Privacy Detection and Protection for Intelligent Transportation Shared Travel Service

Hui Ma, Yong Zhang*
School of Information Engineer, XuChang University, Xuchang, Henan, China
**Corresponding Author.*

Abstract

There is a large amount of privacy data in the big data environment. This paper studies the privacy detection and protection for intelligent transportation shared travel service. On the basis of traditional security policy, through the operation of anonymization of private data, the purpose of keeping private data secret under the premise of protecting data characteristics is realized. The anonymized private data can be used by data engineers to upgrade system functions and improve system user experience. In this paper, through the use of Hadoop project HBase tools, complete the data cleaning of privacy data, data desensitization operation. This paper designs and implements a privacy protection scheme for intelligent shared transportation system in big data environment. This paper uses K anonymous protection technology, data encryption technology to achieve the protection of privacy data. In this paper, HBase technology is used to desensitize the privacy data of the server. The experimental results show that the privacy protection scheme proposed in this paper can meet the security requirements of privacy data in transmission, storage and application.

Keywords*:Big data, intelligent transportation, shared travel, privacy protection.*

I. Introduction

With the development of intelligent transportation system, advanced computer processing technology, information technology, data communication transmission technology, automatic control technology, artificial intelligence and electronic technology are effectively integrated into the transportation management system. With the development of intelligent transportation, a large number of data are stored in the system. In the big data environment, the challenges include the credibility of big data, how to realize the access control of big data, privacy protection in big data and so on [1-2]. Among them, big data privacy protection is directly related to the security of user privacy information, so privacy protection of big data has been applied to a variety of fields [3-5]. In the big data environment, privacy protection can be achieved by anonymous protection. K-anonymity algorithm protects user privacy by anonymizing data [6]. The privacy information after anonymous processing can keep the data characteristics without disclosing the user's privacy.

Using the traditional encryption algorithm and PKI technology, the privacy protection scheme of computer environment, network communication, intelligent transportation system and other dimensions can effectively protect the privacy data [7]. Privacy protection is carried out through the reproduction of privacy data life cycle under the big data platform of intelligent transportation system [8]. However, the above privacy protection scheme will lead to privacy data can not use big data technology to explore new value. In the transportation industry, private data could have been found new value under big data technology [9-10].

The protection of privacy data in intelligent transportation system, on the one hand, ensures the data security of privacy data through the traditional encryption algorithm and authentication, audit mechanism. On the other hand, anonymization algorithm is used to anonymize private data on the premise of preserving data characteristics. The anonymized private data can continue to be used for data mining, so as to achieve the purpose of exploring its commercial value on the premise of protecting the security of private data.

ISSN: 0010-8189

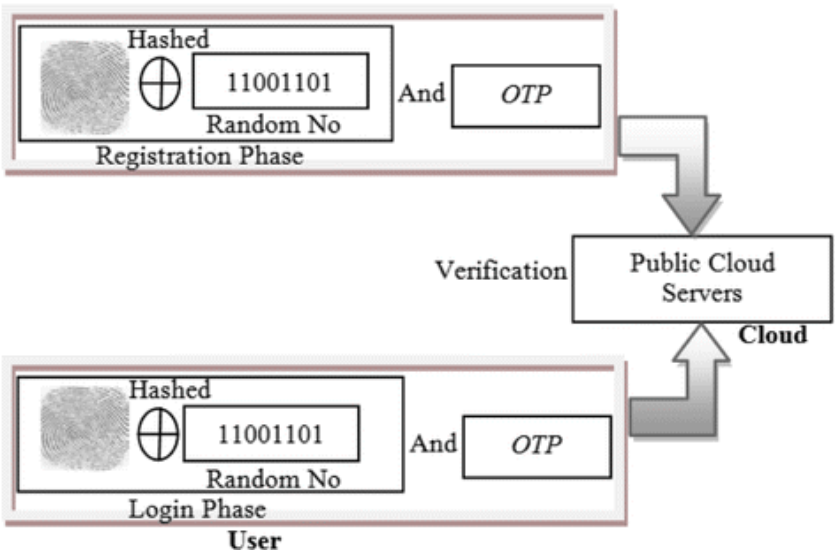

Fig1: Simulation result showing energy efficiency of handover algorithm

As shown in Figure 1, the privacy protection scheme proposed in this paper encrypts the data at the user end through the encryption algorithm to prevent the leakage of privacy data in the transmission process. After the encrypted private data M is transmitted to the server, the data is stored in the database or in the form of data file to the server. This part of the data can only be used when the decryption key is known, which can effectively prevent the privacy leakage event caused by the server-side data leakage. The other data is restored to m after decryption, and the data is processed by the server-side processing program. The processing flow includes data cleaning, data desensitization and data distribution. In the process of data cleaning, the missing fields in the original data are filled, and all the private data are dumped to HBase after unified format.

II. Causes of information data security damage in network communication

A. Network hackers and viruses are rampant

For a long time, hackers and viruses have been rampant in the network environment. Out of the pursuit of material and the challenge to security personnel, a large number of hackers are emerging, and all kinds of viruses are being developed. Nowadays, viruses are likely to enter people's computers through e-mail and other ways, threatening the security of users' network environment. Although the relevant staff has established a firewall system to intercept the virus, its specific risk is still limited [11]. It must be continuously optimized to meet the actual needs of users. The system framework is shown in Figure 2.

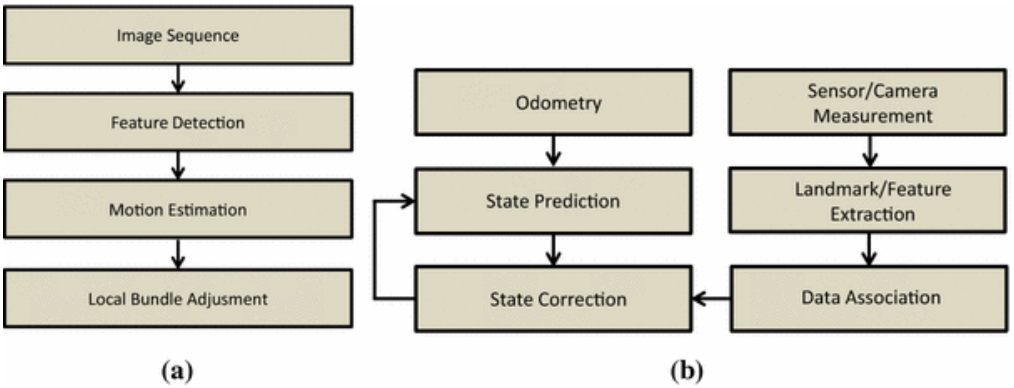

Fig 2: Framework structure diagram of the system

B. The structure of network communication data

At present, in China's communication industry, the primary goal of network security communication is to complete the network data transmission process. Generally, the structure of network security data is tree. The data of this kind of structure will inevitably encounter some risks in the process of transmission, which provides opportunities for the invasion of lawless elements. In addition, in the process of information transmission, the unreasonable application of information encryption technology is also the ultimate cause of the threat to network security. If the establishment of communication system lacks rationality and necessary encryption means, the final communication security will be seriously hindered and the security of the whole communication system will be threatened.

C. Hidden danger of network communication software

Nowadays, most of the basic software used in communication technology is developing towards the trend of openness. This development trend does improve people's use efficiency, but also leads to frequent crises. If the source code and specific information in the process of basic software design are leaked or attacked, a large amount of information will be stolen by criminals. The user's information on the software will be at a glance, and the user's information security will not be guaranteed. In addition, due to the lack of basic security awareness of a large number of users, they set up too simple account and password for themselves, which makes it extremely easy to decipher the password, and the communication security can not be effectively guaranteed.

III. Data security in privacy detection and protection for intelligent transportation shared travel service

A. Firewall or scan

With the gradual upgrading and improvement of communication equipment, it has become a very important part of people's life. Therefore, in order to eliminate the emergence of communication security problems to the greatest extent and reduce the adverse effects as far as possible, the staff should consider from the basic level to ensure that the network security is effectively guaranteed. In this case, a large number of applications of firewall technology can provide effective protection for user data information. Under the application of firewall technology, the attack of illegal personnel can be effectively stopped, and the illegal elements can be prevented from stealing some important data information as much as possible. With the continuous development of network technology, if only rely on managers to manage network security vulnerabilities, there will be a lot of work omissions, and then threaten the security of the entire network environment. Therefore, the staff can take some new ideas to improve the security of the network environment. In today's era, both office life and study life are heavily dependent on network technology. Only on the premise of ensuring the security of the network environment, can people really enjoy the benefits.

B. IP address protection

Staff should do a good job in network exchange to ensure the efficient transmission of network information. At the same time, in the process of information transmission, we should ensure that the location of the switch is in the second layer of the whole IP structure, so as to complete the protection of the IP address. In addition, the router should also be the key object of network security protection, to prevent criminals from obtaining user access address, so as to avoid the threat of criminals to user IP address. In terms of security work, we should start from the physical level, and refer to the specific requirements of information transmission to formulate fixed security standards. For network communication work, both natural factors and human factors will have an important impact on the whole network communication. In the process of information collection, the staff should choose the way of network control chip as far as possible. In the process of information configuration, they should try to think about the way of information collection and the efficiency of information transmission. In the link layer, the communication is usually completed by the physical layer and network layer, and at the same time, the flow control and monitoring data are used to ensure that there is no problem in the transmission of information. The main function of the network layer is to define the network operating system, process all kinds of messages in

groups at the same time, take the optimal path for information transmission, and complete the network communication task. Figure 3 shows the functional structure diagram. The transport layer completes the process of error data processing to ensure the integrity and accuracy of the final data. In this information layer, TCP protocol needs to be applied, so the impact on the whole network layer is relatively large. At the application level, the operation content involves user information and network exchange process, and the whole process should be protected from virus attack.

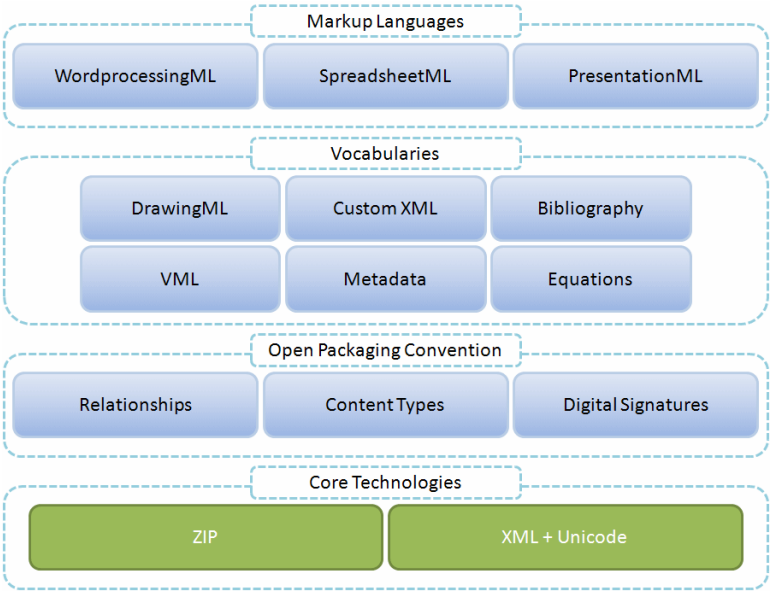

Fig 3: Functional structure diagram

C. Network encryption processing

Encryption processing is widely used in communication network, which usually involves password setting, information encryption, etc. in specific applications, corresponding methods need to be adopted. Link encryption is an effective encryption technology. This technology completes the corresponding encryption processing at both ends through the encryption processing at the physical level. In the process of information transmission, the first thing the staff should do is to encrypt the information. When receiving information, the staff should first decrypt the information, and then take corresponding measures to complete the whole information transmission process. For example, the more advanced way of information transmission in the world is Cisco. In the process of developing special technology, it can make the secret group more simple and convenient for the transmission and sharing of some information, but in this process, the information will be changed. In the actual network communication, the corresponding staff should make corresponding solutions according to the blockchain technology, so as to ensure the security of the whole information. General data block is the main reference for information arrangement.

D. Network authentication

At present, most of the authentication methods adopted by network users are based on user name or password setting. The independence between user name and user password can ensure that the situation of user losing password can be effectively solved. At the same time, there is a one-time user login status. In this login state, the identity information is only valid for a period of time. In some cases, it can also be verified by other auxiliary authentication methods, such as password card and other identity tools. This kind of tool is generally issued by the network management office, and can be combined with the most original identification method. In addition, other high-tech methods such as fingerprint verification and face verification can also ensure the reliability of identity information to the greatest extent. The way of fingerprint verification is to verify the user's identity by verifying the user's fingerprint to ensure that the user's fingerprint information is consistent with the password.

This method has the advantage of no memory, the possibility of identity theft is very small. When the fingerprint information spreads in the network, it will encrypt the fingerprint information and mark the identity information to ensure that it will not be attacked maliciously. In the aspect of security design, staff should try their best to verify and ensure the authenticity and reliability of user information, and do a good job of information encryption in the process of information transmission, so as to avoid the attack of criminals as far as possible.

IV. Implementation of privacy protection scheme

Aiming at the problems of the existing technology, a risk situation prediction and analysis method based on security early warning is proposed

(1) Evaluate the possible impact of the security warning issued by the authority on the network;

(2) This paper analyzes the security attributes of the network, the partition of security domain and the configuration of border firewall strategy, and tries to analyze the possible impact or threat of various security attacks and harmful codes (such as viruses, Trojans, etc.) on the real network.

The technical scheme is divided into the following steps:

A. Collection of safety information: collection of different safety information includes

(1) Collection of security early warning information: automatically synchronize security early warning information from national authority (such as CNCERT) or well-known manufacturers; the invention mainly focuses on vulnerabilities, harmful codes, security threats, etc. in early warning information; after synchronization, it is disassembled into information affecting system (operating system), service or program, attack port, etc;

(2) Collection of information assets: the collection contents include related systems (including versions), vulnerabilities, patches, operation services and external ports provided on the information assets; the security value of each asset;

(3) Network information collection: collect the information about the network topology of each information asset, as well as the protection level of each subnet;

(4) Firewall policy collection: collect the access control policy information of each network boundary firewall, including area, interface, allowed IP address, allowed service port, protocol and other information.

B. Build a model

(1) Early warning model: early warning is tuple as follows:

Warning "warning type, warning name, warning level, {affected system and version}, {affected software version}, {affected port} >

(2) Information assets model: information assets are six tuples, as follows:

Information assets = < system and version information, {vulnerability}, {patch}, {installed software and its version}, {open service and its port}, value >

(3) Firewall policy model: the firewall policy is tuple as follows:

Firewall strategy_ < access direction, source IP address, destination IP address, source port, destination port, protocol, reject I accept >

(4) Network model: the network is tuple as follows:

Network = < information assets}, {border firewall policy}, protection level, {adjacent network}, {subordinate network} >

ISSN: 0010-8189

C. Analysis process

The comparison of privacy data transmission results is shown in Figure 4. According to different types of early warning, the steps of situation analysis include:

(1) First, start with the top-level network (generally the Internet boundary or export);

(2) Analyze whether the relevant access in the early warning will be accepted by its border firewall strategy (mainly analyze the source address, port and protocol of inward access), if not, turn to 5, otherwise turn to 3); when analyzing the lower level network, the relevant firewall strategy in the upper level network should also be selected as a part of its strategy, and the adjacent one should not be used;

(3) Whether the information generated in each vector (0, 1, etc.) will be affected by the matching of the software, vulnerability, or service elements in the system is analyzed as follows:

Matching vector = [matching 1, matching 2, matching n]

According to the matching situation and the weight of each factor, the possibility is calculated

(1) According to the possibility of impact, the risk situation of the affected information assets is comprehensively calculated as the predicted value of the risk situation of the network (regardless of the assets that are not affected at all): the predicted value of the network risk situation;

(2) Obtain adjacent or lower level networks, if there are unanalyzed networks, go to 2), otherwise go to 6);

(3) According to the risk situation prediction value of each network, calculate the overall risk situation prediction value: the overall risk situation prediction value.

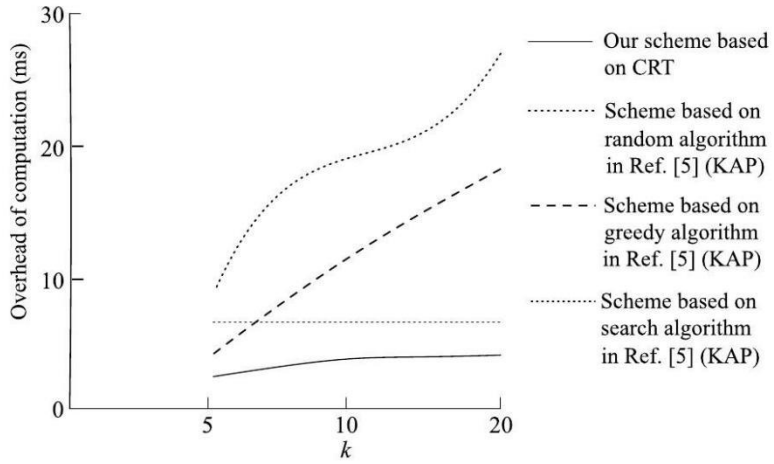

Fig 4: The comparison of privacy data transmission results

V. Conclusion

This paper implements a privacy protection scheme for intelligent transportation system in big data environment. The existing encryption algorithm, authentication and audit mechanism are used to solve the problem of data leakage during the transmission and storage of private data in the intelligent transportation system. After analyzing the characteristics of the private data in the intelligent transportation system, the specific generalization rules are customized, and the k-anonymity algorithm is written to anonymize the private data in the system. The anonymized data retains the data features on the premise of protecting the user's privacy, and can be used by data engineers to explore the value of the data again. The generalization content of anonymization algorithm in this paper only includes geographic location information and time information, while there are a large number of data types in its that can be generalized. In addition, the k-anonymity algorithm can't resist the chain attack when

multiple quasi identifiers are known at the same time.

References

[1] Shi Xinhong, Cai Bogen, Mu Jiancheng. Development of Intelligent Transportation System. Journal of Beijing Jiaotong University, 2002, 26 (001): 29-34

[2] Wang Guofeng, Song Pengfei, Zhang Yunling. Development and Prospect of Intelligent Transportation System. Highway, 2012 (5): 217-222

[3] Yang Ming, Song Xuefeng, Wang Hong. Image Processing for Intelligent Transportation System. Computer Engineering and Application, 2001, 37 (9): 4-7

[4] Ou Haitao, Zhang Weidong, Zhang Wenyuan. Urban Intelligent Traffic Control System Based on Multi-agent Technology. Acta Electronica Sinica, 2000, 28 (12): 52-55

[5] Gao Jingjing, Chen Xia, Tan Zhenhui. Cbtc Technology in Intelligent Transportation System of High Speed Railway. Electrified Railway, 2006 (06): 41-44

[6] Yu X, Jingli K, Wei D. Fuzzy Pattern Recognition Method for Vehicle Classification in Intelligent Transportation System. Journal of Beijing University of Technology: Natural Science Edition, 1999 (2): 171-175

[7] Wang Sheng Nan, Yu Mei, Jiang Gang Yi. Review of Vehicle Detection and Tracking Based on Video Image Processing in Zhitaizhi Transportation System. Computer Application Research, 2005 (09): 15-20

[8] Yao Zhisheng, Shao Chunfu, Xiong Zhihua. Short Term Traffic Flow Prediction of Road Network Based on Principal Component Analysis and Support Vector Machine. Journal of Jilin University (engineering Edition), 2008, 38 (01): 48-52

[9] Wu Bian, Qing Lin Bo, Wang Zheng Yong. Fast Background Modeling Algorithm Based on Moving Objects. Journal of Terahertz Science and Electronic Information Technology, 11 (2): 11-14.

[10] Hang Yue, Yang Rong-jie. the Development of Intelligent Transportation System and the Construction of Its Public Information Platform. Wuhan Daxue Xuebao/journal of Wuhan University, 2005, 29(4):560-563.

ISSN: 0010-8189