A

Seminar Report

A Privacy Security Risk Analysis Method for Medical Big Data in Urban Computing

Submitted in partial fulfillment of the requirements for the Award of the Degree

of

Master of Computer Applications

of

APJ Abdul Kalam Technological University

Submitted by

ALFIN WILLIAM

RegNo: LTVE17MCA058

Department of Computer Applications

COLLEGE OF ENGINEERING TRIVANDRUM

OCTOBER 2019

DEPARTMENT OF COMPUTER APPLICATIONS

COLLEGE OF ENGINEERING TRIVANDRUM

CERTIFICATE

Certified that this Seminar report entitled, ***"A Privacy Security Risk Analysis Method for Medical Big Data in Urban Computing"*** *is the paper presented by* ***" ALFIN WILLIAM" (Reg No: LTVE17MCA058)*** *in partial fulfillment of the requirements for the award of the degree of Master of Computer Applications of APJ Abdul Kalam Technological University during the year 2019.*

Prof. Vinya Vijayan	Dr. Sabitha.S
Co-ordinator	**Head of the Department**

Acknowledgement

First and for most I thank **GOD** almighty and to my parents for the success of this seminar. I owe a sincere gratitude and heart full thanks to everyone who shared their precious time and knowledge for the successful completion of my seminar.

I would like to thank **Dr Jiji C V**, Principal, College of Engineering Trivandrum, who helped me during the entire process of work.

I am extremely grateful to **Dr.Sabitha S**, HOD, Dept of Computer Applications, for providing me with best facilities and atmosphere for the creative work guidance and encouragement.

I would like to thank my coordinator, **Prof.Vinya Vijayan**, Dept of Computer Applications, who motivated me throughout the work of my seminar.

I profusely thank other Asst. Professors in the department and all other staffs of CET, for their guidance and inspirations throughout my course of study.

I owe my thanks to my friends and all others who have directly or indirectly helped me in the successful completion of this seminar. No words can express my humble gratitude to my beloved parents and relatives who have been guiding me in all walks of my journey.

Alfin William

Abstract

As an emerging computing mode, urban computing is mainly used to integrate, analyze and reuse urban resources by using perceptual computing, data mining and intelligent extraction to eliminate the phenomenon of data islands and provide wisdom for people to make decisions. But in the era of big data, the security and privacy leakage of users has become a major obstacle in urban computing. Taking medical big data as an example, this paper analyzed the risk of security and privacy leakage in the collection, transmission, storage, use and sharing of medical big data, and established a medical big data security and privacy leakage risk indicator system with 4 primary indicators and 35 secondary indicators. In addition, the weight of each indicator was calculated by GI method and entropy weight method. Then the fuzzy comprehensive evaluation model was established to verify the risk of medical big data security and privacy disclosure in urban computing. The results show that the risk of medical big data security and privacy leakage in the Grade II Level A hospitals is higher than that in the Grade III Level A hospitals, and in the life cycle of medical big data, the two stages of data storage, data use and sharing may cause more prominent problems of data security and privacy disclosure, while the data collection and data transmission are slightly less. Finally, the comparison of performance further proved the scientificity and effectiveness of this method.

Keywords: Urban Computing, Medical Big Data, Risk Indicators, Fuzzy comprehensive evaluation model

Contents

List of Figures

Chapter 1

Introduction

1.1 Current Scenario and Drawbacks

Driven by the wave of information technology, urban computing with urban backgrounds has emerged as an emerging field, with categories including transportation, environment, economics, social, medical services, and urban planning. It mainly provides data acquisition and analysis of various types of data in urban through intelligent extraction, data mining, and sensing technology to provide predictions and references for urban traffic conditions, disease spread, and house price trends, etc. From the perspective of the Internet, the core problem of urban computing is the use of intelligent sensors to collect and transmit data in the urban, while managing and analyzing the data. From the perspective of the Internet, the core problem of urban computing is the use of intelligent sensors to collect and transmit data in the urban, while managing and analyzing the data. However, in the category of urban computing research, the field of medical services is quite special, because the data it involves is basically human-based . As shown in Figure 1, with the development of science and technology, the medical information data flow in urban computing contains four aspects: medical resources (electronic medical records, clinical testing, doctor-patient behavior, etc.), subject-related data resources (such as life sciences, demography, etc.), industry related data resources (such as medical insurance government), Internet data resources (such as social media), involves data on physiology, psychology, disease prevention and medical management generated during the whole life cycle of people's birth, aging, illness and death, basic necessities, transportation, industry, agriculture and business, etc.

These data have important implications for medical research, commercial development, and more. According to the report, if hospitals can use the technology of intelligent extraction and data mining in urban computing to fully realize the value of these data, it will bring hundreds of billions of profits to the medical industry every year. For example, by means of big data analysis tools in urban computing to help doctors to develop more scientific and effective diagnostic programs. In addition, insurance companies can use these data to update the forecast model in time [9]. It is worth noting that the development of medical digitalization in urban computing can indeed bring some value to people's medical treatment and medical research, but we must also consider the security problems brought by technology while enjoying the convenience. At present, medical data leakage incidents are not uncommon, and there are many problems in the sharing and use of medical data etc. in urban computing. At the same time, the informationization of the medical industry is in a critical period of development, and various new forms of Internet and medical are booming. The widespread application of new technologies such as cloud computing, mobile Internet and Internet of things has brought new challenges and threats to medical big data. Therefore, it is urgent to solve the information security issues throughout the life cycle of medical big data production, collection, storage, sharing, exchange, and use in urban computing.

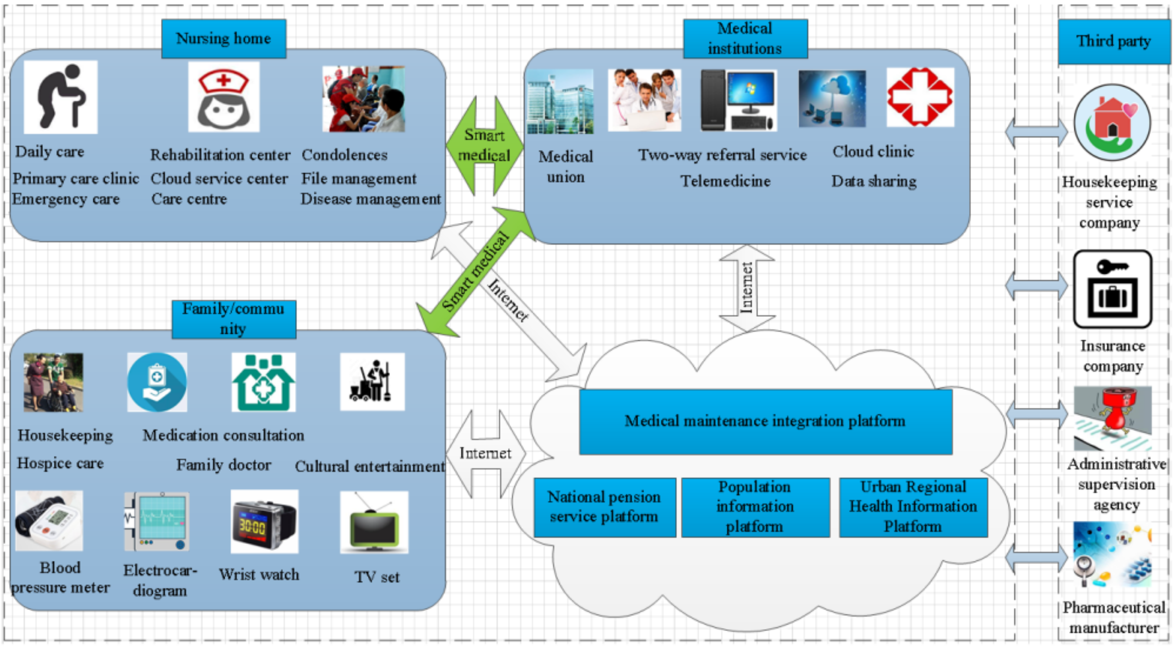


Figure 1.1: Pattern of medical information data flow and relationship between various fields in urban computing

Medical data including three kinds of physical states: pictures and files, video and data flow, language and text, but no matter which scene application is hidden, it is closely related to human's privacy information, must ensure the user's personal privacy and information security. If these private data are leaked, it will cause great harm to the patient's reputation and life, and even bring serious moral and ethical problems to the hospital. In addition, the information leakage in the medical big data environment is not only the data itself, but more serious is that hackers steal patients' social security accounts and personal finance by mining the hidden information behind the data, endangering the patient's personal and property security. Therefore, it is imperative to establish and improve the Internet and medical service security working mechanism in urban computing, and improve data security risk prevention measures and privacy protection.

1.2 Proposed Solution and Related Advantages

In this study, we firstly analyzes the security and privacy leakage risk indicators of medical big data collection, transmission, storage, use and sharing through brainstorming, Delphi, questionnaire, interview and field research. The combination of methods can reduce the influence of subjective factors to some extent. Then, the comprehensive weight of secondary indicators is calculated through the combination of GI method and entropy weight method. The combination of these two methods not only avoids the influence of subjective factors, but also avoids the loss of data hidden in objective information, making the evaluation results more accurate. Next, the fuzzy comprehensive evaluation model is used to evaluate the risk level of medical big data security and privacy disclosure. Finally, the risk reduction strategies corresponding to the four stages of data collection, transmission, storage, use and sharing in the life cycle of medical big data are presented.

The rest of this paper is organized as follows. Section 2 describes the progress and deficiencies in research on information security and privacy protection. Section 3 introduces the establishment process of medical big data security and privacy leakage risk evaluation index system and the calculation method of each index weight. Section 4 introduces the risk quantification method of medical big data security and privacy disclosure. Section 5 conducts experimental analysis with examples, and gives corresponding risk reduction strategies. Section 6 summarizes this article.

Chapter 2

Related Work

At present, few scholars specifically study the medical information security and privacy leakage in the urban At present, few scholars specifically study the medical information security and privacy leakage in the urban 2 VOLUME XX, 2019 This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2943547, IEEE R. Access JIANG, M. Y. SHI et al.: A Privacy Security Risk Analysis Method for Medical Big Data in Urban Computing computing environment, but some scholars have analyzed the information security issues in the construction of smart cities. For example, the literature mainly studied the problem from the aspects of management mechanism, technology and infrastructure, and clarified the new information security management mode and measures. K. Wang analyzed the personal information security issues in the context of smart cities from the legal level, and focused on the contents of legal documents. F. S. Ferraz et al. [16] focused on the issue of citizen privacy violations in a smart city environment. Literature analyzed the relationship between information security risk factors in smart cities, and established a set of information security risk assessment indicators and risk assessment methods. The above studies are only a macroscopic discussion of information security issues in smart cities and do not address specific data security and privacy issues. If the background of urban computing is abandoned, domestic and foreign scholars have already carried out relevant research on data security and privacy protection issues, including the following key technologies: access control technology, secure retrieval, and secure computing. Among them, access control technology is a research hotspot, and this method can be summarized into two categories: one is the technical route based on cryptography, and the other is the access control based on role and risk. The latter is more flexible and suitable for the complex environment of medical big data. A report published by foreign scholar JASON in 2004 is the first to introduce the concept of risk into the field of access control. The report gave some guiding principles and recommendations that should be met based on risk access control and defined the risk quantitative concept. Literature introduced the concept of risk into fuzzy theory, increasing the flexibility of the authorization strategy. There are also some scholars who have proposed a risk-based access control model for medical systems to solve the problem of patient privacy leakage caused by excessive authorization or illegal access in medical systems. Literature analyzed the information security issues in the cloud environment from different levels, and established a cloud computing security evaluation index system. Literature studied the privacy leakage of users, and established a user privacy risk assessment index system from different levels. G. Zhu et al.Constructed a privacy risk analysis indicator system from 3 aspects: social network platform, user behavior and external threats. Finally, the fuzzy risk analysis of social network privacy risk is carried out by AHP and entropy method. Y. Li researched the typical risks and information protection problems faced by personal information protection in the era of big data from the aspects of theory and justice. Meanwhile, corresponding countermeasures and suggestions were given.

In summary, it is not difficult to find that the current research on information security and privacy protection most focuses on the individuals and network systems in the cloud environment. In contrast, there are few studies on data security and privacy protection issues specific to the medical industry. Although some scholars have established risk indicators for data security and privacy leakage in some aspects of the medical big data life cycle, they do not give a specific risk assessment method for the medical industry. At the same time, the establishment of the risk indicator system and the calculation process of the weights of each index are not rigorous, and the results are greatly affected by subjective factors. Therefore, this paper is a necessary supplement to the research on security and privacy protection in the medical big data environment.

Chapter 3

Design Of Risk Evaluation System

3.1 Establish Risk Analysis Indicator System

n the production, collection, transmission, storage, use and destruction of urban medical big data, there are different security and privacy leakage threats in each link. After consulting the medical industry and information security related professionals and analyzing a large number of domestic and foreign references, it is found that the medical big data industry security issues are most prominent in the four stages of data collection, data transmission, data storage, data usage and sharing, and the relationship between them is shown in Figure 2.

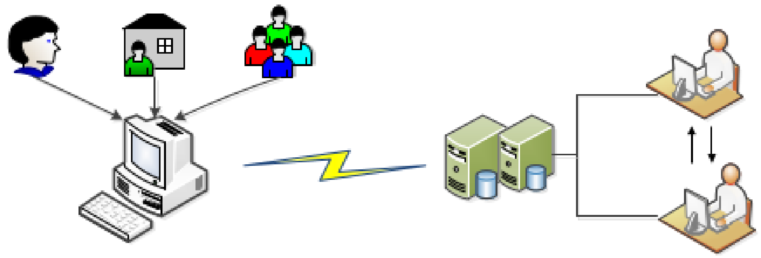


Figure 3.1: The relationship between data collection, transmission, storage, use and sharing in the medical big data life cycle

If the indicators of security and privacy leakage risk in the medical big data collection, transmission, storage, use and sharing can be correctly combed, controlled and managed, the probability of security and privacy leakage problems can be reduced to a large extent. Due to the rapid development of Internet and medical, the risk of security and privacy leakage of medical big data is partly intersected with the risk of network security and privacy leakage. In addition, the research achievements of the academic circle in network security have been quite rich, and the relevant indicator system is relatively mature. Therefore, the initial indicators of this paper will be extracted from the literature [27, 30, 32, 35], but in order to guarantee risk factors is more perfect, objective and accurate, this article will further by consulting relevant experts and relevant personnel of medical institutions(such as doctors, nurses, management personnel, technical personnel, etc.).In addition, by combining field research, Brain Storming, Delphi, Interview, questionnaire and other research methods, we modified and improved the preliminarily established risk index system, deleted some indexes of low importance, and dealt with repeated indexes in an interactive way. Finally, the risk factors of security and privacy disclosure of medical big data are identified and condensed from the aspects of data collection, transmission, storage, use and sharing. Combined with the indicator system construction process shown in Figure 3, a risk assessment indicator system including 4 first-level indicators and 35 second-level indicators is established. The specific indicators and contents are as follows. Medical big data security and privacy leakage caused by the data collection phase (A1),which includes 9 secondary indicators: Lack of data collection specification (B1),Patients lack the right to know when collecting data through smart devices (B2), Data ownership is not clear (B3), Medical staff operation error (B4), Lack of professional ethics staff (B5), Third party malicious behavior (B6), Wearable device positioning function(B7), Lack of supporting policies and supervision mechanism (B8), Lack of special laws and regulations (B9). Medical big data security and privacy leakage caused by the data transmission phase(A2), which includes 8 secondary indicators: Third party malicious behavior (B10), Lack of unified data transfer protocol standard (B11), Encryption and key management weak (B12), Service engine vulnerability (B13), Hardware security (B14), Software security(B15), Virus intrusion (B16), Hacker attacks (B17). Medical big data security and privacy breaches caused by the data storage phase(A3), which includes 17 secondary indicators: Medical staff operation error (B18), Encryption and key management weak (B19), Hardware security (B20), Software security (B21), Virus intrusion(B22), Hacker attacks(B23), Internal personnel stealing information (B24), Physical environment (B25), Virtual vulnerability (B26), Firewall vulnerability(B27), Access control mechanism is not perfect(B28), Identity authentication technology is not complete (B29), Safety audit (B30), Data monitoring (B31), Digital certificate reliability (B32), IDS reliability (B33), Data backup and recovery(B34). Medical big data security and privacy breaches caused by the data usage and sharing phase (A4), which includes 19 secondary indicators: Data ownership is not clear (B35), Lack of supporting policies and supervision mechanism (B36), Lack of special laws and regulations(B37), Encryption and key management weak (B38), Internal personnel stealing information (B39), Firewall vulnerability (B40), Access control mechanism is not perfect(B41), Safety audit (B42), Data monitoring (B43), Digital certificate reliability (B44), Data backup and recovery(B45), Data acquirer's dishonest behavior (B46), Electronic certification service is not perfect (B47), Electronic medical record sharing standards are not perfect (B48), Hospital information platform interaction standard is not standardized (B49), Telemedicine equipment and unified communication interaction standards are not standardized (B50), Data usage management system lacks (B51), Data sharing standard is not perfect (B52), Application management organization system is not sound enough (B53).

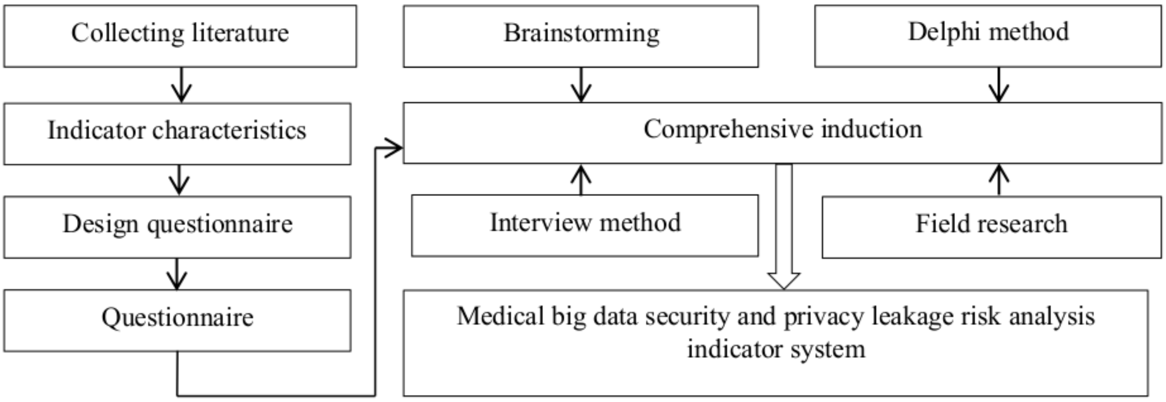


Figure 3.2: Steps for establishing risk indicators of medical bigdata security and privacy leakage

3.2 Calculate the Weight of Risk Analysis Indicators

3.2.1 Subjective Weight Calculation Based on GI Method

The GI method was first proposed by Guo Yajun as a new algorithm for a kind of decision problem. It overcomes the difficulty in constructing a two-two judgment matrix by AHP and ANP methods and the difficulty in passing the consistency test due to too many indicators. The central idea is also to compare the importance of the indicators in each indicator layer, but not to compare all the indicators in pairs. The specific calculation steps are as follows:

1. Sorting the importance of Indicators.

2. Determine the relative importance of each indicator.

3. Determine weight of indicator

4. Cluster decision result

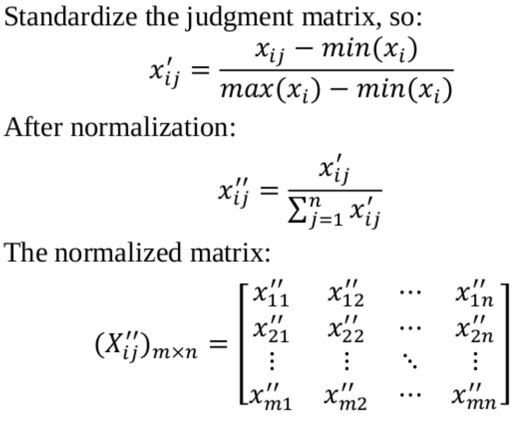
3.2.2 Objective Weight Calculation Based on Entropy Method

The concept of entropy first appeared in thermodynamics. In 1948, Shannon proposed the concept of information entropy and solved the problem of quantitative measurement of information. In information theory, information entropy represents the uncertainty measure of the system in the disordered state. The basic idea of entropy weight method is to determine the objective weight of indicators according to their variability. The larger the entropy is, the smaller the variability of the indicator is, the less information it will provide. On the contrary, the more information it will provide, the greater the role it will play in the comprehensive evaluation and the greater its weight will be. Since the weights of the indicators calculated by GI method will be affected by experts' personal subjective factors, while the entropy weight method is an objective weight determined according to the variability of the indicators, and the subjective weights calculated by the GI method can be corrected to make the evaluation result more scientific and reasonable. Therefore, this section focuses on the specific steps of calculating indicator weight by entropy weight method.

1. Construct judgement matrix X according to expert score table

$$X = (x_{ij})_{m\times n} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}$$

2. Data Standardisation and Normalization



Standardize the judgment matrix, so:

$$x'_{ij} = \frac{x_{ij} - min(x_i)}{max(x_i) - min(x_i)}$$

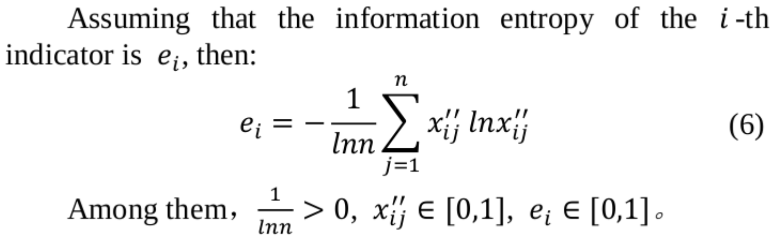
After normalization:

$$x''_{ij} = \frac{x'_{ij}}{\sum_{j=1}^{n} x'_{ij}}$$

The normalized matrix:

$$(X''_{ij})_{m\times n} = \begin{bmatrix} x''_{11} & x''_{12} & \cdots & x''_{1n} \\ x''_{21} & x''_{22} & \cdots & x''_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x''_{m1} & x''_{m2} & \cdots & x''_{mn} \end{bmatrix}$$

3. Calculate Entropy of each Indicator

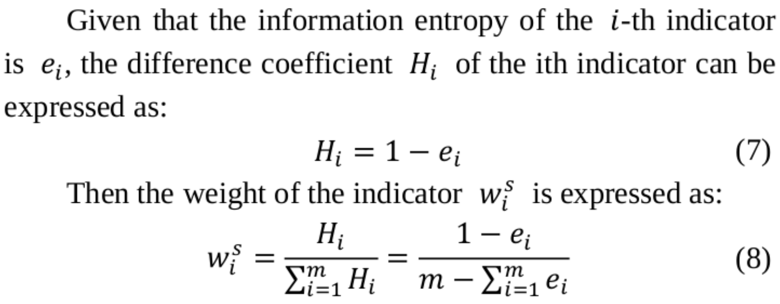


Assuming that the information entropy of the i-th indicator is e_i, then:

$$e_i = -\frac{1}{lnn}\sum_{j=1}^{n} x''_{ij}\, lnx''_{ij} \qquad (6)$$

Among them， $\frac{1}{lnn} > 0,\ x''_{ij} \in [0,1],\ e_i \in [0,1]$。

4. Calculate Weight of each Indicator



Given that the information entropy of the i-th indicator is e_i, the difference coefficient H_i of the ith indicator can be expressed as:

$$H_i = 1 - e_i \qquad (7)$$

Then the weight of the indicator w_i^s is expressed as:

$$w_i^s = \frac{H_i}{\sum_{i=1}^{m} H_i} = \frac{1 - e_i}{m - \sum_{i=1}^{m} e_i} \qquad (8)$$

3.2.3 Fuzzy Comprehensive Evaluation Method

The fuzzy comprehensive evaluation method is a method based on fuzzy mathematics to quantify some factors whose boundaries are not clear and difficult to quantify. The main idea is to comprehensively evaluate the membership level of the evaluation object from a number of factors. The basic principle of the fuzzy comprehensive evaluation method can be roughly summarized into three steps:

- Determine the indicator evaluation set of the evaluated object.
- Determine the weights and membership vectors of each indicator to obtain a fuzzy evaluation matrix.
- Fuzzy evaluation matrix and indicator weight vector for fuzzy operation to obtain fuzzy comprehensive evaluation results.

When quantifying the security and privacy leakage risk of medical big data through fuzzy comprehensive evaluation method, it is mainly carried out from the following six aspects.

1. Determining the indicator domain of the evaluation object.

   U = u1 , u2 .. um m = no. Of evaluation indicators

2. Determine the weight vector of each vector.

   Weights of indictors have been calculated by GI method and Entropy method on step 2.

3. Determine the evaluation level domain.

   Set of various evaluation results experts take on evaluation object, represented by

   V: V = v1 , v2 .... vn

   n = no. Of intervals for risk grades.

4. Single Factor Fuzzy Evaluation

   Fuzzy relation matrix R is determined :

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix}$$

$$r_{ij}(i = 1,2,\cdots,m; j = 1,2,\cdots,n)$$

5. Multi-Factor Fuzzy evaluation

   Fuzzy comprehensive evaluation result vector B can be obtained by performing an intergrated operation on each of indicator weight vector A and the fuzzy relation matrix R.

$$B = A \circ R$$

$$= [a_1, a_2, \cdots, a_m] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix}$$

$$= [b_1, b_2, \cdots, b_n]$$

6. Analyze fuzzy comprehensive evaluation results.

   Since the result of the fuzzy comprehensive evaluation is a vector rather than a specific value, we need to further process the result. According to the relevant literature, we know that the current analysis of fuzzy comprehensive evaluation results mainly includes the following two methods:

   - Principle of maximum membership degree
   - Principle of weighted average

   When dealing with problems, it is necessary to select appropriate methods according to actual conditions.

Chapter 4

Result Analysis

This paper mainly evaluates the risks of medical information security and privacy disclosure involved in urban computing, and selects two Grade III Level A hospitals and two Grade II Level A hospitals respectively. The results show that the Grade III Level A hospital slightly lower than the Grade II Level A hospital in terms of data security and privacy risks. When selecting the hospital, it is mainly considered whether the key risk indicators during the data collection, storage, transmission, analysis and use phases are consistent. The survey shows that the key indicators of the four hospitals are roughly the same, and there is not much difference. Therefore, this also explains the reliability of the evaluation method in this paper to some extent. In view of the above results, it is recommended that the Grade II Level A hospitals should strengthen management in medical data, conduct risk analysis on a regular basis, and adopt certain risk reduction strategies timely to ensure the security of medical data and user privacy. However, in order to further strengthen the persuasiveness, we conducted a comparative analysis of the accuracy of the method in this paper. The results are shown in Figure.

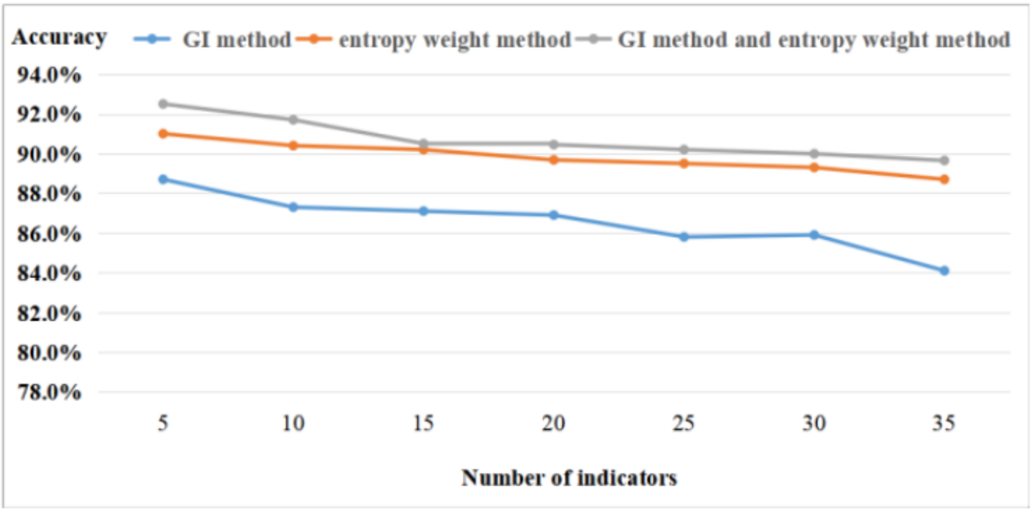


Figure 4.1: Accuracy Comparison Result

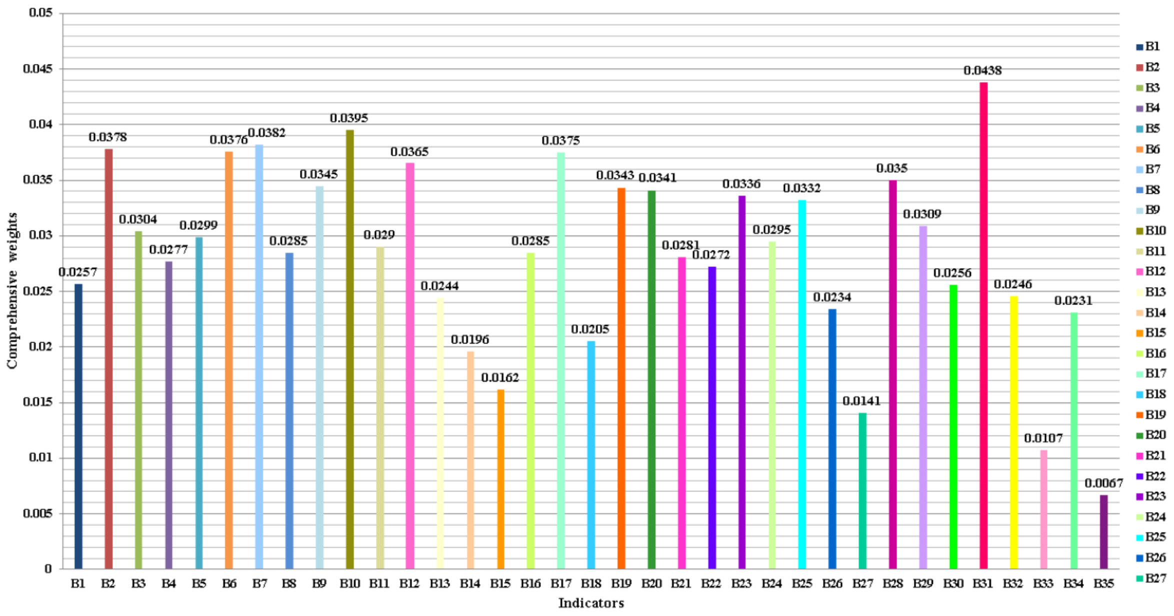


Figure 4.2: Comprehensive weight of secondary indicators

4.1 Preventive Solution

The weights of the primary indicators can be obtained from Figure: WA1,WA2,WA3,WA4 = 0.192,0.153,0.304,0.352. .As shown in Figure, in the four stages of the collection, transmission, storage, analysis and use of medical big data, the weights occupied by the data collection and data transmission stages are relatively small, while the weights of the data storage and data analysis and use phases are relatively higher, according to this situation, we specifically analyze the reasons, and combine with the comprehensive weight of secondary indicators in Figure 9 to give the corresponding risk reduction strategy.

4.1.1 Data Collection Phase

The data collection phase mainly collects data generated by sensors, smart devices, etc., most of which are raw data that has not been processed, resulting in the user's privacy completely out of the user's own control, so this stage of privacy issues is even more prominent. The key factors in the corresponding secondary indicators include: wearable device positioning function (0.0382), the patient's lack of right to know (0.0378), and third party malicious behavior (0.0376).

At this stage, privacy protection technologies such as cryptography, local differential privacy, social graph privacy, and location track privacy are recommended to prevent privacy breaches

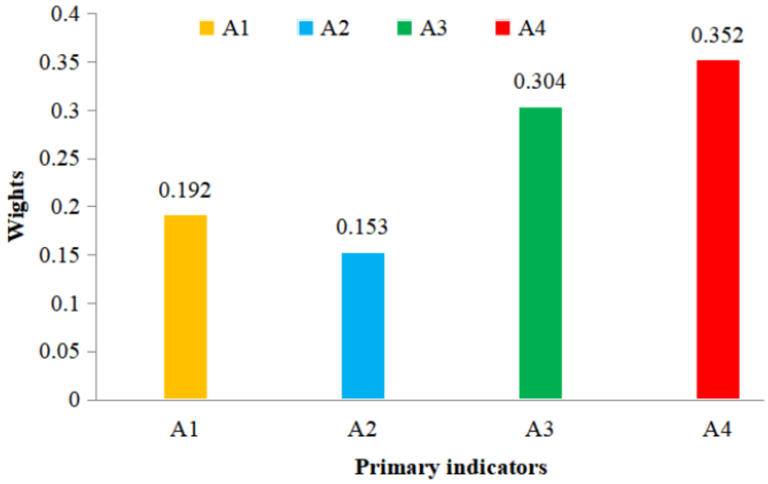


Figure 4.3: Weight of primary indicator

caused by malicious actions of wearable devices and third parties. In addition, an informed consent approach should be established that is consistent with the cultural characteristics of our country.

4.1.2 Data Transmission Phase

In the data transmission phase, the collected data is transmitted to a large database through terminal devices such as smart devices and sensors. The data security problem at this stage is more prominent. The Key factors in the corresponding secondary indicators include: lack of a unified data transfer protocol standard (0.395), service engine vulnerabilities (0.0365), and hacker attacks (0.0285).

At present, there are many kinds of data transmission protocols in the market, such as Bluetooth Medical Device Profile, IEEE 11073-104xx specification, etc. Only when a compatible data transmission protocol is established can data security be guaranteed to a certain extent. At the same time, it is recommended to use VPN technology or SSL communication protocol in the data transmission phase to prevent hackers from attacking during data transmission, and regularly scan for security vulnerabilities to prevent security problems from happening in time.

4.1.3 Data Storage Phase

A large amount of valuable data is stored together, which will not only become the target of external hackers, but also become the main target of internal personnel to steal information, and also include the unauthorized use of some data. Therefore, the security issues facing the data storage hierarchy are multifaceted, including data security, platform security, privacy security and other security requirements. The key factors in the corresponding secondary indicators include: internal personnel stealing information (0.0375), virtual vulnerability (0.0343) and firewall vulnerability (0.0341).

In response to these problems, it is recommended to establish a management system of "multi-level authorization and consistent responsibility", and strict implementation of medical data confidentiality regulations. Establish and improve the personal privacy information protection mechanism, severely punish the illegal stealing, trafficking of medical information. At the same time, it is necessary to update the system patch in time, fundamentally solve the vulnerability problem, strengthen the firewall configuration, and improve the defense capability. In addition, third-party software can be used to provide a certain security guarantee for data storage.
Finally, in order to further ensure the security and privacy protection of data storage, technologies such as access control, secure retrieval, and secure computing can be adopted for big data security. Privacy protection can use differential privacy, k-anonymity, etc.

4.1.4 Data Use and Sharing Phase

The data collection, transmission, and storage stages are all for data usage and analysis services. This stage mainly uses data mining and other technologies to extract information hidden inside the data. Strictly speaking, the privacy issue at this stage is more prominent, but since the data is in the hands of people, the purpose of using the data is not controlled. There are also no specific laws and regulations. So, it is not surprising that data security and privacy issues at this stage are most prominent. The corresponding secondary key indicators include: hospital information platform interaction standard is not standardized (0.0438), data acquirer's dishonest behavior (0.035), lack of special laws and regulations (0.0345), security audit (0.0336), digital certificate reliability (0.0332) and so on.

In view of these problems, medical institutions need to further supplement and improve the standardization of hospital platform interaction and electronic medical record interaction to prevent data security problems from occurring due to inconsistent standards in the interaction process. In addition, China should strengthen laws, regulations and supervision mechanisms on the application of medical big data to prevent the illegal use of data by users. Finally, medical care organizations should establish a trusted digital identity management system to ensure that access to medical data is manageable, controllable, and traceable. At the same time, each user of the access system should implement unified identity identification and management to prevent unauthorized access and behavioral repudiation.

Chapter 5

Conclusion

As an emerging cross-cutting area, urban computing has generated a large amount of data in urban space using sensing technology and large-scale computer infrastructure. Within the domain of computer science, urban computing draws from the domains of wireless and sensor networks, information science, and human-computer interaction. Urban computing now further uses many of the paradigms introduced by ubiquitous computing in that collections of devices are used to gather data about the urban environment to help improve the quality of life for people affected by cities. What further differentiates urban computing in the present scenario from traditional remote sensing networks is the variety of devices, inputs, and human interaction involved. In traditional sensor networks, devices are often purposefully built and specifically deployed for monitoring certain phenomenon such as temperature, noise, and light.As an interdisciplinary field, urban computing also has practitioners and applications in fields including civil engineering, anthropology, public history, health care, urban planning, and energy, among others.
At the same time, the arrival of the era of big data has brought challenges and opportunities for urban computing. This paper analyzed the security and privacy risk assessment methods of medical big data by taking the medical service industry in urban computing as an example. However, in a complex cross-domain environment, ensuring that users' access to data is manageable, controllable, and traceable will be the next step of research.

Bibliography

[1] R. Y. Yu, Y. Yang, L. Y. Yang, and G. J. Han *RAQ – a random forest approach for predicting air quality in urban sensing systems.* Sensors, vol. 16, no. 1, pp. 1-18, Jan. 2016.

[2] N. J. Yuan, Y. Zheng, X. Xie, Y. Z. Wang, K. Zheng, and H. Xiong *Discovering urban functional zones using latent activity trajectories* IEEE Trans, Knowl. Data Eng.vol. 27, no. 3, pp. 71-725, Mar. 2015.

[3] DY. Zheng, L. Capra, O. Wolfson, and H. Yang *Urban computing:concepts, methodologies, and applications* ACM Transactions on Intelligent Systems and Technology (TIST), vol. 5, no. 3, pp. 38, Sep. 2014.

[4] A. Amairah, B. N. Al-Tamimi, M. Anbar, and K. Alou. *Cloud computing and Internet of Things integration systems: A review " Data Science and Soft Computing, Cham.* Switzerland:Springer, vol. 843, pp.406-414, Sep. 2018.

[5] Z. Huang, J. Tang, G. Shan, J. Ni, Y. Chen and C. Wang. *An Efficient Passenger-Hunting Recommendation Framework with Multi-Task Deep Learning.*IEEE Internet of Things Journal, to be published, DOI: 10.1109/JIOT.2019.2901759.

[6] Y. Niu, M. M. Yan, H. Zheng, J. J. Yang. *" Research review on medical data access control in cloud computing environment*vol. 1, pp. 23-27,Feb. 2016.