sampling data, there are many data to be obtained through other channels)

	Criminal investigation	public security	Counter -terrorism	Narcotics Control	Household register	Data conclusion
Element 1	A	A			A	A
------		B	C		B	B
Element N	D		D	D		D

Figure 8 Horizontal dimension table.

2.6.2. Vertical Dimension

When applying this model, it should be used according to different levels and division of functions. Grass-roots units are mainly responsible for the collection of clues. The actual case handling unit is responsible for the process of clue collection and analysis and judgment. The provincial department directly uses the results of analysis and judgment to analyze and assess the public security situation and police work in various regions.

	Basic police station	Branch Office	city bureau	Provincial Office
Application surface	collection	collection	Analysis	Statistical evaluation
		Analysis	collection	

Figure 9 Vertical dimension table.

2.7. Difficulties In Construction

Although the Internet of Things and big data are no longer unfamiliar technologies, they have

also been widely used in various industries, ranging from "item tags" to "Baidu big data analysis plan". It seems that these two types of technology are omnipotent. In fact, in order to take full advantage of technology, certain conditions are needed.One is the choice of high performance and scalable distributed parallel data mining algorithm. Only by choosing the right algorithm of data mining in the Internet of things and adopting the right parallel strategy can we guarantee the accuracy and efficiency of mining. Second, the establishment of Internet of things data exchange standards. Internet of things data mining needs to deal with massive data with various forms and complex types, and these dynamic and heterogeneous data grow exponentially, so it is difficult to formulate the standards of Internet of things data exchange.[8] Three is uncertainty. This is the biggest challenge of mining technology in the Internet of things, which is manifested in the uncertainty of data collection and preprocessing, the uncertainty of mining task description, the uncertainty of mining methods and results, and the uncertainty of mining results evaluation caused by the different ultimate mining objectives of different users. In addition, timeliness is an important feature of police intelligence. The more timely the information is obtained and used, the higher the value of the information is. Police intelligence personnel should have a strong sense of timeliness, and make full use of the Internet of things as a new platform to obtain real-time information, so as to make the most of police intelligence.

3. Conclusion

Although the author has summarized this system model through theoretical basis and work practice, the model has to be applied to practice and a complete and practical software system still has a long way to go. This requires police officers and related industry technical practitioners work together.

Technical means cannot be directly used to solve cases, but if it can be applied scientifically and reasonably, and its role can be fully exerted, the efficiency of police handling cases can be greatly improved. These are all based on a complete management mechanism and personnel understanding, and require the joint efforts of the police, scientific workers, and society.

In the context of Internet of Things technology and big data analysis, strengthening the construction of police security informatization needs to do the following three things:[9] First, improve the understanding of the construction of public security informatization. In the context of big data, the prevention and control system needs to efficiently process and mine massive amounts of information such as text, voice, pictures, and videos. Therefore, it is necessary to change the consciousness of passive analysis of data in the past. The second is to improve the talent mechanism for public security informatization construction. The construction of public security informatization under the background of big data requires a high-quality and high-level talent team, which is a favorable starting point for in-depth promotion of China's informatization construction. The third is to increase the integration of resources for police security informatization. Due to the urgent needs of police work for information resources, it is necessary to break the barriers of barriers, build information sharing platforms of different police types, different departments, and different units, and accelerate the improvement of information resource service platforms and big data cloud center application service systems. Gather various information resources that can be used publicly by the police security organs' internal and external networks.

The rapid development of the construction of police informatization has not only achieved the interconnection and communication between information, allowed information to be shared and co-constructed, and has become a new type of information-oriented, efficient, scientific and other police mechanism. At present, in the face of the advent and in-depth era of the Internet of Things technology and big data, relevant departments of police organs at all levels need to do a good job of collecting and organizing data information, scientifically designing and structuring information systems, and establishing data-based consideration and analysis. Carry out new thinking and new models of police informatization construction, carry out comprehensive unified planning for information systems in the era of big data, and gradually improve the police information system guided by big data, cloud computing, and mobile Internet etc., long-term development lays a solid

foundation.

Acknowledgements

This research was financially supported by the Jilin Police College.

References

[1]. Zhang Liang. C.N. (2017) research on Internet of things information mining technology based on public security information [J]. Journal of Jiangsu Police Academy, No. 3, 103-106.

[2]. Li Fuxiang. C.N. (2015) concept of police big data analysis model based on Internet of things technology [J]. China security, No. 6, 84-88.

[3]. Wang Zhanhong. C.N. (2017) police Internet of things technology research and application [J]. Journal of Zhongzhou University, No. 1, 110-113.

[4]. Li Jialong. C.N. (2014) some problems in the construction and application of police Internet of things [J]. Police technology, No. 1, 47-49.

[5]. Wei Guowei. C.N. (2018) public security information construction problems and Countermeasures under the background of big data [J]. Theoretical observation, No. 9, 87-89.

[6]. Zhao Yuan. C.N. (2018) design and implementation of data management system of big data management platform for public security work [D]. Shanghai: Donghua University, 15-20.

[7]. Yu Long. C.N. (2019) characteristics, advantages and practical application of public security big data [J]. Journal of Liaoning Police College, No. 5, 48-55.

[8]. Niu Jin. C.N. (2019) analysis of data analysis and innovation and development of public security [C]. Excellent paper collection of 2019 national public security communication academic seminar, 3-6.

[9]. Li Xiang. C.N. (2018) analysis of investigation thinking in the context of public security big data [J]. Digital communication world, 263.