Application of Big Data in Security Precaution by Network Technology and Video Image Database

Yougang Yao*, Linling Zou
Chengdu Neusoft University, Chengdu, 611844 China

*Corresponding author e-mail: yaoyougang@nsu.edu.cn

Abstract. The security industry is generated by the needs of modern public security precaution. With the popularization of hardware and network technology, the amount of data in the field of security increases rapidly. Under such a condition, big data technology in the field of security arises. This paper discusses the current situation of the security industry, and discusses the related security problems in the application process of big data in the field of public security precaution. Taking the video image database application platform as an example, combined with the national standards 28181, 25724, 35114, this paper discusses the data security problems involved in the process of data acquisition and data transmission. Based on the Kerberos authentication mechanism of Hadoop, this paper introduces the big data security and access control technology, and forms relevant solutions in the whole process of data acquisition, data transmission and data access.

Keywords: Field of Public Security Precaution, Big Data Technology, Application Security.

1. Introduction
With the rapid development of Internet and cloud computing technology, the amount of global data is growing exponentially. In this environment, big data technology arises. According to the definition given by McKinsey Global Research Institute, a large-scale data set, which is far beyond the capability of traditional database software tools in terms of acquisition, storage, management and analysis, is characterized by massive data scale, rapid data flow, diverse data types and low value density. The field of security is undoubtedly a combination of these complex data. In the field of security, data types can be divided into unstructured data, semi-structured data and structured data. The main types of unstructured data are video data and picture data obtained by various monitoring video cameras. The main types of semi-structured data are face data, fingerprint data, iris data and so on. The main types of structured data are all kinds of identity information, log files, alarm data[1]. The architecture of security big data is shown in Fig 1. The big data platform in the figure is used to manage various types of data.
Fig. 1 Architecture of Security Big Data

The security big data system is mainly built on the video monitoring system, including acquisition system, transmission system and processing system. The acquisition system not only collects the video data collected by the existing monitoring video cameras, but also collects the electronic police data of the bayonet. In the transmission system, various kinds of wired and wireless networks become the main transmission ways. The processing system has the video monitoring platform of safe city built by the public security department; Food and drug supervision department established the “bright kitchen” system; Besides the video monitoring platform built by urban management/Transportation Committee and other departments, all data will be connected to the security big data platform after various complicated platforms, and all kinds of information after analysis and mining will be provided according to the specific needs of various departments and industries.

In this system, the huge data is video data. In the field of security, more applications are carried out around the extraction of image information, and image information often contains more information, involving personal privacy, which is more likely to directly affect the public security. With the improvement of machine learning and deep learning technology, all kinds of feature recognition methods are constantly improved, and the data processing becomes more rapid and accurate. Many domestic enterprises such as Hikvision and Dahua corporation have their own processing platforms. The public security video image data database application platform independently developed by the research institute has also achieved preliminary results in the processing of face, iris, gait and other data. Relying on massive data, it can carry out query and retrieval, time-space analysis, control and alarm. All of this is based on the authenticity of the data. But how to ensure that the data in the acquisition, transmission, storage, processing process is reliable? This paper puts forward the risks faced by big data in the field of security, and gives the corresponding solutions. Taking the application platform of public security video image data database as an example, in Chapter 3.2, it introduces in detail the solutions to the related security problems involved in the platform design, and according to the current big data in the security field of the application level of relevant security technology, this paper proposes some ideas for further research.

2. Security of Big Data
With the development of "high definition, networking and intelligent" in the security industry, the security industry has entered the era of big data, and the security problems of big data in the security industry are also prominent in many aspects. This section mainly introduces the related problems in the acquisition, transmission and storage of big data.
2.1. Data Acquisition
The main source of security data is video data. According to the latest data from IHS Markit, which is an industry research company, 176 million monitoring video cameras were installed in the public and private fields (including airports, railway stations and streets) in China. For such a large front-end data set, the data collected is likely unreliable. Those with ulterior motives may deliberately forge the data by modifying the data acquisition software, falsifying the data or ID or by cloning attacks, or they may modify some key attribute information in the data (such as data size, creation time, etc.)[2], so that analysts can draw wrong conclusions after analyzing these data, so as to achieve the purpose of those with ulterior motives. For example, in June 2017, CCTV reported that a large number of home cameras were invaded, and also in 2017, an enterprise's monitoring equipment was controlled by overseas IP [3]. These incidents will not only damage the privacy of citizens, but also cause immeasurable losses to public security.

2.2. Data Transmission
Compared with the analog signal transmission mode in the past, digital signal transmission mode is more used in the security field, such as data transmission through wired network or wireless network. The process of data transmission is bound to be applied to various IP protocols in the network, and these protocols usually do not have strong security functions. Therefore, if we use network protocol to transmit data without necessary data encryption, those with ulterior motives are likely to intercept and leak relevant information in the process of data transmission, this will make the monitoring equipment ineffective.

2.3. Data Access
Big data is a kind of super large-scale and highly concurrent unstructured data. The traditional relational database is no longer applicable, so now the common way is to store data in non-relational databases, such as HBase, Cassandra, Mongodlb, Redis, etc [4]. However, compared with the more mature relational database, the development of non-relational database is still in the initial stage, and its supporting security mechanism needs to be further improved. Firstly, the verification and authorization mechanism is relatively weak, which makes the database vulnerable to brute force cracking and internal attacks. Attackers may steal or falsify data, which may cause sensitive data to be leaked. Secondly, the storage of big data itself is a kind of distributed storage, which makes the use of Multi-Tenancy technology inevitable. Multi tenancy technology refers to the sharing of physical devices and virtualization resources among multiple independent tenants. Using this technology, cloud service providers can use the characteristics of Multi-Tenancy technology to build an effective scalable resource allocation mechanism to meet the needs of multiple tenants, but it also means that those with ulterior motives can more easily access the target data through resource sharing [5].

3. Big Data Security Technology
The data security of big data is a comprehensive subject, which involves many subjects such as cryptography, system control and so on. In this section, we will focus on the big data security issues in the field of security, and introduce the related security technology in detail.

3.1. Data Acquisition and Transmission Encryption Technology
Cryptography is the core of information security and other related issues, such as authentication, access control, authority and integrity. Cryptography also promotes computer science, especially the technologies used in computer and network security, such as access control and information confidentiality. The widespread use of cryptography makes it possible to use bank chip cards, access passwords for computer users, mobile payment and e-commerce in daily life. Until modern times, cryptography almost exclusively refers to encryption algorithm: It is the process of transforming ordinary information (plaintext) into incomprehensible data (ciphertext); Decryption algorithm is the opposite process: It converts ciphertext back to plaintext; Modern cryptography is the foundation of
big data security. Without cryptography, data encryption and data access control are impossible. At present, the main encryption algorithms in the field of security are symmetric and asymmetric encryption.

3.1.1. Symmetric Encryption. Symmetric encryption algorithm is an early encryption algorithm, the technology is mature. In this kind of encryption algorithm, the sender will send the plaintext (raw data) and the encryption key (key) together through a special encryption algorithm to make it into a complex encrypted ciphertext, and then send it out as a complex encrypted ciphertext. After receiving the ciphertext, the receiver decrypts the ciphertext by using the used key and the inverse algorithm of the same algorithm. In this process, the receiver must know the encrypted key in advance. The commonly used symmetric encryption algorithms are DES, RC4, AES and so on. China has developed its own symmetric algorithms SM1 and SM4. The security of this algorithm is a little poor, but the speed of encryption and decryption is fast.

3.1.2. Asymmetric Encryption. Unlike symmetric encryption algorithm, asymmetric encryption algorithm needs two keys, which were considered as public key and private key. The idea of this algorithm is to use two different but related keys to perform the encryption and decryption process that the public key and private key were appeared in pairs, we can use the public key to encrypt the data, and only the holder of the private key can decrypt the encrypted data. The commonly used asymmetric encryption algorithms are RSA, ECC and so on. China has also developed the corresponding asymmetric encryption algorithm SM2. The advantage of this algorithm is high security, but its encryption and decryption speed is slow.

3.1.3. Application of Encryption Algorithm. In the field of security, video data contains the greatest value. Nowadays, when the monitoring video cameras generally reach 1080p resolution, the capacity of video information is usually huge, and many applications need real-time video, which puts higher requirements for the security of security data. In reference [6], a variety of lightweight video data encryption methods based on symmetric and asymmetric encryption algorithms are proposed to balance the contradiction between security and efficiency. In reference [7], for the common video compression coding method H.264, the asymmetric encryption of AES and Slasha20 algorithm is used to encrypt the video data. In addition to the method of video image direct encryption, a special transmission security protocol can also be added. In reference [8], the security is raised to the system level. Relying on the GB/T 28181 standard [9], not only the video data encryption technology, but also the security measures include the scheme of equipment authentication and signaling integrity protection. In the domestic research field, a SVAC2.0 standard [10] is proposed, which is a coding algorithm specially used in the security field. Compared with H.264 or h.265 coding algorithm, a large number of security options are added. This set of standards has the support of symmetric and asymmetric cryptography technology. In the standard, the key and digital certificate are defined to carry relevant information, which has high security. In reference [11], a video data security management authentication mechanism based on SVAC2.0 technology is proposed. SM1 and SM4 algorithms are used to encrypt the data to improve the encryption efficiency, and SM2 algorithm is used to transfer the key for key management. From the video data acquisition, to the encryption of the transmission process, and then to the final storage, the system realizes the data encryption and key management through the domestic own symmetric and asymmetric encryption technology. In recent years, the release of GB35114 standard [12] has strengthened the related fields of video security. GB35114 is the epitome of public safety video field, and its whole topology is shown in Figure 2. By adding signaling server and media server, the original big data system can access encrypted signaling and video stream. The supporting key management server, password server and CA server generate and manage a complete set of password information.
In the gb35114 docking standard, the camera can adopt the equipment conforming to the standard, and the old camera equipment can be upgraded by adding hardware module or software password module to increase the security of the camera equipment. The terminal customer can browse the encrypted video stream through identity authentication. By adding UKEY, the public security video image database system (PSVIDS) developed by the research institute realizes the functions of device identity authentication, video signature verification and encrypted video decryption. It can effectively verify whether the video content has been falsified, and solve the data security problem in the process of video data acquisition and transmission, which is the most critical of security big data.

3.2. Data Security Access Control
After solving the security problems of data acquisition and transmission, we will start to use the data. Taking the public security video image database system as an example, fusioninsights80, Huawei’s big data platform, integrates a complete set of security authentication framework, it helps users unify authentication and record all submitted tasks, which can not only prevent illegal users from logging in, but also effectively monitor user’s operation, which improves the security of the system. Access control for big data components is an important control about data security. When hundreds of millions of data are aggregated and stored in big data components, it is very likely that there will be leakage. To solve this problem, Huawei’s FusionInsight platform adopts Kerberos authentication technology. Kerberos is a third-party protocol used for security authentication. It adopts the traditional way of sharing key, and realizes the communication between client and server in the network environment which does not necessarily guarantee security. It is suitable for the client/server model [13][14]. FusionInsight platform can also give different operation permissions to different components according to different users. To a certain extent, it limits the components that each user can operate, and especially the related authorization of components storing data, which directly affects the security and maintainability of data.
Fig. 3 shows the Kerberos authentication process, which is divided into four steps. After receiving the request, the AS (authentication server) randomly generates a password ASkey, and generates two tickets (ticket1 and ticket2) to return to the client. After the client gets two tickets, firstly it uses its own password to unlock one of the tickets, ticket1, to get ASkey, and then generates an authenticator, which mainly includes the current time and the check code of ticket2, and encrypts it with ASkey. The client sends the authenticator and ticket2 to the server at the same time. The server first uses its own password to unlock ticket2 and get ASkey, then uses ASkey to unlock the authenticator, and checks whether the timestamp in the authenticator is within 5 minutes of the current time, and whether the timestamp appears for the first time. If the timestamp does not appear for the first time, it indicates that someone intercepted the content sent by the client. If everything is correct, the client passes the authentication, and then the client can initiate access to the server.

Through Kerberos authentication, different users can be classified, data access permission can be restricted, and unauthorized users can not access the data. Kerberos ensures the security of the data.

4. Conclusion
Based on the application of security industry, this paper analyzes the security problems faced by big data on the level of data acquisition, transmission and access control, and expounds the related problems of big data encountered in the field of security. Then, taking the existing public security video image database system as an example, this paper discusses the research results of solving the security problems of big data itself - data encryption technology and data access control technology. At the same time, we can see that encryption technology and coding technology with independent intellectual property rights also have a certain application mode in big data security applications. In the future, in the field of security, with the gradual establishment of a unified video big data networking platform, in addition to the security of the data, further research should be carried out on the use of relevant data related to personal privacy.

Acknowledgments
Our thanks to all of the people who have contributed to this paper. My paper is supported by the project of the Ministry of Education's Fund for Educational Research Projects (2018A03007); Sichuan Provincial Department of Education's Fund for Educational Research and Reform Projects (SCJG20A004-4).
References

[1] Yao Jianbo. Security big data security and privacy status and challenges [C]. Seminar on security classified protection of security system in big data environment: 2018

[2] Chen Wenjie, Cai Lizhi. Big data security and its evaluation [J]. Computer applications and software. 2016, 33 (4): 34-38

[3] Fu Kai, Ni Ping. Research on security risks and protection of network camera system [J]. Modern telecommunication technology. 2017, 47 (6): 54-58

[4] Zhang Yin, Chen Min, Liao Xiaofei. Current situation and Prospect of big data application [J]. Computer research and development. 2013, 50:216-233

[5] Shan Suthaharan. Big Data Classification: Problems and Challenges in Network Intrusion Prediction with Machine Learning[J]. Performance Evaluation Review. 2014, 41(4): 70-73.

[6] Zhang Xiaoliang, Wang Mengmeng, Zhu Guiliang. New progress of image encryption algorithm [J]. Computer engineering and science. 2012, 5:1-6

[7] Hou Senke, Ji Xiaoyong. Asymmetric video encryption scheme based on h.264/avc [J]. Electronic measurement technology. 2014, 9:50-52

[8] Luo Yunfeng, Ding Han, Xu Qingguang. Research on architecture design of trusted video surveillance system [J]. Software guide. 2017, 10:135-138

[9] GB/T 28181-2016, technical requirements for information transport, switch and control in video surveillance network system for public security[S]

[10] GB / T 25724-2017, technical specifications for surveillance video and audio coding[S]

[11] Pei Jing, Liu Na, Wu Yangqi. Application of svac2.0 in video security system [J]. China security. 2017144:80-84

[12] GB /T 35114-2017, Technical requirements for information security of video surveillance network system for public security[S]

[13] Bhat S, Damle S, Chaudhari P, et al. KERBEROS: An Authentication Protocol[J]. International Journal, 2014, 2(2): 200-204.

[14] Jason S. Hadoop Security Overview-From Security Infrastructure Deployment to High-level Services[EB/OL]. http://hbtc2012.Hadooper.cn/subject/keynotep&shihongliang.pdf, 2012-11-30.