Novel Data Fusion Scheme of WBAN for Medical Monitoring

Jian Yan[1], Qingxu Deng[1], Guangjie Han[2], Yuhan Lin[1]

[1] School of Computer Science and Engineering, Northeastern University, China
[2] Department of Information and Communication System, Hohai University, China
jimyancn@qq.com, dengqx@mail.neu.edu.cn, hanguangjie@gmail.com, linyuhancs@gmail.com

Abstract

The aging of the global population and increasing incidence of chronic diseases have exacerbated social problems such as the shortage of medical resources and increased medical costs. There is a growing interest in medical monitoring services based on wireless body area network (WBAN). In WBAN, due to limitations on processing capacity, battery, and storage capacity of the sensors, it is difficult to guarantee the low-latency, high-reliability requirements, and better user experience of medical monitoring services. Therefore, a novel data fusion method of WBAN for medical monitoring services is proposed in this study. According to the redundancy and complementarity of WBAN data collection in time or space, the proposed method realizes the fusion of single source and multisource data in time and space, which obtains a consistent interpretation and description of the measured object, as well as more effective collection results than a single sensor. Furthermore, to meet the requirements of real-time of medical monitoring services, a hierarchical data fusion model based on edge computing is proposed. In this model, to realize the load balance of the entire system and maximize system utility, different types of data fusion tasks are scheduled for execution on the sensor, sink node, or edge node. The simulation results show that the proposed method effectively improves the accuracy and reliability of WBAN data collection when the algorithm execution time is acceptable and meets the real-time requirements of medical monitoring services.

Keywords: WBAN, Data fusion, Edge computing, Medical monitoring

1 Introduction

Currently, hundreds of millions of people worldwide are suffering from chronic diseases. In China alone, the number of people who die from chronic diseases such as cerebrovascular diseases, asthma, obesity, and diabetes each year has reached one million [1]. In addition, problems such as shortage of medical resources, lack of medical staff, increased medical costs, and uneven medical allocation have become a major bottleneck to the rapid development of the health industry [2]. As a result, the wireless body area network (WBAN) [3] technology for medical services has been introduced. It is a human-centered short-distance, highly reliable, and low-power wireless communication network technology that consists of a central control node and multiple sensor nodes connected to each other on the surface, in the body, or near the human body. The sensor collects the user's physiological (heartbeat, body temperature, blood pressure, ECG signal, etc.), behavioral (such as running frequency, distance, duration etc.), and other health-related information and sends it to the central node or to the remote monitoring center through the external network to provide users with timely medical services [4].

WBAN has become the core technology in the medical monitoring field, it is widely used in scenarios such as health monitoring, real-time diagnosis, and remote assistance, effectively alleviating problems such as the lack of high-quality medical resources [5]. WBAN for medical monitoring mainly includes the following key technologies: data acquisition, data transmission, and data fusion technologies. Accurate data collection, reliable data transmission, and real-time data fusion are prerequisites of medical monitoring. In addition, the accuracy, reliability, and real-time nature of the data are crucial for medical monitoring. If some life-threatening emergency data contains errors, it threatens human life and may lead to catastrophic consequences [6]. Currently, this technology mainly faces the following challenges:

(1) Due to limitations of WBAN sensor sophistication and sampling accuracy, the data collected may not always meet the accuracy requirements.

(2) The wireless communication of WBAN will be interfered with by the external environment, mainly through electromagnetic interference and other communication technologies such as ZigBee and WiFi, which will reduce the reliability of data.

*Corresponding Author: Guangjie Han; E-mail: hanguangjie@gmail.com
DOI: 10.53106/160792642021072204015

(3) For portability and comfort, the wearable or implanted sensor devices in WBAN should be as compact as possible. Therefore, the communication, storage, and computing capabilities as well as the battery capacity are limited, making it difficult to run complex data fusion algorithms to improve the real-time performance of collected data.

Data fusion technology [7] is an effective method of solving these problems. The term data fusion here refers to abnormal judgment, data compensation and analysis, and prediction of data from multiple sensors [8]. At present, literature 10-15 propose various data fusion methods based on D-S, maximum likelihood estimation, Kalman filter, support vector machine (SVM), ordered weighted aggregation (OWA), etc. However, these researches mainly focus on two aspects: First, the data fusion methods for WBAN mostly analyze single data, and there is relatively little research on multisource and heterogeneous data analysis of multisensor systems. Second, most of the current fusion algorithms require high-performance hardware or cloud platform support, whereas the processing power of WBAN nodes for medical monitoring is low. If the data is transmitted to the cloud for data fusion, it will cause a large processing delay, which makes it difficult to meet the real-time requirements of medical monitoring services.

The goal of this research is to make full use of the data resources collected from different multiple sensors in time and space. We use computer technology to analyze, synthesize, control, and utilize the observation data obtained in time series under certain criteria to arrive at a consistent interpretation and description of the measured object. Therefore, the corresponding decision making and estimation can be realized, while the system obtains sufficient information from its various components. Edge computing adequately meets the low energy consumption, low latency, and high reliability communication and computing requirements of WBAN [9]. This study proposes a data fusion model based on edge computing, scientifically scheduling WBAN sensor nodes, aggregation nodes and edge nodes and, other computing resources for redundancy and complementary integration. In addition, the proposed method improves data accuracy, reliability, and real-time performance, while reducing network load and system power consumption and improving user experience. The main contributions of this study are as follows:

(1) A single source data fusion algorithm is designed using the time correlation of the data that are collected by the same sensor to perform abnormality judgment, correction, and compensation for the sampled data to improve the accuracy and reliability of the data.

(2) A multisource data fusion algorithm is designed, including similar and heterogeneous multisources, using the spatial redundancy and complementarity of multiple sensor data. The data is then fused by calculating the correlation between the data to obtain more accurate results than those of a single sensor node.

(3) A data fusion model based on edge computing is designed to scientifically schedule data collection, data aggregation, single source data fusion, and multisource data fusion to different locations for execution, which helps realize low-latency and high-performance real-time fusion.

The rest of this paper is arranged as follows: Section 2 reviews the related research on data fusion in WBAN. Section 3 establishes mathematical models of single source and multisource data fusion. Section 4 proposes a data fusion model based on edge computing. Single-source, similar multisource, and heterogeneous multisource data fusion (HMDF) algorithms are realized using this model. Section 5 performs system simulation and data analysis. Section 6 concludes this paper.

2 Related Research

Research on data fusion based on WBAN mainly focuses on improving accuracy, target recognition, analysis and prediction, and data transmission. Literature [10] proposed a cross feature fusion neural network for the enhancement of collaborative filtering by constructing a cross feature fusion network that enables the fusion of user features and item features. Then, the user’s preference for various item features was realized by designing a feature extraction layer with multiple multilayer perceptron (MLP) modules to extract both user features and item features. Compared to existing models, CFFNN has obvious advantages in hit ratio and normalized discounted cumulative gain. In [11], with regard to the shortcomings of using DS combination rules in multisensor data fusion, an improved evidence combination method based on information gain and fuzzy preference relationship was proposed. In addition, by considering both historical and real-time data, more accurate fusion results were obtained. Literature [12] proposed a method of signal conversion using A/D conversion and a microcontroller that realized real-time observation and control of sensor data. It used the maximum likelihood estimation to fuse multiple odometer measurement data to obtain the optimal fusion value. To achieve a low-cost solution that accurately evaluates the displacement under external influence, the literature [13] proposed a multiresponse data fusion using Kalman filtering, which effectively overcomes the expensive use of wired solutions. Literature [14] addressed the problem of traditional fault diagnosis methods not being accurate enough in the data fusion process. It is difficult to distinguish the fault types with nondimensional indicators by performing dimensionless calculations on the original collected data. Literature [14] proposed a data fusion method based on the SVM model, which could more effectively solve the aforementioned problem. Literature [15] proposed a novel method based on the OWA operator through the improvement of the Q function in the OWA. The influence of time intervals on the final fusion result was reduced, and the application of a target recognition based on time series data fusion showed the effectiveness of their method. Literature [16] proposed a new method of multisensor data fusion based on the Bayesian algorithm and compressed sensing (CS), which can realize data fusion and reconstruct sparse signals, the method can identify the damage accurately on the aviation aluminum plate and let the detection error within 0.82 mm. Literature [17] studied the big data analysis application of WBAN in disease prediction through collecting electromyography (EMG), electroencephalogram (EEG) and electrocardiogram (ECG), and other human physiological data and performing data fusion via cloud computing for disease predictions. Literature [18] proposed a method for monitoring biomarkers in athletes' saliva or sweat using wearable devices. It achieves a noninvasive and continuous detection of key indicators such as athletes' stress and physical condition and customized recovery or training programs for each athlete. A study in literature [19] proposed a scalable system using wearable sensor functions through real-time collection of human body temperature, ECG, body position, current geographic location, and other related information. It used decision-level integration to identify the user's posture and health status. Literature [20] focused on biometrics and medical monitoring applications. It proposes a method to support medical diagnosis and treatment services based on multisensor decision fusion systems and using ECG, temperature, accelerometer and other sensors to collect data in real time for fusion analysis. Literature [21] used smartphone accelerometers and gyroscope sensors to collect data on 41 continuous tasks from 44 participants. It calculated 76 signal characteristics, based on three general classifiers (Bayes, SVM, decision tree), and designed a method of categorizing patients into healthy, elderly and stroke patients. Literature [22] was based on Dempster-Shafer theory that fuses the data of two different modal sensors (including depth camera and inertial human sensor) and used the complementarity of sensor data to realize human motion recognition. The results showed that this method could increase the recognition rate by 2% to 23%. In literature [23], a distributed scheduling mechanism for wireless sensor network VD-CSMA is proposed in response to the data fusion problem caused by the various time characteristics of Internet of Things environments. This mechanism considers the data fusion value and delay constraints of the data packet when determining the priority of data fusion. Simulation results show that, compared with the typical solutions, VD-CSMA can improve the throughput and reduce the data transmission delay.

3 System Model

The WBAN topology based on edge computation is shown in Figure 1. Each user carries a number of sensors arranged on the surface of or implanted in their bodies. A sink node forms a star network, and the sink nodes of a group of people form a multihop network that wirelessly accesses the allocated edge nodes. The edge nodes are used to serve multiple neighboring users. One edge node is connected to many sink nodes, and the edge nodes access the Internet through wired or wireless gateways to upload the processed data to the cloud platform. Sensors are responsible for periodically collecting human body data, while sink nodes are responsible for collecting sensor data and carrying out simple data fusion. Simultaneously, edge nodes carry out complex data fusions and upload them to the cloud.

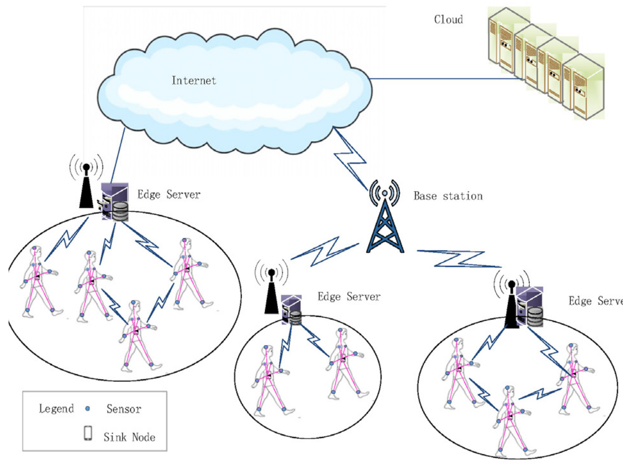

Figure 1. System topological structure

Assuming that a certain edge node is responsible for the data of N users, that is, corresponding to N sink nodes, for any sink node $i(i \in [1, N])$ there are a total of c_i sensor nodes below. For convenience, the following definition is given: the *j-th* sensor of the *i-th* sink node is denoted as $<i, j>$, and the data collected at time t is represented as $d_i^j(t)$ $i(i \in [1, N]), j \in [1, c^i], t \in \mathbb{N}^+)$. If the data collection period of the i-th sink node is T and the data flow from time 1 to time T is $d_i^j(1), d_i^j(2), ..., d_i^j(T)$, then the data of all associated sensors on the sink node is recorded as the data matrix D_i^T, as shown in Formula (1).

$$D_i^T = \begin{bmatrix} d_i^1(1) & d_i^1(2) & ... & d_i^1(T) \\ d_i^2(1) & d_i^2(2) & ... & d_i^2(T) \\ ... & ... & ... & ... \\ d_i^{c_i}(1) & d_i^{c_i}(2) & ... & d_i^{c_i}(T) \end{bmatrix} \quad (1)$$

Due to the different types of data collected by different types of sensors and the different start-up times, the duration it takes to send out data is different. Therefore, at a certain time, some sensors may not send data, resulting in the corresponding $d_i^j(t)$ in D_i^T being empty. As a result, the actual data stream is not a time-continuous matrix. To save storage space, the sink node uses a queue $Q_i = \{d_i^j(t) \mid i \in [1, N], j \in [1, c^i], t \in \mathbb{N}^+\}$ to store the nonempty data collected by all associated sensor nodes.

3.1 Single Source Data Fusion

SSDF uses the data stream collected by a certain sensor node $< i, j >$ within a period of time to correct the measurement result $d_i^j(t)$ at time *t*. Using the correlation of collected data can improve the accuracy of data. Sensors may be affected by abrupt large impulse noise and small-amplitude high-frequency noise. Thus, errors in the data at certain instances is inevitable. Because of the continuous change of human body data over time, the data are strongly correlated with time. Therefore, the accuracy of the collected data can be improved by fusing it with the data collected previously. Here, a single source data fusion method that comprehensively considers both the abrupt large impulse noise and small-amplitude high-frequency noise is proposed. The proposed method is divided into two parts:

For sudden large impulse noise, a method that limits the process is adopted: First, the sampling result should exceed the upper or lower bound data, and if so, the upper or lower bound data are used; Second, the variance amplitude of the measurement results of the two adjacent times cannot exceed the variance threshold. If the difference between the measurement value and the last measurement value is greater than the threshold value, the current value is invalid and gets abandoned. In this case, the last valid value is used instead of the current value to reduce the influence of sudden large impulse noise on the measurement results.

Let the upper limit of the data collected by node $< i, j >$ be $U_{i,j}$, the lower limit be $L_{i,j}$, the threshold value of data variance be $\alpha_{i,j}$. . For $\forall i \in [1, N]$, $j \in [1, c^i], t \in \mathbb{N}^+$, $d_i^j(t)$ is calculated as follows:

$$d_i^j(t) = \begin{cases} U_{i,j}, d_i^j(t) > U_{i,j} \\ L_{i,j}, d_i^j(t) < L_{i,j} \\ d_i^j(t-1), \left| d_i^j(t) - d_i^j(t-1) \right| > \alpha_{ij} \\ d_i^j(t), Otherwise \end{cases} \quad \textbf{(2)}$$

For a high-frequency and small noise, the moving average fusion method is used to save the first θ-1 adjacent measurement data of the current measurement result to the queue *Q*. According to the first-in-first-out principle, every time a new sample is taken, the measurement result is inserted at the end of the queue, so that there are always θ latest data in the queue. The θ data in the queue are arithmetically averaged, and the average is taken as the current time. The measurement result reduces the influence of the high-frequency and small-amplitude noise on the measurement result. Assuming that at time t, the data to be fused is $d_i^j(t-\theta+1), ..., d_i^j(t-2), d_i^j(t-1), d_i^j(t)$, obviously, the larger the value of θ, the more accurate the judgment of the change trend of sensor data. However, a too-large value will increase the processing delay of the data fusion; thus, the value of θ needs to meet the delay constraint. Assuming that the data processing speed ω is constant, the upper limit of the data processing delay is τ for the node $< i, j >$ and the processing delay at time *t* is $\frac{d_i^j(t)}{\omega}$. For the first θ_i^j time, including time *t*, the total processing delay satisfies:

$$\begin{gathered} \sum_{x=t-\theta_i^j+1}^{t} \frac{d_i^j(x)}{\omega} \le \tau \\ \Rightarrow \theta_i^j * \min_{x=1}^{t} \frac{d_i^j(x)}{\omega} \le \tau \Rightarrow \theta_i^j \le \frac{\tau * \omega}{\min_{x=1}^{t} d_i^j(x)} \\ \forall i \in \left[1, N\right], j \in \left[1, c^i\right], t \in \mathbb{N}^+, \theta_i^j \in \mathbb{N}^+ \wedge \theta_i^j < t \end{gathered} \quad \textbf{(3)}$$

Because the sink node uses a cache queue to store the data collected by each sensor node, the data storage capacity cannot exceed the queue length. Let the buffer queue length of sink node *i* be $Len(Q_i)$ and p_j^i be the sampling frequency of node $< i, j >$. If Q_i can store up to $\frac{Len(Q_i)}{p_i^j}$ data at the same time, then:

$$\theta_i^j \le \frac{Len\left(Q_i\right)}{p_i^j} \quad \textbf{(4)}$$

According to Formula (3) and (4), θ can take the value:

$$\theta_i^j = \min\left(\frac{Cnt(Q_i)}{p_i^j}, \frac{\tau * \omega}{\min_{x=1}^{t} d_i^j(x)}\right) \quad \textbf{(5)}$$

Then, the fusion result of node $< i, j >$ at time *t* is

$$d_i^j(t) = \frac{\sum_{x=t-\theta+1}^{t} d_i^j(x)}{\theta_i^j} \quad \textbf{(6)}$$

3.2 Multisource Data Fusion

Data correlation analysis refers to the analysis of two or more correlated variable elements to measure the closeness of two variables, and is used to discover the correlation and strength between different variables. Correlation analysis is a prerequisite for the development of the multi-source data fusion because the accuracy of measurement data can be further improved by fusing relevant data. The various human body parameters collected by sensor nodes reflect the condition of that human body from different perspectives; thus, all data have a strong correlation. The correlation analysis and data fusion based on it have strong practical significance. This research uses the correlation coefficient to measure the correlation between physiological parameters, and the specific processes are as follows:

Node $<i, j>$ is recorded as a random variable X according to the time-varying data flow $d_i^j(t)$ and node $<p, q>$ is recorded as a random variable Y according to the time-varying data flow $d_p^q(t)$, where n is the number of collected data. Then, the correlation coefficient $\rho_{(i,j)(p,q)}$ between nodes $<i, j>$ and $<p, q>$ is defined as:

$$\rho_{(i,j)(p,q)} = \frac{\sum_{t=1}^{n}(X_t - \bar{X})(Y_t - \bar{Y})}{\sqrt{\frac{(X_t - \bar{X})^2}{n}} \times \sqrt{\frac{(Y_t - \bar{Y})^2}{n}}} \quad (7)$$

$$\forall i \in [1, N], \forall j \in [1, c^i], \forall p \in [1, N], \forall q \in [1, c^i]$$

It is observed from the Formula (7) that the range of the correlation coefficient $\rho_{(i,j)(p,q)}$ is [-1,1]. The magnitude indicates the strength of the relationship, and the sign indicates the direction of the correlation. When the nodes are completely uncorrelated, the coefficient is 0, when completely correlated, it is 1. Because the data of different correlations have different weights when fused, to facilitate the algorithm design, we divide the correlation coefficient into three levels: strong, medium, and weak, denoted as $\rho 1$, $\rho 2$ and $\rho 3$ respectively. $\rho 3$ is the lower limit of the correlation coefficient that can be used for fusion, which is determined by the actual application. We only consider the case of $|\rho_{(i,j)(p,q)}| > \phi$ which gives the following definition:

$$\rho_{(i,j)(p,q)} = \begin{cases} \rho 1, \left|\rho_{(i,j)(p,q)}\right| \in \left[1, (\varphi+2)/3\right) \\ \rho 2, \left|\rho_{(i,j)(p,q)}\right| \in \left[(\varphi+2)/3, (2\varphi+1)/3\right) \\ \rho 3, \left|\rho_{(i,j)(p,q)}\right| \in \left[(2\varphi+1)/3, \varphi\right) \end{cases} \quad (8)$$

3.2.1 Similar Multisource Data Fusion

In a specific scenario of WBAN, multiple sensors of the same type may collect identical human body parameters, which is redundant. However, when multiple nodes collect the same data, the correlation coefficient $\rho_{(i,j)(p,q)} = 1$, which is a strong correlation that satisfies the prerequisites for data fusion. Therefore, the data accuracy can be improved through redundant fusion.

Let $H1(i, j)$ be a set of similar sensor nodes of node $<i, j>$, defined as follows:

$$H1(i, j) = \{<p, q> | \rho_{(i,j)(p,q)} = 1 \quad p \in [1, N], q \in [1, c^p]\} \quad (9)$$

In the actual collection process, it is difficult to ensure that the data is simultaneously collected by the same type of sensor. If the time gap is too large, the data correlation will be weakened, and the effectiveness of fusion will gradually decrease. Therefore, it is necessary to compare the data collection time before fusion. Only the data that meets the specific time requirements can be used for fusion. As a result, the data collected by similar sensors $<p, q>$ can be expressed as $\{d_p^q(1), d_p^q(2), ..., d_p^q(t') | p \in [1, N], q \in [1, c^p]\}$, where its latest collected data is $d_p^q(t')$, $d_i^j(t)$ acquisition time is *t*, β is the upper limit of the fusion time, if $|t - t'| < \beta$, $d_p^q(t')$ can be used for the fusion of $d_i^j(t)$, the fusion method is as follows:

$$d_i^j(t) = \frac{\sum_{((p,q)\in H(i,j)) \cap |t'-t| < \beta} d_p^q(t') + d_i^j(t)}{Len(H1(i, j)) + 1} \quad (10)$$

3.2.2 Heterogeneous Multisource Data Fusion

In addition to similar sensors in WBAN, other sensors collect different types of data, such as respiratory rate, body temperature, pulse and blood oxygen. Although the data types are different, the collection source is the same, which reflects the physical condition of the same person from different perspectives. Therefore, these data also have a strong correlation, and their accuracy can be improved through complementary fusion. However, the correlation between different types of physiological parameters is also different, and the sensor used for fusion needs to be determined based on the actual application. A typical fusion process is divided into three steps:

Step 1: Calculate the correlation coefficient

According to the application characteristics, collected physiological data may have a correlation. Assume $d_i^j(t)$, the data from node $<i, j>$, needs to be fused, and $d_p^q(t)$ is the data from node $<p, q>$, which may have a correlation with the previous data. We calculate the correlation coefficient ρ between the two nodes according to Formula (7) and determine the type of ρ according to Formula (8).

Step 2: Establish a linear regression model

The data collected by node $<i, j>$ is represented as $d_i^j(t_1), d_i^j(t_2), ..., d_i^j(t_m)$ in time series, denoted as Y_m, and the data collected by the fused node $<p, q>$ is

$d_p^q(t_1), d_p^q(t_2), ..., d_p^q(t_m)$, denoted as X_m.

Let $f(x_m) = ax_m + b$, which is the regression model between X_m and Y_m, satisfying that the gap between $f(x_m)$ and Y_m is the smallest, then the calculation methods of parameters a and b are as follows:

Let $\varepsilon = \Sigma(f(x_m) - y_m)^2 = \Sigma(ax_m + b - y_m)^2$, ε is the deviation between the regression value $f(x_m)$ and the actual value Y_m, to obtain the minimum deviation, a and b need to satisfy:

$$1. \frac{\partial \varepsilon}{\partial a} = 2\sum(ax_m + b - y_m)x_m = 0$$

$$2. \frac{\partial \varepsilon}{\partial b} = 2\sum(ax_m + b - y_m) = 0$$

$$\Rightarrow \begin{cases} a = \dfrac{\sum(x_m - \bar{x})(y_m - \bar{y})}{\sum(x_m - \bar{x})^2} \\ b = \bar{y} - \dfrac{\sum(x_m - \bar{x})(y_m - \bar{y})}{\sum(x_m - \bar{x})^2}\bar{x} \end{cases}$$

Step 3: Data fusion

Let $H2<i,j>$ be the set of correlation between nodes with node $<i,j>$, the specific form is:

$$H2(i,j) = \{(<p,q> | \rho_{(i,j)(p,q)} = \rho 1 \cup \rho 2 \cup \rho 3, \quad p \in [1, N], q \in [1, c^p]\} \tag{11}$$

Same as Similar Multi-Source Data Fusion, for the correlated nodes $<p,q>$, the acquisition time t' of $d_p^q(t')$ and the acquisition time t of $d_p^q(t)$ need to satisfy $|t-t'| < \beta$ before they can be used for fusion.In view of the strength of the correlation between nodes, three different fusion factors $\gamma 1, \gamma 2, \gamma 3 (\gamma 1, \gamma 2, \gamma 3 \in (0,1)) \wedge (\gamma 1 > \gamma 2 > \gamma 3)$ are defined, corresponding to $\rho 1$, $\rho 2$ and $\rho 3$ respectively. According to different fusion factors, $d_i^j(t)$ and $H2(i,j)$ are combined with all the data that meet the requirements to realize the complementarity of different types of data, the specific calculation method is shown in Formula (12).

$$d_i^j(t) = d_i^j(t) \times (1 - \gamma 1 - \gamma 2 - \gamma 3) + F1 + F2 + F3 \tag{12}$$

$$F1 = \frac{\sum_{(p,q) \in H(i,j) \wedge (\rho_{(i,j)(p,q)} = \rho 1 \wedge (|t-t'| < \beta)} f(d_p^q(t'))}{Len(H2(i,j))} \times \gamma 1 \tag{13}$$

$$F2 = \frac{\sum_{(p,q) \in H(i,j) \wedge (\rho_{(i,j)(p,q)} = \rho 2 \wedge (|t-t'| < \beta)} f(d_p^q(t'))}{Len(H2(i,j))} \times \gamma 2 \tag{14}$$

$$F3 = \frac{\sum_{(p,q) \in H(i,j) \wedge (\rho_{(i,j)(p,q)} = \rho 3 \wedge (|t-t'| < \beta)} f(d_p^q(t'))}{Len(H2(i,j))} \times \gamma 3 \tag{15}$$

4 Design of the Data Fusion Scheme

WBAN nodes are limited by their computing performance, storage space, and energy consumption. Traditional WBAN data processing algorithms are usually executed in the cloud, utilizing their powerful data processing capabilities to achieve data fusion. However, WBAN needs to respond to various emergency situations in real time to locally address abnormal physiological parameters in a timely manner and improve the survival chance of the patient. In this regard, it is difficult for the cloud to meet WBAN real-time requirements. Our study takes full advantage of edge computing and combines WBAN topology to design a hierarchical data fusion model based on edge computing.

4.1 Hierarchical Data Fusion Model Based on Edge Computing

Edge computing provides a solution to the low-latency, high-reliability communication, and computing requirements of WBAN data fusion for medical monitoring. By deploying computing, storage, and processing functions at the edge of WBAN and taking advantage of its distribution and proximity to the collection end, we manage to reduce the amount of data transmission, decrease system energy consumption, improve user experience, and meet the real-time requirements of data fusion. In the hierarchical data fusion model based on edge computing, data collection, single source data fusion, similar multi-source data fusion and heterogeneous multi-source data fusion will be scheduled to sensors, aggregation nodes, edge or cloud for execution. To focus local sensors and sink nodes on simple calculations, edge nodes focus on real-time, short-period data processing, whereas the cloud focuses on non-real-time complex tasks such as big data analysis and realization of mutual cooperation. This effectively supports the real-time processing and execution of local application, as shown in Figure 2. The details are as follows:

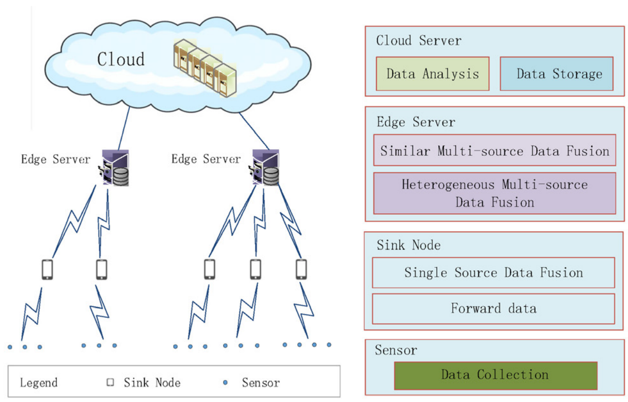

Figure 2. Hierarchical data fusion model based on edge computing

(1) The processing capacity, storage space, and energy of sensors are strictly limited. Generally, they only have simple information processing and wireless transmission functions; therefore, they can only carry out simple data acquisition. They periodically collect data and send the data to sink nodes.

(2) The processing capacity of the sink node is stronger than that of the sensor node, but it is generally battery-powered and limited by the computing storage capacity. As a result, it can be used for simple computing tasks but has difficulties handling complex computing tasks. The sink nodes wait to receive the data from the sensor node and then perform the single source fusion of the data collected by identical sensor nodes. Finally, they send the fusion result and the original data to the edge node.

(3) The edge node has strong computing power and is powered externally. It is close to the body and can undertake more complex real-time computing tasks. By offloading the computationally large multisource data fusion task to the edge node, we reduce the data flow from the device to the cloud and ensure efficient and real-time data fusion. After receiving the data, the edge node uses data redundancy and complementarity of sensors to realize both the fusion of similar multisource data and heterogeneous multisource data. As a result, it obtains more accurate results than a single sensor node and sends the fusion result and original data to the cloud to support the big data application.

(4) The cloud is used to deal with complex long-term data analysis tasks. It is also responsible for data storage, analysis, mining, decision-making assistance, etc.

Figure 3 shows the data processing flow of the sensor nodes, sink nodes, edge nodes, and cloud.

4.2 Data Frame Structure

The data frame of the sensor node includes sensor number *sid*, collection time *time*, original collection data *data*, and data service type *type* and is recorded as *SDF*. The sink node will add the sink node number *siid* and the result *sdata* of the SSDF with respect to *SDF* and is recorded as *SIDF*. The data frame of the edge node contains *SIDF*, the edge node number *eid*, the fusion result of similar multisource data *edata1*, the result of heterogeneous multisource data fusion *edata2* and is denoted as *EDF*, as shown in Figure 4.

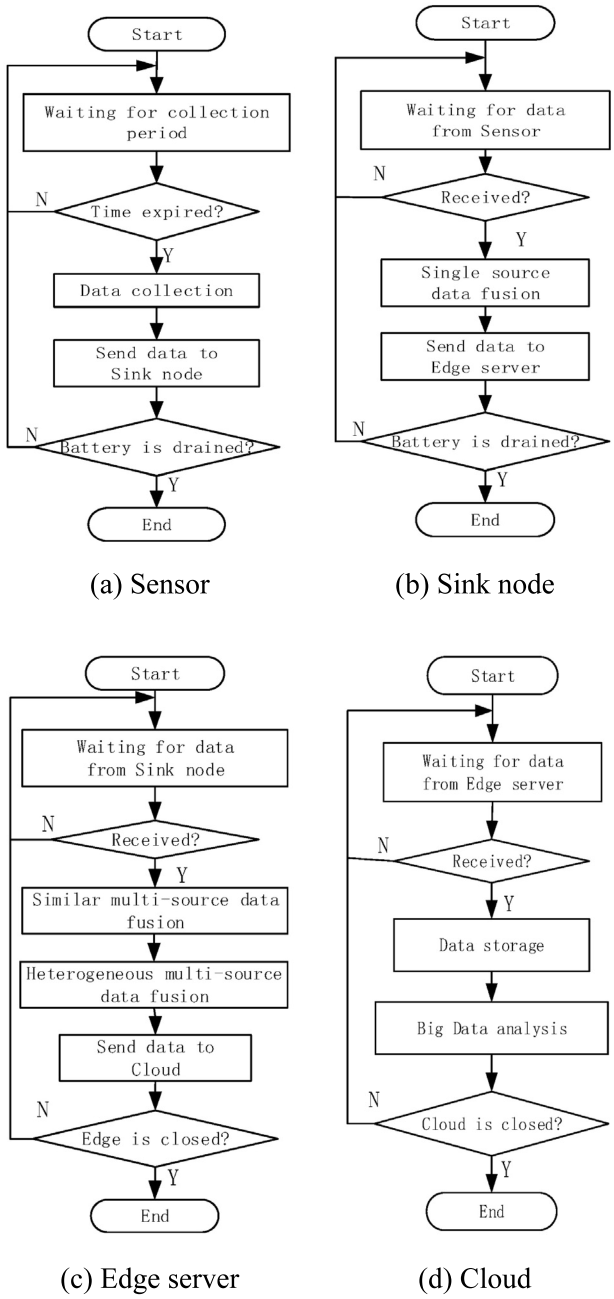

Figure 3. Data processing flowchart

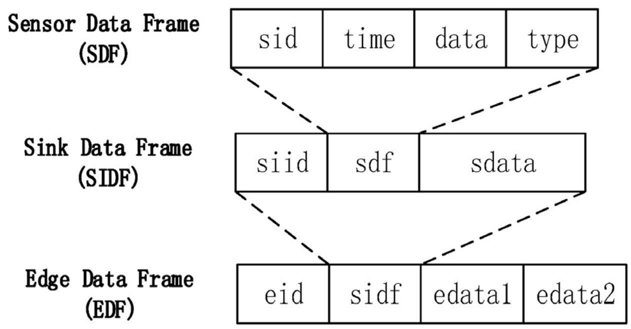

Figure 4. Data Frame structure

4.3 Algorithm of the Single Source Data Fusion

For a given sensor node $<i, j>$, the SSDF algorithm at time t uses Formulas (1)-(6) to fuse the data of a certain sensor in the time dimension. It also uses the queue Q_i to store the data sent by the associated sensor node. The implementation of SSDF is described as follows:

Algorithm 1. SSDF

Input:

sd is used to receive the data from the sensor, the data type is SDF

u and *l* are the upper and lower bound arrays of the data collected by the sensor

a is the variance threshold array of the data collected by the sensor

θ is the number of valid data in the queue

Output:

Q is the buffer queue of the sink node

```
SSDF(sd,Q,u,l,a,θ)
{
InitSIDF(fs); //Init SIDF
if (sd.data > u[sd.sid] or sd.data < l[sd.sid]) {
//Exceed the upper or lower bound data
fs.sdata = u[sd.sid] or l[sd.sid]; }
if (abs(sd.data-Q[rear].SDF.data) > a[sd.sid]) {
// Changes exceed the threshold
fs.sdata = Q[rear].SDF.data; }
if (Q.len<θ) {// Not enough data can be fused
EnQuene (Q,fs); return;}
for (i = Q.front to Q.rear) {// Moving average fusion
sum = sum+Q[i].sdata; }
fs.sdata = (sd+sum)/( θ+1);
EnQuene (Q, fs);
}
```

4.4 Multisource Data Fusion

Multisource data fusion is divided into SMDF and HMDF. The edge node receives the SIDF data sent by the sink node and stores it then executes SMDF and HMDF to generate edge data frame (EDF).

4.4.1 Algorithm of the Similar Multisource Data Fusion

For sensors with redundant nodes, the algorithm uses Formulas (7)-(10) to improve the accuracy and reliability of the collected data by fusing redundant data. The implementation of SMDF is described as follows:

Algorithm 2. SMDF

Input:

s is the SIDF that needs to be fused

D store SIDF queues of all sink nodes

ß is the maximum collection time interval

Output:

fe is used to store the fusion result

```
SMDF (s, D, ß, fe)
{
InitSIDF(si);
InitH1 (h1);
for (i=0 to D.len) {
// Establish a collection of similar sensors
for (j=0 to D[i].len) {
if (s.SIDF.SDF.type == D[i][j].type) {
EnQueue (h1, i, j); }
}
}
for (i = 0 to h1.len) {
// Find collected data that meets the constraint ß
rear = D[h1.m][h1.n].len-1;
if ((D.SDF.time-s.SDF.time) < ß){
EnQueue (si, D); }
}
for (i = 0 to si.len){
// Fusion of data that meets the conditions
sum = sum+si[i].SIDF.sdata;
}
fe.SIDF = s.SIDF;
fe.edata1 = sum/s.len;
}
```

4.4.2 Algorithm of the Heterogeneous Multisource Data Fusion

According to Formulas (7), (8), (11), (12), the HMDF algorithm uses the correlation strength between sensor nodes to perform a complementary fusion and improve the accuracy and reliability of collected data. The process of determining the presence of related nodes is combined with specific application scenarios. Therefore, before executing HMDF, it is necessary to configure the related node sequence *H2* according to Formula (11) in combination with actual applications. The implementation of HMDF is described as follows:

Algorithm 3. HMDF

Input:

s is the SIDF that needs to be fused

*h*2 stores heterogeneous associated sensor nodes

D store SIDF queues of all sink nodes

ß is the maximum collection time interval

γ1, γ2, γ3 are fusion factors

Output:

fe is used to store the fusion result

```
HMDF(s, h2, D, ß, fe, γ1, γ2, γ3)
{
```

```
InitSIDF(si);
for (i = 0 to h2.len){
//Query the data in h2 that meets the time constraints ß
rear = D[h2[i].m][h2[i].n].len-1;
if ((D.SDF.time-s.SDF.time)< ß){
EnQueue(si, D[h1.m][h1.n] [rear]); }
}
for(i = 0 to si.len){
//Fusion of the data according the fusion factors γ1, γ2, γ3
for (j = 0 to h2.len){
//According to the correlation choose different fusion factors
if (si.[i].SIDF.SDF.sid == j){
if (h[j].ρ ==ρ1){// Strong correlation
sum1 = sum1+si[i].SIDF.sdata;count1++;
} else if(h[j].ρ ==ρ2){// Moderate correlation
sum2 = sum2+si[i].SIDF.sdata;count2++;
} else if(h[j].ρ ==ρ3){// Weak correlation
sum3 = sum3+si[i].SIDF.sdata;count3++;}
sum = (sum1/count1)×γ1+(sum2/count2)×γ2+
(sum3/count3)×γ3;
}
fe.SIDF = s.SIDF;
fe.edata2 = sum+s×(1 - γ1 - γ2 - γ3);
}
```

5 Simulation and Analysis

The simulation is based on MSP430 and CC2530 chips for completing the hardware design of sensors and sink nodes. We use IBM System x3650 as an edge node, C language to implement various data fusion algorithms, and Python to analyze data. To verify the effectiveness of the simulation, we use three kinds of sensors to collect respiratory rate, pulse rate, and temperature data, with sampling periods 1000, 2000, and 3000 ms, respectively. To test the system performance under different parameter conditions, the simulation is carried out at six settings with regard to the number of sensors associated with the sink node. Each simulation time is 300 s, and the test data manually sets the abnormal data. The data follows the normal distribution $N(\mu,\sigma^2)$, $\mu=(U-L)/2$, $\sigma=\sqrt{(U+L)/2}$, where U and L refer to the upper and lower limits of the collected data. The simulation platform and specific parameter settings are shown in Table 1. The number of simulations indicates the total number of times each group of experiments is carried, whereas the simulation time refers to the duration of each group. Other simulation parameters have been explained in detail in Section 3.

This article designs the simulation from two perspectives: the first is to analyze the improvement in the accuracy of the measurement results given by the fusion algorithm and measure the percentage of abnormal data in the total abnormal data that can be effectively removed by the fusion algorithm. The second is to analyze the efficiency of the hierarchical data fusion model based on edge computing and verify it by comparing the execution time of the algorithm in various situations to each other. The execution time of the algorithm refers to the time required to execute the data fusion algorithm. To facilitate the data analysis, we will simply process the parameters that have no substantial impact on the simulation results according to a given situation. For example, for the processor processing rate ω in Formula (3), we use a constant value equal to 1 in the calculation because ω has little effect on the simulation results.

The simulation first analyzed the abnormal data detection rate of SSDF, SMDF, and HMDF algorithms in the same data set. It then compared the fusion effectivity of the three algorithms. SSDF is used to fuse data from a single sensor, SMDF is used to fuse data from multiple sensors of the same type, and HMDF is used to fuse two types of sensor data: respiratory rate and pulse rate.

Figure 5 shows the comparison between the abnormal data detection rates of the three fusion algorithms under six different sensor configurations: from 1, 1, 1 to 6, 6, 6 of the respiratory, pulse, and body temperature sensors. Notably, HMDF has highest average abnormal data detection rate as well as the most effective fusion. SMDF has the second-best results, whereas SSDF has the lowest. This indicates that the data fusion algorithm proposed here can effectively improve the accuracy and reliability of the collected data. Furthermore, when the number of sensors is only one, the detection rates of the three algorithms are not significantly different. However, as the number of sensors increases, the detection rates of SMDF and HMDF also increase, which presents an approximately linear growth curve, whereas the detection accuracy of SSDF does not change significantly. This is a result of an increase in sensor nodes, also increasing the amount of redundant and complementary data, such that SMDF and HMDF algorithms can make full use of the redundancy or complementarity of similar and heterogeneous sensor data to improve performance. On the other hand, SSDF mainly relies on the data of a single node, and an increase in sensor nodes cannot improve the fusion effect.

Table 1. Simulation parameters

		Edge	Sink Node	Sensor
Hardware platform	Processor	IBM System x3650	MSP430	CC2530 (8051CPU)
	Wireless Link	2.4GHz IEEE 802.15.4	2.4GHz IEEE 802.15.4	2.4GHz IEEE 802.15.4
Programming language		C and Python		
Simulaiton parameters	Type of sensors	Breathing, Pulse, Body Temperature		
	Number of sensors	1, 1, 1; 2, 2, 2; 3, 3, 3; 4, 4, 4; 5, 5, 5; 6, 6, 6		
	Correlation sensor	Respiratory rate, Pulse rate		
	Regression coefficients *a*, *b*	4, 0 [24-25]		
	Sampling period	1000ms, 2000ms, 3000ms		
	Upper limit *U*	30, 120, 40		
	Lower limit *L*	12, 48, 30		
	Data change threshold *a*	1, 2, 0.5		
	Queue length θ	10		
	Lower limit of correlation coefficient ϕ	0.3		
	Fusion time limit β	6000ms		
	Fusion factor $\lambda 1$, $\lambda 2$, $\lambda 3$	0.3, 0.2, 0.1		
	The simulation time	300s		
	Packet size	20 Bytes		
	The amount of data	Respiration rate 300 times Pulse rate 150 times Body temperature 100 times		
	The number of simulaitons	30 times		

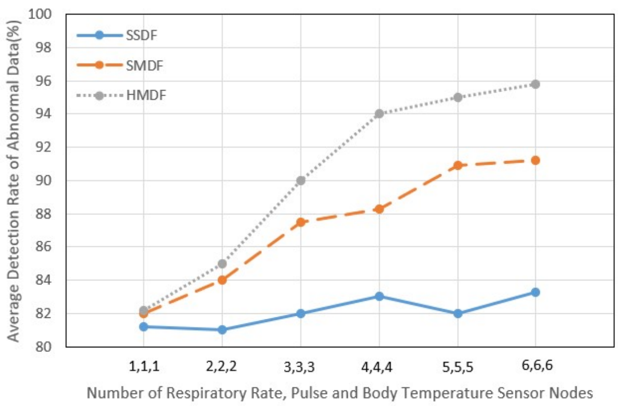

Figure 5. Comparison between the accuracies of the three algorithms

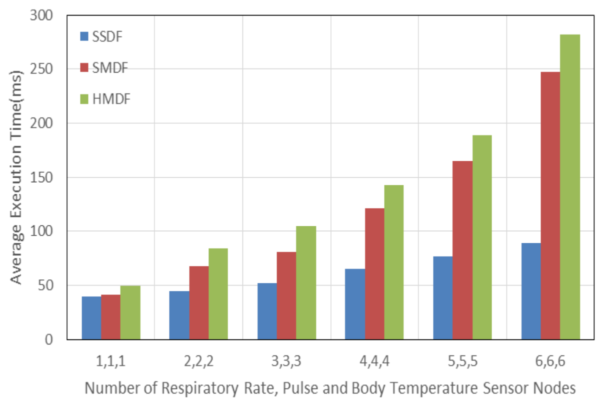

Figure 6. Comparison between the execution times of the three algorithms without edge computing architecture

Figure 6 is the simulation results of the three fusion algorithms SSDF, HMDF, and SMDF executed on the sink node. Figure 7 is the simulation result of SMDF executed on the sink node with HMDF and SMDF scheduled to be executed on the edge node. As the number of sensors increases, the execution time of HMDF and SMDF, as shown in Figure 6, has a larger increase than that of SSDF. However, the increase in Figure 7 is relatively monotonous. It is observed that the fusion algorithm based on the layered edge computing architecture designed in this study can effectively reduce the execution time of the algorithm, improving the real-time performance of the data fusion. When there is only one instance of each of the three types of sensors, the execution times of SSDF and SMDF are almost similar, whereas HMDF takes

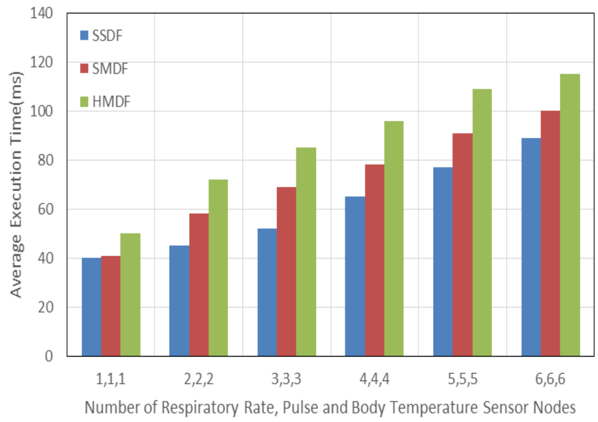

Figure 7. Comparison between the execution times of the three algorithms with edge computing architecture

significantly more time than both. This is because each type of sensor has only one instance, but SMDF fusion requires at least two similar sensor nodes. Thus, this type of fusion cannot be performed, and the execution time remains almost the same as that of SSDF. On the other hand, when there is only one respiratory and pulse rate, the fusion requirements of HMDF are met. Therefore, the execution time of HMDF is significantly higher than those of SSDF and SMDF. When the data size is not very large, the difference between different algorithms is small. However, when the number of sensors increases and the data size increases accordingly, the increase in the execution times of HMDF and SMDF algorithms is greater than that of SSDF. This is because the two algorithms need to fuse data collected by multiple similar and heterogeneous sensors, whereas SSDF only fuses data from a single sensor at a given time. As a result, the increase in execution time of SSDF is relatively low.

From a comprehensive analysis point of view, the fusion algorithm based on the layered edge computing architecture proposed here has high data fusion performance and effectively improves the accuracy of the detection results. However, the execution time increases synchronously with increasing number of nodes this increase is controllable and acceptable because it is limited by the node scale of WBAN and the architecture of edge computing.

The Kalman filter algorithm is a highly versatile data fusion algorithm with a wide range of applications. Therefore, we choose to compare HDMF to Kalman filter to analyze the performance differences between the two algorithms in terms of accuracy and execution time.

Figure 8 is a boxplot diagram of the abnormal data detection rates of HDMF and Kalman filter, showing 30 simulation datasets under 6 different sensor number configurations. It is observed that there is not much difference between the two when the sensors are few. However, with an increase in the number of sensors of each type, especially when there are more than three, the detection rate of HDMF is significantly higher than that of the Kalman filter. This is because HDMF requires sufficient nodes to fully leverage the advantages of multisource fusion, make full use of the correlation between sensor data, and effectively improve the accuracy and reliability of measurement results.

Figure 9 is a boxplot diagram of HDMF and Kalman filter execution times, showing 30 simulation datasets under 6 different sensor configurations. It is observed that with an increase in the number of sensors, the execution time required by the Kalman filter shows a slower increase trend. This is because the number of sensors to be processed by the algorithm also increases, and the data processing time of a single sensor remains unchanged. Therefore, there is not much increase. However, in the same case, the execution time required

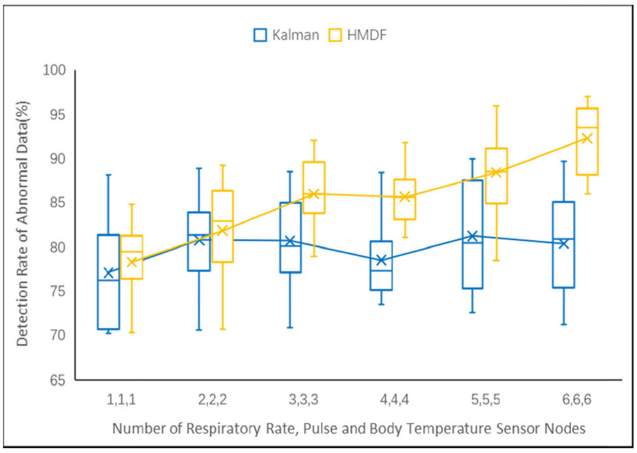

Figure 8. Comparison between the accuracies of HDMF and Kalman Filter

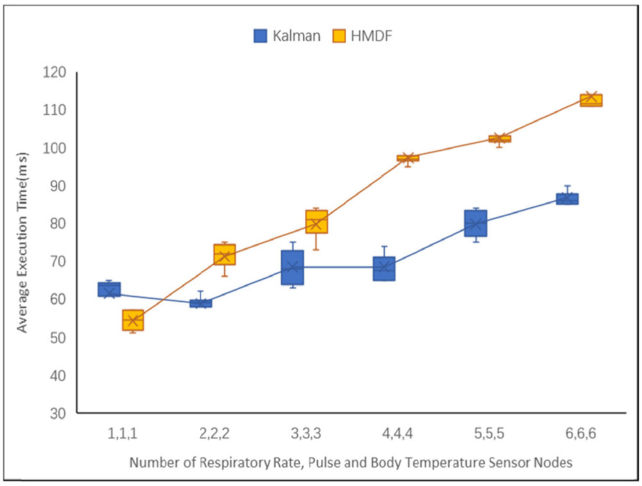

Figure 9. Comparison between the execution times of HDMF and Kalman Filter

by HDMF shows an increasing trend that is significantly faster than that of Kalman filter because increasing the number of sensors increases redundancy and complementary fusion. This results in a faster increase in the algorithm execution time compared to that of the Kalman filter.

Considering the accuracy and execution time of fusion results, the abnormal data detection rate of HDMF is significantly higher than that of the Kalman filter with an increasing number of sensors. This indicates that HDMF has better accuracy and reliability in WBAN. At the same time, when the number of sensors is small, the difference in execution time is very small. Although the increase in HDMF execution time is greater than that of Kalman Filter with an increase in the number of sensors, considering that the sensors in WBAN are mainly wearable devices, the number of sensors is not too large. When the architecture based on edge calculation is adopted, the increase in algorithm execution time is controllable. Therefore, compared to the improvement in detection accuracy, the cost of such execution is justified.

The simulation results show that the data fusion algorithm based on hierarchical edge computing architecture proposed here can effectively improve the accuracy, reliability, and real-time nature of sensor measurement results under the condition that the execution time is acceptable.

6 Conclusions

This research studies the application of WBAN in the medical monitoring field and analyzes the problems and challenges faced by sensor data fusion. To solve existing problems, a hierarchical data fusion model based on edge computing and corresponding data fusion algorithm are proposed. First, to address the real-time requirement of sensor data fusion, an effective scheduling of data fusion tasks is proposed regarding the sensors, sink nodes, edge nodes, and cloud. Through reasonable task offloading, the real-time performance of data fusion is improved. Second, utilizing the redundancy and complementarity of the sensor data in time and space, three data fusion algorithms—SSDF, SMDF and HMDF—are proposed, which effectively improve the accuracy and reliability of the collected data. Finally, a simulation platform is built, which is used to analyze the performance of our prosed algorithm with regard to the accuracy, reliability and execution time parameters of the data fusion. The simulation results prove the effectiveness of the proposed algorithm.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62072085 and Liaoning Revitalization Talent Program under Grant XLYC1902017.

References

[1] F. Bragg, M. V. Holmes, A. Iona, Y. Guo, H. Du, Y. Chen, Z. Bian, L. Yang, W. Herrington, D. Bennett, I. Turnbull, Y. Liu, S. Feng, J. Chen, R. Clarke, R. Collins, R. Peto, L. Li, Z. Chen, Association between diabetes and cause-specific mortality in rural and urban areas of China, *The Journal of the American Medical Association*, Vol. 317, No. 3, pp. 280-289, January, 2017.

[2] S. Ullah, H. Higgins, B. Braem, B. Latre, C. Blondia, I. Moerman, S. Saleem, Z. Rahman, K. S. Kwak, A comprehensive survey of wireless body area networks, *Journal of Medical Systems*, Vol. 36, No. 3, pp. 1065-1094, June, 2012.

[3] IEEE Standard for Local and metropolitan area networks - Part 15.6: Wireless Body Area Networks, *IEEE Std 802.15.6-2012*, pp. 1-271, February, 2012.

[4] R. Chavez-Santiago, K. Sayrafian-Pour, A. Khaleghi, K. Takizawa, J. Wang, I. Balasingham, H.-B. Li, Propagation models for IEEE 802.15.6 standardization of implant communication in body area networks, *IEEE Communications Magazine*, Vol. 51, No. 8, pp. 80-87, August, 2013.

[5] K. Hasan, K. Biswas, K. Ahmed, N. Nafi, M. Islam, A comprehensive review of wireless body area network, *Journal of Network and Computer Applications*, Vol. 143, pp. 178-198, October, 2019.

[6] M. Baig, S. Afifi, H. Gholamhosseini, F. Mirza, A systematic review of wearable sensors and iot-based monitoring applications for older adults – a focus on ageing population and independent living, *Journal of Medical Systems*, Vol. 43, No. 8, pp. 233:1-233:11, August, 2019.

[7] B. Khaleghi, A. Khamis, F. Karray, S. Razavi, Multisensor data fusion: a review of the state-of-the-art, *Information Fusion*, Vol. 14, No. 1, pp. 28-44, January, 2013.

[8] R. Gravina, P. Alinia, H. Ghasemzadeh, G. Fortino, Multi-sensor fusion in body sensor networks: state-of-the-art and research challenges, *Information Fusion*, Vol. 35, pp. 68-80, May, 2017.

[9] W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu, Edge Computing: Vision and Challenges, *IEEE Internet of Things Journal*, Vol. 3, No. 5, pp. 637-646, October, 2016.

[10] R. Yu, D. Ye, Z. Wang, B. Zhang, A. M. Oguti, J. Li, B. Jin, F. Kurdahi, Cffnn: cross feature fusion neural network for collaborative filtering, *IEEE Transactions on Knowledge and Data Engineering*, doi: 10.1109/TKDE.2020.3048788, January, 2021.

[11] Y. Zhou, H. Xu, J. Sun, L. Pan, B. Du, M. Chen, Multisensor Data Fusion Based on Modified D-S Evidence Theory, *2018 International Conference on Computer Modeling, Simulation and Algorithm (CMSA2018)*, Beijing, China, 2018, pp. 324-327.

[12] W. Sung, J. Chen, D. Huang, Y. Ju, Multisensors realtime data fusion optimization for IOT systems, *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, San Diego, CA, USA, 2014, pp. 2299-2304.

[13] F. Casciati, S. Casciati, L. Faravelli, M. Vece, Validation of a data-fusion based solution in view of the real-time monitoring of cable-stayed bridges, *Procedia Engineering*, Vol. 199, pp. 2288-2293, 2017.

[14] J. Xiong, Q. Zhang, J. Wan, L. Liang, P. Cheng, Q. Liang, Data fusion method based on mutual dimensionless, *IEEE/ASME Transactions on Mechatronics*, Vol. 23, No. 2, pp. 506-517, April, 2018.

[15] G. Liu, F. Xiao, Time Series Data Fusion Based on Evidence Theory and OWA Operator, *Sensors*, Vol. 19, No. 5, Article No. 1171, March, 2019.

[16] S. Ji, Z. Y. Chen, P. Guo, Y. J. Sun, J. Shen, J. Wang, C. F. Lai, Bayesian approach for multi-sensor data fusion based on compressed sensing for wireless structural damage signal, *Journal of Internet Technology*, Vol. 17, No. 7, pp. 1363-1371, December, 2016.

[17] D. Sethi, J. Anand, Big data and WBAN: prediction and analysis of the patient health condition in a remote area, *Engineering and Applied Science Research*, Vol. 46, No. 3,

pp. 248-255, July-September, 2019.
[18] D. Seshadri, R. Li, J. Voos, J. Rowbottom, C. Alfes, C. Zorman, C. Drummond, Wearable sensors for monitoring the physiological and biochemical profile of the athlete, *Digital Medicine*, Vol. 2, No. 1, Article No. 72, July, 2019.
[19] W. Yi, O. Sarkar, S. Mathavan, J. Saniie, Wearable sensor data fusion for remote health assessment and fall detection, *IEEE International Conference on Electro/Information Technology*, Milwaukee, WI, USA, 2014, pp. 303-307.
[20] F. Sanfilippo, K. Pettersen, A sensor fusion wearable health-monitoring system with haptic feedback, *11th International Conference on Innovations in Information Technology (IIT)*, Dubai, United Arab Emirates, 2015, pp. 262-266.
[21] N. Capela, E. Lemaire, N. Baddour, Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients, *PLos One*, Vol. 10, No. 4, Article No. e0124414, April, 2015.
[22] C. Chen, R. Jafari, N. Kehtarnavaz, Improving human action recognition using fusion of depth camera and inertial sensors, *IEEE Transactions on Human-Machine Systems*, Vol. 45, No. 1, pp. 51-61, February, 2015.
[23] J. Xu, S. Liu, X. Lu, L. Li, H. Liang, Q. Duan, R. Liu, A Temporal Adaptive Access Mechanism for Data Fusion in an IoT Environment, *Sensors*, Vol. 18, No. 12, Article No. 4205, December, 2018.
[24] A. Iliff, V. Lee, Pulse rate, respiratory rate, and body temperature of children between two months and eighteen years of age, *Child Development*, Vol. 23, No. 4, pp. 237-245, December, 1952.
[25] P. Davies, I. Maconochie, The relationship between body temperature, heart rate and respiratory rate in children, *Emergency Medicine Journal*, Vol. 26, No. 9, pp. 641-643, September, 2009.

Biographies

Jian Yan is a Ph.D. candidate at the School of Computer Science and Engineering of Northeastern University, China. He received the MS degree in computer science and technology from Soochow University, China in 2009. His research interests include WBAN, embedded system, and edge computing.

Qingxu Deng is a professor of the School of Computer Science and Engineering, Northeastern University, China, where he serves as the Director of institute of Cyber Physical Systems. His main research interests include Cyber-Physical systems, embedded systems, and real-time systems.

Guangjie Han received the Ph.D. degree from Northeastern University, Shenyang, China, in 2004. From October 2010 to 2011, he was a Visiting Research Scholar with Osaka University, Suita, Japan. He is currently a Professor with the Department of Information and Communication System, Hohai University, Nanjing, China.

Yuhan Lin is a Ph.D. candidate at School of Computer Science and Engineering of Northeastern University, China. He received the BS and MS degree in computer science and technology from Northeastern University, China in 2007 and 2013. His research interests include embedded system, scheduling theory, and internet of things.