Energy Efficient Wireless Body Area Network Design in Health Monitoring Scenarios

by

Yang Zhou

B.Sc., Beihang University, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Electrical and Computer Engineering)

The University of British Columbia

(Vancouver)

March 2017

Abstract

Wireless body area networks (WBANs) are one of the key technologies that support the development of ubiquitous health monitoring, which has attracted increasing attention in recent years. Wireless on-body sensors free the patients from countless tangled wires, and wireless implanted sensors make it possible for the doctors to monitor an extensive range of critical bio-information continuously, which is crucial for a quick reaction when emergency happens. Due to the size limitation on the sensor nodes and the importance of the life signals transmitted, compared with general wireless sensor networks (WSNs), WBANs have more stringent requirements on reliability and energy efficiency during data's collection and transmission. This thesis aims to propose effective network designs to increase packet delivery rate, reduce energy consumption and prolong network lifetime for WBANs.

In order to solve the major challenges faced by WBANs, due to the energy efficiency and reliable data transmission requirements, in this thesis, network design over multiple layers are considered, including physical layer, medium access control (MAC) layer and routing layer. Network topology design that is suitable for WBANs is also considered. Specifically this thesis:

1. Investigates the design of MAC protocols and proposes an opportunistic scheduling scheme by applying heuristic scheduling and dynamic superframe length adjustment to improve the packet delivery rate and improve transmission reliability;

2. Formulates and solves a mathematical optimization problem to maximize network lifetime, which jointly considers network topology design, transmission power control and routing strategy. Multilevel primal and dual decomposition methods are employed to solve the proposed non-convex mixed-integer optimization problem. A solution with fast convergence rate based on binary search is provided.

Simulations have been conducted to show that our proposed network design increases network performance to a large extent compared with existing solutions.

Preface

This thesis is original, independent work conducted by the author Yang Zhou. A portion of this research has been published in conference proceedings, and the rest has been submitted to IEEE journal. The publications are revised by my supervisor Dr. Victor C.M. Leung, co-supervisor Dr. Peyman Servati, and my colleague Dr. Zhengguo Sheng, Dr. Chinmaya Mahapatra.

A version of Chapter 3 has been published. Yang Zhou, Zhengguo Sheng, Victor C.M. Leung, Peyman Servati, "Beacon-based Opportunistic Scheduling in Wireless Body Area Network", *Proc. of 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBC), 2016, August 16-20, Orlando, USA. Dr. Leung and Dr. Servati provide helpful discussions. Dr. Sheng help with the revisions and proofreading the manuscript.

Materials in Chapter 4 has been accepted by journal of Ad Hoc Networks. Yang Zhou, Zhengguo Sheng, Chinmaya Mahapatra, Victor C.M. Leung, Peyman Servati, "Topology Design and Cross-layer Optimization for Wireless In-Body Sensor Networks". Dr. Leung and Dr. Servati provide helpful discussions. Dr. Mahapatra revised the manuscript. Dr. Sheng help with the revisions and proofreading the manuscript.

Table of Contents

List of Tables

List of Figures

Glossary

The following are the abbreviations and acronyms used in this thesis, listed in alphabetical order:

AAL   Ambient Assisted Living

AP   Access Point

CSMA/CA   Carrier Sense Multiple Access with Collision Avoidance

DoS   Denial of Service

EEG   Electroencephalogram

EMG   Electromyography

HBC   Human Body Communication

ISM   Industrial Scientific Medical

MAC   Medium Access Control

PAN   Personal Area Network

PDA   Personal Digital Assistant

POMDP   Partially Observed Markov Decision Process

QoS   Quality of Service

SAR   Specific Absorption Rate

SMS   Short Message Service

UWB   Ultra Wide Band

WBAN   Wireless Body Area Network

WLAN   Wireless Local Area Network

WSN   Wireless Sensor Network

Acknowledgments

First I would like to express my sincere appreciation to my supervisor Dr. Victor C.M. Leung, and co-supervisor Dr. Peyman Servati, from whom I got amazingly patient guidance and constructive help. Working on a crossing field of wireless communication and body sensors is challenging, I meet hard situations frequently. However, luckily, my supervisor and co-supervisor are the top experts in these two fields and I always regain the courage to continue my research after meeting with them. Looking back on what I have done for these two years, I am quite amazed by how far I have gone on my research topics and one thing is for sure that I could not do these by myself without the help from my supervisors. Additionally, I want to acknowledge the financial support from Canadian Natural Sciences and Engineering Research Council, strategic project for "wearable health monitoring" led by Dr. Peyman Servati, which gives me a huge relief on the burden of daily life.

Secondly, I am grateful to Dr. Dave Michelson, who dedicates his valuable time to serve on my examining committee. The invaluable feedback and comments he provides are essential for improving the quality of my thesis and perfecting my work.

Also, I would like to express my gratitude to Dr. Zhengguo Sheng, who dedicates plenty of his working time to guiding my research work. We have routine meetings weekly and the discussions with him serve as the foundation for my research. Even when Dr. Zhengguo Sheng is working as an instructor in Britain recently, we still schedule phone and skype meetings discussing plenty of research topics. He is a big help for my research and good friend in daily life too.

Finally, I want to express my gratitude to my parents and my girlfriend. My parents always tell me to conquer the challenges and believe in myself. They are my mentor and always encouraging me to try what I want. And I want to thank my girlfriend Shanshan, who is so supportive to my work and choice. I am so lucky to meet such an amazing girl. She gives me the courage to complete my research and continue my work in industry.

Chapter 1

Introduction

1.1 Motivation

Our world is facing a globally aging population, as introduced in [1–5], which results from both lifetime extension and Baby Boomers demographic peak. The current healthcare system suffers from severe challenges that has never met before. The healthcare expenditure has been growing enormously. As a rough estimate, the total expense spent on healthcare is expected to reach 20% of the Gross Domestic Product (GDP) in United States [1] in 2022, which will be an extremely heavy burden for the country's economy.

However, at the same time, there are huge amount of people that receive inadequate healthcare monitoring because of scarce medical resources compared with the great volume of demands. There are currently 180 million people worldwide that suffer from diabetes and the number is expected to grow to 360 million by 2030. Debilitating neuro-degenerative disease such as Parkinson's affects even more [2]. These patients' life can be greatly relieved if a health monitoring system is established with low cost and provide continuous services to the general public. The growing pressure imposed on the public healthcare expenditure can be alleviated as well. As indicated by the recent researches, wireless body area network (WBAN) is a promising choice for building such an health monitor system [4].

With the fast development of tiny-shaped, low power consumption, and low cost biosensors [6–9], WBAN now has the potential to fulfill the monitoring, diagnostic and therapeutic functions with barely noticeable influences to people's normal lives. These tiny medical sensors locates on the surface of or implanted to the human bodies, monitoring the health status and sending the bio-signals wirelessly to the data centre. Specifically, there should be a coordinator, and generally a

1

cell phone is suitable for this role, to communicate with the body sensors. The coordinator will then send the collected data to the cloud through cellular network, WIFI or any other equivalents. With the instant data collected from body sensors, doctors will have a much better picture of the health condition of the patients. This health status monitoring system allows for fast reactions to health problems of a wide range of populations, regardless of where these people locate. Since the sensors are of low cost and can be manufactured extensively, compared with the expensive and scare equipments in the hospital, a much greater population can be covered by this WBAN monitoring system than traditional medical care system.

However, the wide application of WBANs in medical scenarios is seriously hindered by the concerns on reliable data transmission and short lifetime span of the network. Though limited within an area of several meters wide, the channel condition on or in the human body is highly unstable and unpredictable [10–12]. One major problem is that people change their types of physical activities constantly, which results in significantly different and unpredictable channel attenuations. The complexity of the body tissues makes it even harder to set up a channel model and come up with effective transmission strategies to ensure reliable data transmissions. Thus, as mentioned in [10], since the emergency signals in WBANs may lead to the difference between life and death, the reliability is one of the most important considerations.

Another major problem arises from energy limitation for the tiny sensors, especially for the implantable ones. On one hand, the volume of the bio-sensor nodes needs to be tiny enough so that they can be wearable or implantable. The size limitation on the battery leads to very limited energy storage and the sensors need to be extremely frugal during their energy usage since their batteries are not likely to be replaced very often. Effective data transmission strategies need to be proposed to reduce the power consumption of the sensor nodes as much as possible, while maintaining reliability of the transmissions.

In this thesis, effective network designs across multiple protocol layers to achieve higher reliability and longer network lifetime are proposed, providing feasible approaches to solve the major problems for WBANs mentioned above in medical scenarios.

1.2 Wireless Body Area Network

1.2.1 Overview

A WBAN is the composition of a group of energy efficient, miniatured, invasive/non-invasive lightweighted wireless sensors that monitors human bodies' health condition and surrounding en-

2

vironments [13]. Medical care system is one of the most crucial application scenarios for WBANs as described in IEEE 802.15.6 standard [14]. There are 3 subcategories for the WBANs in medical applications: wearable WBAN, implant WBAN and remote control of medical devices [1]. A wearable WBAN is capable of monitoring sleep disorder, asthma, assessing soldier fatigue and battle readiness, aiding sport training to avoid injuries. For implant WBANs, the sensors are either implanted below the surface skin or reside in the blood stream. Functions such as diabetes control, cardiovascular diseases monitoring and cancer detection can be fulfilled by implant WBANs. The remote control of medical devices allows for Ambient Assisted Living (AAL), in which each WBAN exchange information with a back-end medical network [15]. AAL allows for automatic medical care for the patients through the WBANs, freeing patients from intensive personal care.

According to the role in network, the nodes in WBANs can be classified into 3 categories: coordinator, end nodes and relays. The coordinator fulfills the function of communicating with both the nodes in WBAN and the outside world. It receives human body monitoring information from the sensors, and transmit the information to the medical care centre through wireless networks, such as Wireless Local Area Network (WLAN), cellular network, Zigbee or Bluetooth. The coordinator may also send direct notifications to the users if necessary. Generally, the coordinator is a Personal Digital Assistant (PDA). The end nodes and the relay nodes are miniaturized, low power sensor motes, which consists of sensing unit, power unit, communication unit, storage unit and processing unit. These nodes communicate with low power, short range wireless communication protocols, such as Bluetooth Low Energy, Zigbee, or Ultra Wide Band (UWB) protocol [16]. The end nodes are limited to collecting the sensing information and transmitting the data to the coordinator or relay nodes. The relay nodes act as intermediate nodes when the end nodes are far away from the coordinator. The relay nodes may fulfill the function of collecting monitoring data as well. Sometimes, there are nodes called "actuator" incorporated in the WBAN. The actuators react with certain operations to the patient when receiving certain signals from the sensors [17]. For instance, the actuator can functioning as pumping the proper dose of medicine into patient's body at the right moment [18].

A unique identifier, which is an one-octet byte, is assigned to the sensor nodes and coordinator for frame exchanges. Following this rule, a typical medical network based on WBANs can support up to 256 nodes [19, 20]. Except for this general requirement, the network size will also be limited by other network parameters, such as superframe length or the number of channels available [17]. According to IEEE 802.15.6 standard, one-hop and two-hop tree topology networks with coordinator at the centre are supported. The coordinator locates at the location such as the waist [21, 22]. Both beacon mode and non-beacon mode exist for the data communication. In the beacon mode,

the coordinator will broadcast the control signal at the start of the superframe with the beacon frame. The control signal includes resource allocation information and synchronization information. The superframe length, or the length of beacon period, can be specified by the user and based on WBAN's standard [23, 24]. In the non-beacon mode, the sensors applied Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) to send data to the coordinator. A beacon based TDMA scheme is preferred by WBAN for the reason of low communication overhead, low transmission delay and free of collisions [11].

1.2.2 Architecture

Generally, WBAN-based healthcare monitoring systems are composed of three tiers of communications [1, 4, 13, 17]:

- Tier 1: Intra-WBAN communication

- Tier 2: Inter-WBAN communication

- Tier 3: Beyond-WBAN communication

Figure 1.1 gives an illustration on the working process of these three components. In tier 1 communication, bio-sensors, such as Electromyography (EMG) sensors, pulse oximeter, electroencephalogram (EEG) sensors, collect the life signals of human body and transmit these signals to a coordinator. In WBAN, this coordinator acts as a sink node, and communicates with all of the sensor nodes. A cell phone, or any other PDA, is generally a good choice for the role of coordinator. The coordinator should have access to the internet Access Point (AP). Once coordinator collects the data from the sensors, it will forward the data to the AP in tier 2 communication. In tier 3 communication, the life signal message will start from AP and be routed through the internet to the medical data centre, where doctors or auto-diagnose system will react to the abnormal situations.

Tier 1: Intra-WBAN communication

In this level, the communications among the body sensors and the communications between the body sensors and the coordinator is considered. The data transmission range is around 2 meters in or around human body. The network design in tier 1 is one of the most crucial research focus in WBAN, since the Intra-WBAN communication needs to solve special challenges brought up by the human body wireless communication environment.

Generally, for WBAN, star topology will work since the size of network is tiny [25–27]. However, owing to the low transmission power and irregular body movements, great challenges exist
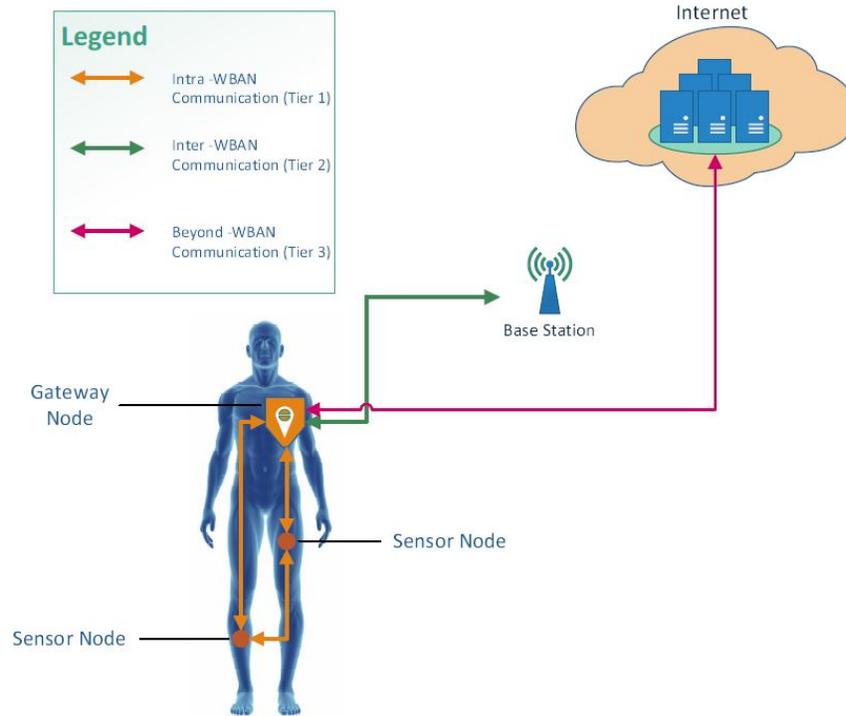
Figure 1.1: WBAN-based healthcare monitoring architecture

for reliable wireless communications [10]. Multi-hop network results in lower transmission power, however higher transmission delays [1]. Because the multi-hop network shortens the transmission distance, the data transmission reliability can be improved. As proposed in the IEEE WBAN standard [14], there can be at most two hops in IEEE WBAN standards compliant communication. The reason is that, with the increase of the hopping number, the communication complexity and overhead will be increased as well. For a network with moderate size, a network design with more than 2 levels may not incur extra benefits.

Tier 2: Inter-WBAN communication

The aim for the communications in this layer is to connect the WBAN with the broader networks that we use everyday in our daily lives, such as the cellular network and Internet. This connection is achieved by the communication between the coordinator defined in intra-WBAN, and one or more Access Points (APs). The APs can be integrated as a part of the infrastructure, which is

5
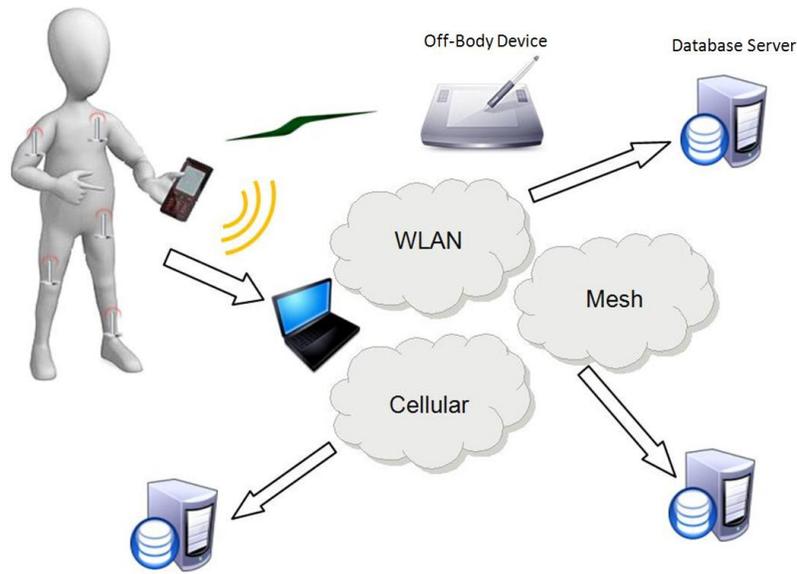
Figure 1.2: Inter-WBAN Communication: Infrastructure-based architecture

the infrastructure-based architecture, or can be placed dynamically, which is the ad hoc-based architecture.

Infrastructure-based architecture: the infrastructure-based architecture is centralized, as is shown in Figure 1.2, and all of the BANs in this area communicates with the same AP. Generally, Inter-WBAN communication is confined within a limited space, for example, a waiting room in the hospital. The major advantage of the infrastructure-based architecture is that it allows for centralized management and security control.

Ad hoc-based architecture: as is shown in Figure 1.3, the ad hoc-based architecture is composed of multiple APs, which forms a mesh structure. Because of the characteristic of ad hoc network, it allows for dynamic and flexible deployment. The network can be enlarged with minor influence to the rest of the network. The multi-hop network ensures a larger area coverage compared with the infrastructure-based architecture, which greatly facilitates users's mobility.

Tier 3: Beyond-WBAN communication

Beyond-WBAN communication considers the communication between AP and the outside world, including internet and remote electronic medical care centres. One of the cornerstones of Tier-3 is the database that stores user's profile and medical history. The doctor will access to the patient's

6
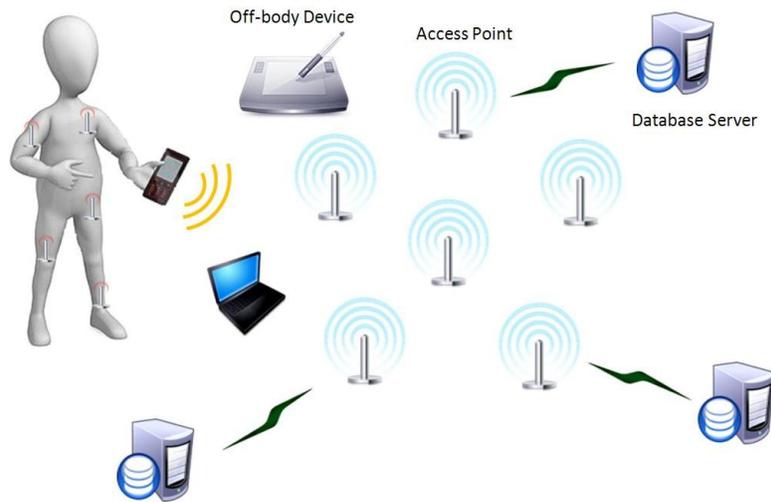
Figure 1.3: Inter-WBAN Communication: Ad hoc-based architecture

information when needed. Automatic notifications can be set to send emergency alarms to both doctors and patients when emergency status appears through Internet or short message service (SMS). Another potential application scenario is remote disease diagnose as indicated in [3, 4, 13, 28]. The doctor can remotely obtain all of the information needed from the wireless sensors worn by patient and the historical data stored in the database.

1.2.3 Challenges

WBANs compose a subset of general WSNs and they share many common challenges [29]. Nevertheless, a great many differences exist between these two types of networks with respect to extreme reliability requirement, extreme energy efficiency, special designs of the bio-sensor nodes, bio-channel complexity, security and mobility issues [10, 12, 30–33]. For WBANs in medical scenario, reliability is of vital importance since the signal collected and transmitted can be life-critical. At the same time, other challenges listed below need to be tackled efficiently.

Extreme energy efficiency

The sensor nodes in WBANs are very different from the ones in WSNs. In WBANs, the sensor nodes need to be extremely tiny in order to be wearable and implantable [31]. Large sensors on-body or in-body will heavily influence patients' daily lives or even body functions. The size

requirement sets a strict limitation on the battery size of the nodes. However these nodes are expected to work for years, especially for the implanted ones, which cannot be recharged or replaced without great pain. So the sensor nodes need to be extremely frugal in consuming their energy. The energy consumption of the sensor nodes includes data collection, data processing and data transmission. In typical sensor applications, energy consumption is dominated by sensors' radio consumption [34]. So the data transmission need be managed wisely, which is one of the topics discussed in this thesis.

Unique characteristic of the bio-channel

To ensure the energy efficiency, the transmission power of the sensor nodes are generally low. At the same time, because of the relatively small antennas and simple energy-efficient designs, the receivers' sensitivity level is not high either. The severe attenuation, which is over 100dB, that can occur around or within human body will cause transmission outages [10].

The body channel model is a very complex research issue, which is still an open research topic [12]. One difficulty results from the complexity of body tissues' characteristics and their effects on wireless signals, which is still not a well researched field. Another difficult results from the fact that in WBANs, unlike in static networks such as WSNs, the received signal strength is mostly affected by users' body movements instead of fixed distances. Since human body movement is generally unpredictable and relates to users' personal habits, except for some simple scenarios such as running or walking, an effective model is very hard to be established.

To tackle this problem, in this thesis, I apply reinforcement learning techniques, which are model-free methods, to the WBAN and make it possible to react to the channel fluctuation.

Interference

In a WBAN, the coordinator controls the data transmissions with the body sensors centrally to avoid collisions. This works fine when no external sensors goes inside or outside of the WBAN. Things become pretty complicated when multiple persons, all wearing body sensors, come close together. In this case, the coordination strategy designed fails. Generally, people's actions are unpredictable, which means the coordinator is kind of blind to the nodes coming in or going out of the range it controls. Because of this reason, an effective interference mitigation scheme is expected to be able to adapt faster than the actual topology changes [10].

Security

Considering the specialty of the data processed in WBANs, which contains potentially important health information and private information that people do not want to share, data security is a serious topic. The data stored in the healthcare data centre cannot be modified illegally, and should be always available to the doctors even under attacks, such as denial-of-service (DoS) attacks. According to [35], as is demonstrated in McAfee conference, it is possible to inject a deadly dose of insulin to the pumps by hacking the insulin pump without any knowledge of the device. Also a pacemaker transmitter could be reverse engineered and hacked to deliver a deadly electric shock with a maximum voltage of 830 volts, resulting in a simulated cardiac arrest as well as continuous shocks. The attack of WBANs can result in severe results and should be considered with top priority. However, the problem is that the body sensor are limited in energy, data processing ability, storage and data transmission ability, so public key encryption system is not suitable for WBANs [12]. Though, a bunch of WBAN specific security designs have been proposed [36–39], it is still a crucial open research field.

1.2.4 IEEE 802.15.6 BAN Protocol

In November 2007, the IEEE 802.15.6 working group was created and started to draft the first international WBAN standard. IEEE 802.15.6 standard aims to support the in-body and near-body wireless communication applications. The existing industrial scientific medical (ISM) bands as well as frequency bands approved by national medical and/or regulatory authorities are used in the protocol. The support of various carrier frequency bands is crucial, since some applications have better performances or are required to operate at certain frequencies. For instance, the implanted sensors can only operate at the band 402-405MHz worldwide. The standard takes human influences into consideration, such as the effects on the portable antennas due to the presence of a person (varying with male, female, skinny, heavy, etc.), radiation pattern shaping to minimize the specific absorption rate (SAR) into the body, and communication behaviour reactions to the user motions [14].

Flexibility is the most appealing feature of IEEE 802.15.6 standard [10]. The standard supports three different PHY layers: narrowband PHY, ultra-wideband (UWB) PHY and human body communication (HBC) PHY. Two types of MAC layer are offered, which are the contention-based and scheduling-based channel accessing schemes [40]. Each device only needs to choose the proper PHY layer and accessing method according to its specific requirements, such as energy efficiency, transmission delay, device costs etc.

As is concluded in [10], despite of the suitability of IEEE 802.15.6 for WBAN, a successful network design should be combined with the investigations on the body channels. The network protocol should be able to react and handle the highly dynamic body channel environments, which is the major focus of this thesis.

1.3 Research Problems

Based on the challenges introduced in the previous section, specifically, we focus on achieving reliable and energy efficient wireless communications among the body sensors and the coordinator. Draft IEEE 802.15.6 standard is a major reference for WBAN system. However, for flexibility, it leaves a bunch of high level open questions such as the choice of scheduling method, how should relay be used and what is the proper transmitting power for the sensors [10]. These are the vital considerations when designing a reliable and energy efficient WBAN, which are also the major focuses for this thesis.

On MAC layer, because of the fluctuating characteristic of the body channel, traditional static scheduling that schedules transmissions regardless of channel fluctuation will lead to bad performance since the bad channels are not avoided and the good ones are not fully exploited. In this thesis, we discuss the effectiveness of applying opportunistic scheduling with low overhead in WBAN background, where the length of superframe and the transmission scheduling are adjusted according to the channel status. By scheduling the data transmission more wisely, lower outage rate is achieved as illustrated by the simulations in Chapter 3.

Beyond MAC layer design, relay nodes and cross-layer optimization are potentially effective approaches to optimize the network's energy consumption and prolong network's lifetime [41–49]. However, under WBAN context, a cross-layer optimization scheme that jointly considers relay positioning, transmission power control and routing strategy is still missing. A cross-layer approach leads to a much larger feasible set for the optimal solution and yields better solution than the existing methods that optimize each layer independently as will be illustrated in Chapter 4. One challenge for such an optimization problem is that the formulated problem is non-convex and mixed-integer so the optimal solution cannot be obtained from solver directly. Special algorithms need to be designed to derive the optimal solutions efficiently.

1.3.1 Medium Access Control Design

For the majority of the existing MAC protocols, TDMA scheme with master/slave star topology is adopted [50]. For example, both IEEE 802.15.4 and IEEE 802.15.6, which are widely used for

10
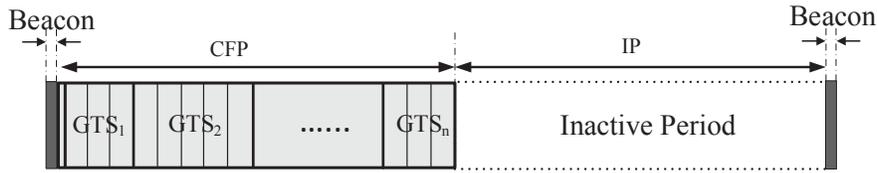
Figure 1.4: Superframe structure of TDMA MAC design

WBANs suggest a beacon-enabled TDMA working mode. A typical superframe is composed of a beacon frame, a Contention Free Period (CFP) and an Inactive Period (IP) as is shown in Figure 1.4. The beacon frame is used for synchronization and broadcasting resource allocation information. In CFP the coordinator allocates Guaranteed Time Slots (GTS) for periodic data transmissions. During IP, all sensor nodes will stop transmissions and sleep for energy saving.

The general TDMA scheme may serve well in a static network such as WSNs. However since mobility is a key characteristic for WBAN, the scheduler should be smart and able to learn the characteristics of the channel fluctuations [10]. For instance, the retransmission should not happen in the immediate future after the time when outage happens [11], because the channel condition may not yet recover from the previous outage. As in Figure 1.4, common schedulers operate regardless of the channel condition. Opportunistic scheduling, on the other hand, is able to take channel state into consideration. Since channel fluctuation is a major influence on the performance of WBANs, opportunistic scheduling has the potential of improving the system performance to a large extent by avoiding the bad links and take advantage of the good ones.

The Gilbert model [51], which is a two-state Markov chain model, has been proved to be powerful and simple enough to model the body channels [11, 52, 53]. There are two channel states defined for each time slot, namely "good" and "bad" that are represented by "1" and "0" respectively. Although there exists more complicated Finite-State Markov Channel (FSMC) model [54] and it can model Rayleigh fading channels quite well, it is not necessarily helpful for WBANs. Actually, according to the 802.15.6 group [55] and research [56], on-body channels can barely be modelled by Rayleigh fading.

Because energy is extremely limited in WBANs, probing the channels is not suitable. The only chance to observe the channel status for the coordinator is when the node is transmitting. At the other time slots, we use state belief instead of the exact channel status to represent the channel condition, that is, the channels are partially observed. If the transition matrix at time slot k for node

i is $\mathbf{P}^i(k) = \begin{pmatrix} p_{00}^i(k) & p_{01}^i(k) \\ p_{10}^i(k) & p_{11}^i(k) \end{pmatrix}$ and the initial state belief of node i to be $\begin{bmatrix} p_0^i(0) \\ p_1^i(0) \end{bmatrix}$, the state belief followed at the time slot n can be calculated as

$$\begin{bmatrix} p_0^i(n) \\ p_1^i(n) \end{bmatrix}^T = \begin{bmatrix} p_0^i(0) \\ p_1^i(0) \end{bmatrix}^T \prod_{k=1}^{n} \begin{pmatrix} p_{00}^i(k) & p_{01}^i(k) \\ p_{10}^i(k) & p_{11}^i(k) \end{pmatrix} \tag{1.1}$$

where $p_0^i(0)$ is the probability of having a bad channel at the beginning and $p_1^i(0)$ is that of having a good channel. Normally "slot 0" is when transmission happens and the channel condition is known. The channel belief far away from the initial state is useless because it converges to the steady state.

Though it is safe to model WBAN channels using Markov process, great difficulties exist. The first problem is that the process is partially observable. We are only able to obtain the exact channel condition when a node is transmitting. Actually the problem is even more complicated than Partially Observed Markov Decision Process (POMDP) for which the transition model exists. Here we have no transition model and the transition probability may not even be stationary because humans change their body gestures constantly. These problems are the major reasons why opportunistic scheduling in WBANs is different from that of the other systems, such as cellular network.

As will be illustrated in Chapter 3, certain parameters has profound influences on the system performances. One parameter we focus on is the superframe length. It turns out that superframe length is a very crucial parameter that decides the outage rate of the data transmissions. Since opportunistic scheduling is based on Markov process as introduced above, an excessively large superframe length will lead to unreliable channel state estimations, which means that the state belief converges to the steady state and provides basically no useful information. However, on the other hand, if the superframe length is small, the channel does not change too much within a superframe. Then scheduling becomes trivial since it does not make too much difference where each node is scheduled. Furthermore, the suitable superframe length changes if the object modifies his/her body movements, so a "self-learning" scheduler is required in this context.

1.3.2 Network Cross-layer Optimization

There are many optimization literatures for WSNs [44–49]. However owing to the unique characteristics of WBANs as illustrated above, and according to the suggestion of IEEE WBAN standard [14] that the network should be two hops maximum, the optimization designs for WBANs can be further customized and achieve a better performance.

We focus on prolonging the lifetime of in-body implanted WBANs, where on-body relay nodes are applied. Relay nodes can shorten the transmission range of sensor nodes so that the sensor nodes are able to transmit with lower power and higher data rate. Specifically, we jointly consider the optimal locations of the relay nodes placed on the body surface, and the optimal transmission power and transmission time of the implanted sensors to achieve the maximum network lifetime. The decision of the optimal data routing on the routing layer, and the decision of transmission power and transmission time control on the physical layer are considered simultaneously. It can be seen further from problem model (4.4) that the considerations in these two layers cannot be decoupled.

The formulated problem is a mixed-integer and non-convex optimization problem, which can not be solved by existing solvers directly. In Chapter 4, we propose multilevel primal and dual decomposition approach to separate the original problem into independent smaller ones, and manage to find the solution to the original problem by solving the decomposed ones. More practical approach based on binary search algorithm that yields faster convergence rate is further developed and proposed in the thesis.

1.4 Contributions

In this thesis, a network model that covers network topology design, physical layer design, MAC layer design and routing layer design is proposed, ensuring a highly reliable and energy efficient WBAN for medical care scenarios. An effective opportunistic scheduling with low overhead that takes human body movements into consideration is proposed and shown to achieve lower outage rate than the existing scheduling schemes. Furthermore, a cross-layer optimization model that joint considers power control and relay selection is proposed to achieve the maximum network lifetime.

For the proposed opportunistic scheduling scheme, Gilbert model, which is a two state Markov chain model, is applied. Instead of adopting traditional method of setting a fixed transition matrix based on past experience, a novel estimation rule is proposed to achieve real-time updates to the transition matrix based on channel status. Since people change their physical movements constantly, the real-time updates on the transition matrix can reflect the channel dynamics more accurately. An effective heuristic scheduling scheme, which incur low extra overhead to the coordinator and no extra overhead to the sensors, is proposed and achieves lower outage rate than fixed scheduling and other opportunistic scheduling schemes. Based on the proposed heuristic function and channel dynamics estimation, the proposed opportunistic scheduling scheme is proved to be optimal mathematically.

Beyond mere scheduling strategy, during our research, superframe length is found to be a key parameter that significantly influences the outage rate performance of data transmission. As is explained in the previous section, the proper superframe length is decided by user's body channel dynamics. A long superframe is proper for a slowly changing channel and for a fast changing channel environment, a short superframe would be suitable. In order to interact with the body channel environment, a scheduler in the proposed system is designed as a decision maker that applies reinforcement learning strategies. Specifically, the scheduler utilizes Q-learning techniques. It observes rewards from the environment and updates Q values for the state-action pairs. The optimal action will be chosen according to the Q values when scheduler observes a certain state. The proposed reinforcement learning techniques for the superframe length adjustment makes opportunistic scheduling much more effective at avoiding the bad channel status and take advantage of the good one.

In order to achieve the maximum network lifetime of WBAN, an optimization model that jointly considers power control and relay selection strategy is proposed. The formulated optimization problem turns out to be a mixed-integer non-convex problem. In order to solve the proposed problem, multilevel primal and dual decomposition methods are applied. The original problem is decomposed into independent sub-problems confined in each body area, which can be solved by the commercial solvers effectively. Furthermore, in order to expedite the convergence rate, a method based on binary search is proposed to replace the original sub-gradient descent method to find the optimal values. Observations on the influence of data rates, the number of sensors, the number of relays and the body gestures shows the proposed optimization model yields consistently better results than the existing WBAN models.

Chapter 2

Literature Review

This chapter aims to provide a comprehensive review of the previous studies and researches conducted on energy efficient WBAN designs. In Section 2.1, Medium Access Control (MAC) techniques are reviewed with their advantages and disadvantages discussed. Specifically, in Section 2.1.1, the energy efficient MAC designs in WSNs and WBANs are presented. In Section 2.1.2, opportunistic scheduling is introduced, which has been proved to be greatly helpful for improving the network performance. Beyond MAC layer, in Section 2.2, the low energy designs in other layers are further discussed. In Section 2.3, an illustration on the dataset used to conduct simulations in this research is provided.

2.1   Medium Access Control Design in Wireless Body Area Networks

Lots of energy efficient MAC protocols have been developed for WSNs to make them live longer. Actually, some of these designs serve as a foundation for the WBAN MAC designs, such as the idea of periodic sleep mechanism. However, some fundamental differences exist between WSNs and WBANs, so that the existing MAC designs for WSNs are not proper for WBANs. The first problem is that the channel condition of WBANs is much less stable than that of WSNs, which means a fixed scheduling will result in bad performance since channel may fluctuate significantly within the scope of one superframe. The second issue is that the volume of bio-sensors are very tiny, especially for the implantable ones. They are much more power constrained than the other sensors and transmit with lower transmission power, which means these bio-sensors are quite vulnerable to the body channel fluctuation. However at the same time, the data transmissions are required to be highly reliable. So the requirement of avoiding the bad link status and take advantage of the good

15

channels for WBAN is very urgent.

Thus we conclude that opportunistic scheduling can be promising at promoting the performance of WBANs. However, most of the existing designs on opportunistic scheduling are for cellular network, which differs significantly from the WBANs. One major difference is that cellular network is much less power constrained compared with WBANs, so that channel probing can be extensively applied, which means the channel status is always available. The objective are different too. The cellular networks generally focus on throughput and fairness, while WBANs focus on reliability and energy consumption. So special opportunistic scheduling algorithms need to be developed for WBANs so that they are able to estimate channel status without probing it and make wise scheduling accordingly.

2.1.1   Energy Efficient Medium Access Control Design

For WSNs

The energy efficient MAC design emerges when WSNs become widely used. The most widely used method for saving energy is periodical sleeping and waking up schedule. Basically there are three types of MAC layer design for wireless sensor network: frame-slotted, synchronous and asynchronous [57].

Frame-slotted MAC: is designed for high throughput and low delay networks. It allocates time slots in a way that no two nodes within the two-hop communication neighbourhood are assigned to the same slot. This addresses collision and hidden terminal problem, providing a collision free data transmission environment. IEEE 802.15.4 (Zigbee) and TRAMA are two examples.

IEEE 802.15.4, which is a standard for low-power and low-cost Personal Area Networks (PANs), is based on frame-slotted MAC [24]. Before the outcome of IEEE 802.15.6 protocol, IEEE 802.15.4 is applied widely in Body Area Networks. A PAN is formed by one PAN coordinator which is in charge of managing the whole network, and by one or more coordinators which are responsible for a subset of nodes in the network. Ordinary nodes must associate with a coordinator in order to communicate. The supported network topologies are star, cluster-tree, and mesh topology. The standard supports two channel access methods: a beacon-enabled mode and a nonbeacon-enabled mode. The beacon-enabled mode provides a power management mechanism based on a duty cycle. It uses a superframe structure which is bounded by beacons. In the nonbeacon-enabled mode there is no superframe, nodes are always active (energy conservation is delegated to the layers above the MAC protocol) and use the unslotted CSMA/CA algorithm for channel access. As is

16
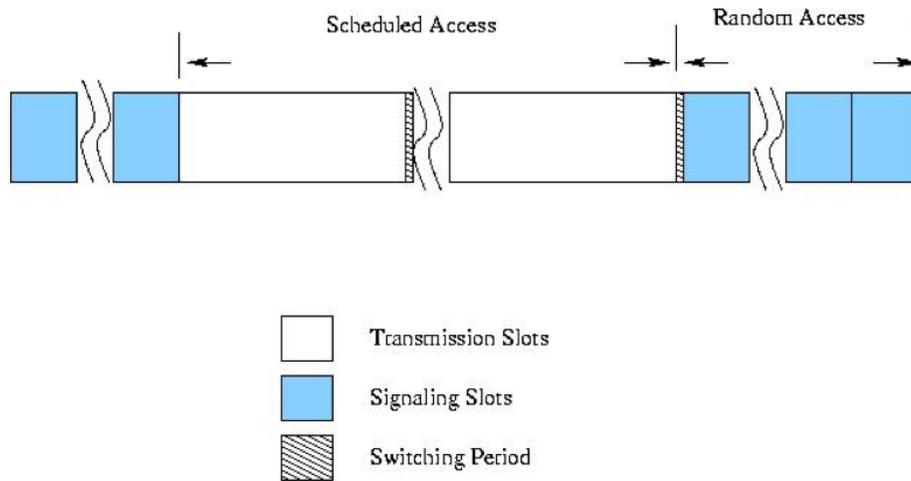
Figure 2.1: Time slot organization of TRAMA

describe, PAN is kind of similar to BAN in some aspects, and that is why IEEE 802.15.4 is used for BAN at an early stage. In fact, PAN is an extension of the general WSN, whose network size if smaller and concentrate solely on human body. BAN is a step further compared with PAN. While PAN focuses more on large wearable and portable devices for entertainment, BAN focuses more on miniature sensor devices that can be attached or implanted to human body for medical purpose. BANs need better solutions on energy efficiency, reliability and security issues.

The traffic-adaptive medium access protocol (TRAMA) is introduced for energy-efficient collision-free channel access in wireless sensor networks. TRAMA reduces energy consumption by ensuring that unicast, multicast, and broadcast transmissions have no collisions, and by allowing nodes to switch to a low-power, idle state whenever they are not transmitting or receiving. TRAMA assumes that time is slotted and uses a distributed election scheme based on information about the traffic at each node to determine which node can transmit at a particular time slot. Transmission scheduling is based on two-hop neighbourhood information and one-hop traffic information. TRAMA avoids the assignment of time slots to nodes with no traffic to send, and also allows nodes to determine when they can become idle and not listen to the channel using traffic information [58]. The working process of TRAMA is composed of alternative random access periods and scheduled access periods as in Figure 2.1. Random access period is used for synchronization and updating two-hop neighbour information. Scheduled access period is used for contention-free data exchange between nodes, which supports unicast, multicast and broadcast communication.

TRAMA protocol includes three parts: neighbour protocol, schedule exchange protocol and

adaptive election algorithm. TRAMA achieves a data transmission with no collision and no allocation of time slots to nodes with no data to transmit, which successfully ensures both low energy consumption and high data transmission rate. However the memory requirement of TRAMA is large to store the topology information of the nodes and scheduling information of neighbours. The computational complexity is large too. Each node has to calculate its two-hop neighbours' priorities and run AEA algorithm, which imposes a heavy burden on hardware requirement.

Synchronous MAC: Synchronizing active time of neighbouring nodes is a natural solution to establish communication between two nodes. However, it brings the cost of additional synchronization overhead.

S-MAC is contention based. It originates from 802.11 MAC protocol, aiming to achieve the requirement of wireless communication network [59]. It is one of the earliest WSN protocol that applies periodical sleeping techniques to reduce energy consumption, and supports self-configurations. It enables low-duty-cycle operation in a multi-hop network. Nodes form virtual clusters based on common sleep schedules to reduce control overhead and enable traffic-adaptive wake-up. S-MAC uses in-channel signalling to avoid overhearing unnecessary traffic. Finally, S-MAC applies message passing to reduce contention latency for applications that require in-network data processing. By lowering the duty cycle of each node together with overhearing avoidance and message passing, S-MAC has achieved a very low power consumption compared with other protocols with no sleep and wake-up mechanism. However, to trade for lower energy consumption, the transmission latency is increased and the throughput of the network is decreased. Also since the periodic sleeping scheme is unchanged, it means that S-MAC is not able to handle a workload-changeable situation.

T-MAC is developed from S-MAC and designed to cope with the issue of fixed duty cycle in S-MAC and handle the workload-changeable situation. Like S-MAC it is also a contention-based Medium Access Control protocol for wireless sensor networks. To handle load variations in time and location T-MAC introduces an adaptive duty cycle in a novel way: by dynamically ending the active part of it. A comparison between the scheduling in S-MAC and T-MAC is introduced in Figure 2.2. T-MAC reduces the amount of energy wasted on idle listening, in which nodes wait for potentially incoming messages, while still maintaining a reasonable throughput. However due to the mechanism of dynamic duty length adjustment, T-MAC suffers from early-sleep problem. This simply because when some node fails to detect other nodes' activity in Time Activity (TA) length time, these nodes will go to sleep, making signal transmission delayed.

Because S-MAC and T-MAC adopt a periodical sleep mechanism, the data transmission between the nodes will stop once in a while. Thus the delay will increase sharply when the number of

Figure 2.2: Scheduling comparison between S-MAC and T-MAC
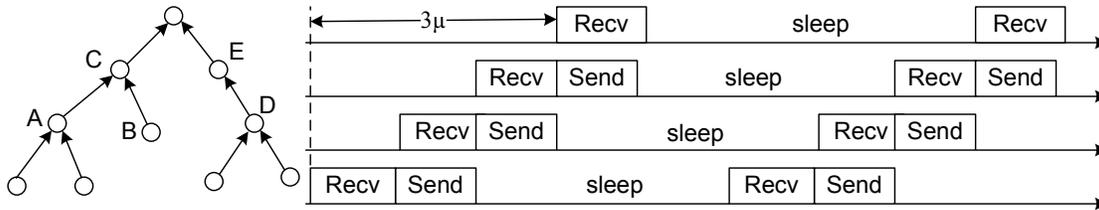
Figure 2.3: DMAC in a data gathering tree

nodes the message passes by increases. The principal aim of DMAC is to achieve very low latency, while keeping the energy efficiency high. For DMAC, slots are assigned to the sets of nodes based on a data gathering tree as shown in Figure 2.3. Hence, during the receiving period of a node, all of its children hold the transmission periods and contend for the medium. Low latency is achieved by assigning subsequent slots to the nodes that are successive in the data transmission path [60].

For DMAC, data can be transmitted non-stop along the multi-hop path, reducing the delay caused by periodical sleep. Furthermore, data rate can be adjusted by modifying the active time of nodes along the path. By data estimation, the interference among different children of the same parent node can be solved. All of these will contribute to a reduction on energy consumption as well as latency. However, in this method, the synchronization is strictly required, which can be a potential challenge. Collision avoidance methods are not utilized, hence when a number of nodes that have the same schedule (same level in the tree) try to send to the same node, collisions will occur. This is a possible scenario in event-triggered sensor networks. Besides, the data transmission paths may not be known in advance, which precludes the formation of data gathering tree.

Asynchronous MAC: Each node chooses its active schedule autonomously. Without paying the price for synchronizing neighbours' schedules. Asynchronous MAC protocols can achieve ultra-low duty cycle but have to search efficient ways to establish communication between two nodes.

WiseMAC is an optimized version of spatial TDMA and CSMA with preamble sampling. For
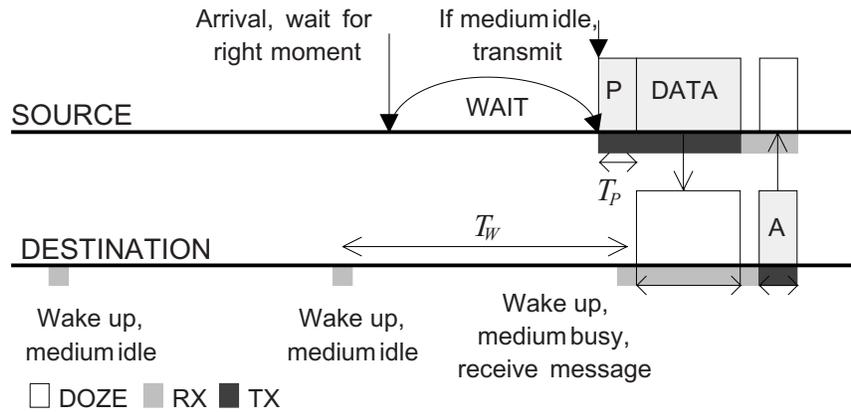
Figure 2.4: WiseMAC

the traditional version of preamble sampling method, the length of the preamble should be at least equal to sampling period. To reduce the power consumption incurred by the predetermined fixed-length preamble, WiseMAC offers a method to dynamically determine the length of the preamble. That method uses the knowledge of the sleep schedules of the transmitter node's neighbours. The nodes learn and refresh their neighbour's sleep schedule during every data exchange as part of the Acknowledgment message. In that way, every node keeps a sleep schedule table of its neighbours. Based on the neighbours' sleep schedule tables, WiseMAC schedules transmissions so that the destination node's sampling time corresponds to the middle of the sender's preamble Figure 2.4. To decrease the possibility of collisions caused by that specific start time of a wake-up preamble, a random wake-up preamble is preferred [61].

Another parameter affecting the choice of the wake-up preamble length is the potential clock drift between the source and the destination. A lower bound for the preamble length is calculated as the minimum of destination's sampling period, Tw, and the potential clock drift with the destination, which is a multiple of the time since the last ACK packet arrived. Considering this lower bound, a preamble length (Tp) is chosen randomly.

WiseMAC does has a uniform sleep-listen schedule. Nodes in a neighbourhood has different schedules. The broadcasted packets will be buffered for the neighbours in sleep mode and will be delivered many times. This will result in extra energy consumption. What is more, the hidden terminal issue is not coped in WiseMAC. It will lead to collisions when one node starts to transmit the preamble to a node that is already receiving another node's transmission where the preamble sender is not within range.

20

S wakes up shortly before predicted time on predicted channel
S sends DATA packet in response to beacon

time →

R sends another beacon, which also serves as ACK
R wakes up on pseudorandomly chosen (channel, time) and sends beacon

PKT Transmit   PKT Receive   ▢ Node active
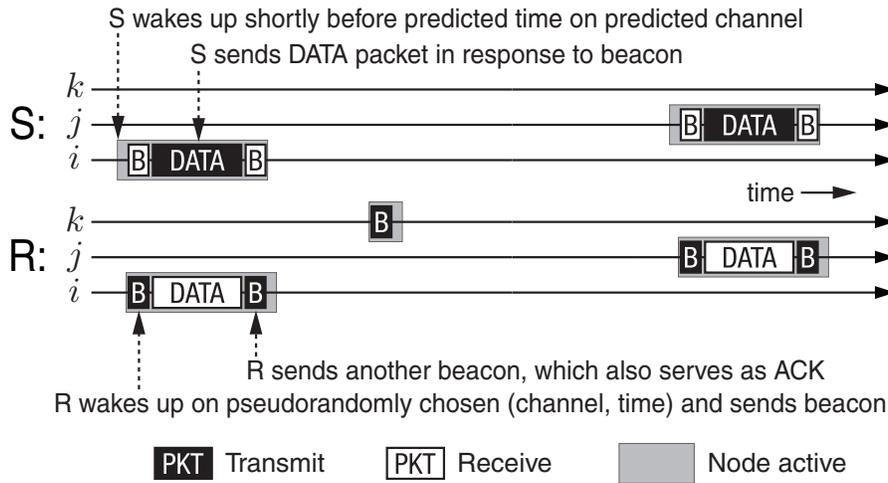
Figure 2.5: EM-MAC

EM-MAC is a multichannel asynchronous duty-cycling MAC protocol. It doesn't require nodes to synchronize their clock, does not use a common control channel, and does not explicitly exchange channel and wake-up schedules. Instead, every node independently decides its own pseudorandom channel-switching behaviour and wake-up times [62].

EM-MAC achieves high energy efficiency by enabling senders to accurately predict the wake-up channel and wake-up time of a receiver Figure 2.5. In particular, each time a node using EM-MAC wakes up, it independently selects its own wake-up time and channel according to a pseudo-random function, while avoiding undesirable channels on which it has detected high traffic loads or excessive wireless interference. The independent pseudorandom wake-up scheduling of EM-MAC aims to spread the traffic load to different channels, reducing wireless collisions caused by nodes waking up at the same time on the same channel. EM-MAC efficiently achieves multichannel rendezvous between a sender and a receiver with the prediction of a receiver's wake-up channel and wake-up time. A EM-MAC sender wakes up on the receiver's current wake-up channel right before the receiver does, finishes the packet transmission, and quickly goes back to sleeping state, achieving high energy efficiency through minimizing idle listening and overhearing.

Even though the performance of EM-MAC is better than other MAC protocols, the workload of EM-MAC is high. The sender has to estimate the channel condition to figure out suitable channels to choose. Also before sending data, the sender has to acquire the schedule of the aiming receiver, which requires lots of calculation and memory to process the data.
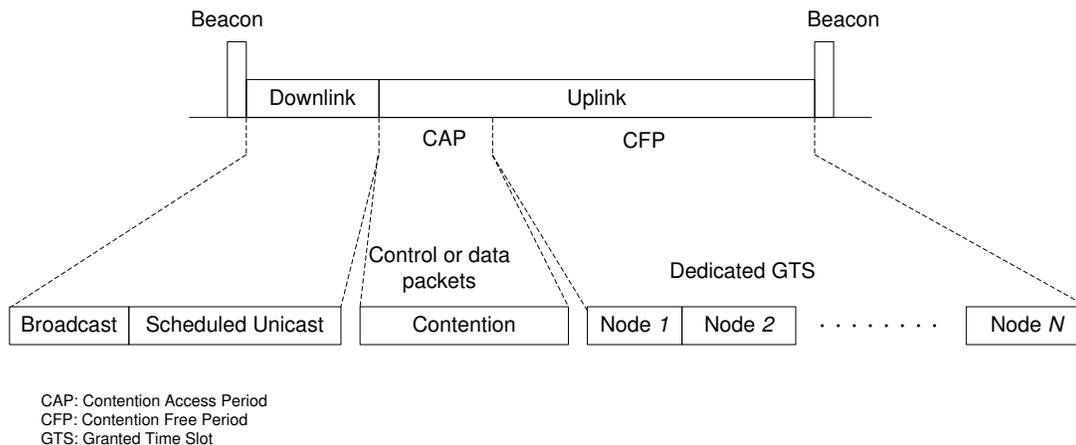
Figure 2.6: BodyMAC frame structure

For WBANs

As has been introduced previously, different from that of WSNs, the MAC design in WBANs needs further considerations on energy conservation, reliability and security issues.

BodyMAC, which is introduced in [26], is an MAC design suitable for WBANs. This work focuses on MAC frame structure, bandwidth management and sleep mode mechanisms by taking the above considerations into account. The MAC frame of BodyMAC can be divided into three parts: Beacon, Contention Access Period (CAP) and Contention Free Period (CFP). CAP is used for both bandwidth request and small size data transmission. In order to save energy, the node will not conduct synchronization unless the node meets a transmission failure in Guaranteed Time Slots (GTS). Since TDMA is applied in BodyMAC, the data transmission in CFP is collision free, which ensures energy efficiency. However, one potential drawback that is not fully discussed is the energy waste due to the failed transmissions caused by clock drift.

In [27], the authors proposed MedMAC aiming for Medical body area networks. It is also of single-hop start topology and TDMA-based. The main contribution is the proposition of Adaptive Guard Band Algorithm (AGBA) and Drift Adjustment Factor (DAF) to adjust the length of guard bands, which enables the nodes to sleep through a number of beacon periods and only listen for the beacons bounding the multi-superframes.

Addressing the QoS issue, authors of [63] have brought up with a hybrid and secure priority-guaranteed MAC protocol. This MAC design is adapted from IEEE 802.15.6 and assures security and priority issues.

However, these designs mentioned above neglect the profound influence of human body on the channel condition. High Packet Error Rate (PER) or even outage is expected to occur when the direct link is blocked by body parts [64]. Retransmission schemes can be used to deal with bad channel conditions. However it is only effective when the duration of bad channel is significantly shorter than the packet length. Retransmitting when suffering a long term bad channel condition leads to waste of energy, serious delay and low throughput. Relay thus can play a significant role in these scenarios where continuous transmission outage occurs.

In [65], a distributed dynamic scheduling scheme, which is based on Centralized BAN Access Scheme (CBAS) described in [66], is proposed to ensure the satisfaction of packet delivery rate constraint. This is achieved by letting the neighbouring nodes to copy the packet of the transmitting node during transmission. And once this transmitting node suffers a transmission failure, the same message can be relayed to the coordinator by one of its neighbouring nodes. According to the results, the dynamic scheduling scheme can achieve a very high packet delivery rate.

A cooperative relaying scheme that takes RSSI into account has been proposed in [64] is shown to achieve good outage rate performance. Another joint relay selection scheme in [42] has also shown good performance on outage rate. In this design, there are three sensors, two relays and a hub in the system. Each time, the sensor node will choose the best route from two relay routes and one direct route to transmit.

In [67], a multi-hop network has been proposed and achieves good throughput and delay performance. The traffic of the network is controlled by setting up a spanning tree. The scheme messages will be broadcasted and used by the parent and children of each node in the tree. The performance of this multi-hop network is better than single-hop network, however the improvement of performance is based on much higher complexity and resources consumption.

2.1.2   Opportunistic Scheduling

Common schedulers operate regardless of channel condition. Opportunistic scheduling, on the other hand, is able to take channel state into consideration. Since the channel fluctuation is a major influence on the performance of WBANs, opportunistic scheduling has the potential of improving the system performance to a large extent by avoiding the bad links and take advantage of the good ones.

A survey [68] provides a comprehensive summary on opportunistic scheduling. So far, the opportunistic scheduling researches largely focus on cellular networks, which differs significantly from WBANs. The most crucial difference is the requirement of extreme energy efficiency in

WBANs. In cellular networks, each user is able to spend extra energy for probing the channel before transmitting. However sensor nodes in WBANs are normally quite energy-limited and cannot afford this energy consumption. So the scheduler is kind of "blind" to the channel condition. The second significant feature of WBANs that makes the existing opportunistic scheduling method ineffective is that WBANs operate in a beacon-enabled TDMA mode, which is suggested by IEEE 802.15.6 protocol [69]. The coordinator will schedule the transmissions of all nodes at the beginning of each superframe, or each "round" of transmissions as will be called in the following. However most of the existing opportunistic scheduling designs can schedule only one transmission at a time. The objectives of these scheduling algorithms are generally fairness or stability, which are not the major concerns in WBANs.

Because of these special characteristics, several solutions have been proposed focusing on how to apply opportunistic scheduling to WBANs. [11, 52, 53] focus on applying Markov chain model to forecast the channel condition based on the transmission result from the previous round. Markov model has been proved to be effective by articles, such as [51, 70, 71], at modeling radio communication channels with burst errors. And in [11, 52, 53] the two-state Gilbert model is used to analyze the link status of WBANs in the future. In [11, 52], a flipping strategy is proposed with the assumption of homogeneous link status among all sensor nodes. The flipping strategy can be implemented with little overhead and it requires no priori knowledge on the statistical information of the channel condition. Each round the sensor node will be divided into two groups: nodes with "good" channel state and nodes with "bad" channel state in the previous round. The nodes with "good" channel will be scheduled ahead of the "bad" ones. Within the "good" group, the relative order of transmissions will be reversed while the relative order in the "bad" group will remain unchanged. The flipping strategy has been proved to perform better than the other alternatives. However the assumption of homogeneous link status among all sensor nodes is not realistic. The flipping strategy can be misleading when the speed of different channels converging to the steady states varies greatly. In [53], similarly, the author proposed a threshold-based scheduling scheme, which is also based on Gilbert model and monotonicity property. Each node is assigned a threshold on the delivery probability and the delivery probabilities of the time slots in a round will be calculated according to the Markov process. The time slots are assigned according to the thresholds and the delivery probabilities of the nodes. However, the transition matrix is actually unknown to the scheduler. In [53], the transition matrix is randomly chosen in a certain range. As has been mentioned previously, the stochastic characteristics of the channels differ and the transition probability matrix of the same link is likely to change instead of being stationary through time. The design in [53] is unable to figure out the real stochastic parameters online and the schedule made accordingly

24

cannot be reliable.

The fact that many daily activities of human body are periodic provides new methods for channel prediction and variable scheduling schemes. In [72], BANMAC is proposed and takes advantage of periodic human movements, such as walking and running. FFT is applied to transform RSSI time series into frequency domain and find the dominant frequency. A tight band-pass filter centered at the dominant frequency is used to filter the raw RSSI data. The transmission of each node will be scheduled near the peak of the predicted channel gain in the future and the pendulum task model is used. In BANMAC, the authors assume that the nodes exhibiting periodic channel fluctuations can be divided into two groups, nodes on the right side of the body and nodes on the left. The channel fluctuations within one group share the same fluctuation tendency (i.e. same trough and peak). The channel fluctuations between these two groups, on the other hand, exhibit opposite fluctuation tendency. However this assumption may be true for very limited number of occasions. The channel condition of the sensor node on the right foot may exhibit totally different fluctuation tendency from the sensor node on the right hand though they may share the same fluctuation frequency. Furthermore, BANMAC is unable to deal with extensive other activities with irregular body movements.

2.2   Cross-layer Design in Wireless Body Area Networks

There are many optimization literatures for Wireless Sensor Networks (WSNs) [44–49]. However WBANs have certain unique properties so that the conclusions for general WSNs may not be applied in WBANs. According to [10], the energy constraint for WBANs is even harsher than that of WSNs because the volume of on-body or in-body sensor nodes needs to be tiny and the sensors are hard to be recharged generally. Since excessive radiation absorption will cause damages to the vulnerable body tissues, the transmission power of the sensor nodes and relay nodes need to be constrained. Furthermore, the body channels are highly unstable because of the irregularity of people's movements and the signal absorption incurred by body tissues. All of these constraints in WBANs make data transmissions hardly reliable.

In order to deal with unstable channel status, relay has been proposed as one of the solutions in [10]. In [42, 64], the authors consider using relays to improve the network performance. According to simulation results, relays can reduce the outage rate and power consumption effectively. However, the decision of relay locations, which has profound influence on the performance, has not been discussed. In [73], an Energy-Aware WBAN Design Model (EAWD) is proposed to address the on-body relay positioning problem. However, only network layer is considered and no

cross-layer design is included, which undermines its effectiveness in prolonging network lifetime. Furthermore, the proposed objective is minimizing total energy consumption and may lead to energy shortage for the heavily used nodes quickly.

In [74], the authors propose solutions to the energy minimization problem and network lifetime maximization problem based on intelligent time and power resource allocation in WBAN context. Both problems are formulated and solved as geometric programming. The network in [74] is single-hop star topology, where each node communicates directly with the coordinator. This proposed design is not proper for In-Body Sensor Networks since the transmission power of the implanted sensor nodes needs to be extremely low for both energy saving and body tissue protection. The star topology is not reliable enough considering the channel fluctuation around human body. Furthermore, since relay nodes can decrease the transmission power of sensor nodes by shortening the transmission distance, the network lifetime can be greatly prolonged by deploying relays.

In [43], the authors propose a relay based routing protocol for Wireless In-Body Sensor Network. Network lifetime maximization and end-to-end-delay minimization problems are formulated and solved with linear programming. However, no systematical scheme is proposed to address the optimal relay location consideration. Furthermore, since the model proposed only considers routing layer, it fails to formulate a cross-layer optimization problem. Considerations on other layers, such as power control technique, are neglected.

So far, a WBAN optimization model that jointly considers relay positioning, transmission power control and routing strategy, which are the key considerations for an energy efficient network design, is still missing. One potential problem that is hard to solve is that the formulated optimization problem is non-convex and mixed-integer when multiple considerations are included simultaneously. Effective solutions need to be proposed before the model can be applied.

2.3   On-body Channel Dataset

On-body channel dataset [75] is a public dataset provided by National Information Communications Technology Australia (NICTA) and collected using a NICTA developed wearable channel sounder/radio device. The dataset contains WBAN channel gain data (majority on-body, but some off-body, none in-body) in various environment, carrier frequencies and measurement scenarios.

Except, where otherwise indicated, the measurement environment for the wearable radio data is "everyday" mixed-activity data collected in multiple measurement locations. There are however, certain dataset for indoor activities, such as sleeping, sitting, walking and running, and outdoor activities, such as running and driving, specifically. Across all measurements there are 22 adult
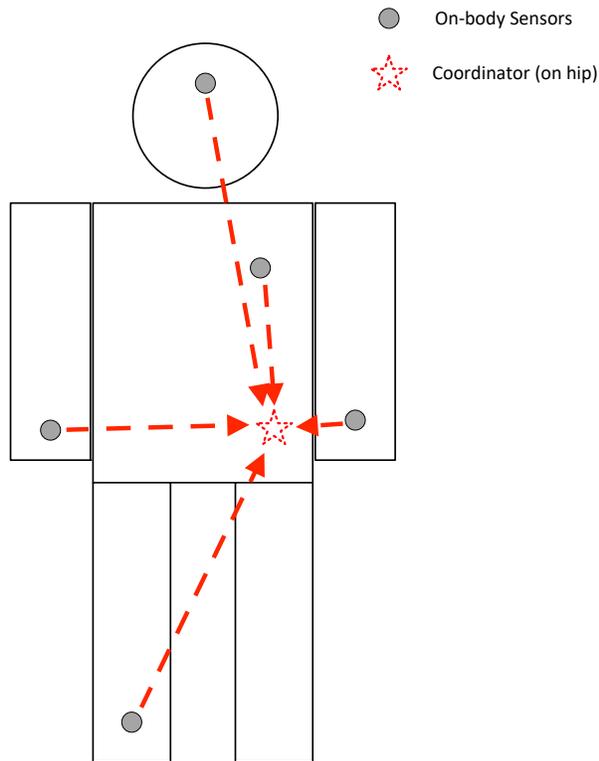
Figure 2.7: Location of the sensors and coordinator in the simulation of Chapter 3

subjects used to capture this WBAN channel data. In .mat files, each adult subject is referred as anonymized, with gender, i.e. male or female, followed by a number identifier. Thus, e.g., any filename that contains "Male1" implies the BAN link data captured for the same male person. The detailed channel information are described in the name of the channel data variable as well as the info.txt file, containing the origin and destination locations of the channel link.

On-body channel dataset [75] is used in Chapter 3 to conduct simulations of on-body WBAN. 8 datasets for 6 individuals are included in the simulation. For each adult subject, channel gain data of 5 channels, which are the channels from left hip to head, chest pocket, left wrist, right wrist and right ankle, are used to conduct the simulations. The coordinator locates at the left hip of the experiment object. The location of the sensors and coordinator is plotted in Figure 2.7. Compared with Monte Carlo simulation, the simulation based on real-life data is more compelling since no satisfactory channel model for on-body channels has been proposed so far.

In Chapter 4, since the numerical results are obtained based on the optimization model pro-

posed, the dataset is not used.

2.4 Summary

In this chapter, an extensive summary of the previous studies and researches on energy efficient WBAN designs has been presented. Both the advantages and disadvantages of each design have been analyzed and discussed. Specifically, the existing open problems need to be solved include 1) an effective opportunistic MAC layer scheduling scheme based on body channel dynamic characteristics with low overhead is still missing 2) an effective cross-layer optimization scheme that focuses on human body environment and channel conditions, while jointly considering relay positioning, transmission power control and routing strategy.

Chapter 3

Proposed Opportunistic Scheduling

3.1 Introduction

In this chapter, to address the problem that the current WBANs with fixed scheduling fail to react to the channel fluctuations, an opportunistic scheduling scheme that actively adapts to user's body channel dynamics is proposed. A two-state Markov model is adopted to analyze on-body channel conditions and beacon-based TDMA scheme is utilized. To address the problem that literatures [11, 52, 53] fail to provide a reliable method to calculate the channel dynamics, or the transition probabilities, we propose a novel estimation method. By counting the samples of the state transitions, it is possible to estimate the corresponding probability distribution. Considering the fact that on-body channels are highly dynamic and versatile, we emphasize more on the recent samples and use a discounting factor ρ to control how fast the coordinator forgets. By modeling the channel with more accurate transition probabilities, the Markov model provides much better state belief estimations than that of the literatures. Furthermore, we propose a heuristic scheduling method that is proved to be optimal under some trivial assumptions. The existing flipping method in [11] is revealed to be a special case of the proposed method. Based on the observation that the superframe length plays a significant role on the performance of opportunistic scheduling, we further propose to dynamically adjust the superframe length according to the channel status. The principle behind this design is that if the superframe length is too small, scheduling does not influence too much because the channel condition does not change too much within the range of a superframe. On the other hand, the superframe length cannot be too large, or the estimation of channel condition is unreliable. As is shown by the simulations, the design of heuristic scheduling and dynamic superframe length adjustment turns out to perform much better than the fixed scheduling

and other existing WBAN opportunistic scheduling schemes. Specifically, the proposed method can avoid around 9% of the outage occurrence compared with the fixed scheduling and achieve around 30% better performance than the flipping method [11] and more than 35% better than the random method.

The rest of this section is organized as follows. In Section 3.2, we introduce the system model and formulate the scheduling problem in WBAN context. In Section 3.3 and Section 3.4, we present the proposed heuristic scheduling scheme and dynamic superframe length adjustment approach. The simulation results are presented in Section 3.5 and Section 3.6 summarizes this chapter.

3.2 System Model and Problem Formulation

3.2.1 System Model

We focus on scenarios such as health care and fitness. Sensor nodes locate on the surface of human body and transmit the data of life signals to the coordinator periodically. A single-hop star topology is adopted. Sensor nodes are generally powered by tiny batteries and of limited capacity. The coordinator is an ordinary node, such as cell phone. Both IEEE 802.15.4 and IEEE 802.15.6, which are widely used for WBANs suggest a beacon-enabled TDMA working mode. A beacon frame is broadcasted by coordinator for synchronization and resource allocation at the beginning of each superframe. Each node will wake up and transmit in assigned time slots in each superframe periodically.

The Gilbert model [51] is adopted, where there are two channel states defined for each time slot, namely "good" and "bad". They are represented by "1" and "0" respectively. Although there exists more complicated Finite-State Markov Channel (FSMC) model [54] and it can model Rayleigh fading channel quite well, it is not necessarily helpful for WBANs. According to 802.15.6 group [55] and the research [56], on-body channels can barely be modeled by Rayleigh fading. Moreover, the simple Gilbert model has been proved to be effective at modeling on-body channels' dynamics [11, 52, 53].

Because energy is extremely limited in WBANs, probing the channels is not suitable. The only chance to observe the channel status for the coordinator is when the node is transmitting. At the other time slots, we use state belief instead of the exact channel status to measure the channel condition, that is, the channels are partially observed. If the transition matrix at time slot k for node i is $\mathbf{P}^i(k) = \begin{pmatrix} p_{00}^i(k) & p_{01}^i(k) \\ p_{10}^i(k) & p_{11}^i(k) \end{pmatrix}$ and the initial state belief of node i to be $\begin{bmatrix} p_0^i(0) \\ p_1^i(0) \end{bmatrix}$, the state

belief followed at time slot n can be calculated as

$$\begin{bmatrix} p_0^i(n) \\ p_1^i(n) \end{bmatrix}^T = \begin{bmatrix} p_0^i(0) \\ p_1^i(0) \end{bmatrix}^T \prod_{k=1}^n \begin{pmatrix} p_{00}^i(k) & p_{01}^i(k) \\ p_{10}^i(k) & p_{11}^i(k) \end{pmatrix} \tag{3.1}$$

where $p_0^i(0)$ is the probability of having a bad channel at the beginning and $p_1^i(0)$ is that of having a good channel. Normally "slot 0" is when transmission happens and the channel condition is known. However, the channel belief far away from the initial state is useless because it converges to the steady state.

Though it is safe to model WBAN channels using Markov process, great difficulties exist. The first problem is the process is partially observable. We are only able to obtain the exact channel condition when a node is transmitting. Actually the problem is even more complicated than Partially Observed Markov Decision Process (POMDP) for which the transition model exists. Here we have no transition model and the transition probability may not even be stationary because humans change their body gestures constantly. These problems are the major reasons why opportunistic scheduling in WBANs is different from that of other systems, such as cellular network.

3.2.2 Problem Formulation

The objective is to maximize the expectation of the number of successfully transmitted packets. Assume $h_i^r(\tau)$ represents the channel quality belief, which is the probability of node i having a good channel at time slot τ in round r. Thus we can express the problem as

$$\begin{aligned} \text{maximize} \quad & \sum_{1 \le i \le M, 1 \le \tau \le T, 1 \le r < \infty} h_i^r(\tau) I_i^r(\tau) \\ \text{subject to} \quad & \sum_{1 \le i \le M} I_i^r(\tau) \le 1 \\ & \sum_{1 \le \tau \le T} I_i^r(\tau) = x_i \end{aligned} \tag{3.2}$$

where $I_i^r(\tau)$ represents whether node i is scheduled at slot τ in round r and x_i represents the number of packets node i needs to transmit in a round. M represents the number of nodes. T is the number of slots in a round. Here we assume only one packet is transmitted in a slot.

As mentioned in the previous section, how to calculate channel belief $h_i^r(\tau)$ is still a critical problem to solve. Assume we have the reliable estimation of $h_i^r(\tau)$, then the solution can be solved in each round independently. The local solutions that maximize the number of successful transmis-

sions in each round jointly compose the global optimal solution. In each round, this problem can be turned into maximum weighted bipartite matching problem and further solved as assignment problem. Hungarian algorithm gives a solution of complexity $O(T^3)$. However, this can be quite large for the computation in a round considering T can be as large as hundreds of time slots.

3.3 Heuristic Scheduling

3.3.1 Transition Matrix Estimation

The existing Markov models that focus on WBANs assume stationary channels, that is, the transition probabilities are constants, such as [11, 52, 53] and many works mentioned in [68]. However, considering the highly dynamic channel environment of WBANs, we take the fluctuation of transition matrix into consideration and estimate the probabilities by state transition frequencies. We only assume the transition matrix stays the same during the time range of scheduling at the beginning of each superframe.

The coordinator observes and calculates the frequency of each node transitioning from one state to another to approximate the transition probabilities. For node i, the coordinator will calculate $W^i_{S_1 S_2}(r)$, which is the weighted sum of the times that node i transits from state S_1 to state S_2 until the rth superframe, $S_1, S_2 \in \{0, 1\}$. Specifically, we define a discounting factor ρ and a time period length L. $W^i_{S_1 S_2}(r)$ will be discounted by a factor of ρ every L superframes. The weighted sum is used so that we give more trust to the recent samples and less to the remote ones. Then for the rth superframe, we have the estimation $p^i_{S_1 S_2}(r) = \frac{W^i_{S_1 S_2}(r)}{\sum_S W^i_{S_1 S}(r)}$, where $S \in \{0, 1\}$.

This is a more accurate model than the globally stationary assumption for the reason that a person may change his/her body movement from time to time, and thus the channel dynamics will change over time. Furthermore, the dynamically updated transition matrix makes it unnecessary to measure the channel dynamics specially by the experts, and instead only some reasonable initialization of $W^i_{S_1 S_2}(r)$ at the very beginning is required.

Note that the node should transmit their packets in consecutive slots because the transmission order is decided by the nodes, which will be illustrated by Lemma 1 and Appendix B in the following. So the coordinator is able to monitor the channel changes between the adjacent time slots and calculate $W^i_{S_1, S_2}(r)$.

3.3.2 Heuristic Scheduling

In order to ensure an acceptable complexity when scheduling, a greedy method which performs better than the literature and is of low complexity is considered.

One basic idea when scheduling is that we can divide the nodes into two groups. One group contains nodes that transmit successfully in the previous round and the other contains the ones that fail their transmissions. If a node transmits more than one packet, which group it belongs to is decided by the last transmission in that round. Apparently, the "successful" nodes should be scheduled before the "failed" nodes. In this way, the node with a good channel is able to take advantage of the channel before it turns worse, and the node with a bad channel will have enough time to recover. This assertion is true regardless whether the dynamics, or the transition matrix, of the channels among all nodes are identical or not. A proof can be found in [11].

Furthermore, since we have the estimation of transition matrix and the transmission results in the previous round, we are able to calculate the channel state belief in any time slot in the current round (any time slot in the future theoretically, however in the far future the state belief will converge to the steady state and provide basically no useful information for scheduling). According to [56], the coherence time of everyday BAN channel is around 400ms, which is larger than the superframe length in our design. So the channel belief we estimate is reliable.

According to Section II, we can calculate p_1^i to approximate h_i^r in the objective function, where the last transmission in the previous round is the initial state belief of p_1^i. It can be proved that p_1^i evolves as an exponential function as (4) (Appendix A). Here to simplify the expression and according to the fact that the channel dynamics during scheduling can be regarded as stationary, we omit the round number r and use ε_i to represent the probability of transiting from bad channel to good channel $p_{01}^i(r)$ and δ_i the probability of transiting from good channel to bad channel $p_{10}^i(r)$. The estimation of the belief of having a good channel at time t can be expressed as Equation 3.3. t is the number of time slots away from the initial state, which is the result of the previous transmission.

$$
p_1^i(t) = \begin{cases} [1 + \frac{\delta_i}{\varepsilon_i}(1 - \varepsilon_i - \delta_i)^t]\frac{\varepsilon_i}{\varepsilon_i + \delta_i}, \text{successful node} \\ [1 - (1 - \varepsilon_i - \delta_i)^t]\frac{\varepsilon_i}{\varepsilon_i + \delta_i}, \quad \text{failed node} \end{cases} \tag{3.3}
$$

Consider the fluctuation range of the channel belief to be the utility function or the heuristic function. Specifically, for successful nodes, we calculate the channel belief difference of each node between the start of a round and the end of all successful nodes' transmissions. Assume N_i is the number of time slots between the previous transmission of node i and the beginning of next round and N_{Good} is the number of time slots assigned for successful nodes' transmissions. The utility

33
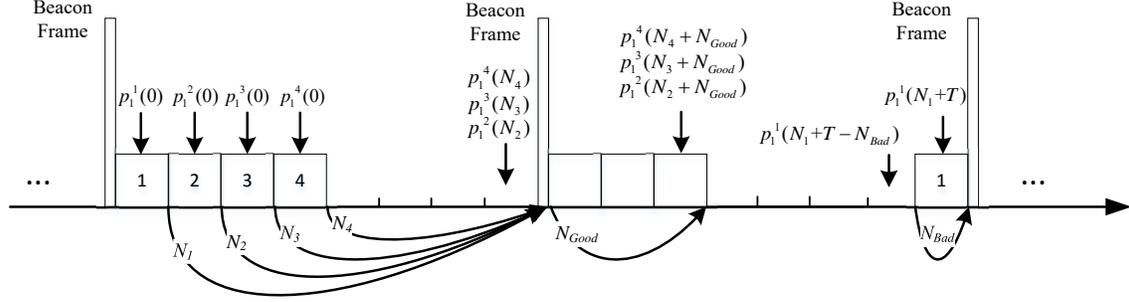
Figure 3.1: Illustration on utility function calculation

function for the successful nodes can be expressed as

$$U_i = p_1^i(N_i) - p_1^i(N_i + N_{Good}) \tag{3.4}$$

The node with a larger utility function will be scheduled at the front. This design can be illustrated as in Figure 3.1. Assume there are 4 nodes and node 1 fails the transmission while the others succeed.

Similarly, the utility function for the failed nodes can be calculated as below

$$U_i = p_1^i(N_i + T - N_{Bad}) - p_1^i(N_i + T) \tag{3.5}$$

where N_{Bad} is the number of time slots assigned to the failed nodes and T is the number of time slots in a round.

Lemma 1 *Given that node i and node j are in the same group (both fail or succeed in their previous transmissions), if $\varepsilon + \delta$ are identical among all nodes and $U_i > U_j$, the slots of node i should be scheduled in front of the slots of node j.*

See Appendix B for detailed proof of Lemma 1. This utility function makes sense in that when the channel environment of a node that transmits successfully in the previous round deteriorates fast, this node should be scheduled as early as possible so that it can transmit before the channel becomes bad. Similarly, when the channel environment improves slowly for a node that fails its previous transmission, this node should be scheduled as early as possible in its group. The pseudo code for this approach is illustrated in Algorithm 1.

Assume each node transmits N_p packets in each round on average, the complexity of the scheduling scheme proposed is $O(M \times N_p + M log M)$. This computational complexity is trivial for a powerful coordinator (such as a cell phone) in WBANs, especially that normally we have a

---

Algorithm 1 Heuristic scheduling

Initialize: $r = 1$, $W^i_{S_1,S_2}(r) = \theta$ (θ is a small positive value). Schedule randomly for the first round.

while $r \geq 1$ **do**

    Transmit according to the schedule;

    for $i = 1, 2, ..., M$ **do**

        if S_1 is observed at one slot and S_2 at the next **then**

            $W^i_{S_1,S_2}(r) = W^i_{S_1,S_2}(r) + 1$;

        end if

        if $r \equiv 0 \pmod{L}$ **then**

            $W^i_{S_1,S_2}(r) = \rho W^i_{S_1,S_2}(r)$;

        end if

    end for

    for $i = 1, 2, ..., M$ **do**

        $p^i_{S_1,S_2}(r) = \frac{W^i_{S_1,S_2}(r)}{\sum_S W^i_{S_1,S}(r)}$;

    end for

    Schedule the successful nodes at the front and the failed nodes at the end;

    Compute U_i and sort each group by decreasing order;

    $r = r + 1$;

end while

---

limited number of sensor nodes and low working load. The requirement that all nodes share the same $\varepsilon + \delta$ value is mostly true in practice and our method achieves better performance than other existing methods. The flipping method in [11] is actually a special case of our proposed method, where both ε and δ are required to be identical for all nodes (see Appendix B).

3.4 Dynamic Superframe Length Adjustment

Based on our observation, the influence of scheduling within a certain round is quite limited with an improper superframe length. The reason is that if the superframe length is small, the channel does not change too much within this round, which makes scheduling unimportant. On the other hand, if the superframe is too long, the channel state belief converges to the steady state and barely provides any information for scheduling. Furthermore, the suitable superframe length changes if the objective modifies his/her body gesture.

So here we propose a dynamic superframe length adjustment method based on Temporal-Difference (TD) error, which will improve the performance further.

3.4.1 Problem Formulation

The problem of deciding the superframe length can be formulated as a decision problem, where the agent is the coordinator and it makes decisions based on the states and rewards. The states, actions and rewards can be defined as

1. State: Packet Delivery Rate (PDR) in the previous round

2. Action: Choice of the superframe length T and the packet number to deliver in a round x_i

3. Reward: rw = - (1 - PDR) + PDR in the current round

Note here, in order to maintain a constant data rate, when adjusting the superframe length, which will always be the multiple of a minimum length T_0, the number of the packets to deliver x_i will be adjusted proportionally to T.

We use action-value function $Q(s,a)$, which represents the expected value of doing action a in state s and then following the optimal policy, to make decisions. In each state, we will pick the action with the largest Q value.

If we use dynamic programming to solve this problem directly, we will have the Bellman equation as below

$$Q(s,a) = \sum_{s'} p(s'|s,a)[rw(s,a,s') + \gamma \max_{a'} Q(s',a')] \tag{3.6}$$

However, the transition probability is not easy to obtain. We are unable to do iterations and get the optimal values or policy.

Here we choose to use TD learning, specifically Q-learning, which is a model free method to calculate Q values. The updating rule can be expressed as [76]

$$Q(s,a) \leftarrow Q(s,a) + \alpha[rw + \gamma \max_{a'} Q(s',a') - Q(s,a)] \tag{3.7}$$

The pseudo code for this design is in Algorithm 2. The dominant computational complexity is still the portion of the proposed heuristic scheduling as has been analyzed above.

The bonus of applying the variable superframe length is that when the superframe length becomes larger, the energy consumed by receiving beacon frames becomes less for that the nodes receive less beacon frames. Though this may incur synchronization problem because of the drift, it can be addressed by Adaptive Guard Band Algorithm (AGBA) and Drift Adjustment Factor (DAF) method proposed in [27].

Algorithm 2 Dynamic superframe length adjustment

Initialize $Q(s,a)$ arbitrarily, default superframe length value T_0, default packet number x_{i0} in one superframe, $r = 1$;

while $r \geq 1$ **do**

    Step 1 : Choose action a according to $Q(s,a)$ values being ε-greedy. $T = a \times T_0$, $x_i = a \times x_{i0}$;

    Step 2 : Perform heuristic scheduling and transmit with superframe length T and packet number x_i;

    Step 3 : Observe transmission results and calculate the reward rw and the new state s';

    Step 4 : Update the $Q(s,a)$ belief using $Q(s,a) \leftarrow Q(s,a) + \alpha \left[rw + \gamma \max_{a'} Q(s',a') - Q(s,a) \right]$;

    Step 5 : $s \leftarrow s'$, $r = r + 1$;

end while

3.5 Performance Evaluations

3.5.1 Simulation Methodology and Settings

In the simulations, there are five sensor nodes and one coordinator, which are on the surface of human body. The coordinator locates at the left hip pocket and the sensors locate at head, chest pocket, left wrist, right wrist and right ankle. Each time slot is 5ms long. Assume the number of packets transmitted by each node in each round is identical. The channel gain samples of on-body channels from data set [75] are used to conduct the simulations. Specifically, 8 datasets for 6 individuals are included, which are '20090421_Male1.mat', '20090416_Male2.mat', '20090429_Male2.mat', '20090417_Male4.mat', '20090422_Male4.mat', '20090421_Male5.mat', '20090428_Female1.mat', '20090430_Female2.mat'. Both man and woman adult subjects are covered to ensure a good generalization of the result. For each adult subject, channel gain data of 5 channels, which are the channels from left hip to head, chest pocket, left wrist, right wrist and right ankle, are used to conduct the simulations. The coordinator locates at the left hip of the experiment object. Note that instead of assuming a certain distribution and randomly generating the channel gain data, we use real life data in the simulations. Since there is no satisfactory model for on-body channels so far, the real-life data should be an accurate reflection of the real channel status. When the channel gain is lower than the outage threshold, outage is regarded to happen and the transmission fails. The final results are averaged over these 8 datasets.

    Two other opportunistic scheduling designs are used as comparisons to evaluate the perfor-

mance of our proposed methods. One is the random scheduling, which divides the nodes into "successful" and "failed" groups as introduced above and schedules the transmissions randomly within each group. The other one is flipping strategy proposed in [11]. The performance of fixed scheduling is used as the baseline and the performance of the other scheduling schemes are measured by their improvements over the fixed scheduling. Specifically, the percentage of outage avoided compared with fixed scheduling is used. The simulations are based on Matlab platform.

3.5.2 Heuristic Scheduling with Fixed Superframe Length

In this section, we present the performance of only using heuristic scheduling without dynamic superframe length strategy. In Figure 3.2, we show the influence of each node's packet delivery number in each round. In order to consider a wide range of possible packet number values, the superframe length is set to 80 slots, which is 400ms long and around the maximum coherence time for on-body channel. The outage threshold for the channel gain samples is -80 dB. It can be seen that using the simple trick of putting the successful nodes at the front and the failed nodes at the end in a round (random scheduling method) can already achieve a good improvement on the packet delivery rate compared with the fixed scheduling. It can be concluded that our proposed heuristic scheduling design (without dynamic superframe length) results in a better performance than the flipping method [11] consistently.

In Figure 3.3, we show the influence of superframe length on the performance. Each node will transmit 2 packets when the superframe length is 20 slots and the number of the packets transmitted grows proportionally with the superframe length to maintain a constant data rate. The outage threshold is still -80 dB. According to Figure 3, the proposed heuristic scheduling has a clear improvement over the performance of other methods. The curve is consistent with our analysis above: neither too large nor too small a superframe length is suitable. A medium superframe length is a good balance between the channel state estimation accuracy and the effect of scheduling.

In Figure 3.4, we show the influence of the outage threshold. The superframe length is 40 slots and each node transmits 4 packets per round. A medium outage threshold results in the best performance and the reason has been introduced in [11].

Though the improvement brought by pure scheduling is not significant, the scheme proposed is quite beneficial. The reason is that there is no extra overhead added to the sensors. All calculation is done by the coordinator, which is not power constrained and with strong computational ability.

38
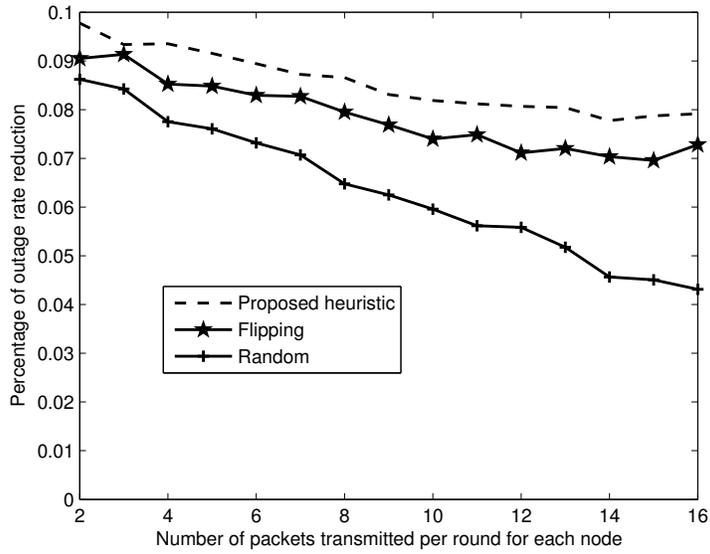
Figure 3.2: Influence of packet number in fixed superframe length scheduling
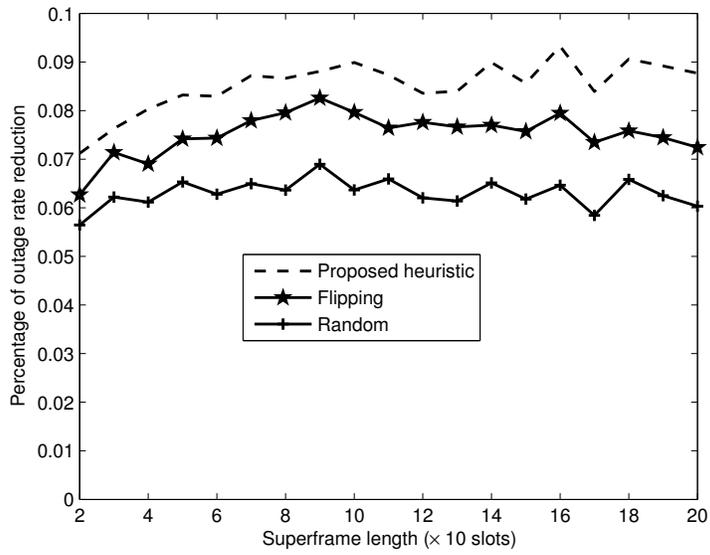
Figure 3.3: Influence of superframe length
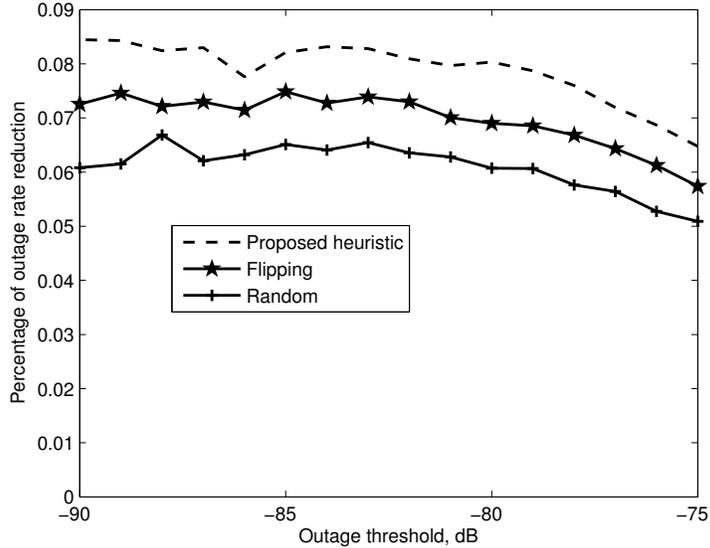
Figure 3.4: Influence of outage threshold in fixed superframe length scheduling

3.5.3 Heuristic Scheduling with Dynamic Superframe Length

In this section, we show the performance of the dynamic superframe length adjustment method combined with the heuristic scheduling scheme proposed. Since too large superframe length leads to unreliable channel belief estimation thus deteriorates performance, to avoid the superframe becomes unreasonably large, we set the default value T_0 of the proposed method, which equals to the fixed superframe length of the other methods, to be 40 slots. There are five superframe length values defined for dynamic superframe length scheduling. Specifically, it can be 40, 80, 120, 160 or 200 slots.

In Figure 3.5, the influence of packet number is shown. The packet number per round of the fixed superframe length methods ranges from 2 to 7 and so is the default packet number x_{i0} of the proposed dynamic superframe length method. The outage threshold is -80 dB. A combination of the proposed heuristic scheduling and the variable superframe length scheme yields a clear improvement over the other scheduling approach.

In Figure 3.6, the influence of outage threshold is presented, which is consistent with the rest of the results.
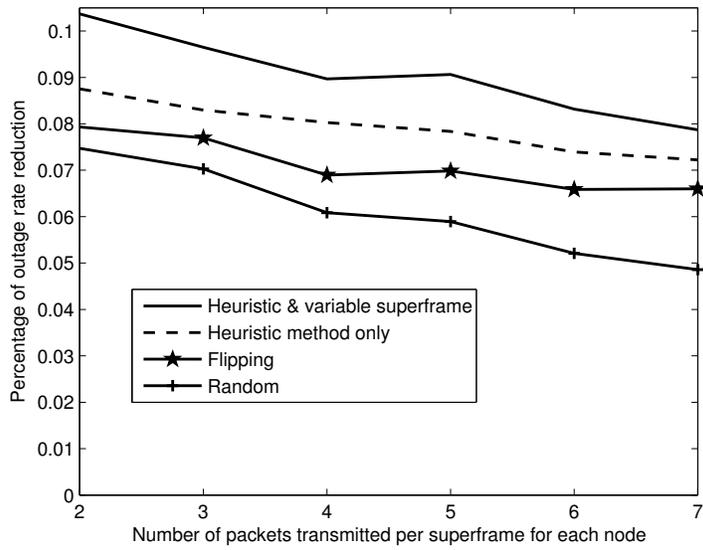
Figure 3.5: Influence of packet number
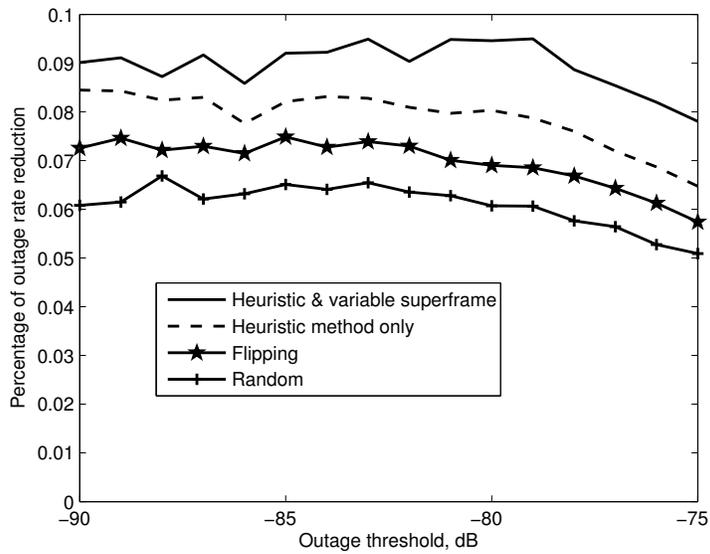
Figure 3.6: Influence of outage threshold

41

3.6 Summary

There are two major contributions in this chapter. Firstly, a method to estimate the channel dynamics has been proposed and further, based on this estimation, a simple scheduling scheme with good performance using the proposed heuristic function has been designed. Secondly, the fundamental effect of a proper superframe length in opportunistic scheduling has been revealed. The proposed combined scheduling yields a much better performance on the outage rate compared with the literature.

Chapter 4

Cross-layer Optimization

4.1   Introduction

In this chapter, to address the problem that the key considerations to prolong the network lifetime are considered separately in the existing models, a cross-layer design that jointly considers relay positioning, transmission power control and routing strategy is proposed. A single model enables us to search the optimal solutions over a larger feasible set and achieve longer lifetime than the literatures. However the formulated optimization problem turns out to be hard to solve. It is both non-convex and mixed-integer. A series of multilevel primal and dual decomposition approaches are developed to effectively find the optimal solutions to the formulated problem.

Note that the design in this chapter does not include MAC layer scheme and is compatible with the MAC layer scheduling scheme discussed previously.

The rest of this chapter is organized as follows. In Section 4.2, the problem scenario and problem model are formulated. In Section 4.3 and Section 4.4, effective solutions are designed to solve the proposed optimization problem. In Section 4.5, extensive numerical analysis is conducted to illustrate the performance of the presented algorithms and Section 4.6 concludes this chapter.

4.2   System Model and Problem Formulation

A Wireless In-Body Sensor Network with in-body sensors is considered in this chapter. The locations of the sensors are decided by doctors according to the patient's health status. The sensor set is represented by S. The relays locate on the surface of human body. For example, they can be worn on the clothes. The potential locations of on-body relay nodes are called the Candidate Sites

43

Table 4.1: Notation Definition

Notation	Meaning
S	Sensor node set
R	Set of relay candidate sites
c	Coordinator
L_i	Lifetime for node i
T_i	Transmission time for node i per superframe
T_{frame}	Duration of a superframe
B_i	Initial battery level for sensor node i
r_{ij}	Channel capacity for link (i, j)
P_i	Transmission power for sensor node i
z_j	Decision variable for relay candidate site j
x_i	Data generation rate for sensor node i
P_{max}	Maximum transmission power for each node
P_{min}	Minimum transmission power for each node

(CSs) as in [73, 77]. The relay CS set is represented by R. A decision variable z_j, $j \in R$ is used to denote whether a relay CS is chosen. $z_j = 1$ represents that the jth location of the relay CSs is used. Notations used are explained in Table 4.1. Figure 4.1a gives a quick view of the WBAN that is considered. Grey dots represent the implanted sensor nodes. The blue crosses represent the relay CSs. The red star represents the coordinator. Each node is labelled with an ID, which is used in the numerical analysis.

We divide the human body into six regions: left arm, right arm, left leg, right leg, head and torso. The node set in each region, which includes both implanted sensors and wearable relays, are denoted as G_1, G_2, G_3, G_4, G_5 and G_6. Define total node set $G = \{G_1, G_2, G_3, G_4, G_5, G_6\}$.

Specifically, we consider the following scenario:

- The sensor nodes only utilize the relay CSs that are in the same body region, which will result in a stable network topology and communication performance. Though the relative distance among the nodes in different regions fluctuates significantly during the process of body movements, the relative distance among the nodes in the same region will be much less affected. So the requirement that each sensor only communicates with the relays in the same body region indicates that the data transmission of the sensors will remain stable. The relays are equipped with larger battery volume and are with higher transmission power, which are less sensitive to the distance fluctuation. Based on this consideration, a static network, where the distance among the nodes is fixed, is assumed as in [43].

44

(a) WBAN nodes   **(b)** WBAN topology
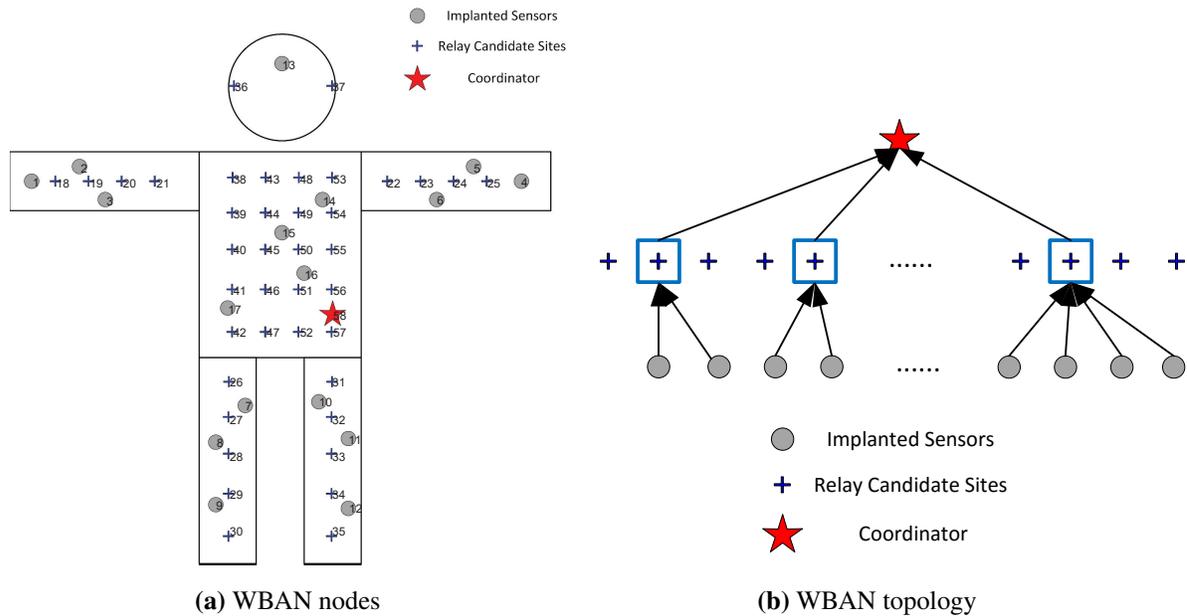
Figure 4.1: WBAN nodes and topology

- The positions of the relay candidate sites are chosen depending on the locations of the implanted sensors, which are decided by the patient's actual health condition in the real application. Only the body regions with sensors implanted should contain relay candidate sites. In this article, it is assumed that the patient has randomly placed implanted sensors in each body region. In order to figure out the optimal places for the relays in the body regions, each body region is divided by grid with proper grid line spacing. The relay candidate sites are placed at the intersections of the grid. A grid with too large spacing may lead to inaccurate positions, and too small spacing will cause unnecessarily high computational complexity and little accuracy improvement. The grid line spacing used in this article turns out to be a good balance between complexity and accuracy, which results in reasonable computational time consumption and good network performance.

- For each body region, only one relay is used. As the number of relays increases, the performance will improve. However, according to [78], the absorption of radiation will increase the temperature of body tissues. This can cause damage to sensitive organs by reducing the blood flow and growing certain type of bacteria. Because of the health consideration, the number of relays applied need to be low. The optimal number of relays to place is out of the

scope of this article and it is assumed that only one relay is used in each region.

- Relays are placed on-body and they can be easily replaced or recharged compared with the implanted sensors. Thus as in [43], we do not include the considerations on relay's energy consumption and assume the relay nodes always transmit with the highest transmission power permitted.

- Multi-hopping among in-body sensors is not allowed and the in-body sensors transmit directly to the relays. The relay nodes transmit their data directly to the coordinator as well since they have a high transmission power. Thus the whole network is of a two-layer tree topology as is shown in Figure 4.1b. The blue boxes represent the selected relay CS locations.

- As suggested by standard IEEE 802.15.6 protocol [79] that is widely used by WBAN designs, TDMA scheme is adopted. Among all the nodes in WBAN, each node transmits in its protected time slot and no other nodes transmit simultaneously. Thus there is no interference among the transmissions.

To achieve low-delay and collision-free packet transmissions, TDMA scheme is used. Both IEEE 802.15.4 and IEEE 802.15.6, which are widely adopted by WBANs, suggest Beacon-enabled TDMA scheme. A typical TDMA superframe structure has been shown in Figure 1.4 previously.

Since there is no interference, the channel capacity r_{ij} can be expressed as a function of P_i

$$r_{ij} = W \log \left(1 + \frac{\alpha_{ij} P_i}{N_j} \right) \tag{4.1}$$

where the base of the logarithmic function is 2, W is the bandwidth, N_j is the power of noise and α_{ij} is the channel gain between node i and j.

The network lifetime is defined as the duration until the first node of entire body runs out of energy, which has been adopted in [80]. This definition is a reasonable choice for WBAN. If some node dies, the whole WBAN can no longer function well and guarantee the patient's safety because some critical health information may be lost.

For each node, its lifetime can be calculated as

$$L_i = \frac{B_i}{P_i T_i / T_{frame}} = \frac{B_i T_{frame}}{P_i T_i} \tag{4.2}$$

The network lifetime is

$$\min_{i \in S} L_i = \min_{i \in S} \frac{B_i T_{frame}}{P_i T_i} \tag{4.3}$$

The locations of the relay candidates sites, the sensor nodes and the coordinator are described by two dimensional coordinates in a reference system that contains the human body shape. The distance among the nodes and hence the path loss are calculated based on these coordinates.

We focus on maximizing the network lifetime by jointly considering the transmission power (P_i), transmission time (T_i) and relay location decisions (z_j). An offline approach is proposed to derive the optimal network parameters. The proposed model is defined as

$$\underset{P,T,z}{\text{maximize}} \quad \min_{i \in S} \frac{B_i T_{frame}}{P_i T_i} \tag{4.4a}$$

$$\text{subject to} \quad \sum_{i \in S} T_i + \sum_{j \in R} T_j z_j \leq T_{frame} \tag{4.4b}$$

$$T_i \sum_{j \in G_k \cap R} r_{ij} z_j \geq x_i T_{frame}, \ i \in G_k \cap S, \ G_k \in G \tag{4.4c}$$

$$\sum_{j \in G_k \cap R} r_{jc} T_j z_j \geq \sum_{i \in G_k \cap S} \sum_{j \in G_k \cap R} T_i r_{ij} z_j, \ G_k \in G \tag{4.4d}$$

$$\sum_{j \in G_k \cap R} z_j = 1, \ G_k \in G \tag{4.4e}$$

$$P_{\min} \leq P_i \leq P_{\max}, \ i \in S \tag{4.4f}$$

$$z_j \in \{0,1\}, \ j \in R \tag{4.4g}$$

where objective function (4.4a) represents that the goal of the proposed algorithm is to maximize the network lifetime. Since TDMA-based scheme and superframe is adopted, all nodes will transmit their data periodically round by round. For inequality (4.4b), it states that the sum of the transmission time assigned to each node must be less than the time range of a superframe, which is the total time for a round of transmissions. Here we consider the situation that no time slot reuse is applied, which means no two nodes are allowed to transmit simultaneously or collision will happen. For inequality (4.4c), the flow balance of each sensor node i is considered. The data transmission speed of a sensor node need to be larger than the speed of its data generation. x_i is the average data generating speed for sensor node i. r_{ij} is the channel capacity and represents the maximum data rate possible between the sensor i and relay j. Note here only one relay j can be chosen in each body region G_k, which is constrained by (4.4e). For inequality (4.4d), similar to (4.4c), it describes the flow balance requirement for the relay nodes. If a relay candidate site j is

47

used for body region G_k, then its output data rate to the coordinator need to be larger than the input data stream, which is the sum of all sensors' data stream in that region. If the relay candidate site j is not chosen, z_j would be 0, this inequality does not constrain anything. (4.4e) forces that within each body region, only one relay CS is chosen and actually used to install the relay. (4.4f) sets the lower and upper bounds on the transmission power of the sensor nodes. (4.4g) requires that the value of the decision variable z_j is limited to $\{0, 1\}$.

Note that (4.4c)-(4.4e) and (4.4g) are the constraints that state which relay to use for the data transmission of each sensor and it is a routing layer consideration. (4.4b) and (4.4f) are the constraints that state the transmission power and how much time is assigned to each sensor node in each superframe, which is a physical layer consideration. The physical layer and routing layer considerations are correlated since z appears in (4.4b)-(4.4e), and data rate r_{ij} in (4.4c)-(4.4d) is a function of power P_i in (4.4f). So the proposed problem is a cross layer problem that needs to be solved on physical layer and routing layer simultaneously.

4.3 Multilevel Primal and Dual Decomposition Algorithm

The proposed problem (4.4) is not a convex problem. One difficulty results from the multiplication with the integer variable z_j, which makes (4.4b)-(4.4d) not convex. So the optimal solution of (4.4) cannot be solved directly. Therefore, we propose to use multilevel decomposition techniques and transform the original problem into several mixed integer sub-problems. Each sub-problem is confined within a body region and the optimal relay location can be derived by exhaustive search. Specifically, for a sub-problem in a certain body region, each relay CS will be tested and the one with the best objective function value will be selected. With decision variable z obtained first, the proposed problem can be solved.

4.3.1 Decomposition Algorithm

By introducing a new variable $t = \min_{i \in S} \frac{B_i T_{frame}}{P_i T_i}$, (4.4a) can be replaced by t. A new constraint should be added as

$$t \leq \frac{B_i T_{frame}}{P_i T_i}, \ i \in S \qquad (4.5)$$

By using (4.4c), inequality (4.4d) can be transformed as

$$\sum_{j \in G_k \cap R} r_{jc} T_j z_j \geq \sum_{i \in G_k \cap S} T_i \sum_{j \in G_k \cap R} r_{ij} z_j \geq \sum_{i \in G_k \cap S} x_i T_{frame} \qquad (4.6)$$

Apply log transformation to the variables with base e as $\tilde{t} = \ln t$, $\tilde{P}_i = \ln P_i$, $\tilde{T}_i = \ln T_i$ and the original problem can be transformed into

$$\underset{\tilde{P},\tilde{T},\tilde{t},z}{\text{maximize}} \quad \tilde{t} \tag{4.7a}$$

$$\text{subject to} \quad \tilde{t} + \tilde{P}_i + \tilde{T}_i \leq \ln\left(T_{frame} B_i\right), \ i \in S \tag{4.7b}$$

$$\sum_{i \in S} e^{\tilde{T}_i} + \sum_{j \in R} e^{\tilde{T}_j} z_j \leq T_{frame} \tag{4.7c}$$

$$\tilde{T}_i + \ln\left(\sum_{j \in G_k \cap R} W \log\left(1 + \frac{\alpha_{ij} e^{\tilde{P}_i}}{N_j}\right) z_j\right) \geq \ln\left(x_i T_{frame}\right), \ i \in G_k \cap S, \ G_k \in G \tag{4.7d}$$

$$\ln\left(\sum_{j \in G_k \cap R} r_{jc} e^{\tilde{T}_j} z_j\right) \geq \ln\left(\sum_{i \in G_k \cap S} x_i T_{frame}\right), G_k \in G \tag{4.7e}$$

$$\sum_{j \in G_k \cap R} z_j = 1, \ G_k \in G \tag{4.7f}$$

$$\tilde{P}_{\min} \leq \tilde{P}_i \leq \tilde{P}_{\max}, \ i \in S \tag{4.7g}$$

$$z_j \in \{0,1\}, \ j \in R \tag{4.7h}$$

Because of the coupling variable $\tilde{t}$ in (4.7a)-(4.7b) and coupling constraint (4.7c), the problem is still central and cannot be solved locally in each body region. To mitigate the coupling effect of (4.7c), the Lagrangian of problem (4.7) is considered as

$$L\left(\tilde{t}, \tilde{T}, \tilde{P}, z, \lambda\right) = \tilde{t} - \lambda\left(\sum_{i \in S} e^{\tilde{T}_i} + \sum_{j \in R} e^{\tilde{T}_j} z_j - T_{frame}\right) \tag{4.8}$$

The Lagrange dual function $g\left(\lambda\right)$ can be obtained from

$$\underset{\tilde{P},\tilde{T},\tilde{t},z}{\text{maximize}} \qquad\qquad L\left(\tilde{t}, \tilde{T}, \tilde{P}, z, \lambda\right) \tag{4.9a}$$

$$\text{subject to} \qquad\qquad (7b), (7d) \text{ - } (7h) $$

If we call problem (4.7) as the master problem, the dual problem of the master problem is

$$\underset{\lambda}{\text{minimize}} \qquad\qquad g\left(\lambda\right) \tag{4.10a}$$

$$\text{subject to} \qquad\qquad \lambda \geq 0 \tag{4.10b}$$

In order to obtain $g(\lambda)$ and further solve (4.10), we apply primal decomposition to problem (4.9). Specifically we fix the coupling variable $\tilde{t}$ and divide problem (4.9) into two levels. At the lower level, we have

$$\underset{\tilde{P},\tilde{T},z}{\text{maximize}} \qquad \tilde{t} - \lambda \left(\sum_{i \in S} e^{\tilde{T}_i} + \sum_{j \in R} e^{\tilde{T}_j} z_j - T_{frame} \right) \qquad (4.11a)$$

$$\text{subject to} \qquad (4.7b), (4.7d) \text{ - } (4.7h) \qquad$$

whose difference from problem (4.9) is that $\tilde{t}$ is a constant for problem (4.11).

At the higher level we have the secondary master problem as

$$\underset{\tilde{t}}{\text{maximize}} \qquad f^*(\tilde{t},\lambda) \qquad (4.12)$$

$f^*(\tilde{t},\lambda)$ is the optimal value of problem (4.11) given $\tilde{t}$ and λ. The result of (4.12) gives $g(\lambda)$.

Problem (4.11) can be decomposed into sub-problems that can be solve independently within each body region. For body region with node set G_k, the sub-problem is

$$\underset{\tilde{P},\tilde{T},z}{\text{maximize}} \qquad -\lambda \left(\sum_{i \in S \cap G_k} e^{\tilde{T}_i} + \sum_{j \in R \cap G_k} e^{\tilde{T}_j} z_j \right) \qquad (4.13a)$$

$$\text{subject to} \qquad \tilde{t} + \tilde{P}_i + \tilde{T}_i \leq \ln\left(T_{frame} B_i\right), \; i \in S \cap G_k \qquad (4.13b)$$

$$\tilde{T}_i + \ln \left(\sum_{j \in G_k \cap R} W \log\left(1 + \frac{\alpha_{ij} e^{\tilde{P}_i}}{N_j} \right) z_j \right) \geq \ln\left(x_i T_{frame}\right), \; i \in G_k \cap S \qquad (4.13c)$$

$$\ln \left(\sum_{j \in G_k \cap R} r_{jc} e^{\tilde{T}_j} z_j \right) \geq \ln \left(\sum_{i \in G_k \cap S} x_i T_{frame} \right) \qquad (4.13d)$$

$$\sum_{j \in G_k \cap R} z_j = 1 \qquad (4.13e)$$

$$\tilde{P}_{\min} \leq \tilde{P}_i \leq \tilde{P}_{\max}, \; i \in S \cap G_k \qquad (4.13f)$$

$$z_j \in \{0,1\}, \; j \in R \cap G_k \qquad (4.13g)$$

If we call the maximum value of sub-problem (4.13) as $f_k^*(\tilde{t},\lambda)$, then $f^*(\tilde{t},\lambda) = \tilde{t} + \lambda T_{frame} + \sum_{k=1}^{6} f_k^*(\tilde{t},\lambda)$. Since only one relay can be selected in a certain body region, which results from the constraints (4.13e) and (4.13g), the above mixed integer sub-problem (4.13) can be solved by enumerating all possible values of z_j in this region. That is, we try each relay CS in that region

50

and select the one that results in the optimal objective function value. With integer variable z_j obtained and removed, the sub-problem (4.13) becomes a regular convex optimization problem and can be solved efficiently. To further convert the problem so that it can be solved by CVX, (4.13c) is transformed into $\tilde{T}_i + \ln\left(\sum\limits_{j \in G_k \cap R} W \log\left(\frac{\alpha_{ij}e^{\tilde{P}_i}}{N_j}\right) z_j\right) \geq \ln(x_i T_{frame})$. In our model, the condition $\frac{\alpha_{ij}e^{\tilde{P}_i}}{N_j} \gg 1$ is ensured. The reason is that the sensors are only allowed to transmit to the relays in the same body region, which leads to a short communication distance. Even with the lowest transmission power, $\frac{\alpha_{ij}e^{\tilde{P}_i}}{N_j}$ is guaranteed to be above 40.

By solving each sub-problem, the optimal value for the binary decision variable $z^*(\tilde{t}, \lambda)$ in problem (4.11) can be obtained. Specifically, consider body region with node set G_k, if the jth relay CS in this region achieves the highest value of (4.13a), then $z_j = 1$ and $z_l = 0$ for all $l \neq j$, where $j, l \in R \cap G_k$. With $z^*(\tilde{t}, \lambda)$ obtained, problem (4.11) can be transformed into

$$\underset{\tilde{P}, \tilde{T}}{\text{maximize}} \quad \tilde{t} - \lambda \left(\sum_{i \in S} e^{\tilde{T}_i} + \sum_{j \in R} e^{\tilde{T}_j} z_j^*(\tilde{t}, \lambda) - T_{frame}\right) \tag{4.14a}$$

$$\text{subject to} \quad \tilde{t} + \tilde{P}_i + \tilde{T}_i \leq \ln(T_{frame}B_i), \ i \in S \tag{4.14b}$$

$$\tilde{T}_i + \ln\left(\sum_{j \in G_k \cap R} W \log\left(\frac{\alpha_{ij}e^{\tilde{P}_i}}{N_j}\right) z_j^*(\tilde{t}, \lambda)\right) \geq \ln(x_i T_{frame}), \ i \in G_k \cap S, \ G_k \in G \tag{4.14c}$$

$$\ln\left(\sum_{j \in G_k \cap R} r_{jc} e^{\tilde{T}_j} z_j^*(\tilde{t}, \lambda)\right) \geq \ln\left(\sum_{i \in G_k \cap S} x_i T_{frame}\right), \ G_k \in G \tag{4.14d}$$

$$\tilde{P}_{\min} \leq \tilde{P}_i \leq \tilde{P}_{\max}, \ i \in S \tag{4.14e}$$

Note that (4.14) is a convex problem.

In order to derive the gradient of $f^*(\tilde{t}, \lambda)$ over $\tilde{t}$ and conduct gradient ascent method to obtain the solution of (4.12), the Lagrange dual function $d(\gamma, \tilde{t}, \lambda)$ of (4.14) is considered as

$$\underset{\tilde{P}, \tilde{T}}{\text{maximize}} \quad \tilde{t} - \lambda \left(\sum_{i \in S} e^{\tilde{T}_i} + \sum_{j \in R} e^{\tilde{T}_j} z_j^*(\tilde{t}, \lambda) - T_{frame}\right) - \sum_{i \in S} \gamma_i \left(\tilde{t} + \tilde{P}_i + \tilde{T}_i - \ln(T_{frame}B_i)\right) \tag{4.15a}$$

$$\text{subject to} \quad (4.14c) \text{ - } (4.14e)$$

The dual problem of (4.14) can be expressed as

$$\underset{\gamma}{\text{minimize}} \qquad\qquad d\left(\gamma, \tilde{t}, \lambda\right) \tag{4.16a}$$

$$\text{subject to} \qquad\qquad \gamma \geq 0 \tag{4.16b}$$

Because the optimal primal variables $\tilde{P}^*\left(\tilde{t}, \lambda\right)$ and $\tilde{T}^*\left(\tilde{t}, \lambda\right)$ of (4.14) are already obtained and problem (4.14) is convex, the optimal solution $\gamma^*\left(\tilde{t}, \lambda\right)$ of the dual problem (4.16) can be derived according to the KKT conditions as is illustrated in the Appendix C.

Since problem (4.14) is a convex problem, its duality gap with the dual problem (4.16) is zero and we have

$$
\begin{aligned}
f^*\left(\tilde{t}, \lambda\right) &= \inf_{\gamma \geq 0} d\left(\gamma, \tilde{t}, \lambda\right) \\
&= \tilde{t} - \lambda \left(\sum_{i \in S} e^{\tilde{T}_i^*\left(\tilde{t}, \lambda\right)} + \sum_{j \in R} e^{\tilde{T}_j^*\left(\tilde{t}, \lambda\right)} z_j^*\left(\tilde{t}, \lambda\right) - T_{frame} \right) \\
&\quad - \sum_{i \in S} \gamma_i^*\left(\tilde{t}, \lambda\right) \left(\tilde{t} + \tilde{P}_i^*\left(\tilde{t}, \lambda\right) + \tilde{T}_i^*\left(\tilde{t}, \lambda\right) - \ln\left(T_{frame} B_i\right) \right)
\end{aligned} \tag{4.17}
$$

where $\tilde{T}^*\left(\tilde{t}, \lambda\right)$ and $\tilde{P}^*\left(\tilde{t}, \lambda\right)$ are the optimal solutions of (4.14), and $\gamma^*\left(\tilde{t}, \lambda\right)$ is the optimal solution of (4.16).

Thus the subgradient of $f^*\left(\tilde{t}, \lambda\right)$ over $\tilde{t}$ can be calculated as

$$\frac{\partial f^*\left(\tilde{t}, \lambda\right)}{\partial \tilde{t}} = 1 - \sum_{i \in S} \gamma_i^*\left(\tilde{t}, \lambda\right) \tag{4.18}$$

From (4.15), for each certain γ, $d\left(\gamma, \tilde{t}, \lambda\right)$ is an affine function of $\tilde{t}$. Since pointwise minimum preserves concavity, $f^*\left(\tilde{t}, \lambda\right) = \inf_{\gamma \geq 0} d\left(\gamma, \tilde{t}, \lambda\right)$ is a concave function of $\tilde{t}$ given λ. Thus (4.12) can be solved by subgradient ascent method using the updating rule for $\tilde{t}$

$$\tilde{t} = \tilde{t} + \theta \left(1 - \sum_{i \in S} \gamma_i^*\left(\tilde{t}, \lambda\right) \right) \tag{4.19}$$

Once the optimal value $\tilde{t}^*\left(\lambda\right)$ is obtained, $g\left(\lambda\right)$ can be calculated by

$$g\left(\lambda\right) = \tilde{t}^*\left(\lambda\right) - \lambda \left(\sum_{i \in S} e^{\tilde{T}_i^*\left(\tilde{t}^*\left(\lambda\right), \lambda\right)} + \sum_{j \in R} e^{\tilde{T}_j^*\left(\tilde{t}^*\left(\lambda\right), \lambda\right)} z_j^*\left(\tilde{t}^*\left(\lambda\right), \lambda\right) - T_{frame} \right) \tag{4.20}$$

For the same reason that $f^*(\tilde{t}, \lambda)$ is a concave function of $\tilde{t}$, $g(\lambda)$ is a convex function of λ. The subgradient descent update for λ is

$$\lambda = \left[\lambda + \theta \left(\sum_{i \in S} e^{\tilde{T}_i^*(\tilde{t}^*(\lambda), \lambda)} + \sum_{j \in R} e^{\tilde{T}_j^*(\tilde{t}^*(\lambda), \lambda)} z_j^* (\tilde{t}^*(\lambda), \lambda) - T_{frame} \right) \right]^+ \qquad (4.21)$$

A flow chart of the proposed method is shown in Figure 4.2 and the pseudocode code is illustrated in Algorithm 3.

Algorithm 3 Joint relay location control and cross-layer optimization algorithm

---

Initialize: step size θ, threshold TH, $m = 1$, λ^0 and λ^1.
while $\left| \lambda^m - \lambda^{m-1} \right| \geq TH$ **do**
   Initialize $\tilde{t}^0$, $\tilde{t}^1$, $n = 1$;
   while $\left| \tilde{t}^n - \tilde{t}^{n-1} \right| \geq TH$ **do**
     for G_k in each sub-region **do**
       for each possible z_j, $j \in G_k$ **do**
         Solve the convex sub-problem (4.13);
       end for
       Obtain the optimal solutions given $\tilde{t}^n, \lambda^m$: $z_k^* (\tilde{t}^n, \lambda^m)$, $\tilde{P}_k^* (\tilde{t}^n, \lambda^m)$, $\tilde{T}_k^* (\tilde{t}^n, \lambda^m)$;
     end for
     Obtain the optimal Lagrange multiplier $\gamma^* (\tilde{t}^n, \lambda^m)$ of (4.14b) by KKT conditions;
     $\tilde{t}^{n+1} = \tilde{t}^n + \theta \frac{\partial f^*(\tilde{t}, \lambda^m)}{\partial \tilde{t}}$, $n = n+1$;
   end while
   $\lambda^{m+1} = \lambda^m - \theta \frac{dg(\lambda)}{d\lambda}$, $m = m+1$;
end while
Return $\tilde{P}^* (\tilde{t}^*(\lambda^*), \lambda^*)$, $\tilde{T}^* (\tilde{t}^*(\lambda^*), \lambda^*)$, $z^* (\tilde{t}^*(\lambda^*), \lambda^*)$, λ^*, $\tilde{t}^*(\lambda^*)$, $\gamma^* (\tilde{t}^*(\lambda^*), \lambda^*)$.

4.3.2 Duality Gap

In the proposed algorithm, dual decomposition is applied to the master problem (4.7), which is a non-convex problem. Thus the optimal solutions of the dual problem (4.10) formulated is potentially sub-optimal to the original problem proposed. There might be duality gap between the optimal values of the primal problem (4.7) and that of the dual problem (4.10).

To research on the influence of the duality gap, in the experiment section, we simulate a WBAN with small network size and obtain the optimal values of both primal problem (4.7) and dual problem (4.10). As will be illustrated in the numerical results in Section VI (Figure 4.6, Figure 4.7 and Figure 4.8), for the network we consider, the optimal network lifetime of the primal and dual
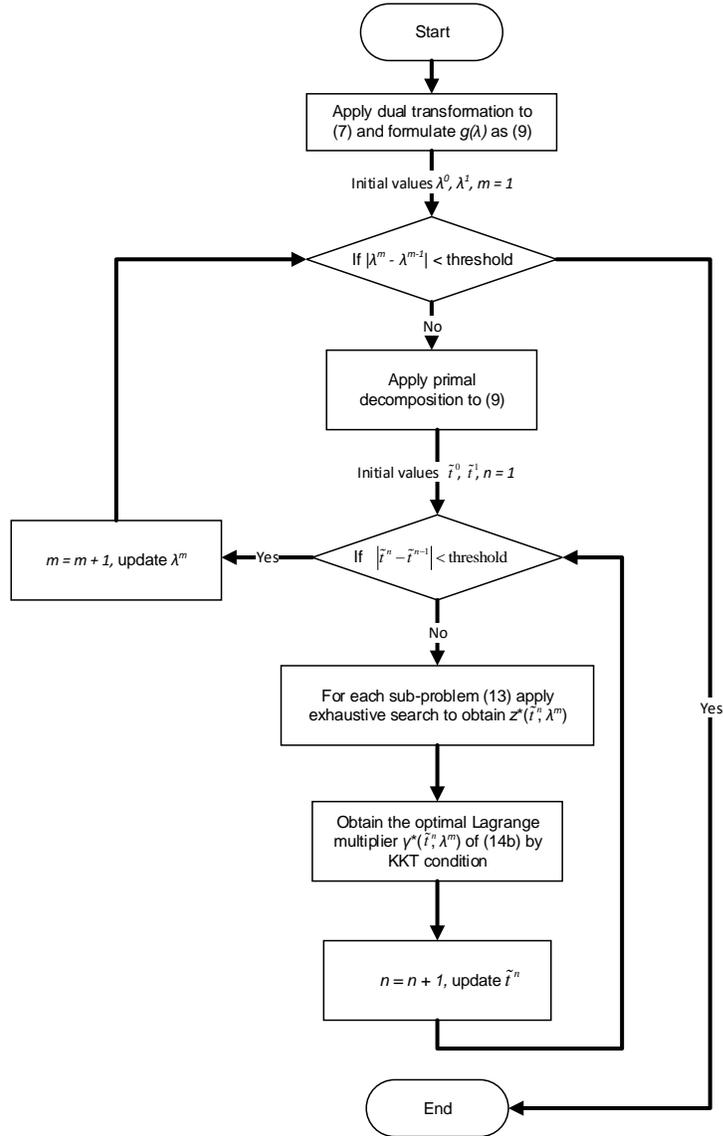
```
                    ┌─────────┐
                    │  Start  │
                    └────┬────┘
                         │
              ┌──────────▼──────────┐
              │ Apply dual          │
              │ transformation to   │
              │ (7) and formulate   │
              │ g(λ) as (9)         │
              └──────────┬──────────┘
              Initial values λ⁰, λ¹, m = 1
                         │
                    ◇ If |λᵐ - λᵐ⁻¹| < threshold ◇
                         │ No
              ┌──────────▼──────────┐
              │ Apply primal        │
              │ decomposition to (9)│
              └──────────┬──────────┘
```

Start

Apply dual transformation to (7) and formulate $g(\lambda)$ as (9)

Initial values λ^0, λ^1, $m = 1$

If $|\lambda^m - \lambda^{m-1}| <$ threshold

No

Apply primal decomposition to (9)

Initial values $\tilde{t}^0$, $\tilde{t}^1$, $n = 1$

$m = m + 1$, update λ^m — Yes — If $\left|\tilde{t}^n - \tilde{t}^{n-1}\right| <$ threshold

No

For each sub-problem (13) apply exhaustive search to obtain $z^*(\tilde{t}^n, \lambda^m)$

Obtain the optimal Lagrange multiplier $\gamma^*(\tilde{t}^n, \lambda^m)$ of (14b) by KKT condition

$n = n + 1$, update $\tilde{t}^n$

Yes

End

Figure 4.2: Flow chart of multilevel primal and dual decomposition algorithm

problems are very close, which indicates the effectiveness of the proposed method.

4.4 Fast Convergence Rate Method Based on Binary Search

Generally speaking, the convergence rate of subgradient method is slow [81]. In this section, a solution based on binary search for the optimal values is proposed and achieves much faster convergence rate compared with pure subgradient method.

Consider the updating rules (4.20) and (4.21) for $\tilde{t}$ and λ. Since it is known that $f^*(\tilde{t}, \lambda)$ is a concave function of $\tilde{t}$ and $g(\lambda)$ is a convex function of λ, the convergence rate can be increased by observing the signs of the derivatives.

For λ, the curve of $g(\lambda)$ can be illustrated as in Figure 4.3. Since λ is larger than or equal to zero, the lower bound λ_{Lower} is set to zero at the start. In order to find a proper upper bound, a step size $\lambda_{Interval}$ is defined. The initial λ_{Upper} can be obtained by

$$\lambda_{Upper} = \lambda_{Lower} + N \times \lambda_{Interval} \tag{4.22}$$

where N is the smallest integer that satisfies $\frac{dg(\lambda)}{d\lambda}\Big|_{\lambda=\lambda_{Upper}} > 0$.

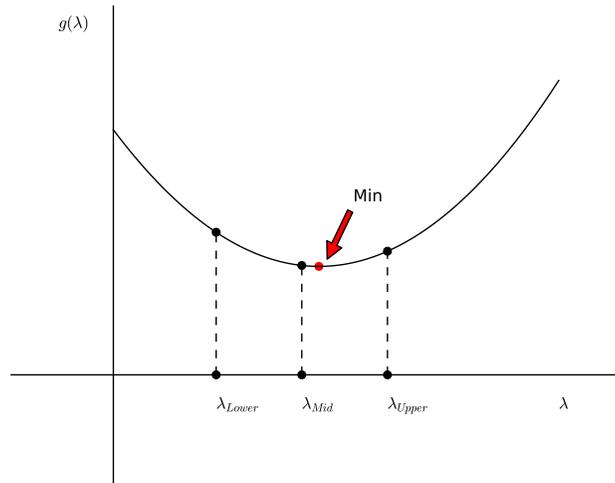**Figure 4.3:** Binary search on optimal λ

Once the initial upper bound and lower bound of λ are obtained, binary search for the optimal solution λ^* can be continued by checking the sign of the derivative at $\lambda_{Mid} = \frac{\lambda_{Upper}+\lambda_{Lower}}{2}$ and updating the value of λ_{Upper} and λ_{Lower} iteratively.

For $\tilde{t}$, there is a threshold imposed by constraint (4.14b). When $\tilde{t}$ grows too large, the problem (4.14) becomes infeasible. So there is a threshold $\tilde{t}_{TH}$ for $\tilde{t}$, beyond which feasible $\tilde{t}$ does not exist. $f^*(\tilde{t}, \lambda)$ can be plotted as in Figure 4.4.
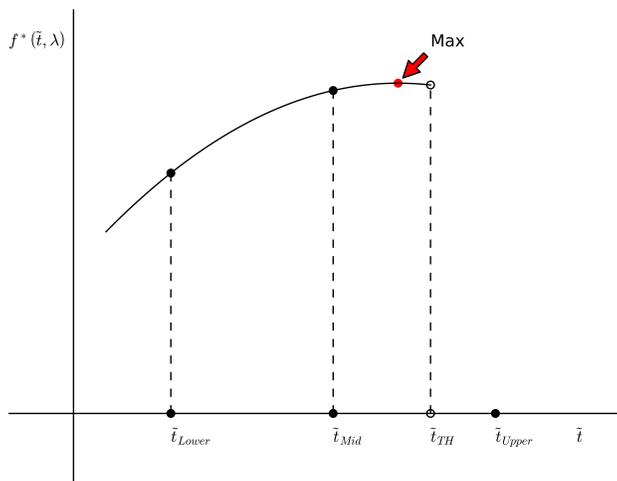

Figure 4.4: Binary search on optimal $\tilde{t}$

At the beginning, a $\tilde{t}_{Lower}$ small enough and a proper $\tilde{t}_{Interval}$ are chosen. The initial $\tilde{t}_{Upper}$ can be derived by

$$\tilde{t}_{Upper} = \tilde{t}_{Lower} + N' \times \tilde{t}_{Interval} \tag{4.23}$$

where N' is the smallest integer that satisfies $\left.\frac{df^*(\tilde{t}, \lambda)}{d\tilde{t}}\right|_{\tilde{t}=\tilde{t}_{Upper}} < 0$ or no solution exists for $\tilde{t} = \tilde{t}_{Upper}$.

The pseudocode for the proposed binary search algorithm is given in Algorithm 4.

4.5 Numerical Results

4.5.1 Evaluation Methodology and Settings

We assume there are 17 implanted sensor nodes and 40 relay CSs in the Wireless In-Body Sensor Network that we consider. The sensor nodes and relay CSs are placed as is shown in Figure 4.1a. The initial battery level of the sensor nodes are set to be 1J. The maximum transmission power is $P_{max} = 0$dBm and the minimum is $P_{min} = -15$dBm. A typical noise density value -174dBm/Hz is

---

Algorithm 4 Proposed algorithm based on binary search

Initialize: λ_{Lower}, $\lambda_{Interval}$, threshold TH.

Find λ_{Upper};

while $\lambda_{Upper} - \lambda_{Lower} \geq TH$ **do**

  $\lambda_{Mid} = \frac{\lambda_{Upper} + \lambda_{Lower}}{2}$;

  Initialize $\tilde{t}_{Lower}$, $\tilde{t}_{Interval}$ and find $\tilde{t}_{Upper}$;

  while $\tilde{t}_{Upper} - \tilde{t}_{Lower} \geq TH$ **do**

    $\tilde{t}_{Mid} = \frac{\tilde{t}_{Upper} + \tilde{t}_{Lower}}{2}$ and calculate $\left. \frac{df^*(\tilde{t},\lambda)}{d\tilde{t}} \right|_{\tilde{t}=\tilde{t}_{Mid}}$

    if $\left. \frac{df^*(\tilde{t},\lambda)}{d\tilde{t}} \right|_{\tilde{t}=\tilde{t}_{Mid}}$ does not exist **then**

      $\tilde{t}_{Upper} = \tilde{t}_{Mid}$;

    else if $\left. \frac{df^*(\tilde{t},\lambda)}{d\tilde{t}} \right|_{\tilde{t}=\tilde{t}_{Mid}} < 0$ **then**

      $\tilde{t}_{Upper} = \tilde{t}_{Mid}$;

    else

      $\tilde{t}_{Lower} = \tilde{t}_{Mid}$;

    end if

  end while

  $\tilde{t}^*(\lambda_{Mid}) = \frac{\tilde{t}_{Upper} + \tilde{t}_{Lower}}{2}$;

  Obtain $\tilde{P}^*(\tilde{t}^*(\lambda_{Mid}), \lambda_{Mid})$, $\tilde{T}^*(\tilde{t}^*(\lambda_{Mid}), \lambda_{Mid})$, $z^*(\tilde{t}^*(\lambda_{Mid}), \lambda_{Mid})$;

  if $\left. \frac{dg(\lambda)}{d\lambda} \right|_{\lambda=\lambda_{Mid}} > 0$ **then**

    $\lambda_{Upper} = \lambda_{Mid}$;

  else

    $\lambda_{Lower} = \lambda_{Mid}$;

  end if

end while

$\lambda^* = \frac{\lambda_{Upper} + \lambda_{Lower}}{2}$;

Return $\tilde{P}^*(\tilde{t}^*(\lambda^*), \lambda^*)$, $\tilde{T}^*(\tilde{t}^*(\lambda^*), \lambda^*)$, $z^*(\tilde{t}^*(\lambda^*), \lambda^*)$, λ^*, $\tilde{t}^*(\lambda^*)$, $\gamma^*(\tilde{t}^*(\lambda^*), \lambda^*)$.

---

used. According to IEEE standard 802.15.6 [79], the transceivers of body sensor networks work at frequency band 402MHz-405MHz and there are 10 channels. In the numerical analysis, only one channel is used and the bandwidth W = 0.3MHz.

The path loss in dB at distance d with reference distance d_0 can be calculated as

$$PL(d) = PL(d_0) + 10n\log\left(\frac{d}{d_0}\right) \tag{4.24}$$

For in-body channels between the implanted sensors and the on-body relays, we use the channel

model parameters from [82]. The reference distance d_0 is 5cm. Path loss at the reference location is 47.14dB and the path loss exponent n is 4.26. For on-body channels between the relays and the coordinator, channel model parameters from [83] are applied. The reference distance d_0 is 10cm. Path loss at the reference location is 35.20dB and the path loss exponent n is 3.11.

To evaluate the proposed algorithm, multi-hop network with only transmission power control as in [43] and multi-hop network with only relay location optimization as in [73] are used as comparisons. In the multi-hop network with only transmission power control and no relay location optimization, one relay CS is selected randomly within each body region and is used to relay the messages of the implanted sensors in that region. For the multi-hop network with only relay location control, the transmission power for each sensor node is set to be -10dBmW.

Furthermore, in order to check the duality gap between the primal problem (4.7) and the dual problem (4.10), the optimal network lifetime from solving the centralized problem (4.7) directly is also plotted. Problem (4.7) is solved by enumerating all possible values of z directly and selecting the z value that results in the maximum objective value (4.7a). Note that the reason problem (4.7) can be solved directly using a centralized method is that we assume a small number of relay CSs in the numerical analysis. As the number of relay CSs increases, the number of the cases to enumerate for variable z in the centralized solution will increase with a power of 6 according to (4.7f) and (4.7h). The running time for the centralized method can become unreasonably long very easily. For example, if the number of relay CSs is doubled in each body region, the running time for the centralized solution will become 2^6 times as much as that of the original problem. For comparison, the running time for the proposed decomposition method only increases linearly to the number of relay CSs according to sub-problem (4.13).

The optimization problems are solved by CVX 2.1 with Mosek solver in Windows/64-X86 environment. The platform is equipped with Xeon(R) processor with 4 cores and the operating frequency is 3.5GHz.

4.5.2 Results Analysis

Figure 4.5 shows a typical result of the optimal network topology and data flow. The data rate for each sensor is 40Kbps and the superframe length is 400ms. The black lines represent the optimal traffic paths among the in-body sensors, the on-body relays and the coordinator. The thickness of the lines is proportional to the channel capacity.
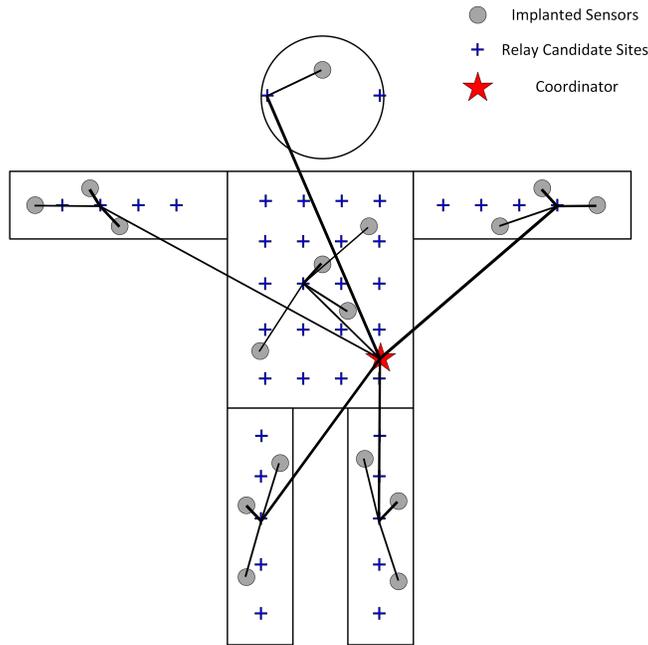
58

Figure 4.5: The optimal network topology and channel capacity

The influence of the traffic generated by sensors

The influence of the traffic generated by sensors on the network lifetime is shown in Figure 4.6. It is assumed that all sensors generate data streams with the same speed. The superframe length T_{frame} is set to be 400ms. The proposed network optimization method, which jointly considers the relay location control and network cross-layer optimization, achieves much longer network lifetime than the alternatives. Specifically, compared with the other multi-hop networks simulated, the proposed network achieves around 20% longer network lifetime. As data rate increases, the energy consumption per superframe grows larger and the sensors' energy is consumed faster. So the network lifetime becomes shorter when traffic of the sensors becomes larger as is illustrated in Figure 4.6.

Note that when the data rate for each sensor node is higher than 40Kbps, the network with fixed transmission power fails to provide feasible solutions. When the data rate for each sensor node is higher than 45Kbps, the network without relay location optimization is unable to provide feasible solutions. The reason is that these two methods search over a smaller feasible set compared with the

59

proposed approach. In this way, even when the average traffic for the sensor nodes grows to around 50Kbps, the proposed approach is still able to guarantee that all sensors finish their transmissions within T_{frame} without collisions and provides feasible solutions.

From Figure 4.6, it can be further concluded that compared with solving the centralized problem (4.7) directly, the proposed decomposition method yields the optimal network lifetime values that are very close. So the duality gap between the optimal values of problem (4.7) and (4.10) is trivial and the solution of the dual problem (4.10) is a good approximation to that of the primal problem (4.7).

The gap between the results of the network without relay location optimization and that of the proposed network grows larger when traffic of the sensors increases. The reason is that when the sensors' data rate is high, the optimal transmission power approaches to the maximum value and power control becomes less helpful. The effect of relay location optimization becomes more profound.
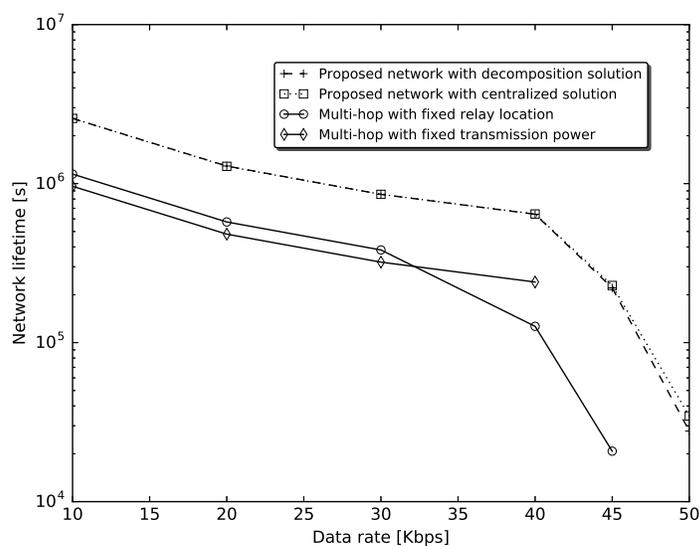

Figure 4.6: The influence of the traffic generated by sensors

The influence of the number of sensors

The optimal network lifetime of deploying different number of sensors are plotted in Figure 4.7. The traffic generated by each sensor is set to be 40Kbps and the superframe length is set to be

400ms. There are originally 17 sensors as in Figure 4.1a and we consider the influence of removing some of the sensors. Though the sequence of removing or adding the sensors can influence the final results, the characteristics exhibited stay the same. As is shown in Figure 4.7, the proposed network optimization method yields the highest network lifetime, which is consistent with the previous results.

Note that the curves exhibit step shape. For example, when the number of the sensor nodes increases from 11 to 12, there is an obvious network lifetime decrease for all curves. Around this decrease, the values of the curves remain stable. In our experiment, this is when the 12th sensor node is added to the torso and a new "bottleneck" node shows up. Here the "bottleneck" indicates the sensor node with the lowest lifetime value in the network, which equals to the value of the whole network's lifetime. If the number of sensors becomes large, because the total time for the transmissions in a round is limited and constrained according to (4.4b), the transmission power of the sensors needs to be improved to transmit faster and meet the time constraint. The energy spent in each superframe for each node can be calculated as

$$E_i = T_i P_i = \frac{x_i T_{frame}}{r_{ij}} P_i = \frac{x_i T_{frame}}{W \log \left(1 + \frac{\alpha_{ij} P_i}{N_j} \right)} P_i \tag{4.25}$$

so the higher the transmission power is, the higher energy consumption will be in each superframe and the lifetime of the sensor nodes will decrease.

Since the objective of the optimization problem is to maximize the shortest sensor lifetime, when a new sensor is added, the bottleneck sensor will not be affected if it is possible to add the new sensor and only decrease the non-bottleneck sensors' lifetime. So during the process of adding more sensors, we are expected to see, when only non-bottleneck sensors' lifetime is shortened and the bottleneck sensor's lifetime remains the same, the network lifetime does not change. Only when the lifetime of all sensors are close enough, after adding a new sensor, a new bottleneck sensor appears and the network lifetime decreases. This is the reason why step shaped curves show up.

Energy consumption comparison

Figure 4.8 shows the total network energy consumption in each superframe of the methods considered. The proposed network optimization approach achieves the lowest energy consumption, which is consistent with the analysis of the network lifetime. The results of the proposed decomposition solution and centralized solution are still almost identical. When traffic of each node is
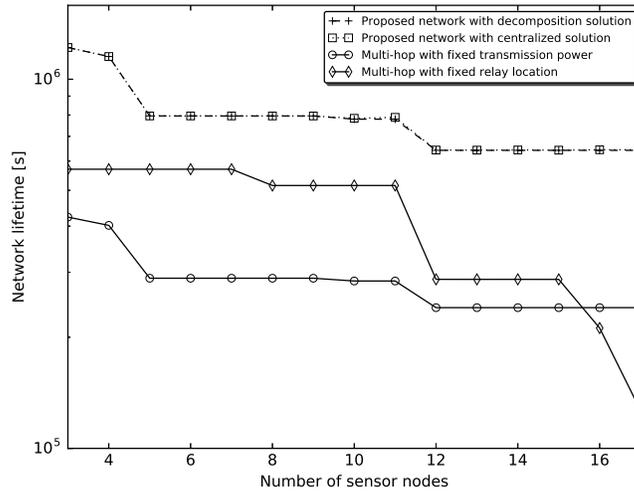
Figure 4.7: The influence of the number of sensors

larger than 40Kbps, the energy consumption of all methods grows up sharply, which coincides with the results of network lifetime in Figure 4.6. From Figure 4.6, it can be seen that the network lifetime decreases fast when data rate is beyond 40Kbps.

Lifetime of the sensor nodes

Figure 4.9 shows the lifetime of all sensor nodes derived from the proposed decomposition method. Still the traffic for each sensor is 40Kbps and the superframe length is 400ms. From Figure 4.9, it can be concluded that the lifetime results of the sensors in different regions are quite even. This indicates that no body region is the bottleneck of the whole network's lifetime. If the network lifetime is to be increased, more relays need to be added to all regions at the same time.

The influence of the number of relays

Here we provide a brief discussion on how the number of relays influences the network lifetime. In the numerical analysis, there are originally 6 relays as in Figure 4.5 and we choose to add relays to the body regions evenly. The number of relays on the head is kept to be one since there is only one sensor node and the network lifetime will not be influenced by adding more relays there. Each sensor node will choose the nearest relay that is in the same region.

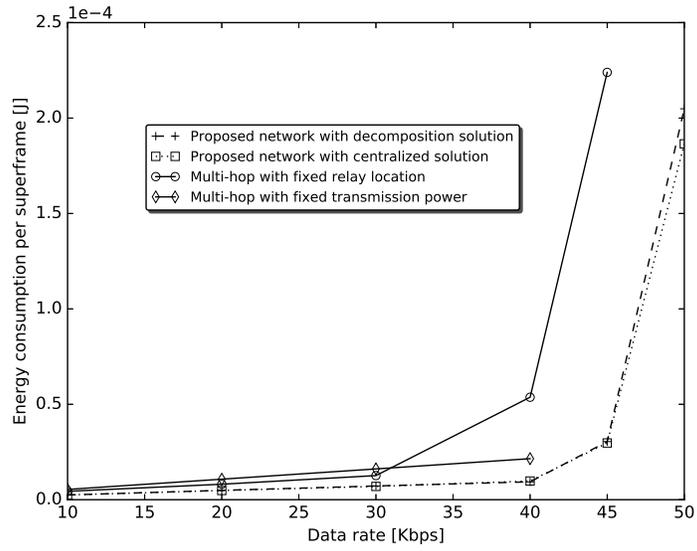In this section, we only provide the results of the proposed decomposition method based on
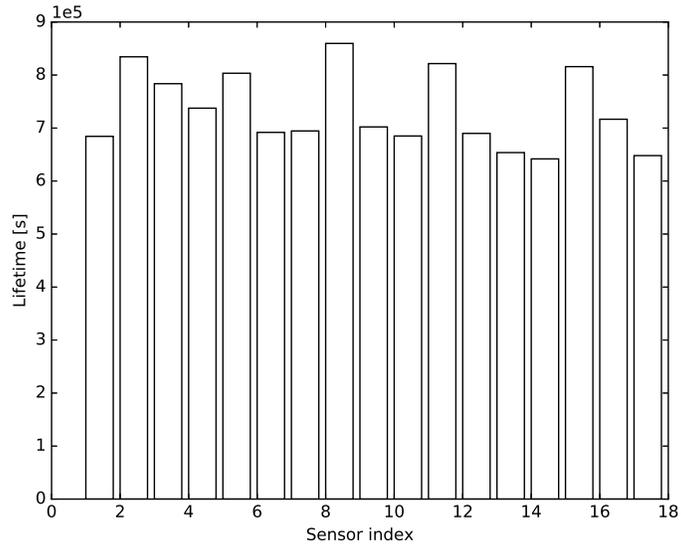
Figure 4.8: Energy consumption comparison

Figure 4.9: Lifetime of the sensors

binary search. The reason is that when it is allowed to choose more than one relays in each body

region, the number of occasions to enumerate in the centralized method grows very fast as mentioned previously. The amount of time consumed to solve the centralized problem soon becomes unreasonable. In comparison, the time consumed by the decomposition method grows much slower with the increase of relay number.
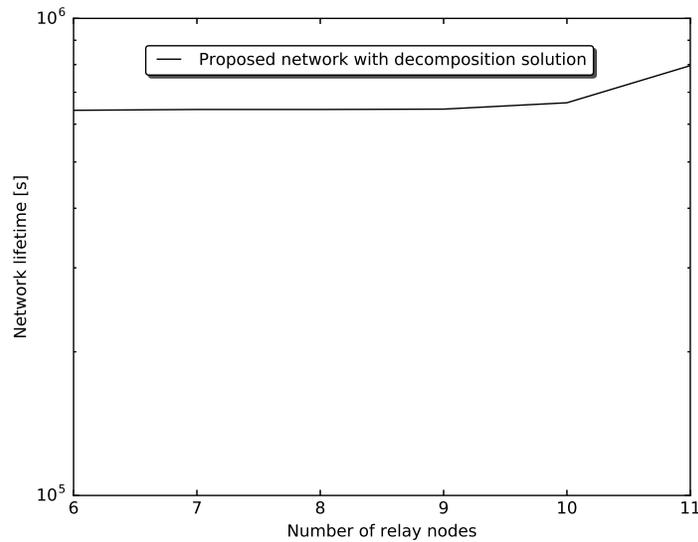

Figure 4.10: The influence of the number of relays

As is shown in Figure 4.10, the largest increase of the network lifetime happens between the 10th and the 11th relay, before which the network lifetime only increases moderately. This observation can be explained by the results shown in Figure 4.9. In Figure 4.9 there are 6 relays, and for each region, there is some sensor node whose lifetime is almost as low as that of the bottleneck sensor node. That is to say, in order to increase the whole network lifetime, more relays should be added to nearly all the regions, so that each of these short-lived sensors are tackled and their lifetimes are prolonged. Before that, the network lifetime only increases moderately as the number of relays grows.

The influence of body gesture

Body gesture change is one of the key features that distinguish WBSN from other general WSNs. In order to look into the influence of body movement on the network performance, we conduct numerical analysis on three different standing positions as shown in Figure 4.11. According to
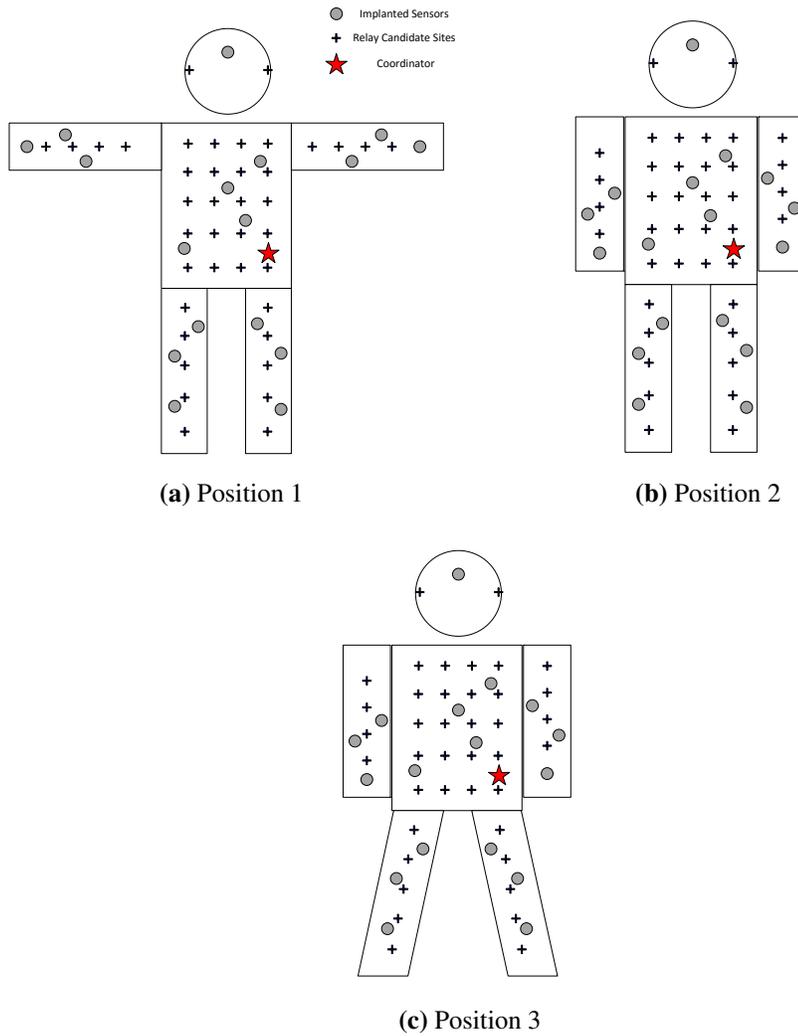
(a) Position 1

(b) Position 2

(c) Position 3

Figure 4.11: Three different standing positions in the numerical analysis

the analysis in Section III, by dividing body into several body regions, the relative distances between the sensors and relays are kept stable and the influence of body position change on network performance should be alleviated.

The numerical results are shown in Table 4.2 with average data generation rate for each sensor as an independent variable. The network lifetimes in seconds for each data rate and scenario composition are listed. "NaN" represents no feasible solution is found. The proposed model provides consistently better network lifetime performance than the rest two alternative models. From the

Table 4.2: Network lifetime [s] results for three body positions

		10Kbps	20Kbps	30Kbps	40Kbps	45Kbps	50Kbps
Position 1	Proposed decomposition	2.56923e+06	1.29017e+06	8.56168e+05	6.41688e+05	2.21018e+05	2.78254e+04
	Proposed centralized	2.57099e+06	1.28446e+06	8.56446e+05	6.44264e+05	2.28773e+05	3.45601e+04
	Fixed relay location	1.14800e+06	5.73944e+05	3.82667e+05	1.26615e+05	2.07722e+04	NaN
	Fixed transmission power	9.62167e+05	4.81084e+05	3.20722e+05	2.40542e+05	NaN	NaN
Position 2	Proposed decomposition	2.56655e+06	1.28378e+06	8.55537e+05	6.41629e+05	3.06432e+05	5.37266e+04
	Proposed centralized	2.56864e+06	1.28457e+06	8.56274e+05	6.42574e+05	3.16206e+05	5.87761e+04
	Fixed relay location	1.14799e+06	5.74002e+05	3.82592e+05	1.35012e+05	2.79811e+04	NaN
	Fixed transmission power	9.62167e+05	4.81084e+05	3.20722e+05	2.40542e+05	2.13815e+05	NaN
Position 3	Proposed decomposition	2.56655e+06	1.28320e+06	8.55537e+05	6.41628e+05	1.87684e+05	3.38611e+04
	Proposed centralized	2.56903e+06	1.28532e+06	8.56575e+05	6.42618e+05	2.24166e+05	3.84971e+04
	Fixed relay location	1.14785e+06	5.73218e+05	3.82615e+05	1.05530e+05	NaN	NaN
	Fixed transmission power	9.62167e+05	4.81084e+05	3.20722e+05	2.40542e+05	NaN	NaN

table, it can be concluded that when data rate is low, the performance of different body positions are very close. When data rate is higher, body gesture change results in more obvious influence on the lifetime performance. The proposed model with both decomposition and centralized methods are able to find feasible solutions for all the data rates that are researched in the numerical analysis. However, the rest two comparisons often fail to find feasible solutions when data rate is high.

Running time comparison

As is shown in Figure 4.12, in the numerical analysis, the decomposition algorithm based on binary search consumes less time compared with the other methods. The rectangular bars represent the average values and the line segments on the top represent the value range of one standard deviation. The decomposition method based on subgradient descent/ascent approach is time consuming and binary search is able to expedite the running process greatly. The centralized approach requires about twice as much running time compared with the binary search method. As is shown by the figure, the decomposition methods are influenced by the initial values of the parameters, so the running time varies for different runs a lot. In contrast, the centralized method has a more stable running time. However, note that when the number of relay CSs becomes larger, the running time consumed by the centralized method grows much faster than that of the decomposition method.
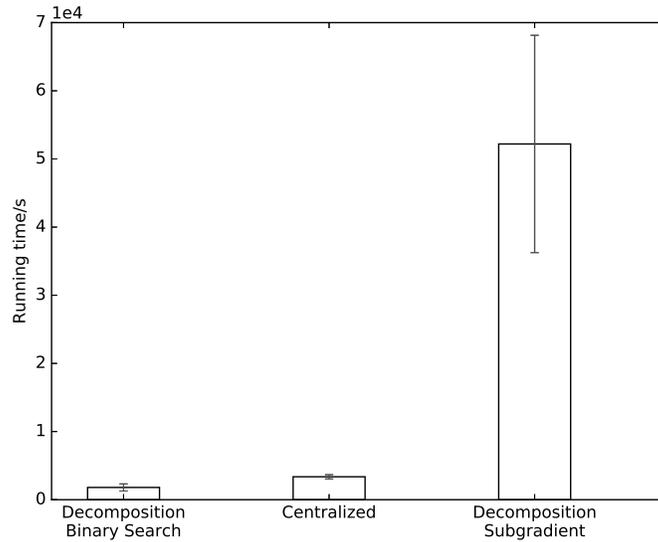
Figure 4.12: Running time comparison

4.6 Summary

In this chapter, we have proposed a network optimization approach that jointly considers the relay location control and network cross-layer optimization, which is suitable for the WBAN context. Multilevel primal and dual decomposition methods have been proposed to transform the original non-convex problem into convex ones, which can be solved by CVX solvers. To expedite the running process of optimization, an effective binary search approach has been proposed. According to the numerical results, compared with the literature, the proposed decomposition method yields clear improvements on the network performances.

Chapter 5

Conclusions and Future Work

5.1 Summary

In this thesis, focusing on fulfilling a highly reliable and energy efficient WBAN model in medial care scenario, network designs across physical layer, MAC layer and routing layer have been proposed.

Considering the specialty of applications in medical care scenario, the data transmission should be highly reliable since the data transmitted can be life-critical. However, the body channels are generally highly unstable owing to the frequent and unpredictable body movements. Furthermore, the characteristics of human body parts are complex, making it extremely challenging to build an effective model for on-body and in-body channels and predict the influence of body parts on the transmission signals. On the MAC layer, the common fixed scheduling that schedules data transmissions for each node regardless their channel conditions may result in failing to avoid bad channels and missing the good ones. The reliability performance of the WBAN will thus be compromised. It is urgent to come up with an effective MAC layer scheduling scheme that can adapt to the channel status, which is able to adjust the scheduling and its parameters according the channel condition.

In Chapter 3, a novel opportunistic scheduling scheme, which can effectively avoid bad channels and utilize the good ones, has been proposed. A two state Markov model, Gilbert model, is adopted to analyze the channel status migration. Considering that people constantly modify their body movement, an effective estimation approach has been proposed to trace the variations of the transition matrix, providing a more accurate estimation for the channel dynamics than the traditional fixed transition matrix approach. A heuristic scheduling scheme causing no extra overhead

to sensor nodes has been proposed and proved to be optimal based on the channel status estimation.

The fundamental effect of a proper superframe length in opportunistic scheduling has been revealed. In order to adjust the superframe length and ensure an optimal value for the current channel dynamics, the scheduler is treated as a decision maker and trained with reinforcement learning techniques. As explained in Chapter 3, a proper superframe length is found crucial at making opportunistic scheduling effective at avoiding the bad channels and take advantage of the good ones.

The combined scheduling scheme yields a much better performance on the outage rate compared with the existing methods according to the simulation results. Specifically, the proposed method can avoid around 9% of the outage occurrence compared with the fixed scheduling and achieve around 30% better performance than the flipping method [11] and more than 35% better than the random method.

Another characteristic of WBANs, especially for the implanted ones, is that the sensor nodes are extremely tiny and energy constrained. At the same time, the sensor nodes or their batteries are generally not easy to be replaced or recharged. In order to prolong the network lifetime, beyond MAC layer design, a cross-layer network optimization approach has been further presented in Chapter 4. The optimization model that jointly considers power control and relay selection strategies serves as a supplement to the proposed MAC layer scheduling scheme and is designed for WBAN in medical care context.

In order to solve the formulated optimization problem, multilevel primal and dual decomposition methods have been developed to transform the original non-convex problem into independent convex subproblems, which can be solved by CVX solvers efficiently. To expedite the running process of optimization, an effective binary search approach has been proposed. Since the proposed optimization approach searches the optimal solution within a larger feasible set than doing optimization in each layer separately, compared with the literature, the proposed decomposition method is able to achieve better solutions for higher network lifetime performances. The influences of sensor node number, relay node number, body gesture and data rate have been explored and according to the simulation results, the proposed network model yields consistently longer network lifetime than the existing methods.

Through this thesis, I have achieved a thorough understanding of the design caveats of WBANs. I have got a clear knowledge of what are the major differences between WBANs and other similar networks, such as WSNs. In order to promote the performances on reliability and power consumption, I have gradually learned which layers need to be optimized and what are the major factors that can be improved. My researching abilities have been enhanced greatly while doing this thesis.

5.2 Future Work

The following are suggested areas for future work and improvement to this thesis:

1. For the current proposed MAC design, the analysis for the retransmission approach is not included. In the next step, an effective retransmission strategy can be researched to ensure both the reliability and low latency for sensor nodes' data transmissions, which are the two major considerations for WBANs in medical care scenario.

2. For the optimization model proposed in Chapter 4, only 2-dimensional coordinates are applied for the simplicity of analysis. However, in the real life, locations of the sensors on or in human body should be depicted by 3-dimensional coordinates. With 3-dimensional coordinates, other body gestures, such as sitting on a chair can be further researched to verify the effectiveness of the proposed algorithm under more extensive occasions. More interesting conclusions may be obtained by examining other body gestures.

3. In this thesis, the results obtained are from Matlab simulations. Though channel data used is from real life, still, the WBANs are simulated. For the next step, the algorithms designed in this article can be implemented on the testbed. For the on-body network designs, volunteers can be hired to wear the sensor nodes to examine the performance of the proposed algorithms in real life scenarios. For the implanted WBANs, collaborations with the hospitals need to be achieved to test the effectiveness of the proposed algorithm.

Bibliography

[1] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1658–1686, 2014.

[2] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *IEEE Wireless Communications Magazine*, vol. 17, no. 1, pp. 80–88, 2010.

[3] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov, "System architecture of a wireless body area sensor network for ubiquitous health monitoring," *Journal of mobile multimedia*, vol. 1, no. 4, pp. 307–326, 2006.

[4] M. Chen, S. Gonzalez, A. Vasilakos, H. Cao, and V. C. Leung, "Body area networks: A survey," *Mobile networks and applications*, vol. 16, no. 2, pp. 171–193, 2011.

[5] A. Milenković, C. Otto, and E. Jovanov, "Wireless sensor networks for personal health monitoring: Issues and an implementation," *Computer communications*, vol. 29, no. 13, pp. 2521–2533, 2006.

[6] K. J. Cash and H. A. Clark, "Nanosensors and nanomaterials for monitoring glucose in diabetes," *Trends in molecular medicine*, vol. 16, no. 12, pp. 584–593, 2010.

[7] F. Moussy, "Implantable glucose sensor: progress and problems," in *Proceedings of IEEE Sensors*, vol. 1, pp. 270–273, IEEE, 2002.

[8] J. Wang, X. W. Sun, A. Wei, Y. Lei, X. Cai, C. M. Li, and Z. L. Dong, "Zinc oxide nanocomb biosensor for glucose detection," *Applied Physics Letters*, vol. 88, no. 23, p. 233106, 2006.

[9] K. S. McKeating, A. Aubé, and J.-F. Masson, "Biosensors and nanobiosensors for therapeutic drug and response monitoring," *Analyst*, vol. 141, no. 2, pp. 429–449, 2016.

[10] A. Boulis, D. Smith, D. Miniutti, L. Libman, and Y. Tselishchev, "Challenges in body area networks for healthcare: The MAC," *IEEE Communications Magazine*, vol. 50, no. 5, 2012.

[11] Y. Tselishchev, A. Boulis, and L. Libman, "Variable scheduling to mitigate channel losses in energy-efficient body area networks," *Sensors*, vol. 12, no. 11, pp. 14692–14710, 2012.

[12] S. Li, F. Hu, and G. Li, "Advances and challenges in body area network," in *International Conference on Applied Informatics and Communication*, pp. 58–65, Springer, 2011.

[13] S. Ullah, H. Higgins, B. Braem, B. Latre, C. Blondia, I. Moerman, S. Saleem, Z. Rahman, and K. S. Kwak, "A comprehensive survey of wireless body area networks," *Journal of medical systems*, vol. 36, no. 3, pp. 1065–1094, 2012.

[14] "IEEE standard for local and metropolitan area networks - part 15.6: Wireless body area networks," *IEEE Std 802.15.6-2012*, pp. 1–271, Feb 2012.

[15] M. Lipprandt, M. Eichelberg, W. Thronicke, J. Kruger, I. Druke, D. Willemsen, C. Busch, C. Fiehe, E. Zeeb, and A. Hein, "OSAMI-D: an open service platform for healthcare monitoring applications," in *2009 2nd Conference on Human System Interactions*, pp. 139–145, IEEE, 2009.

[16] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.

[17] B. Latré, B. Braem, I. Moerman, C. Blondia, and P. Demeester, "A survey on wireless body area networks," *Wireless Networks*, vol. 17, no. 1, pp. 1–18, 2011.

[18] S. Wang and J.-T. Park, "Modeling and analysis of multi-type failures in wireless body area networks with semi-markov model," *IEEE Communications Letters*, vol. 14, no. 1, pp. 6–8, 2010.

[19] L. Hanlen, D. Miniutti, D. Smith, D. Rodda, and B. Gilbert, "Co-channel interference in body area networks with indoor measurements at 2.4 GHz: Distance-to-interferer is a poor estimate of received interference power," *International Journal of Wireless Information Networks*, vol. 17, no. 3-4, pp. 113–125, 2010.

[20] B. Zhen, M. Patel, S. Lee, E. Won, and A. Astrin, "Tg6 technical requirements document (TRD) IEEE p802.15-08-0644-09-0006," Sept. 2008.

[21] R. C. Shah and M. Yarvis, "Characteristics of on-body 802.15. 4 networks," in *2006 2nd IEEE Workshop on Wireless Mesh Networks*, pp. 138–139, IEEE, 2006.

[22] C. Tachtatzis, F. Di Franco, D. C. Tracey, N. F. Timmons, and J. Morrison, "An energy analysis of ieee 802.15. 6 scheduled access modes," in *2010 IEEE Globecom Workshops*, pp. 1270–1275, IEEE, 2010.

[23] M. Sukor, S. Ariffin, N. Fisal, S. S. Yusof, and A. Abdallah, "Performance study of wireless body area network in medical environment," in *Second Asia International Conference on Modeling & Simulation, 2008*, pp. 202–206, IEEE, 2008.

[24] "IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks specific requirements part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (LR-WPANs)," *IEEE Std 802.15.4-2003*, 2003.

[25] N. Torabi and V. C. Leung, "Robust license-free body area network access for reliable public m-health services," in *2011 13th IEEE International Conference on e-Health Networking Applications and Services (Healthcom)*, pp. 312–319, IEEE, 2011.

[26] G. Fang and E. Dutkiewicz, "BodyMAC: Energy efficient TDMA-based MAC protocol for wireless body area networks," in *9th International Symposium on Communications and Information Technology, 2009*, pp. 1455–1459, IEEE, 2009.

[27] N. F. Timmons and W. G. Scanlon, "An adaptive energy efficient MAC protocol for the medical body area network," in *1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, 2009*, pp. 587–593, IEEE, 2009.

[28] K. Bilstrup, "A preliminary study of wireless body area networks," 2008.

[29] S. Gowrishankar, T. G. Basavaraju, M. D.H., and S. K. Sarkar, "Issues in wireless sensor networks," *Proceedings of The World Congress on Engineering 2008*, pp. 176–187.

[30] D. M. Barakah and M. Ammad-uddin, "A survey of challenges and applications of wireless body area network (WBAN) and role of a virtual doctor server in existing architecture," in *2012 Third International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, pp. 214–219, IEEE, 2012.

[31] B. Antonescu and S. Basagni, "Wireless body area networks: challenges, trends and emerging technologies," in *Proceedings of the 8th international conference on body area networks*, pp. 1–7, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013.

[32] M. A. Hanson, H. C. Powell Jr, A. T. Barth, K. Ringgenberg, B. H. Calhoun, J. H. Aylor, and J. Lach, "Body area sensor networks: Challenges and opportunities," *Computer*, vol. 42, no. 1, p. 58, 2009.

[33] G. Ragesh and K. Baskaran, "An overview of applications, standards and challenges in futuristic wireless body area networks," 2012.

[34] A. Bachir, M. Dohler, T. Watteyne, and K. K. Leung, "MAC essentials for wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 222–248, 2010.

[35] J. Kang and S. Adibi, "A review of security protocols in mhealth wireless body area networks (WBAN)," in *International Conference on Future Network Systems and Security*, pp. 61–83, Springer, 2015.

[36] S.-D. Bao and Y.-T. Zhang, "A design proposal of security architecture for medical body sensor networks," in *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)*, pp. 4–pp, IEEE, 2006.

[37] O. G. Morchon and H. Baldus, "Efficient distributed security for wireless medical sensor networks," in *2008 International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 249–254, IEEE, 2008.

[38] M. Li, W. Lou, and K. Ren, "Data security and privacy in wireless body area networks," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 51–58, 2010.

[39] J. D. Lee, T. S. Yoon, S. H. Chung, and H. S. Cha, "Service-oriented security framework for remote medical services in the internet of things environment," *Healthcare informatics research*, vol. 21, no. 4, pp. 271–282, 2015.

[40] S. Ullah, M. Mohaisen, and M. A. Alnuem, "A review of IEEE 802.15. 6 MAC, PHY, and security specifications," *International Journal of Distributed Sensor Networks*, vol. 2013, 2013.

[41] R. Pan, D. Chua, J. S. Pathmasuntharam, and Y. P. Xu, "An opportunistic relay protocol with dynamic scheduling in wireless body area sensor network," *IEEE Sensors Journal*, vol. 15, no. 7, pp. 3743–3750, 2015.

[42] J. Dong and D. Smith, "Joint relay selection and transmit power control for wireless body area networks coexistence," in *2014 IEEE International Conference on Communications (ICC)*, pp. 5676–5681, IEEE, 2014.

[43] N. Javaid, A. Ahmad, Y. Khan, Z. A. Khan, and T. A. Alghamdi, "A relay based routing protocol for wireless in-body sensor networks," *Wireless Personal Communications*, vol. 80, no. 3, pp. 1063–1078, 2015.

[44] A. L. Stolyar, "Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm," *Queueing Systems*, vol. 50, no. 4, pp. 401–457, 2005.

[45] X. Lin and N. B. Shroff, "Joint rate control and scheduling in multihop wireless networks," in *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, vol. 2, pp. 1484–1489, IEEE, 2004.

[46] X. Lin and N. B. Shroff, "The impact of imperfect scheduling on cross-layer rate control in wireless networks," in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 3, pp. 1804–1814, IEEE, 2005.

[47] A. Eryilmaz and R. Srikant, "Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control," in *IEEE/ACM Transactions on Networking*, vol. 3, pp. 1794–1803, IEEE, 2005.

[48] A. Eryilmaz and R. Srikant, "Joint congestion control, routing, and MAC for stability and fairness in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1514–1524, 2006.

[49] M. Johansson and L. Xiao, "Scheduling, routing and power allocation for fairness in wireless networks," in *2004 IEEE 59th Vehicular Technology Conference. VTC 2004-Spring (IEEE Cat. No.04CH37514)*, vol. 3, pp. 1355–1360, IEEE, 2004.

[50] S. A. Gopalan and J.-T. Park, "Energy-efficient MAC protocols for wireless body area networks: survey," in *International Congress on Ultra Modern Telecommunications and Control Systems*, pp. 739–744, IEEE, 2010.

[51] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell system technical journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[52] Y. Tselishchev, L. Libman, and A. Boulis, "Reducing transmission losses in body area networks using variable TDMA scheduling," in *2011 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, pp. 1–10, IEEE, 2011.

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

[54] H. S. Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 1, pp. 163–171, 1995.

[55] K. Y. Yazdandoost, K. Sayrafian-Pour, *et al.*, "Channel model for body area network (BAN)," *IEEE P802*, vol. 15, 2009.

[56] D. B. Smith, A. Boulis, and Y. Tselishchev, "Efficient conditional-probability link modeling capturing temporal variations in body area networks," in *Proceedings of the 15th ACM international conference on Modeling, analysis and simulation of wireless and mobile systems*, pp. 271–276, ACM, 2012.

[57] P. Huang, L. Xiao, S. Soltani, M. W. Mutka, and N. Xi, "The evolution of MAC protocols in wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 101–120, 2013.

[58] V. Rajendran, K. Obraczka, and J. J. Garcia-Luna-Aceves, "Energy-efficient, collision-free medium access control for wireless sensor networks," *Wireless Networks*, vol. 12, no. 1, pp. 63–78, 2006.

[59] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Transactions on networking*, vol. 12, no. 3, pp. 493–506, 2004.

[60] G. Lu, B. Krishnamachari, and C. S. Raghavendra, "An adaptive energy-efficient and low-latency MAC for data gathering in wireless sensor networks," in *18th International Parallel and Distributed Processing Symposium, 2004. Proceedings.*, p. 224, IEEE, 2004.

[61] A. El-Hoiydi and J.-D. Decotignie, "WiseMAC: an ultra low power MAC protocol for the downlink of infrastructure wireless sensor networks," in *Proceedings. ISCC 2004. Ninth International Symposium on Computers And Communications (IEEE Cat. No.04TH8769)*, vol. 1, pp. 244–251, IEEE, 2004.

[62] L. Tang, Y. Sun, O. Gurewitz, and D. B. Johnson, "Em-MAC: a dynamic multichannel energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, p. 23, ACM, 2011.

[63] S. Ullah, M. Imran, and M. Alnuem, "A hybrid and secure priority-guaranteed MAC protocol for wireless body area network," *International Journal of Distributed Sensor Networks*, vol. 2014, 2014.

[64] M. Momoda and S. Hara, "A cooperative relaying scheme for real-time vital data gathering in a wearable wireless body area network," in *2013 7th International Symposium on Medical Information and Communication Technology (ISMICT)*, pp. 38–41, IEEE, 2013.

[65] N. Torabi, C. Kaur, and V. C. Leung, "Distributed dynamic scheduling for body area networks," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pp. 3177–3182, IEEE, 2012.

[66] N. Torabi and V. C. 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, 2013.

[67] B. Braem, B. Latre, I. Moerman, C. Blondia, and P. Demeester, "The wireless autonomous spanning tree protocol for multihop wireless body area networks," in *2006 Third Annual International Conference on Mobile and Ubiquitous Systems: Networking Services*, pp. 1–8, IEEE, 2006.

[68] A. Asadi and V. Mancuso, "A survey on opportunistic scheduling in wireless communications," *IEEE Communications Surveys Tutorials*, vol. 15, no. 4, pp. 1671–1688, 2013.

[69] I. . T. G. . A. online. http://www.ieee802.org/15/pub/TG6.html.

[70] B. D. Fritchman, "A binary channel characterization using partitioned markov chains," *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 221–227, 1967.

[71] S. Srinivas and K. Shanmugan, "Markov models for burst errors in radio communications channels," in *Wireless Personal Communications*, pp. 175–184, Springer, 1994.

[72] K. S. Prabh and J.-H. Hauer, "Opportunistic packet scheduling in body area networks," in *Wireless Sensor Networks*, pp. 114–129, Springer, 2011.

[73] J. Elias and A. Mehaoua, "Energy-aware topology design for wireless body area networks," in *2012 IEEE International Conference on Communications (ICC)*, pp. 3409–3410, IEEE, 2012.

[74] X. Zhou, T. Zhang, L. Song, and Q. Zhang, "Energy efficiency optimization by resource allocation in wireless body area networks," in *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, pp. 1–6, IEEE, 2014.

[75] D. R. B. G. J. D. D. Smith, L. Hanlen and V. Chaganti, "Body area network radio channel measurement set." http://filestore.nicta.com.au/Comms/OpenNICTA/ NICTA_BodyAreaNetwork_RadioChannel_Data/. Accessed 10-Dec-2015.

[76] C. J. C. H. Watkins, *Learning from delayed rewards*. PhD thesis, University of Cambridge England, 1989.

[77] R. R. Boorstyn and H. Frank, "Large-scale network topological optimization," *IEEE Transactions on Communications*, vol. 25, no. 1, pp. 29–47, 1977.

[78] S. Movassaghi, M. Abolhasan, and J. Lipman, "A review of routing protocols in wireless body area networks," *Journal of Networks*, vol. 8, no. 3, pp. 559–575, 2013.

[79] A. Astrin *et al.*, "IEEE standard for local and metropolitan area networks part 15.6: Wireless body area networks: IEEE std 802.15. 6-2012," *The document is available at IEEE Xplore*, 2012.

[80] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2185–2193, 2006.

[81] N. Z. Shor, *Minimization methods for non-differentiable functions*, vol. 3. Springer Science & Business Media, 2012.

[82] K. Sayrafian-Pour, W.-B. Yang, J. Hagedorn, J. Terrill, K. Y. Yazdandoost, and K. Hamaguchi, "Channel models for medical implant communication," *International Journal of Wireless Information Networks*, vol. 17, no. 3-4, pp. 105–112, 2010.

[83] E. Reusens, W. Joseph, B. Latré, B. Braem, G. Vermeeren, E. Tanghe, L. Martens, I. Moerman, and C. Blondia, "Characterization of on-body communication channel and energy efficient topology design for wireless body area networks," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 6, pp. 933–945, 2009.

Appendix A

Channel Belief Calculation in Chapter 3

ε is the probability of transiting from bad channel to good channel and δ is the probability of transiting from good channel to bad channel. Assume the initial channel state belief is $\begin{bmatrix} 0 & 1 \end{bmatrix}^T$. According to our definition, this represents a good channel, where the probability of having a good channel is 1 and the probability of having a bad channel is 0. So the channel belief after t time slots would be

$$
\begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \times \begin{bmatrix} 1-\varepsilon & \varepsilon \\ \delta & 1-\delta \end{bmatrix}^t
$$
$$
= \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \times \begin{bmatrix} 1 & 1 \\ 1 & -\frac{\delta}{\varepsilon} \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 0 & 1-\varepsilon-\delta \end{bmatrix}^t \times \begin{bmatrix} \frac{\delta}{\varepsilon} & 1 \\ 1 & -1 \end{bmatrix} \times \frac{1}{1+\frac{\delta}{\varepsilon}} \qquad \text{(A.1)}
$$
$$
= \begin{bmatrix} [1-(1-\varepsilon-\delta)^t] \times \frac{\delta}{\varepsilon+\delta} \\ [1+\frac{\delta}{\varepsilon} \times (1-\varepsilon-\delta)^t] \times \frac{\varepsilon}{\varepsilon+\delta} \end{bmatrix}^T
$$

So the probability of having a good channel after t time slots for node i is $p_1^i(t) = [1 + \frac{\delta_i}{\varepsilon_i} \times (1-\varepsilon_i-\delta_i)^t] \times \frac{\varepsilon_i}{\varepsilon_i+\delta_i}$ when the initial state is $\begin{bmatrix} 0 & 1 \end{bmatrix}^T$. This is an exponential function. Following the same process, we can derive the probability of having a good channel when the initial state is $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$.

Appendix B

Proof of Lemma 1 in Chapter 3

Consider two nodes, where node i and node j are the two nodes that are both in the "successful" group (the failed nodes can be proved with the same process). Node i has a larger utility function than node j, which means

$$U_i = p_1^i(N_i) - p_1^i(N_i + N_{Good}) > U_j = p_1^j(N_j) - p_1^j(N_j + N_{Good}) \tag{B.1}$$

Using the results in Appendix A, we have

$$[1 + \frac{\delta_i}{\varepsilon_i} \times (1 - \varepsilon_i - \delta_i)^{N_i}] \times \frac{\varepsilon_i}{\varepsilon_i + \delta_i} - [1 + \frac{\delta_i}{\varepsilon_i} \times (1 - \varepsilon_i - \delta_i)^{N_i + N_{Good}}] \times \frac{\varepsilon_i}{\varepsilon_i + \delta_i}$$
$$> [1 + \frac{\delta_j}{\varepsilon_j} \times (1 - \varepsilon_j - \delta_j)^{N_j}] \times \frac{\varepsilon_j}{\varepsilon_j + \delta_j} - [1 + \frac{\delta_j}{\varepsilon_j} \times (1 - \varepsilon_j - \delta_j)^{N_j + N_{Good}}] \times \frac{\varepsilon_j}{\varepsilon_j + \delta_j} \tag{B.2}$$

And further

$$(1 - \varepsilon_i - \delta_i)^{N_i} \times \frac{\delta_i}{\varepsilon_i + \delta_i} \times [1 - (1 - \varepsilon_i - \delta_i)^{N_{Good}}]$$
$$> (1 - \varepsilon_j - \delta_j)^{N_j} \times \frac{\delta_j}{\varepsilon_j + \delta_j} \times [1 - (1 - \varepsilon_j - \delta_j)^{N_{Good}}] \tag{B.3}$$

If $\varepsilon_i + \delta_i = \varepsilon_j + \delta_j$, we have

$$(1 - \varepsilon_i - \delta_i)^{N_i} \times \frac{\delta_i}{\varepsilon_i + \delta_i} > (1 - \varepsilon_j - \delta_j)^{N_j} \times \frac{\delta_j}{\varepsilon_j + \delta_j} \tag{B.4}$$

Note that N_{Good} does not matter actually. Using $p_1^i(N_i + 1)$ has the same effect as using $p_1^i(N_i + N_{Good})$. In the article we use $p_1^i(N_i + N_{Good})$ for easier illustration.

Assume node i and node j are scheduled at time slot t_i and t_j in the current round respectively. If $t_i > t_j$, that is, node i is scheduled behind node j, then

$$
\begin{aligned}
&\left[p_1^i(N_i+t_j) + p_1^j(N_j+t_i) \right] - \left[p_1^i(N_i+t_i) + p_1^j(N_j+t_j) \right] \\
&= [1 + \tfrac{\delta_i}{\varepsilon_i} \times (1-\varepsilon_i-\delta_i)^{N_i+t_j}] \times \tfrac{\varepsilon_i}{\varepsilon_i+\delta_i} \\
&\quad + [1 + \tfrac{\delta_j}{\varepsilon_j} \times (1-\varepsilon_j-\delta_j)^{N_j+t_i}] \times \tfrac{\varepsilon_j}{\varepsilon_j+\delta_j} \\
&\quad - [1 + \tfrac{\delta_i}{\varepsilon_i} \times (1-\varepsilon_i-\delta_i)^{N_i+t_i}] \times \tfrac{\varepsilon_i}{\varepsilon_i+\delta_i} \\
&\quad - [1 + \tfrac{\delta_j}{\varepsilon_j} \times (1-\varepsilon_j-\delta_j)^{N_j+t_j}] \times \tfrac{\varepsilon_j}{\varepsilon_j+\delta_j} \\
&= \tfrac{\delta_i}{\varepsilon_i+\delta_i}(1-\varepsilon_i-\delta_i)^{N_i} \left[(1-\varepsilon_i-\delta_i)^{t_j} - (1-\varepsilon_i-\delta_i)^{t_i} \right] \\
&\quad + \tfrac{\delta_j}{\varepsilon_j+\delta_j}(1-\varepsilon_j-\delta_j)^{N_j} \left[(1-\varepsilon_j-\delta_j)^{t_i} - (1-\varepsilon_j-\delta_j)^{t_j} \right] \\
&= \left[\tfrac{\delta_i}{\varepsilon_i+\delta_i} \times (1-\varepsilon_i-\delta_i)^{N_i} - \tfrac{\delta_j}{\varepsilon_j+\delta_j} \times (1-\varepsilon_j-\delta_j)^{N_j} \right] \\
&\quad \times \left[(1-\varepsilon_i-\delta_i)^{t_j} - (1-\varepsilon_i-\delta_i)^{t_i} \right]
\end{aligned}
\tag{B.5}
$$

From the previous result we know $(1-\varepsilon_i-\delta_i)^{N_i} \times \tfrac{\delta_i}{\varepsilon_i+\delta_i} > (1-\varepsilon_j-\delta_j)^{N_j} \times \tfrac{\delta_j}{\varepsilon_j+\delta_j}$ and because $t_i > t_j$, we have $(1-\varepsilon_i-\delta_i)^{t_j} - (1-\varepsilon_i-\delta_i)^{t_i} > 0$. Thus we can conclude

$$
p_1^i(N_i+t_j) + p_1^j(N_j+t_i) > p_1^i(N_i+t_i) + p_1^j(N_j+t_j)
\tag{B.6}
$$

This means swapping the slots assigned to node i and node j can improve the expectation of the amount of successfully transmitted packets when node i is scheduled behind node j and $U_i > U_j$. So node i should be scheduled in front of node j instead. This also indicates that all slots assigned to a node should be consecutive because the slot locations are only decided by the order of the utility values of the nodes. Furthermore, if $\varepsilon_i = \varepsilon_j$ and $\delta_i = \delta_j$, (13) becomes $N_i < N_j$ and the proposed method is equivalent to the flipping method [11].

Appendix C

KKT Condition in Chapter 4

Since problem (14) is convex and apparently slater condition can be satisfied, thus strong duality holds. The optimal value of the Lagrange multiplier of (14b) can be derived by first forming the Lagrangian including all inequality constraints of problem (14) as

$$
\begin{aligned}
&L\left(\gamma, \eta, \nu, \omega, \tau, \tilde{P}, \tilde{T}\right) \\
&= \tilde{t} - \lambda \left(\sum_{i \in S} e^{\tilde{T}_i} + \sum_{i \in R} e^{\tilde{T}_j} z_j^* \right) \\
&\quad - \sum_{i \in S} \gamma_i \left(\tilde{t} + \tilde{P}_i + \tilde{T}_i - \ln\left(T_{frame} B_i\right) \right) \\
&\quad - \sum_{i \in S} \eta_i \left(\ln\left(x_i T_{frame}\right) - \left(\tilde{T}_i + \ln\left(\sum_{j \in G_k \cap R,\, G_k \ni i} W \log\left(\frac{\alpha_{ij} e^{\tilde{P}_i}}{N_j} \right) z_j^* \right) \right) \right) \\
&\quad - \sum_{G_k \in G} \nu_k \left(\ln\left(\sum_{i \in G_k \cap S} x_i T_{frame} \right) - \ln\left(\sum_{j \in G_k \cap R} r_{jc} e^{\tilde{T}_j} z_j^* \right) \right) \\
&\quad - \sum_{i \in S} \omega_i \left(\tilde{P}_i - \tilde{P}_{\max} \right) \\
&\quad - \sum_{i \in S} \tau_i \left(\tilde{P}_{\min} - \tilde{P}_i \right)
\end{aligned}
$$

Then according to the KKT conditions

$$\tilde{t} + \tilde{P}_i + \tilde{T}_i - \ln\left(T_{frame} B_i\right) \leq 0, \ i \in S$$

$$\ln\left(x_i T_{frame}\right) \leq \tilde{T}_i + \ln\left(\sum_{j \in G_k \cap R} W \log\left(\frac{\alpha_{ij} e^{\tilde{P}_i}}{N_j}\right) z_j^*\right), \ i \in G_k \cap S, \ G_k \in G$$

$$\ln\left(\sum_{i \in G_k \cap S} x_i T_{frame}\right) \leq \ln\left(\sum_{j \in G_k \cap R} r_{jc} e^{\tilde{T}_j} z_j^*\right), \ G_k \in G$$

$$\tilde{P}_i - \tilde{P}_{\max} \leq 0, \ \tilde{P}_{\min} - \tilde{P}_i \leq 0, \ i \in S$$

$$\gamma \succeq 0, \ \eta \succeq 0, \ \nu \succeq 0, \ \omega \succeq 0, \ \tau \succeq 0$$

$$\gamma_i\left(\tilde{t} + \tilde{P}_i + \tilde{T}_i - \ln\left(T_{frame} B_i\right)\right) = 0, \ i \in S$$

$$\eta_i \ln\left(x_i T_{frame}\right) = \eta_i\left(\tilde{T}_i + \ln\left(\sum_{j \in G_k \cap R, \ G_k \ni i} W \log\left(\frac{\alpha_{ij} e^{\tilde{P}_i}}{N_j}\right) z_j^*\right)\right), \ i \in S$$

$$\nu_k\left(\ln\left(\sum_{i \in G_k \cap S} x_i T_{frame}\right) - \ln\left(\sum_{j \in G_k \cap R} r_{jc} e^{\tilde{T}_j} z_j^*\right)\right) = 0, \ G_k \in G$$

$$\omega_i\left(\tilde{P}_i - \tilde{P}_{\max}\right) = 0, \ i \in S$$

$$\tau_i\left(\tilde{P}_{\min} - \tilde{P}_i\right) = 0, \ i \in S$$

$$\frac{\partial L}{\partial \tilde{P}_i} = -\gamma_i + \eta_i \frac{\log e}{\log\left(\frac{\alpha_{ij}}{N_j}\right) + \tilde{P}_i \log e} - \omega_i + \tau_i, \ i \in S, \ j \in G_k \cap R, \ G_k \ni i, \ z_j^* = 1$$

$$\frac{\partial L}{\partial \tilde{T}_i} = -\lambda e^{\tilde{T}_i} - \gamma_i + \eta_i, \ i \in S$$

$$\frac{\partial L}{\partial \tilde{T}_j} = -\lambda e^{\tilde{T}_j} + \nu_k, \ z_j^* = 1, \ j \in G_k$$

the optimal solution $\gamma^*\left(\tilde{t}, \lambda\right)$ can be solved given $\tilde{P}^*\left(\tilde{t}, \lambda\right)$ and $\tilde{T}^*\left(\tilde{t}, \lambda\right)$, which are derived from solving the convex problem (14) directly.