International Journal of Data Warehousing and Mining
Volume 19 • Issue 5

Cybersecurity of Medical Data Based on Big Data and Privacy Protection Method

Jianhong Li, Second Affiliated Hospital of Wenzhou Medical University, China*
An Pan, Second Affiliated Hospital of Wenzhou Medical University, China
Tongxing Zheng, Second Affiliated Hospital of Wenzhou Medical University, China

ABSTRACT

Big data brings new opportunities to discover the new value of healthcare industry, since it can help us understand the hidden value of data deeply. This also brings new challenges: how to effectively manage and organize these datasets. Throughout the whole life cycle of publishing, storing, mining, and using big data in health care, different users are involved, so there are corresponding privacy protection methods and technologies for different life cycles. Data usage is the last and most important part of the whole life cycle. Therefore, this article proposes a privacy protection method for large medical data: an access control based on credibility of the requesting user. This model evaluates and quantifies doctors' credibility from the perspective of behavioral trust. Comparative experiments show that under the background of linear, geometric and exponential distribution trends and mixed trends, the regression model in this article is better than the existing methods in predicting trust accuracy and trust trends.

KEYWORDS

Access Control, Big Data, Medical Data, Privacy Protection

INTRODUCTION

As the next level of mobile Internet and the Internet of Things develops quickly, the cost of computing and storage of medical data is decreasing, while the efficiency of data processing is also increasing. In particular, high-tech technologies, such as artificial intelligence and sensing devices, are gradually being integrated into the medical industry, generating and accumulating a substantial quantity of medical data, both organized and unorganized. Because of its clear boundaries and consistent format for creation and storage, structured data may be integrated and analyzed automatically. Conversely, unstructured data must undergo extensive pre-processing to be usable by analysis tools because it cannot be read by machines. Meanwhile, the development and improvement of health-care information systems (HIS) has organized and summarized the scattered data from numerous departments of medical facilities or cooperation among facilities. An HIS facilitates accessibility, analysis, and sharing of

DOI: 10.4018/IJDWM.325222 *Corresponding Author

medical data, which has played a great role in improving the efficiency of medical treatment and modern management of medical institutions and has become an essential technical tool in medical activities (Huo et al., 2014). An HIS enables health-care institutions to gather, store, manage, analyze, and optimize patient treatment histories and other important data. Additionally, these technologies make it simple for medical professionals to obtain data regarding large-scale environments, such as trends in community health. Traditionally, the source of medical big data is mainly the vast quantity of information produced by individuals seeking medical care at communities and health-care facilities.

In the big data environment, how to minimize the misuse of information or misuse of permissions originating from within the system and thus causing damage to patient privacy is an important issue facing users of medical big data. Unauthorized access to patient medical records happens when a person accesses data, including protected health information (PHI), that is contained in those records without the proper consent, authorization, or other legal authority. Medical records for patients can be unlawfully accessed from a variety of places. Access control technology secures medical data at the use stage from the root cause of data leakage (an access rights assignment problem). This technology ensures that the data can be accessed by the right users with legal privileges through certain condition constraints. Access control technology ultimately serves the purpose of protecting the privacy and security of medical big data. Therefore, this paper proposes a big data access control model that meets the needs of the medical industry by combining the characteristics of the medical industry and the behavior patterns of doctors. On the basis of this model, a set of trust-based evaluation methods and constraint mechanisms are designed to evaluate users' trust from their behaviors and improve the access control granularity of the model.

Policymakers face a challenge to establish if the material accessible by physicians is required from a personal standpoint to create effective access control rules because of the profession, difficulty, and cost of learning clinical qualifications. The traditional coarse-grained access control policy can no longer meet the highly complex medical big data system. This paper proposes a trust-based access control system for medical big data that combines the traditional task-based access control (TBAC) model and role-based access control (RBAC) model and introduces trust assessment as a constraint to build a trust-based T-RBAC model to meet the needs of the medical industry. The model combines the advantages of TBAC and RBAC and introduces a behavior alert module that dynamically monitors user access behavior in real time and gives certain warnings to illegal users. Finally, it is experimentally proven that the warning mechanism can effectively reduce the frequency of malicious access to the system and improve the overall behavior of users.

In the rest of the paper we explain existing work in detail, discuss the process of cybersecurity of medical data based on big data, and share the details of the experiment and result analysis.

RELATED WORK

Traditional Access Control

Access control technology originated in the 1970s to address the necessity to handle authorized access to shared information on mainframe systems. This technology was primarily used to control access to critical resources by explicitly granting or limiting the scope and capability of access to data and information through some means after verifying the legitimate identity of the access user, thus preventing intrusion by illegal users and damage caused by inadvertent operations of legitimate users (Wang et al., 2015). In short, the technique was mainly used to decide which users access which information with which privileges, while ensuring the privacy and security of data. Lampson (2014) first proposed a formalization of access control as well as a description of the mechanism in 1974 and detailed the idea of accessing the subject and object. Subsequently, in 1985 the U.S. Department of Defense published the *Trusted Computer Security Evaluation Criteria (TCSEC)* that clarified the importance of access control in computer security systems (Ferraiolo et al., 2003). *TCSEC* pointed

out that there are two general access control mechanisms: autonomous access control and mandatory access control. With the continuous development of society, storage systems have become larger and more complex, and consequently, the corresponding access control systems have evolved on the traditional base model. To enhance the applicability of the model to make it more flexible on complex systems, Ferraiolo et al. (2003) first proposed the idea of simplifying security management by using roles, hierarchies, and constraints to organize privileges; this proposal was the earliest RBAC model.

Despite the majority of businesses having established cybersecurity processes, cybersecurity maturity scoring revealed that awareness and education levels were generally low (He et al., 2020). Recognizing key cybersecurity issues, options used by the health-care system, and emerging issues to address the huge rise in cyberwarfare is critical. Attackers used cyberwarfare to exploit weaknesses in individuals' work behavior and use of technology that businesses created by policies they decided to implement in reaction to the COVID-19 disease outbreak (O'Brien et al., 2021). Because of the heterogeneous distributed nature of big data and other characteristics, traditional access control techniques can no longer meet the needs in the context of big data. Chakraborty et al. (2006) for the first time acted on trust degree as a constraint in the access control mechanism. First, trust is graded based on the historical behavior records of the users to be evaluated, and roles are assigned according to the trust level. To reduce the effect of subjectivity in trust assessment, Zhang et al. (2014) divided trust assessment into two parts—direct trust and indirect trust—and assigned trust levels and privileges according to the final comprehensive trust.

Medical Big Data Access Control

Access control technology, one of the key technologies to ensure secure sharing of big data, has emerged as a data analysis hotspot. Although people appreciate the convenience that changes to health-care data offer, how to safeguard the protection of physician data as part of big data have also risen to prominence as a concern. Most of the research on access control conducted by scholars at home and abroad has been conducted in the context of big data, and less relevant research has been applied to the medical field (Wang et al., 2011). However, some scholars have also done research on big data in the medical field. Khan et al. (2015) analyzed the problems and potential threats faced by health-care digitization and proposed a fine-grained context-dependent access control approach for health-care big data based on conventional free access control, such as discretionary access control (DAC) and RBAC, and these researchers' proposed eTRON architecture effectively solves some problems in a priori authentication. Vawdrey et al. (2004) introduced trust into health-care information systems and constructed a framework for authentication and access control services based on trust negotiation. Shakhovska et al. (2019) analyzed the importance of health-care providers and security domains to establish trust between users, providers, and medical staff. These researchers used a privacy-preserving trust negotiation protocol to improve the security of health-care data transmission and sharing and showed how it can be used to automatically establish mutual trust between interacting parties. As health-care informatization continues to accelerate, the trust relationship between data providers, data managers, and data use in the context of medical big data becomes increasingly complex, so integrating trust assessment models with big data access control techniques is a trend for future research and a better solution for improving security and reliability when storing, accessing, and sharing medical data.

Medical Big Data

The use of big data is changing various fields of the medical industry, especially personalized medical services, and the corresponding data analysis services as a benchmark model in the field of big data are also starting to heat up. For example, in 2013, the Society of Clinical Oncology, an American nonprofit organization, proposed a medical project to help treat cancer with the help of big data. The project involves collecting and analyzing the treatment data of a large number of cancer patients and then using it to guide and assist physicians in the treatment of cancer-related patients.

Meanwhile, the hospital side is also actively exploring and practicing big data in health care. For example, Beijing Capital Medical University and Pfizer Investment established the first big data project applied to cardiovascular treatment in China to explore and test the value and operation mode of big data application in the field of cardiovascular diseases in China (Powsner et al., 2014). The current research on medical big data is mainly in the following areas.

Research on Data Storage, Integration, and Mining of Medical Big Data

As the process of medical informatization continues to accelerate, the massive amount of data requires higher storage performance for the database, and the structured database used in the past can no longer meet the storage requirements in the big data environment. In 2014, the U.S. Food and Drug Administration (FDA) established a public data open project. The project provided 3 million reports of raw data of desensitized adverse drug reaction records from 2004 to 2013. This data can be downloaded by companies, organizations, or even individuals for relevant mining and analysis to discover the value of this data.

Medical Big Data Security and Privacy Protection Research

The big data industry has driven the development of medical big data, and medical big data has gradually become one of the important strategic deployments of the country. Li analyzed the security risks of medical big data from the legal perspective, introduced the concept of "right to be forgotten," and gave the corresponding legal protection measures. Singh et al. (2020) investigated patients' and physicians' perceptions of the benefits and risks of informatization of medical records and found that patients were generally concerned that the medical information repository did not have sufficiently secure access constraints and authority control.

CYBERSECURITY OF MEDICAL DATA BASED ON BIG DATA AND PRIVACY PROTECTION METHODS

Medical Big Data Physician Behavior Privacy Protection Trust Assessment

The doctor's medical record is a valuable resource for health-related information and serves as a "credential" created by the doctor-patient relationship. It serves as a reflection of the doctor's commercial acumen, ethical guidelines, and capacity to provide complete care. It also serves as a crucial foundation for determining if medical conduct is standardized. To determine if a doctor is reliable or not, several access control models for big data in medicine use past treatment records among the key pillars. Although the current trust-based authentication mechanisms are more flexible and have a finer level of detail than conventional network access, a number of drawbacks remain. For example, current research puts the emphasis more on enhancing the model's performance level through the advancement and creativity of the accept quantitative method without recognizing the significance of the alerting system in the connectivity management system, as well as the trust quantification neglects to take into consideration.

Physician Behavior Model

To fulfill their job responsibilities, physicians' visits can be divided into two stages. The most frequent reason for medical errors is a breakdown in communication. These problems can occur verbally or in writing between a doctor, nurse, member of the health-care team, or a patient in a medical office or health-care system. Medical errors are sometimes caused by poor communication. First, after interacting with the patient, the physician identifies an initial target G0 for the patient (that is, the cause of the physician's suspicion based on the patient's relevant symptoms and combined with his or her own experience. LePort asserts that the key to a successful doctor-patient relationship is the capacity for empathy. This empathy is the most efficient technique to win a patient's trust and

establish a genuinely honest, respectful connection. Disruptive physician conduct is a pattern of personality qualities that interferes with the physician's ability to work effectively in the clinical setting. Subsequently, the physician selects the set of medical records related to the current target to confirm the correctness of that target as shown in Figure 1.

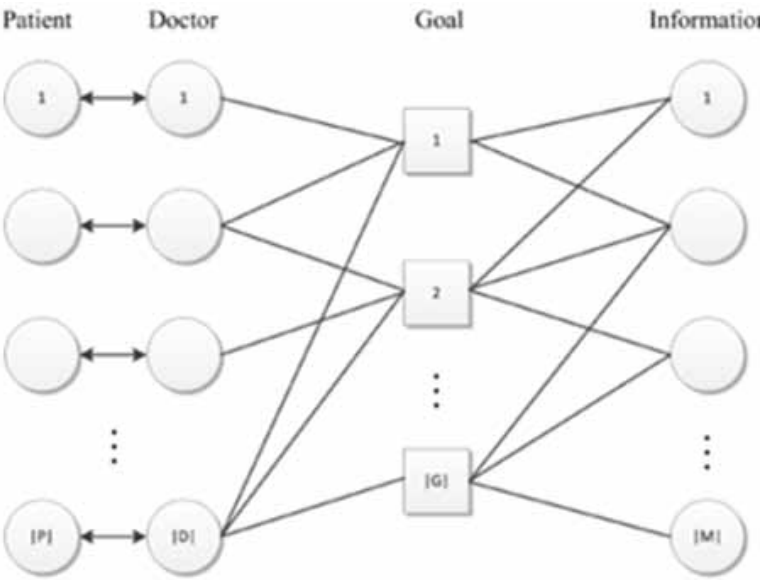
If the answer is accurate, the work is finished; if it is inaccurate, the cycle is continued until Liam leaves by himself or the ailment is verified, or the procedure is documented in the HIS as a full treatment lifetime, creating a health history.

Behavioral trust parameter construction

Node trust (NT), sometimes referred as instant trust, is the confidence in the participant's node level accessible behavior. This accepts variable is unrelated to the user's past context and simply reflects the characteristics of the recipient's present behavior. In reality, the method recognizes the reliability of customers' queries depending on the kind and quantity of information the customers have asked to acquire whenever they wish to connect to the HIS and receive assets. Determining the individual's behavioral trust (network trustworthiness) will not take much time because the kind and quantity of services demanded by the client are available. For the behavior warning module (BWM) in the process of carrying out recognition warning and knowledge return activities on the user's behavior, the network trustworthiness NT is also a crucial element. By designating each nearby node in the network map as either internal or external in reference to the local node of that network map, the Trusted Node Security feature enables administators to impose more stringent security standards when interacting with specific nodes in their networks. Individual nodal credibility is thus significant and essential as an element in the evaluation of customers' aggregate trustworthiness. The algorithm for calculating the individual's source vertex accessing behavior's trustworthiness is shown in equation (1).

$$T_{node} = N\left(M_{bool'}EM_{bool}\right) = 1 - \frac{1}{\sqrt{n}}\left(\sum_{i-1}^{n}\left(\mu_{ij} - e\mu_{ij}\right)^2\right)^{\frac{1}{2}} \tag{1}$$

Figure 1. Physician behavior model



The technology gives a mark to every connection to denote the type of that contact, and historically, relationship confidence is determined depending on the user's interactions with the program. The PIR calculation, which is the proportion of the number of affirmative permissions towards the overall number of approvals in the doctor's history accessing behavior, may be used to immediately acquire this data by accessing the ATDC-DB and doing so as shown in equation (2).

$$PIR = \begin{cases} 1, \#\xi > 1\#\# \\ \dfrac{AUTH_{\delta^+}}{AUTH_{\delta^+} + AUTH_{\delta^-}} \end{cases} \tag{2}$$

For each interaction between a user and the system, the behavioral trust assessment system performs a dynamic trust assessment and generates a corresponding comprehensive user trust CT, which is stored in the ATDC. A psychological instrument used to monitor, describe, explain, and predict behavior is a behavioral assessment test. This test is done to gauge and assess the many behavioral indicators of a candidate's cognitive capabilities. It should be noted that the CT is not used as the final trust output for access control policy invocation, but the trust value corrected by regression analysis is used for the final access control policy assignment.

Trust-Based Privacy-Preserving Health-Care Big Data Access Control Model

The central concept of the access management model is that dynamic resource evaluation can refresh consumer reputation in real time, but this approach to combating and discouraging malicious visitors by lowering user reputation is an after-the-fact punishment, meaning that the effects of malicious network consultations won't be seen until the following process of formative evaluation. The consequences of a user's malicious visit will be felt only after the next cycle of dynamic evaluation. Post-facto penalties have little to no effect on users who are engaging in malicious behavior and are not as effective in combating malicious behavior. Malicious visitors can be categorized into two categories: purposeful malicious visitors, who steal patient information for profit, and purposeless malicious visitors, who pry into patients' privacy out of curiosity as well as boredom. Early warning mechanisms can effectively address this issue because they have the psychologically deterring impact of timely warnings when malevolent visitors attempt to commit a "crime," which always raises the "crime cost" for visitors. The idea that penalizing someone for doing something bad, especially something bad like committing a crime, will stop malicious visitors and others from doing it again. Visitors' cost of crime is always increased by this self-perceived emotional penalty, which is especially potent for careless and malevolent viewers. Consequently, a realistic notification method can significantly increase the authorization model's overall efficacy.

Model Framework

Currently, access control mechanisms based on roles, identities, and attributes are generally used for health-care big data, where different responsibilities (e.g., pharmacists, nurses, or doctors) are allocated anomaly privileges, but such access control models do not systematically consider malicious access and intrusive behavior of the roles themselves. However, these approaches do not take into account mobile applications and have narrow authorization resolution, similar to the security solution in research that employs single-value quantifying data (value is taken, risk value) and subsequently integrates it with a predetermined authentication and authorization. The fine-grained network management paradigm FTOACM, which is specialist provides and fits the needs of security mechanisms for medical big data, is proposed in this work based on previous research. The proactive alert mechanism improves general user behavior in the network by raising the "price of offense" for unauthorized attackers, thus,

to some level, lowering the likelihood of unauthorized access to an HIS. A technique for limiting who has access to particular data is fine-grained access control. Fine-grained access control, as opposed to broad data access control, or coarse-grained access control, employs more subtle and flexible techniques for granting access. The FTOACM adopts a quantitative methodology, particularly the trust assessment concept of overall and individual proximity analogy, which is applicable to a variety of responsibilities (doctors, nurses, pharmacists, etc.) as long as there is a personal correlation, significantly improving the model's adaptability. Supervisors can change the guidelines using the ability to converse that the platform provides at the appropriate period. As a result, the framework for this research offers strong applicability and adaptability to accommodate various organizational designs and demands in the health-care sector (Figure 2).

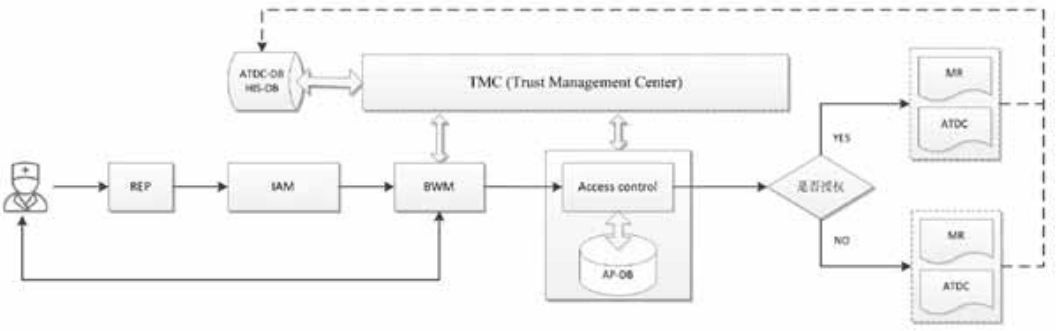
Doctors who want to use the HIS databases as well as medical information apps make up the majority of the model's entities. The prototype also comprises the following components:

- an aware subsystem that gives input and precautions predicated on the confidence level of the physician's centralized access behavior
- a protocol component that is in charge of overseeing and attempting to control access permissions according to predetermined permissions
- a component for the collection of approval and believe certifications

The validation and recognition control unit supports authentication whenever the doctor pertains to the collection and sharing of medical data. The detailed steps are as follows:

1. The user's request is approved by the request execution point (REP), which then transmits it to the identity authentication module (IAM).
2. The IAM delivers its authorization to the behavior warning mechanism after first verifying that the doctor's connection period, login IP, and other identification credentials are accurate (BWM).
3. The trust management center (TMC) receives a request from the BWM asking for the necessary behavioral trust.
4. After receiving the necessary data from the TMC, the BWM classifies the user's behavior according to its degree of trust and issues the associated alert or warning message.
5. The users will decide either to maintain the connection after obtaining the data from the BWM and will then transmit that decision back to the BWM (if the connection is interrupted, the operation will be aborted; if not, the procedure will proceed).
6. When the TMC receives a request from the BWM to resume connection, it contacts the BWM and calls up the Trusted Digital Certificate Database (ATDC-DB) to additionally analyze the user's level of confidence.

Figure 2. ATDC and medical record generation process

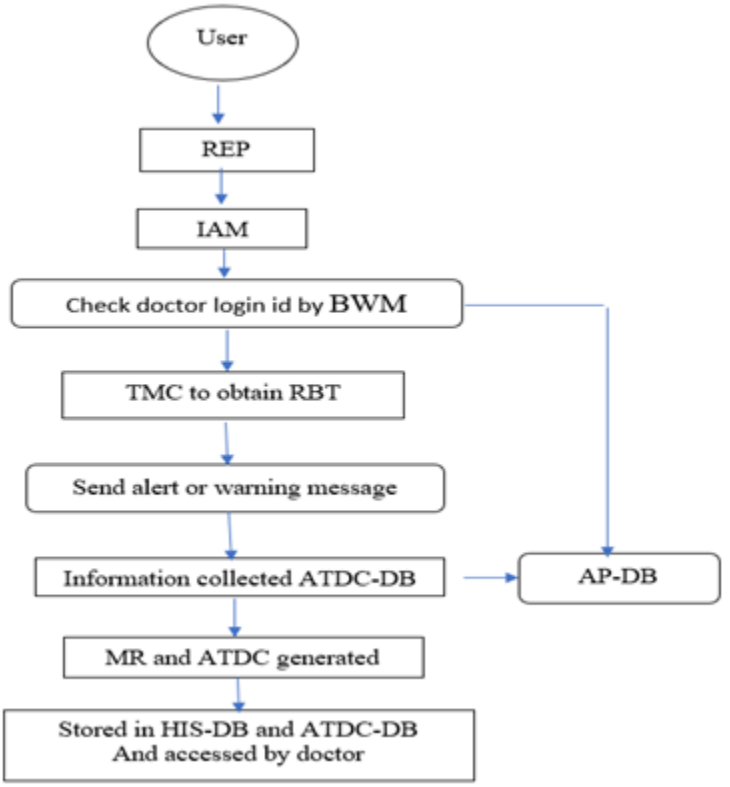


7. The access policy database (AP-DB) then communicates with the BWM to determine the individual's ultimate trustworthiness and renders a decision based on the access restrictions.
8. The most reliable resource is chosen from all those that are accessible and transmitted to the doctor if any are permitted. Then, the matching MR and ATDC are created and saved in the HIS-DB and ATDC-DB.
9. The related MRs and ATDCs will be created and recorded in the HIS-DB and ATDC-DB if the doctor is not permitted, and the client won't be capable of accessing the appropriate resources.
10. The doctor can now use its services and carry out its tasks or processes (figure 3).

Authentication Module

The initial line of defense for FTPACM is the identification component IAM, which offers fundamental coarse-grained network access via pertinent identification. The IAM is the particular virtual or physical setting in which a service instance is installed within an execution zone. An authentication module is a plug-in that gathers user data, such as a user ID and password, and compares it with entries in a database. The IAM is in charge of verifying the veracity of the requested user's identification and newly added enrollment. In a separate database, the component keeps track of enrolled users' identity and verification data. The user credentials are compared with the login credentials kept in the platform's internal database whenever a user wishes to log in. If the desired account is legitimate, the IAM gets the logged-in user's IP location as well as connection time variables and validates the accuracy of the values. If the demand is valid, it is sent to the user request handling (URH) for additional user

Figure 3. Flow chart for process of health information



process execution. The functions that manage client requests and create responses are known as request handlers. The servlet base classes define the request handler functions. A derivative servlet implements the function or functions that process the client request. Individual permissions are not able to change the preregistered customer IP and communication bandwidth settings in the model. Employment hours, essential assigned shift, vacation duty time, as well as duration of stay, are the four segments that make up the network delay. The algorithms for checking user legitimacy and user registration are shown below.

```
Algorithm 1. check_user_legitimacy()
11: else
12: User not registered.
13: Do you want to register?
14: if (user response == yes) then
15: user_registration ()
16: else
17: Exit.
18: END
Algorithm 2. user_registration()
1: Produce INPUT: (U_Id, U_PW, UI)/* U_Pw is the user password and
the UI is the user other informations. */
2: OUTPUT: (Success, Failure) of the user registration.
3: BEGIN
4: Enter your information and password.
5: if (User detials and password are valid) then
6: Success of the user registration
7: else
Enter valid user information.
END
```

Behavior Alert Module

This module is for behavior warning. The task of recognizing and evaluating users' access behavior falls to the BWM, which must then provide analysis and information in line with the behavior features. Users submit access requests to the BWM, which then transmits particular decrypt messages to the TMC and uses them to determine the trustworthiness of pertinent access behavior. Trust management can be viewed in two different ways: (1.) as the process of developing one's own reliability and (2.) as the process of determining the dependability of others. In the context of managing trust, both types of trust are regarded as being equally important. Python's warnings module manages them. With the warn() function, we may display user-generated warnings. The filterwarnings() function allows us to act on particular warnings. The module fuzzes the returned node trust value, fuzzy classifies its trust level, and finally provides corresponding information feedback to users according to the output trust level. The corresponding information is fed back to the user.

In this paper, according to the actual trust level and the characteristics of user access behavior, we established the corresponding fuzzy distribution function to fuzzify the node trust value for output, and according to the different output values, different prompt information was fed back to the user. The corresponding fuzzified affiliation function is shown in equation (3).

$$\mu_1(x) = \begin{cases} 1, & x \le a \\ \dfrac{b-x}{b-a}, & a \le x \le b \\ & x > b \end{cases}$$

$$\mu_2(x)=\begin{cases}\dfrac{x-a}{b-a}, & a\le x\le b\\ 1, & b\le x\le c\\ \dfrac{d-x}{d-c} & c\le x\le d\\ 0, & x\langle a,x\rangle d\end{cases}$$

$$\mu_3(x)=\begin{cases}0, & x<a\\ \dfrac{x-c}{d-c}, & c\le x\le d\\ 1, & x>d\end{cases} \tag{3}$$

According to the experience subjective selection of certain forms of fuzzy distribution, and then according to the actual measurement data to confirm the value of parameters, we finally determined, after a large number of data test simulation and reference to expert opinion, that a = 0.3, b = 0.5, c = 0.7, and d = 0.9. The warning affiliation function is shown in Figure 4.

$$\mu_1(x)=\begin{cases}1, & x\le 0.3\\ \dfrac{0.5-x}{0.2}, & 0.3\le x\le 0.5\\ 0, & x>0.5\end{cases}$$

$$\mu_2(x)=\begin{cases}\dfrac{x-0.3}{0.2}, & 0.3\le x\le 0.5\\ 1, & 0.5\le x\le 0.7\\ \dfrac{0.9-x}{0.2}, & 0.7\le x\le 0.9\\ 0, & x\langle 0.3,x\rangle 0.9\end{cases}$$

$$\mu_3(x)=\begin{cases}0, & x<0.7\\ \dfrac{x-0.7}{0.2} & 0.7\le x\le 0.9\\ 1, & x>0.9\end{cases}$$

As per the information criteria contained in the database, the feedback involves sending feedback notifications to the customer. Table 1 lists the particular feedback mechanism criteria.

Access Control Policy

We provide real legislation for FTPACM permission behavior management that entails the use of user identity and the feeling of security to regulate accessing characteristics depending on an access control collection. We categorize user access possibilities into three categories—normal access, additional

Figure 4. The affiliation function of BWM

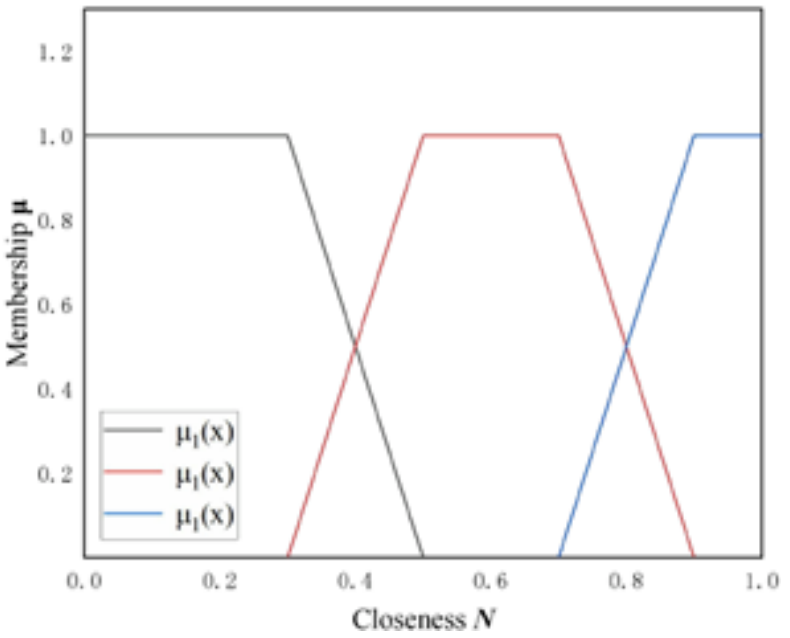


Table 1. Feedback mechanism of BWM

Feedback Type	Mapping Relationships	Trigger Conditions	Feedback Action
P_1	$P_1 \rightarrow \mu_1$	$\max\limits_{1\le x\le n}\left\{\dfrac{\mu_i(x)}{x}\right\}=\mu_1$	Send M_1 to the doctor
P_2	$P_2 \rightarrow \mu_2$	$\max\limits_{1\le x\le n}\left\{\dfrac{\mu_i(x)}{x}\right\}=\mu_2$	Send M_2 to the doctor
P_3	$P_3 \rightarrow \mu_3$	$\max\limits_{1\le x\le n}\left\{\dfrac{\mu_i(x)}{x}\right\}=\mu_3$	Send M_3 to the doctor

security, and undetermined access—to improve appropriate access in the FTPACM architecture. The access control guidelines for FTPACM are shown in Figure 5.

These conjunction, disjunction, and sequence pattern elements are obtained from the system design method. A and B are the initial and final edges of the system design, E is the order's path, and G provides extremely fast delivery.

EXPERIMENT RESULTS AND ANALYSIS

We randomly generated 100 system users and divided them into four categories according to their behavior levels (25% are high-level malicious, 25% are medium-level malicious, 25% are low-level

Figure 5. FTPACM rules

Rule1 user.id = {true} ∧ ac.time = {legal} ∧ ac.situation = {normal} ∧ $t_{range}(T_{min}, T_{max})$ = {true} ⟶ H_{sign} = $\{\delta^+\}$ ∪ H_{avl} = {full view} ∪ H_{record} = {Generated}

Rule2 user.id = {true} ∧ ac.time = {illegal} ∧ ac.situation = {normal} ∧ $t_{range}(T_{min}, T_{max})$ = {true} ⟶ H_{sign} = $\{\delta^-\}$ ∪ H_{avl} = {no view} ∪ H_{record} = {null}

Rule3 user.id = {false} ∧ ac.time = {legal} ∧ ac.situation = {normal} ∧ $t_{range}(T_{min}, T_{max})$ = {true} ⟶ H_{sign} = $\{\delta^-\}$ ∪ H_{avl} = {no view} ∪ H_{record} = {null}

Rule4 user.id = {true} ∧ ac.time = {legal} ∧ ac.situation = {normal} ∧ $t_{range}(T_{min}, T_{max})$ = {false} ⟶ H_{sign} = $\{\delta^-\}$ ∪ H_{avl} = {no view} ∪ H_{record} = {Generated}

Rule5 user.id = {true} ∧ ac.time = {legal} ∧ ac.situation = {emergency} ∧ $t_{range}(T_{min}, T_{max})$ = {true} ⟶ H_{sign} = $\{\delta^+\}$ ∪ H_{avl} = {full view} ∪ H_{record} = {Generated}

Rule6 user.id = {true} ∧ ac.time = {illegal} ∧ ac.situation = {emergency} ∧ $t_{range}(T_{min}, T_{max})$ = {true} ⟶ H_{sign} = $\{\delta^+\}$ ∪ H_{avl} = {full view} ∪ H_{record} = {null}

Rule7 user.id = {false} ∧ ac.time = {legal} ∧ ac.situation = {emergency} ∧ $t_{range}(T_{min}, T_{max})$ = {true} ⟶ H_{sign} = $\{\delta^-\}$ ∪ H_{avl} = {no view} ∪ H_{record} = {null}

Rule8 user.id = {true} ∧ ac.time = {legal} ∧ ac.situation = {emergency} ∧ $t_{range}(T_{min}, T_{max})$ = {false} ⟶ H_{sign} = $\{\delta^-\}$ ∪ H_{avl} = {part view} ∪ H_{record} = {Generated}

Rule9 user.id = {true} ∧ ac.time = {legal} ∧ ac.situation = {undefined} ∧ $t_{range}(T_{min}, T_{max})$ = {true} ⟶ H_{sign} = $\{\delta^-\}$ ∪ H_{avl} = {no view} ∪ H_{record} = {null}

malicious) and then the remaining 25% are legitimate users. Each user has 100 history records corresponding to the randomly generated data and meets the initial PIR of 0.2500, 0.5500, 0.7500, and 0.9500, respectively, in the forms shown in Tables 2 and 3.

We generated all the above simulation data using Java and ran the simulation experiments on a physical host with 64-bit Windows 10 as the operating system. The host configuration was an AMD Ryzen 5 3500U with a Radeon Vega Mobile Gfx with a 2.10GHz CPU and 8 GB of RAM. A midrange processor for laptops in the mainstream class was the AMD Ryzen 5 3500U. The Ryzen 5 3500U is a superb processor for home and business computing tasks, and its four cores and maximum clock speed of 3.7 GHz enable it to run numerous apps at once.

The goal of this series of trials is to determine how the warning module's involvement will affect people's general access to doctors. We assume that the warning mechanism has a 30% blocking effect on malicious behavior; that is, after receiving a notification from the warning module, users have a 30% chance of immediately stopping their access activity and a 70% chance of continuing it. For ease of discussion, we refer to the concept used in this study as the W model and the conventional model without the warning mechanism as the T model.

Whenever individuals in the W model communicate with one another, 30% of the visits that BWM deems "malicious" will be arbitrarily stopped. We averaged the results of three consecutive encounters to provide three test trust levels. This study used the growth ratio R index to reflect. The R implies that the greater the influence on users, the stronger the effects of the BWM process are. to precisely illustrate the influence of the W model on consumers with varying degrees of malice. The reputation growth ratio for prior interactions as well as the trust growth ratio are shown in equation (4).

$$R_{CCT} = \frac{CCT_{W-model} - CCT_{T-model}}{CCT_{T-model}}$$
$$R_{PIR} = \frac{PIR_{W-model} - PIR_{T-model}}{PIR_{T-model}} \quad (4)$$

The distribution in Figure 6a compares the credibility levels for the four different user groups in the two approaches, and it is clear that the W model has greater trustworthiness than the T model.

Table 2. CCT dataset for users

	CCT_1	CCT_2	CCT_3	…	CCT_{99}	CCT_{100}
User1	0.8387	0.7745	0.7928	…	0.8451	0.8988
User2	0.6003	0.6128	0.6172	…	0.5947	0.5996
			…	…		
User100	0.1700	0.1800	0.1900	…	0.3100	0.3200

Table 3. User's PIR Database

	PIR_1	PIR_2	PIR_3	…	PIR_{99}	PIR_{100}
User1	0.7600	0.7700	0.7600	…	0.7300	0.7200
User2	0.4500	0.4600	0.4700	…	0.5800	0.5700
			…	…		
User100	0.1700	0.1800	0.1900	…	0.3100	0.3200

According to the CCT market growth chart, the more mischievous the customer conduct, the faster trust grows (high: 0.1484, normal: 0.0208) and the more pronounced the influence of BWM is. The impact of BWM is clear. It is evident from the frequency distribution in Figure 6b that the PIR of the W model is superior to that of the T model, so it is completely apparent from the PIR overall growth graph that the more maliciously inclined the consumers are, the more effectively the consequence is. The histogram compares the PIR for the four different forms of user groups in the two approaches. The Protein Information Database (PIR) is a comprehensive, open-access resource for protein informatics that aids in scientific discovery and genomic and proteomic research. The Protein Sequence Database (PSD), a database of more than 283,000 annotated protein sequences that spans the whole taxonomic spectrum, is maintained by PIR. In other words, the dynamic authentication scheme with the BWM may successfully block and curtail some dangerous user activity without harming regular user access, and it can aggressively enhance the total access rate of users through recognizing and making announcements. Figure 5 illustrates how the approach further enhances system users' general behavior while preserving standard access control features.

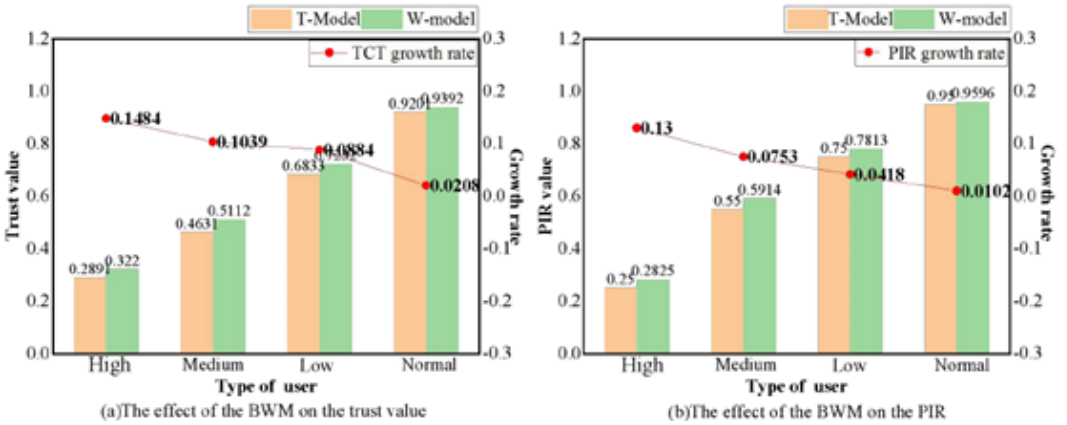
The model architecture consists of several modules; namely, the authentication besides request processing module, the behavior alert module, and the access policy module. This architecture introduces the functions and implementation principles of each module, respectively. Finally, we proved through experiments that the addition of the early warning mechanism can enhance the total behavior of users by enhancing the user's and indicators in the system.

CONCLUSION

In this paper we analyzed the sources and characteristics of big medical data as well as studies and the literature on the background of big medical data. We found that one of the key factors affecting the popularization and development of big medical data is the safety of medical data use. To ensure the privacy and security of medical data in the use phase, we proposed an access control model for the large medical data field based on the characteristics of the medical industry. Based on this model, we evaluated trust of doctors from the behavioral dimension, which improved the granularity of access control of the model.

In this paper, many aspects of the trust-based access control model for big medical data are in the research exploration stage, so future work will be considered for them. The selection of parameters of the fuzzy membership function in the fuzzy classifier of the early warning mechanism is based on expert experience, and there are specific factors directly related. Future work should be analyzed according to actual user data, and the influence of the main related factor should be minimized.

Figure 6. Performance comparison of CCT and PRI



REFERENCES

Chakraborty, S., & Ray, I. (2006, June). TrustBAC: Integrating trust relationships into the RBAC model for access control in open systems. In *Proceedings of The Eleventh ACM Symposium on Access Control Models and Technologies*, (pp. 49–58). Association for Computing Machinery. doi:10.1145/1133058.1133067

Ferraiolo, D. F., Kuhn, D. R., & Chandramouli, R. (2003). *Role-Based Access Control*. Artech House. Inc.

He, Y., Aliyu, A., Evans, M., & Luo, C. (2021). Health care cybersecurity challenges and solutions under the climate of COVID-19: Scoping review. *Journal of Medical Internet Research*, *23*(4), e21747. doi:10.2196/21747 PMID:33764885

Huo, C., & Zhenqiang, W. (2014). Patient-oriented access control model for privacy protection in medical information systems. *Computer Applications and Software*, *11*, 75–77.

Khan, M. F. F., & Sakamura, K. (2015, May). Fine-grained access control to medical records in digital healthcare enterprises. In *Proceedings of the* 2015 *International Symposium on Networks, Computers and Communications (ISNCC)*, pp. 1–6. IEEE. doi:10.1109/ISNCC.2015.7238590

Lampson, B. W. (2014). Protection. *ACM SIGOPS Operating Systems Review*, , *8*(1), 18–24. 10.1145/775265.775268

O'BrienN.MartinG.GrassE.DurkinM.DarziA.GhafurS. (2020). *Cybersecurity in healthcare: Comparing cybersecurity maturity and experiences across global healthcare organizations*. SSRN. 10.2139/ssrn.3688885

Powsner, S. M., Wyatt, J. C., & Wright, P. (1998). Opportunities for and challenges of computerisation. *Lancet*, *352*(9140), 1617–1622. doi:10.1016/S0140-6736(98)08309-3 PMID:9843122

Shakhovska, N., Fedushko, S., Melnykova, N., Shvorob, I., & Syerov, Y. (2019). Big data analysis in development of personalized medical system. *Procedia Computer Science*, *160*, 229–234. doi:10.1016/j.procs.2019.09.461

Singh, A., & Chatterjee, K. (2020). An adaptive mutual trust based access control model for electronic healthcare system. *Journal of Ambient Intelligence and Humanized Computing*, *11*(5), 2117–2136. doi:10.1007/s12652-019-01240-2

Vawdrey, D. K., Sundelin, T. L., Seamons, K. E., & Knutson, C. D. (2003, September). Trust negotiation for authentication and authorization in healthcare information systems. *In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. (IEEE Cat. No. 03CH37439), (pp. 1406–1409). IEEE. doi:10.1109/IEMBS.2003.1279579

Wang, Q., & Jin, H. (2011, March). Quantified risk-adaptive access control for patient privacy protection in health information systems. *In ASIACCSA '11: Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security,* (pp. 406–410). Association for Computing Machinery. doi:10.1145/1966913.1966969

Wang, Y., Yang, J., & Xu, C. (2015). A review of cloud computing access control technology research. *Software Journal*, *26*(05), 1129–1150.

Zhang, K., & Pan, X. Z. (2014). An access control model based on user behavior trust under cloud computing. *Computer Applications (Nottingham)*, *34*(04), 1–4.

APPENDIX

Abbreviation and expansion

Abbreviation	Expansion
HIS	Health-care information system
RBAC	Role-based access control
FDA	Food and Drug Administration
NT	Node trust
BWM	Behavior warning module
PIR	Protein information database
REP	Request execution point
IAM	Identity authentication module
TMC	Trust management center
AP-DB	Access policy database
ATDC-DB	Trusted Digital Certificate Database
URH	User request handler
PSD	Protein sequence database

Jianhong Li, Information Center, Big Data Center, the Second Affiliated Hospital of Wenzhou Medical University, Wenzhou 3250027, Zhejiang, China.

An Pan, Information Center, Big Data Center, the Second Affiliated Hospital of Wenzhou Medical University.

Tongxing Zheng, Information Center, Big Data Center, the Second Affiliated Hospital of Wenzhou Medical University.