




Computer-Aided Secure Access and Management of Wireless Medical Devices using Internet of Things and Biometric Technology

Haifeng Chen^{1*} 

¹School of Information Engineering, Lianyungang Technical College, Lianyungang 222006, Jiangsu, China, 1994060014@lygtc.edu.cn

Correspondence author: Haifeng Chen, 1994060014@lygtc.edu.cn

Abstract. In order to improve the management effect and security of wireless medical devices, this paper proposes a method to realize the secure access and management of wireless medical devices by using Internet of Things and biometric technology, which can not only improve the management efficiency of medical devices, but also effectively improve the security of medical devices. Moreover, aiming at the outlier problem of medical diagnosis data, an outlier detection method based on the law of gravity is proposed. In addition, aiming at the problem that there are a large number of default values in medical diagnosis data, a default value processing method based on ignoring-deleting-filling is proposed. Finally, this paper combines biometric technology and Internet of Things technology to build the system. Through the analysis, it can be seen that the proposed method of using Internet of Things and biometric technology to achieve secure access and management of wireless medical devices can not only improve the management efficiency of medical devices, but also effectively improve the security of medical devices.

Keywords: Internet of Things; biometrics; wireless; medical equipment; safety; Computer-Aided Secure Access

DOI: <https://doi.org/10.14733/cadaps.2024.S9.82-103>

1 INTRODUCTION

Intelligent medical treatment refers to the regional medical and health service mode with personal health records as the core, which is established by using diagnosis and treatment technology, modern information technology and relies on the information technology platform. Intelligent medical treatment is to build a regional medical information platform for health records and use the

most advanced Internet of Things technology to realize the interaction between patients and medical personnel, medical institutions and medical equipment, and gradually achieve informationization. It can also be said that informationization has created intelligent medical care. Moreover, the intelligent medical information system integrates all kinds of patient information, which can be used by medical staff, government administrators, patients and even other medical service personnel. The integration of patients' various information makes medical and health services more convenient, scientific and even economical. It is mainly manifested in better availability of medical data, prevention of medical errors and unnecessary repeated examinations, cost saving, reduction of waiting time and medical expenses, etc. In addition, the overall model of the medical and health industry is constantly optimized, so that all components of the medical ecosystem can benefit.

The industry characteristics of the Internet of Things are mainly reflected in its application field, and its application layer combines with industry requirements to achieve intelligent applications of the Internet of Things. In the medical field, remote medical applications of the Internet of Things technology can achieve intelligent healthcare, which utilizes information recognition technologies such as sensors to achieve interaction between patients, medical personnel, medical institutions, and medical equipment through the network, achieving intelligent medical services. Intelligent healthcare can make hospitals more information-based and improve the efficiency of patients' internal settlement and treatment [14]; Being able to achieve medical and health information sharing, medical business collaboration, and health business linkage through regional medical collaboration, solve the problem of difficult and expensive medical treatment for residents, and improve the efficiency of cross regional medical treatment; By focusing on patients and organizing remote collaboration between medical personnel and equipment, intelligent healthcare can be implemented in all time and space to address issues such as chronic disease prevention and treatment, aging society, remote areas, and first aid. Intelligent healthcare provides a solution for China's relatively scarce medical human resources and centralized deployment of medical resources [12].

Biometric recognition has its own natural advantages. The biological characteristics of individuals, especially their physiological characteristics, are inherently immutable and have natural characteristics such as uniqueness and permanence. Therefore, biometric recognition not only has higher accuracy, but also can identify whether someone has multiple identity documents under different names without the need for memory, making it difficult to forge. Compared to traditional recognition methods, the biometric recognition process can be completed in seconds, making it safer and more convenient [3]. Computer-Aided technologies also assist in capturing biometric data accurately. Devices such as fingerprint scanners, iris scanners, and facial recognition cameras are equipped with sensors and image processing capabilities to capture high-quality biometric samples. These devices ensure that the data collected is reliable and suitable for accurate recognition. Furthermore, computer-aided systems can provide real-time feedback and guidance during the data capture process, ensuring that individuals present their biometric features correctly and optimizing the quality of the collected data.

There are many shortcomings in traditional identity recognition methods. Traditional identity recognition methods, such as username and password, have many shortcomings: they are prone to theft, loss, or forgetting, and are prone to personal information leakage, identity theft, fraud, and so on. Especially in recent years, telecommunications fraud has been frequent, and personal information leakage and trading have emerged one after another. Traditional identity recognition methods can no longer meet people's needs for identity authentication. Compared to traditional identity recognition methods, biometric recognition overcomes the above shortcomings and requires the real-time appearance of individuals, which has a strong trend of replacing traditional identity recognition methods [18].

Face recognition refers to the use of computer technology to analyze and compare facial videos or images, extract effective facial recognition features or information from them, and ultimately recognize an individual's identity based on these features or information. Face recognition can be divided into two types: visible light face recognition and active light face recognition [10].

There is a circular area between the black pupils and the white sclera on the surface of the human eye, which is called the iris. The iris exhibits rich texture information under infrared light, such as detailed features such as spots, stripes, filaments, crowns, and dimples. Generally speaking, an individual's iris begins to develop from the third month of embryonic stage and its main texture structure is basically formed by the eighth month. Once an individual's iris is formed, it remains almost unchanged for a lifetime thereafter, unless they suffer from certain diseases or undergo surgery that endangers the eyes [7]. Iris recognition is the identification of individuals by comparing the similarity between their iris features. The higher the similarity, the greater the likelihood that the two being compared are the same person. Among all biometric recognition technologies, except for DNA recognition, which has not yet been widely used, iris recognition is the most accurate and error prone recognition technology. Due to its high accuracy, iris recognition is currently mainly used in areas or places with high security requirements, such as the entry and exit of military hospitals [15].

In order to achieve secure storage and sharing of medical data, many scholars have conducted a series of research and exploration. In traditional public key cryptosystems [13], data owners can encrypt plaintext with the public key, so that only data users with the corresponding private key can decrypt the ciphertext, effectively solving the problem of key distribution in traditional symmetric encryption. However, when a large amount of medical data needs to be transmitted, certificate management issues in public key cryptography systems will seriously affect the operational efficiency of the system [19]. Meanwhile, in most cases, the data owner cannot predict the exact identity of the recipient. Therefore, the use of public key cryptography has certain limitations in solving medical data access and sharing in cloud storage [16]. Reference [8] proposes the concept of Identity Based Encryption (IBE). The proposal of IBE solves the problem of certificate management in public key encryption. In IBE, the public key of users is actually related to their social identity information, such as their citizen ID card number or telephone number. The encryptor can use a public key created with the recipient's social identity to encrypt plaintext. Reference [5] replaces personal identity information in IBE with a set of user attributes, dividing ABE into two types: Key. PolicyAttribute BasedEncryption (KP-ABE) scheme based on key policy, and Ciphertext. PolicyAttribute BasedEncryption (CP-ABE) scheme based on ciphertext policy. In the KP-ABE scheme, ciphertext is associated with an attribute set, and the user's keys are associated with an access policy. Correspondingly to KP-ABE, in the CP-ABE mechanism, the user's key is associated with an attribute set, and the ciphertext is associated with an access policy. Therefore, the difference between KP-ABE and CP-ABE depends on who determines the access policy in the encryption system [11]. In KP-ABE, the access policy is specified by the key, and the ciphertext specifies the set of attributes. Only when the set of attributes of the ciphertext meets the access policy specified by the key can it be decrypted. In CP-ABE, when a message is encrypted, the user can only decrypt the given ciphertext if the attribute set in the user key meets the access policy specified in the ciphertext [2]. In order to achieve more flexible access control, reference [4] proposed a KP-ABE scheme to support the expression of non-monotonic access policies in keys. However, in KP.ABE, once a user's key is determined, their access policy will also be determined, and subsequent encryptors need to compare the decryption user's access policy with the access policies of all other users in order to select an appropriate set of attributes for the ciphertext. Reference [1] proposed the first CP-ABE scheme. In CP-ABE, as data encryptors can freely define access policies, CP-ABE can achieve more flexible access control.

The difference between biometric recognition and traditional recognition methods lies in the fact that biometric information can reveal individual medical information. At present, there is evidence to suggest that certain diseases are related to certain special types of biological characteristics, which can be used to understand whether individuals have certain diseases or the probability or trend of future illnesses. Measurement of palm shape can identify special types of palm shape caused by certain diseases, such as gout and arthritis; Certain special types of fingerprints are associated with certain chromosomal abnormalities (such as Down syndrome, Turner syndrome, and Klinefelter syndrome) [9]; The changes of retinal microvessels may be related to type 2 diabetes and hypertension, as well as stroke and cardiovascular death; Infrared cameras used to generate biometric templates can also monitor modifications to certain surgical procedures in the body [17]. On the one hand, this information can be used to monitor, predict, or assist in diagnosing diseases, but if used improperly and leaked to third parties (such as employers, insurance companies, etc.), individuals may face discrimination and even stigmatization [6].

In this paper, the Internet of Things and biometric technology are combined together, and with the support of wireless technology, the safe access of medical equipment is realized, and the safety management of equipment data is carried out to improve the stable operation and technical management of medical Internet of Things equipment.

2 THEORETICAL BASIS OF DIAGNOSTIC INFORMATION FUSION BASED ON MEDICAL INTERNET OF THINGS

2.1 Information Foundation of Medical Internet of Things

In this paper, the research on medical diagnosis assistant decision-making based on evidence theory is carried out, and the basic framework of evidence theory based on evidence representation, evidence fusion and evidence decision-making is mainly applied, as shown in Figure 1.

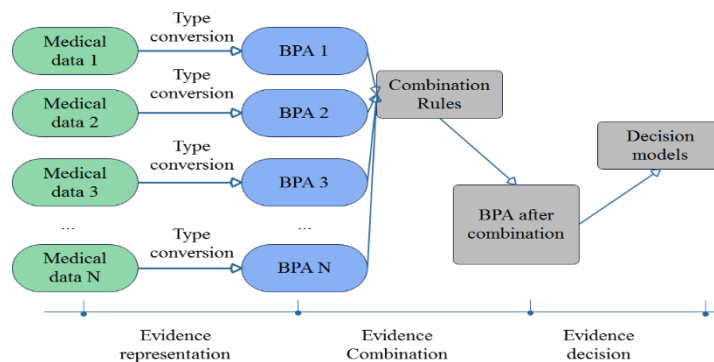


Figure 1: Basic Framework of Evidence Theory Application Under the Background of Medical Diagnosis.

The scope of discussion of evidence theory is expressed as a frame of discernment, and the elements in it are the research objects of evidence theory.

If we assume that a finite set $\Theta = \{\theta_1, \theta_2, \dots, \theta_N\}$ consisting of N elements contains all the propositions or hypotheses to be discussed in the system, and the elements in it are mutually

exclusive, then Θ can be regarded as the frame of discernment of the system. The power set of Θ is expressed as $P(\Theta)$, which consists of 2^N elements and represents the set of all propositions of the system:

$$P(\Theta) = \{\phi, \{\theta_1\}, \{\theta_2\}, \dots, \{\theta_N\}, \{\theta_1 \cup \theta_2\}, \{\theta_1 \cup \theta_3\}, \dots, \Theta\} \quad (1)$$

Basic Probability Assignment (BPA): We set the system frame of discrimination as Θ , and the basic probability assignment is defined as a function $m: P(\Theta) \rightarrow [0, 1]$, satisfying the following conditions:

$$m(\phi) = 0 \quad (2)$$

$$\sum_{A \subseteq \Theta} m(A) = 1 \quad (3)$$

For each $A \in (\Theta)$, $m(A)$ is called the basic probability number, which represents the reliability that completely contributes to proposition A, that is, $m(A)$ only represents the reliability of A, but does not include a subset of A, etc. If $m(A)$ is 0, then proposition A is called focal element of m , and the set of all focal elements in m is called Core of m .

Belief function: When a basic probability number $m(A)$ is given, the reliability function is defined as:

$$Bel(A) = \sum \{m(B) \mid B \subseteq A\} \quad (4)$$

Plausibility function: If a proposition is set as A, its inverse proposition is \bar{A} , which means non-A. According to the definition of reliability function, proposition \bar{A} represents the reliability of negative proposition A, and the plausibility function of proposition A is defined as:

$$Pl(A) = 1 - Bel(\bar{A}) \quad (5)$$

The plausibility function is further expressed as:

$$Pl(A) = \sum \{m(B) \mid A \cap B \neq \phi\} \quad (6)$$

Under the same frame of discernment Θ , two independent basic probability assignments are expressed as m_1 and m_2 , and their joint basic probability assignment $m_{1,2}$ is obtained by using Dempster fusion rule:

$$m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K} \quad (7)$$

Among them,

$$K = \sum_{B \cap C = \phi} m_1(B)m_2(C) \quad (8)$$

The above is a combination formula between two independent evidences. Considering that there are usually multiple evidences in real decision-making systems, the following fusion formula between n evidences is given:

$$m(A) = (1-K)^{-1} \sum_{\cap A_j = A} \prod_{i=1}^n m_i(A_j) \quad (9)$$

$$m(\phi) = 0 \quad (10)$$

Among them,

$$K = \sum_{\cap A_j = \phi} \prod_{i=1}^n m_i(A_j) \quad (11)$$

Dempster combinatorial rule has many excellent mathematical properties, which is the reason why it has been widely concerned. In the following, several important mathematical properties are introduced:

1. Commutative law: The sequence of evidence does not affect the fusion result, and the formula is expressed as $m_1 \oplus m_2 = m_2 \oplus m_1$, which is the orthogonal sum.
2. Associative law: Changing the fusion order of evidence does not affect the final fusion result, and the formula is expressed as $(m_1 \oplus m_2) \oplus m_3 = m_1 \oplus (m_2 \oplus m_3)$.
3. Focus: Reliability can focus on a proposition through the fusion of evidence.

If the basic probability assignment under frame of discernment Θ is m , and α represents the discounting factor, which satisfies $\alpha \in [0,1]$, the discounted evidence body m^α is defined as:

$$m^\alpha(\Theta) = \alpha \cdot m(\Theta) + (1-\alpha) \quad (12)$$

$$m^\alpha(A) = \alpha \cdot m(A), \forall A \subseteq \Theta, A \neq \Theta \quad (13)$$

If the basic probability on the frame of discernment Θ is assigned to m , the gambling probability transformation $BetaP_m : \Theta \rightarrow [0,1]$ of m is defined as:

$$betP_m(w) = \sum_{A \subseteq \Theta, w \in A} \frac{1}{|A|} \frac{m(A)}{1-m(\phi)} \quad (14)$$

Because of $m(\phi)=0$, the above definition can be abbreviated as:

$$setP_m(w) = \sum_{A \subseteq \Theta, w \in A} \frac{m(A)}{|A|} \quad (15)$$

The main function of gambling probability transformation is to transform basic probability assignment into probability distribution, so as to facilitate the decision-making process in medical diagnosis.

The above completes the introduction of the relevant knowledge in the application framework of evidence theory. In view of the research needs of this paper, the following supplements are made to another relevant knowledge.

If the two evidence bodies defined on the frame of discernment Θ are m_1 and m_2 , the evidence distance between them is defined as:

$$d(m_1, m_2) = \sqrt{\frac{1}{2} \cdot (\overline{m_1} - \overline{m_2})^T \underline{D} (\overline{m_1} - \overline{m_2})} \quad (16)$$

Among them, $\overline{m_1}$ and $\overline{m_2}$ are vector forms of evidence bodies m_1 and m_2 , and \underline{D} is a $2^\Theta \times 2^\Theta$ -dimensional matrix where the elements are:

$$\underline{D}(A, B) = \frac{A \cap B}{A \cup B} \quad (17)$$

It is proved that the evidence distance satisfies all distance axioms and satisfies $d \in [0, 1]$. In particular, if propositions A and B have no common elements, then $|A \cap B| = 0$. The reason is that A and B are highly conflicted at this time.

2.2 Uncertainty of Medical Diagnosis Information-Reliability Entropy

If an evidence body on the frame of discernment Θ set by the reliability entropy is m , its reliability entropy is defined as:

$$E_d = - \sum_{A \subseteq \Theta} m(A) \log \frac{m(A)}{2^{|A|} - 1} \quad (18)$$

Among them, A is a focal element in m and $|A|$ is the potential of A. The reason is that when the basic probability assignment degenerates into probability distribution, the isentropy degenerates into Shannon entropy accordingly.

By analyzing Formula 18, it can be found that isentropy is actually a combined measurement method, which can be expressed in the following form:

$$E_d = \sum_{A \subseteq \Theta} m(A) \log(2^{|A|} - 1) - \sum_{A \subseteq \Theta} m(A) \log m(A) \quad (19)$$

For the convenience of expression, the description of output information of single classifier is given:

For a classification problem, all classes form a set $c = \{c_1, c_2, \dots, c_m\}$, each sample corresponds to n attributes, which is represented as $X = \{x_1, x_2, \dots, x_n\}$, all classifiers in the system form a set

$C = \{C_1, C_2, \dots, C_L\}$, and classifier C_j outputs the class label of judging sample X as $C_j(X) \in \{c_1, c_2, \dots, c_m\}$. According to different types of output information of single classifier, the classification results provided by single classifier are divided into the following levels:

1. Abstraction layer: Each single classifier C_j generates only one classification label $C_j(X)$ for a sample X ;
2. Sorting layer: each single classifier C_j generates a classification label sequence for a sample X , and the higher the ranking, the greater the possibility of becoming the final classification result;
3. Metric layer: Each single classifier C_j generates a classification label set for a sample X , which identifies the degree to which the sample X belongs to various labels.

According to this classification information, the fusion types of multi-classifier system can be divided into cascade and parallel from the structural point of view. For cascade mode, the input of the next classifier is the output of the previous classifier, while for parallel mode, each classifier is independent of each other. Schematic diagrams of the two fusion structures are given in figure 2. As can be seen from the cascade mode on the left side, the output of the upper classifier serves as the input of the next classifier, thus guiding the classification. For this structure, there are two main fusion methods: redefinition method and set reduction method. Because this paper mainly focuses on parallel fusion mode, these two methods will not be introduced too much.

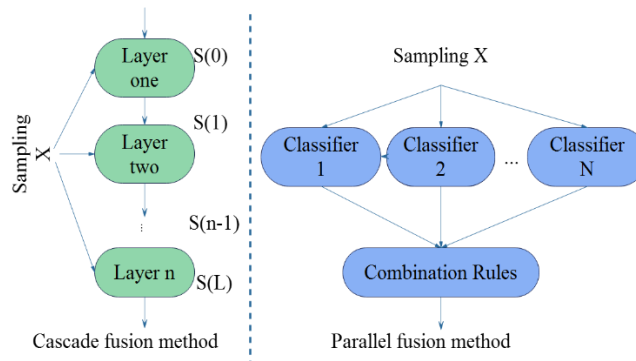


Figure 2: Cascade Classifier Fusion and Parallel Classifier Fusion.

Average method (AF): If the set composed of all classifiers in the system is $C = \{C_1, C_2, \dots, C_L\}$ and P_i represents the output result of classifier C_i , the average fusion method is defined as:

$$P = \frac{1}{L} \sum_{n=1}^L P_i \quad (20)$$

Weighted average method (WAF): If the set composed of all classifiers in the system is $C = \{C_1, C_2, \dots, C_L\}$ and P_i represents the output result of classifier C_i , the weighted average fusion method is defined as:

$$P = \sum_{n=1}^L w_n P_n \quad (21)$$

Among them, w_i is the weight of the i -th classifier.

Voting method (MV): If the set composed of all classifiers in the system is $C = \{C_1, C_2, \dots, C_L\}$, and I_i represents the decision value of the output result P_i of classifier C_i , the voting method is defined as:

$$I = \frac{1}{L} \sum_{n=1}^L I_n \quad (22)$$

Weighted voting method (WMV): If the set composed of all classifiers in the system is $C = \{C_1, C_2, \dots, C_L\}$ and I_i represents the decision value of the output result P_i of classifier C_i , the weighted voting method is defined as:

$$I = \sum_{n=1}^L w_n I_n \quad (23)$$

Among them, w_i is the weight of the i -th classifier.

2.3 Standardized Management of Diagnostic Information Based on Medical Internet of Things

We set the original data set as $X_{ij}, i = \{1, 2, \dots, n\}$, n as the number of samples, $j = \{1, 2, \dots, m\}$, and m as the number of attributes. For any two samples x_i and x_e , the gravitation between them is defined as:

$$F_{ie} = \frac{M_i \times M_e}{d_{ie}^2} \quad (24)$$

Among them, M_i represents the mass of the sample x_i and d_{ie} represents the distance between the samples x_i and x_e . The specific calculation method will be described below.

For a sample, its quality is defined as the sum of the number of samples with the same attribute value under different dimensions, and the mathematical form is defined as:

$$M_i = \sum_{j=1}^m \sum_{k=1}^n \text{count}(x_{kj}, x_{ij}) \quad (25)$$

Among them, the count (x, y) function means that if $x=y$, the count is added once. The distance d_{ie} between samples x_i and x_e is defined as:

$$d_{ie} = \sum_{j=1}^m d(x_{ij}, x_{ej}) \quad (26)$$

Among them, $d(x_{ij}, x_{ej})$ represents the distance between the attribute values x_{ij} and x_{ej} . Based on the above definition, the overall gravity between sample x_i and other samples can be obtained as follows:

$$F_i = \sum_{k=1}^n F_{ik} \quad (27)$$

The idea of outlier detection method based on gravity is as follows: Firstly, the method sets a parameter γ , and if the gravitational force between sample x_i and other samples is less than parameter γ , x_i is judged as an outlier. The flow chart of the algorithm is given below (Figure 3).

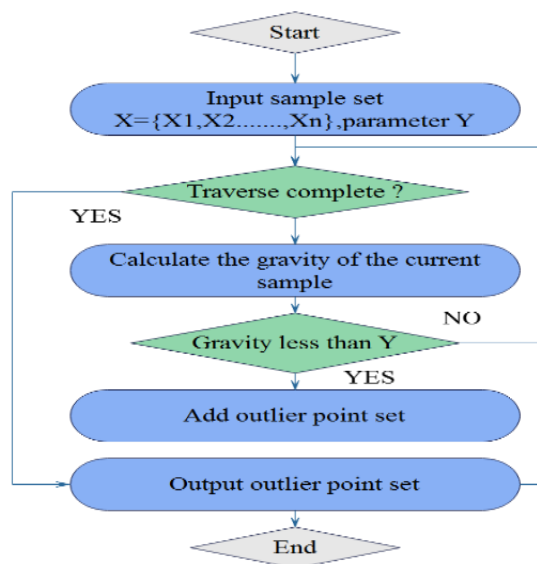


Figure 3: Flow Chart of Outlier Detection Method Based on Gravity.

For a given original dataset X_{ij} , $i = \{1, 2, \dots, n\}$ represents the number of samples n , $j = \{1, 2, \dots, m\}$ represents the number of attributes m , and the number of missing attributes of sample S_i is N_i^A , then the attribute value missing rate of the sample S_i is defined as:

$$R_i^A = \frac{N_i^A}{m} \quad (28)$$

For a given raw dataset X_{ij} , $i = \{1, 2, \dots, n\}$ represents the number of samples m , and $j = \{1, 2, \dots, m\}$ represents the number of attributes m . If the number of samples with missing information under attribute A_j is N_j^S , the sample missing rate of attribute A_j is defined as:

$$R_j^S = \frac{N_j^S}{n} \quad (29)$$

The missing rate of attribute value defines the missing degree of each sample attribute value, and the missing rate of sample defines the missing degree of samples under different attributes. After defining the missing rate of attribute value and missing rate of sample, the missing threshold of attribute value σ_A and missing rate of sample σ_s are used to express the degree that the algorithm users can tolerate attribute value missing and sample missing subjectively, respectively. When the missing rate of attribute value and missing rate of sample are higher than the corresponding thresholds σ_A and σ_s , the corresponding samples or attributes will be deleted to ensure the validity of the data. Based on the above description, the rules of the deletion method are expressed as follows:

$$Y_{ij} = \begin{cases} \neg X_{i*}, R_i^A > \sigma_A \\ \neg X_{*j}, R_j^S > \sigma_s \end{cases} \quad (30)$$

For a given raw dataset X_{ij} , $i = \{1, 2, \dots, n\}$ represents n samples and $j = \{1, 2, \dots, m\}$ represents m attributes. If the set of attribute values not missing under attribute A_j is $\tilde{X}_{*j} = \{X_{1j}, X_{2j}, \dots, X_{lj}\}$, the missing attribute value \bar{X}_{*j} is defined as:

$$\bar{X}_{*j} = \frac{1}{l} \sum_{i=1}^l X_{ij} \quad (31)$$

For a given original dataset X_{ij} , $i = \{1, 2, \dots, n\}$ represents the number of samples n , $j = \{1, 2, \dots, m\}$ represents the number of attributes m , and the missing attribute value of sample S_i under attribute A_j represent \bar{X}_{*j} . If the attribute A_j is not considered, a new dataset $(X_{ij})_{n \times (m-1)}$ is obtained. It is worth noting that this dataset does not contain the default value of the attribute. We

choose the appropriate distance measure to calculate K samples closest to the sample S_i , which are represented as $\{S_1, S_2, \dots, S_K\}$, and the attribute value of each sample under the attribute A_j is represented as $\{X_{1j}, X_{2j}, \dots, X_{Kj}\}$. According to common sense, the influence of the selected K samples on sample S_i will vary with different distances, so the weight of the selected K samples is defined as:

$$\zeta_k = e^{-\gamma \cdot d_k} \quad (32)$$

$$d_k = \frac{d(S_i, S_k)}{\min_{k \in [1, K]} d(S_i, S_k)} \quad (33)$$

The distance weight is normalized below:

$$\omega_k = \frac{\zeta_k}{\sum_{k=1}^K \zeta_k} \quad (34)$$

The weights of K samples are obtained, and the following weighted average operation is carried out on K attribute values to obtain the missing attribute values of sample S_i under attribute A_j :

$$\vec{X}_{ij} = \sum_{k=1}^K \omega_k X_{kj} \quad (35)$$

We randomly select a sample S_i under a certain attribute A_j , and the true attribute value of the sample is X_{ij} . At the same time, we assume that the attribute value is the default value, and use the average method and the nearest neighbor method to evaluate the missing attribute value, and the evaluation results are expressed as X_{ij}^{mean} and X_{ij}^{knn} .

The real attribute value of sample S_i under attribute A_j is X_{ij} , and the evaluation value of this attribute value is expressed as \vec{X}_{ij} , then the proximity ψ of the evaluation result is defined as:

$$\psi = \frac{|\vec{X}_{ij} - X_{ij}|}{|X_{ij}|} \quad (36)$$

The molecular part represents the difference between the estimated value and the true value. It can be seen that the smaller the difference value, the smaller the proximity ψ , indicating the better the evaluation effect, and vice versa.

After defining the proximity index of the evaluation results, we can calculate the proximity index of the evaluation results obtained by the average method and the nearest neighbor method:

$\psi_{mean} = |X_{ij}^{mean} - X_{ij}| / |X_{ij}|$ and $\psi_{kn} = |X_{ij}^{kn} - X_{ij}| / |X_{ij}|$. In this experiment, five attributes are randomly selected for the experiment, and 100 samples are randomly selected under each attribute.

R-type clustering method is widely used in attribute reduction, especially in medical diagnosis. Therefore, this paper proposes a new attribute reduction method based on R-type clustering. The algorithm is described in detail below.

(1) Data preprocessing

Attribute standardization: In order to eliminate the influence of dimensions on clustering effect and improve the convergence speed and accuracy of clustering, the following formula is used to standardize the original data set:

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (37)$$

Among them, $i = \{1, 2, \dots, n\}$ represents samples, $j = \{1, 2, \dots, m\}$ represents attributes, and the mean value \bar{x}_j of x_j is calculated by the following formula:

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (38)$$

The standard deviation s_j of x_j is calculated by the following formula:

$$s_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2} \quad (39)$$

2. The coefficient of variation can be used to measure the discrete degree of data. The larger the coefficient of variation, the higher the discrete degree of data, the stronger the ability to distinguish information, and vice versa. The coefficient of variation of attribute A_j is defined as:

$$v_j = \frac{s_j}{\bar{x}_j} \quad (40)$$

The coefficient of variation threshold is σ_j . If $v_j = \sigma_j$, the attribute A_j is eliminated, otherwise, it is retained.

(2) R-type clustering

The normalized dataset is represented as x'_{ij} , where $i = \{1, 2, \dots, n\}$ represents a total of n samples and $j = \{1, 2, \dots, m\}$ represents a total of m attributes. The correlation coefficient between attributes A_j and A_k is defined as:

$$R_{jk} = \frac{\sum_{l=1}^m (x'_{jl} - \bar{x}_{*j})(x'_{kl} - \bar{x}_{*k})}{\sqrt{\sum_{l=1}^m (x'_{jl} - \bar{x}_{*j})^2 \sum_{l=1}^m (x'_{kl} - \bar{x}_{*k})^2}} \quad (41)$$

Among them, $\bar{x}_{*j} = \frac{1}{m} \sum_{l=1}^m x'_{jl}$, $\bar{x}_{*k} = \frac{1}{m} \sum_{l=1}^m x'_{kl}$. The correlation coefficient matrix is obtained by calculating the correlation coefficient between two attributes, which is represented as $C_0 = (R_{jk})_{m \times m}$. We choose the largest correlation coefficient in C_0 and set it as R_{pq} , which means that the correlation between A_p and A_q is the greatest at this time, then we aggregate them into one class, and recalculate the correlation coefficient matrix between the new class and other attributes to get $C_1 = (R_{jk})_{(m-1) \times (m-1)}$. The above steps are repeated until all attributes are clustered into one class. To solve this problem, this chapter puts forward the weighted average coefficient method.

The weighted average coefficient method assumes that the maximum value of non-diagonal elements in matrix C_* is R_{pq} , and the new attribute after aggregation of attribute A_p and A_q is recorded as A_r , then the correlation coefficient between A_r and other attributes in matrix C_{*+1} is defined as:

$$R_{jr} = \frac{1}{n_p + n_q} (n_p R_{jp} + n_q R_{jq}) \quad (42)$$

Among them, n_p and n_q are the number of attributes in attribute (set) A_p and A_q , respectively. The following is an example to demonstrate the use process of weighted average coefficient method.

According to formula 20, the correlation coefficient between attribute A_{14} and other attributes in matrix C_1 is calculated:

$$R_{1.14} = \frac{1}{n_6 + n_7} (n_6 R_{16} + n_7 R_{17}) = \frac{1}{2} \times (0.3270 + 0.2713) = 0.2992 \quad (43)$$

This section selects a confidence level of 0.3, as shown in Figure 4, so the attributes can be divided into five categories, $(A_6, A_7, A_9, A_{11}, A_{12})$, (A_2, A_8) , (A_1, A_{10}, A_3) , (A_5) and (A_3, A_4) . Next, the

representative elements in each cluster will be extracted according to the clustering results as the result of the final attribute reduction.

In the same way, the correlation coefficient between A_{14} and other attributes can be obtained, and the matrix C_1 can be obtained. This operation is continued until all attributes are clustered into one class. Figure 5 shows the result of attribute R clustering of heart disease data set.

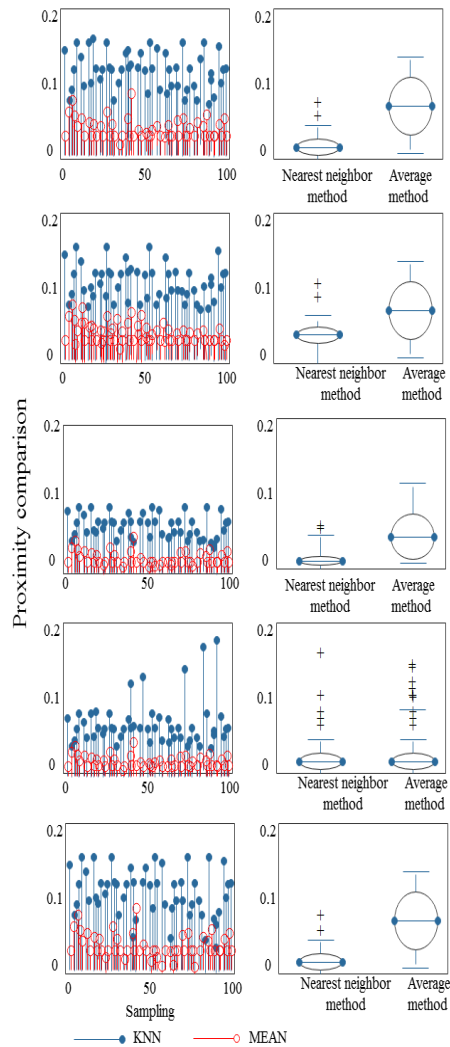


Figure 4: The Filling Effect of Attribute Default Values Based on Average Method and Nearest Neighbor Method. The Left Side Shows the Proximity Comparison of the Two Filling Results, and the Right Side Shows the Statistical Results of Box Chart.

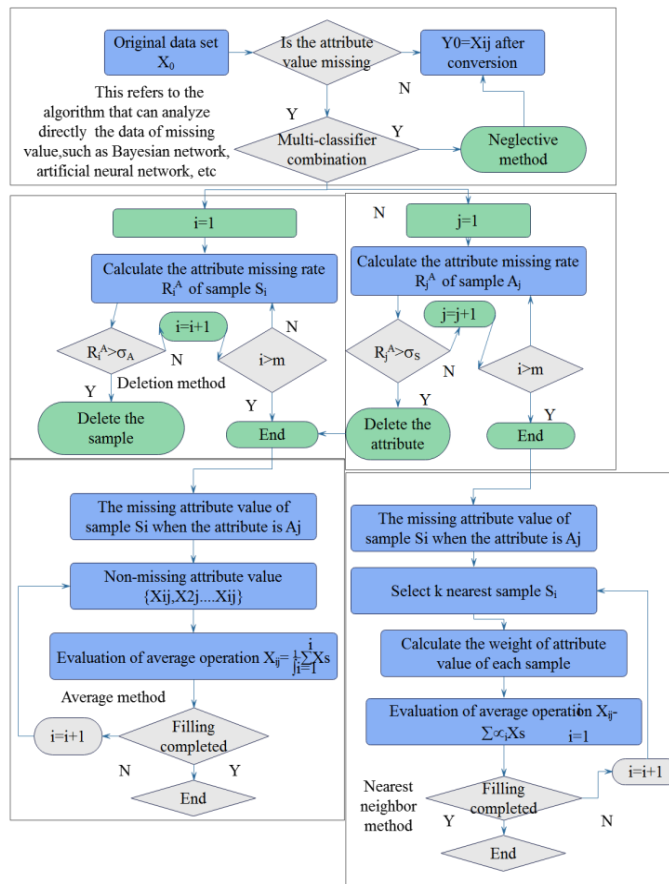


Figure 5: Processing Flow of Default Value of Medical Diagnosis Data.

(3) Extracting representative element

Taking cluster $(A_6, A_7, A_9, A_{11}, A_{12})$ as an example, the complex correlation coefficients between each element in the class and other combined elements are calculated respectively, and $R^2_{A_6, (A_7, A_9, A_{11}, A_{12})}^c$ and $R^2_{A_7 (A_6, A_9, A_{11}, A_{12})}$, $R^2_{A_9 (A_6, A_7, A_{11}, A_{12})}$, $R^2_{A_{11} (A_6, A_7, A_9, A_{12})}$ are obtained. Therefore, the attribute A_6 with the largest complex correlation coefficient is selected as the representative element, because it can represent the whole cluster to the greatest extent. In the same way, the representatives of several other clusters are A_6, A_2, A_{10}, A_5 and A_3 , respectively. The R-clustering renderings of attributes of the heart disease data set are shown in Figure 6.

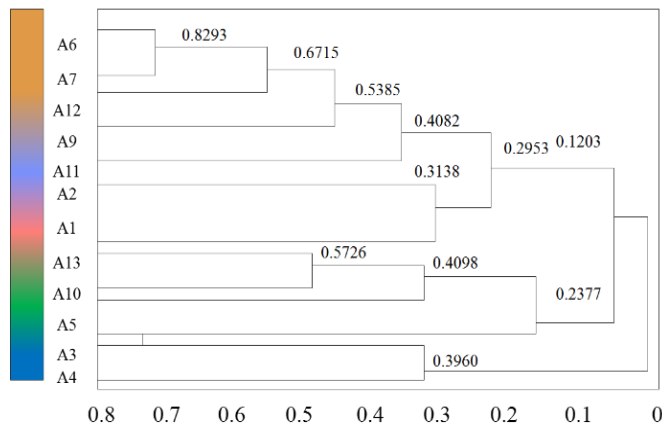
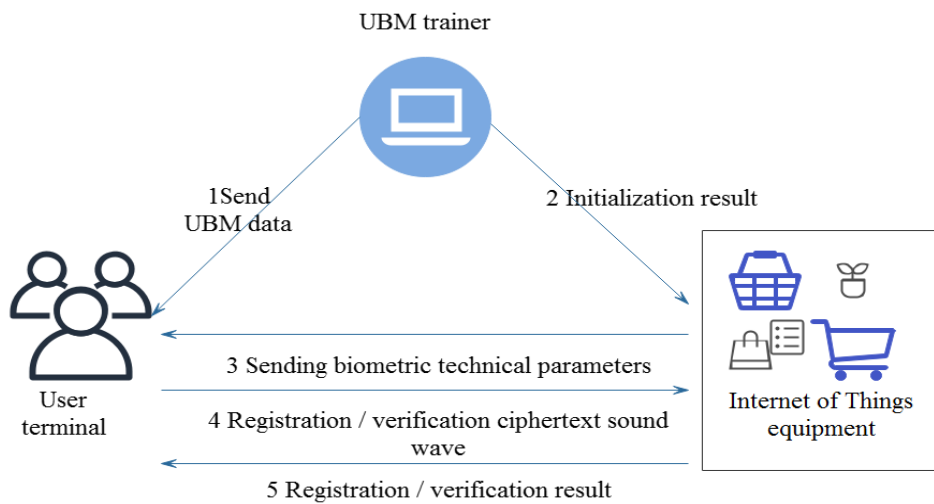


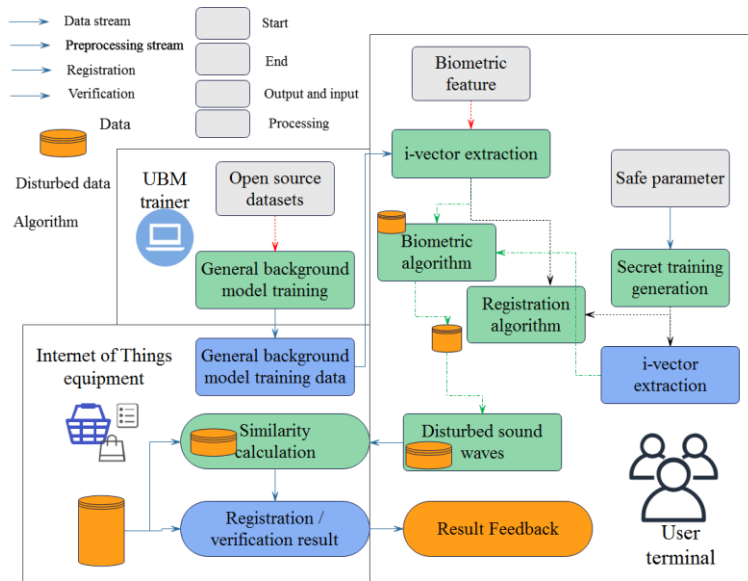
Figure 6: R-Clustering Renderings of Attributes of Heart Disease Data Sets.

3 SECURE ACCESS AND MANAGEMENT OF WIRELESS MEDICAL DEVICES USING INTERNET OF THINGS AND BIOMETRIC TECHNOLOGY

In this system, we focus on protecting users' privacy, and abstract the system model from the intelligent medical Internet of Things scene. Specifically, there are three main entities in the system model, namely IoT device, client and UBM trainer, as shown in Figure 7 (a). The overall flow is summarized in Figure 7 (b).



(a) Medical Device Management System Model using Internet of Things and Biometric Technology



(b) Overview of the overall process

Figure 7: System Model.

When the parameter is updated, the previous gradient will be retained and the current gradient will be added. The schematic diagram is shown in Figure 8. The whole surface in the graph represents the loss function. The loss function in the graph is convex, but in fact the loss function may have multiple local minimum points.

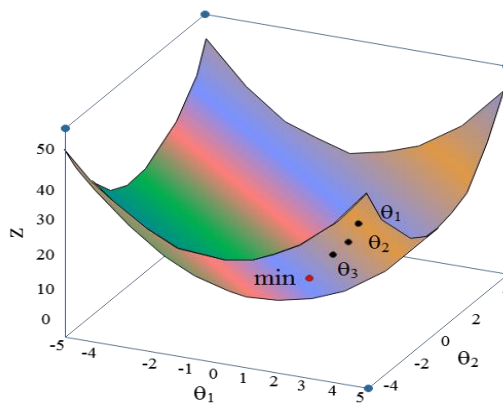


Figure 8: Schematic Diagram of Signal Gradient Processing.

Due to the limitation of wireless channel quality, hardware design or standard, high-order modulation may not be suitable in some situations. Therefore, in order to make the transceiver adapt to different channel conditions and modulation methods, this paper proposes a symbol segmentation and

compression method, and uses the algorithm in the second part to perform experimental analysis. When generating 6 symbol streams, the symbols in each symbol stream will have different values before modulation (as shown in Figure 9) and after IFFT (as shown in Figure 10). Therefore, the threshold value of each symbol stream can be set separately so as to increase the compression ratio as much as possible while satisfying the maximum allowable distortion requirement.

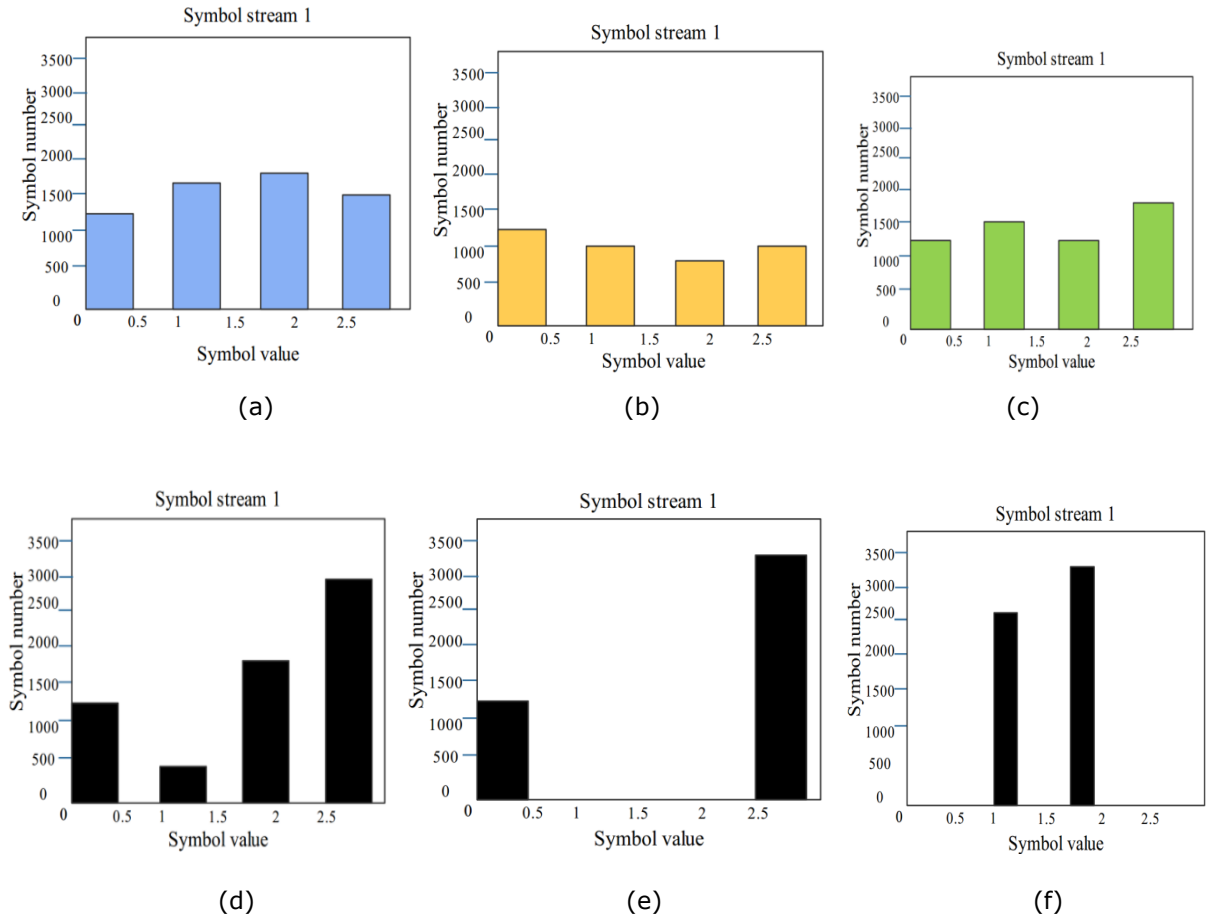
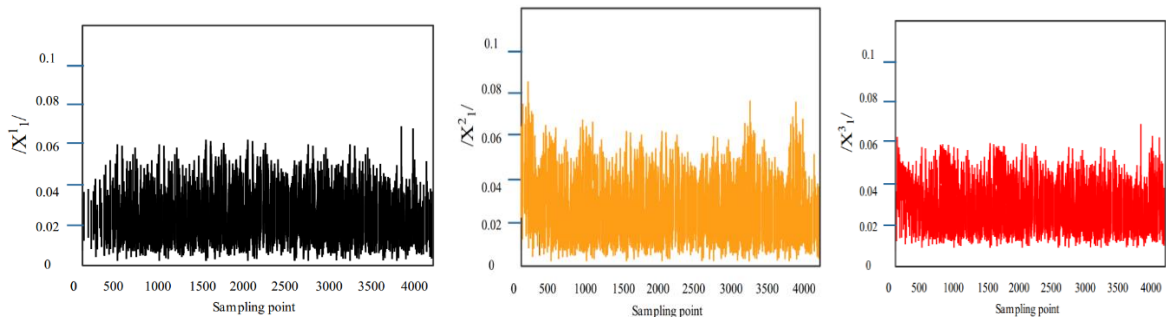


Figure 9: Symbol Stream Generated Before Modulation.



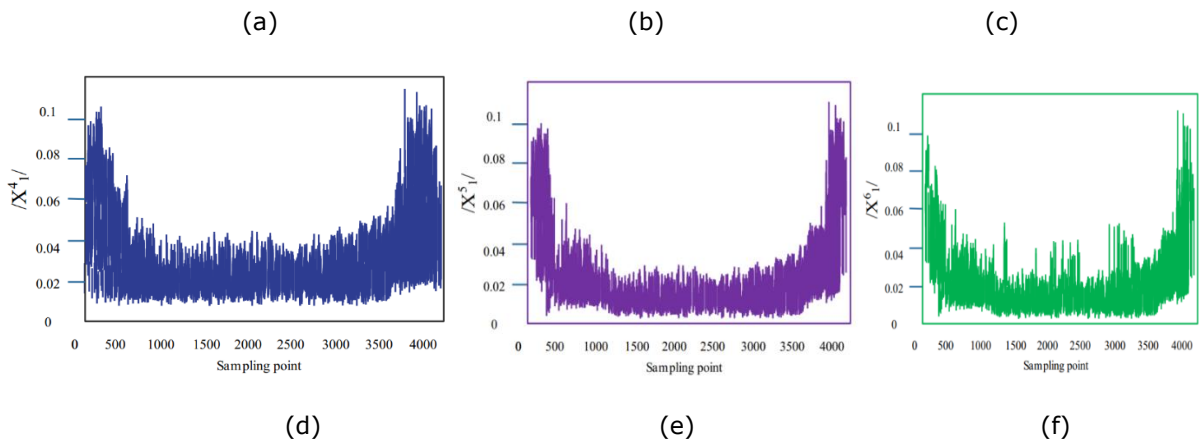


Figure 10: Generation of Symbol Stream After IFFT During Modulation.

From the above analysis, it can be seen that the proposed method of using Internet of Things and biometric technology to achieve secure access and management of wireless medical devices can not only improve the management efficiency of medical devices, but also effectively improve the security of medical devices.

4 CONCLUSION

Medical Internet of Things applies the concepts, tools and technologies of Internet of Things to the medical care industry, uses big data, edge computing and other technologies to realize the intelligence and personalization of medical processes, and establishes all feasible network services to connect available medical resources with various medical services. Moreover, the medical Internet of Things makes hospital management safer and more convenient by tracking and monitoring drug circulation and managing inventory. In addition, medical equipment is not only advanced, but also its data has certain value, so it is necessary to ensure the security of wireless medical equipment. This paper combines the Internet of Things with biometric technology, and realizes the safe access of medical equipment with the support of wireless technology, and manages the equipment data safely, so as to improve the stable operation and technical management of medical Internet of Things equipment. After building the model, it can be seen that the proposed method of using Internet of Things and biometric technology to achieve secure access and management of wireless medical devices can not only improve the management efficiency of medical devices, but also effectively improve the security of medical devices.

Haifeng Chen, <https://orcid.org/0009-0007-6432-701X>

ACKNOWLEDGE

Haifeng Chen was born in Shanghai, P.R. China, in 1970. He obtained a master's degree from Nanjing University of Technology, P.R. China. His research interests include the Internet of Things, embedded systems, and artificial intelligence.

REFERENCES

- [1] Anggriawan, R.; Salim, A. A.; Gunawan, Y.; Arumbinang, M. H.: Passenger Name Record Data Protection under European Union and United States Agreement: Security over Privacy?, *Hasanuddin Law Review*, 8(2), 2022, 95-110. <https://doi.org/10.20956/halrev.v8i2.2844>
- [2] Ansari, M. T. J.; Agrawal, A.; Khan, R. A.: DURASec: Durable Security Blueprints for Web-Applications Empowering Digital India Initiative, EAI Endorsed Transactions on Scalable Information Systems, 9(4), 2022, e7-e7.
- [3] Bakhtina, M.; Matulevicius, R.: Information Security Risks Analysis and Assessment in the Passenger-Autonomous Vehicle Interaction, *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, 13(1), 2022, 87-111.
- [4] Cao, L.: Decentralized ai: Edge Intelligence and Smart Blockchain, Metaverse, web3, and descI, *IEEE Intelligent Systems*, 37(3), 2022, 6-19. <https://doi.org/10.1109/MIS.2022.3181504>
- [5] Chen, Z.; Liu, J.; Shen, Y.; Simsek, M.; Kantarci, B.; Mouftah, H. T.; Djukic, P.: Machine Learning-Enabled Iot Security: Open Issues and Challenges Under Advanced Persistent Threats, *ACM Computing Surveys*, 55(5), 2022, 1-37. <https://doi.org/10.1145/3530812>
- [6] Ghelani, D.; Hua, T. K.; Koduru, S. K. R.: A Model-Driven Approach for Online Banking Application Using AngularJS Framework, *American Journal of Information Science and Technology*, 6(3), 2022, 52-63.
- [7] Indrasari, A.; Nadjmie, N.; Endri, E.: Determinants of Satisfaction and Loyalty of E-Banking Users During the COVID-19 Pandemic, *International Journal of Data and Network Science*, 6(2), 2022, 497-508. <https://doi.org/10.5267/j.ijdns.2021.12.004>
- [8] Jain, A. K.; Gupta, B. B.: A Survey of Phishing Attack Techniques, Defence Mechanisms and Open Research Challenges, *Enterprise Information Systems*, 16(4), 2022, 527-565. <https://doi.org/10.1080/17517575.2021.1896786>
- [9] Jia, H.; Teng, Y.; Li, N.; Li, D.; Dong, Y.; Zhang, D.; Qin, W.: Dual Stimuli-Responsive Inks Based on Orthogonal Upconversion Three-Primary-Color Luminescence for Advanced Anticounterfeiting Applications, *ACS Materials Letters*, 4(7), 2022, 1306-1313. <https://doi.org/10.1021/acsmaterialslett.2c00328>
- [10] Kuo, T. T.; Jiang, X.; Tang, H.; Wang, X.; Harmanci, A.; Kim, M.; Ohno-Machado, L.: The Evolving Privacy and Security Concerns for Genomic Data Analysis and Sharing as Observed from the Idash Competition, *Journal of the American Medical Informatics Association*, 29(12), 2022, 2182-2190. <https://doi.org/10.1093/jamia/ocac165>
- [11] Nikkhah, H. R.; Sabherwal, R.: Information Disclosure Willingness and Mobile Cloud Computing Collaboration Apps: The Impact of Security and Assurance Mechanisms, *Information Technology & People*, 35(7), 2022, 1855-1883. <https://doi.org/10.1108/ITP-12-2019-0630>
- [12] Ouda, A. J.; Yousif, A. N.; Hasan, A. S.; Ibrahim, H. M.; Shyaa, M. A.: The Impact of Cloud Computing on Network Security and the Risk for Organization Behaviors, *Webology*, 19(1), 2022, 195-206. <https://doi.org/10.14704/WEB/V19I1/WEB19015>
- [13] Sahu, A. K.; Gutub, A.: Improving Grayscale Steganography to Protect Personal Information Disclosure within Hotel Services, *Multimedia Tools and Applications*, 81(21), 2022, 30663-30683. <https://doi.org/10.1007/s11042-022-13015-7>
- [14] Srivatanakul, T.; Annansingh, F.: Incorporating Active Learning Activities to the Design and Development of an Undergraduate Software and Web Security Course, *Journal of Computers in Education*, 9(1), 2022, 25-50. <https://doi.org/10.1007/s40692-021-00194-9>
- [15] Sudarwanto, A. S.; Kharisma, D. B. B.: Comparative study of Personal Data Protection Regulations in Indonesia, Hong Kong and Malaysia, *Journal of Financial Crime*, 29(4), 2022, 1443-1457. <https://doi.org/10.1108/JFC-09-2021-0193>
- [16] Sun, H.; Samad, S.; Rehman, S. U.; Usman, M.: Clean and Green: the Relevance of Hotels' Website Quality and Environmental Management Initiatives for Green Customer Loyalty, *British Food Journal*, 124(12), 2022, 4266-4285. <https://doi.org/10.1108/BFJ-09-2021-1002>

- [17] Tahaei, M.; Li, T.; Vaniea, K.: Understanding Privacy-Related Advice on Stack Overflow. Proceedings on Privacy Enhancing Technologies, 2022(2), 2022, 114-131. <https://doi.org/10.2478/popets-2022-0038>
- [18] Torres-Hernández, N.; Gallego-Arrufat, M. J.: Indicators to Assess Preservice Teachers' Digital Competence in Security: A Systematic Review, Education and Information Technologies, 27(6), 2022, 8583-8602. <https://doi.org/10.1007/s10639-022-10978-w>
- [19] Zhao, H.; Liu, Z.; Yao, X.; Yang, Q.: A Machine Learning-Based Sentiment Analysis of Online Product Reviews with a Novel Term Weighting and Feature Selection Approach, Information Processing & Management, 58(5), 2021, 102656. <https://doi.org/10.1016/j.ipm.2021.102656>