

# Joint Optimization of Data Urgency and Freshness in Wireless Body Area Networks for Enhanced eHealth Monitoring

Zheng Zhang, Zhangyong Li, Chun Sing Lai, *Senior Member, IEEE*, Ruiheng Wu, *Senior Member, IEEE*, Xinwei Li, Jinzhao Lin, Xinxing Ren and Huidong Qiao

**Abstract**—Wireless Body Area Networks (WBANs), as an effective technology for electronic health monitoring, have transformed traditional consumer electronics (CE) into the next generation of devices with enhanced connectivity and intelligence. The improved interconnectivity between sensor nodes, coordinators, and other consumer devices has increased data availability and enabled autonomous monitoring within CE networks. However, due to the time-sensitive nature of physiological data transmission in WBANs and the urgency of sensor node data, addressing real-time data transmission under dynamic link conditions remains a significant challenge. To tackle this issue, we propose a joint optimization scheduling strategy that considers both data urgency and freshness. Our proposed strategy consists of two key components: a Sink Channel Allocation (SCA) strategy and a Node Scheduling Selection (NSS) strategy. By integrating deep reinforcement learning (DRL), we overcome the challenges posed by the large action space in channel allocation and timeslot selection, thereby improving scheduling efficiency. Both theoretical analysis and simulation results demonstrate that our method significantly outperforms traditional approaches in terms of real-time data transmission and scheduling optimization.

**Index Terms**— Consumer Electronics, Wireless Body Area Network, Data Urgency, Age of Information, Scheduling, Reinforcement Learning.

## I. INTRODUCTION

CONSUMER electronics (CE) are now integral to various aspects of modern society. These include electronic entertainment, daily health monitoring, medical fitness tracking, and individual combat operations, as shown in Fig. 1. A

This work was supported in part by the National Natural Science Foundation of China under Grant U21A20447, 62471077, 62473140 and 62173134; in part by the Chongqing Natural Science Foundation under Grant CSTB2022NSCQ-LZX0064 and CSTB2022NSCQ-LZX0069; in part by the Natural Science Foundation of Hunan Province under Grant 2023JJ50029, 2023JJ60151 and 2022JJ30197. (*Corresponding author: Zhangyong Li; Chun Sing Lai.*)

Zheng Zhang and Jinzhao Lin are with the School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: Zheng.Zhang2@brunel.ac.uk; linjz@cqupt.edu.cn).

Zhangyong Li and Xinwei Li are with the Chongqing Engineering Research Center of Medical Electronics and Information Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: lizy@cqupt.edu.cn; lixinwei@cqupt.edu.cn).

Chun Sing Lai, Ruiheng Wu, and Xinxing Ren are with the College of Engineering, Design and Physical Sciences, Brunel University London, UB8 3PH, United Kingdom (e-mail: chunsing.lai@brunel.ac.uk; ruiheng.wu@brunel.ac.uk; xinxing.ren@brunel.ac.uk).

Huidong Qiao is with the National University of Defense Technology, Changsha, 410000, China (e-mail: east\_qiao@126.com).

Digital Object Identifier xx

common characteristic of these applications is the integration of sensing devices with the human body. The rapid development of the Internet of Things (IoT) has enabled seamless connectivity between devices and people [1]. Simultaneously, the increasing number of chronic disease patients globally has made it difficult for individuals to receive timely medical evaluations and advice, driving the demand for daily health monitoring services [2]. Wireless Body Area Networks (WBANs), as a key physiological data monitoring solution, have garnered significant attention in the personal consumer electronics sector. WBANs are now recognized as a central theme in next-generation CE and consumer technology (CT) research and are widely regarded as a critical solution for future human-machine integration [3], [4].

As the preferred technology for future CE, WBANs are networks specifically designed for human-centric applications, enabling real-time, continuous, and long-term monitoring of physiological data. These networks facilitate timely diagnosis and provide valuable information to healthcare professionals [5], [6]. By utilizing low-power, miniaturized physiological sensor nodes, WBANs collect various physiological data and vital signs from the body [7]. Depending on the type of application, this data is transmitted via a single-hop star topology to a coordinator (Sink) within the WBAN (intra-WBAN), which then forwards the information to Remote Monitoring Center (RMC) or cloud platforms for health monitoring services (beyond-WBAN).

The time sensitivity of physiological data transmission is of paramount importance in WBAN systems, similar to other key performance indicators in CE. Communication metrics, particularly real-time data transmission and the urgency of data from sensor nodes, are critical [8]. Delayed physiological data can lead to unnecessary energy consumption or severe consequences, such as impacting medical decisions or causing accidents in emergencies. However, traditional delay metrics are insufficient to meet the stringent time-sensitivity requirements of WBANs [9].

To address this issue, the Age of Information (AoI) has been introduced as a novel metric [10]. AoI considers transmission time, queue waiting time, and the duration that data remains after being received to measure the freshness of information [11], [12]. It has been widely adopted to mitigate queuing delays in the transmission of physiological data packets [13], [14]. Numerous studies [15]–[20] use AoI to evaluate wireless network performance, including scheduling algorithms that

minimize average AoI [15], cooperative scheduling methods for sensor nodes [16], and strategies that minimize average AoI while maintaining throughput constraints [17], [18]. These studies also explore transmission scheduling strategies, link scheduling strategies in noisy channels [19], and considerations of average peak AoI in IoT contexts with deadlines [20].

Despite the progress in AoI research, most studies focus on enhancing network system performance without fully considering the specific scenarios of WBANs. Moreover, existing methods often lack to address the urgency of physiological data in WBAN systems. For instance, transmitting abnormal Electrocardiogram (ECG) signals immediately to RMC is critical for emergency decision making [21]. However, most scheduling strategies simply categorize physiological data into broad classes such as urgent or periodic, leading to resource wastage from transmitting irrelevant data. This exacerbates the delay problem and highlights the need to prioritize and improve the freshness of received physiological data in WBAN systems.

To overcome this difficulty, we propose a scheduling strategy that jointly optimizes data urgency and freshness, with the goal of maximizing the urgency of received data while minimizing AoI. The novelty and contribution of this study are highlighted as follows.

- We formulated a joint optimization problem that focuses on both data urgency and AoI. By considering the inter-relationship between these two factors, we introduce the Urgency Level to AoI Ratio (ULAR) as our performance metric. ULAR measures the system's ability to transmit urgent data while maintaining data freshness. Our primary objective is to maximize ULAR, thereby ensuring timely and prioritized delivery of critical physiological data.
- We decomposed the scheduling strategy into two inter-related strategies: the Sink Channel Allocation (SCA) strategy and the Node Scheduling Selection (NSS) strategy, to efficiently solve this optimization problem. This decomposition facilitates an effective solution. Considering the dynamic nature of the AoI, we introduced a two-stage scheduling strategy based on deep reinforcement learning (DRL) to tackle the challenge of a large action space and effectively reduce its size.
- To evaluate the performance of the proposed scheduling algorithm, we conducted comparative experiments with different algorithms to ensure a comprehensive assessment of our simulation results. The results validate the effectiveness of the proposed method.

The remainder of this paper is organized as follows: Section II reviews related work on scheduling strategies. Section III presents the system model. The problem formulation is introduced in Section IV. Section V details the joint scheduling policy. Section VI evaluates the performance of the proposed algorithms through simulations. Finally, Section VII concludes the paper.

## II. RELATED WORK

The field of health monitoring has seen extensive research, and this section aims to briefly review and summarize the related work on transmission scheduling strategies in WBANs.

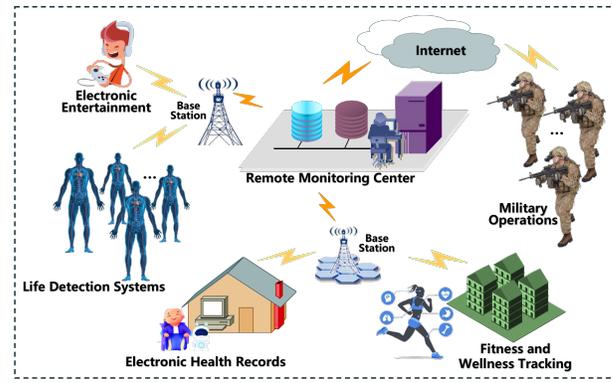


Fig. 1: Remote Monitoring System for Consumer Electronics.

Early transmission scheduling schemes typically involve prioritizing and ranking nodes based on various criteria [22]–[27]. For example, The authors [22] designed a utility function for nodes, which calculates the utility to improve network performance. The authors [23] considered different parameters of sensor nodes to calculate their importance, achieving a reduction in packet loss probability. Liu *et al.* [24] dynamically adjusted the transmission order and duration of nodes based on channel conditions and application context to optimize quality of service. Ullah *et al.* [25] proposed a link scheduling method based on sensor clustering and cooperative routing protocols. In [26] developed a data forwarding strategy that uses data compression to balance the network performance of sensor nodes. In [27] introduced a mixed cost parameter to evaluate the effectiveness of quality of service. Additionally, some studies focused on prioritizing and scheduling based on health parameters and data types, Liang *et al.* [28] designed a multi-level priority scheme and used a step-based timeslot allocation method for scheduling node data. Kim *et al.* [29] proposed a multi-criteria decision-making method for link scheduling of node data. However, these studies mainly focus on node prioritization and do not adequately consider the urgency of the data.

In terms of scheduling strategies for urgent data, existing research often involves redesigning superframe structures [30]–[37]. For instance, Deepak *et al.* [30] proposed a channel access scheme for nodes carrying urgent data frames to enhance network performance. Misra *et al.* [31] addressed urgent data scheduling by adjusting the data rates of different nodes. In [32] introduced an adaptive medium access control (MAC) algorithm that adjusts the MAC frame payload of WBAN sensor nodes according to the severity of the sensed health parameters, in compliance with the IEEE 802.15.4 protocol. In [33] developed an energy-efficient medium access control protocol that modifies the superframe structure and allocates priority levels to detect and handle urgent events from in-body sensors. The author [34] designed an emergency-prioritized timeslot allocation scheme where relay nodes perform channel sensing and handle urgent data. The author [35], [36] proposed coordinated superframe duty cycle hybrid MAC (SDC-HYMAC) and multi-channel hybrid MAC (MC-HYMAC) protocols to ensure the network performance of urgent critical

TABLE I: Description of Key Notations

Notation	Description
$C$	Number of available channels
$S$	Number of Sinks
$s$	Index of the Sink
$\chi$	Set of all Sinks
$\gamma_s$	Set of sensor nodes associated with Sink $s$
$N_s$	Number of sensor nodes associated with Sink $s$
$n_s$	Index of a sensor node associated with Sink $s$
$t$	Index of time slot
$\mathbf{Y}$	Set of data packet sizes for sensor nodes
$Y_j$	Specific data packet size in $\mathbf{Y}$
$Y_{n_s}(t)$	Actual data packet size transmitted by sensor node $n_s$ in time slot $t$
$p_{n_s}^{Y_j}$	Probability that sensor node $n_s$ samples a data packet of size $Y_j$
$p_{n_s}^d$	Probability that sensor node $n_s$ samples data with urgency level $d$
$\varrho_{n_s}(t)$	Whether the sensor node $n_s$ occupies the channel for data transmission in time slot $t$
$\xi_{n_s}(t)$	Sampling timeslot of data transmission for sensor node $n_s$ in time slot $t$
$l_{n_s}(t)$	Urgency of physiological data transmitted by sensor node $n_s$ in time slot $t$
$d_N$	Number of slots required for sensor node $n_s$ to transmit a data packet
$R_s$	Symbol rate of the transmission channel
$pASlotRes$	Length of allocated unit slots for a sensor node
$m_{n_s}(t)$	Number of slots needed for sensor node $n_s$ to complete the transmission
$N_{Pre}$	Length of the preamble
$N_{Header}$	Length of the PLCP header
$N_j$	Length of the PSDU
$SPLCP_{Header}$	Spreading factor of the PLCP header
$M$	Cardinality of the constellation in a modulation scheme
$\pi$	A feasible scheduling strategy
$\bar{\Lambda}_\pi$	Average urgency of received physiological data under scheduling policy $\pi$
$\bar{A}_\pi$	Average AoI of received physiological data under scheduling policy $\pi$

physiological data transmission. In [37] introduced a deep deterministic policy gradient algorithm inspired by random graphs, formulating the problem as a Markov decision process to ensure the transmission performance of urgent critical physiological data. These methods have achieved some success in the urgent transmission of anomalous data and have considered performance metrics such as delay. However, they rarely address the freshness of the data during the scheduling process. Although some studies mention data freshness (e.g., [38], [39]), they do not comprehensively consider the urgency of the data.

Despite the progress made in transmission scheduling, there remains a lack of strategies that jointly optimize data urgency and freshness. Unlike the aforementioned studies, our work comprehensively considers the health monitoring needs of WBANs. By integrating data urgency and freshness, we achieve adaptive transmission. Additionally, to address the issue of large action spaces in existing algorithms, we decompose the scheduling strategy into two parts: one for Sink channel allocation and the other for timeslot selection, aiming to balance data urgency with the average AoI in the system.

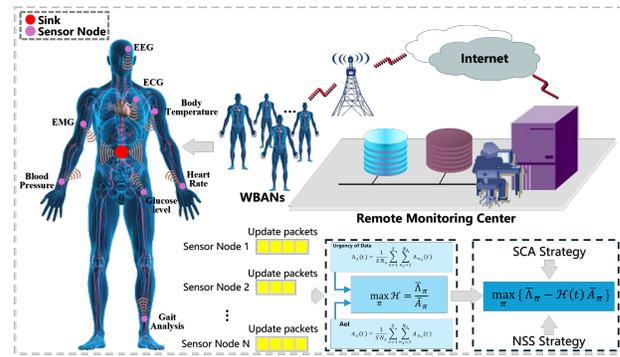


Fig. 2: System Model.

### III. SYSTEM MODEL

We investigate a typical slot-based WBAN system. The system consists of a RMC and  $S$  WBANs. Each WBAN includes a Sink, denoted by  $s \in \chi = \{1, 2, \dots, S\}$ , where each Sink is equipped with  $N_s$  sensor nodes tasked with collecting physiological data. Each sensor node is indexed by  $n_s \in \gamma_s = \{1, 2, \dots, N_s\}$ . These sensor nodes transmit data wirelessly to the Sink in a star topology configuration. Subsequently, the Sink forwards the physiological data to the RMC using WiFi or 5G wireless channels, as illustrated in Fig. 2. Given that data forwarding from Sink to RMC falls under the category of beyond-WBAN, our primary focus is on the process of sensor nodes transmitting data to the Sink (intra-WBAN). Some key notations in the optimization problem are listed in Table I.

In this system, we assume that the number of available wireless channels is  $C$ . During each time slot, up to  $C$  channels can be allocated for uplink data transmission, where the number of active Sinks  $S$  satisfies  $C \leq S$ . To avoid transmission conflicts among sensor nodes within a single WBAN, we adopt a Time Division Multiple Access (TDMA) protocol. Time is divided into discrete and equal length slots denoted as  $t \in \{1, 2, \dots, T\}$ . In each time slot, each Sink selects only one sensor node  $n_s$  for data transmission. Regarding the upload of physiological data from sensor nodes, not all sensor nodes access the wireless channel in every time slot. Therefore, we adopt a generalized model similar to that in [15]. Specifically, for sensor nodes assigned a channel in time slot  $t$ , data sampled before time slot  $t$  is discarded, and only the most recent data sampled in time slot  $t$  is uploaded. This process continues until the data is successfully transmitted, at which point the Sink sends an acknowledgment (ACK) to release the allocated channel.

Due to the heterogeneity of sensor nodes, different sensor nodes exhibit different traffic patterns, as listed in Table II, including continuous or periodic data transmission, or transmission triggered by sudden events (e.g., heart attack). Therefore, the Sink must allocate slots that precisely match the traffic generation time of each node to avoid delays or wasted bandwidth. Additionally, accurate calculations of the required round-trip time for data transmission, as well as physical layer (PHY) functions and data lengths, are essential. To model the data packet sizes for sensor nodes, let  $j \in \{1, 2, \dots, J\}$  denote

TABLE II: Different Types of Medical Sensors

Sensing type	Data rate	Bandwidth
Electromyography (EMG)	320 kbps	0-10000 Hz
Electrocardiogram (ECG)	288 kbps	100-1000 Hz
Electroencephalogram (EEG)	42.3 kbps	0-150 Hz
Glucose	1600 bps	0-50 Hz
Temperature	120 bps	0-1 Hz
Cochlear implant	100 kbps	-
Voice	50-100 kbps	-

the index of packet sizes, where  $Y_j \in \mathbf{Y}$  represents a specific size within the set  $\mathbf{Y} = \{Y_1, Y_2, \dots, Y_J\}$ . The data packet size  $Y_j$  can be expressed as

$$Y_j = N_{Pre} + N_{Header} \cdot S_{PLCPHeader} + \frac{N_j}{\log_2 M} \quad (1)$$

where  $N_{Pre}$ ,  $N_{Header}$ , and  $N_j$  represent the length of the preamble, the physical layer convergence protocol (PLCP) header, and the PLCP service data unit (PSDU), respectively.  $S_{PLCPHeader}$  and  $M$  denote the spreading factor of the PLCP header and the cardinality of the constellation of a given modulation scheme, respectively.

Let  $d_N$  denote the number of slots required for sensor node  $n_s$  to complete the upload of physiological data packets, i.e.,  $d_N = \lceil \frac{Y_j}{R_s \cdot pASlotRes} \rceil$ , where  $R_s$  and  $pASlotRes$  denote the symbol rate and the length of the allocated unit slots requested by the sensor node, respectively. Assume that sensor node  $n_s$  has a probability  $p_{n_s}^{Y_j}$  of sampling data packets of size  $Y_j$ , where  $\mathbf{Y} = \{Y_1, Y_2, \dots, Y_J\}$  is the set of packet sizes. The actual data packet size sampled or transmitted by sensor node  $n_s$  during a specific time slot is denoted as  $Y_{n_s}(t)$ , which corresponds to one value  $Y_j$  in the set  $\mathbf{Y}$ , determined by the probability distribution  $p_{n_s}^{Y_j}$ . By definition, the probabilities satisfy  $\sum_{j=1}^J p_{n_s}^{Y_j} = 1, \forall n_s \in \gamma_s$ .

For each sensor node  $n_s$ , let  $\varrho_{n_s}(t) \in \{0, 1\}$  indicate whether it occupies the channel for data transmission in a given timeslot. Specifically, if sensor node  $n_s$  transmits data during timeslot  $t$ ,  $\varrho_{n_s}(t)$  is set to 1; otherwise, it is set to 0. The variable  $\xi_{n_s}(t)$  denotes the sampling timeslot for data transmission of sensor node  $n_s$  in timeslot  $t$ . In this context,  $m_{n_s}(t)$  represents the number of timeslots needed for the complete transmission of the data packet  $Y_{n_s}(t)$ . The required number of timeslots to complete the transmission is given by

$$m_{n_s}(t) = \lceil \frac{Y_{n_s}(t)}{R_s \cdot pASlotRes} \rceil \quad (2)$$

This formula calculates the minimum number of timeslots required for sensor node  $n_s$  to successfully upload a data packet of size  $Y_{n_s}(t)$ , considering the allocated channel resources. The precise allocation of these timeslots ensures efficient utilization of the channel and minimizes potential delays. By integrating this calculation into the overall scheduling framework, the system can dynamically adapt to the varying data sizes and transmission requirements of individual sensor nodes.

## IV. PROBLEM FORMULATION

### A. Urgency of Data

We categorize the urgency of physiological data within the WBAN system into  $D$  levels. The urgency of physiological data transmitted by sensor node  $n_s$  in timeslot  $t$  is denoted by  $l_{n_s}(t) \in \{1, 2, \dots, D\}$ , where higher levels indicate greater urgency of the physiological data. The urgency of data received by the Sink from sensor node  $n_s$  in timeslot  $t$  is represented by  $\Lambda_{n_s}(t) \in \{0, 1, 2, \dots, D\}$ . When  $\varrho_{n_s}(t) = 1$ , sensor node  $n_s$  begins sampling data in timeslot  $t$ , and the probability distribution of the urgency level of the sampled data is denoted as  $\Pr(l_{n_s}(t) = d) = p_{n_s}^d$ , where  $d \in \{1, 2, \dots, D\}$  and  $\sum_{d=1}^D p_{n_s}^d = 1, \forall n_s \in \gamma_s$ . If the data from the sensor node is successfully transmitted to the Sink by the end of timeslot  $t$ , then  $\Lambda_{n_s}(t)$  equal  $l_{n_s}(t)$ . The dynamic nature of  $\Lambda_{n_s}(t)$  is expressed as

$$\Lambda_{n_s}(t) = \begin{cases} l_{n_s}(t)\varrho_{n_s}(t), & t+1 - \xi_{n_s}(t) = m_{n_s}(t) \\ 0, & t+1 - \xi_{n_s}(t) \neq m_{n_s}(t) \end{cases} \quad (3)$$

Therefore, the urgency level of physiological data received by the Sink at timeslot  $t$  is represented as

$$\Lambda_s(t) = \frac{1}{SN_s} \sum_{s=1}^S \sum_{n_s=1}^{N_s} \Lambda_{n_s}(t) \quad (4)$$

### B. Age of Information

AoI is a critical metric for the evaluation of time-sensitive networks handling physiological data in WBAN systems. In this context, AoI is defined as the time elapsed since the most recent packet was received by the Sink at time slot  $t$ , which was generated at time slot  $\bar{t}$ . Specifically, AoI is the difference between the current time slot  $t$  and the time slot  $\bar{t}$  when the last physiological data packet was generated by the sensor node, represented as  $A_{n_s}(t) = t - \bar{t}$ . Considering that the sampling and scheduling of physiological data from sensor node  $n_s$  occurs in time slot  $t$ , and sensor node  $n_s$  completes data transmission by the end of time slot  $t$ , then the AoI at time slot  $t+1$  for the Sink can be represented as  $m_{n_s}(t) = t+1 - \xi_{n_s}(t)$ , otherwise, it increases by one unit. The expression for  $A_{n_s}(t)$  is as follows

$$A_{n_s}(t+1) = \begin{cases} m_{n_s}(t), & t+1 - \xi_{n_s}(t) = m_{n_s}(t) \\ & \text{and } \varrho_{n_s}(t) = 1 \\ A_{n_s}(t) + 1, & t+1 - \xi_{n_s}(t) \neq m_{n_s}(t) \\ & \text{or } \varrho_{n_s}(t) = 0 \end{cases} \quad (5)$$

Therefore, the AoI of the physiological data received by the Sink at time slot  $t$  is given by

$$A_s(t) = \frac{1}{SN_s} \sum_{s=1}^S \sum_{n_s=1}^{N_s} A_{n_s}(t) \quad (6)$$

### C. Optimization Problem

The aim of this study is to devise a scheduling strategy that jointly optimizes the urgency of physiological data

and the AoI. Let  $\bar{\Lambda}_\pi$  and  $\bar{A}_\pi$  denote, respectively, the average urgency and average AoI of received physiological data under a feasible scheduling strategy  $\pi$ , where  $\bar{\Lambda}_\pi = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{E}[\Lambda_s(t)]$  and  $\bar{A}_\pi = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{E}[A_s(t)]$ . The performance metric utilized is the ULAR, which signifies the ratio of the urgency of the received data to the long-term average AoI. A higher ULAR value indicates that the received physiological data in each time slot is more urgent, leading to higher data freshness. The scheduling strategy  $\pi$  is represented as a feasible scheduling policy, where at the beginning of each time slot  $t$ ,  $\pi$  assigns channels to sensor nodes, samples data, and initiates data transmission. Consequently, the optimization objective can be formulated as follows

$$\begin{aligned} \mathbf{P1} : \max_{\pi} \quad & \mathcal{H} = \frac{\bar{\Lambda}_\pi}{\bar{A}_\pi} \\ \text{s.t.} \quad & C1 : \sum_{s=1}^S \sum_{n_s=1}^{N_s} \varrho_{n_s}(t) \leq C, \quad \forall t \\ & C2 : \sum_{n_s=1}^{N_s} \varrho_{n_s}(t) \leq 1, \quad \forall s, t \end{aligned} \quad (7)$$

where  $C1$  denotes the constraint on the number of channels, ensuring that where at most  $C$  channels are allocated per time slot for uplink transmission of physiological data from sensor nodes.  $C2$  represents the link contention constraint within each Sink, allowing at most one sensor node to transmit physiological data to the Sink in each time slot. Next, we present the proposed scheduling strategy.

## V. JOINT SCHEDULING POLICY

Let  $\mathcal{H}^*$  represent the optimized value of  $\mathcal{H}$ . Therefore,  $\mathcal{H}^* = \frac{\bar{\Lambda}_{\pi^*}}{\bar{A}_{\pi^*}}$ , where  $\bar{\Lambda}_{\pi^*}$  and  $\bar{A}_{\pi^*}$  are the values corresponding to the optimal policy  $\pi^*$ . For the nonlinear fractional programming problem defined by the objective function  $\mathbf{P1}$ , the optimality, denoted as  $\mathcal{H}^*$ , is attained if and only if the condition stipulated in [40] is satisfied

$$\max_{\pi} \{ \bar{\Lambda}_\pi - \mathcal{H}^* \bar{A}_\pi \} = \bar{\Lambda}_{\pi^*} - \mathcal{H}^* \bar{A}_{\pi^*} = 0 \quad (8)$$

Hence, problem  $\mathbf{P1}$  can be reformulated equivalently as equation (9)

$$\mathbf{P2} : \max_{\pi} \{ \bar{\Lambda}_\pi - \mathcal{H}^* \bar{A}_\pi \}, \quad \text{s.t.} \quad C1, C2 \quad (9)$$

Due to the unknown value of  $\mathcal{H}^*$ ,  $\mathbf{P2}$  remains challenging to solve. To simplify the solution, we introduce equation (10) as a surrogate for the unknown  $\mathcal{H}^*$  in  $\mathbf{P2}$

$$\mathcal{H}(t) = \frac{\frac{1}{t} \sum_{\tau=1}^t \mathbf{E}[\Lambda_s(\tau)]}{\frac{1}{t} \sum_{\tau=1}^t \mathbf{E}[A_s(\tau)]} = \frac{\bar{\Lambda}_\pi(t)}{\bar{A}_\pi(t)} \quad (10)$$

where  $\mathcal{H}(1) = 0$ , and the value of  $\mathcal{H}(t)$  depends on prior scheduling decisions. By utilizing  $\mathcal{H}(t)$  instead of  $\mathcal{H}^*$ , problem  $\mathbf{P2}$  can be reformulated as

$$\mathbf{P3} : \max_{\pi} \{ \bar{\Lambda}_\pi - \mathcal{H}(t) \bar{A}_\pi \}, \quad \text{s.t.} \quad C1, C2 \quad (11)$$

In problem  $\mathbf{P3}$ , the number of possible scheduling decisions for each timeslot is given by  $\sum_{c=1}^C \binom{S}{c} N_s^c$ , which results in a large action space due to the dynamic nature of AoI. This

poses a challenge for existing algorithms to efficiently solve the optimization problem. To tackle this issue, we decompose the scheduling policy into two strategies: SCA Strategy and NSS Strategy. The SCA strategy is responsible for determining the allocation of idle channels to Sink, while the NSS strategy focuses on selecting time slot for sensor nodes.

Algorithm 1 illustrates the NSS-SCA strategy. The algorithm begins by initializing all necessary parameters and networks, including the state, experience buffer, online network, target network, update interval, discount factor, and exploration rate (Step 1). For each time slot, the algorithm retrieves the current state, employs the NSS scheduling strategy to obtain the action and reward, and stores the transition information. Subsequently, it updates the Q-network using the SCA scheduling strategy (Steps 2-7). This iterative process in a dynamic environment progressively optimizes scheduling decisions, enhancing data transmission efficiency and resource utilization.

---

### Algorithm 1 NSS-SCA Strategy

---

- 1: **Initialize:** State  $\mathbf{s}(t)$ , experience buffer, the online network  $Q$  with weights  $\delta$ , target network  $\bar{Q}$  with weights  $\bar{\delta}$ , update interval  $\mathbb{Z}$ , discount factor  $\phi$ , and exploration rate  $\epsilon$ .
  - 2: **for** each time slot  $t = 1, 2, \dots, T$  **do**
  - 3:   Obtain current state  $\mathbf{s}(t)$ ,
  - 4:   Execute **NSS Scheduling Strategy** (Algorithm 3) to get action  $w(t)$  and reward  $\varepsilon(t)$
  - 5:   Perform action  $w(t)$  and observe new state  $\mathbf{s}(t+1)$
  - 6:   Store transition  $(\mathbf{s}(t), w(t), \varepsilon(t), \mathbf{s}(t+1))$  in experience buffer
  - 7:   Execute **SCA Scheduling Strategy** (Algorithm 2) to update Q-network
  - 8: **end for**
- 

#### A. SCA Strategy

SCA strategy employs DRL techniques for optimal channel allocation among Sink. Specifically, it utilizes a method known as Dueling Double Deep Q-learning Network (D3QN), which learns the optimal policy through iterative interaction between the agent and the environment to maximize long-term discounted rewards. D3QN takes the system's state as input and generates Q-values for each state-action pair. Compared to traditional Deep Q-Networks (DQN), D3QN offers enhanced training stability by mitigating issues like Q-value overestimation, which can occur in standard DQN models. This results in more accurate value estimates and improved performance in decision-making tasks [14]. It enhances learning efficiency by enabling the agent to simultaneously learn the value of being in a specific state and the advantage of taking a particular action. Moreover, D3QN facilitates the discovery of optimal strategies by exploring the state-action space effectively, utilizing a combination of the state-action value function and the advantage function to output Q-values.

Our goal is to maximize the objective function. We define the system state, action and cost function as follows.

- **System state:** We define the state space of the system at time slot  $t$  as  $\mathbf{s}(t) = (\mathbf{A}(t), \hat{\boldsymbol{\xi}}(t), \mathbf{m}(t), \mathbf{l}(t), \mathcal{H}(t))$ , which

consists of five elements: freshness, sampling time slot, allocated time slots, urgency level, and optimal value. Each state vector is listed as follows

$$\mathbf{A}(t) = (\mathbf{A}_s(t))_{s \in \mathcal{X}} \quad (12)$$

$$\hat{\boldsymbol{\xi}}(t) = (\hat{\boldsymbol{\xi}}_s(t))_{s \in \mathcal{X}} \quad (13)$$

$$\mathbf{m}(t) = (\mathbf{m}_s(t))_{s \in \mathcal{X}} \quad (14)$$

$$\mathbf{l}(t) = (\mathbf{l}_s(t))_{s \in \mathcal{X}} \quad (15)$$

For each sub-state vector,  $\mathbf{A}_s(t)$ ,  $\hat{\boldsymbol{\xi}}_s(t)$ ,  $\mathbf{m}_s(t)$ , and  $\mathbf{l}_s(t)$  can be expressed as follows

$$\mathbf{A}_s(t) = [A_{n_1}(t), A_{n_2}(t), \dots, A_{n_s}(t)] \quad (16)$$

$$\hat{\boldsymbol{\xi}}_s(t) = (t - 1 + \hat{\xi}_{n_s}(t))_{s \in \mathcal{X}} \quad (17)$$

$$\mathbf{m}_s(t) = [m_{n_1}(t), m_{n_2}(t), \dots, m_{n_s}(t)] \quad (18)$$

$$\mathbf{l}_s(t) = [l_{n_1}(t), l_{n_2}(t), \dots, l_{n_s}(t)] \quad (19)$$

- **Action:** Let  $w(t) = (w_s(t))_{s \in \mathcal{X}}$  denote the channel allocation action at timeslot  $t$ , where  $w_s(t) = \{0, 1\}$  and  $w_s(t) = 1$  indicates the allocation of an idle channel to Sink at timeslot  $t$ . With  $C$  available channels, the feasible actions that can be selected at each decision are limited to at most  $\sum_{c=1}^C \binom{S}{c}$ , representing the size of the action space.
- **Cost function:** The reward obtained in timeslot  $t$  is then defined as

$$\varepsilon(t) = \sum_{s=1}^S \sum_{n_s=1}^{N_s} \Lambda_{n_s}(t) - \mathcal{H}(t) \sum_{s=1}^S \sum_{n_s=1}^{N_s} A_{n_s}(t) \quad (20)$$

The value function  $V(\mathbf{s}, w)$  and the advantage function  $\mathcal{A}(\mathbf{s}, w)$  are derived from the output of the hidden layer of the D3QN network. Subsequently, following equation (20) in [41], the Q-value for each state-action pair is determined

$$Q(\mathbf{s}, w) = V(\mathbf{s}) + \left( \mathcal{A}(\mathbf{s}, w) - \frac{1}{\sum_{c=1}^C \binom{S}{c}} \sum_w \mathcal{A}(\mathbf{s}, w) \right) \quad (21)$$

D3QN architecture comprises an online network  $Q$  and a target network  $\bar{Q}$  [14]. To improve training stability, the weight vectors of the target network  $\bar{Q}$ , denoted as  $\delta$  and  $\bar{\delta}$ , are periodically updated from the online network  $Q$  at intervals of  $\mathbb{Z}$  steps.  $Q(\mathbf{s}, w|\delta)$  and  $\bar{Q}(\mathbf{s}, w|\bar{\delta})$  represent the value functions for state-action pairs in the online network  $Q$  and the target network  $\bar{Q}$ , respectively. The target values are calculated as follows

$$\mathcal{F}(t) = \varepsilon(t) + \phi \bar{Q} \left( \mathbf{s}(t+1), \arg \max_w Q(\mathbf{s}(t+1), w|\delta_t) | \bar{\delta}_t \right) \quad (22)$$

where  $0 < \phi \leq 1$  is the discount factor, and the loss function is defined as

$$L(\delta_t) = (\mathcal{F}(t) - Q(\mathbf{s}(t), w(t)|\delta_t))^2 \quad (23)$$

to update the loss function, gradient descent is employed

$$\nabla_{\delta_t} L(\delta_t) = (\mathcal{F}(t) - Q(\mathbf{s}(t), w(t)|\delta_t)) \times \nabla_{\delta_t} (Q(\mathbf{s}(t), w(t)|\delta_t)) \quad (24)$$

During the learning process of D3QN, a balanced exploration-exploitation strategy is employed through an  $\epsilon$ -greedy approach to avoid local optima. With probability  $\epsilon$ , the agent randomly selects one of all feasible actions to explore the policy. For the remaining probability  $1 - \epsilon$ , the agent exploits by choosing the action with the minimum Q-value among the feasible actions [42]. To ensure sufficient exploration, an initially high value is assigned to the exploration parameter  $\epsilon$  at the outset of the training process. This value gradually decreases towards 0 as the number of iterations increases to facilitate convergence. Additionally, historical experiences are stored in an experience buffer to enhance training effectiveness. During each training step, a batch of historical experiences is randomly selected for learning.

Algorithm 2 illustrates the SCA strategy. This algorithm takes as input the current state, action, reward, and next state. It begins by randomly sampling a batch of experiences from the experience buffer (lines 1-2). For each sampled experience, the target value is computed to update the Q-network. Specifically, the loss function  $L(\delta_t)$ , which measures the discrepancy between the current network output and the target value, is calculated. The network weights  $\delta_t$  are then updated using gradient descent, and the target network is periodically synchronized with the updated weights (lines 3-7). Finally, the exploration parameter  $\epsilon$  is decayed according to a predefined schedule to ensure effective policy convergence and long-term optimization.

---

#### Algorithm 2 SCA Scheduling Strategy

---

- 1: **Input:** State  $\mathbf{s}(t)$ , action  $w(t)$ , reward  $\varepsilon(t)$ , next state  $\mathbf{s}(t+1)$
  - 2: Sample a batch from experience buffer
  - 3: **for** each sample  $(\mathbf{s}, w, \varepsilon, \mathbf{s}')$  **do**
  - 4:     Compute target value  $\mathcal{F}(t)$  by (22)
  - 5:     Compute loss function  $L(\delta_t)$  by (23)
  - 6:     Perform gradient descent to update  $\delta_t$  by (24)
  - 7:     **if** the current step is a multiple of  $\mathbb{Z}$  **then**
  - 8:         Update target network weights:  $\bar{\delta} = \delta$
  - 9:     **end if**
  - 10:     Decay the exploration parameter  $\epsilon$  according to the exploration schedule
  - 11: **end for**
- 

#### B. NSS Strategy

The NSS strategy devises a combined scheduling policy to optimize the objective after the SCA strategy determines the channel allocation at the Sink. This policy comprises scheduling selection decisions of the Sink aimed at allocating idle channels. In each schedulable time slot  $t$ , the NSS strategy selects scheduling decisions that maximize the expected value of  $\bar{\Lambda}_\pi(t+1) - \mathcal{H}(t) \bar{A}_\pi(t+1)$ . However, due to the transmission delay of physiological data, executing scheduling decisions may not immediately reduce the AoI in subsequent time slots. Therefore, directly maximizing the expected value of  $\bar{\Lambda}_\pi(t+1) - \mathcal{H}(t) \bar{A}_\pi(t+1)$  is infeasible. To address this issue, the expected reduction in AoI is designed to be related

to  $\eta_{n_s}(t)$ , where  $\eta_{n_s}(t) = m_{n_s}(t) - t + \xi_{n_s}(t)$ , and  $\eta_{n_s}(t)$  represents the remaining time slots until the corresponding AoI decreases. Assuming that the transmission of sensor nodes in time slot  $t$  can reduce the AoI by  $\alpha_{n_s}(t)$  in time slot  $t + 1$ , this is expressed as follows

$$\alpha_{n_s}(t) = \begin{cases} \frac{A_{n_s}(\xi_{n_s}(t))}{\eta_{n_s}(t)}, & \varrho_{n_s}(t) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

where  $A_{n_s}(\xi_{n_s}(t))$  represents the expected reduction in AoI following the completion of the corresponding physiological data packet transmission. Note that if the respective data packet does not complete uplink transmission in time slot  $t$ , then in time slot  $t + 1$ , the Sink's AoI does not actually decrease. Therefore,  $A_{n_s}(\xi_{n_s}(t))$  can be regarded as a virtual AoI reduction. The Sink's AoI then needs to be adjusted by the corresponding bias term  $\zeta_{n_s}(t + 1) = \alpha_{n_s}(t)$  in the subsequent time slot  $t + 2$ . If the corresponding physiological data packet completes uplink transmission in time slot  $t$ , then  $\alpha_{n_s}(t) = A_{n_s}(\xi_{n_s}(t))$  denotes the actual AoI reduction at the Sink in time slot  $t + 1$ , with a bias term of 0. This is expressed as follows

$$\zeta_{n_s}(t + 1) = \begin{cases} 0, & t + 1 - \xi_{n_s}(t) = m_{n_s}(t) \\ & \text{and } \varrho_{n_s}(t) = 1 \\ \alpha_{n_s}(t), & \text{otherwise} \end{cases} \quad (26)$$

Based on the above analysis, we can construct a virtual queue as defined in Equation

$$q_{n_s}(t + 1) = A_{n_s}(t) + 1 + \zeta_{n_s}(t) - \alpha_{n_s}(t) \quad (27)$$

According to Theorem 1 in [13], as  $t$  tends to infinity,  $\frac{1}{t} \sum_{\tau=1}^t A_{n_s}(\tau)$  and  $\frac{1}{t} \sum_{\tau=1}^t q_{n_s}(\tau)$  become equal. This implies that under any feasible scheduling strategy, the time-averaged virtual queue length is equivalent to the time-average AoI. Hence, we can substitute  $\bar{q}_\pi(t)$  for  $\bar{A}_\pi(t)$  in Equation (28), denoted as

$$\bar{q}_\pi(t) = \frac{1}{t} \sum_{\tau=1}^t \mathbf{E} \left[ \frac{1}{SN_s} \sum_{s=1}^S \sum_{n_s=1}^{N_s} q_{n_s}(\tau) \right] \quad (28)$$

It is also noteworthy that if a sensor node occupies the wireless channel for uplink transmission in time slot  $t$ , the urgent level  $\Lambda_{n_s}(t)$  of physiological data packets received by the Sink remains zero during the data transmission process, and only becomes  $l_{n_s}(t)$  after the completion of packet transmission in time slot  $t + m_{n_s}(t) - 1$ . It is important to emphasize that the urgent level  $l_{n_s}(t)$  remains constant during the transmission of physiological data packets. Therefore, to evaluate the scheduling strategy in the corresponding scheduling time slot  $\varrho_{n_s}(t)$ , an equivalent variable can be constructed to assign the urgent level of its physiological data packets ahead of time in time slot  $\varrho_{n_s}(t) + 1$ . The expression for this equivalent variable is as follows

$$\sigma_{n_s}(t + 1) = (1 - \varrho_{n_s}(t - 1))\varrho_{n_s}(t)l_{n_s}(t) \quad (29)$$

it can be observed that as  $t$  approaches infinity, we can obtain equation (30)

$$\begin{aligned} \bar{\sigma}_\pi(t) &= \frac{1}{t} \sum_{\tau=1}^T \mathbf{E} \left[ \frac{1}{SN_s} \sum_{s=1}^S \sum_{n_s=1}^{N_s} \sigma_{n_s}(\tau) \right] \\ &= \frac{1}{t} \sum_{\tau=1}^T \mathbf{E} \left[ \frac{1}{SN_s} \sum_{s=1}^S \sum_{n_s=1}^{N_s} l_{n_s}(\tau) \right] \end{aligned} \quad (30)$$

As each Sink's scheduling decision is mutually independent and unrelated to others, the NSS strategy needs to select the strategy combination in each time slot  $t$  that maximizes the value of  $\bar{\sigma}_\pi(t + 1) - \mathcal{H}(t)\bar{q}_\pi(t + 1)$ . Based on this principle, the NSS strategy pairs sensor nodes within the Sink with their corresponding idle channels as illustrated below.

$$n_s^* = \arg \max_{n_s \in \gamma_s} \sum_{d=1}^D p_{n_s}^d d + \mathcal{H}_s(t) \frac{A_{n_s}(t)}{\sum_{j=1}^Y p_{n_s}^j d_N} \quad (31)$$

Algorithm 3 outlines the detailed steps of the NSS strategy. Initially, the algorithm takes the current state  $\mathbf{s}(t)$  as input and selects an action using the  $\epsilon$ -greedy strategy (steps 1-2). Then, it schedules each sensor node to maximize the objective function (steps 3-7). Finally, it calculates the relevant values and outputs the action  $w(t)$  and reward  $\varepsilon(t)$  for the current time slot (steps 10-14).

The link selection process for each Sink operates independently and without interdependency. According to equations (25)-(27), during the scheduling process, when an idle channel is assigned to a Sink, the sensor node  $n_s$  with the highest value of  $\sum_{d=1}^D p_{n_s}^d d + \mathcal{H}_s(t) \frac{A_{n_s}(t)}{\sum_{j=1}^Y p_{n_s}^j d_N}$  is selected to maximize the objective function  $\bar{\sigma}_\pi(t + 1) - \mathcal{H}(t)\bar{q}_\pi(t + 1)$ . The time complexity for each Sink is primarily determined by the process of selecting maximum value. If a binary search method is used to find the maximum value, the time complexity becomes  $O(\log N_s)$ . Therefore, after  $T$  iterations, the total computational complexity of the NSS strategy for the entire system is  $O(T \log \hat{N}_s)$ , where  $\hat{N}_s$  represents the maximum number of sensor nodes served by each Sink, i.e.,  $\hat{N}_s = \max(N_1, N_2, \dots, N_S)$ .

Based on the above analysis, we incorporate the NSS strategy into the SCA framework. After each bandwidth allocation action, the Sink assigned to the bandwidth schedules sensor nodes for sampling and data transmission according to the NSS strategy, thereby constructing the proposed NSS-SCA strategy. Compared to directly applying the D3QN algorithm to solve problem P3, the NSS-SCA strategy reduces the feasible action space of the neural network from  $\sum_{c=1}^C \binom{S}{c} N_s^c$  to  $\sum_{c=1}^C \binom{S}{c}$ , significantly decreasing the problem's complexity. This enhancement enables the DRL-based scheduling strategy to solve problem P3 more efficiently, improving the overall performance of the algorithm.

## VI. SIMULATION RESULTS

The proposed scheme is simulated using Python 3.8.2. The computational environment includes a Windows 10 operating system, an Intel Core i7-10700 CPU, and 32GB of RAM. The NSS-SCA algorithm is implemented with PyTorch 1.6.0,

### Algorithm 3 NSS Scheduling Strategy

```

1: Input: Current state  $\mathbf{s}(t)$ 
2: Select action  $w(t)$  using  $\epsilon$ -greedy strategy
3: for each sensor node  $s = 1, 2, \dots, S$  do
4:   if  $w_s(t) = 1$  then
5:     Schedule node  $n_s^*$  to maximize  $\sum_{d=1}^D p_{n_s}^d d + \mathcal{H}_s(t) \frac{A_{n_s}(t)}{\sum_{j=1}^Y p_{n_s}^j d_N}$ 
6:   else
7:     Do not schedule sensor node
8:   end if
9: end for
10: Calculate AoI reduction  $\alpha_{n_s}(t)$ 
11: Compute bias term  $\zeta_{n_s}(t+1)$ 
12: Update virtual queue  $q_{n_s}(t+1)$ 
13: Calculate equivalent variable  $\sigma_{n_s}(t+1)$ 
14: Compute reward  $\varepsilon(t)$ 
15: Output: Action  $w(t)$ , reward  $\varepsilon(t)$ 
    
```

TABLE III: D3QN Hyper-Parameters

Parameter	Value	Parameter	Value
Discount factor $\phi$	0.98	Replay memory size	2000
Learning rate	0.0001	Initial exploration rate $\epsilon$	0.3
Number of timeslot $T$	$5 \times 10^4$	Batch size	32
Update step $\mathbb{Z}$	10	Activation function	ReLU

TABLE IV: Common Simulation Parameters

Parameter	Value	Parameter	Value
Frequency band	2.4 GHz	$R_s$	600 Kbps
Modulation	DBPSK	$M$	4
ACK policy	I-ACK	$D$	4
$pASlotRes$	500 $\mu$ s	$J$	5
$S_{PLCPHeader}$	4	$N_{pre}$	90 bit
$N_j$	50-250 bytes	$N_{Header}$	31 bit

featuring a network architecture with two hidden layers containing 80 and 40 ReLU activation units, respectively. The specific system parameters utilized in the analysis are detailed in Table III.

To validate the effectiveness of the proposed scheme in optimizing the three key performance indicators (data urgency, AoI, and ULAR), we first compare two scheduling mechanisms: the AoI-aware scheduling mechanism (AoI-AS) [38], which focuses on AoI, and the i-MAC sorting algorithm (i-MAC) [33], which considers only data urgency. Specifically, Fig. 3 illustrates the performance differences between these two algorithms as the number of sensor nodes increases in a single Sink scenario. We set the number of WBAN users to  $S = 1$  and the available channels to  $C = 1$ , with the number of sensor nodes ranging from  $n_s \in \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ . The data urgency level was set to  $D = 4$ , and the probability parameters  $p_{n_s}^d$  and  $p_{n_s}^{Y_j}$  were generated using random functions. Other relevant parameters are shown in Table IV.

As shown in Fig. 3(a), when considering only data urgency, the data urgency values of all three algorithms decrease as the number of sensor nodes  $N_s$  increases. This decline occurs because the increase in sensor nodes reduces each

node's scheduling opportunities. However, the i-MAC from [33] outperforms the other methods in terms of data urgency optimization, followed by our proposed algorithm, while the AOI-AS from [38] performs the worst. This is because the i-MAC algorithm is specifically designed to optimize data urgency, whereas the AOI-AS does not consider data urgency, leading to suboptimal results. Our proposed algorithm achieves a balanced performance by considering both AoI and data urgency simultaneously.

Fig. 3(b) presents the comparison of AoI performance, with the same experimental parameters as Fig. 3(a). As the number of sensor nodes  $N_s$  increases, the AoI values of all algorithms exhibit an upward trend. However, the AOI-AS achieves the lowest AoI values due to its explicit consideration of data expiration rates, while the i-MAC maintains consistently higher AoI values as it disregards AoI. Our proposed algorithm strikes a balance between AoI and data urgency, positioning its AoI performance at an intermediate level, thereby achieving a more balanced overall performance.

Fig. 3(c) shows the comparison of average ULAR values across different numbers of sensor nodes, with experimental settings consistent with those in Fig. 3(a) and Fig. 3(b). As the number of sensor nodes increases, the average ULAR values of all algorithms gradually decline, which is expected since more nodes reduce the scheduling opportunities for individual nodes to perform uplink transmission. Nonetheless, our proposed strategy demonstrates superior average ULAR performance, primarily due to our joint optimization of data urgency and AoI, which significantly enhances the overall system performance. Compared with the i-MAC, which focuses only on urgency, and the AOI-AS, which prioritizes AoI, our method exhibits a more comprehensive performance advantage in WBAN system.

At the same time, we compare the proposed scheduling strategy with several existing mechanisms to analyze their respective strengths, weaknesses, and distinguishing characteristics. This comparative analysis not only demonstrates the effectiveness of our approach but also highlights the performance differences of each strategy in specific application scenarios.

- 1) **Greedy Strategy:** This strategy prioritizes scheduling the sensor node with the highest  $A_{n_s}(t)$  value at each time slot  $t$  to ensure optimal data freshness.
- 2) **Max-Ratio (MR) Strategy:** This strategy accounts for uplink transmission delay by prioritizing the scheduling of the sensor node with the highest  $\frac{A_{n_s}(t)}{\sum_{j=1}^Y p_{n_s}^j d_N}$  value at each time slot  $t$ , aiming to optimize data freshness under transmission delay conditions.
- 3) **MRUD Strategy:** The Maximum Ratio based on Urgency Degree (MRUD) strategy further considers the urgency of physiological data, building upon the MR strategy. It prioritizes scheduling the sensor node with the highest  $\frac{A_{n_s}(t) \sum_{d=1}^D p_{n_s}^d d}{\sum_{j=1}^Y p_{n_s}^j d_N}$  value at each time slot  $t$ , taking into account the impact of data urgency on the average AoI.

Fig. 4 shows the comparison of data urgency levels under

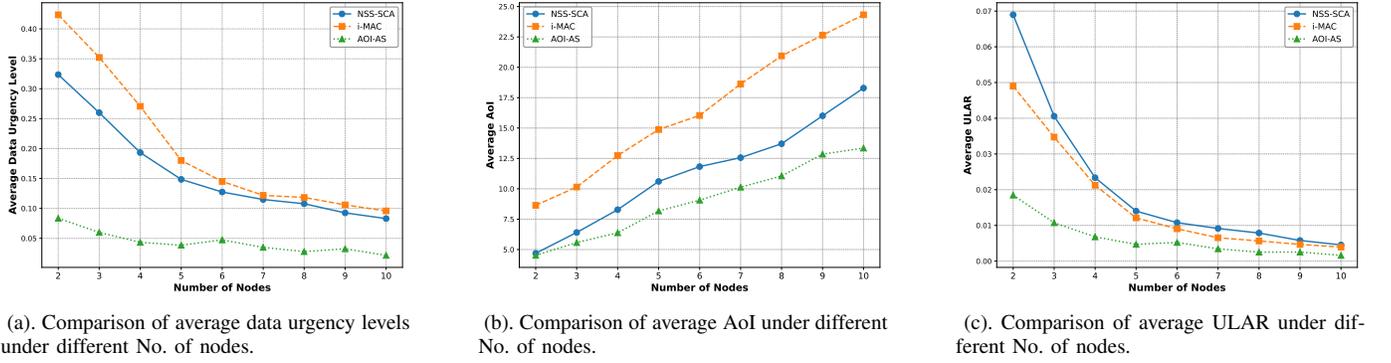


Fig. 3: Comparison of algorithm performance in data urgency, AoI and ULAR.

different strategies with different numbers of sensor nodes per Sink. The experimental setup includes  $S = 3$  Sinks,  $C = 2$  available channels, and varying numbers of sensor nodes in each Sink  $n_s \in \{3, 4, 5, 6, 7, 8, 9\}$ . The level of urgency is divided into  $D = 4$  levels. The probability parameters  $p_{n_s}^d$  and  $p_{n_s}^{Y_j}$  for the sensor nodes are generated using a random function, with other relevant parameters detailed in Table IV. The figure illustrates that as the number of sensor nodes  $N_s$  increases across all algorithms, the data urgency levels gradually decrease. This decline is attributed to the increased number of sensor nodes, which reduces the scheduling opportunities for each individual node.

Furthermore, the NSS-SCA strategy demonstrates superior performance in terms of average data urgency over time compared to other strategies. The greedy and MR strategies do not consider data urgency in the system, so their performance in terms of data urgency is similar to random decisions, resulting in the lowest time-averaged urgency among the four strategies. Although the MRUD strategy considers data urgency, the dynamic changes in the AoI affect the corresponding scheduling decisions, thereby affecting the dynamic changes in the received data urgency. Consequently, the MRUD strategy achieves a lower time-averaged data urgency compared to the NSS-SCA strategy because it does not consider the impact of these dynamic changes.

Additionally, we examined the effect of varying the number of Sinks on data urgency, as shown in Fig. 5. In the setting where each Sink has 4 sensor nodes, the NSS-SCA strategy again demonstrates superior performance in terms of average data urgency over time. In summary, by thoroughly considering data urgency and system dynamics, the NSS-SCA strategy exhibits high efficiency and superior performance in various settings.

In Fig. 6, we compare the average AoI for different strategies under different numbers of sensor nodes per Sink. The experimental parameters are the same as in Fig. 4. It is evident from the figure that as the number of sensor nodes  $N_s$  increases, the average AoI also increases for all algorithms. However, the NSS-SCA strategy consistently shows superior performance in terms of average AoI compared to other strategies. This is primarily because the Greedy strategy considers the reduction in AoI that would result from completing the

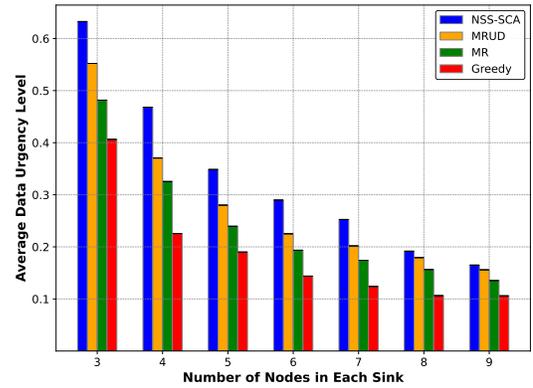


Fig. 4: Comparison of average data urgency levels under different No. of nodes in each Sink.

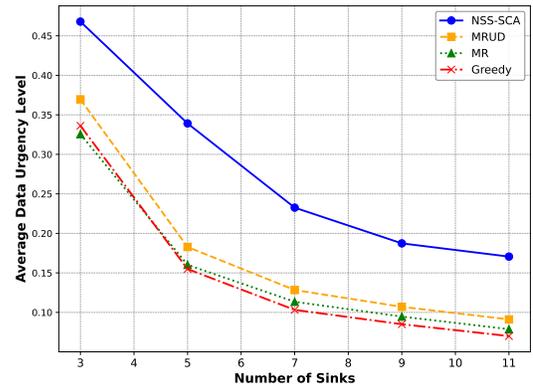


Fig. 5: Comparison of average data urgency levels under different No. of Sinks.

data transmission, but neglects the effect of the transmission delay on AoI, resulting in poorer AoI performance. The MR and MRUD strategies fail to account for the dynamic changes in AoI, resulting in their lower time-averaged AoI performance compared to the NSS-SCA strategy, which is based on deep reinforcement learning.

Additionally, we examined the effect of varying the number of Sinks on the average AoI, as shown in Fig. 7. In a setting where each Sink has 4 sensor nodes, the NSS-SCA

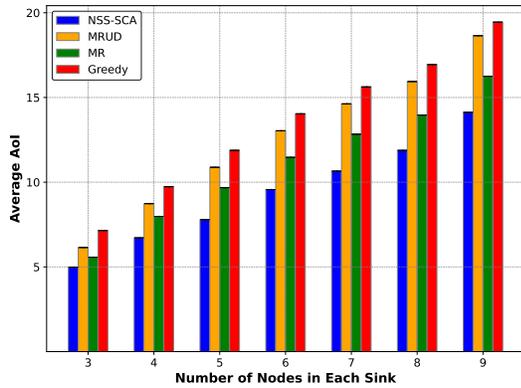


Fig. 6: Comparison of average AoI under different No. of nodes in each Sink.

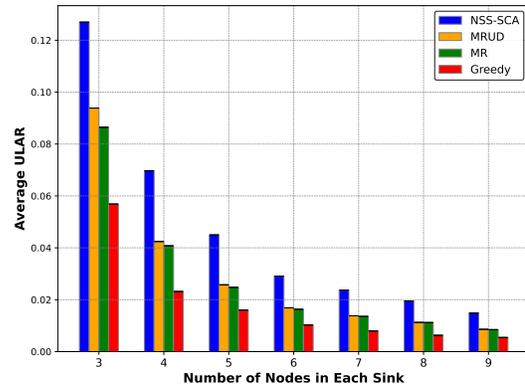


Fig. 8: Comparison of average ULAR under different No. of nodes in each Sink.

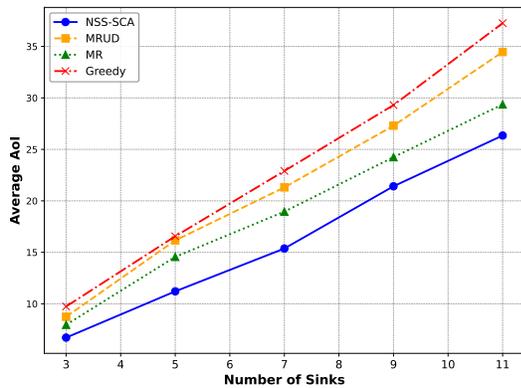


Fig. 7: Comparison of average AoI under different No. of Sinks.

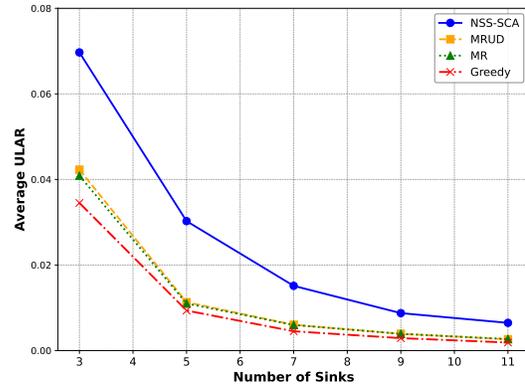


Fig. 9: Comparison of average ULAR under different No. of Sinks.

strategy again demonstrates superior average AoI performance. Overall, the NSS-SCA strategy exhibits high efficiency and superior performance in various settings due to its comprehensive consideration of data urgency and system dynamics. This highlights the importance of accounting for dynamic changes in AoI to ensure optimal system performance, and underscores the superiority of the NSS-SCA strategy in achieving the best time-averaged AoI.

Fig. 8 illustrates the comparative results of the average ULAR values for each unit sink under different strategies. The experimental parameters remain the same as in Fig. 4. Observing the graph, we note a gradual decrease in the average ULAR as the number of sensor nodes  $N_s$  increases across all algorithms. This decline stems from the increased number of sensor nodes, which leads to a higher count of sensor nodes per unit Sink, consequently reducing the opportunities for each sensor node to be scheduled for uplink transmission.

However, the NSS-SCA strategy outperforms other strategies in terms of average ULAR performance. This superiority can be attributed to two main factors: Firstly, the NSS-SCA strategy directly optimizes the time-averaged ULAR. Hence, compared to the myopic Greedy and MR strategies, which solely consider AoI, NSS-SCA yields superior ULAR values. Secondly, while the MRUD strategy jointly considers data

urgency and AoI, it neglects the dynamic nature of AoI, resulting in inferior performance in terms of time-averaged ULAR compared to the NSS-SCA strategy based on DRL.

Similarly, a horizontal comparison of the effect of varying the number of Sinks on the average ULAR is shown in Fig. 9. In the setting where the number of sensor nodes per unit Sink is 4, the NSS-SCA strategy also shows superior performance. In summary, the NSS-SCA strategy exhibits high efficiency and superior performance in various settings due to its comprehensive consideration of data urgency and system dynamics.

## VII. CONCLUSION AND FUTURE WORK

To support long-term, efficient electronic health monitoring while integrating with CE, this paper proposes a joint optimization scheduling strategy based on data urgency and freshness. Specifically, the strategy consists of two components: SCA and NSS. Our DRL-based approach efficiently handles large action spaces in channel allocation and timeslot selection, leading to improved scheduling performance. Simulation results demonstrate that the proposed strategy outperforms existing methods by enhancing the average urgency of received data and reducing the average AoI.

While the proposed joint optimization strategy has demonstrated significant performance gains in simulated environments, further validation under real-world conditions is essential to fully assess its practical applicability. WBAN systems operate in dynamic and often unpredictable environments, where factors such as interference, energy constraints, and patient mobility can significantly impact performance. Implementing the proposed method in real-world WBAN setups, using commercially available wearable sensors and healthcare systems, is a critical next step. However, this requires access to specialized hardware and collaboration with healthcare institutions or research laboratories. Future work will focus on addressing these challenges by designing experiments that evaluate the method's performance in practical scenarios. Metrics such as data transmission latency, energy efficiency, and robustness to interference will be key to understanding its behavior under realistic conditions. Furthermore, we plan to explore how the proposed approach can be adapted to account for the dynamic nature of WBANs. For instance, adjustments in the scheduling algorithm could ensure robustness in the presence of node mobility or unpredictable network conditions. By addressing these real-world challenges, we aim to further refine the proposed optimization strategy and enhance its applicability in practical WBAN environments.

## REFERENCES

- [1] C. K. Wu, C.-T. Cheng, Y. Uwate, G. Chen, S. Mumtaz, and K. F. Tsang, "State-of-the-art and research opportunities for next-generation consumer electronics," *IEEE Trans. Consum. Electron.*, vol. 69, no. 4, pp. 937–948, Nov. 2023.
- [2] U. Islam, G. Mehmood, A. A. Al-Atawi, F. Khan, H. S. Alwageed, and L. Cascone, "Neurohealth guardian: A novel hybrid approach for precision brain stroke prediction and healthcare analytics," *J. Neurosci. Methods.*, vol. 409, p. 110210, Jul. 2024.
- [3] M. Shahid, M. Tariq, Z. Iqbal, H. M. Albarakati, N. Fatima, M. A. Khan, and M. Shabaz, "Link-quality-based energy-efficient routing protocol for WSN in IoT," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 4645–4653, Feb. 2024.
- [4] S. P. Mohanty and F. Pescador, "Introduction consumer technologies for smart healthcare," *IEEE Trans. Consum. Electron.*, vol. 67, no. 1, pp. 1–2, Feb. 2021.
- [5] A. Javadpour, A. K. Sangaiah, F. Ja'fari, P. Pinto, H. Memarzadeh-Tehran, S. Rezaei, and F. Saghafi, "Toward a secure industrial wireless body area network focusing MAC layer protocols: An analytical review," *IEEE Trans. Ind. Inform.*, vol. 19, no. 2, pp. 2028–2038, Feb. 2023.
- [6] R. Cavallari, F. Martelli, R. Rosini, C. Buratti, and R. Verdone, "A survey on wireless body area networks: Technologies and design challenges," *IEEE Commun. Surv. Tutor.*, vol. 16, no. 3, pp. 1635–1657, Feb. 2014.
- [7] G. Mehmood, M. Z. Khan, A. K. Bashir, Y. D. Al-Otaibi, and S. Khan, "An efficient QoS-based multi-path routing scheme for smart healthcare monitoring in wireless body area networks," *Comput. Electr. Eng.*, vol. 109, p. 108517, May 2023.
- [8] Z. Lin, X. Liu, H. Zhou, and J. Wu, "Adaptive time-varying routing for energy saving and load balancing in wireless body area networks," *IEEE Trans. Mobile Comput.*, vol. 23, no. 1, pp. 90–101, Oct. 2024.
- [9] S. Verma, Y. Kawamoto, Z. M. Fadlullah, H. Nishiyama, and N. Kato, "A survey on network methodologies for real-time analytics of massive IoT data and open research issues," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 3, pp. 1457–1477, Apr. 2017.
- [10] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Orlando, FL, USA, Apr. 2012, pp. 2731–2735.
- [11] Kahraman, A. Köse, M. Koca, and E. Anarım, "Age of information in internet of things: A survey," *IEEE Internet Things J.*, vol. 11, no. 6, pp. 9896–9914, Mar. 2024.
- [12] G. Z. X. W. Y. Liu, X. Wang and Z. Ning, "An aoi-aware data transmission algorithm in blockchain-based intelligent healthcare systems," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 1180–1190, Feb. 2024.
- [13] X. Xie, H. Wang, and X. Liu, "Scheduling for minimizing the age of information in multisensor multiserver industrial internet of things systems," *IEEE Trans. Ind. Inform.*, vol. 20, no. 1, pp. 573–582, Jan. 2024.
- [14] H. Wang, X. Xie, and J. Yang, "Optimizing average age of information in industrial IoT systems under delay constraint," *IEEE Trans. Ind. Inform.*, vol. 19, no. 10, pp. 10244–10253, Jan. 2023.
- [15] H. Tang, J. Wang, L. Song, and J. Song, "Minimizing age of information with power constraints: Multi-user opportunistic scheduling in multi-state time-varying channels," *IEEE J. Select. Areas Commun.*, vol. 38, no. 5, pp. 854–868, Mar. 2020.
- [16] B. Yao, H. Gao, Y. Zhang, J. Wang, and J. Li, "Maximum AoI minimization for target monitoring in battery-free wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 22, no. 8, pp. 4754–4772, Mar. 2023.
- [17] E. Fountoulakis, T. Charalambous, A. Ephremides, and N. Pappas, "Scheduling policies for AoI minimization with timely throughput constraints," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 3905–3917, May. 2023.
- [18] A. Zakeri, M. Moltafet, M. Leinonen, and M. Codreanu, "Minimizing the AoI in resource-constrained multi-source relaying systems: Dynamic and learning-based scheduling," *IEEE Trans. Wireless Commun.*, vol. 23, no. 1, pp. 450–466, Jun. 2024.
- [19] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM J. Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, Oct. 2018.
- [20] J. Li, Y. Zhou, and H. Chen, "Age of information for multicast transmission with fixed and random deadlines in IoT systems," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8178–8191, Mar. 2020.
- [21] Z. Zhang, J. Lin, and Z. Li, "Resource scheduling based on multi-factor priority for high performance requirements in WBANs," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 6487–6492.
- [22] G. Sun, K. Wang, H. Yu, X. Du, and M. Guizani, "Priority-based medium access control for wireless body area networks with high-performance design," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5363–5375, Feb. 2019.
- [23] K. Das, S. Moulik, and C.-Y. Chang, "Priority-based dedicated slot allocation with dynamic superframe structure in IEEE 802.15.6-based wireless body area networks," *IEEE Internet Things J.*, vol. 9, no. 6, pp. 4497–4506, Aug. 2022.
- [24] B. Liu, Z. Yan, and C. W. Chen, "Medium access control for wireless body area networks with QoS provisioning and energy efficient design," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 422–434, Mar. 2017.
- [25] Z. Ullah, I. Ahmed, F. A. Khan, M. Asif, M. Nawaz, T. Ali, M. Khalid, and F. Niaz, "Energy-efficient harvested-aware clustering and cooperative routing protocol for WBAN (E-HARP)," *IEEE Access*, vol. 7, pp. 100036–100050, Jul. 2019.
- [26] D. Wu, B. Yang, H. Wang, D. Wu, and R. Wang, "An energy-efficient data forwarding strategy for heterogeneous WBANs," *IEEE Access*, vol. 4, pp. 7251–7261, Sep. 2016.
- [27] Z. Liu, B. Liu, and C. W. Chen, "Buffer-aware resource allocation scheme with energy efficiency and QoS effectiveness in wireless body area networks," *IEEE Access*, vol. 5, pp. 20763–20776, Oct. 2017.
- [28] B. Liang, M. S. Obaidat, X. Liu, H. Zhou, and M. Dong, "Resource scheduling based on priority ladders for multiple performance requirements in wireless body area networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7027–7036, Jul. 2021.
- [29] B.-S. Kim and K.-I. Kim, "A priority-based dynamic link scheduling algorithm using multi-criteria decision making in wireless body area networks," in *Proc. IEEE 28th Int. Conf. (MASCOTS)*, Nice, France, Nov. 2020, pp. 1–8.
- [30] K. S. Deepak and A. V. Babu, "Improving reliability of emergency data frame transmission in IEEE 802.15.6 wireless body area networks," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2082–2093, Sep. 2018.
- [31] S. Misra, S. Moulik, and H.-C. Chao, "A cooperative bargaining solution for priority-based data-rate tuning in a wireless body area network," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2769–2777, May. 2015.
- [32] S. Moulik, S. Misra, and D. Das, "AT-MAC: Adaptive MAC-frame payload tuning for reliable communication in wireless body area networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 6, pp. 1516–1529, Jun. 2017.

- [33] S. Misra, P. K. Bishoyi, and S. Sarkar, "i-MAC: In-body sensor MAC in wireless body area networks for healthcare IoT," *IEEE Syst. J.*, vol. 15, no. 3, pp. 4413–4420, Sep. 2021.
- [34] B. Liang, X. Liu, H. Zhou, V. C. M. Leung, A. Liu, and K. Chi, "Channel resource scheduling for stringent demand of emergency data transmission in WBANs," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2341–2352, Apr. 2021.
- [35] D. D. Olatinwo, A. M. Abu-Mahfouz, G. P. Hancke, and H. C. Myburgh, "Energy efficient priority-based hybrid MAC protocol for IoT-enabled WBAN systems," *IEEE Sensors J.*, vol. 23, no. 12, pp. 13 524–13 538, Jun. 2023.
- [36] D. D. Olatinwo, A. M. Abu-Mahfouz, and Z. Hancke, "Energy-efficient multichannel hybrid MAC protocol for IoT-enabled WBAN systems," *IEEE Sensors J.*, vol. 23, no. 22, pp. 27 967–27 983, Nov. 2023.
- [37] Y.-H. Xu, L. Suo, W. Zhou, and G. Yu, "A graph-learning-inspired resource optimization for digital-twin-empowered wireless body area networks," *IEEE Sens. Lett.*, vol. 7, no. 11, pp. 1–4, Nov. 2023.
- [38] Z. Zhou, F. Ke, K. Liu, and J. Tan, "Age of information-aware scheduling for dynamic wireless body area networks," *IEEE Sensors J.*, vol. 23, no. 16, pp. 17 832–17 841, Aug. 2023.
- [39] Z. Ning, P. Dong, X. Wang, X. Hu, L. Guo, B. Hu, Y. Guo, T. Qiu, and R. Y. K. Kwok, "Mobile edge computing enabled 5G health monitoring for Internet of Medical Things: A decentralized game theoretic approach," *IEEE J. Select. Areas Commun.*, vol. 39, no. 2, pp. 463–478, Feb. 2021.
- [40] Y. Li, M. Sheng, Y. Shi, X. Ma, and W. Jiao, "Energy efficiency and delay tradeoff for time-varying and interference-free wireless networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 5921–5931, Nov. 2014.
- [41] H. Yang, J. Zhao, K.-Y. Lam, S. Garg, Q. Wu, and Z. Xiong, "Deep reinforcement learning based resource allocation for heterogeneous networks," in *Proc. 17th Int. Conf. Wireless Mobile Comput., Netw. and Commun. (WiMob)*, Bologna, Italy, Oct. 2021, pp. 253–258.
- [42] N. Kiran, C. Pan, S. Wang, and C. Yin, "Joint resource allocation and computation offloading in mobile edge computing for SDN based wireless networks," *J Commun Netw-S KOR.*, vol. 22, no. 1, pp. 1–11, Feb. 2020.



**Zheng Zhang** (Member, IEEE) received the M.S. degree in Electronics and Communications Engineering from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2015. He is currently pursuing a Ph.D. degree in Communication and Information Engineering at the same institution while serving as a Visiting Ph.D. Scholar at Brunel University of London, UK. His research interests include WBANs, smart healthcare, wearable devices, and human-computer interaction.



**Zhangyong Li** received the Ph.D. degree from Chongqing Medical University, Chongqing, China, in 2004. He is currently a Professor with the School of Bioinformatics, Chongqing University of Posts and Telecommunications, Chongqing. His main research interests include biomedical information processing, digital medical instrument design, electromagnetic radiation protection technology, environmental assessment, and reliability of medical instruments.



**Chun Sing Lai** (Senior Member, IEEE) received the B.Eng. (First Class Hons.) in electronic and electrical engineering from Brunel University of London, UK, in 2013, and the D.Phil. degree in engineering science from the University of Oxford, UK, in 2019. He is currently a Senior Lecturer with the Department of Electronic and Electrical Engineering and Course Director of MSc Electric Vehicle Systems at Brunel University of London. His research interests are in power system optimization and electric vehicle systems. Dr. Lai was a Technical Program Co-Chair for 2022 IEEE International Smart Cities Conference. He is the Vice-Chair of the IEEE Smart Cities Publications Committee. He is an Associate Editor for IEEE Transactions on Systems, Man, and Cybernetics: Systems, IEEE Transactions on Consumer Electronics and IET Energy Conversion and Economics. He is the Working Group Chair for IEEE P2814 and P3166 Standards, an Associate Vice President, Systems Science and Engineering of the IEEE Systems, Man, and Cybernetics Society (IEEE/SMCS) and Co-Chair of the IEEE SMC Intelligent Power and Energy Systems Technical Committee. He is a recipient of the 2022 Meritorious Service Award from the IEEE SMC Society for "meritorious and significant service to IEEE SMC Society technical activities and standards development". He is an IET Member, Chartered Engineer, and Fellow of the Higher Education Academy.



**Ruiheng Wu** (Senior Member, IEEE) received his BEng and MEng degrees from Tianjin University, China, in 1982 and 1986 respectively. Following 10 years as a faculty member in the Department of Electronic Engineering at Tianjin University and two years research at the City University of Hong Kong, he joined the Analogue Circuit Design Research Group at Oxford Brookes University, U.K., as a PhD student in January 1999, where he completed his PhD thesis entitled 'Design of Wide Bandwidth High Linearity Amplifiers'. Currently, Dr Wu is a senior lecturer in Electronic and Electrical Engineering at Brunel University London. Dr Wu's general research interests are in the area of analogue signal processing and integrated circuit design; RF circuits and systems; remote health monitoring and applications of modern electronics in life search and rescue. In recent years, as a principal investigator, he has been involved in many industrial and research projects supported by EU, NHS, HEFCE etc. and has several tens publications in his research areas.



**Xinwei Li** received her B.S. and Ph.D. degrees from Beihang University in 2011 and 2018, respectively. In 2016, she was invited as a visiting student to the Beckman Institute, University of Illinois Urbana-Champaign, USA. From 2018 to 2021, she served as an Assistant Professor at the School of Bioinformatics, Chongqing University of Posts and Telecommunications, China. From 2022 to 2023, she was promoted to Associate Professor. In 2024, she became a Wenfeng Professor at the School of Life Health Information Science and Engineering, Chongqing University of Posts and Telecommunications, China. Her research focuses on medical image processing, brain science, and artificial intelligence.



**Jinzhao Lin** received the master's and Ph.D. degrees from Chongqing University, Chongqing, China, in 1995 and 2001, respectively. He is currently a Professor with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing. His main research interests include theoretical, technical, methodological, and applied research on smart medical devices and their intersections.



**Xinxing Ren** (Student Member, IEEE) received the B.Eng. degree (with First-Class Hons.) in electronic and electrical engineering from the Brunel University of London, Uxbridge, U.K., in 2022, where he is currently pursuing the Ph.D. degree. He is a member of the Brunel Interdisciplinary Power Systems Research Centre, Brunel University of London. His research interests include connected vehicles, wireless body area networks, and embodied intelligence.



**Huidong Qiao** received the M.S. degree from Central South University and Ph.D. from the National University of Defense Technology. His research interests include information security, cryptography, wireless body area networks, and embodied intelligence.