# Medium Access Control for Wireless Body Area Networks with QoS Provisioning and Energy Efficient Design

Bin Liu, *Member, IEEE,* Zhisheng Yan, and Chang Wen Chen, *Fellow, IEEE*

**Abstract**—With the promising applications in e-Health and entertainment services, wireless body area network (WBAN) has attracted significant interest. One critical challenge for WBAN is to track and maintain the quality of service (QoS), e.g., delivery probability and latency, under the dynamic environment dictated by human mobility. Another important issue is to ensure the energy efficiency within such a resource-constrained network. In this paper, a new medium access control (MAC) protocol is proposed to tackle these two important challenges. We adopt a TDMA-based protocol and dynamically adjust the transmission order and transmission duration of the nodes based on channel status and application context of WBAN. The slot allocation is optimized by minimizing energy consumption of the nodes, subject to the delivery probability and throughput constraints. Moreover, we design a new synchronization scheme to reduce the synchronization overhead. Through developing an analytical model, we analyze how the protocol can adapt to different latency requirements in the healthcare monitoring service. Simulations results show that the proposed protocol outperforms CA-MAC and IEEE 802.15.6 MAC in terms of QoS and energy efficiency under extensive conditions. It also demonstrates more effective performance in highly heterogeneous WBAN.

**Index Terms**—Context aware, energy-efficient, MAC protocol, QoS, slot allocation, WBAN

## 1 Introduction

Thanks to advanced radio technologies and sensor hardware, wireless body area network (WBAN) has become practically feasible. WBAN has attracted strong interest due to its potential economic impacts in numerous applications, such as vital sign monitoring, interactive gaming, and telemedicine. The system architecture of a typical WBAN-based system is shown in Fig. 1. WBAN is composed of multiple on-body sensor nodes, and a personal server (PS) equipped on human body. Since the communication range of sensor nodes is 2-3 m [2], [3], these nodes are usually organized as a star-topology mobile network. Thus the body sensor nodes can gather a variety of physiological information, as well as delivering them to the PS for displaying and processing. The quality of service (QoS) requirement of the data delivery is adjusted by the PS when it detects the change of monitoring contexts. Alternatively, the emergency users, e.g., doctors, and normal users, e.g., family members, can proactively send requests to the PS to change the delivery requirement. This way, body area information can be effectively collected without interrupting people's normal activities.

One critical research task for WBAN is the medium access control (MAC), that is, how should the sensor nodes access the limited wireless channel resources in order to ensure efficient and reliable data transmission. Although MAC protocols for wireless sensor network (WSN) have been widely studied, they cannot be directly applied in WBAN applications. First, a WSN is relatively stationary. However, a WBAN worn by a person is highly mobile, which results in the unique features of on-body channel and traffic [4]. Second, MAC protocols for WSN usually target a large number of sensors, which makes them distributed in nature. Nevertheless, this is not the case for WBAN [4] and thus the design principle can be totally different. We identify the unique challenges of WBAN MAC design as follows.

- Ensuring reliable transmission under the lossy on-body channel: Body movement may cause *deep fading* that immediately leads to frame loss [5], e.g., when the moving arms or legs obstruct a communication path. This is further complicated by the fact that the deep fading of on-body channel lasts much longer (up to 400 ms) than traditional wireless networks [6]. Hence, an accurate yet mathematically tractable channel model is needed to design the MAC and thus guarantee the delivery probability.
- Allocating appropriate channel resources for heterogeneous and dynamic traffic: The data traffic

*The preliminary results were presented at HealthCom 2012 [1].*
*Bin Liu and Zhisheng Yan contributed equally to this work (co-first author in alphabetical order).*
*Bin Liu is with the Key Laboratory of Electromagnetic Space Information, Chinese Academy of Sciences, P. R. China and School of Information Science and Technology, University of Science and Technology of China, P. R. China (email: flowice@ustc.edu.cn)*
*Zhisheng Yan and Chang Wen Chen are with Computer Science and Engineering Department, State University of New York at Buffalo, Buffalo, NY (email: {zyan3, chencw}@buffalo.edu)*
*This work is supported by the National Natural Science Foundation of China (Grant No. 61202406).*

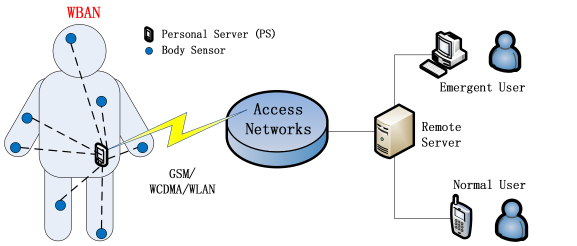



Fig. 1. General architecture of WBAN-based system.

of different sensors are dynamically dictated by the sensor functionalities, body movements, and environment status [7]. For example, when a cardiomyopathy patient is exercising, sampling rate of heart-related data should be increased and more data should be sent with high priority in this abnormal context. Meanwhile, the QoS demand for heart-unrelated data, such as electromyography (EMG), can be relaxed.

- Maximizing the energy efficiency for the power-hungry body sensor nodes: Body sensors usually need to work for months or even years without interruption. It is a non-trivial task to strike a balance between minimized energy and guaranteed QoS. For example, more transmission might result in higher throughput and lower average delay, but will inevitably consume more energy.

Recently, MAC protocols specifically for WBAN have been proposed to address the dynamic traffic pattern [8], [9] or the energy efficiency [10], [11]. However, they focus on tackling each individual WBAN issue. The unique features of WBAN channel, traffic and energy have not been comprehensively studied in the context of MAC designs. The goal of this research is to jointly explore the unique features of WBAN at the physical and application layer, and to design a MAC protocol for star-topology WBANs that can guarantee both QoS and energy efficiency.

In particular, we propose a general time division multiple access (TDMA) based MAC protocol that can flexibly address the QoS of heterogeneous sensors in different monitoring contexts (see Section 4.3 for definition) and can maximize the energy efficiency. The proposed MAC protocol can adapt to the time-varying channel and traffic of WBAN by dynamically optimizing the *transmission schedule*, i.e., *transmission order* and *transmission duration*, of each sensor node. Besides, it proactively minimizes the energy consumption of sensor nodes while ensuring desirable delivery probability and throughput. To further reduce energy consumption, we embed a new synchronization scheme in the proposed MAC protocol without impacting its optimality. Moreover, we develop an analytical model to investigate the latency performance of the proposed MAC protocol. We explore the parameters space and analyze how the proposed protocol can adapt to different latency requirements, which theoretically guides the real-world deployment. Finally, we evaluate the QoS and energy of the proposed MAC protocol and therein validate the designs and analysis by extensive simulations. Our results show that the proposed MAC protocol is especially suitable for highly heterogeneous WBAN and that we should choose the largest possible superframe length to maximize energy efficiency when a set of values can satisfy the QoS.

In the following, we focus the description on healthcare monitoring, the most important WBAN application. The design principles however are generic and are applicable to other WBAN applications.

## 2 Related Work

IEEE has finalized the 802.15.6 WBAN standard [2], where multiple MAC techniques, such as scheduled access, polling, and contention access, are supported in a beacon-based superframe. Another related standard is Bluetooth Low Energy (BLE), where the MAC layer adopts a TDMA based structure. Sensors periodically wake up to initiate/schedule and transmit the data via the advertising and data channel, respectively. However, both 802.15.6 and BLE MAC introduce large overhead in the periodic synchronization. Besides, there is no dedicated design for the dynamic traffic and channel in WBAN, which could significantly degrade QoS. We now review prior works that address these issues.

**Energy Efficiency.** Exiting WBAN MAC protocols enhanced the energy efficiency primarily by TDMA multiplexing [8], [10]–[12] or reducing the communication frequency of beacons [13]. However, the synchronization of such superframe-based structures would consume extra energy. Therefore, energy-efficient synchronization have been studied. In [11], the sensors were synchronized via detecting their own signal peaks driven by the heartbeat. However, this cannot always be robust since the change of heartbeat rhythm may not be reflected simultaneously on all the sensors and some sensors, e.g., accelerometers, may not be used to extract the heartbeat. Our work differs from all these works in that we propose a new synchronization scheme to dynamically adjust synchronizing frequency and maximize the interval between two synchronization, which could largely reduce the overhead. More importantly, we do not independently consider the energy issue. Instead, we minimize the energy while satisfying QoS constraints via optimal slot allocation.

**Lossy On-body Channel.** The long-lasting deep fading of WBAN is especially troublesome for static TDMA assignment [8], [10]–[12]. Using static TDMA, a node will obtain consecutive slots and thus its frames are transmitted one by one. If a frame is lost due to deep fading, the subsequent transmission would be, most probably, dropped again because the deep fading of WBAN lasts for such a long interval (up to 400 ms) that multiple frames could be scheduled for transmission by WBAN radio [6]. Consequently, repeated frame loss would occur.

In [14], the authors proposed to reduce frame loss using a variable TDMA scheduling algorithm. They dynamically moved forward or backward nodes' transmission in such a way that node average delivery probability could be improved. Nevertheless, since only one single slot was allocated to a node and the goal of the algorithm was limited to decreasing node average loss rate, it was highly likely for high-priority nodes to undergo unacceptable throughput and frame loss. Recently, CA-MAC protocol that adopted a hybrid TDMA and contention based superframe was proposed to address the WBAN deep fading [3], [15]. Thanks to the random backoff in contention period, the frame transmission from nodes/links are interleaved. Therefore, when a node regains its opportunity of transmission after a delivery failure and subsequent random backoffs, the link condition might have improved and accordingly the retransmission might succeed. However, the usage of contention period could negatively impact the energy efficiency and channel utilization. In this paper, we instead utilize the collision-free TDMA and then formulate a slot allocation optimization in order to guarantee the different QoS requirements among heterogeneous sensors. We also mitigate the impact of fading by optimally adjusting transmission order, which is based on formal WBAN channel and traffic model rather than simply depending on the contention to blindly spread the transmission in time, as is done in [3], [15].

**Dynamic Traffic**. The variation of monitoring contexts brings time-varying sensor sampling rate and WBAN traffic. This could deteriorate the performance of protocols using fixed resource allocation [10]–[13]. As such, researchers adaptively adjusted TDMA slots [16] or separated the slots into data and control purpose for prompt transmission [8]. In [9], a contention-based WBAN system was designed for falls assessment, where high-priority nodes in current context can interrupt low-priority nodes and aggregate the data for fast transmission. However, this design could introduce severe collisions due to its contention nature. Besides, it is computationally impractical for sensor nodes to decide what is the current context and which types of data needs to be gathered with high priority, which is usually complicated analysis rather than simple out-of-bound computation. In [17], the authors used wakeup radio signals to activate the sensor nodes from sleep state and to allocate channel resources when abnormal traffic comes. However, an extra wakeup circuit is added in traditional sensor nodes, which introduces additional hardware complexity. Our work, instead, utilizes the centralized nature of WBANs and allows the PS to adaptively allocate the slots based on traffic demand, which adapts to the dynamic contexts. Thanks to this asymmetric architecture, we distribute the computation of context switching and slot allocation to the powerful PS.

**Cross-layer Designs.** The authors in [18] studied the two-tier sensor-to-PS and PS-to-network communications and improved delay, reliability, and energy performance by properly using radio frequency (RF) and body coupled communications (BCC). However, the body area communication is restricted to BCC rather than the more commonly deployed RF WBAN. In [19], a receiver-initiated MAC protocol was developed for rapid and robust sensor data delivery. The authors focused on spectrum access to handle the inter-BAN interference while not examining the body movement pattern and the resulting challenges for sensor-to-PS transmission. In summary, the proposed MAC protocol is essentially different from these schemes because it targets the body area transmission by comprehensively exploring the mobility-incurred channel and traffic characteristics of WBAN. It also formulates an optimal slot allocation to minimize the energy subject to QoS constraints. The new synchronization scheme and analytical performance model further compliment the energy efficiency and QoS guarantee.

## 3 System Model

### 3.1 Basic Framework

We consider the star-topology WBAN, as shown in Fig. 1. The PS is not power hungry and hence we only consider power consumption of sensors. The PS manages the major operations of WBAN, such as resynchronization, context recognition, and computing tranmission schedule. We assume there are $N$ sensor nodes in the WBAN and each node $i$ belongs to the sensor nodes set $\mathcal{N}$.

We employ a TDMA-based MAC protocol to prevent collisions, idle listening, and overhearing. As shown in Fig. 2a, we divide the time axis into periodic time intervals, called *superframes*, with constant length $T$. A superframe consists of beacon, active, and inactive part. The beacon part is used by the PS to broadcast messages to the sensor nodes. The active part is comprised of consecutive slots and each slot can accommodate one frame transmission (including data transmission, ACK reception, and radio transition time) from one node. Since the radio turnaround time between transmitting and receiving state is much smaller than the time for one frame transmission, it is disregarded in this paper. The *guard time slot* ($T_g$) is added in order to prevent slot overlapping and perform the proposed synchronization scheme. After sufficient slots have been reserved for the sensor nodes according to sampling rate requirement, the rest of the superframe is considered as inactive part, during which no transmission is scheduled. In order to achieve these functions, we define the MAC frame format of the proposed protocol based on IEEE 802.15.6 standard [2], as shown in Fig. 2b.

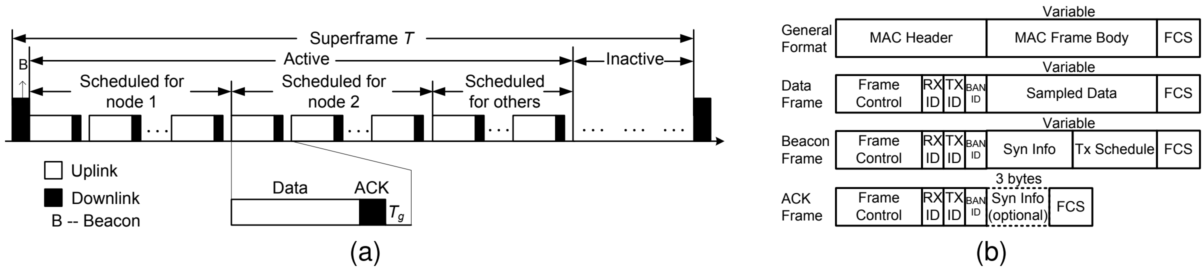



Fig. 2. a) TDMA-based superframe structure of the proposed MAC protocol; b) MAC frame format of the proposed MAC protocol.

### 3.2 Channel Model

IEEE 802.15.6 BAN standard has suggested the use of Markov channel models since experiments have verified that Markov models can effectively characterize the WBAN channel dynamics [6], [20]. In this paper, we model the on-body channel at each link (between each sensor node and the PS) during a particular mobility pattern (called *activity*) as a two-state Markov process. A frame can be successfully received in *good* state while it will be dropped in *bad* state. Since measurement studies have shown that channel condition is directly correlated to human activity [6], [21], the model parameters differ in different activities. Such modeling is feasible since human body presents a stable moving pattern under a given activity (e.g., walking) and thereby the Markov characterization exists in each link. Besides, the unchanged channel condition during one slot is especially true for WBAN due to its long-lasting nature of channel state [6], e.g., up to 400 ms (dozens of frames) in bad state.

Specifically, the Markov transition probability matrix $\mathbf{P}$ can be denoted as,

$$\mathbf{P} = \begin{bmatrix} 1 - P_{BG} & P_{BG} \\ P_{GB} & 1 - P_{GB} \end{bmatrix} \qquad (1)$$

where $P_{GB}$ represents the probability of a good state transits to a bad state, i.e., the probability of a node's frame loss in current slot given its successful transmission in the previous slot. Similarly, $P_{BG}$ represents the probability of a bad state transits to a good state. Note that a link may have different transition probability matrix under different human activities, and different links may have different $\mathbf{P}$ due to their heterogeneous locations. Under a particular activity, however, each link's Markov characterization ($\mathbf{P}$) is fixed [14], [21]. We denote the probability of a link being in good state at a given slot as $p_0$ and the sum of $P_{BG}$ and $P_{GB}$ as $Q$. Then the probability for this link to be good state after $\tau$ slots (during the same activity) is [14]

$$\begin{aligned} p(\tau) &= \begin{bmatrix} 1 - p_0 & p_0 \end{bmatrix} \cdot (\mathbf{P})^{\tau} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \frac{P_{BG}}{Q} + \frac{p_0 P_{GB} - [1 - p_0] P_{BG}}{Q} (1 - Q)^{\tau} \\ &= \begin{cases} \frac{P_{BG}}{Q} - \frac{P_{BG}(1-Q)^{\tau}}{Q} & p_0 = 0 \\ \frac{P_{BG}}{Q} + \frac{P_{GB}(1-Q)^{\tau}}{Q} & p_0 = 1 \end{cases} \end{aligned} \qquad (2)$$

If the channel state of the given slot is known ($p_0 = 0$ or $p_0 = 1$), we then can estimate the future channel condition using (2). We next present two propositions that shed light on the important properties of the Markov channel model.

*Proposition 1:* When $\tau \rightarrow \infty$, $p(\tau) \rightarrow \frac{P_{BG}}{Q} = \frac{P_{BG}}{P_{BG}+P_{GB}}$. The larger $Q$ is, the more $(1-Q)^{\tau}/Q$ shrinks and the faster $p(\tau)$ converges.

*Proposition 2:* Let $\Delta\tau$ be a small time interval. Given $Q < 1$, we have $p(\tau) < p(\tau + \Delta\tau)$ if $p_0 = 0$ and $p(\tau) > p(\tau + \Delta\tau)$ if $p_0 = 1$.

*Remark 1:* Proposition 1 emphasizes that under a given activity the probability to be good eventually converges to a value that represents the steady fraction of good slots, i.e., the steady delivery probability. Furthermore, $Q$ signifies the channel variation speed. That is, the channel would vary from good/bad state to steady state faster with a larger $Q$. Proposition 2 manifests that the delivery probability is either monotonically increasing or decreasing when the initial channel is bad or good, respectively. The constraint ($Q < 1$) has been shown to hold under extensive measurements of typical WBAN settings [14] and hence it is assumed to be true in the following.

## 4 Proposed MAC Protocol

### 4.1 Basic Operations

We first describe the basic operations conducted at the PS and sensor nodes.

1) PS broadcasts a beacon frame (as the format in Fig. 2b) with the initial synchronization message, including clock information, superframe length and slot length. The initial sampling rate and initial time to transmit/sleep for each sensor are also derived via initial transmission schedule (random order and equal duration) and embedded in the beacon.
2) Sensors synchronize with PS via initial beacon.
3) Sensors receive the sampling rate and transmission schedule in the beacon frame. Then they collect the data using received sampling rate and schedule the next time to transmit/sleep before entering sleep.
4) Sensors wake up at the scheduled slots and transmit one or more data frames based on the

transmission duration. These data will be used by the PS to recognize the current context.

5) PS replies an ACK upon receiving a data frame from a sensor. The resynchronization information is optionally embedded in the ACK only if the clock drift would cause frame collision.
6) Sensors power off and save energy during others' scheduled slots.
7) PS analyzes the received sensor data and only sends out updated sampling rate and transmission schedule via the beacon frame at the upcoming superframe. The transmission order is first decided by two ordering rules (Section 4.2) and the transmission duration is then obtained by solving the optimal slot allocation problem (Section 4.3).
8) Repeat step 3) through step 7).

This way, a sensor $i$ is able to periodically acquire these scheduled access information. If it misses a beacon, it will simply keep asleep during the superframe and wake up at the next beacon period. We have found in the evaluations that by increasing the transmit power of the non-power-hungry PS to -10 dBm, only negligible amount ($\sim 1\%$) of beacons were lost, which did not cause any significant impacts on the overall performance. Hence, we adopt this simple strategy to handle beacon loss.

### 4.2 Adjusting Transmissin Order

In this section, we estimate the dynamic channel status and thus adjust the transmission order of nodes to achieve relatively high delivery probability. Note that although the channel model parameters is fixed for a certain activity, e.g., walking, the link status is still time-varying. The fixed activity only means $\mathbf{P}$ is time-invariant. When the activity is changed, which can be easily detected by context recognition algorithms using sensor data (e.g., from accelerometer and/or inertial sensor) [22]–[24], the PS only needs to utilize the updated channel model with new parameters and conduct the channel estimation. The channel model parameters for typical activities can be pre-collected from experimental results. Hence, we focus on the introduction of channel estimation and transmission order adjustment under a given activity.

**Between Set Ordering.** Given that the PS only obtains outdated link status when listening to sensors in their respective slots, the most recent channel status available to the PS is the outcome of the last allocated transmission for each node in the previous superframe. These most recent outcomes are termed as *initial channel state*, $X_0$. We denote sensor nodes with initially good and bad channel state as the set $GOOD = \{i|X_{0,i} = 1\} = \{i|p_{0,i} = 1\}$ and $BAD = \{i|X_{0,i} = 0\} = \{i|p_{0,i} = 0\}$, respectively. Considering the monotonically varying $p(\tau)$ in (2), we observe that the delivery probability for $GOOD$ nodes will decrease from 1.0 to $\frac{P_{BG}}{Q}$ with regard to (w.r.t.) time, and increase from 0 to $\frac{P_{BG}}{Q}$ w.r.t. time for $BAD$ nodes.

*Remark 2 (Between Set Ordering):* Due to such monotonicity, $GOOD$ ($BAD$) nodes would enjoy a higher (lower) delivery probability if they are scheduled to transmit sooner in the upcoming superframe. Thus, our first ordering rule is to assign all $GOOD$ nodes in the former part of a superframe before all $BAD$ nodes.

**Within Set Ordering.** As sensor nodes have heterogeneous link status and QoS demands, it is important to ensure that the delivery probability of node $i$ satisfies a threshold $TH_i$. We aim at adjusting the order of nodes within each set such that those individual thresholds can be approached as close as possible.

For node $i \in GOOD$, if the delivery probability of its last assigned transmission in the upcoming superframe is larger than the threshold $TH_i$, its frames scheduled before this last assigned slot can also enjoy reliable transmission due to its monotonically decreasing delivery probability. We can therefore give the delivery probability constraint for node $i$, $i \in GOOD$,

$$p(D_i + x_i) \geq TH_i \Rightarrow x_i \leq a_i \tag{3}$$

where $D_i$ indicates how many slots have passed since the last transmission of node $i$ in the previous superframe, $x_i$ is the index showing which slot is its last allocated slot in the upcoming superframe, and $a_i$ is the upper bound of this index. That is, we should schedule all the transmissions for node $i, i \in GOOD$, before the $a_i$th slot of the upcoming superframe.

Similarly, for node $i \in BAD$, if its first scheduled transmission in the upcoming superframe enjoys a delivery probability larger than the threshold, its other transmissions can be all reliable. The corresponding delivery probability constraint of node $i$, $i \in BAD$, is

$$p(D_i + y_i) \geq TH_i \Rightarrow y_i \geq b_i \tag{4}$$

where $y_i$ is the index that shows which slot is the first allocated slot for node $i$ in the upcoming superframe, and $b_i$ is the lower bound of such index. That is, we should schedule all the transmissions for node $i, i \in BAD$, after the $b_i$th slot of the upcoming superframe.

*Remark 3 (Within Set Ordering):* To help each individual $GOOD$ ($BAD$) node to achieve their delivery probability threshold, we need to order them within the $GOOD$ ($BAD$) set based on the upper bound $a_i$ (lower bound $b_i$) of their assigned slots. Thus, the second ordering rule is that a $GOOD$ node with a larger upper bound should be scheduled after the one with a smaller upper bound. Similarly, a $BAD$ node $i$ should be scheduled before node $j$ if $b_i < b_j$.

By computing all $a_i$'s for node $i \in GOOD$ and all $b_i$'s for node $i \in BAD$, as well as carrying out the two ordering rules, we can decide the transmission order in a channel-aware way.

### 4.3 Optimizing Transmission Duration

We have introduced the adjustment of transmission order that facilitates reliable transmission. However, in order to guarantee the delivery probability and throughput, it is still essential to quantitatively study the transmission duration of nodes to complete the slot allocation. In this section, we formulate the *optimal slot allocation problem*, in order to seek the optimal transmission duration, i.e., the number of assigned slots $n_i$, for each node in the upcoming superframe.

**Delivery Probability Constraint.** We have derived the constraint of delivery probability for the proposed WBAN in (3-4). Now we convert them into the formulas w.r.t. the number of assigned slots $n_i$. By using (3), we have explained that all allocated slots for node $i \in GOOD$ should be ordered before the $a_i$th slot so as to meet the threshold $TH_i$. Thus the total number of slots scheduled for node $i$ and the nodes before $i$ should not exceed $a_i$. The delivery probability constraint for node $i$, $i \in GOOD$, can be converted as

$$n_i + \sum_{\{j|a_j<a_i\}} n_j \leq a_i \tag{5}$$

Similarly, we can also re-write the constraint of delivery probability for node $i \in BAD$ as

$$\sum_{\{g|a_g \leq a_{\max}\}} n_g + \sum_{\{j|b_j<b_i\}} n_j \geq b_i \tag{6}$$

where $a_{\max} = \max\{a_i | i \in GOOD\}$.

**Throughput Constraint.** To handle the various types of traffic and the dynamic nature of WBAN, we introduce the concept of monitoring contexts.

*Definition 1 (Monitoring Context):* A monitoring context is a medical scenario that is decided by diagnosis and human mobility, and is predefined by doctors. For example, three contexts can be defined for a healthy 20 year-old person, i.e., normal (walking/80-120 heart rate), alert (exercising/120-200 heart rate), and emergency (maximum efforts/more than 200 heart rate) [25]. In different contexts, doctors require different sampling rate to assist the treatment.

We aim at adjusting the transmission duration (number of slots) of each sensor such that their traffic requirement in current context can be satisfied. For example, the slots for heartbeat data should be increased during emergency context in order to meet the increased sampling rate. This feature is termed as context-awareness. To recognize context variation of the person, the PS collects and analyzes data frames from multiple sensors, e.g., via heart rate and motion data. Notice that context recognition algorithms have been widely studied [22]–[24]. The protocol can simply borrow existing algorithms, which will not affect the design of the proposed MAC.

In order to guarantee the context-aware transmission, the amount of transmitted data of each type of sensor nodes should be always larger than the minimum throughput requirement in the current context. Formally, the corresponding constraint for node $i$ to achieve desirable throughput is given by

$$n_i(RT_{data} - L_{oh}) \geq S_i T \tag{7}$$

where $R$ is the transmission rate of the radio platform, $T_{data}$ is the duration of transmitting a data frame, $L_{oh}$ is the overhead bits per frame, and $S_i$ is the minimum throughput (derived from sampling rate) currently required by application-layer services for node $i$.

Accordingly, those sensor nodes more related to the new monitoring context (larger $S_i$) would obtain more slots for data transfer. However, other irrelevant nodes (smaller $S_i$) would release some or all of the slots, reducing the energy cost.

**Optimal Slot Allocation.** Since the PS can be easily recharged, the optimization objective is to minimize the total energy consumption of the power-hungry sensors, subject to the constraints of delivery probability and throughput. The sensors' energy consumption primarily results from data transmission, ACK reception, radio wakeup transition (from sleeping to transmitting/receiving), and reception of synchronization information. Other energy consumption factors are not considered because they are usually multiple orders of magnitude smaller than the energy consumption caused by transmission or reception. Since the frequency of wakeup transition is fixed for sensors and the synchronization is impacted by clock drift rather than slot allocation, the objective is equivalent to minimizing the energy consumption of data communications, i.e., the total number of assigned slots, in the upcoming superframe. Then we can formulate the optimal slot allocation problem as

$$\begin{array}{lll} \min & \sum_{i \in \mathcal{N}} n_i & \\ \text{s. t.} & n_i(RT_{data} - L_{oh}) \geq S_i T & \forall i \in \mathcal{N} \\ & n_i + \sum_{\{j|a_j<a_i\}} n_j \leq a_i & i \in GOOD \\ & \sum_{\{g|a_g \leq a_{\max}\}} n_g + \sum_{\{j|b_j<b_i\}} n_j \geq b_i & i \in BAD \\ & \sum_{i \in \mathcal{N}} n_i \leq \frac{T}{T_{slot}} & \forall i \in \mathcal{N} \end{array} \tag{8}$$

where $T_{slot}$ is the duration of a frame transmission slot. By using the first constraint, we ensure the transmitted data is no less than the application-level traffic load. We also guarantee the reliable transmission of $GOOD$ and $BAD$ nodes via the second and third constraint, respectively. Finally, with the last constraint, the total number of allocated slots can be limited to the length of the superframe. Considering the fact that the constraints and the objective of (8) are all linear, and the number of assigned slots shall be a integer, the formulated problem is an Integer Linear Programming problem, which can be solved efficiently [26]. By solving (8), the PS can periodically compute the optimal transmission duration for each node after the transmission order adjustment.

### 4.4 Synchronization Scheme

The synchronization overhead is one of the main obstacles for traditional beacon-enabled protocols. In [2], [15], for example, all the sensor nodes have to receive synchronization information periodically through the beacon frame even if some of them are still synchronized with the PS.

In this section, we propose a new synchronization scheme to decrease the synchronization overhead, where the sensor nodes only need to resynchronize through ACK frames when the clock drift may cause overlapped transmission and they do not have to synchronize in every superframe.

After a certain synchronization, the clock drift between the clocks of each sensor node and the PS increases w.r.t. time. A sensor node might transmit a frame before or after the beginning of the assigned slot, which might cause overlapped transmission with other nodes. Thereby, we incorporate a guard time slot into a frame transmission slot after the ACK reception in order to avoid such frame collisions.

The length of the guard time slot $T_g$ should be deliberately set to accommodate both forward and backward overlapping. In the proposed MAC, the transmission order of sensor nodes is dynamically adjusted in every superframe. The maximum duration of a node not receiving an ACK from the PS is $2T - T_{slot} - T_b$, where $T_b$ is the duration of a beacon slot. The maximum duration is achieved when the node transmits at the first slot in a superframe and then transmits at the last slot in the next superframe. The maximum clock drift over the maximum duration between a node's two receptions of ACK is then $\theta(2T - T_{slot} - T_b)$, where $\theta$ is the frequency tolerance. Since the clock drift could be either forward or backward, the transmission of a senor could be delayed or advanced. To avoid the transmission overlapping, the guard time slot $T_g$ is therefore given by

$$T_g = 2\theta(2T - T_{slot} - T_b) \tag{9}$$

Note that $\frac{T_g}{2}$ is the upper bound of the clock drift over the time interval between two receptions of ACK, i.e, only one node is likely yet unnecessary to accumulate such a drift. In reality, the clock drift of a node between two ACK receptions could be any value in the range from 0 to $\frac{T_g}{2}$.

Regarding the operation procedure of the proposed synchronization scheme, a sensor node should first encapsulate the local clock information (time-stamp) into the frame control field of a data frame. When the PS receives the data frame, it takes the clock information and compares that with its own clock information. If the time interval between the two clocks exceeds half the guard time slot (i.e., $\frac{T_g}{2}$), the transmission overlapping may occur. Then the control field Syn Info is set as the time interval. Otherwise, no operation is needed, which means the sensor node can wait more time before resynchronization. The PS then encapsulates Syn Info field, if any, into an ACK and sends it to the corresponding sensor node. The sensor can accordingly adjust the local clock relying on the value of Syn Info after receiving the ACK frame.

Since the times of resynchronization only depends on the property of nodes' clocks and is not related to transmission order or duration, the proposed synchronization scheme can be directly added into the proposed MAC protocol to reduce overhead whereas not negatively impacting the optimal slot allocation.

## 5 Latency Analysis

Recall that the transmission order of sensor nodes are dynamically moved forward and backward in every superframe to enure reliability. This mechanism results in a distinct latency performance compared with conventional protocols with fixed transmission order. Therefore, in this section, we analyze the proposed MAC protocol in terms of latency, which is another important QoS requirement, especially in heathcare monitoring application. More importantly, we demonstrate the impacts of protocol parameters on latency performance and how we can adjust the parameters in order to meet different WBAN requirements.

### 5.1 Latency of Normal Context

In a normal monitoring context for WBAN, data frames are generated with a constant inter-arrival time. Due to the throughput constraint in (7), periodically generated frames during a superframe can be successfully transmitted by the end of the next superframe. Accordingly, the maximum latency for an arbitrary frame in a normal context is thus $2T - T_b - T_{ACK} - T_g$, where $T_{ACK}$ is the duration for receiving a complete ACK frame (including Syn Info field). This case occurs when the frame arrives at the transmission queue just after the node's assigned slot which happens to be the first slot of a superframe, and then is transmitted at the last slot of next superframe. This latency bound (e.g., $\sim 300$ms in our evaluations) can generally satisfy the requirement in normal context [7]. We can also tune the values of the protocol parameters ($T$, $T_b$, $T_{ACK}$, etc.) offline before the system starts in order to meet more strict latency bound.

However, coarse analysis of delay bound is not sufficient for an abnormal context because its traffic pattern and latency sensitivity is quite diverse. That is, a fine-grained and exact delay analysis for each sensor is needed. Hence, we primarily focus on the latency analysis of abnormal contexts in the following section. Moreover, when there is a context switch, we assume the queued frames at a node will be dropped so that the PS can receive the most recent sensor data for the new context. Accordingly, the latency performance of those frames is not evaluated in this paper.

### 5.2 Latency of Other Abnormal Contexts

Due to the time-varying sampling requirement and body movements, WBAN may be in different types of abnormal contexts, e.g., alert or emergency context. In an abnormal context, the inter-arrival time between two frames in a sensor node's transmission queue is a random variable. We model the arrivals of frames at a sensor node in abnormal contexts as a Poisson process as some existing WBAN systems [27], [28]. Notice that the frame arrival rate of a sensor at different abnormal contexts is different. We thus denote the frame arrival rate of node $i$ at the time $t$ as $\lambda_i(t)$. In a TDMA-based system, after a given node finishes the scheduled transmission slots, it would enter sleep state when other nodes are scheduled for transmission. For the given node, the experienced sleep can be viewed as the server's vacation in queueing theory. Hence, we model each sensor node as an M/D/1 queueing system with vacations [29]. We assume frames at each node are served in the first come first served manner. Since it takes $T_{slot}$ to finish one frame transmission, the service rate of node $i$ at the time $t$ is given by $\mu_i(t) = \frac{1}{T_{slot}}$. Thus the utilization factor $\rho_i(t)$ ($\rho_i(t) < 1$) is given by [29]

$$\rho_i(t) = \frac{\lambda_i(t)}{\mu_i(t)} = T_{slot}\lambda_i(t) \tag{10}$$

We now derive the expected frame latency, which consists of queueing and transmission latency. To simplify the notation, we consider an arbitrary frame arrival, regardless of the context the system belongs to (i.e., regardless of $t$). We first derive the queueing latency. For a given node, a frame may arrive when the node is experiencing the frame transmission slot or the sleep period. The frame has to wait until the frame transmission slot or the sleep period ends. Then it also needs to wait for the transmission of the frames that arrive at the queue before it. Hence, the expected queueing latency of an arbitrary frame $T_q$ is given by

$$T_q = T_r + \frac{E[N]}{\mu} \tag{11}$$

where $T_r$ is the expected remaining time for the transmission slot or sleep period and $E[N]$ is the expected number of frames that are already in the queue. According to Little's law [29], the expected number of frames in the queue seen by the given node is $E[N] = \lambda T_q$. Therefore, $T_q$ can be expressed as

$$\begin{aligned} T_q &= T_r + \frac{\lambda}{\mu}T_q = T_r + \rho T_q \\ \Longrightarrow T_q &= \frac{T_r}{1-\rho} \end{aligned} \tag{12}$$

Now we derive the expected remaining time $T_r$. We denote the probability that a frame arrives during the given node's transmission slot and sleep period as $p_{slot}$ and $p_{sleep}$, respectively. Since the utilization factor $\rho$ represents the fraction of time for frame transmission slots, we have $p_{slot} = \rho$. Note that a node would not necessarily transmit a frame during a frame transmission slot since the queue might be empty then. Besides, we also have $p_{slot} + p_{sleep} = 1$. Therefore, the expected remaining time $T_r$ is given by

$$T_r = p_{slot}\overline{T}_{r,slot} + p_{sleep}\overline{T}_{r,sleep} = \rho\overline{T}_{r,slot} + (1-\rho)\overline{T}_{r,sleep} \tag{13}$$

where $\overline{T}_{r,slot}$ and $\overline{T}_{r,sleep}$ is the expected remaining time of the transmission slot and sleep period, respectively. Since a frame tends to arrive halfway through the transmission slot on average and the duration of a slot is fixed at $T_{slot}$, we have

$$\overline{T}_{r,slot} = \frac{T_{slot}}{2}. \tag{14}$$

In contrast, for a given node, the duration of the sleep period is highly variable depending on many factors, such as the most recent transmission outcome of all nodes and the slot allocation of the previous superframe. Thus the durations of consecutive sleep periods seen by the node can be approximated as independent. Therefore, the expected remaining time of the sleep period for a node is approximated as uniformly distributed in the possible range. In Section 6, we will demonstrate the closeness with which the analytical approximation matches the simulation results. Due to the adjustment of transmission order, the duration of the remaining sleep period for node $i$ ranges from $0$ to $2T - T_b - 2n_{min,i}T_{slot}$, where $n_{min,i}$ is the minimum number of allocated slots for node $i$ to satisfy traffic requirement, i.e., $n_{min,i} = \frac{S_i T}{RT_{data} - L_{oh}}$. The maximum remaining sleep period is achieved when the node is ordered as the first one in the current superframe while as the last one in the next superframe. Thus the expected remaining sleep period for node $i$ is approximated as

$$\overline{T}_{r,sleep} \approx \frac{2T - T_b - 2n_{min,i}T_{slot}}{2} \tag{15}$$

By using (12-15), the expected queueing latency of an arbitrary frame $T_q$ for node $i$ can be obtained. Finally, the expected latency of a frame $T_{total}$ for node $i$ is the sum of the expected queueing latency and the expected transmission latency, i.e.,

$$\begin{aligned} T_{total} =& T_q + T_{data} \\ \approx& \frac{\rho T_{slot}}{2(1-\rho)} + \frac{2T - T_b - 2n_{min,i}T_{slot}}{2} + T_{data} \\ =& (\frac{\rho}{2(1-\rho)} - n_{min,i})T_{slot} + \frac{2T - T_b}{2} + T_{data} \end{aligned} \tag{16}$$

According to (16), when the proposed MAC protocol is deployed (i.e., when the protocol parameters including $T$, $T_{slot}$, $T_b$, $T_{data}$ and $L_{oh}$ are fixed), the expected frame latency of node $i$ in an abnormal context is determined by $\frac{\rho}{2(1-\rho)}$ and $n_{min,i}$. Notice that nodes with higher traffic requirement have larger value of

$\frac{\rho}{2(1-\rho)}$ and $n_{min,i}$. We also find that $n_{min,i}$ changes more significantly than $\frac{\rho}{2(1-\rho)}$ when the traffic requirement varies, which indicates that the variation trend of $n_{min,i}$ will decide the trend of $T_{total}$. Such relation always holds under practical settings of QoS parameters as we will demonstrate in Section 6.

*Remark 4:* Nodes with higher traffic requirement would achieve lower frame latency, which inherently meets the heterogeneous latency requirement of WBAN. Besides, when a given node's traffic requirement dynamically increases or decreases, based on (16), the frame latency of this node would adaptively decreases or increases, respectively.

We have showed above how the proposed MAC protocol addressing heterogeneous and dynamic traffic. However, $T_{total}$ may not always be able to satisfy the individual latency bound under a specific abnormal context. Although we may increase the traffic requirement of a particular node to decrease its expected latency, such operation is usually not allowed and may not work well. Hence, we sometimes need to carefully configure the protocol parameters before the deployment of the proposed MAC protocol. Further converting (16), $T_{total}$ can be derived as follows.

$$\begin{aligned} T_{total} \approx & \frac{\rho T_{slot}}{2(1-\rho)} + T - \frac{T_b}{2} - n_{min,i} T_{slot} + T_{data} \\ = & \frac{\rho T_{slot}}{2(1-\rho)} + T - \frac{T_b}{2} + (T_{slot} - T_{ACK} - T_g) \\ & - \frac{S_i T T_{slot}}{R(T_{slot} - T_{ACK} - T_g) - L_{oh}} \end{aligned} \tag{17}$$

According to (17), the expected frame latency of node $i$ is determined by the protocol parameters, such as $T$, $T_{slot}$, $T_b$, $T_{ACK}$, $T_g$ and $L_{oh}$, under a certain abnormal monitoring context and WBAN hardware platform (i.e., the traffic requirement and transmission rate is known). Since $T_g = 2\theta(2T - T_{slot} - T_b)$, in this case we can easily adjust the values of $T$, $T_{slot}$, $T_b$, $T_{ACK}$, and $L_{oh}$ for the protocol in order to meet a specific latency bound of that context. Note that the proposed MAC protocol is based on a Markov channel model, where the channel state is unchanged during a frame transmission slot. Hence, the duration of a slot $T_{slot}$ should be carefully set based on empirical channel measurements, which makes $T_{slot}$ improper to be changed as we want. The value of $L_{oh}$ would also better not be changed because a stable amount of overhead is needed for achieving protocol functions. Furthermore, since the possible values of $T_b$ and $T_{ACK}$ are relatively small, even very significant change of these parameters will not have desirable impacts on the latency. Therefore, we arrive at the following remark regarding MAC deployment guideline.

*Remark 5:* The superframe length $T$ should be changed to achieve the latency bound of a given context before the MAC deployment.

In Section 6, we will show that the latency of abnormal frames is monotonically and smoothly increasing as $T$ increases, which assists us to choose appropriate protocol parameters when deploying applications with different latency bounds.

# 6 Performance Evaluations

In this section, we evaluate the performance of the proposed MAC protocol in three aspects: reliability, energy efficiency, and latency.

**Simulation Modules.** We built a event-driven WBAN system using MATLAB. The application module generates constant-bit rate and Poisson traffic depending on the context. The MAC module supports multiple basic access schemes, such as scheduled access and random access. By combining or modifying these schemes, one can implement a specific MAC protocol for evaluations. For the proposed protocol, we also embed YALMIP [30], an MATLAB-based optimization toolbox, to solve the optimization in (8). The channel module simulates the deep-fading on-body channel by first generating a channel state of a slot based on the Markov model in Section 3.2 and then computing a received power satisfying that channel state. In other words, if the simulated channel state is good, the channel module would signal a received power higher than the receiver sensitivity to the receiver and vice versa for bad state. The receiver would finally compare the received power against sensitivity to drop/receive the frame. Note that we adopt the values of $\frac{P_{BG}}{Q}$ and $Q$ rather than the transition probabilities to generate the Markov channel states as they have realistic indication, i.e., steady delivery probability and channel variation speed.

**Parameters Setting.** We configure a typical WBAN that has one PS and five sensor nodes [7]. The maximum transmission rate of the transceivers is 220.1931 Kbps. We set the frame transmission slot as 10 ms, because the deep fading of on-body channel can last as short as 10 ms [6]. One superframe lasts for $T = 150$ ms. Considering the physical layer header, the length of a beacon frame and a complete ACK frame (including Syn Info field) is set to be 32 bytes and 16 bytes, respectively. The overhead for synchronization, i.e., the length of Syn Info field, is set to be 3 bytes. Thus the duration of transmitting a beacon frame, receiving a complete ACK frame, and receiving an ACK frame without Syn Info field can be obtained by dividing their corresponding length by the transmission rate. Due to the constant $T_{slot}$ and $T_{ACK}$, the duration reserved for transmitting a data frame $T_{data}$ is dictated by $T_g$, which is given by (9). The overhead of a frame $L_{oh}$ is set to be 13 bytes. The frequency tolerance $\theta$ is 100 parts per million, which means the clock will lose 100 time units over one million time units.

Regarding the QoS requirements, we assume there are one normal context and four abnormal contexts, i.e., alert, semi-urgent, urgent, emergency. We configure the data generation rate of different sensor under different contexts based on [7] in TABLE 1. We assume the first three sensors are less relevant to this patient (termed as group one, G1) and their rate demands decrease w.r.t the emergency level. For example, when a patient after heart surgery undergoes a sudden heart problem, the monitor of body recovery, e.g., EMG data, can be relaxed. This is vice versa for the other two nodes (termed as group two, G2). We set the threshold of delivery probability $TH_i$ for G1 and G2 nodes as 90% and 95% (frame loss rate 10% and 5%), respectively [7]. The default monitoring context is set to be semi-urgent in order to simulate the traffic load that is neither too heavy nor too light. The battery voltage, current draw, radio wakeup transition time of nodes are based on MICAz mote datasheet [31].

**Reference Protocol.** We compare the proposed MAC protocol with CA-MAC [15] introduced in Section 2 and IEEE 802.15.6 MAC with beacon-enabled mode [2]. We simulate CA-MAC using the same parameters except for several design-particular parameters, which are configured to the typical values presented in [15]. For the IEEE 802.15.6 MAC implementation, we include contention, schedule, and polling based access periods in the superframe structure. The contention period adopts CSMA/CA while the schedule-based periods allow a sensor to transmit frames periodically. The polling periods are used to schedule temporary slots in current superframe in case that retransmission or extra data is needed.

**Results Collection.** We will evaluate the system performance versus a serial of parameters. The curve in each figure shows the performance versus different values of the parameter of interest, while all the other parameters are remained at the typical values as defined above. Each point in the figures represents the average of 16 runs that each lasts for 10000 periods of superframe and the error bar showing the standard error. In each run, we uniformly choose $\frac{P_{BG,i}}{Q_i}$ and $Q_i$ for each link from the two ranges $[\min(\frac{P_{BG,i}}{Q_i}), \max(\frac{P_{BG,i}}{Q_i})]$ and $[\min(Q_i), \max(Q_i)]$, respectively, in order to simulate different channel conditions. From a public dataset that contains the link measurements of WBAN over the course of several days [21], $\frac{P_{BG}}{Q}$ was found to lie between 0.90 and 0.99 and $Q$ was found to lie between 0.05 and 0.5 [14]. Details of getting the ranges can be found in [14].

### 6.1 Reliability

We evaluate the reliability of a protocol by measuring its improvement over the fixed TDMA protocol. Here we denote $FLR_{reduc}$ as the reduction of frame loss rate over the fixed TDMA protocol, i.e.,

$$FLR_{reduc} = \frac{FLR_{fix} - FLR}{FLR_{fix}} \tag{18}$$

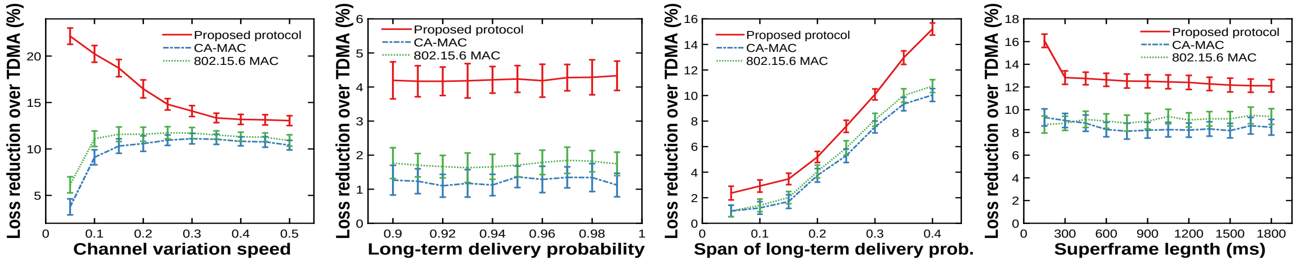
where $FLR$ and $FLR_{fix}$ is the frame loss rate for the protocol under evaluation and the fixed TDMA protocol, respectively.

TABLE 1
Data Generation Rate of Different Contexts (Kbps)

| Contexts | G1 | | | G2 | |
|---|---|---|---|---|---|
| | Node 1 | Node 2 | Node 3 | Node 4 | Node 5 |
| Normal | 12.96 | 12.96 | 25.92 | 25.92 | 51.84 |
| Alert | 25.92 | 25.92 | 25.92 | 38.88 | 38.88 |
| Semi-urgent | 19.44 | 19.44 | 19.44 | 51.84 | 51.84 |
| Urgent | 12.96 | 12.96 | 12.96 | 58.32 | 58.32 |
| Emergency | 6.48 | 6.48 | 6.48 | 71.28 | 71.28 |

We first evaluate the reliability performance under various channel parameters. We vary the channel variation speed $Q$ of all nodes from a practical range [0.05 to 0.5]. Note that, in this case, we only individually choose the $\frac{P_{BG}}{Q}$ for each link in the 16 runs. We can see from Fig. 3a that all protocols achieve a reduction of frame loss rate over the fixed TDMA and the proposed protocol brings the most improvement. This is because the proposed protocol exploits the model of WBAN deep fading and accordingly allocates the slots via link estimation rather than blindly interleaving the channel access of nodes as in the contention period of CA-MAC or 802.15.6 MAC. Besides, there are a large portion of schedule-based TDMA slots in CA-MAC or 802.15.6 MAC, which could still cause consecutive frame loss for nodes. We also see that $FLR_{reduc}$ of the proposed protocol first lowers down and then converges to a stable value with $Q$ increasing. In the proposed protocol, we utilize the most recent channel status to accurately estimate future channel. However, with a larger $Q$, the delivery probability of each node approaches to $\frac{P_{BG}}{Q}$ faster in the upcoming superframe no matter what the most recent link status is. This implies that the obtained link information would be less useful and hence the benefit of scheduling is mitigated as $Q$ increases. If $Q$ continues to increase, the actual delivery probability would tend to approximate the steady delivery probability $\frac{P_{BG}}{Q}$, and the impact of varying $Q$ on the scheduling benefit becomes trivial. In contrast, if channel varies slower (smaller $Q$), the channel in the upcoming superframe would be more relevant to that of the most recent transmission and thus slot allocation could introduce more benefits. However, we observe that $FLR_{reduc}$ of CA-MAC and 802.15.6 MAC show a different trend. This is because a larger portion of slots may have the same channel state with a slower-changing channel and thus the benefits of timely spreading the transmission become slight. Furthermore, 802.15.6 MAC is slightly better than CA-MAC since it uses polling slots to recover some of the frame loss.

We also evaluate the reliability metric under different steady delivery probability $\frac{P_{BG}}{Q}$ (from 0.9 to 0.99), and show the results in Fig. 3b. Similarly, only $Q$ is individually chosen for each link in the 16 runs in this evaluation. We can observe that the reduction of frame loss rate of the three protocols stabilizes at around 4%, 1% and 1.5%, respectively. From this re-



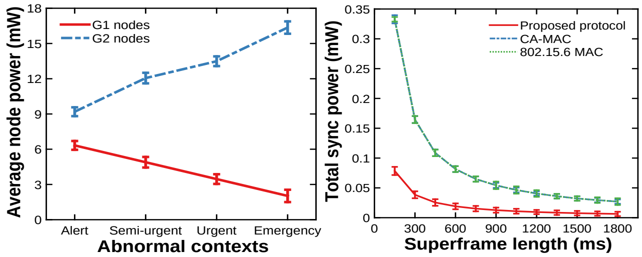

(a) (b) (c) (d)

Fig. 3. The reduction of frame loss rate over fixed TDMA protocol of different protocols versus a) channel variation; b) steady delivery probability; c) span of steady delivery probability; d) superframe length.

sult, we conclude that $\frac{P_{BG}}{Q}$ poses a negligible effect on $FLR_{reduc}$. It is true that the absolute value of loss rate for both protocols may reduce with $\frac{P_{BG}}{Q}$ increasing. However, the fixed TDMA can also achieve a lower loss rate because $\frac{P_{BG}}{Q}$ actually defines the long-term delivery probability. Furthermore, it is interesting to notice that the loss reduction are not significant in this evaluation. In fact, it is attributed to the lack of diversity among the link conditions, i.e., the steady channel status of all links are identical. We will discuss this in the following evaluation.

According to the above observations, we now attempt to explore the performance when the diversity of link conditions increases. We fix the average $\frac{P_{BG}}{Q}$ of each link at 0.95 while changing the span of each $\frac{P_{BG}}{Q}$ (denoted by $SP$). In other words, we uniformly choose the value of $\frac{P_{BG}}{Q}$ for all nodes from [0.95-$SP$, 0.95+$SP$]. With a larger $SP$, the diversity between link conditions would be larger. In Fig. 3c, $FLR_{reduc}$ grows with $SP$ increasing. This is because the increased $SP$ enlarges the channel status diversity among all the links. Thereby, some nodes may possesses a $\frac{P_{BG}}{Q}$ lower than 0.95 in the fixed TDMA. Nevertheless, we can still ensure the delivery threshold for such nodes by adopting the proposed protocol, which implies the reliability gains would rise. From these results, we can conclude that the proposed protocol is especially efficient for highly heterogeneous WBANs.

We now evaluate the reliability when varying superframe length $T$ from 150ms to 1800ms. In Fig. 3d, $FLR_{reduc}$ for the proposed protocol keeps dropping down w.r.t $T$ until a relatively stable value is reached. This phenomenon is very similar to the case of larger $Q$ shown in Fig. 3a. Given a larger $T$, the sensor nodes would tend to deliver a frame with a probability approaching $\frac{P_{BG}}{Q}$ in the upcoming superframe, regardless of the channel history. Therefore, the gain introduced by smart scheduling turns out to be less and eventually trivial. We also found that the loss reduction of CA-MAC and 802.15.6 MAC are essentially not impacted by $T$. As the interleaving benefits in contention is dictated by the extent of link diversity or variation rather than superframe length, the total loss rate could be nearly unchanged even though the contention length may rise with $T$ increasing.



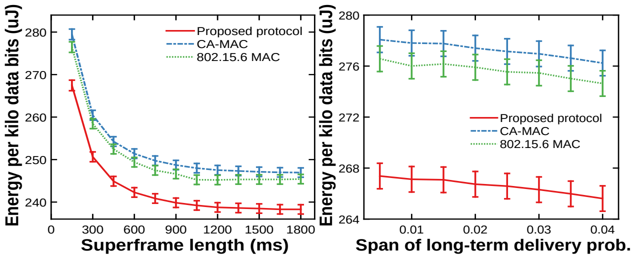

(a) (b)

Fig. 4. a) Average node power consumption of G1 and G2 nodes in different abnormal contexts; b) Power consumption for synchronization of different protocols versus superframe length.

### 6.2 Energy Efficiency

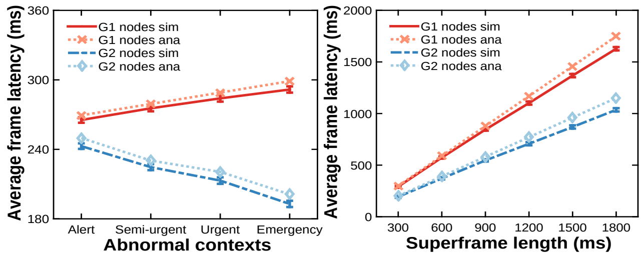
We first evaluate how the power consumption of the proposed MAC changes when the context varies. In Fig. 4a, we demonstrate the average node power consumption of G1 and G2 nodes under different abnormal contexts. We observe that the power of G2 nodes increases as the context becomes emergent. This is due to the increasing importance of G2 data, which results in more assigned slots and transmission attempts for G2 nodes to meet the larger throughput requirement. However, G1 nodes will reduce the transmission and the power because G1 data becomes less useful. Such heterogeneous and dynamic power dissipation accords with the features of WBAN.

Now we compare the power consumption for synchronization introduced by the proposed protocol and two reference protocols in order to validate the effectiveness of our design. We show the results in Fig. 4b. It can be seen that the proposed protocol consumes the least synchronization power. This is because the proposed synchronization scheme does not have to synchronize in every superframe, which is however required for CA-MAC and 802.15.6 MAC. We also observe that the synchronization power of the proposed protocol decreases as the superframe length increases. According to (9), the guard time slot increases as $T$ increases, which means the proposed protocol can tolerate a larger clock drift before resynchronization and thus needs a less frequent synchronization. We




(a) (b)

Fig. 5. Energy consumption per kilo data bits of different protocols versus a) superframe length; b) span of steady delivery probability.




(a) (b)

Fig. 6. Average frame latency of G1 and G2 nodes versus a) abnormal contexts; b) superframe length.

found that the average node synchronization times drops from 2600 to 2507 with increasing $T$. Furthermore, CA-MAC and 802.15.6 MAC consume almost the same power because they both use the periodic synchronization via beacon frames.

To further measure the energy efficiency, we also evaluate the energy consumption per kilo data bits. We compare the energy efficiency under a set of $T$ in Fig. 5a. The proposed protocol outperforms CA-MAC and 802.15.6 MAC since it guarantees the delivery probability and thus mitigates the energy wastage from frame loss. However, CA-MAC and 802.15.6 MAC exploit random backoff mechanism, which spoils the bandwidth usage and may incur frame collisions. Besides, the synchronization overhead of the proposed protocol is also much smaller. It is interesting to note that the energy efficiency of proposed protocol improves as $T$ increases despite of its degraded reliability performance (Fig. 3d). This is because, with $T$ increasing, the energy benefits introduced by using the proposed synchronization and the periodically fixed-number wakeup is greater than the energy wastage caused by using the less useful channel estimation. The same rationale also applies to the trend of CA-MAC and 802.15.6 MAC.

We also explore how the energy efficiency is impacted by varying channel parameters, i.e., $\frac{P_{BG}}{Q}$, $SP$, or $Q$. We arrive at that the proposed protocol achieves a superior energy efficiency under extensive channel conditions. More importantly, the curve of energy efficiency is highly depended on the reliability curve. An example for the energy efficiency with changing $SP$ is demonstrated in Fig. 5b. We observe that the energy per kilo data bits decreases w.r.t. the diversity among channels. Such trend of increasing energy efficiency accords with the trend of loss rate reduction with $SP$ increasing. This is because the energy is primarily dictated by data transmission when synchronization and wakeup power is fixed. The similar rationale can also be applied when $\frac{P_{BG}}{Q}$ or $Q$ is changed.

### 6.3 Latency

We use the time interval between the arrival time of a frame at a sensor node and the reception time of the frame at the PS to signify the frame latency. Since WBAN is implemented in a short range, the relatively low propagation delay is disregarded. The average latency of all successfully transmitted data frames during a simulation is used as the metric.

We now demonstrate that the proposed protocol can achieve latency adaptivity as analyzed in (16) and latency scalability as analyzed in (17), which satisfies the heterogeneous latency requirement of sensors in different contexts. In Fig. 6a, we show the average frame latency of the proposed protocol when context becomes more abnormal. To highlight the WBAN heterogeneity, we enhance the difference of $n_{min,i}$ (i.e., minimum number of allocated slots) among the sensors by increasing the superframe length to 300 ms. It can be seen that the average frame latency for G1 nodes and G2 nodes increases and decreases, respectively. This is because the proposed protocol dynamically re-allocates the slots by using the context-aware adjustment of transmission duration. Thus when G2 nodes have higher traffic requirement and more critical delivery tasks with more emergent context, they will have a greater $n_{min,i}$ and can transmit more frames per unit time, which reduces the average frame latency. Similarly, some slots of G1 nodes are released when the context becomes emergent (i.e., G1 nodes' traffic requirement becomes lower), which results in the latency increase. We also observe that such a trend matches the analytical results obtained from (16). Hence, the proposed protocol can properly switch among different contexts to handle the traffic by re-allocating the transmission duration of nodes.

Then we proceed to evaluate the abnormal latency performance of the proposed protocol under different superframe length $T$. The monitoring context is set to "emergency" in order to highlight the latency difference between the two priorities. As shown in Fig. 6b, the average frame latency of both G1 and G2 nodes increase as $T$ increases. Although $T_{data}$ decreases due to the enlarged $T_g$, the average length of sleep period experienced by the sensor nodes is increased and contributes more to the total latency. Hence, according to (16), the average frame latency increases. We can observe that the gap between the simulation results and the analytical values increases with increasing

$T$. This trend attributes to fact that the number of assigned slots has to adjust with a greater extent in order to satisfy the reliability constraint in (8) under a larger $T$. Thus the actual residual sleep time would deviate greater from the approximated value in (15). Finally, based on such monotonically and smoothly increasing trend of latency, we are able to properly adjust the superframe length so as to meet a specific latency bound when deploying the system.

### 6.4 Scalability

We now evaluate the system performance when the number of sensor nodes is varied. According to IEEE 802.15.6 application summary [7], the maximum number of sensors among all wearable BAN applications is 24. Therefore, we range the sensor number from 5 to 25. We also configure the same data generation rate for all nodes such that the inclusion of one new node would introduce the same level of traffic. Default values are used for all other parameters. We show the results in Fig. 7. We observe that the performance of the proposed protocol is relatively stable. It only shows a degradation (less loss reduction and higher energy per kilo bits) when the number of sensors reaches 25. This is because the slot allocation will have less flexibility with more sensors and sometimes the optimal slots of a node could be taken by another node, which leads to the sub-optimal performance. Since WBAN does not need more sensors for wearable applications, we believe the proposed protocol is sufficiently scalable. On the other hand, the performance of CA-MAC and 802.15.6 decrease with increasing WBAN size since more collisions would be introduced in the contention period.

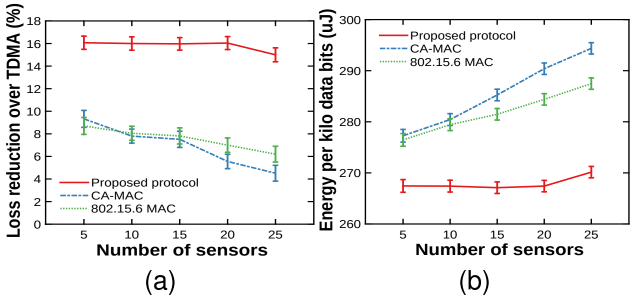



Fig. 7. a) Reduction of frame loss rate over fixed TDMA protocol versus number of users; b) Energy consumption per kilo data bits versus number of users.

## 7 Discussion

**Topology.** We focus on the most typical WBAN, a star-topology deployment [2], and do not use multi-hop relaying to maintain the reliability. This is because we can ensure the reliability by carefully scheduling nodes' transmission without using relay. Besides, relay could lead to additional latency and energy cost.

**WBAN Channel and Loss.** We focus on addressing the deep fading and frame loss caused by body movement via intelligently scheduling the sensor transmission, because this is the unique feature that sets WBAN apart from other wireless networks. Other conventional reliability issues can be handled by exiting techniques, e.g., noisy data can be recovered by error correction schemes.

We also do not study inter-BAN interference because we would like to focus on the reliability issues within one WBAN area. Despite its importance, a separate full-scale study is definitely needed because people may have distinct movement pattern when others are around and how the inter-BAN interference would impact the deep fading is unknown.

**Markov Channel.** To enhance the accuracy of Markov channel model, one may introduce more channel states. Besides, designing a online learning process of model parameters might be also beneficial. However, the proposed protocol has been shown to be robust against modeling inaccuracy since it shows satisfactory performance in extensive channel settings.

**Practical Issues.** To realize the proposed WBAN, context information can be first entered into the PS. Then the MAC of sensor nodes will dynamically configure the received sampling rate and radio schedule into the sensor hardware and wireless interface, respectively. Such level of programmability is widely supported in modern embedded systems. For example, programmable sampling rate is implemented using commercial sensors in [9].

## 8 Conclusions

In this paper, we propose a QoS-driven and energy-efficient MAC protocol to handle the technical challenges resulting from energy consumptions, time-varying channel and context variations in WBAN. We are able to achieve energy-efficient delivery with QoS provision, by adopting the optimal slot allocation based on both traffic nature and channel condition. The energy efficiency of WBAN is further enhanced by using the proposed synchronization scheme. The proposed protocol accomplishes more effective performance than CA-MAC and IEEE 802.15.6 MAC under a diverse set of practical settings. Furthermore, we show that the proposed protocol is especially suitable to the dynamic and heterogeneous environment of WBAN. More importantly, we provide practical guide on how to appropriately configure the superframe length in order to achieve QoS and energy efficiency.

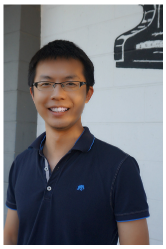
## References

[1] Z. Yan, B. Liu, and C. W. Chen, "Qos-driven scheduling approach using optimal slot allocation for wireless body area networks," in *Proc. IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Oct. 2012, pp. 267 –272.

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

[3] B. Liu, Z. Yan, and C. W. Chen, "Mac protocol in wireless body area networks for e-health: challenges and a context-aware design," *IEEE Wireless Commun. Mag.*, vol. 20, pp. 64–72, Aug. 2013.
[4] S. Ullah *et al.*, "A comprehensive survey of wireless body area networks: on phy, mac, and network layers solutions," *Journal of Medical Systems*, vol. 36, no. 3, pp. 1065–1094, 2012.
[5] P. Hall *et al.*, "Antennas and propagation for on-body communication systems," *IEEE Antennas Propag. Mag.*, vol. 49, no. 3, pp. 41–58, Jun. 2007.
[6] K. Y. Yazdandoost and K. Sayrafian-Pour, "Channel model for body area network (BAN)," IEEE P802.15-08-0780-11-0006, Sep. 2010.
[7] B. Zhen, M. Patel, S. Lee, and E. Won, "Body area network (BAN) technical requirements," IEEE P802.15-08-0037-04-0006, Jul. 2008.
[8] L. Lin, K.-J. Wong, A. Kumar, S.-L. Tan, and S. J. Phee, "A novel tdma-based mac protocol for mobile in-vivo body sensor networks," in *Proc. 14th IEEE Healthcom*, Beijing, China, Oct. 2012, pp. 273–278.
[9] T. O'Donovan, J. O'Donoghue, C. Sreenan, D. Sammon, P. O'Reilly, and K. O'Connor, "A context aware wireless body area network (BAN)," in *Proc. 3rd International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, Apr. 2009, pp. 1–8.
[10] O. Omeni, A. Wong, A. Burdett, and C. Toumazou, "Energy efficient medium access protocol for wireless medical body area sensor networks," *IEEE Trans. Biomed. Circuits Syst.*, vol. 2, pp. 251–259, Dec. 2008.
[11] H. Li and J. Tan, "Heartbeat-driven medium-access control for body sensor networks," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 1, pp. 44–51, Jan. 2010.
[12] S. Marinkovic, E. Popovici, C. Spagnol, S. Faul, and W. Marnane, "Energy-efficient low duty cycle mac protocol for wireless body area networks," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, pp. 915–925, Nov. 2009.
[13] B. Kang, J. Im, C. Chung, and J. Kim, "A energy-efficient system by reducing beacon listening for periodic vital sign monitoring," in *Proc. IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Seoul, South Korea, Aug. 2011, pp. 1–4.
[14] Y. Tselishchev, L. Libman, and A. Boulis, "Reducing transmission losses in body area networks using variable tdma scheduling," in *Proc. 2011 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Jun. 2011, pp. 1–10.
[15] B. Liu, Z. Yan, and C. W. Chen, "CA-MAC: A hybrid context-aware mac protocol for wireless body area networks," in *Proc. 13th IEEE International Conference on e-Health Networking Applications and Services (Healthcom)*, Columbia, MO, Jun. 2011, pp. 213–216.
[16] Z. Yan and B. Liu, "A context aware mac protocol for medical wireless body area network," in *Proc. 7th International Wireless Communications and Mobile Computing Conference (IWCMC 2011)*, Istanbul, Turkey, Jul. 2011, pp. 2133–2138.
[17] M. Ameen, J. Liu, S. Ullah, and K. S. Kwak, "A power efficient mac protocol for implant device communication in wireless body area networks," in *Proc. 2011 IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, USA, Jan. 2011, pp. 1155–1160.
[18] A. Argyriou, A. Breva, and M. Aoun, "Optimizing data forwarding from body area networks in the presence of body shadowing with dual wireless technology nodes," *IEEE Trans. Mobile Comput.*, early access, 2014.
[19] N. Torabi and V. Leung, "Realization of public m-health service in license-free spectrum," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 1, pp. 19–29, Jan. 2013.
[20] V. Chaganti, L. Hanlen, and T. Lamahewa, "Semi-markov modeling for body area networks," in *Proc. 2011 IEEE International Conference on Communications (ICC)*, Jun. 2011, pp. 1–5.
[21] D. B. Smith, L. W. Hanlen, J. A. Zhang, D. Miniutti, D. Rodda, and B. Gilbert, "First-and second-order statistical characterizations of the dynamic body area propagation channel of various bandwidths," *Annals of Telecommunications*, vol. 66, no. 3-4, pp. 187–203, Apr. 2011.
[22] B. T. Korel and S. G. Koo, "A survey on context-aware sensing for body sensor networks," *Wireless Sensor Network*, vol. 2, no. 8, pp. 571–583, 2010.
[23] B. Korel and S. Koo, "Addressing context awareness techniques in body sensor networks," in *Proc. 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW)*, May 2007, pp. 798–803.
[24] D. Roggen *et al.*, "OPPORTUNITY: Towards opportunistic activity and context recognition systems," in *Proc. IEEE WoWMoM*, Jun. 2009, pp. 1–6.
[25] S. Fox, J. Naughton, and W. Haskell, "Physical activity and the prevention of coronary heart disease," *Annals of Clinical Research*, vol. 3, pp. 404–432, 1971.
[26] G. L. Nemmhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*, 1st ed. Wiley-Interscience, 1988.
[27] Y. He, W. Zhu, and L. Guan, "Optimal resource allocation for pervasive health monitoring systems with body sensor networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 11, pp. 1558–1575, Nov. 2011.
[28] H. Su and X. Zhang, "Battery-dynamics driven tdma mac protocols for wireless body-area monitoring networks in healthcare applications," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 4, pp. 424–434, May 2009.
[29] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992.
[30] J. Lofberg, "YALMIP : a toolbox for modeling and optimization in MATLAB," in *Proc. 2004 IEEE CACSD Conference*, Sep. 2004, pp. 284–289.
[31] Crossbow technology,crossbow datasheet on micaz. [Online]. Available: http://www.xbow.com

**Bin Liu** received the B.S. and M.S. degrees, both in electrical engineering, from University of Science and Technology of China, Hefei, Anhui, China, in 1998 and 2001, respectively, and the Ph.D. degree in electrical engineering from Syracuse University, Syracuse, NY, in 2006. Currently, he is an Associate Professor with the School of Information Science and Technology, University of Science and Technology of China. His research interests are signal processing and communications in wireless sensor and body area networks.



**Zhisheng Yan** is a PhD student at Computer Science and Engineering Department, State University of New York at Buffalo. His research interests lie in perception, processing and networking of multimedia data. Currently, his research is focused on energy-saving mobile display and wireless HTTP adaptive streaming.

**Chang Wen Chen (F'04)** is a Professor of Computer Science and Engineering at the State University of New York at Buffalo, USA. Previously, he was Allen S. Henry Endowed Chair Professor at Florida Institute of Technology from 2003 to 2007, a faculty member at the University of Missouri - Columbia from 1996 to 2003 and at the University of Rochester, Rochester, NY, from 1992 to 1996. He has been the Editor-in-Chief for IEEE Trans. Multimedia since 2014. He has also served as the Editor-in-Chief for IEEE Trans. Circuits and Systems for Video Technology from January 2006 to December 2009 and an Editor for Proceedings of IEEE, IEEE T-MM, IEEE JSAC, IEEE JETCAS, and IEEE Multimedia Magazine. He and his students have received eight (8) Best Paper Awards or Best Student Paper Awards and have been placed among Best Paper Award finalists many times. He is a recipient of Sigma Xi Excellence in Graduate Research Mentoring Award in 2003, Alexander von Humboldt Research Award in 2009, and SUNY-Buffalo Exceptional Scholar - Sustained Achievements Award in 2012. He is an IEEE Fellow and an SPIE Fellow.