UTS
UNIVERSITY OF TECHNOLOGY, SYDNEY

Enhanced Reliability of Wireless Body Area Network with abnormality detection and intelligent Medium Access Control

Faculty of Engineering and Information Technology
School of Electrical, Mechanical and Mechatronic Systems

PhD

Student: Jing Zhou

Supervisors:
Associate Professor Steven Su
Professor Hung Nguyen
Dr. Li Li

September, 2015

CERTIFICATE OF ORIGINAL AUTHORSHIP

I certify that the work in this thesis has not been submitted for a degree previously or submitted as part of requirements for a degree.

I also certify that the thesis is all written by me. Any help that I have received in my research work has been acknowledged. Finally, I certify that all reference sources and literature used in the thesis are indicated.

Signature of Student:

Date:

Acknowledgement

First of all, I would like to thank my supervisors Associate Professor Steven Su, Prof. Hung Nguyen and Dr Li Li for giving me this opportunity to pursue Doctor Degree at University of Technology, Sydney.

Secondly, I really appreciate Associate Professor Steven Su for his invaluable guidance, support, patience, encouragement, and life teachings. He always gives me guidance and assistance immediately when I encounter obstacles.

Then I am thankful to my friends and colleagues at FEIT of UTS. I am very appreciated and proud to work and study with all of you. Also, I have sincere thanks to all the staff in the UGS and FEIT at University of Technology, Sydney for their kindness and valued assistance. I also thank all the friends who have been very helpful with comments and suggestions.

Finally I would like to show my deepest gratitude to my family for their love, support and company.

Abstract

Wireless Body Area Network (WBAN) is an emerging technology which utilizes wireless sensors on, around or in a human body to exchange information between a person and a control centre, which is connected to remote stations for further analysis. WBAN will generate more reliable and efficient services in health care and sports training facilities with the developing technology in information and communication. But the reliability is still a challenge for widespread applications due to some constraints. In order to enhance the reliability some faults control mechanisms can be considered from two aspects: fault management and fault prevention. The former remedies the data after detecting abnormal events and the later reduces the possibility of fail packets in transmission. So in this thesis two approaches are proposed for high reliability and better performances of WBAN.

Firstly a comprehensive and effective method, which merges statistical analysis and intelligent computational model, is proposed to detect and identify abnormality in wireless communication. The principal component analysis (PCA) is utilized to acquire main features of the measured data and the K-means is combined to cluster various processes for abnormalities detection. And a back propagation (BP) neural network is constructed to identify and isolate faults.

In addition to detect abnormality for fault information, the scheme of fault prevention is generated by intelligent management of Medium Access Control by suppressing sources of faults with smart strategies. A centralized scheduling mechanism is generated with a fuzzy control mechanism based on input parameters for better performance. It controls the contention window in contention access period (CAP) and slots allocation in contention free period (CFP) respectively according to nodes' priorities and status. The improved performance of throughput, latency, and packets breakdown are acquired by efficient usage of bandwidth and avoidance of collision.

Furthermore this thesis proposes an intelligent management of multiple access schemes for contention access and contention free access simultaneously in MAC with game theory. An improved strategy is generated to manage various access schemes under different situations based on services and link states. Nodes in a network choose various strategies for best payoff according to different conditions including normal status, packets-oriented services, emergent situation and bad link states with specific utility functions. The novel management model supports flexible and efficient allocation of total resources and brings benefits for WBAN in different applications.

The effectiveness and feasibility of the proposed approaches are evaluated with experimental data, simulation and analysis, which illustrates the enhanced reliability in WBAN and can be used to promote the development of WBAN application.

Keywords: Wireless Body Area Networks; Reliability; Fault Detection and Identification; Medium Access Control; Fuzzy control; Game theory

Contents

List of Abbreviations

ARQ Automatic repeat request

ANN Artificial Neural Network

AWGN Additive white Gaussian noise

BAN Body area network

CAP Contention access phase

CFP Contention free access phase

CCA Clear channel assessment

CW Contention window

CPV Cumulative percentage variance

ECG Electrocardiography

EEG Electroencephalography

EMI Electromagnetic Interference

EMG Electromyography

GTS Guaranteed Time Slots

HR Heart rate

LAN Local Area Networks

MAC Medium Access Control

MAN Metropolitan Area Networks

MEMS Microelectromechanical Systems

PAN Personal Area Networks

PHY Physical Layer

QoS Quality of Service

SINR Signal to interference plus noise ratio

SPE Squared prediction error

SpO2 Oxygen

SVD Singular value decomposition

TDMA Time division multiple access

Vo2 Oxygen uptake

WBAN Wireless Body Area Network

WSN Wireless Sensor Network

List of Publications

Journal Papers

1. Jing Zhou, Aihuang Guo, Hung Nguyen and Steven Su, Intelligent management of multiple access schemes in Wireless Body Area Network，Journal of Networks, 2015.10(2) , p.108-116.
2. Mitchell Yuwonoa, Jing Zhou, Ying Guo, Branko G. Celler, Steven W. Su. Automatic bearing fault diagnosis using particle swarm clustering and Hidden Markov Model., Engineering Applications of Artificial Intelligence. 2015.
3. Jing Zhou, Aihuang Guo, Juan Xu, Steven Su ，An optimal Fuzzy Control Medium Access in Wireless Body Area Networks, Neurocomputing，142(2014), p.107-114.
4. Jing Zhou, Aihuang Guo, Branko Celler, Steven Su, Fault detection and identification spanning multiple processes by integrating PCA with neural network, Applied soft computing, 2014. 14(A), p.4-11.
5. XU Juan, WU Yurun, L(U) Binbin, ZHOU Jing, Asymptotic Bandwidth of Asynchronous Impulse Radio UWB System, Journal of Tongji University (Natural Science) 2013. 41（4）, p.591-595.
6. Jing Zhou, Steven Su, Aihuang Guo and Weidong Chen, Abnormalities Detection of IMU based on PCA in motion monitoring, Applied Mechanics and Materials, 2012. Vol.224, p.535-538.
7. Jing Zhou, Aihuang Guo, Research of Online-Learning in Higher Engineering Education，Industry and Information Technology Education, 2013（8），p. 83-86.
8. Jing Zhou, Aihuang Guo, Cultivate leadership in Higher Engineering Education，Computer Education，2013（5），p. 106-110.
9. Jing Zhou, Aihuang Guo, Implications of Higher Engineering Education in Australia，Computer Education，2012（21），p. 98-102.

Conference Papers

1. Jing Zhou ,Aihuang Guo, Juan Xu , Hung Nguyen, and Steven Su, A Game theory control scheme in medium access for Wireless Body Area Network, Proceedings of the 10th International Conference on Wireless communications, Networking and Mobile Computing 2014, p.404-409
2. Jing Zhou, Aihuang Guo, Juan Xu, Branko Celler, and Steven Su, A reliable medium access mechanism based on priorities for Wireless Body Sensor Networks，Conference Proceeding IEEE Eng Med Biol Soc. 2013; 1855-8. doi:10.1109/EMBC.2013.6609885.
3. Xujuan, Cheng Jiujun, Han Lina, Zhou Jing, Scaling Law of Ordinary Clustering Ultra-wide band Sensor Network, IWCMC 2013.
4. Jing Zhou, Steven Su ,Aihuang Guo, Fault detection and identification of COSMED K4b2 based on PCA and neural network, ICICSP 2012.
5. Jing Zhou, Fault tolerance in Wireless Body Sensor Networks. IEEE International Conference on Information Theory and Information Security 2011.p.590-594.
6. Jing Zhou, Security Analysis of Wireless Sensor Networks in Intelligent Transportation Systems. National Conference on Networks and Information Security 2011. p.131-134.
7. Jing Zhou, Aihuang Guo, Building a comprehensive Laboratory system in Higher Engineering Education，ISPR 2013.p.191-194.
8. Jing Zhou, Aihuang Guo, Bilingual Teaching in Telecommunication Curriculum. Proceedings, Higher Education Branch of China Education Society of Electronics.2011. p. 96-99.

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Overview of WBAN

The rising proportion of older people has significant increased not only in developed countries but also in developing countries. The worldwide aging population is projected to increase rapidly, almost doubling by 2020 and more than tripling by 2050 [1]. The global demographic trend creates evident consequences on economic, political and social processes. First and the most priority requirement is to promote the well-being and health with enabling environment [2]. In most countries health care spending is almost on the top of the political agenda. It will take 20–30 percentage of GDP in some economies by 2050 [3]. Simultaneously with this constantly increasing number of elderly people, health costs are also increasing demanding for novel technological solutions that enhance elderly independence, address health monitoring and lower healthcare bill. In Germany, the overall hospital expenditure will increase from about 24.7 billion Euros in 2000 to more than 46 billion Euros in 2041 [4]. Heart failure affects nearly five million people in the Unites States and it is a key factor of deaths every year[5]. How can we deliver better and less expensive healthcare services to this growing population of seniors? One possible solution could be to shift health monitoring from hospital systems to those places where elderly people live and move. This novel solution allows seniors to stay longer at home with good services.

Recent progresses in wireless networking, microelectronics and wireless sensors help to deal with the imminent problems and social challenges of a growing aging population. Advances in these technologies can make the pervasive wireless body area network into a reality and are likely to change every aspect of our daily life. In recent years, there has been increasing interests in researches investigating elderly health monitoring systems. Wireless Body Area Networks (WBAN) is an emerging technology which utilizes wireless sensors placed around, on or in a human body that are used to exchange information between a person and remote stations. WBAN can produce a sensing and communication platform to health monitoring, offer the chance to diagnose early symptoms of diseases and have the ability to monitor disease progression and patient response to treatment. WBAN also would allow doctors to observe patients during their daily life and it can be used in "closed feedback loop" to guide treatment delivery.

Ubiquitous healthcare promises continuous gathering and analysis of physiological, behavioural and other such health-related information and either acting on it, providing feedback, or delivering it to health professionals. The aim is to provide better healthcare for people via continuous monitoring of small autonomous wireless body-worn sensors, often outside of typical healthcare settings such as hospitals. If it becomes commonplace, the gathering of such health-related information needs to be unobtrusive and cheap, yet its data delivery must ensure certain reliability and timeliness guarantees. Wireless body area network (also can be called as Wireless body sensor network), aims to satisfy these requirements and set a standard for the development of a common model towards pervasive monitoring. A WBAN contains one or more small sensors (e.g., temperature, blood pressure) that communicate their readings back to a hub device on or near the individual (e.g., WBAN enabled smart phone), which may in turn communicate to others via another longer-range network (e.g., WiFi, 3G). These sensors send data to a personal device (e.g. a PDA PC or a smart phone) which acts as a sink, a coordinator, a personal sever or a gateway to health care.

WBAN will play a critical role in telemedicine-based health services in the future because it can bring lots of benefits in health caring including collecting medical data of people from their home or workplace with a cost-effective method, reducing the patients' mobility, decreasing the care expenses, improving the care quality with availability of efficient and immediately information. For example sometimes it is very difficult to acquire transient data for different diagnostic purposes and patients have to be taken observation at clinics or in the hospital, which is very resource intensive. The applying of WBAN could reduce the cost and increase the treatment because professionals can obtain data from anywhere and anytime.

The recent development of intelligent sensors improves the application of WBANs that can be worn on or implanted in the human body to monitor health status. These systems reduce the enormous costs of patients in hospitals as monitoring can occur real-time and over a longer period of time, even at home. WBANs, for example, provide novel uses in healthcare, fitness, and entertainment. WBANs nodes must be extremely non-invasive, and WBANs must have fewer and smaller nodes relative to a conventional WSN. Although WBAN comes from wireless sensor network (WSN), it is a special type with specific features and QoS requirements. The major different characteristics between WBAN and WSN are listed in table 1.1. To support the healthcare needs of people, the WBAN should transmit both routine vital signs and alerting signals when vital signs cross one or more individualized thresholds. Usually WBAN needs low packet transmission delay, almost no packet loss, low energy usage, which depend on many factors such as hardware design, application requirements and network structure and protocols. A health monitoring sensor obtains, samples, and digitizes the vital signs for transmission as network packets [6].In addition to vital signs, the device must also monitor skin breakdowns, gait and balance (to reduce the number of falls), motor activity, agitation, patient location, cigarette smoke, and the amount of moisture in the patient's clothes. Monitoring many physiological signals from a large number of patients at the same time is one of the current needs in order to deploy a complete wireless sensor network system in medical centres.

The comprehensive WBANs could cover scenarios involving patients with different levels of mobility, locations, and timeliness of monitoring. In passive monitoring, the vital signs are recorded for a subsequent analysis by healthcare professionals, while active monitoring involves generation, transmission and analysis of live vital signs and related information. These can be further divided in continuous or event-driven, based on whether the monitoring is continuous or event-driven such as passage of time or patient intervention.

1.2 Terms and Cloud architecture

1.2.1 Definition of Terms

Some terms related to this research are outline as follows:

Sensing: Sensing is a technique used to gather information about a physical object or process, including the occurrence of events (i.e., changes in state such as a drop in temperature or pressure)[7]. An example of the steps performed in a sensing (or data acquisition) task is shown in Figure 1.1.

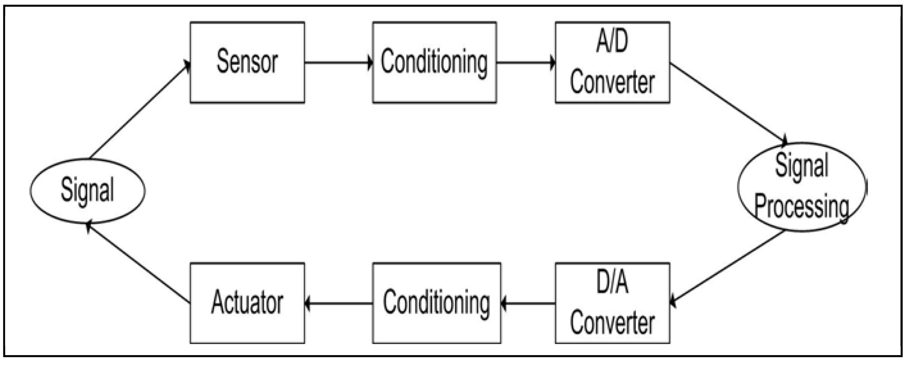


Figure 1.1 Data Acquisition and actuation

Sensor: There are no uniform descriptions of sensors or the process of sensing. In many cases, the definitions available are driven by application perspectives. Taking a general perspective, a sensor can be defined as: Senor is a device that receives a stimulus and responds with an electrical signal[8]. When considering the observational elements of measurements, a sensor can be described as: a device that converts a physical measure into a signal that is read by an observer or by an instrument[9].

Early sensors were simple devices, which measured a quantity of interest and produced some forms of mechanical, electrical, or optical output signals. In the last decade computing, pervasive communications, connectivity to the Web, mobile smart devices, and cloud integration have been added immensely to the capabilities of sensors. Table 1-1 shows various types of sensors.

Biosensors: Biosensors use biochemical mechanisms to identify analyse in chemical, environmental (air, soil, and water), and biological samples (blood, saliva, and urine). The sensor uses an immobilized biological material, which could be an enzyme, antibody, nucleic acid, or hormone, in a self-contained device. The biological material used in the biosensor device is immobilized in a manner that maintains its bioactivity. The immobilization process results in contacts between the immobilized biological material and the transducer. When an analyse comes into contact with the immobilized biological material, the transducer produces a measurable output, such as a current, change in mass, or a change in colour[10].

Table 1. 1 Types of Sensors

Type	Examples
Temperature	Thermistors, thermocouples
Pressure	Pressure gauges, barometers, ionization gauges
Optical	Photodiodes, phototransistors, infrared sensors, CCD sensors
Acoustic	Piezoelectric resonators, microphones
Mechanical	Strain gauges, capacitive diaphragms
Motion, Vibration	Accelerometers, gyroscopes, photo sensors
Flow	Anemometers, mass air flow sensors
Position	GPS, ultrasound-based sensors
Electromagnetic	Hall-effect sensors, magnetometers
Humidity	Hygrometers, MEMS-based humidity sensors

Biosensors monitor heart rate, oxygen level, blood flow, respiratory rate, muscle activities, movement patterns, body inclination, oxygen uptake (VO2) and so on. There are some examples of the commercially available wireless biosensors for health monitoring:

- Pulse oxygen saturation sensors: They measure the percentage of hemoglobin (Hb) saturated with oxygen (SpO2) and heart rate (HR);

- Blood pressure sensors: non-invasive devices designed to measure human blood pressure;
- Electroencephalography (EEG): Measure voltage fluctuations resulting from ionic current within the neurons of the brain.
- Electrocardiogram (ECG): Record the electrical activity of the heart;
- Electromyogram (EMG) : Measure muscle activities;
- Temperature sensors : Measure core body temperature and skin temperature;
- Respiration sensors: Monitor respiration rate, waveform and amplitude;
- Blood oxygen level sensor: Measure cardiovascular exertion.

Wireless Sensor Network: Wireless sensor networks (WSNs) are a subset of sensor networks, and normally a sink node, which is usually called a "base station," and a number of wireless, battery-powered sensor nodes. The base station typically has significantly higher processing and data storage capabilities than the other nodes on the network[11].A number of sensors communicate the collected data wirelessly to a centralized processing station. Therefore, a wireless sensor has not only a sensing component, but also on-board processing, communication, and storage capabilities. With these enhancements, a sensor node is often not only responsible for data collection, but also for in-network analysis, correlation, and fusion of its own sensor data and data from other sensor nodes.

Wireless Body Area Network: When Wireless Sensor Network is deployed on human body it is called wireless Body Area Network (WBAN), which has more challenging than other type of WSNs and extensive research is focusing on biocompatibility, reliability, signal propagation, power management and other properties. A WBAN contains one or more small sensors placed on or inside the human body that communicate their readings to a hub device on or near the individual.

1.2.2 Sensors and the Cloud

Cloud computing has become one of the most active areas in information technology. It already begins to transform the businesses manage and use their computing and storage infrastructure. The cloud-based model provides flexibility and scalability for computing, storage, and application resources, optimal utilization of infrastructure, and reduced costs. Cloud computing has the potential to provide storage, processing, and visualization capabilities for sensors and sensor networks, as shown in Figure 1.2. The sensors can be discrete or part of a geographically distributed network, and feature highly dynamic data throughputs. This cloud-based integration is able to accommodate dynamic loads and sharing of sensor resources by different users and applications, in flexible usage scenarios.

A cloud-based approach has the potential to reduce overall cost in a sensor deployment, as it can easily support elastic consumption of resources. Clouds are typically based on usage models, allowing application developers to optimize frequency and resolution of data against cost. Vast amounts of sensor data can be processed, analysed, and stored by computational and storage resources of the cloud. Once the sensor data is stored in the cloud, new applications based on aggregated data sets can be created from different organizational resources or from crowd sourcing. Essentially, the cloud can be used to create virtualized sensors which could be accessed by any customers or applications. This will break down the siloing that currently exists in many sensor applications.

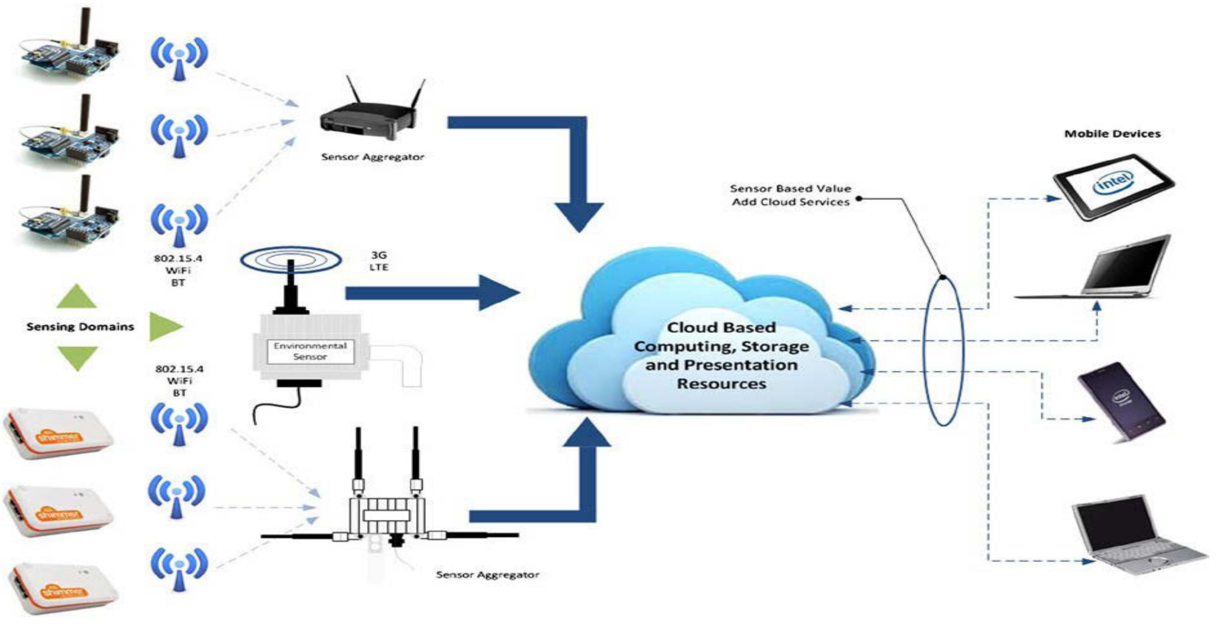


Figure1.2 A sensor cloud architecture

Users and applications do not need to concern about the physical location of sensing resources when using cloud-based functions, because they are essentially virtualized. End users can dynamically provision and group virtual sensors based on need, and terminate them easily when no longer required. New physical sensors can usually be added to the cloud through a registration process.

1.3 WBAN application

WBAN for healthcare application is showed in figure 1.3. Currently the demanding on healthcare is continually growing. The longer is the human being's life, the higher is the occurrence of age-associated chronic diseases. Scientific development will allow some terminal conditions such as a cancer to be transferred and treated as chronic diseases. Therefore, it is urgent to invent a new way to provide care while reducing the already high healthcare costs, which will bring disaster in most countries.

Modern lifestyles are seen to play an important role in people's health. Some "lifestyle diseases" are becoming the most significant causes of death. There are many lifestyle factors that affect diseases such as obesity, stress, high fat/sugar/salt/processed meal diets, cigarette smoking, and high levels of alcohol consumption. People try to find new technologies, mainly in the activity monitoring, to provide more awareness of our sedentary lifestyles. Sensing technologies and supporting software could be applied to push people to build habits that help them have better health in later life. A lot of new products, including Nike Fuel and Fitbit, already direct to make people more active and provide mechanisms to keep them engaged in fitness over the longer term.

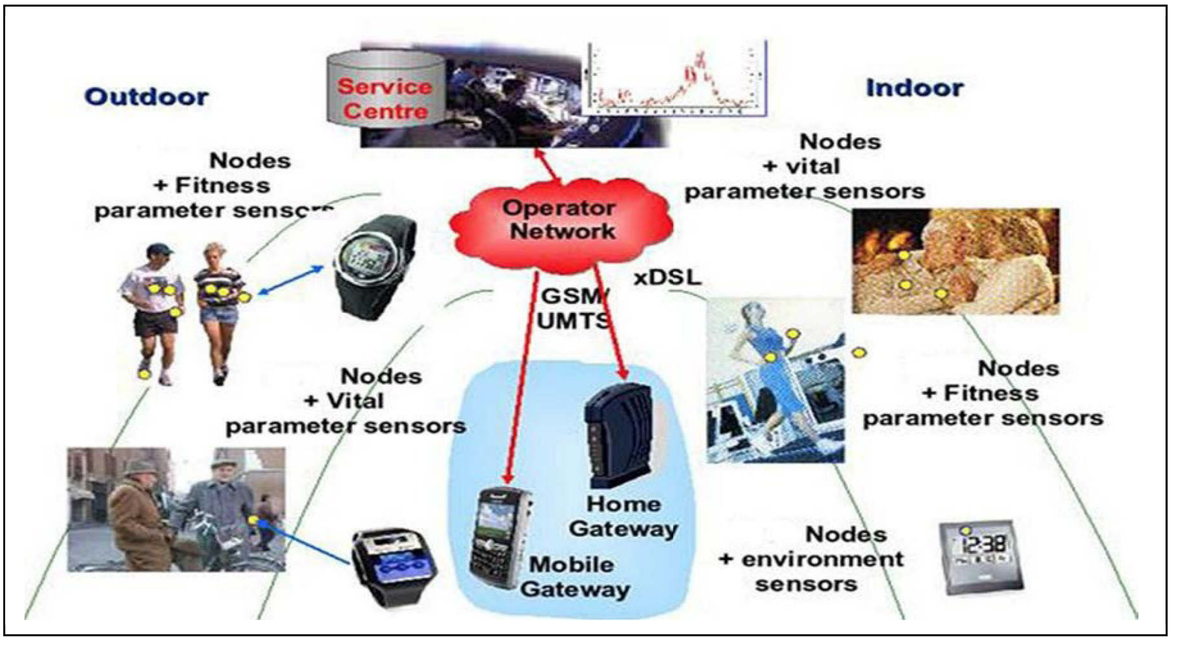


Figure 1. 3 WBAN health application

Occurrence of disease, whether it is associated with lifestyle-related or aging, invokes us to change the healthcare model in the future. Modifying the healthcare services model will need multiple approaches and require adjustment from the following aspects:

- Using remote assessment technology at home and community center to reduce hospitals' pressure;

- Transferring from a reactive model to a preventative model;
- Training healthcare to individuals such as: risk factor identification, preventative intervention, and treatment;
- Involving individuals to a greater extent in monitoring and maintaining their health and wellness.
- Using technology to enable better management of clinical workloads and to allow health professionals to prioritize the patients of greatest need.

1.3.1 Sensing for health Monitoring and Diagnosis in hospital and community

Sensing is pervasive everywhere in hospital-based care, from a digital thermometer to complicated laser-guided surgical devices. Many imaging equipment, such as x-ray, magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), and ultrasound provide doctors with non-invasive insight into the human body. Generally these images allow doctors to pinpoint areas of injury or abnormality, perform minimally invasive surgery, and evaluate the success or failure of a procedure. These large, non-discrete sensors require careful sample preparation by skilled professionals to ensure optimum results. Sensors also play an important role in treatment technologies. They can sense events, such as missed heartbeats, that can be acted upon by clinicians or actuators. They can optimize drug delivery devices by identifying the optimum time to administer a drug. And they can continuously track a patient's vital signs to ensure that a treatment is delivered safely. Imaging, implantable, and drug-delivery devices are all rich sources of sensor application. .

1.3.2 Contextual information in Health Applications

Context is critical in determining the value of sensor data. For example, some isolated sensor readings may have limitations, which can partly be addressed by related contextual information. For instance: What was a person doing during the measurement? Where was the person? What was the local environment that could affect the measurement? Context is very important when acquiring physiological measurements, such as activities to heart rate. Usually Context can be captured with additional sensors. For example, accelerometers are used to determine a person's movement when taking a physiological measurement.

Contextual sensing information can generally be used in three ways:

- The most common one is for a clinician to manually review the contextual sensor information to support the interpretation process.

- Alternatively, contextual information can be overlaid with the sensor measurements of interest in the same graphical representation or in an adjacently chart. This overlapping of data visually can be very helpful in the interpretative process.
- Finally, the most sophisticated approach is the intelligent and automated fusion of contextual data with the data source of interest. Algorithmic data fusion can be very useful in reducing the dimensionality of data and inferring higher-level information. It can also be useful to determine whether the context of measurement is valid based on a defined measurement process.

Sensors can also be used to provide context to the success or failure of a treatment process. By knowing what happens between outpatient visits, a clinician can optimize an existing treatment protocol or develop a new protocol. Data acquired from ambient and body-worn sensors can also be combined with context-detection algorithms to generate support messages. Such messages are generated at appropriate times to support a patient on a specific treatment process, for example to prompt taking medication, remind about exercise, or advise about the consumption of food and drink.

1.3.3 Vital Signals in healthcare

Hospital care and treatment mostly depends on many vital signals. A patient's vital signals represent the status of the main body functions—typically body temperature, heart rate, blood pressure, and respiratory rate. But other measurements may be included, depending on the context. For example, in a first-aid situation, skin and level of consciousness are also examined. In intensive-care or post-operative units, blood pressure, heart rate, blood oxygenation, and many other variables are continuously monitored using sensing technology. In lower-dependency units, these variables are intermittently monitored by nursing staff, who manually measure these variables using portable monitors. Disposable, wearable vital sign sensors are beginning to emerge that permit personal, low-cost, continuous monitoring of vital signs for all patients, regardless of health status or location.

Similarly, sensor technologies that target both elite athletes and ordinary individuals, such as body-worn pulse-rate monitors or smart clothing with integrated sensing capabilities, are also emerging to determine performance and fitness levels.

Heart Rate: Heart rate refers to the speed of the heartbeat. It is typically measured in beats per minute. Heart rate is measured to detect slow heart rate, fast heart rate or irregular heart rate and rhythm, any of which can indicate illness. Like many vital signs, heart rate is age-dependent: an infant's heart rate is fine between130 and 150 beats per minute, whereas an adult's should fall between 50 and 80 beats per minute. Heart rate is also highly dependent on context: it is raised along with exercise and stress, and the resting heart rate of an endurance athlete is much lower than a non-athlete. The non-technical method to measure heart rate is to feel artery pulsations at a pulse point using the index and middle fingers. The number of pulsations is counted to find the heart rate. In a hospital setting, heart rate is measured continuously using an electrocardiograph (called an ECG or an EKG).An ECG measures the electrical activity of the heart, using electrodes attached to the surface of the skin, filtering circuitry, and a data logger. The heart rate can be determined by measuring the interval between one R-wave and the next R-wave of the ECG signal, called the R-R interval. Variability in heart rate can also be predictive of many issues, including congestive heart failure. The timing between different points of an ECG signal can indicate a number of conditions, including hypocalcemia or coronary ischemia. In a hospital setting, 12-lead ECGs are used for diagnostic purposes, and 5-lead and 3-lead ECGs are used for continuous monitoring. The number of leads describes the number of electrodes attached to the body. Each electrode is connected with cables to a data-filtering and logging circuit. Expertise is required to attach the ECG electrodes correctly and interpret the resulting ECG strip.

Fitness and in-home heart-rate sensors are beginning to move toward integrated wireless ECG devices, which simplify the placement of electrodes and interpretation of data. These devices are used by obstetric and community clinicians, but are increasingly being sold for personal use. Heart rate can also be measured using pulse oximeters or body vibrations (seismo cardiography).

Blood Pressure: Blood pressure is the pressure exerted by the blood on the walls of the large arteries, such as the brachial artery in the arm. High blood pressure is a risk factor for stroke, heart attack and chronic kidney failure; therefore, it is essential to not only diagnose it but also to continuously monitor the impact of treatment. Low pressure can be problematic if it results in fainting or dizziness. Blood pressure is typically described as a systolic value over a diastolic value, and is measured in millimetres of mercury (mm Hg). Systolic pressure is the peak pressure in the arteries during the cardiac cycle and the diastolic pressure is the lowest pressure at the resting phase of the cardiac cycle.

The blood pressure of a resting, healthy adult human is approximately 120 mm Hg systolic and 80 mm Hgdiastolic (written as 120/80 mm Hg). Arterial blood pressure (BP) is the most accurate method to measure BP. This invasive method involves placing a cannula into a blood vessel and connecting it to an electronic pressure transducer.

Non-invasive methods are simpler and quicker and require less expertise, but are slightly less accurate. These methods measure the pressure of an inflated cuff at the points when it just occludes blood flow (systolic pressure), and again when it just permits unrestricted flow (diastolic pressure). There are three non-invasive methods commonly used for routine examinations and monitoring blood pressure:

- The common method requires a clinician to manually compress the artery in the upper arm using an inflatable cuff. The clinician listens to the artery using a stethoscope to recognize when the blood just begins to flow back in the artery (systolic blood pressure) and when no sound can be heard (diastolic blood pressure). The blood pressure at both of these points can be read from the mercury or aneroid manometer that is connected to the cuff. This method is considered the gold standard by many, despite its high reliance on human hearing and interpretation.
- Oscillometric methods are used in long-term measurement, home measurement, and sometimes in general practice. The equipment is functionally the same as for the auscultatory method, but with an electronic pressure sensor (transducer) fitted in the cuff to detect blood flow, instead of using the stethoscope and the expert's ear. These devices use a method, called mean arterial pressure (MAP) to calculate blood pressure. The accuracy of the algorithms can vary greatly among different devices. It is therefore essential to confirm worrying home readings with a clinician, who will have a more accurate BP device.
- Continuous noninvasive arterial pressure (CNAP) is used in research, critical care, and anesthesia to understand blood pressure at a more granular level than can be achieved using auscultatory or oscillometric methods. There are three common methods to do this: arterialtonometry is a technique for measuring blood pressure in which an array of pressure sensors is pressed against the skin over an artery. The second method, pulse transit time (PTT) is the time it takes a pulse wave to travel between two arterial sites. Blood pressure can be determined from the inverse of the PTT. In the final method, which involves clamping the finger, blood volume in the finger is measured using a light transmitter and receiver. The pressure of a finger cuff is adjusted to maintain constant blood volume in the finger. This pressure corresponds to the patient's blood pressure. CNAP methods were traditionally expensive and limited to hospital settings, but community-based PTT devices are currently available, and home-based devices, such as the Scandu Scout, will soon be available.

While annual blood pressure measurements in a doctor's office may be appropriate for healthy young adults with no cardiovascular risk factors, there are many situations in which blood pressure requires more detailed or long-term investigation. Older adults with impaired blood pressure regulation may experience sudden drops in blood pressure when they go from sitting to standing.

This is a known risk factor for falls, and can be diagnosed only using beat-to-beat CNAP techniques in a clinical setting. There are also situations in which the blood pressure reading in a doctor's office is not representative of the patient's actual blood pressure. "White coat hypertension" can cause a rise in some patients' blood pressure while they are in a clinical setting. This is a very common occurrence and may lead to clinician's prescribing blood pressure medicine to those who do not need it. Circadian variations in blood pressure can also result in incorrect prescription of medication. Some patients, who have high or normal blood pressure during the day, may have low blood pressure at night. Prescription of blood pressure-reducing medication to manage the high blood pressure measured in the doctor's office during the day could drop the patient's blood pressure to dangerously low levels overnight. Ambulatory blood pressure monitoring, using a wearable oscillatory device for a 24-hour period, is the only way to identify circadian blood pressure rhythms. As form-factor and interpretation technology evolve, it is expected that monitoring this valuable vital sign outside of a hospital setting will become increasingly smaller, cheaper, and more accurate.

Temperature: The body employs several strategies to maintain body temperature. When the body is too hot, the blood vessels in the skin dilate (expand) to carry the excess heat to the skin's surface, where it is removed through perspiration. When the body is too cold, it conserves body heat by narrowing blood vessels to reduce blood flow to the skin and generates heat by shivering. An abnormally low (hypothermia) or high (hyperthermia) body temperature can be serious and even life threatening. Temperature sensors are therefore an essential, low-cost, reliable method to indicate health status. Body temperature can be measured in many locations on the body, including the mouth, ear, armpit, rectum, forehead, bladder, skin, and esophagus. The traditional method to measure temperature, the mercury thermometer, has been replaced by contact and non-contact sensors. Contact temperature sensors reach thermal equilibrium with any objects they are in contact with. They can measure the temperature of that object by measuring their own temperature. Non-contact temperature sensors measure radiated heat from the measured brightness or spectral radiance of an object.

Respiratory Rate: Respiratory rate is the amount of taken breaths in one minute. Just as the heart rate, the normal respiratory rate relies on some factors, such as age, emotional state and sleeping. Tachypnea is abnormally high respiratory rate and can be generated by something as simple as exercise, or as serious as carbon monoxide poisoning. Bradypnea is abnormally low respiratory rate and can indicate problems such as heart tissue damage or high blood pressure. Respiratory rate is very commonly applied to diagnose a condition named as sleep apnea, which means pauses, shallow or in frequent breathing in sleeping. This condition affects not only adults but also children and can cause daytime sleepiness, while attention and memory issues. Sleep apnea can be diagnosed in an overnight sleep study in a hospital lab, where respiration, brainwaves, eye and muscle movement are monitored. The non-technical method to measure respiration rate is counting the frequency of chest rises and falls in a fixed time. This method is easy to error. Chest straps and smart clothing can measure respiration by measuring changes in tension on the fabric as the chest expands and contracts. Contact and non-contact acoustic and optical methods have also been developed to measure respiration. Pressure-sensitive mattresses have been used to measure the respiration rate of sleeping babies or at-risk adults. These sensors detect respiration by measuring changes in pressure on the mattress as the monitored person inhales and exhales. In a hospital setting, respiration rate can be derived from an existing ECG signal or from the electrodes used to measure ECG.

Blood Oxygenation: A body needs a certain level of oxygen in the blood circulating to cells and tissues for correct working. When an oxygen level decreases below a certain point, a person may have shortness of breath. The amount of oxygen travelling in an artery can be measured by testing a sample of blood from an artery. It can also be estimated non-invasively using a pulse oximeter, a small device that clips on the finger, earlobe or across a baby's foot. Pulse oximetry is a non-invasive method for monitoring a patient's O2 saturation. It is used throughout the healthcare domain, particularly for assessing patients with respiratory complaints or in respiratory-related procedures. Normal pulse oximeter readings range from 95 to 100 percent.

Transmissive pulse oximeters measure the saturation of oxygen in the blood, which is a proxy measure of blood oxygen levels. They operate by passing light of two wavelengths from one side of the clip, through the patient, to a photo detector on the other side of the clip. The pulsing arterial flow can be determined by measuring the changing absorbance at each of the wavelengths. This method can be used to measure both oxygenated and deoxygenated hemoglobin at the peripherals. Reflectance pulse oximetry can be used on the feet, forehead, and chest. In this method, the detector lies adjacent to the light source on a flat surface such as the forehead. A number of situations can cause an erroneous SpO2 reading, especially with the use of transmission probes. These include skin pigment, nail polish, motion, ambient or excessive light, hypoperfusion, cardiac arrest, hypoxia, malposition of the probe, and intravenous dyes. These variables must be considered and blood tests should be used to confirm hypoxia if an erroneous pulse oxygenation reading is suspected.

1.3.4 Mobility and Behavioral Monitoring

The ubiquitous sensing is based on distributed and networked sensors that monitor users' activities while remaining transparent to the users. For behavioural monitoring, wireless sensors and other methods such as RFID are used to detect interactions between humans and their environment. Behavioural monitoring has attracted attention due to the potential, which provides clinicians and care givers with an impartial view of patients in their home environments. With behavioural information a clinician can decide if an individual is able to live independently by tracking the ability to complete all activities of a daily life (ADLs). This data could also be useful to detect early signs of diseases, such as dementia or diabetes, by changes in behaviour patterns.

Reliable and accurate behavioural information will support family doctors or public health professionals to ask questions informed about the patient's recent history, thus optimizing patient visits. Sensors are also useful for supporting technologies. For example, PIR sensors used in a kitchen environment can detect when the person is there; readers and RFID tags can recognize what items they are using. The data can be deployed to actuate indications (audio, video, or visual) to help users during every process in a task like preparing a meal.

Another approach to behavioural monitoring that does not require the installation of application-specific sensors is based around the use of smart meter data. The data can be used to identify patterns such as cooking, sleeping, when the home is unoccupied, and so on. The detecting of these events has the capability for smart meters to act as basic health monitors, particularly for older adults. As smart meters become more ubiquitous over the next five years there is significant possibility to explore the data for widespread behavioural monitoring applications.

1.4 Wireless Communication Standards and Medium Access Control

1.4.1 Wireless Communication Standards

In most networks, multiple nodes share a communication medium for transmitting their data packets. The IEEE 802 Standard comprises a family of networking standards that cover the physical layer specifications of technologies from Ethernet to wireless. For wireless communication the specifications include 802.11 for Wi-Fi, 802.15 for Wireless Personal Area Networks, and 802.16 for Wireless Metropolitan Area Networks and 802.20 for Mobile Broadband Wireless Access as shown in figure 1.4.

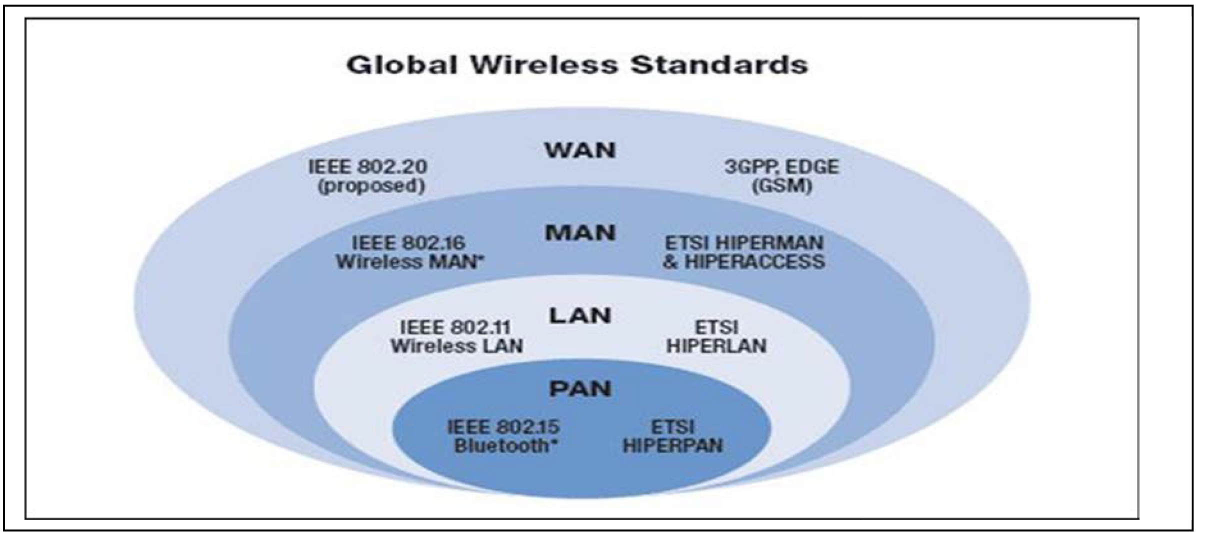


Figure1.4 Wireless standards

1.4.2 Medium Access Control

The wireless medium must be shared by multiple network devices, therefore a mechanism is required to control access to the medium. This responsibility is carried out by the second layer of the OSI reference model (Figure 1.5), called the data link layer. According to the IEEE 802 reference model (also shown in Figure 1.5), this layer is further divided into the logical link control layer and the medium access control layer. The MAC layer operates directly on top of the physical layer, thereby assuming full control over the medium. (MAC) protocol is primarily responsible for regulating access to the common medium. The choice of MAC protocol has a direct bearing on the reliability and efficiency of network transmissions due to these errors and interferences in wireless communications and to other challenges such as the hidden-terminal and exposed-terminal problems. The main function of the MAC layer is to decide when a node accesses a shared medium and to resolve any potential conflicts between competing nodes. It is also responsible for correcting communication errors occurring at the physical layer and performing other activities such as framing, addressing, and flow control.

Existing MAC protocols can be categorized by the way they control access to the medium. Figure 1.6 shows an example of such a categorization. Most MAC protocols fall either into the categories of contention-free or contention-based protocols. In the first category, MAC protocols provide a medium sharing approach that ensures that only one device accesses the wireless medium at any given time. This category can further be divided into fixed and dynamic assignment classes, indicating whether the slot reservations are fixed or on demand.

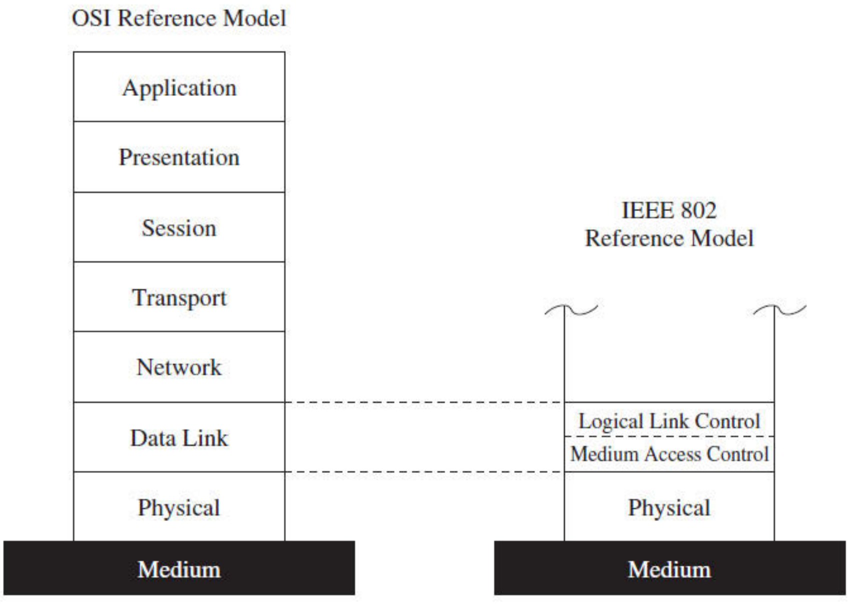


Figure 1.5 OSI reference model and the IEEE 802 reference model

In contrast to contention-free techniques, contention-based protocols allow nodes to access the medium simultaneously, but provide mechanisms to reduce the number of collisions and to recover from such collisions. Finally, some MAC protocols do not easily fit into this classification since they share characteristics of both contention-free and contention-based techniques.

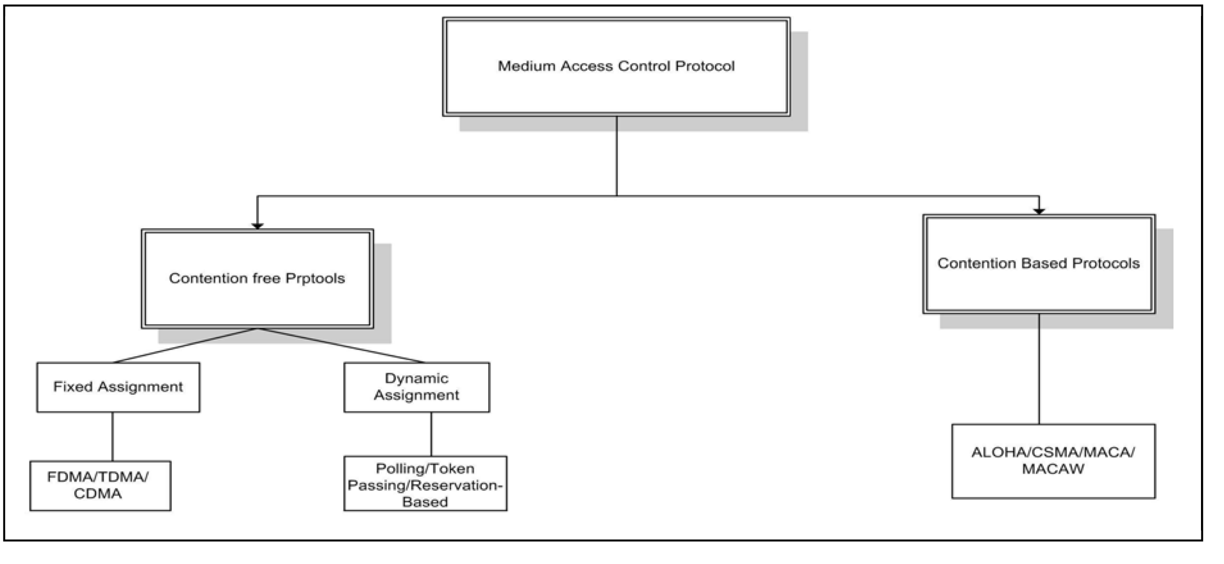


Figure 1.6 Classification of Medium Access Protocols

Collisions can be avoided by allocating resources to nodes such that each node can use its resources exclusively. For example, the frequency division multiple access (FDMA) protocol is one of the oldest methods of sharing a communication medium. In FDMA, the frequency band is divided into several smaller frequency bands, which can be used for data transfer between a pair of nodes, while all other nodes that could potentially interfere with this transfer use a different frequency band. Similarly, the time division multiple access (TDMA) protocol allows multiple devices to use the same frequency band, but it uses periodic time windows (called frames), consisting of a fixed number of transmission slots, to separate the medium accesses of different devices. A time schedule indicates which node may transmit data during a certain slot, that is, each slot is assigned to at most one node.

The main advantage of TDMA is to avoid collisions with the mechanism that nodes do not have to contend to access the medium. The disadvantage of TDMA is that when the network topology is changed and the slot allocations are also changed. Further, TDMA protocols can be inefficient in their bandwidth utilization when slots are of fixed size with different packet size and when allocated slots to a node are not used in each iteration. A third class of MAC protocols is based on the concept of code division multiple access (CDMA), where simultaneous accesses of the wireless medium are supported with different codes. If these codes are orthogonal, it is possible for multiple communications to share the same frequency band, where forward error correction (FEC) at the receiver is used to recover from interferences among these simultaneous communications.

Strategies of fixed assignment cannot be very efficient because it is typically impossible to reallocate slots belonging to one device to other devices when slots are not needed for the designated node in each frame. Also, creating schedules for a whole network could be a tough task and the schedules may need modifications each time when the network topology or traffic characteristics change. Therefore, dynamic assignment strategies can avoid rigid allocations with allowing nodes to access the medium when they need. For example, in polling-based protocols, a controller device (for example a base station in an infrastructure-based wireless network), sends small polling frames and asks each node if it has data to send. If a node doesn't have data to be sent, the controller polls the next station. A modified approach is token passing, in which nodes pass a polling request to each other with a special frame named as a token. A node can transmit data only when it gets the token. In addition, reserved assignment protocols use static time slots to allow nodes to reserve future access to the medium based on demand. For example, a node can show the requirement of data transmission by toggling a reservation bit in a fixed location. These complicate protocols guarantee other possible conflicting nodes can take such a reservation to avoid collisions.

Contention-Based Medium Access

Compared with contention-free techniques, contention-based protocols allow nodes to access the medium simultaneously with the possibility of contention, but they use mechanisms to reduce the possibility of collisions and to recover from collisions. The ALOHA protocol is an example which uses acknowledgments to verify the success of data transmission. Nodes with ALOHA can access the medium immediately, but collisions are addressed with an exponential backoff to enhance the likelihood of successful transmissions.

The slotted-ALOHA protocol tries to reduce the number of collisions by setting that a station may start transmission only at predefined time (the beginning of a time slot). Although the slotted-ALOHA improves the efficiency of ALOHA, it also needs synchronization among nodes.

The Carrier Sense Multiple Access (CSMA) approach is a typical contention-based MAC scheme and it can be divided as Collision Detection (CSMA/CD) and Collision Avoidance (CSMA/CA). In CSMA/CD-based schemes, the sender firstly senses the medium to judge whether it is free or busy. If the medium is busy, the sender does not transmit packets. If the medium is idle, the sender can initiate data transmission. In wired systems, the sender continues to listen to the medium to detect collisions. But in wireless systems, collisions occur at the receiver, and the sender will not know a collision. The hidden-terminal problem is another concern. For example two sender devices A and C are able to reach a receiver device B, but cannot overhear each other's signals. Therefore it is possible for A and C to transmit data to B at the same time, and then cause a collision at B, without directly detecting this collision. A related problem is the exposed-terminal problem, where C wants to transmit data to a fourth node D, but decides to wait because it overhears an ongoing transmission from B to A. However, B's transmission will not interfere with data reception at D because D is outside the transmission range of B. As a consequence, node C's waiting delays its transmission unnecessarily.

1.5 Conclusions and contributions of this thesis

With wireless sensors WBAN creates invisible interconnections of the physical world to measure, monitor, and manage data from body or environment. The system operates a process of data acquisition, distributed processing, wireless communication data storage with high reliability, failure-tolerant security and easily encrypted privacy. WBAN has enormous potential to change our lives into a more convenient and efficient status. But the reliability is a critical challenge to promote the wide application of WBAN. The aim of this thesis is to propose effective schemes to enhance the system reliability from two aspects: fault detection with effective mathematic methods and fault prevention with protocol design and control.

The main contributions of this thesis contain:

1. Propose an automatic abnormality detection methodology with few features of PCA for IMU, which is placed on body and used for falling detection in the experiment.
2. Propose a multiprocessing structure of fault detection, identification and reconstruction with

comprehensive processes based on COSMED K4b^2 sampling data for exercises analysis.

3. Initiate a novel reliable medium access scheme for WBAN based on priority to improve the possibility of successful data transmission.
4. Launch an adaptive medium access control in CSMA/CA and GTS respectively with fuzzy control to avoid collisions and improve the performance.
5. In order to intelligently manage the medium resources game theory is applied and utility functions are defined to acquire best payoff, which means better implementation under various conditions.

1.6 Outline of the Thesis

This thesis is organized as follows. Chapter 1 introduces the research background and basic concepts. Chapter 2 explains a typical structure, main requirements, constraints and reliability challenge of WBAN and illustrates related research. Chapter 3 formally defines the fault diagnosis problem and analyses the theory of principle components of analysis for fault detection and applies the methodology to abnormalities detection of IMU. Chapter 4 demonstrates the approach of intelligent fault detection, identification and reconstruction spanning multiple processes for COSMED K4b2 sampling data. Chapter 5 introduces a mechanism for fault prevention with medium access control based on priority, which is efficient to improve the possibility of successful packets in transmission. A short introduction of general IEEE personal area network protocol is outlined in this chapter. Chapter 6 proposes a fuzzy control medium access scheme respectively in contention access and contention free access. Chapter 7 designs an intelligent methodology for medium resource management based on proposed utility functions to acquire best payoff. The game theory for completion is illustrated and various operation conditions are analysed for better performance. Finally, the future development of WBAN, conclusions and the further research of reliability is addressed in the last chapter.

Chapter 2 Wireless Body Area Network Structure and Reliability

2.1 A system structure of WBAN

WBAN is specifically defined to develop a common model for pervasive monitoring with implantable and wearable body sensors. A diagram of a typical architecture of WBAN is illustrated in Figure 2.1. A number of sensors are attached on, in or around a patient's body and each sensor is connected to a personal server with wireless transmitter which is seamlessly integrated with home or hospital environment.

Each node accurately captures data, carries out low level processing of the data, and then wirelessly transmits this information to a Local Processing Unit. The data from all the sensors are collected by the LPU, further processed, and fused before being transmitted to a central monitoring server either via a wireless LAN, Bluetooth, or mobile phone (GPRS or 3G) network.

The system can be divided into three levels. In the first level each user wears a number of sensor nodes that are strategically placed on her body. These intelligent nodes are cable of sensing, sampling, processing, and communication of physiological signals. For example, an ECG sensor can be used for monitoring heart activity, an EMG sensor for monitoring muscle activity, an EEG sensor for monitoring brain electrical activity, a blood pressure sensor for monitoring blood pressure, and a breathing sensor for monitoring respiration. These sensor nodes unobtrusively collect vital signals and transfer the relevant data to a person server (PS), such as PC, PAD or mobile phone, through wireless network.

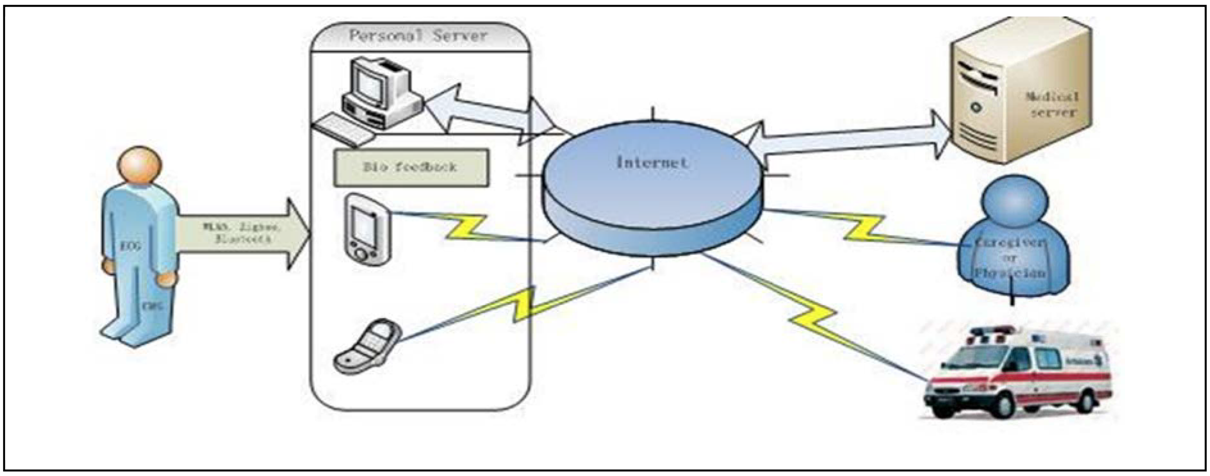


Figure 2.1 A basic Structure of wireless body area network

A lot of topologies are defined in a communication network to describe the geometrical position of nodes such as bus, star, ring, mesh, tree and so on. These topologies are showed in figure 2.2.

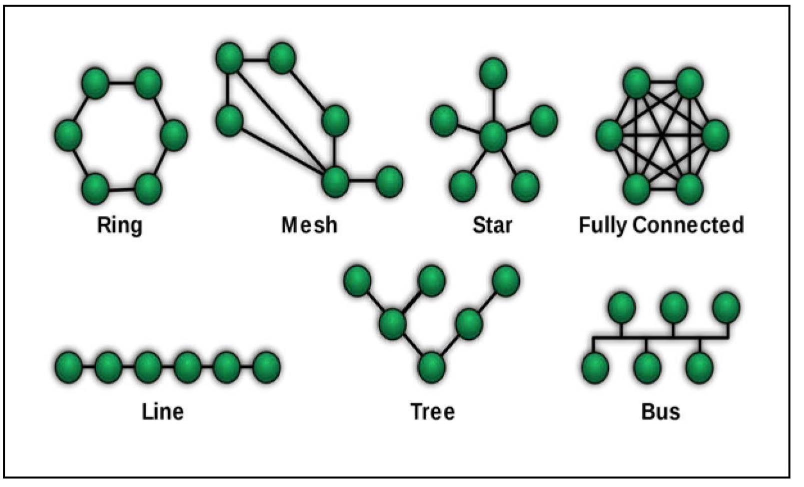


Figure 2.2 Topologies of a communication network

A star topology consists of one central switch, hub or computer, which acts as a coordinator to transmit messages including information and control commands. The central node provides a common connection point for all nodes and every node sends data to the central node. Star topology prevents the passing of data packets through an excessive number of nodes. Each device is inherently isolated by the link and this isolation also prevents any non-centralized failure from affecting the network.

In a bus topology nodes are connected by a linear sequence of buses. Every station receives all network traffic, and the traffic generated by each station has equal transmission priority. Each network segment is a collision domain. In order to transmit information on the same cable simultaneously a media access control technology such as carrier sense multiple access (CSMA) or a bus master is used.

In a ring network each node connects to exactly two other nodes, forming a single continuous pathway for signals through a ring. Data travel from a node to an adjacent node and each node along the way handles every packet.

In a mesh network each node relays data for the network. All nodes cooperate in the distribution of data in the network.

For health monitoring, a person could wear a wearable computing system such as Smart Shirt or its Life Shirt version, or use a hand hold device with sensors. Usually a star topology is preferred for WBAN application scenarios where a central coordinator collects all physiological data, which are acquired and sampled by sensing nodes.

It is the responsibility of a medical server to authenticate users, accept health monitoring session uploads, format and insert this session data into corresponding medical devices, analyse the data patterns, recognize serious health anomalies in order to contact emergency care givers, and forward new instructions to the users, such as physician prescribed exercises. The patient's physician can access the data from his/her office via the Internet and examine it to ensure the patient is within expected health metrics (heart rate, blood pressure, activity), guarantee that the patient is responding to a given treatment or a patient has been performing the given exercises. A server agent may inspect the uploaded data and create an alert in the case of a potential medical condition. The large amount of data collected through these services can also be utilized for knowledge discovery through data mining. Integration of the collected data into research databases and quantitative analysis of conditions and patterns could be valuable for researchers who try to link symptoms and diagnoses with historical changes in health status, physiological data, or other parameters (e.g., gender, age, weight).

2.2 Main requirements for WBAN application

In order to support different medical applications a WBAN acquire periodic and no periodic data and send information to a service node with minimum delay without losing of critical information. At the same time the system should exchange nonmedical data such as environment data and control information with remote control.

- A WABN needs to be operated in a power-constrained condition so that the battery can work for a reasonably long time. Energy saving is very important for implantable nodes compared with other devices. The controlling of energy consumption can be achieved by PHY and MAC design.
- An ideal WBAN should be secure, reliable and has the ability of self-healing.
- A WBAN needs to support various data rate varies from tens of kbs to Mbps due to the images and video transmission.
- The priority services should be considered in QoS management because a system must guarantee critical information for emergency situation or critical patients. So low transmission delay and packets loss are the key features.
- A WBAN will incorporate the transmission with a narrow band and an ultra-wide band technology to cover different applications.
- A WBAN may coexist with other network devices which operate in similar frequency band, where networks with various standards should cooperate with each other.

2.3 Constraints of WBAN

The purpose of WBAN is to build a pervasive health monitoring system. Although in many aspects WBAN are similar to WSN, there are intrinsic differences between them as shown in Table 2.1. Therefore for WBAN there are a number of specific constraints including better sensor design, biocompatibility, power supply, environment awareness, wireless transmission security, and so on.

Table 2. 1 Different characteristics for WSN and WBAN

Characteristic	WSN	WBAN
Scope	Meters/ kilometers	Millimeters/centimeters
Node Number	Large amount	Fewer (limited by space)
Accuracy	Redundant node compensates for accuracy and validation	Require high robust and accuracy
Size	Size is not a major limitation	miniaturization
Environment	Exposed in weather, noise and asynchrony	Motion and body physiology
Variability	Fixed or static structure	Biological variation and complexity of a variable structure
Privacy	Lower level security	High level security to protect patient information.
Power	Power is easily supplied	Energy is difficult to supply
Biocompatibility	Not a consideration in most application	Must support for implantable and some external sensors.
Data	Loss of data can be compensated	Loss of data may require additional measurement for QoS

2.3.1 Sensor design

Advances in manufacturing, biological, electrical, Nano-engineeing techniques and Microelectro mechanical Systems (MEMS) can make wearable and implantable sensors available. Although the scope of these sensors is very wide some examples prove the progress of pervasive patient monitoring. For diabetes implantable glucose sensors are doing the trials to get rid of the need for periodic invasive blood glucose pinprick testing. Beside that the implantable sensor and a fast acting insulin may construct a “closed feedback loop” system to maintain blood glucose in a defined reference range[12, 13].

Improvements in MEMS technology offer the prospect of delivering efficient and precise sensors for no more than a few dollars. Accelerometers are a good example to illustrate the MEMS technology. When it is used in the automobile industry, \]ar airbag releases can be efficiently and reliably triggered during simulated accidents. The development and rapid commercialization of low-cost microelectro mechanical systems (MEMS) sensors, such as 3D accelerometers, has led to their integration into a diverse range of devices extending from cars to smart phones. And the IMU is used for falling detection without interference the normal life style[14]. Although much of this technology is still in experiment, it is believable that these sensors will change therapy methods for chronic diseases in the recent future.

There is obvious improvement in sensor design but it is still have to solve the problems of long-term stability for a more accurate result. Additionally, they need to be simple to be used and comfortable to be worn for long period of time. When large groups of patients carry implanted devices, interference and integration should be considered and a new industrial standard for wireless transmission frequency is required. In addition to that alarm, antenna and comfort are also key points in designing. Alarm option should be included to detect an outage or a sensor node fails. Miniaturized antenna should be considered for specific application. Flexible and stretchable materials should be selected so that sensors can easily be embedded in the clothes or be attached to the body comfortably.

2.3.2 Biocompatibility

Biocompatibility is another problem considered in design and applications. One successful example is the cardiac pacemaker and the Implantable Cardioverter-Defibrillator (ICD). Many patients may benefit from this device since 2000.A range of implantable devices currently is used in clinical. The fact that large groups of the patients already carry implanted devices and the integration of these already implanted sensors and effectors into a larger wireless body area network deserves consideration.

Interference of these devices with each other, as well as with day-to-day technologies used by patients such as mobile phones, is a concern that has been noted and must be addressed. The interference may result in not only sensor malfunction, but also might affect implanted drug delivery systems and stimulators. A new industrial standard for the wireless transmission frequency in body area network is required.

2.3.3 Energy consumption

One of the key considerations for WBAN is energy consumption. Usually WBAN nodes will require power for three main parts: the sensor itself, any signal conditioning or data processing circuitry, and the wireless data link. For all of these functions, the power requirements depend strongly on the nature of the measurement.

Energy consumption is also an important issue because it determines not only the required size of the battery but also the length of time that sensors can work. These factors are critical in the implantable sensors as well as external sensor setups. Big size of sensors is rejected not only in implantable sensors but also external sensors because they require miniaturization of the power source. One useful strategy is the development of micro-fuel cells in the implantable sensors with reducing the size of the power source while increasing the lifetime of the battery. In a WBAN, the wireless communication link is possible the greatest reason for energy consumption. Therefore the development of low-power wireless data paths is important. Reducing the power consumption of the radio transceiver is crucial to the application of WBAN. Ultra wideband radio is also suggested for WBAN with relatively high data rates and low power consumption.

2.3.4 Security

WBAN must have high security with encryption to protect the patient's privacy because sensitive patient information is being transmitted through the wireless network. But strong encryption needs extensive computation and resources which are limited in WBAN nodes. In order to maximize security while minimize resource utilization a robust and efficient security structure is required for application. Unlike general wired or wireless network architectures where the network configuration is mostly static and there is limited constraint on resources. Therefore a dynamic feature of WBAN means that normal static network authentication schemes are not applicable. Even the asymmetric cryptography technique could cause too much computation for WBAN. So a lightweight security infrastructure is required for the practical applications.

2.3.5 Reliability

The network reliability directly affects the quality of people's health care or monitoring. For a worst-case situation, it would be fatal if a life threatening signal has been missed. However, due to the restrictions on communication bandwidth and power consumption, typical network reliability techniques may not be practical for WBAN applications. Reliability can be defined by the application. For instance: reliability for data transfer is the integrity of data and all information are accurately received by the receiver; reliability for audio or video , reliability concerns more of tolerable distortion at the application layer. Usually the reliability in each layer of the communication stack has various requirements and definition of error rates, error burst, delay, error concealment techniques, etc.

The reliability of the network directly affects the quality applications. In a patient monitoring system, data transmission, reliability and latency is extremely critical. A WBAN uses a wireless channel to transmit data, which is inherently unreliable. Therefore it is a critical requirement to enhance reliability for widespread usage. Error checking and correction should be considered in the reliability schemes and the wireless technology should make less interference on other medical devices. The sensor data should be transferred in a safe model.

Usually there are two types of reliability:

Reliability of Infrastructure

WBAN requires a highly reliable network can be accessed anywhere anytime. One way to create reliable access is to utilize multiple wireless networks that may be available at a given location. The vital signs and environmental variables can be divided into several packets, transmitted over networks, and aggregated before being delivered to one or more healthcare professionals. This reliable wireless network should work even under the following circumstance:

1. Failure in the networks: when sink nodes, mobile devices, or databases fail;
2. Coverage limitations: when cellular wireless networks experience coverage problems;
3. Intermittent access: when a device is unable to access a network continuously.

The reliable wireless network architecture will allow users to overcome time and/or location-sensitive dis-connectivity or intermittent connectivity problems and failures in one network by switching to another wireless network. We assume that: (a) Devices can sense the presence of multiple wireless networks, (b) local characteristics of individual wireless networks are available and can be used in deciding which network to switch to, and (c) devices have the hardware to switch among multiple networks. Each of the networks may have its own complexity in terms of total bandwidth, usable bandwidth, coverage and reliability, required power level for transmission, priorities for access and transmission, and, any protocol specific requirements.

Reliability of message delivery

One of the most difficult challenges in WBAN with wireless networks, especially for emergency messages, is the reliability of message delivery. The unpredictable quality and reliability could cause difficulty in achieving continuous health monitoring and losing of monitoring signals. Eventually, missing of life-critical signals may eventually lead to life threatening disasters if patients are in urgent treatments. Usually there are two kinds of faults in message delivery: noisy and faulty sensor node reading due to failure or malicious attacks. To guarantee the reliability requirements of WBAN, it is necessary to create fault models and fault tolerance mechanism to support delivery of vital signs and collected information.

Researchers have proposed many methods to improve the reliability in WBAN. One simple approach is to use limited retransmission until the acquisition of the acknowledgement. However, retransmission often creates additional overhead to the network and requires more energy consumption. So there is still significant interest in exploring the autonomic sensing schemes for developing self-protecting, self-healing, self-optimising, and self-configuring.

2.3.6 Context Awareness

In addition to monitor a person's physiological parameters, researchers also pay attention to the importance of the environment or background of the person. Simple activities such as "sleeping" and "walking" affect on not only vital signals such as blood pressure or heart rate, but also on measurement of activities and mobility. This "Context awareness" is also helpful for errors detection and calibration. In normal conditions, visual monitoring effectively provides the environment or contextual information, but it is impossible in the pervasive healthcare monitoring setting. With fusing multiple sensors' data the environment awareness will bring optimal results for analysis.

Accelerometers are one type of sensors used for environment awareness. It is suggested as appropriate gadget to detect activity state including walking, sleeping, exercising and posture such as lying, sitting, and standing. Audio sensors have showed the potential to determine the level of noise, or if the object is talking. Integration of several sensors is an environment strategy to produce more accurate model.

2.4 Reliability Challenges in WBAN

As mentioned before the reliability is the most important issue for WBAN. The reliability of the network, directly affects the quality of patient monitoring. Because WBAN and traditional Wireless Sensor Networks have differences in communication bandwidth, power consumption, signal processing, and feedback control traditional network reliability techniques may not be available for WBAN applications. There are some challenges as described in the following:

Unpredictable traffic patterns: Data traffic in WBANs is entirely different from traditional WSNs. Sometimes we might experience data burst, sometimes no traffic and sometimes real time responses required in the emergency case of heart failure scenarios.

Network instability: The network topology might change frequently due to link failure, power failure, and mobility of the nodes. Also for the fact that sometimes certain devices need not to be in operational mode and while in sleep mode to save energy. This changes the network structure. The network topology might change frequently. Routing and medium access under these unstable conditions is challenging.

Data redundancy: There might be possibilities of high redundancy in the sensor data. The solutions can be in the form of data fusion and data aggregation. These techniques help to maintain robustness while decreasing redundancy in the data. But it also introduces latency and complicates in design.

Packet criticality: Some data in WBANs are most time critical and it needs real time attentions.

Extreme energy efficiency: To deliver the levels of comfort and unobtrusiveness for wide spread adoption, WBAN sensor nodes must be small and have batteries that last on the order of days to years, depending on the application. The size requirement obviously limits the size of the batteries that will power the nodes, so WBAN nodes must be extremely frugal in their energy usage.

Unique characteristics of the wireless channel: The behaviour of the wireless channel around the human body poses a unique set of challenges to reliable communication. The first of these challenges is severe attenuation of the wireless signal between the sensor and the hub that can and may push the received signal power below the level required for reliable sensitivity, and is typically limited in WBAN nodes due to their relatively small antennas and simple energy-efficient designs. Unlike other longer-range networks where the distance between the transmitter and receiver dominates signal attenuation, the strength of a WBAN signal is most affected by the physical location and orientation of the nodes in relation to each other as well as the human body, which can "shadow" or attenuate the signal. Moreover, people move about, which constantly changes the attenuation at a rate that depends on the type of physical activity.

Managing interference: The nodes in a WBAN can be centrally coordinated by the hub, thus allowing a large number of devices to coexist in a single network without having interfere with each other. Things become more complicated when multiple people wearing BANs come into range of each other. In this case coordination may become impossible. The difficulty comes from the peoples' actions, which are unpredictable from a network's viewpoint, and can result in networks' moving into and out of range of each other. In such a situation there is no natural way of choosing a network coordinator. As a corollary, any interference mitigation scheme will need to adapt faster than the rate at which the network topology changes if it is to be successful in minimizing interference.

Application requirements: WBANs also need to support a wide range of throughput rates (1 kb/s to 10 Mb/s) to accommodate higher throughput applications such as video, while still delivering the high reliability and low-latency required in many medical applications.

2.5 Factors affect reliability

2.5.1 Reliability of Communication protocol Stack

A protocol stack is a complete set of network protocol layers that work together to provide networking capabilities. It is called a stack because it is typically designed as a hierarchy of layers and each layer supports the one above it and uses those below it. Traditionally, each layer of the communication stack addresses reliability at different timescales to fix errors that are not correctable, observable, or too costly to correct at the lower layers. The reliability of WBAN is mainly determined by the design of physical and medium access control layers.

(1) Reliability at the physical layer

In digital communication the PHY is responsible for bit-level sending/receiving of signals between nodes. The transmitted data should be reliably reconstructed at the termination. The whole process of transmission is described as following: firstly the transmitter sends data to a target receiver in the form of bits, and these bits are mapped into signal waveform after modulation and coding. Then the transmitted waveform is converted with a RF oscillator to the desired central frequency, amplified and transmitted by the antenna. After that the waveform propagates through the wireless channel. Once the signal arrives at the receiver, it passes through the low-noise amplifier.

Some of the important metrics that characterize reliability at the physical layer include the signal to interference plus noise ratio (SINR), bit error rate (BER), and outage probability. Certain issues related to the reliability and relevant error sources at the PHY layer may also be explained through the help of basic channel capacity formulation. Achievable capacity for reliable communications may simply be written for additive white Gaussian noise (AWGN) channel as formulation 2.1:

$$C_{awgn} = B\log(1+\frac{P_{rec}}{\sigma_I^2+\sigma_n^2}) \tag{2.1}$$

Where B is the communication bandwidth, P_{rec} is the received power of the signal, captures the variance of different error/interference terms, $\sigma_n^2 = BN_0$ is the noise variance, N_0 is the noise spectral density, and $\frac{P_{rec}}{\sigma_I^2+\sigma_n^2}$ is referred to as the SINR. This equation will be used to discuss the major error sources that may impact the reliability at the PHY layer. Possible techniques may be used to mitigate the undesired effects of these error sources.

Attenuation: P_{rec} should be sufficiently larger than the combination of noise and interference powers for the reliable detection of received bits. Due to path loss, the received power is less than the transmitted power.

Multipath propagation: Apart from path-loss and shadowing, the received signal power is also subject to variation in time, frequency, and space. These three characteristics of the channel have critical impacts on receiver design and the reliability of communications.

Interference sources: Inter factors such as multiuser interference and narrowband interference may take σ_I^2 larger, and hence degrade the SINR and the reliability of the received signals. Short-range wireless communication systems typically have to coexist with various technologies utilizing different frequency bands. Therefore, they may receive interference from other wireless technologies such as the WLANs that operate within the unlicensed bands.

(2) Reliability at the MAC layer

Medium access control is of paramount importance in wireless system: it arranges how the spectrum is shared across users and directly impacts the system throughput, reliability, quality of service and fairness. The MAC layers needs to be designed to suit specific requirements of specialized applications.

At the MAC layer, reliability is traditionally defined from the data integrity point of review and packets erroneously received from the physical layer are dropped. Thus a critical metric at this layer is the packet drop rate (PDR) and at least for point to point unicast transmissions. MAC layer designs aim at marginalizing the packet drops due to link/ channel errors. Collision-free channel access and coded or uncoded packet retransmissions are the main mechanisms employed at this layer to improve the PDR. Limiting the reliability to data integrity and /or packet drop rates at the MAC layer is quite a narrow view once the requirements of several short-range radio applications are considered. Delay and jitter are directly impacted by the MAC layer decisions. The scheduling problem might be quite hard even under fixed channel conditions and error-prone wireless channels coupled with such scheduling decisions lead to an even greater challenge.

2.5.2 Fault tolerance for reliability

The aim of WBANs is sensing, gathering and processing the information in the region covered by the network and transferring the data to sink node [1]. Reliability of health monitoring and message delivery is one of the most difficult challenges in health monitoring using wireless networks, especially for emergency messages. The unpredictable quality and reliability could lead to difficulty in achieving continuous health monitoring and delivery of monitoring signals from a patient to healthcare professionals. In recent years, the researchers have reached a common realization that fault tolerant is a critical issue to control possible faults for reliability. But the traditional double or triple redundancies are not adequate solutions due to their power consumption, space, and cost. Additionally, new types of failure modes are characteristic for wireless sensor networks. These include the areas of sensors, communication, and processing capabilities, performance and power related failures and so on.

So generally for WBAN faults can be controlled from two aspects: fault prevention and fault management. The former one reduces possibility of faulty events, while the latter recovers the data transmission after detecting the faulty event.

(1) Fault Prevention

In order to reduce faulty possibility, we need to design some protocols to prevent the fault from happening, which can be realized by suppressing the sources of the faults. These faults in sensor networks can happen due to energy depletion in nodes, collision on data link and route breakdown. An efficient fault prevention protocol can reduce the occurrence of faults, which saves system resources.

Energy efficiency

Node failure in a sensor network occurs mainly due to energy constrains. In some applications, sensor nodes are deployed in an unattended area, which makes it impossible to have their batteries recharged. To prolong the lifetime of sensor nodes, energy consumption must be minimized when the network is running.

Collision-free MAC

Due to the broadcast nature of channels, all the nodes need to contend for the usage of the channel. Link failure happens when two or more nodes transmit at the same time in the neighbouring area, causing collisions or interferences in the transmission medium. In order to alleviate the severe impacts of collision on data transmission, some TDMA-based or contention-free MAC protocols are adopted. They assign different time slits to different nodes and allow each node to transmit at its own slot.

Coverage and Connectivity

Because of sensor networks' applications in monitoring, coverage area is a critical criterion of the quality of service for a network. Connectivity is the ability of active nodes to stay connected. Designing and deploying a sensor network with high connectivity will avoid occurrence of faults when some nodes are dead.

(2) Fault management

Although the fault occurrence can be reduced by some efficient algorithms and protocols, fault cannot be avoided. Fault management must be implemented to recover data in WBANs. These approaches can be divided into two main phases: fault detection and fault recovery.

Fault detection

Fault detection is a phase where it is recognized that an unexpected event has occurred. Furthermore the pattern of faults will be diagnosed and identified in order to make a decision. Traditionally, fault detection techniques are classified into offline and online detections. Most often, for offline detection, special diagnostic programs are employed, either during idle periods of time, or using multiplexing with a regular mode of operation. Online detection targets real-time fault identification and is performed simultaneously with a real work load. Therefore fault detection is a complex task and can be fulfilled by some special monitoring nodes, sensor nodes themselves or even the cooperation with other nodes.

Fault recovery

After a decision of fault has been made, the next step is to recover the data. Recovery is a stage where an attempt to eliminate the effects of faults is conducted. Two most widely used recovery techniques are fault masking and retry. The fault masking approach is one where redundant correct information is used to eliminate the impact of incorrect information. In retry, after the fault is detected, a new attempt to execute a piece of a program is made in the hope that the fault is transient. Restart is the stage that is invoked after the recovery of correct undamaged information. In cold-restart, a complete resetting of the system is conducted.

2.6 Related reseach of reliability in WBAN

To support the reliability and stability of health monitoring in WBAN, significant works are required to prevent faults, detect errors, manage resources and improve efficiency.

On one hand designing of fault control mechanisms in BANs to ensure reliability is still under development. Some recently fault control schemes are outlined as following: Duk-Jin Kim addresses an issue of identifying and isolating three different types of faults in a Body Sensor Network and proposes fault isolation strategies using history-based and non-history based approaches for remote monitoring of daily living activities [15]. The contributions of that paper are: (1) faulty sensor node identification in a small number of deployed body sensors (accelerometers); and (2) identification of a faulty sensor node using a statically or dynamically bound group of sensor nodes that is sharing similar sensor signal patterns. Zappi, Piero investigated the method of exploiting redundant sensors distributed on the body to balance interconnection failures and jitter in sensors[16]. Heiermann S. did a study to evaluate the reliability and accuracy of the portable Sense Wear armband in determining resting energy expenditure which is important to guarantee the operation [6]. Stefano Galzarano showed a self-healing layer is able to detect and possibly recover faults of sensing readings to improve the recognition accuracy[17].

On the other hand effective mechanisms are proposed in protocols to reduce the possibility of error or packets losing. An adaptive and flexible fault tolerant communication scheme (AFTCS) for BAN is proposed by Guowei Wu[18]. This scheme include: 1) In order to meet the reliability requirement of critical sensors, fault-tolerant priority and queue are employed to adaptively adjust the channel resource allocation. Thus it can adaptively provide the reliability assurance for the sensors with high demand of reliability. (2) A resource reservation method based on dynamic priority queue is presented, including bandwidth measurement, bandwidth requirements calculation and bandwidth allocation methods. In AFTCS, the fault-tolerant priority dynamically changes according to the fault-related information, which can reflect the dynamic change in reliability requirements of sensors. In case of channel impairments, the packet loss rates for critical sensors will be decreased after channel reservation, and the times of retransmission will be reduced, thus lowering the average transmission latency.

A novel quality-of-service fuzzy-rule based cross-layer scheduling algorithm under certain selected medical scenarios for body sensor networks is presented by Otal [19]. To fulfil the reliability requirements, packet transmissions are scheduled with taking into account the channel quality among body sensors and each sensor specific medical constraints. That is to say that sensors waiting time in the accessing system and their residual battery lifetime are also worth bearing in mind. For the fuzzy-logic scheduler implementation, adapted distributed queuing MAC protocol that has recently been proved to be energy-efficient.

Instead of using a circular coverage area, a more accurate model is defined based on the path loss along the human body. Braem proposes improvements to CICADA[20], across-layer multi-hop protocol that handles both medium access and the routing of data in BANs. CICADA is slot-based and uses schemes to allocate these slots. Results for two reliability improvements are given: randomization of the schemes and repeating the schemes received from a parent node. It shows that these improvements positively affect the throughput of the network and lead to fewer retransmissions while the energy consumption of the nodes is hardly influenced. The proposed model uses a lognormal distribution for determining the range of a node instead of a circular coverage area. Doing so, a more realistic view of the network is obtained. This model was subsequently used for evaluating the proposed reliability mechanisms. The scheme randomization does lead to better results, although the number of retransmissions increases for reasons that are not clear. Adding the parent's scheme to the control message increases the reliability even further.

The significant amount of network traffic generated by health-monitoring systems could affect the reliability of message delivery and the end to end monitoring delay. A reliable multi network-based architecture can be designed using cellular/3G, wireless local area networks, ad hoc wireless networks, radio frequency identification, and sensors-based networks. Although individually, a single wireless network might not provide the reliability, coverage, and networking resources to support highly reliable comprehensive health monitoring, the ability to access and switch among multiple networks can create a fault-tolerant architecture, capable of overcoming multiple access and failure problems.

The development of fault models for sensors will be in particularly difficult due to a great variety of their types, environments in which they will be deployed. Note that there is a need to address testing not just on the components and the individual node level, but also at the network and distributed system level. It is important to observe that for each of these applications there will be a variety of approaches for usage. For example, in addition to equation based sensor fusion, we will see graphical based sensor fusion; statistics based sensor fusion and stochastic based sensor fusion. Each of these techniques has a number of unique peculiarities. While we do expect that specific fault management and fault prevention techniques will be developed and advanced reliability methods that can be applied to multiple applications.

Chapter 3 Abnormalities Detection of IMU based on PCA in motion monitoring

3.1 Introduction

Continuous monitoring of persons' activities has become one prosperous part of the research areas in Body Area Network (BAN). In typical wireless tele-monitoring scenery, a person wears portable sensors with reading physiological data and sending them wireless to a central server or health care workers that can take corresponding response according to acquired information[21].Fault detection is a necessary process to recognize unexpected events for reliable services. The main goal is to effectively detect faults and accurately isolate them to a failed component in the shortest time. Furthermore the pattern of faults would be diagnosed and identified in order to make a decision to ignore or recover. Thus, an affordable and effective fault detections and identification method in proposed in the research of sensor signals to monitor daily activities.

3.2 Data Quality and Faults

3.2.1 Data Quality

The success of any sensor applications depends on the quality of the data. Without trust in that quality, the value of the data is significantly undermined, and it's observational, diagnostic, or actionable value is limited. It is important, therefore, to ensure that data quality is an integral part of any sensor application development process. A variety of issues can affect the data quality during the application lifecycle, in all of its phases:

- Sensor system design, development, and validation
- Deployment
- Protocol design
- Data processing and visualization

Some issues can be mediated through careful design of the sensor system. A tightly controlled deployment process or active management strategy can proactively identify issues that affect data quality. Key factors to be considered include data consistency, measurement accuracy, and reliability during data collection, processing, storage, and transmission. The primary objective should be to minimize, or ideally eliminate, both the general and application-specific data deficiencies. A risk matrix can be useful in prioritizing data-quality impacts. In this matrix, priority is normally allocated to both high-impact and high-frequency risks. The rationale for this prioritization is based on for significant impact. These are then followed in order by lesser influences, such as outlier detection. The key factors affecting data quality are outlined in Table3.1 [22].

Table 3. 1 Factors affect Sensor Data Quality

Factor	Impact
Sensor limitation	Operation limitations: over-sensitivity to environment influences; System design Limitation: data through.
Calibration error; drift	Incorrect calibration for the required operational range, or frequent recalibration needed to maintain accuracy; Accuracy deteriorates over time.
Environmental influences	Performance variation due to temperature, humidity, ingress of moisture; Degrading of sensor materials; Malicious damage to the sensor or the measurement environment.
Malfunctioning Sensor	Sensor ceases to function correctly, resulting in erroneous output.
Incorrect Values	Incorrect sensor measurements can arise due to noise.
Unsuitable Protocol	Measurement protocol cannot be utilized correctly due to its complexity, Protocol does not acquire data at required periodicity range
Human Influences	Skewed results due to location of humans in the measurement environment. Incorrect use of the sensor: incorrectly attached electrodes a body-worn application
Incorrect installation	Inaccurate sampling

3.2.2 Types of Faults

A fault is defined as an unpredicted or unexpected vibration or change of system behaviour such that it either deteriorates the performance or demolishes the normal operation of the system. While the former is usually called an incipient fault, the later is usually considered as a total failure[23]. A failure is usually the result of the progression of an incipient fault over time, and could lead to hazardous situations. Faults in a system are usually classified based on their time behaviour and their severity (i.e., their impact on system behaviour). From time behaviour point of view, faults can be classified into the following two groups:

• **Intermittent/Transient faults:** These faults persist for only a bounded period of time after their initiation. It should be noted, however, that even upon their termination the system may not behave in the same manner as before the fault initiation.

• **Permanent faults:** Once occurred, these faults exist forever unless the faulty component is replaced by a redundant one or serviced/repaired.

3.3 Fault diagnosis system

3.3.1 Fault diagnosis Process

A "fault diagnosis system" is defined as a system to detect the presence of faults, decide their locations, and estimate their severities which can be described as three tasks of detection, isolation, and identification of faults in a system. The definition of these three tasks can be described as following:

Fault detection: To make a binary decision whether something has gone wrong or that everything is fine.

Fault isolation: To determine the location of the fault, i.e., to identify which component, sensor, or actuator has become faulty.

Fault identification: To estimate the severity, type or nature of the fault.

The relative importance of the above three tasks highly depends on the application and the system operator's objective of having a fault diagnosis system. However, the detection is essential for any practical system, isolation is almost equally important, and identification is crucial for fault recovery and reconfiguration as well as health monitoring and maintenance purposes.

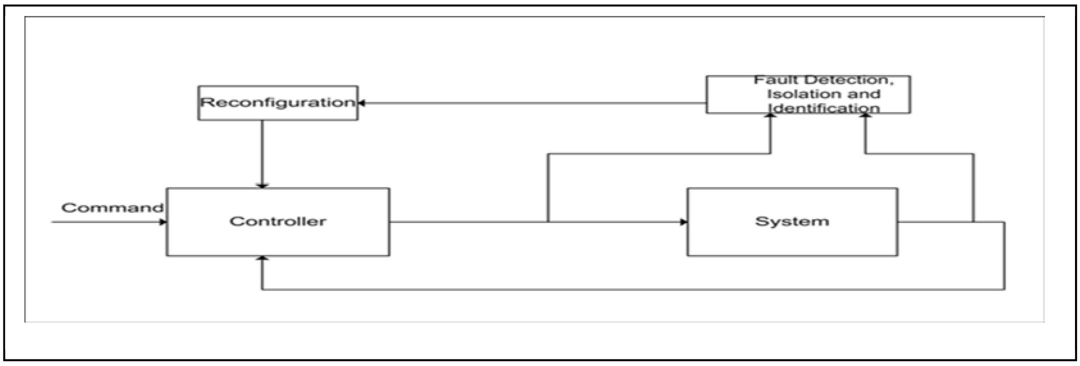


Figure 3.1 Fault Detection and Identification Process

We will mainly focus on sensor data faults since accurate data a WBAN is vital for enhancement of the reliability and safety of the system as well as fault prediction and consequently condition based maintenance (CBM).

3.3.2 Desired requirements of a Fault Diagnosis System

A fault diagnosis system should meet some criteria. Some desirable requirements of a diagnostic system are outlined as following:

• **Early detection and diagnosis:** It means the system has the ability of detecting and isolating incipient faults. Early detection and isolation of faults prior to their full manifestation into a failure is very critical for safety critical systems as well as condition-based maintenance. In addition to be sensitive to incipient faults, the diagnostic system should support false alarms under normal operational modes, which is a challenge in realizing early detection.

• **Isolation:** It means to distinguish the origins of a fault from other potential fault sources or to locate a faulty component among various components of a system. Isolation is also crucial to obtain fault tolerance, since proper counter-measures cannot be taken without knowing the source of an anomaly in system. Isolation of a fault does not depend only on the diagnostic system design but also on the way the fault affects system outputs. Moreover, various sources of uncertainties such as modelling uncertainty/errors and system disturbances pose a serious challenge to achieve a high degree of isolation.

• **Fault Identification:** It means the estimation of the severity, type or nature of the fault. Identification usually is very hard to achieve due to presence of measurement noise, system disturbances, modelling uncertainties, and last but not the least coupling/interactions between potential fault sources in the system being monitored.

• **Robustness**: Uncertainties are inevitable in practical settings. Therefore, robustness to measurements noise, system disturbances, and modelling uncertainties is one of the most highly desirable attributes of a diagnostic system intended for practical implementations. Robustness essentially augments diagnostic system reliability and effectiveness.

• **Novelty Identification:** Although the well-known, industry standard failure analysis tools such as FMEA and its recent extension, FMECA (failure mode, effects, and criticality analysis) provide fruitful information on potential failure modes within a system and their effects/impacts upon it, as well as charting the probability of failure modes against the severity of their consequences (i.e., criticality analysis), there is still a chance of novel anomalies occurring in the system. It is expected from a diagnostic system *not* to wrongly classify novel malfunctions in the system as other *a priori* known type of malfunctions or to treat them as being a healthy operational mode. While detection of novel faults is relatively easy to achieve, isolation and identification of them is extremely difficult to accomplish, especially because these faults cannot be modelled due their unknown nature.

• **Multiple Fault Identification:** This refers to the ability of a diagnostic system to identify and correctly classify multiple faults that may even coexist in a system.

This is a rather difficult requirement mainly due to nonlinearities and coupling/interactions that generally exist between the states and the potential fault sources of a dynamical system. Another reason is that some faults in an engineering system are extremely difficult to model because of their complexity.

3.4 Faults detection for Inertial Measurement Unit

Sensors are basically the output interface of a system to the external world, and convey information about a system's behaviour and internal states. Therefore, sensor faults may cause substantial performance degradation of all decision-making systems or processes that depend on data integrity for making decisions.

The methodology is applied for inertial measurement unit (IMU) based on signal analysis. Inertial measurement unit (IMU) is a sensor system used in health care for motion monitoring of senior people being alone or with disorders such as Parkinson’s and Alzheimer’s because ambulatory accelerometer is capable to detect falls in a setting environment, and it is also available for assessing gait and tremor of Parkinson.

Therefore, the measurements and parameters can be detected and reviewed by a clinician or professional staff. If abnormal results are obtained, the person would be asked for hospital visit or taking emergency treatment. These inertial sensors are mostly micro-electro-mechanical systems (MEMS) and unfortunately the actual sensor outputs can include some abnormalities, such as impulsive noise which is related to sensors or in data acquisition process[24]. These abnormal outputs from the WBAN might cause misinterpretations of regular living activities to persons being monitored. To overcome the anomalies, a method based on 2-CUSUM algorithm was applied by ElcioJeronimo to detect faults in IMU with minimal redundancy of gyros [25]. Shuguang Sun proposed a fault detection and isolation method for gyro errors with six sensors configuration redundant IMU[26]. These works for a complete FDI algorithm based on redundancy of gyros to ensure fault tolerance but it is not available for WBAN because wearing a larger number of sensors can be stressful on a daily basis. Sensors need to be simple to use and comfortable to wear for long period of time and it is better to put less sensors be embedded in the clothes or attached to the body.

In our experiment, we monitor person’s motion and detect falling with an IMU because it is effective to judge a person’s motion by acquired attributes [14]. Redundant sensors are not accepted in body area network. The falling detection by IMU is part of a body sensor network which focuses on health monitoring with sensors that acquire data such as heart rate, EMG, EEG, ECG, oxygen saturation, body temperature, and so on. Unfortunately, abnormalities detection of data collected from the IMU is still a challenging issue in real application because obtained samples are multivariate parameters. The issue is more difficult when the sensor produces data that is erroneous but not obviously wrong. (i.e., data is being communicated from the sensor but the data are wrong).

Since the system is complicated and energy efficiency constrains the processing time to analyse many attributes, a process of attributes selection algorithm is a critical issue in data analysis for fault detection. In this paper, we adopted a method based on PCA (Principal component analysis), which will generate a new small set of artificial variables, for analysis and detection abnormal signals from an IMU.

This chapter focuses on abnormalities detection of an on-body inertial measurement unit. Section two briefly outlines the proposed algorithm and experiment scenario. Section three describes the process of application and validates the result by simulation. Our experimental results show that the approaches indeed help detect abnormalities and provide good confidence in the IMU data that is being used for clustering motions. Section four makes conclusions and introduces future works.

3.5 Methodology of Fault detection for IMU

3.5.1 Experimental Setup

An inertial measurement unit (IMU) has been widely used for inertial navigation of moving vehicles and motion detection. IMUs are commonly referred to as tri-axial sensor clusters since a minimal setup to sense 3-D motion requires three accelerometers and three gyroscopes aligned to each orthogonal axis.

The Shimmer 9DoF expansion module is used as an IMU in the experiment for motion monitoring with 9 degrees of freedom motion capture. The 9DoF board combines a free scale MMA73613-axisaccelerometer, an InvenSense5003-axis gyros, and a Honeywell HMC58433-axis magnetic sensors to provide a comprehensive kinematic sensing solution. Wireless communication is established with a ZigBee protocol in 2.4 GHz to communicate with the receiver board. ZigBee is a specification a suite of high level communication protocols used to create personal area networks with small, low-power digital radios. Though low-powered, ZigBee devices often transmit data over longer distances by passing data through intermediate devices to reach more distant ones. It is used in applications that require a low data rate, long battery life, and secure networking. ZigBee has a defined rate of 250 kbps, best suited for periodic or intermittent data or a single signal transmission from a sensor or input device. ZigBee networks are secured by 128 bit symmetric encryption keys.

In home automation applications, transmission distances range from 10 to 100 meters line-of-sight, depending on power output and environmental characteristics.

At the receiving termination of the wireless communication link, the samples are handled in real time. The accelerometer sensor is placed in the right pocket of a vest. It works by detecting the current rate of acceleration with accelerometers, detecting changes in rotational attributes like pitch, roll and yaw with gyroscopes and detecting direction and magnitude of external magnetic fields with magnetic sensors.

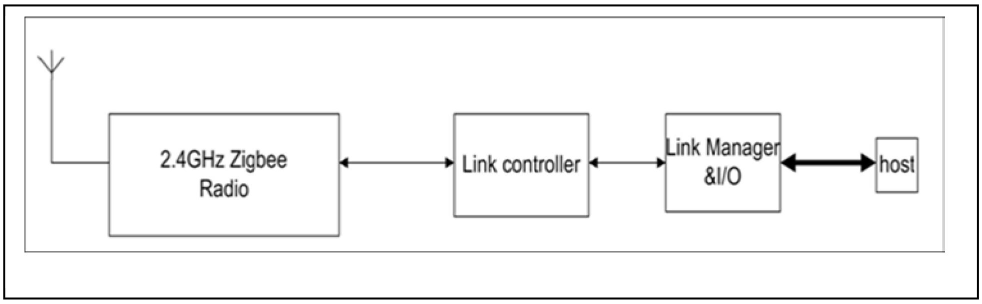


Figure 3.2 ZigBee Wireless communication

During walking, the frequency of torso swing ranges from 0.6 to 2.5 Hz, so we sampled data for processing at 20Hz and each signal has 5 seconds samples in the experiment. We arranged our experiments to detect movements. From the experimentation we acquired a total of 7000 samples of data with normal activity (walk slow, normal and fast) and 500 samples of fault data (under wrong scenarios) . Nine attributes in each sample are measured in motion monitoring with IMU. In order to identify abnormal data, we deployed an attribute selection technique called PCA for dimension reduction.

3.6 Overview of Principal Component Analysis

PCA is a vector space transformation often used to transform multivariable space into a subspace which preserves maximum variance of the original space in minimum number of dimensions because the measured process variables are usually correlated to each other. PCA can be defined as a linear transformation of the original correlated data into a new set of uncorrelated data, which means PCA rotates the original coordinate system along the direction of maximum variance.

The number of principal components usually is less than or equal to the number of original variables. Consequently the first principal component represents the largest possible variance, and each succeeding component has the decreasing variance under the constraint that it is orthogonal to the preceding components. The principal components are the eigenvectors of the symmetric covariance matrix so they are orthogonal.

The modelled component is the distance from the origin to the principal eigenvector. Likewise, the residual component is the distance from original to the residual eigenvector. When a PCA model is applied to a new observation, the two performance statistics, which are normally considered are the squared prediction error (SPE) and Hotelling's. The SPE indicates how much an observation deviates from the model and is defined as SPE. Alternatively, Hotelling's indicates how much an observation deviates inside the model, and is calculated. A process is considered normal as referenced in [25] and [4], respectively. Therefore PCA is a useful method to transform original process variables in a new set of uncorrelated variables which represent the trend of a process. PCA has been used for various multivariate data analysis techniques such as process monitoring, quality control, sensor and process fault detection, fault diagnosis[27, 28] . The model building of PCA relies on eigenvalue decomposition of parameter covariance matrices and the computational scales will reduce dramatically based on principal elements. Usually PCA uses the correlation structure to set error alarms and identify faults. Generally a model is built in the first step to characterize the correlation of the data. New measurements can be compared to the built model. If a parameter of new measurements is significantly deviated from the model's parameter, the measurement is assumed as a fault.

In normal condition, the PCA is established with a collected data matrix $X \in R^{n*m}$, where n is the number of samples and m is the number of variables. This matrix must be standardized to eliminate the effects of different units of variables. So the standard database $\overline{X}$ is firstly normalized. Then construct the covariance matrix R:

$$R = \frac{1}{n-1}\overline{X}^T\overline{X} \tag{3.1}$$

And then perform the SVD decomposition on R:

$$R = UD_\lambda U^T \tag{3.2}$$

Where U_{m*m} is a unitary matrix, and $D = \text{diag}\,(\lambda_{i=1,2,\dots,n})$ is a diagonal matrix. In equation 3.2, $U = [u_1, u_2, \dots u_m]$ is a standard base of R_m and the database $\overline{X}$ is described based upon U. The variances of $\overline{X}$ in the every direction from the new coordinate satisfy $\lambda_1 > \lambda_2 > \dots > \lambda_n$, where λ_i=1, 2,…, n are the diagonal elements of D_λ. The subspace, which is formed with the first k （k<n） vectors without correlation, is called principal component subspace; and the other subspace, which is formed with the n-k vectors $\tilde{S} = [u_{k+1,} u_{k+2}, \dots, u_n]$ are called residual subspace $\tilde{S}$. So the database X with m dimensions is replaced by the principal subspace $\hat{S}$ with k dimension and the residual subspace $\tilde{S}$ with n-k dimension.

The transformation matrix $P \in R^{n*k}$ generated by choosing k eigenvectors or columns of U related to k eigenvalues.

Elements of T, called as scores, are calculated by Columns of matrix P.

$$T = XP \tag{3.3}$$

Scores are the values of the original measured variables that have been transformed into the reduced dimension space. Where $\widehat{X} = TPT, E = X - \widehat{X}$, at last raw data space can be calculated as:

$$X = TPT + E \tag{3.4}$$

It is critical to select the number of principal components, because TPT represents principal elements of variability in the process and E stands for the variability related to process noise. There are various criteria and methods for choosing the appropriate number of principal components. S.Valle presented a method based on the variance of the reconstruction error to select the number of PCs[29]. This method demonstrates a minimum over the number of PCs. Conditions are given under which this minimum corresponds to the true number of PCs.

The popular procedure for choosing components is Cumulative Percent Variance (CPV) approach [30].It is a method of the percent variance CPV (i) captured by the first k principal components:

$$CPV(i) = \frac{\Sigma_{j=1}^{k} \lambda_i}{\Sigma_{j=1}^{n} \lambda_j} \tag{3.5}$$

3.7 Abnormalities detection with PCA

Abnormalities detection can be realized in the subspace with fewer dimensions by Hotelling T^2 statistical variables. A PCA model is established according to historical data and multivariate control diagram set by Hotelling's T^2 and squared prediction error (SPE).

T^2 stands for the major variation in the data and can be counted as the sum of squares of a new process data vector x:

$$T^2 = X^T P U_k^{-1} P^T X \tag{3.6}$$

Where U_k is a squared matrix formed by the first k rows and columns of U .

The data are regarded as right for a given significance level if: $T^2 \le T_\alpha^2$

$$T_\alpha^2 = k*(n-1)/(n-k)*finv(\alpha, K, n-k) \tag{3.7}$$

Where finv(α, k, n-k) is the critical value of the Fisher-Snedecor distribution with n, n-k freedom and α the level of significance. α usually uses values between 90% and 95%.T^2 comes from the first k principal components therefore it offers a detection for derivations in the latent variables which are most critical to the variance in the procedure.

Besides T^2 , SPE can also detect the variation in new events. The square prediction error (SPE), also known as Q, is a statistic that measures the lacking of fit in a model to data. The SPE statistic indicates the difference, or residual, between a sample and its projection into the k components retained in the model.

Mathematically, the SPE is a measurement of the sample to the model and is directly associated with the noise:

$$SPE = X^T (I - PP^T) X \tag{3.8}$$

The upper limit of this statistic can be computed as the next form:

$$\delta_\alpha^2 = \theta_1 \left(\frac{c_{\alpha\sqrt{2\theta_2 h_0^2}}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_2} \right)^{\frac{1}{h_0}} \tag{3.9}$$

Subject to: $\theta_1 = \Sigma_{j=K+1}^{m} \lambda_j$, $\theta_2 = \Sigma_{j=K+1}^{m} \lambda_j^2$, $h_0 = 1 - \frac{2\theta_1\theta_2}{3\theta_2^2}$

where c_α is the value of the normal distribution with α the level of significance.

If SPE $> \delta_\alpha^2$, it means the serious deviation has occurred. So when an abnormality occurs and it causes a change in the covariance structure of the model, it will be judged by a high value of SPE to detect the fault.

3.8 Application of Abnormalities Detection

3.8. 1 Measured Parameters and analysis

The motion monitoring by IMU is measured by nine state variables. The involved variables are combined of: accelerometer-x, accelerometer-y, accelerometer-z, gyros-x, gyros-y, gyros-z, magnetic-x, magnetic-y and magnetic-z. The process is divided into two steps: data acquisition and abnormalities detection. The former one is done with Java the latter one is done with Matlab. In the data collection process the sampled original acceleration signals in discrete time was captured using a tri-axial accelerometer. And then a defined algorithm based on PCA is executed by Matlab for abnormality detection.

In our experiment abnormalities detection procedure is described as follows: Generating the training samples in the reference situation, we chose 2000 and 5000 samples of nine variables in normal process such as walk normally. Some samples are showed in Table 3.2.

Table 3. 2 IMU Samples

ax	ay	az	gx	gy	gz	mx	my	mz
240	-659	-282	0.00959	0.006393	0.006393	531	-800	-575
245	-656	-288	0.003197	-0.00639	0	531	-800	-575
236	-653	-285	0.003197	0.003197	0.003197	540	-806	-580
239	-670	-289	0.003197	-0.0032	0.003197	535	-815	-586
244	-639	-273	0.006393	0	0.003197	544	-814	-584
254	-666	-563	0.003197	0.003197	0.003197	544	-814	-584
237	-652	-300	0.00959	0.003197	0.006393	547	-808	-583

Computing the feature matrix U based on training samples. The eigenvalues U_i and corresponding eigenvectors V of covariance matrix C of training samples X can be obtained, and further decreasingly ordered. The first k eigenvectors are packed to form the feature matrix base on equation (5). A calculation based on CPV reveals the first three components and it represents the high portion of data, which is clearly shown in Table 3.3 and 3.4 with the eigenvalues and the proportion of each eigenvalue of the total data.

Table 3. 3 Eigenvalues and the proportion of each eigenvalue with 2000 samples

Principal Components(PC)	Eigenvalue	Proportion
PC1	85584.94	0.351404
PC2	42430.02	0.174214
PC3	39139.67	0.160704
PC4	34688.71	0.142429
PC5	32727.47	0.134376
PC6	8979.498	0.036869
PC7	0.547325	2.25E-06
PC8	0.190671	7.83E-07
PC9	0.088746	3.64E-07

From the tables we can clearly notice that the first three eigenvalues represent approximately 70% of the total data. Therefore, in our work we will take the first three principal components and use corresponding scores as parameters for further analysis.

Table 3. 4 Eigenvalues and the proportion of each eigenvalue with 5000 samples

Principal Components(PC)	Eigenvalue	Proportion
PC1	57697.89	0.315134
PC2	34871.84	0.190463
PC3	29459.7	0.160903
PC4	26616.48	0.145374
PC5	25665.46	0.140179
PC6	8764.439	0.04787
PC7	13.84535	7.56E-05
PC8	0.443944	2.42E-06
PC9	0.136238	7.44E-07

3.8.2 Abnormalities detection

Abnormalities detection can be made by comparing the control lines in the training state with those in the setting state. Figure 3.3 shows the results of dimensioned data based on three principal components for 2000 normal samples and 500 abnormal samples. "*" represent normal data while "o" represent abnormal data. The axis represents the value of scores of each principal component respectively. It is clearly to see normal data are centralized and abnormal data are far away.

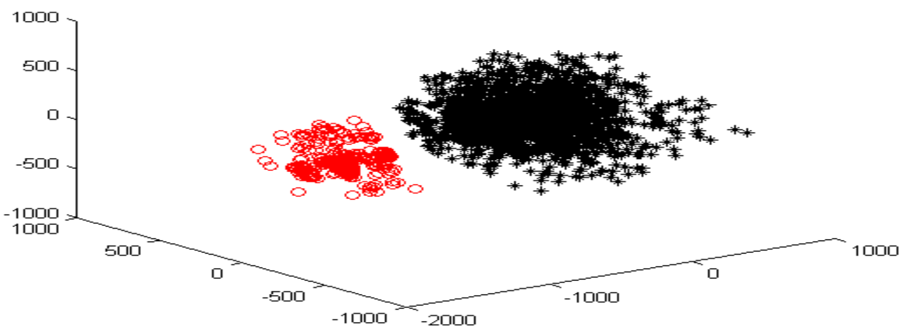


Figure 3.3 Comparison of 2000 normal samples and 500 abnormal samples

In the monitoring experiment we chose 2000 and 5000 normal samples to validate our algorithm. The upper limits T^2 of PCA model are 7.8359 and 7.8253 respectively. Figure 3.3 and Figure 3.4 show that three principal components are used to capture variability of process with T^2. T^2 of normal data are under thresholds while T^2 of abnormal samples exceed the control line. At the same time we detect fault data with SPE. Figure 3.5 and Figure 3.6 show abnormalities are also detected by SPE statistic. SPE $> \delta_{\alpha}^{2}$ means the serious deviation has occurred, that presents the abnormal measurement detected with IMU.

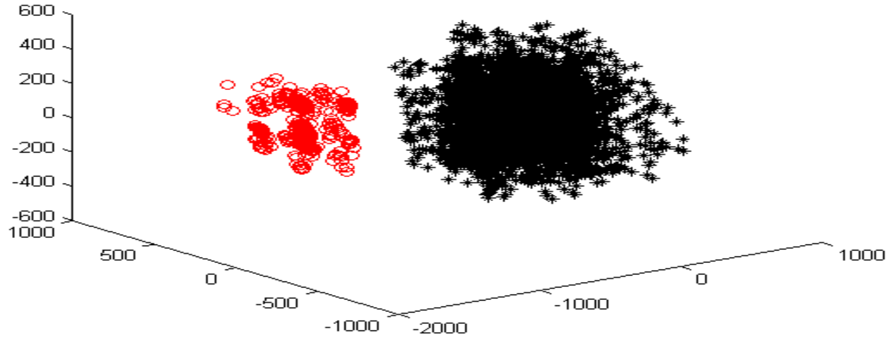


Figure 3.4 Comparison of 5000 normal samples and 500 abnormal samples

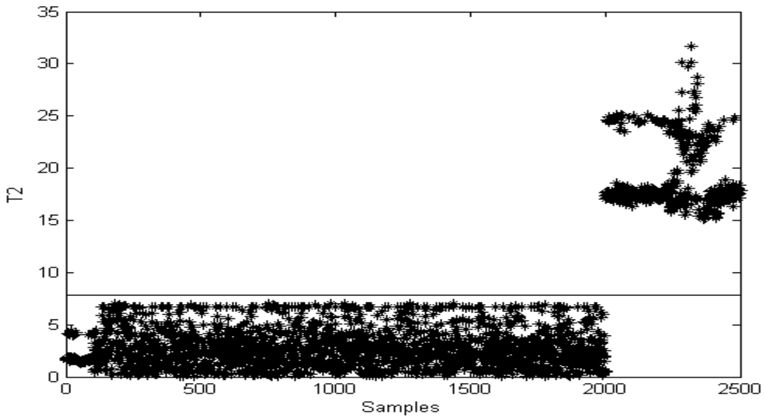


Figure 3.5 T^2 Detection for Abnormalities (2000 normal and 500 abnormal data)

From the above figures, we can see it is easy to detect abnormalities with this method and an alarm can be considered in the process when an abnormal signal occurs. If there are continual alarms in monitoring, an uncommon event may happen and we should replace this sensor or take further diagnosis by other measurement.

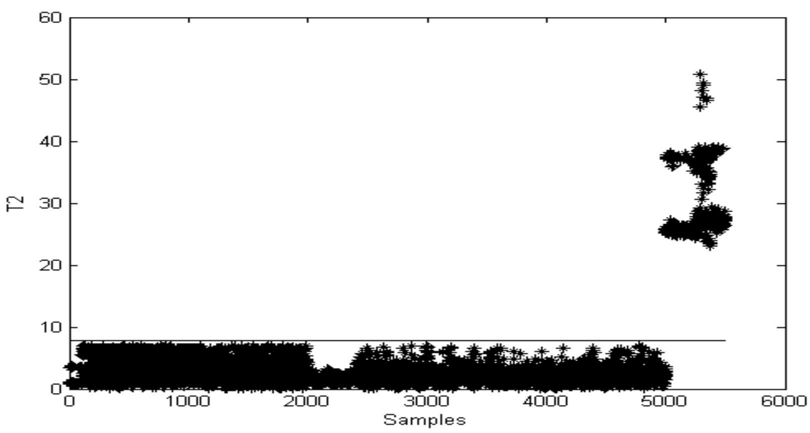


Figure 3.6 T^2 Detection to Abnormalities (5000 normal and 500 abnormal data)

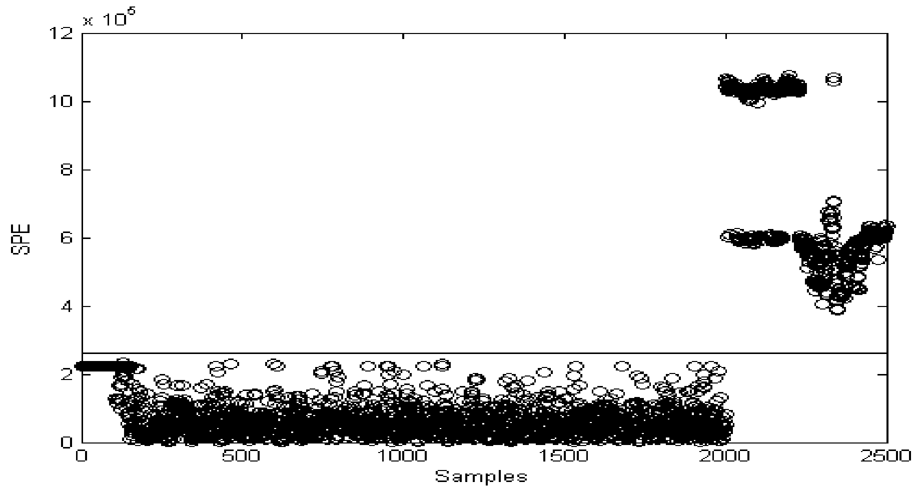


Figure 3.7 SPE Detection to Abnormalities (2000 normal and 500 abnormal data)

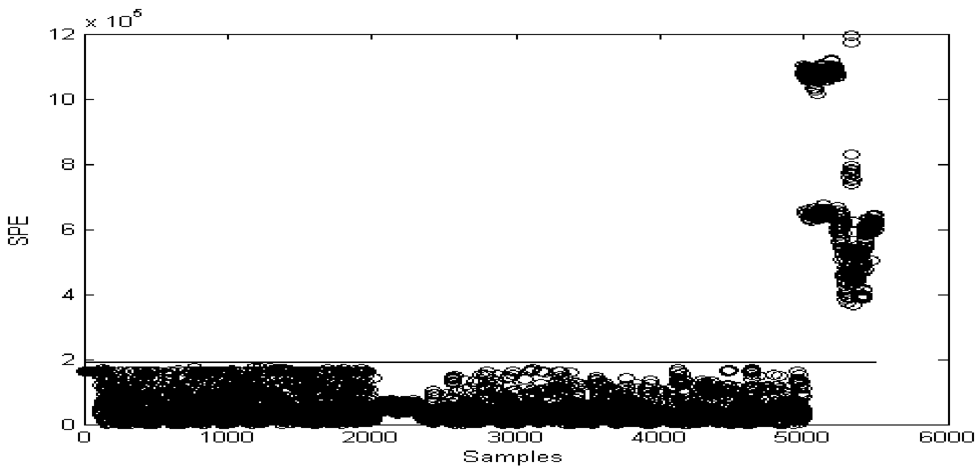


Figure 3.8 SPE Detection to Abnormalities (5000 normal and 500 abnormal data)

3.9 Conclusions

Normal fault diagnosis of IMU based on redundant devices is not available for body sensor networks due to the limitation of sensor numbers and energy saving requirements. A simple and effective algorithm is proposed for abnormalities detection in motion monitoring based on PCA in this chapter. Firstly we acquire training data which represents normal process operations and then scale the training data to build PCA model. By reducing dimensions with PCA and calculate upper control limits of T^2 and SPE to compare normal and abnormal samples. The effectiveness of algorithm has been illustrated through the real measured IMU data from the FEIT, University of Technology, Sydney laboratory.

This PCA based algorithm is simple, feasible and effective to detect abnormal data in IMU for motion monitoring. The method has many merits, such as less computational cost, independence of system model and etc. Although the control line can reflect the decisional criterion whether the error is present or not, fault identification and isolation under different sceneries and networks environment is still explored to achieve better performance combined with other new technologies.

Chapter 4 Fault detection and identification spanning multiple processes with PCA and NN

4.1 Introduction

Cardio-respiratory parameters, such as oxygen consumption rate, carbon dioxide production rate and heart beat are major indicators to evaluate human's physical status and sports performances [31-35]. These parameters are traditionally tested by indirect calorimeter with a metabolic cart. However, most metabolic carts are restricted in laboratories due to their size and weight limitations. Although many activities can be simulated in the laboratory (e.g. walking on a treadmill), some activities involving in occupational, recreational activities and health monitoring are not available in the laboratory. Using heavy equipment to collect expired air in fields can often disturb activities under investigation. Because of the improvement of miniaturized metabolic measurement systems oxygen consumption can be measured in real circumstances by small gadgets. Recently lightweight, portable telemetric gas analysis systems are used to acquire parameters in daily activities or sports for further analysis.

Among several gas analysis devices COSMED K4b^2 metabolic measurement system is a typical and popular instrument, which is designed to perform pulmonary function tests. It is to be used by physicians or by trained personnel on a physician responsibility. It is designed to measure parameters of ventilatory, oxygen consumption and carbon dioxide production with several sensors such as flow meter, oxygen sensor, carbon dioxide sensor, environmental sensor and so on[31, 36-39]. This system is one of the latest portable devices for cardiopulmonary gas exchange analysis base on true breath-by-breath without limitation. Along with the widely usage of COSMED K4b^2, the reliability is very crucial for measurement in metabolic testing or health monitoring. R Duffield assessed the validity and reliability of a COSMED K4b^2 portable telemetric gas analysis system by experiments and reliability is formed in specific steady state and sustained exercise[40]. However, in free living environment, the sensor outputs and data received by base station may be abnormal, which can be caused by impulsive noise, low batteries or environment interference. These abnormal data might cause misinterpretations of exercises or living activities and lead to unreliable results. Although some methods are available to detect faults, normal approach may yield false alarms for multi-processes applications especially for wireless transmission. Therefore, an effective and feasible method is necessary to detect abnormal situation and identify faults in multi-statuses for monitoring. In this chapter, a novel fault detection and identification model is proposed for systems performing in multiple processes with integrating principal component analysis (PCA), K-means and neural networks. Firstly, the PCA is used to detect the potential faults with decreasing the dimension of data set so that the energy consumption in transmission is consequently reduced. Thereafter, in the reduced data space, K-means is applied to cluster data into different groups and detect faults with statistic criteria. Then a neural network is constructed and trained to identify faulty by reconstruction. Based on real data experiments and simulations the proposed fault detection and identification (FDI) model is proved to automatically detect and identify faults. The approach is implemented from the time point that data are generated and alarm is set when faults present.

The main contributions of this chapter contain three aspects: 1) It applies the dimension reducing to decrease energy consumption in transmission which is very important in wireless body sensors networks; 2) With clustering it can automatically detect faults in different processes which is usually different in traditional model; 3) This model works beyond fault detection and it can identify crashed sensors according to sampled data.

The remaining sections of this chapter are presented as following: Section 2 introduces the portable medical device: COSMED K4b^2. Section 3 explains the principles of PCA and k-means for clustering and fault detection. In section 4 the back propagation (BP) neural networks for prediction is outlined. Section 5 verifies the FDI model with demonstration of experimental results. The section 6 makes a conclusion.

4.2 Overview of K4b^2 system

A K4b^2 system is a COSMED portable medical instrument used for testing of pulmonary functions as shown in figure 4.1. It can be worn by people during activities and is capable of delivering real-time measurements into a PC base station. Due to its convenience, it is applied in many fields such as: sports medicine research, gait lab, occupational health, cardiology, cardiac rehabilitation, clinical nutrition and so on. It can measure physiological response to exercise in the field without limitation.

This telemetric gas analysis system contains a soft, flexible face mask to sample expired air and sensor units to test ventilation (VE), oxygen (O_2) and carbon dioxide (CO_2) concentrations in the expired air[41, 42]. These measurements represent energy cost of activities and are highly correlated with each other. So it is available to extract main principal components for reducing transmission amount. The meaning of part measured symbols of K4b^2 in breath by breath exercise testing are showed in Table 4.1[36]. K4b^2 can be an auxiliary instrument used by physicians or by trained personnel on a physician responsibility and provide support in the following aspects:

• The formulation for lung pathology diagnosis;

• Important research related to human physiology;

• The collecting of critical information for sport medicine.

Table 4. 1 Measured parameters in K4b^2

Symbol	UM	Parameter
VO2	Ml/min	Oxygen Uptake
VCO2	Ml/min	Carbon Dioxide production
Vt	1	Tidal Volume
FetO2	%	End Tidal O2
FetCO2	%	End Tidal CO2
R	---	Respiratory Quotient
VE	l/min	Ventilation
HR	l/min	Heart Rate
Qt	l	Cardiac output
AT	---	Anaerobic Threshold
VE	l/min	Ventilation
SV	l/min	Stroke volume
RF	l/min	Respiratory Frequency
FeO2,FeCO2	%	Averaged expiratory concentration of O2 or CO2
VE/VO2	---	ventilatory equivalent for O2
VE/VCO2	---	ventilatory equivalent for CO2
VO2/HR	Ml/beat	Oxygen pulse
VO2/Kg	ml/min/Kg	VO2 per Kg
Ti,Te,Ti/Ttot	sec	time breaths
Vd/Vt	---	Vd/Vt ratio
PaCo2	mmHg	arterial PCO2 (estimated)
P(a-et)CO2	mmHg	Delta PaCO2 – PetCO2

Figure 4.1 Cosmed K4b²

4.3 PCA and K-means

4.3.1 Faults detection with PCA

PCA is often used to transform multivariable space into a subspace which preserves maximum variance of the original space in minimum dimensions[30, 43]. The measured process variables are usually correlated to each other. The measured process variables are usually correlated to each other and data can be disposed into the significant patterns. Then the abnormalities, such as noises or outliers in residual subspace can be identified. PCA transforms the original correlated data into a new set of uncorrelated data set that represent the trend of the process. It is highly useful in analysing state data which contain relationships between variables. It is proved successfully in many applications such as reducing dimensionality, data compression, and fault detection [44-48].

The square prediction error (SPE), also known as Q, indicates the difference, or residual, between a sample and its objects into the k components in the model[49, 50]. The SPE value is used as the main criterion for fault detection because training data are multivariate normal in the feature space.

4.3.2 Clustering with K-means

K-Means Clustering is an algorithm to classify or to group objects into K numbers of group based on attributes or similarities [51-53]. This is a type of vector quantization and is popular for cluster clustering and data mining. The aim of this algorithm is to divide observations into k collections where each observation is classified to one collection with the nearest mean. K is a positive integer number and the grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid. Thus, it is effective to separate the different process modes.

K-means can be embedded in Euclidean vector spaces and the algorithm usually works in two stages. In the first stage, initial k centroids are randomly chosen. In the second stage, each point in the data set is allocated to a cluster which has the nearest centroid. Then values of centroid in each cluster are calculated. Depending on the new centroids' values, stage 2 is repeated until the centroid values converge to the same value. The detailed process can be described as following:

1. Choose K initial cluster centroids $\{z_1, z_2, ... z_K\}$ randomly from the n points $\{x_1, x_2, ..., x_n\}$.

2. Assign points x_i, i = 1, 2, . . . , n to cluster $C_j \in \{1, 2, ..., K\}$, if $D(Z_j, x_i) < D(Z_p, x_i)$,
p = 1, 2, …, K, and $j \neq p$, ties are resolved arbitrarily.

3. Compute new cluster centroids, $Z_1^*, Z_2^*, ... Z_K^*$ as follows:

$$Z_i^* = \frac{1}{n_i} \sum_{x_j \in C_i} x_j, i = 1, 2, ..., K \tag{4.1}$$

where n_i is the number of elements belonging to the cluster C_i.

4. If $Z_i^* = Z_i$, i = 1, 2, . . . , K, then terminate. Otherwise continue from Step 2.

K-means clustering minimizes the global cluster variance J to maximize the compactness of the clusters. The global cluster variance J can be represented as

$$J = \sum_{K=1}^{K} \sum_{x_j \in C_K} D^2(Z_K, x_j) \tag{4.2}$$

Note that if the process does not terminate at step 4 normally, then it is executed for a maximum fixed number of iterations[54].

For the multiple processes fault detection, firstly acquired values are transformed by PCA into another subspace where K-means can be applied efficiently. Data must be classified to specific process to avoid false alarm and SPE is used as the main criterion for fault detection. If the SPE is within the threshold, then data are considered normal. Otherwise it means faults occurrence and further approaches will be executed.

4. 4 Prediction with Artificial Neural Networks

4.4.1 Principles of artificial neural networks

An artificial neural network(ANN) is a mathematical model to simulate some behaviours of biological nervous systems[55]. It is modelled on the interconnection of the neuron in human brain's nervous systems. ANN is constructed from atomic components known as "neurons".

Actually it is a computation tool and a type of non-linear processing system which is appropriate for lots of application. ANN can be trained to solve certain problems using a learning scheme and sample data. It is very useful for circumstances with an abundance of data, but less underlying theory. The data may be hard to be modelled because they may come from non-linear or non-stationary experiments. Importantly, ANN does not require any priori assumptions, not even statistical distribution. The ANN performs the necessary analytical work with training.

Each neural network has three critical components: Neurons, network topology and learning rules. Neurons determine how signals are processed are connected together with weights so that they can deal with information collaboratively. Network topology determines the ways nodes are organized and connected. Learning rules determine how the weights are initialized and adjusted. ANN has been used for time series prediction, pattern recognition, and fault diagnosis [56-58].

Neurons can have different forms and are connected into a particular network pattern to serve specific functional purposes. The connections between neurons are allocated multiple weights, which can be used for calibration or training to generate the proper output. Each layer represents a combination of non-linear functions from the previous layer. Each neuron is a multiple-input, multiple-output (MIMO) system that receives signals from the inputs, generate a complex signal, and send the signal to all outputs. Learning is an important process which typically happens in a specific training stage. Once ANN is trained, it goes into a production stage. Training can be different forms including an association of learning paradigms, learning algorithms and learning rules. A learning paradigm can be classified as supervised, unsupervised or a hybrid with both methods. A learning rule is a model for the types to train the system, and also a goal for what types of results are to be produced. The learning algorithm is a mathematical scheme that is used to adjust weights during each training repeating.

The most common structure of three layers is shown in Figure 4.2. This ANN consists three layers, which are called as the input layer, hidden layer, and output layer. Each layer consists of one or more nodes which represent Neurons. Input neurons in the first layer send data via synapses to the hidden layer and then to the output layer of neurons. The hidden layer is treated as a black box to people who are interfacing with the system.

When design a neural networks many different parameters should be considered, for example: the number of neurons per layer, the number of training iterations, learning rules and so on.

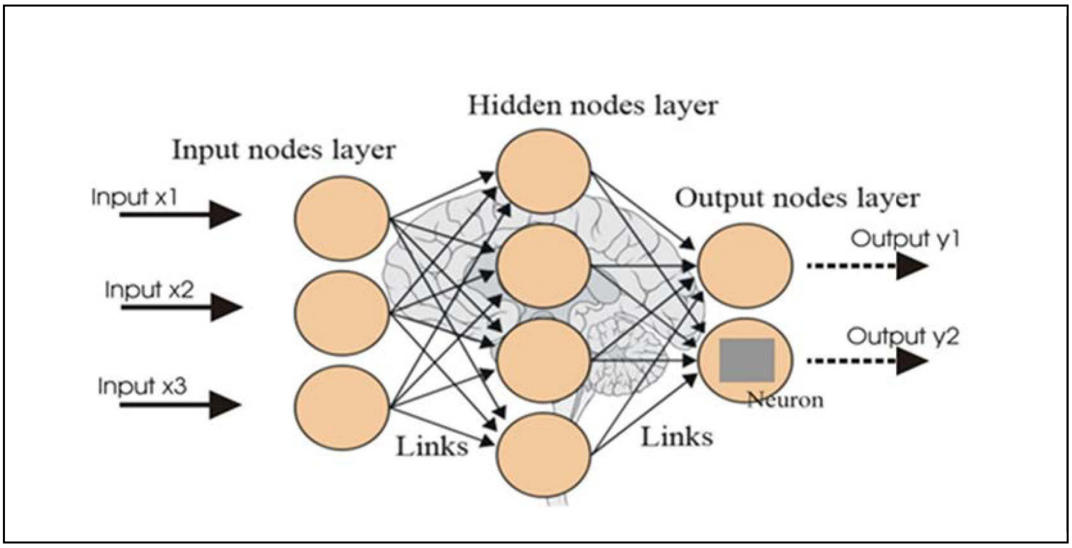


Figure 4.2 Structure of three layers Neural Network

The behaviour of an ANN (Artificial Neural Network) depends on both the weights and the transfer function, such as linear, threshold and sigmoid, that is specified for the units. The inputs are multiplied with weights and then added together to be put into an activation function, which defines the output. The process is outlined in figure 4.3.

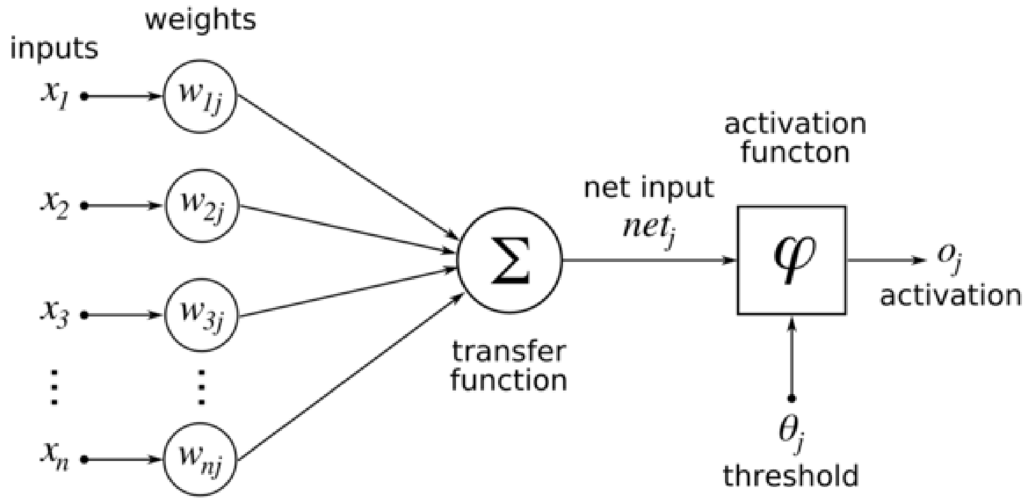


Figure 4.3 Process from inputs to outputs in Neural Network

The back propagation (BP) learning algorithm is one of the most popular ANN for practical applications [59,60].

In BP neural network, the weights of each unit are adjusted to reduce the error between the desired and the actual output. Therefore it must calculate the error derivative of the weights (EW).

A BP neural network is one feed forward neural network, in which neurons are arranged in layers, and each neuron can be connected only with the neurons in the next layer. BP neural networks are trained by a supervised learning algorithm. An error back-propagation algorithm is an iterative procedure typically used to train a BP neural network. Specifically, the process modifies the weights in the network in an iterative method so that the resulting network fits the training data well. A schematic diagram of a single output BP neural network is showed in figure 4.4. This neural network includes three layers: input layer, hidden layer, and output layer. The number of neurons in the input layer can be set according to actual requirement with historical values for prediction. Calculation for data processing takes place in the hidden layer. The output layer has only one neuron and the value is the predicted result. A BP neural network is a useful model for prediction and forecasting proved by many researches. For example, Jiantao Liu used a BP artificial neural network to predict the flow stress of high-speed steel during hot deformation[61].Yong Wang presents an accurate electricity load forecasting algorithm with back propagation neural networks [62]. So we apply a BP neural network for fault reconstruction based on predicted values.

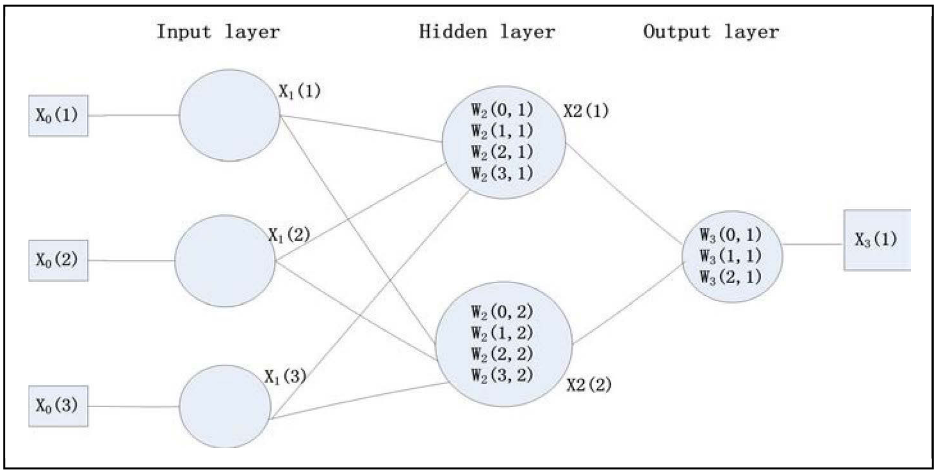


Figure 4.4 A schematic structure of a single output BP neural network

The entire network learning process includes two phases: the first stage calculates from input layer to output layer. Output can be calculated by training samples with initial structure and weight W; the second stage is modifying weights and threshold, and it starts from output layer to input layer, and weights of neurons connect to output can be adjusted according to errors of output. Hidden layer weight can also be modified. The two stages are iterative processes, repeat until convergence. All layers adjust weights through formula:

$$W_i(t+1) = W_i(t) + \eta \delta_i X_j^k \tag{4. 3}$$

Where η is regarded as precision of network learning, which can be used as conditions to judge the network finishing. δ_i is the value of error and can be defined as

$$\delta_i = y_i(1 - y_i)(d_i - y_i) \tag{4.4}$$

Where y_i is the output value and d_i is the desired output.

4.5 Procedure of Proposed Method

Due to the correlation of variables measured with K4b^2, PCA can be used for fault detection. But in health and exercising monitoring people usually perform at different status, such as walking, running or swimming. Classical PCA may fail to extract correction across different activities and then detect excessive faults or miss faults because grade transitions from one to another status can break the correlation between variables. Therefore it is necessary to design a feasible model to detect faults for multiple processes applications. The proposed model of multiple processes FDI is showed in figure 4.4.

Firstly data are acquired and then PCA is implemented to transform original data to subspace with extraction of main features. Consequently energy consumption is reduced in wireless transmission because only main components are transferred. Then K-means with setting reference patterns are adopted to judge various processes. Threshold of SPE is calculated in each process to detect abnormalities. If faults are captured a trained BP neural network is used to predict values for error reconstruction and identification. The whole procedure of FDI is outlined as the following steps:

The first step: Acquire validated data which represents normal process operations; carry out SVD to transform data for a PCA model; determine the number of principal components and build a formation of reference patterns with k-means.

The second step: Transmit main features of data, and implement the process judgment by reference patterns.

The third step: Evaluate SPE statistics to compare with the threshold obtained by PCA model. If the value exceeds the threshold then go to step four. Otherwise go back to obtain new data.

The forth step: When a fault is detected the trained BP neural network is applied to reconstruct the possible value based on the previous measured value.

The fifth step: Compare the reconstructed data with real sampled data and identify faults. If the fault is transient, replace it with the reconstructed data. Otherwise set alarm and identify fault sensors.

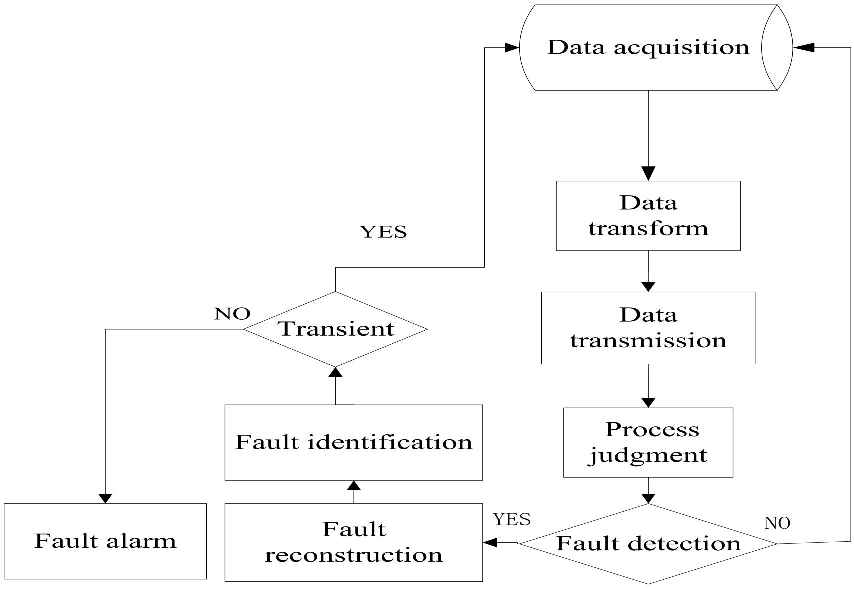


Figure 4.5 Approach of multiple processes FDI

4.6 Experiment and results analysis

4.6.1 Sampled data and clustering

In order to verify the whole procedure of this FDI model we use measured datasets from UTS FEIT laboratory in exercise monitoring. A $K4b^2$ system measured physical values of a person such as VO2, VCO2, FetO2, VE, HR and etc. The data samples in five separate time points are showed in table 4.2. We used 22 recorded values, which demonstrate tidal volume of each breath (VT), averaged expiratory concentration of O2 and CO2, oxygen uptake (VO2), carbon dioxide production (VCO2), and minute ventilation (VE), heart rate (HR) and so on. For example VO2 is calculated as the volume of inspired O2 minus the volume of expired O2 and the measurement at rest and during exercise increases abruptly. These parameters are necessary for pulmonary gas exchange measurement and acquired the feature matrix U by PCA analysis. The eigenvalues Ui and corresponding eigenvectors V of the covariance matrix C of training samples X are obtained, and ordered decreasingly. The first k eigenvectors are packed to form the feature matrix. A calculation based on CPV reveals the principal components represent the high portion of data, which is clearly shown in table 4.4 with the main eigenvalues and the proportion of each eigenvalue in total data.

Table 4. 2 Measured data (part1) by K4b^2

VT	VE	VO2	VCO2	O2exp	CO2exp	R	FeO2	FeCO2	HR	FetO2
l	l/min	ml/min	ml/min	ml	ml	---	%	%	bpm	%
1.24133	10.3015	569.440	483.495	178.805	71.3999	0.84907	14.4042	5.75184	97	13.4
1.30865	68.2778	3115.21	2650.34	203.387	62.3219	0.85077	15.5417	4.76227	98	14.2
1.34333	44.0438	1933.50	1706.40	211.037	63.8519	0.88254	15.7099	4.75322	100	14.48
1.36373	58.8664	2438.84	2181.35	218.075	62.0159	0.89441	15.9910	4.54749	103	14.7
2.11445	70.8756	2763.86	2660.93	342.821	97.4099	0.96275	16.2132	4.60685	107	15.35

Table 4. 3 Measured data (part 2) by K4b^2

FetCO2	FiO2	FiCO2	Ti	Te	Ttot	PetO2	PetCO2	PaCO2	P(a-et)CO2	PAO2
%	%	%	sec	sec	sec	mmHg	mmHg	mmHg	mmHg	mmHg
6.64	20.93	0.03	6.21	1.02	7.23	95	47	47	0	108
6.14	20.93	0.03	0.48	0.67	1.15	100	43	44	1	110
6.08	20.93	0.03	0.89	0.94	1.83	102	43	44	1	109
5.82	20.93	0.03	0.67	0.72	1.39	104	41	42	1	110
5.62	20.93	0.03	0.82	0.97	1.79	108	39	40	1	109

Table 4. 4 Eigenvalues and the proportion

Principal Components	Eigenvalue	Proportion
1	313138.1	0.976558
2	6740.755	0.021022
3	700.2923	0.002184
4	52.16869	0.000163
5	13.65424	4.26E-05

From table 4.4 we can clearly notice that the first two eigenvalues can represent approximately 99% of the total data. They are suitable for fault detection but not ideal for clustering so we take the first three principal components and use corresponding scores as parameters for further analysis.

4.6.2 Experiments for Fault identification

Experiment 1: In an experiment a person walks and runs with a $K4b^2$ in a time period and real data are acquired for indoor exercising monitoring. We select some validated data to illustrate the method. If we do not separate two processes and evaluate all samples together with PCA, the result of excessive faults is achieved and shown in figure 4.6 which incorrectly interpret collected data because PCA does not get the right correlation along different periods. The green line value is δ , which is calculated by upper limit of statistic from equation 3.9 and used as the threshold for abnormal data in figure 4.6, 4.8, 4.9, 4.10. The red lines represent the actual value of square prediction error (SPE) acquired by equation 3.8. Now when SPE $> \delta_{\alpha}^{2}$, it means classical PCA detect excessive fault due to the status transition. Therefore simple and signal process of PCA cannot detect faults correctly.

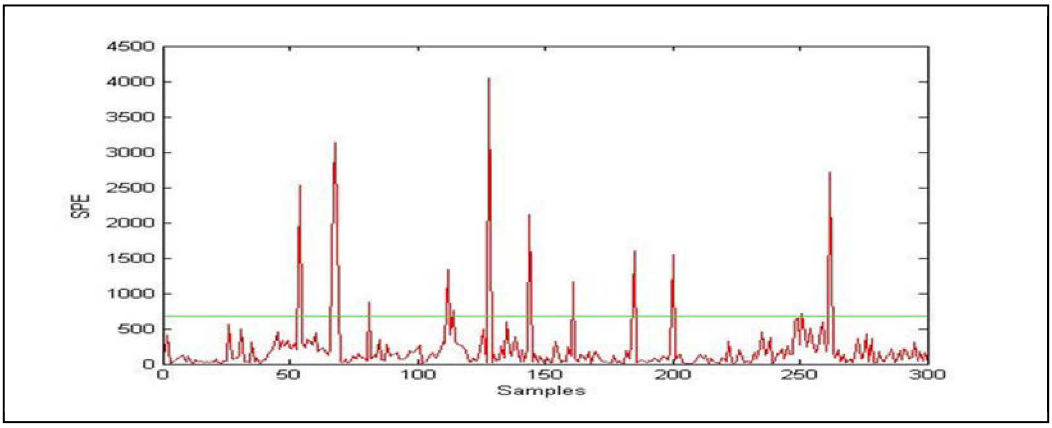


Figure 4.6 Excessive faults detected by SPE

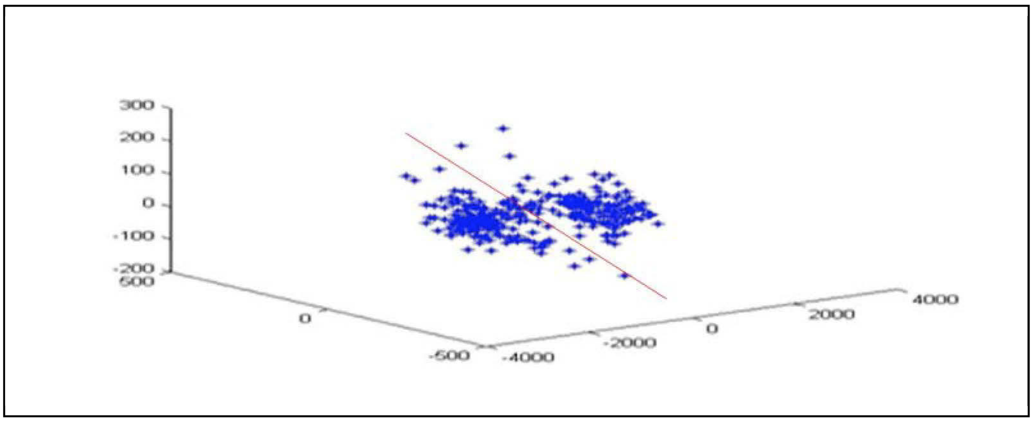


Figure 4.7 Three principal components of samples

Then we use K-means to classify data with three principal components as shown in figure 4.7. After PCA processing the time of calculation can be reduced dramatically. The figure shows samples as shown in blue spots can be classified into two groups at the 130th spot which is corresponded to the real scenarios in our experiments.

After that we divide data into two isolated processes and implement the fault identification with distinct SPE again. These data in each process are detected normal as shown in figure 4.8 and figure 4.9 respectively because they have different properties of correlation in different activities. Therefore using clustering is necessary for faults detection in multi-processes monitoring.

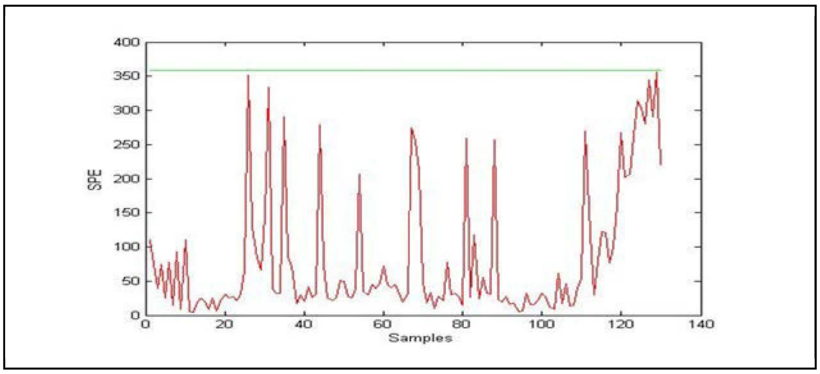
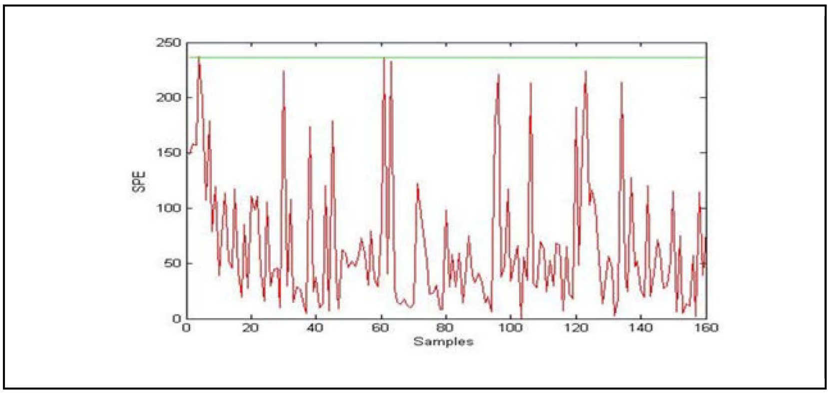


Figure 4.8 The first process of SPE detection



4.9 The second process of SPE detection

Experiment 2: Then we apply this FDI method in new samples, which contains right data and wrong data in the specific period. Firstly data are classified to different process they belong to with K-means based on reference patterns. As we see in figure 4.10 SPE increases abruptly and keeps this status from the 67th sampling point according to the calculated threshold. It means faults occur at that time and are permanent errors in this process. This result is consistent with the pre-designed condition in real situation. It shows faults are correctly detected and the detection accuracy is guaranteed.

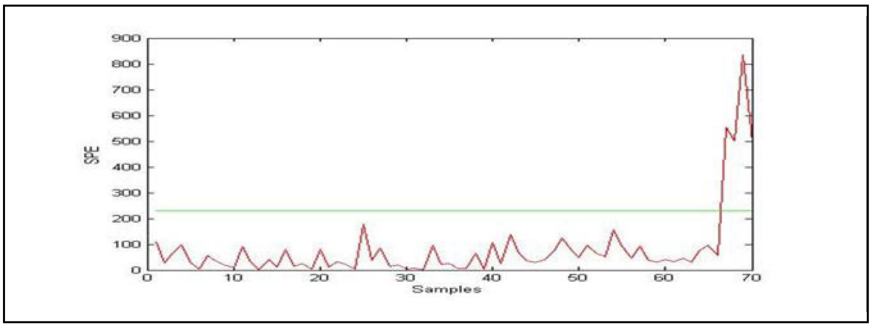


Figure 4.10 Detection by SPE with continuous error

4.6.3 Data Prediction and Faults Reconstruction

Usually detecting faults is not enough for reliability and furthermore it is necessary and essential to find which sensors break or crash if permanent faults happen. In order to identify fault sensors we construct a BP neural network for data prediction and faults identification. The network has one input-layer with 10 neurons (which mean ten time serial sampled data), one hidden-layer neurons with 12 neurons, and one output-layer with a neuron as shown in figure 4.11. The transfer function of the hidden layer is sigmoid and output layer is linear respectively. The following steps are executed to train for reliable prediction:

(1) Train the neural network using 10 continuous normal samples as the input data and the corresponding expected output data.

(2) Calculate the error of the network by given corresponding output data. Then, propagate the error backward to the input layer. In this error back-propagation process, the weights on connections are changed to reduce the error. These steps will be repeated until the network error is very small.

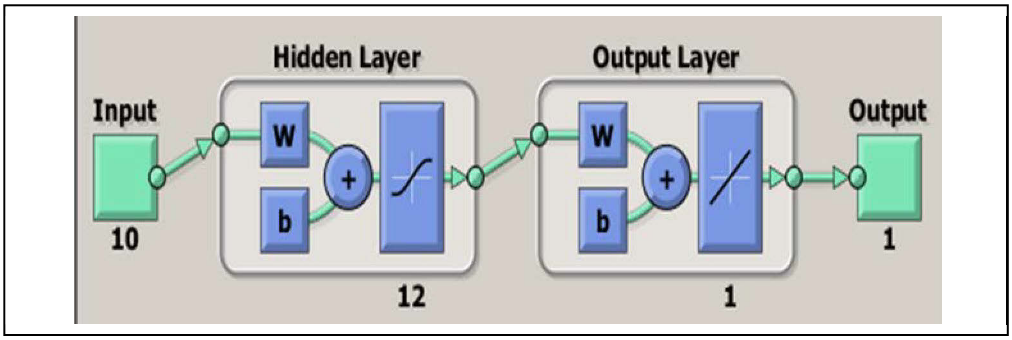


Figure 4.11 Trained neural network

In the training process, we input real time values, which come from 10 consistent steps before the fault sampling point. A predicated value is generated with the trained BP neural network for fault reconstruction. We compare the predicted data with the fault samples and recognize that O_2 related parameters such as Vo2, Feo2 are abnormal as illustrate in figure 4.12 with different symbols for measured data and predicated data. (Numbers in figure 4.12 represents 22 measured parameters.)These parameters are collected by the oxygen sensor in K4b^2. While we use the reconstruction data to replace the detected values, the faulty signs disappear as shown in figure 4.13. Therefore this sensor is indicated as malfunction from the 67th sampling moment. If the fault is transient and recover quickly, we assume it is temporally interfered. If the errors are continuous and several related values are all bias we can accurately determine the faulty sensor and isolate it consequently. At the same time the alarm information is set for further action.

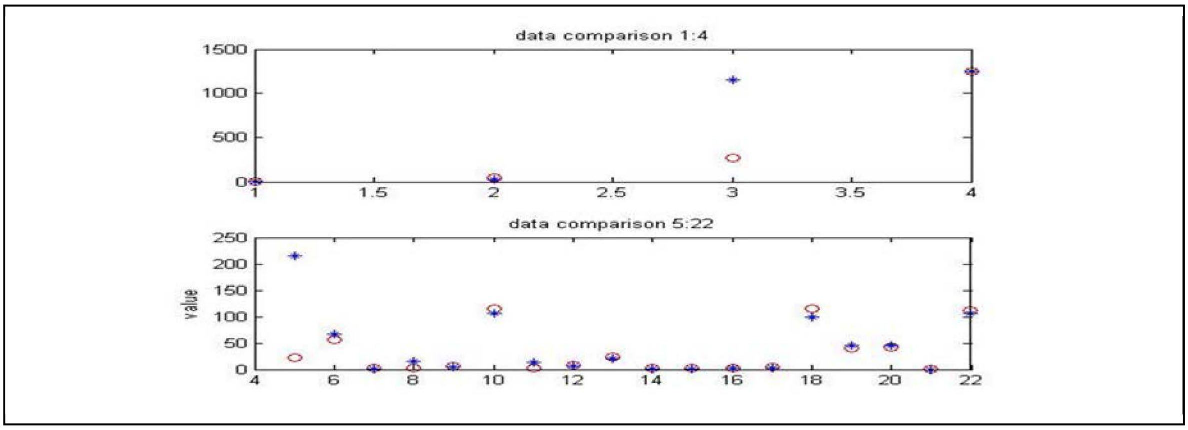


Figure 4.12 Comparison of measured and predicated values

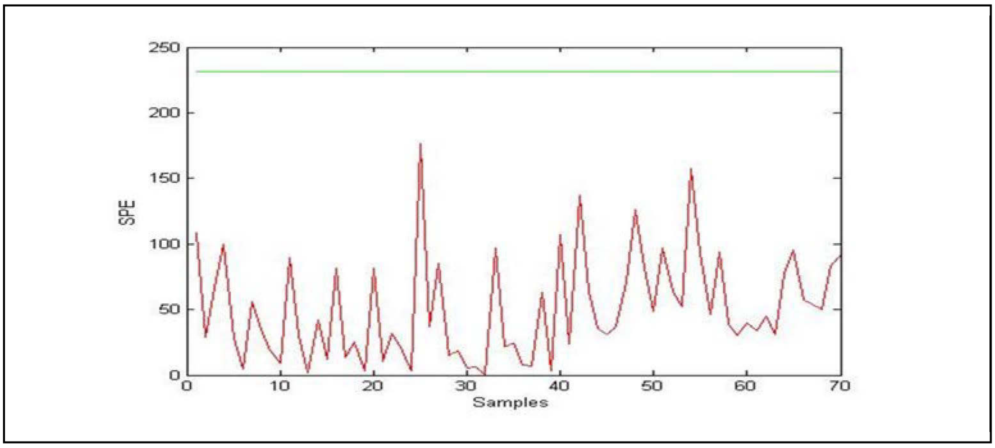


Figure 4.13 Detection by SPE statistic with reconstruction

4.7 Conclusions

A wealth of reports of using COSMED K4b^2 to measure energy expenditure in exercises and its significance has been researched and published. While if this system is applied in multiple processes the FDI process may be complicated. In order to guarantee the accuracy and availability, the reliability is essential in physical activity, sport performance and especially in some life critical situations. When monitored objects are performing in various statuses or activities classical PCA cannot capture correct correlation information from data to detect faults. So we design a comprehensive fault detection and identification procedure based on statistical techniques of PCA, clustering of K-means and machine learning of BP neural network in multiple processes.

This mechanism is verified in experiments and proved to be effective and practical for faults occurring in multiple processes. It can be a generic detection model and easily implemented in sensors with significant features such as dimension reducing for less energy consumption in transmission, clustering for fault detect in different processes and identification of crashed sensors. Due to these advantages, this comprehensive scheme can be widespread in body area networks or other systems which produce large data and perform in various processes. Therefore it will contribute to our future research to enhance the reliability of body area networks.

Chapter 5 Priority based Reliable Medium Access Mechanism in WBAN

Wireless Body Area Network is a novel technology to provide efficient, convenient and safe health, medical and personal entertainment services but it still faces longstanding challenges in the application. Current personal area networks (PANs) do not meet the medical (proximity to human tissue) and relevant communication regulations for some application environments. They also do not support the combination of reliability, QoS, low power, data rate, and non-interference required to broadly address the breadth of body area network (BAN) applications. A smart communication standard is essential to solve some problems in wireless data transmission and optimize system performance.

An effective fault prevention protocol can reduce the occurrence of faults and improve the reliability. Some techniques can be used to prevent network faults such as collision-avoidance, robust data dissemination and congestion prevention. A reliable medium access mechanism is proposed for an effective scheme to avoid faults and increase the efficiency of information delivery.

5.1 WBAN wireless channel

A WBAN uses wireless channel to provide continuous gathering of physiological, behavioural or other health related parameters and forward them to professionals for analysis. Compared with ultra wideband and human body conduction systems, narrowband wireless communication is generally accepted by the greatest number of healthcare applications. At channel bandwidths typical of narrowband BAN systems, the radio channel has been shown to be essentially slow and flat-fading, with an insignificant amount of inter symbol interference from multi-path. That said the movement of the human body has a dramatic effect on the strength of the received signal. Usually, a slow changing channel means that we can make predictions at the MAC level about the channel's future state and use this knowledge to improve the reliability.

WBAN is realized and widely applied in many fields due to recently advanced technologies in microelectronic integrated circuits, wireless communication, data digging and so on. In some innovative and significant utilization such as healthcare, sports, entertainment and military, WBAN is convenient and economical through invasive or non-invasive mini gadgets. However special concerns, and requirements encountered in the application hinder the wide spread usage in dedicated areas. For example different data for health or medical information have different sampling frequency and transmission requirements. Noncritical signs or non-real-time messages are only required to be delivered with the best effort. While critical vital signs, or real-time messages should be transmitted with least delay, and they should not be collided.

Particularly for people in some critical or emergency medical status, highest reliability, shortest access and transmission latency and immediate response are required. The losing of life-critical signals to professionals could eventually lead to life threatening consequences if patients are supervised for urgent treatments. These special features distinguish WBAN to conventional WSN in capability and induce many new issues in the process of utilization. Adaptation should be considered to satisfy network conditions such as alteration the amount of nodes or data rate. In addition effective and feasible schemes of flexible bandwidth request are required to solve constrains of limited capabilities in traditional designed.

To guarantee reliability of diversities data transmission, in recent years, a wealth of research papers are inspired to propose coordination or optimization method on PHY or MAC layers to realize a more reliable and efficient communication environment. Pangun Park derived an adaptive algorithm to formalize the power consumption while guaranteeing reliability and delay constraints in the packet transmission[63]. Yan Zhang used application specific control channels in MAC to provide priority Guarantee [64]. H. Ozgur Sanli and Hasan Çam proposed a practical coverage and rate allocation protocol to exploit this dependency in realistic environments[65]. G. Wu considered resource reservation based on priority queue but it serves sensors firstly when they require less bandwidth reservation with the same priority[18]. These protocols considered more of the power consumption and efficiency of transmission in BAN while the influence of priority is paid less attention in applications of health monitoring. Or the priority is considered but cannot ensure enough bandwidth and first services for emergency information. In fact appropriate priority assignment and bandwidth scheduling schemes are very significant to realize the reliability in applications.

A novel scheme dynamically changes the super frame pattern to schedule channel according to priority requirements of nodes. An optimized solution based on IEEE 802.15.4 standard is introduced in frame structure by allocating slots to ensure reliable transmission with low latency, collision avoidance and optimal bandwidth usage. The performance of the scheme is evaluated with simulations in Castalia-3.2 and OMNet++ 4.22.

5.2 Overview of IEEE 802.15.4

5.2.1 General description

The IEEE 802.15.4 standard specifies the physical layer and the MAC layer to convey information over relatively short distances. It allows small, power-efficient, inexpensive solutions to be implemented for a wide range of devices. Some of the capabilities provided by this standard are as follows:

- Star or peer-to-peer operation
- Unique 64-bit extended address or allocated 16-bit short address
- Optional allocation of guaranteed time slots (GTSs)
- Carrier sense multiple access with collision avoidance (CSMA-CA) or ALOHA channel access
- Fully acknowledged protocol for transfer reliability
- Low power consumption
- Energy detection (ED)
- Link quality indication (LQI)

Two different device types can participate in an IEEE 802.15.4 network: a full-function device (FFD) and a reduced-function device (RFD). An FFD is a device that is capable of serving as a coordinator. An RFD is a device that is not capable of serving as a coordinator. An RFD is intended for applications that are extremely simple, such as a light switch or a passive infrared sensor; it does not have the need to send large amounts of data and only associates with a single FFD at a time. Consequently, the RFD can be implemented using minimal resources and memory capacity.

A well-defined coverage area does not exist for wireless media because propagation characteristics are dynamic and uncertain. Small changes in position or direction often result in drastic differences in the signal strength or quality of the communication link. These effects occur whether a device is stationary or mobile, as moving objects affect station-to-station propagation.

5.2.2 Network topologies

Depending on the application requirements, an IEEE 802.15.4 wireless network operates in either of two topologies: the star topology or the peer-to-peer topology. Both are shown in Figure 5.1.

(1) Star network formation

In the star topology, the communication is established between devices and a single central controller, which is called as the coordinator. A device typically has some associated applications and is either the initiation point or the termination point for network communications. A coordinator can also have a specific application, but it can be used to initiate, terminate, or route communication around the network. The coordinator is the primary controller of the network. All devices operating on a network of either topology have unique address referred to as extended addresses. A device will use either the extended address for direct communication within the network or the short address that was allocated by the coordinator when the device is associated. The coordinator will often be mainly powered, while the devices will most likely be battery powered. Applications that benefit from a star topology include home automation, personal computer (PC) peripherals, games, and personal health care. When an FFD is activated, it can establish its own network and become the coordinator. All star networks operate independently from all other star networks currently in operation. This is achieved by choosing an identifier that is not currently used by any other network within the radio communications range. Once the identifier is chosen, the coordinator allows other devices, potentially both FFDs and RFDs to join its network.

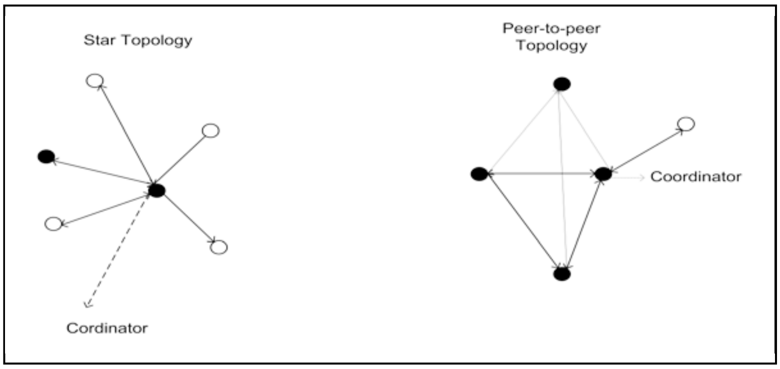


Figure 5.1 Star and Peer-to-Peer Topology

(2) Peer-to-peer network formation

In a peer-to-peer topology, each device is capable of communicating with any other device within its radio communications range. One device is nominated as the PAN coordinator, for instance, by virtue of being the first device to communicate on the channel. Further network structures are constructed out of the peer-to-peer topology, and it is possible to impose topological restrictions on the formation of the network.

The peer-to-peer topology also has a coordinator; however, it differs from the star topology in that any device is able to communicate with any other device as long as they are in range of one another. Peer-to-peer topology allows more complex network formations to be implemented, such as mesh networking topology.

Applications such as industrial control and monitoring, wireless sensor networks, asset and inventory tracking, intelligent agriculture, and security would benefit from such a network topology. A peer-to-peer network allows multiple hops to route messages from any device to any other device on the network. Such functions can be added at the higher layer, but they are not part of this standard.

Each independent network selects a unique identifier. This identifier allows communication between devices within a network using short addresses and enables transmissions between devices across independent networks. The mechanism by which identifiers are chosen is outside the scope of this standard. The network formation is performed by the higher layer, which is not part of this standard.

5.2.3 Architecture

The IEEE 802.15.4 architecture is defined in terms of a number of blocks in order to simplify the standard. These blocks are called layers. Each layer is responsible for one part of the standard and offers services. The interfaces between the layers serve to define the logical links that are described in this standard. A device comprises at least one PHY, which contains the radio frequency (RF) transceiver along with its low-level control mechanism, and a MAC sub-layer that provides access to the physical channel for all types of transfer.

(1) Physical layer (PHY)

The PHY provides two services: the PHY data service and the PHY management service. The PHY data service enables the transmission and reception of PHY protocol data units (PPDUs) across the physical radio channel. The features of the PHY are activation and deactivation of the radio transceiver, ED, LQI, channel selection, clear channel assessment (CCA), and transmitting as well as receiving packets across the physical medium.

(2) MAC sublayer

The MAC sublayer provides two services: the MAC data service and the MAC management service interfacing to the MAC sublayer management entity (MLME) service access point (SAP) (known as MLME-SAP). The MAC data service enables the transmission and reception of MAC protocol data units (MPDUs) across the PHY data service.

The features of the MAC sublayer are beacon management, channel access, GTS management, frame validation, acknowledged frame delivery, association, and disassociation. In addition, the MAC sublayer provides hooks for implementing application-appropriate security mechanisms.

5.2.4 Superframe structure

This standard allows the optional use of a superframe structure. The format of the superframe is defined by the coordinator. The superframe is bounded by network beacons sent by the coordinator and is divided into 16 slots of equal duration. Optionally, the superframe can have an active and an inactive portion. During the inactive portion, the coordinator is able to enter a low-power mode. The beacon frame transmission starts at the beginning of the first slot of each superframe.

If a coordinator does not wish to use a superframe structure, it will turn off the beacon transmissions. The beacons are used to synchronize the attached devices, to identify the network, and to describe the structure of superframes.

Any device wishing to communicate during the contention access period (CAP) between two beacons competes with other devices using a slotted CSMA-CA or ALOHA mechanism, as appropriate. For low-latency applications or applications requiring specific data bandwidth, the coordinator dedicates portions of the active superframe to that application. These portions are called guaranteed time slots (GTSs).The GTSs form the contention-free period (CFP), which always appears at the end of the active superframe and starts at a slot boundary immediately following the CAP. The PAN coordinator allocates up to seven of these GTSs, and a GTS is allowed to occupy more than one slot period. However, a sufficient portion of the CAP remains for contention-based access of other network devices or new devices to join the network. All contention-based transactions are completed before the beginning of CFP. Also each device transmitting in a GTS ensures that its transaction is completed before the time of the next GTS or the end of the CFP.

5.2.5 Data transfer model

Three types of data transfer transactions exist. The first one is the data transfer to a coordinator in which a device transmits the data. The second transaction is the data transfer from a coordinator in which the device receives the data. The third transaction is the data transfer between two peer devices. In star topology, only two of these transactions are used because data is exchanged only between the coordinator and a device. In a peer-to-peer topology, data is exchanged between any two devices on the network; consequently all three transactions are used in this topology.

The mechanisms for each transfer type depend on whether the network supports the transmission of periodic beacons. A beacon-enabled network is used in networks that either require synchronization or support for low latency devices, such as PC peripherals. If the network does not need synchronization or support for low latency devices, it can elect not to use the beacon for normal transfers. However, the beacon is still required for network discovery.

(1) Data transfer to a coordinator

When a device wishes to transfer data to a coordinator in a beacon-enabled PAN, it firstly listens for the network beacon. When the beacon is found, the device synchronizes to the superframe structure. At the appropriate time, the device transmits its data frame to the coordinator. The coordinator will acknowledge the successful reception of the data by transmitting an acknowledgment frame, if requested. When a device wishes to transfer data in a non beacon-enabled network, it simply transmits its data frame to the coordinator. The coordinator acknowledges the successful reception of the data by transmitting an optional acknowledgment frame, completing the transaction.

(2) Data transfer from a coordinator

When the coordinator wishes to transfer data to a device in a beacon-enabled PAN, it indicates in the network beacon that the data message is pending. The device periodically listens to the network beacon and, if a message is pending, transmits a MAC command requesting the data. The coordinator acknowledges the successful reception of the data request by transmitting an acknowledgment frame. The pending data frame is then sent by the coordinator. The device acknowledges the successful reception of the data by transmitting an acknowledgment frame, if requested. The transaction is now completed. Upon successful completion of the data transaction, the message is removed from the list of pending messages in the beacon.

When a coordinator wishes to transfer data to a device in a non beacon-enabled network, it stores the data for the appropriate device to make contact and request the data. A device requests data by transmitting a MAC command requesting the data to its coordinator. The coordinator acknowledges the successful reception of the data request by transmitting an acknowledgment frame. If a data frame is pending, the coordinator transmits the data frame. If a data frame is not pending, the coordinator indicates this fact either in the acknowledgment frame following the data request or in a data frame with a zero-length payload. If requested, the device acknowledges the successful reception of the data frame by transmitting an acknowledgment frame.

(3)Peer-to-peer data transfers

In a peer-to-peer network, every device communicates directly with every other device in its radio communications range. In order to do this effectively, the devices wishing to communicate will need to either receive constantly or synchronize with each other. In the former case, the device can simply transmit its data. In the latter case, other measures need to be taken in order to achieve synchronization.

5.2.6 Frame structure

The frame structures have been designed to keep the complexity to a minimum while at the same time making them sufficiently robust for transmission on a noisy channel. Each successive protocol layer adds to the structure with layer-specific headers and footers. The MAC frames are passed to the PHY as the PHY service data unit (PSDU), which becomes the PHY payload.

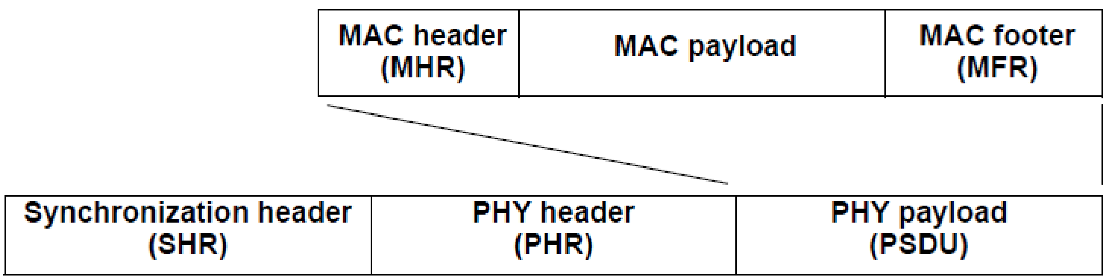


Figure 5.2 Schematic View of the PPDU

This standard defines four MAC frame structures:

- A beacon frame, used by a coordinator to transmit beacons
- A data frame, used for all transfers of data
- An acknowledgment frame, used for confirming successful frame reception

- A MAC command frame, used for handling all MAC peer entity control transfers

5.2.7 Improving probability of successful delivery

The IEEE 802.15.4 employs some mechanisms to realize the probability of successful data transmission. These mechanisms are discussed as following:

(1) CSMA-CA mechanism

The IEEE 802.15.4 uses two types of channel access mechanism, depending on the network configuration. Non beacon-enabled networks use an unslotted CSMA-CA channel access mechanism. Each time a device wishes to transmit data frames or MAC commands, it waits for a random period. If the channel is found to be idle, following the random backoff, the device transmits its data.

If the channel is found to be busy following the random backoff, the device waits for another random period before trying to access the channel again. Acknowledgment frames are sent without using a CSMA-CA mechanism. Beacon-enabled networks use a slotted CSMA-CA channel access mechanism, where the backoff periods are aligned with the start of the beacon transmission. The backoff periods of all devices within one network are aligned to the coordinator. Each time a device wishes to transmit data frames during the CAP, it locates the boundary of the next backoff period and then waits for a random number of backoff periods. If the channel is busy, following this random backoff, the device waits for another random number of backoff periods before trying to access the channel again. If the channel is idle, the device begins transmitting on the next available backoff period boundary. Acknowledgment and beacon frames are sent without using a CSMA-CA mechanism.

The CSMA-CA algorithm shall be used before the transmission of data or MAC command frames transmitted within the CAP, unless the frame can be quickly transmitted following the acknowledgment of a data request command. The CSMA-CA algorithm shall not be used for the transmission of beacon frames in a beacon-enabled network, acknowledgment frames, or data frames transmitted in the CFP.

If periodic beacons are used, the MAC sublayer shall employ the slotted version of the CSMA-CA algorithm for transmissions in the CAP of the superframe. Conversely, if periodic beacons are not being used or if a beacon could not be located in a beacon-enabled network, the MAC sublayer shall transmit packets with the unslotted version of the CSMA-CA algorithm. In both cases, the algorithm is implemented with units of time called backoff periods, where one backoff period shall be equal to a UnitBackoff Period.

In slotted CSMA-CA, the backoff period boundaries of every device in the PAN shall be aligned with the superframe slot boundaries of the PAN coordinator; i.e., the start of the first backoff period of each device is aligned with the start of the beacon transmission. In slotted CSMA-CA, the MAC sublayer shall ensure that the PHY commences all of its transmissions on the boundary of a backoff period. In unslotted CSMA-CA, the backoff periods of one device are not related in time to the backoff periods of any other devices.

Each device shall maintain three variables for each transmission attempt: NB, CW, and BE. NB is the number of times the CSMA-CA algorithm is required to back off while attempting the current transmission; this value shall be initialized to zero before each new transmission attempt. CW is the contention window length, defining the number of backoff periods that need to be clear of channel activity before the transmission can commence. This value shall be initialized to CW_0 before each transmission attempt and reset to CW_0 each time when the channel is assessed to be busy. For operation in the Japanese 950 MHz band, CW_0 shall be set to one; otherwise, CW_0 shall be set to two. The CW variable is only used for slotted CSMA-CA. BE is the backoff exponent, which is related to how many backoff periods a device shall wait before attempting to assess a channel. In unslotted systems, or slotted systems with the received battery life extension (BLE) field, set to zero, BE shall be initialized to the value of macMinBE. In slotted systems with the received BLE field set to one, this value shall be initialized to the lesser of two and the value of macMinBE. Note that if macMinBE is set to zero, collision avoidance will be disabled during the first iteration of this algorithm. Although the receiver of the device is enabled during the CCA analysis portion of this algorithm, the device may discard any frames received during this time.

Figure 5.3 illustrates the steps of the CSMA-CA algorithm. If the algorithm ends in “Success,” the MAC is allowed to begin transmission of the frame. Otherwise, the algorithm terminates with a channel access failure. In a slotted CSMA-CA system with the BLE field set to zero, the MAC sublayer shall ensure that, after the random backoff, the remaining CSMA-CA operations can be undertaken and the entire transaction can be transmitted before the end of the CAP. If the number of backoff periods is greater than the remaining number of backoff periods in the CAP, the MAC sublayer shall pause the backoff countdown at the end of the CAP and resume it at the start of the CAP in the next superframe. If the number of backoff periods is less than or equal to the remaining number of backoff periods in the CAP, the MAC sublayer shall apply its backoff delay and then evaluate whether it can proceed. The MAC sublayer shall proceed if the remaining CSMA-CA algorithm steps, the frame transmission, and any acknowledgment can be completed before the end of the CAP. If the MAC sublayer can proceed, it shall request that the PHY perform the CCA in the current superframe. If the MAC sublayer cannot proceed, it shall wait until the start of the CAP in the next superframe and apply a further random backoff delay before evaluating whether it can proceed again.

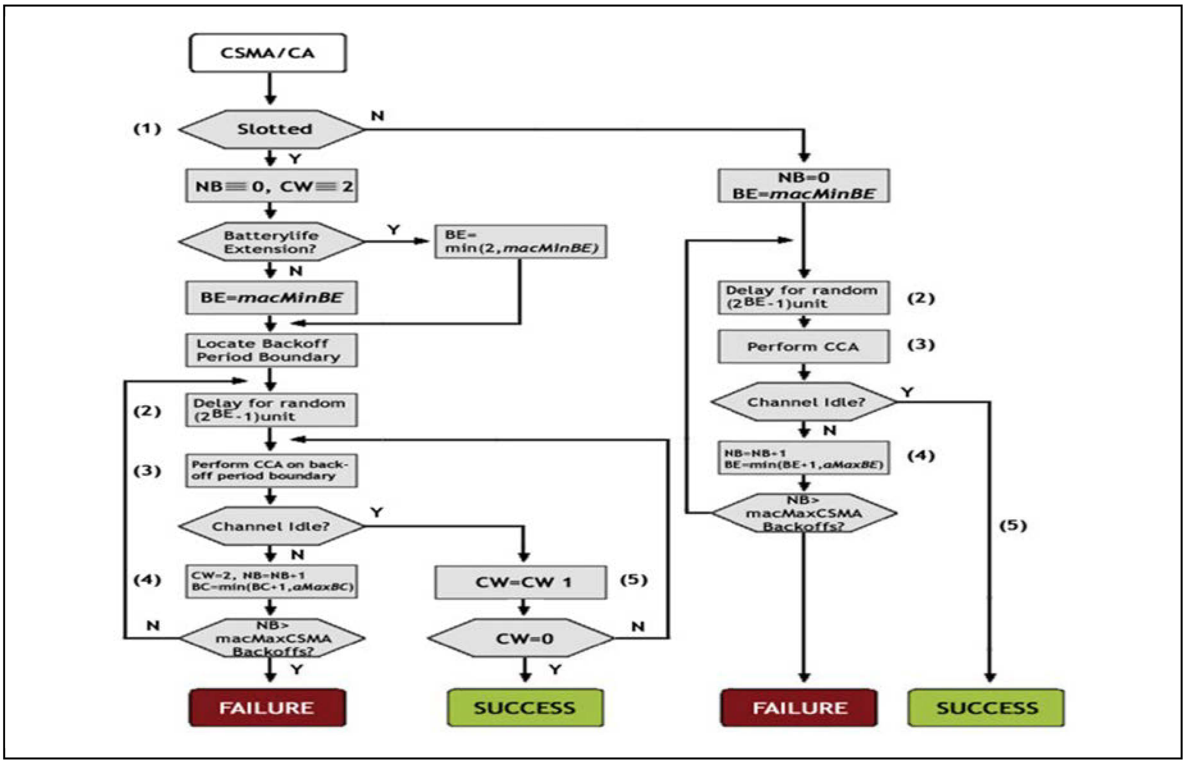


Figure 5.3 CSMA-CA Algorithms

(2) ALOHA mechanism

An ALOHA channel provides access to a common communication channel from multiple independent packet transmitters by the simplest mechanisms. When each transmitter is ready to transmit the packet, it simply transmits the packet burst without any coordination with other transmitters using the shared channel. If each user of the ALOHA channel is required to have a low duty cycle, the probability of a packet from one user overlapping and thus interfering with a packet from another user is small as long as the total number of users on the shared ALOHA channel is not too large. As the number of users on the shared ALOHA channel increases, the number of packet overlaps increase and the probability that a packet will be lost due to an overlap with another packet on the same channel also increases.

In the ALOHA protocol, a device transmits without sensing the medium or waiting for a specific time slot. The ALOHA mechanism is appropriate for lightly loaded networks since the probability of collision is reasonably small if the probability of clear channel is sufficiently large.

(3) Frame acknowledgment

A successful reception and validation of a data or MAC command frame is optionally confirmed with an acknowledgment. If the receiving device is unable to handle the received data frame for any reason, the message is not acknowledged. If the originator does not receive an acknowledgment after some periods, it assumes that the transmission was unsuccessful and retries the frame transmission. If an acknowledgment is still not received after several retries, the originator can choose either to terminate the transaction or to try again. When the acknowledgment is not required, the originator assumes the transmission was successful.

(4) Data verification

A cyclic redundancy check (CRC) is used to detect errors in every PSDU. A device with CRC calculates a short, fixed-length binary sequence which is appended to the original data for each block to be sent or stored. This binary sequence is called as a codeword and the device receives the codeword and compares the resulting check value with an expected residue constant. If these values do not match, then it means a data error. Then rereading or requesting can be done. Otherwise, the data is assumed to be correct.

5.2.8 MAC function and superframe of beacon enabled mode

(1) MAC function

MAC sublayer handles all access to the physical radio channel and is responsible for the following tasks:

- Generating network beacons if the device is a coordinator
- Synchronizing to network beacons
- Supporting association and disassociation
- Supporting device security
- Providing a reliable link between two peer MAC entities
- Frame delimiting and recognition
- Addressing of destination stations (both as individual stations and as groups of stations)

- Conveyance of source-station addressing information
- Transparent data transfer of LLC PDUs
- Protection against errors, generally by means of generating and checking frame check sequences
- Control of access to the physical transmission medium

(2) Superframe structure of beacon enabled mode

For WBANs, a beacon enable mode is widely accepted. A coordinator can optionally bind its channel time using a superframe structure. A superframe is bounded by the transmission of a beacon frame and can have an active portion and an inactive portion. The coordinator regularly forwards the beacon frames in each beacon interval. The coordinator and sampling sensors communicate in duplex during active period, named the superframe duration SD, and go into the low-power status within the inactive period. The structure of the superframe is determined by two parameters, the beacon order BO and the superframe order SO, which determine the length of the superframe and its active period, respectively. The range of the superframe and the range of its active period can be defined as

$$BI = aBaseSlotDuration \times 2^{BO} \tag{5.1}$$

$$SD = aBaseSlotDuration \times 2^{SO} \tag{5.2}$$

where $0 \leq SO \leq BO \leq 14$.

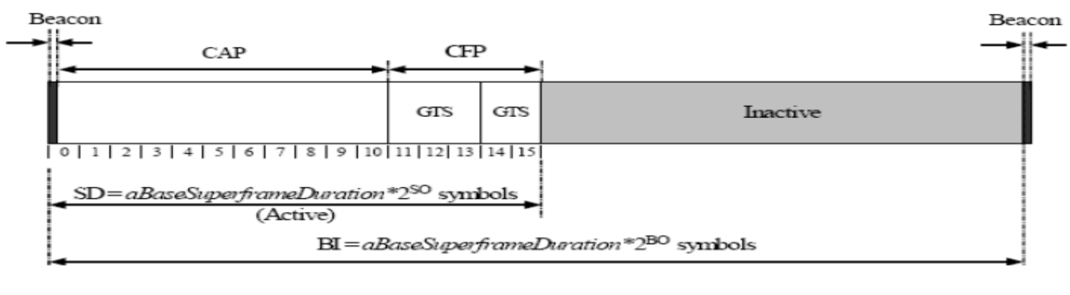


Figure 5.4 Superframe structure of beacon-enabled model

A BaseSuperframeDuration is the number of symbols constituting a basic superframe when SO is equal to 0. In addition, the superframe is divided into 16 equally sized slots. Figure 5.4 shows the superframe structure of IEEE 802.15.4 beacon-enabled mode. The superframe consists three components: a beacon, a contention access period (CAP), and a contention free period (CFP). The beacon is transmitted in Slot 0 without CSMA, and the CAP start immediately after the beacon.

The CFP, if presents, commences after the CAP to the end of the active period. Any allocated GTSs should be located within the CFP. All frames transmitted in the CAP shall use a slotted CSMA-CA mechanism to access the channel. MAC command frames shall always be transmitted in the CAP.

The CFP will start on a slot boundary immediately following the CAP, and it will finish before the end of the active period. If any GTSs have been allocated by the PAN coordinator, they will be located within the CFP and occupy contiguous slots. The CFP duration will increase or decrease depending on the total length of all of calculated GTSs.

5.2.9 Limitations of IEEE802.15.4 in WBANs

Although IEEE802.15.4 is used in short-range wireless systems it has several limitations which contain statically bandwidth utilization of the messages, back off period, and data length in the application of WBANs. The standard is not flexible managed because BO and SO are decided by the network in initialization. The system cannot modify configuration in run time therefore it is unable to deal with changing requirements of messages. Under some situations, fixed setting will lead to the wastage of energy and slots. For example, all the slots' length is equal when BO and DO are set. When assigned slots duration meets demand of nodes with real time information, part of this period for nodes with no real time information is wasted. Also, GTSs' amount is constrained up to 7 in the CFP since a minimum CAP of length 8 slots (aMinCAPLength) has to be kept in each superframe to maintain MAC commands could still be delivered to nodes when GTSs are being used. So CFP can accommodate only a few real time messages in high speed. If nodes require more bandwidth to send high priority data, they won't be allocated enough bandwidth. Then they have to wait with long time and would eventually result data loss. In order to overcome these disadvantages in IEEE 802.15.4 some modifications are considered in our reliable medium access mechanism (RMAM) to achieve automatically reconfiguration of superframe structure.

5.3 Structure of RMAM system model

According to the specific circumstances in WBAN the configuration of RMAM is operated by the coordinator. First of all a device, such as a PAD, a laptop or computer, is selected as a coordinator which is in charge of receiving data from other nodes, analysing, setting priorities, allocating bandwidth and transferring to the service centre. When abnormal or emergency data are detected, the coordinator changes initial priorities and allocate bandwidth to adapt to new circumstances.

Then some sensors change data transmission rates with adjusted allocation to ensure high priority with high reliability. A typical architecture of the system is divided into three logical function modules: Information collection module, priority assign module and bandwidth allocation module. Figure 5.5 illustrates the structure of the RMAM system.

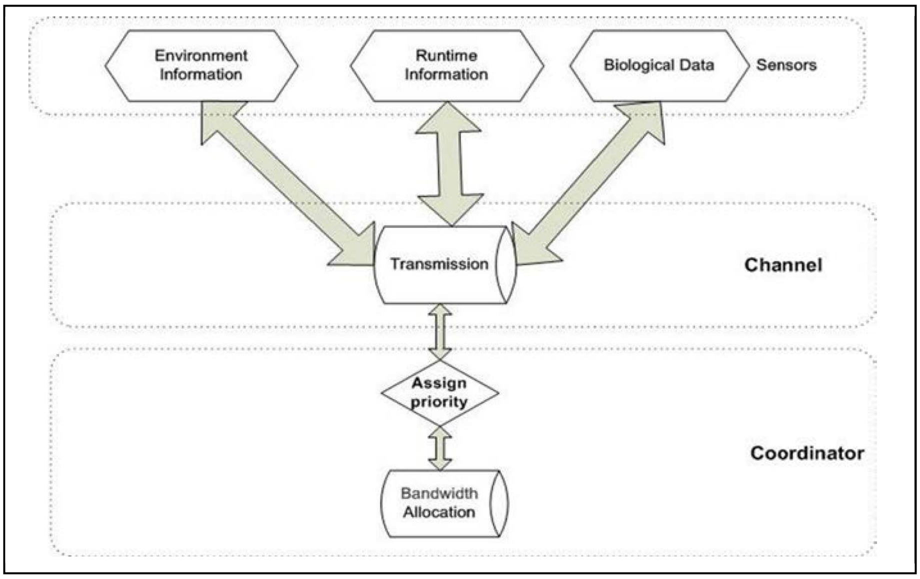


Figure 5. 5 Structure of RMAM

5.3.1 Information collection module

The information collection module consists of nodes which collect environmental, biological and runtime information, such as blood pressure, body temperature, heart rate, SpO2, environment temperature and so on. The sampling frequencies of biomedical sensors usually are no more than 300 Hz since biological signals oscillate at low frequencies. But if it uses a large array of sensors, e.g., EEG 192-channel recording, the data rate will increase up to hundreds of kbps[66]. Table 5.1 shows the data rate of several biomedical sensor measurements. If environment sensors send video information, high data rate is required. Thus these sensors should be classified by critical or non critical nodes with original priorities in different applications. Data acquired by a variety of sensors go through wireless channel to coordinator in WBAN.

Table 5. 1 Sampling frequency of biomedical sensors

Sensors	Sampling frequency (Hz)	Basic Data Rate
ECG 1-point	250	6 kbit/s
ECG 12-point	250	72 kbit/s

EEG 1-channel	200	4.8 kbit/s
EEG 192-channel	200	921.6 kbit/s
Blood pressure	60	1.44 kbit/s
SpO2	300	7.2 kbit
Body temperature	0.1	2.4 bit/s

5.3.2 Priority assign module

A priority assign module is executed by the coordinator. It receives data, analyses and assigns priority. The data classification refers to the parameters collected by various sensors defined as different types. Some data are very critical under certain situation. For example: heart rate, ECG, oxygen and other indicators indicate the important real-time physiological status. Those vital data should be ensured with high end-to-end transmission reliability and lower delivery delay constraints. However non critical data are not time urgent and real time transmission is not necessary for them. So normal data can be transferred at low rate to save energy consumption while abnormal or emergency data should be transferred with high rate for further actions. In our mechanism data are classified into four types: non real- time data, normal real-time data, abnormal data, and emergency data. These four types of priorities are defined as 1,2,3,4. The coordinator can analyse and judge abnormality of information with pre-defined activation threshold or principal components analysis[67]. If abnormal or emergency data are detected, the involved nodes should send information as soon as possible. So there are three possible priorities of each node. For example the priority of nodes which transfer non real time data possibly is 1, 3 or 4 depending on the sampling data. But the priority of a node collecting real time data would be 2, 3, and 4. The priority can be set in MAC frame control reserved 7-9 bit. Each sensor has a weight β, which is allocated at initial configuration by coordinator according to sampling values to avoid collision when maximum bandwidth cannot meet the requirements of several nodes with emergency data simultaneously or slots in collision access periods. When sensors with transient abnormal data go back to normal status, the priority decreases to its original configuration. High priority will be assigned a wider bandwidth and resources and low priority will be allocated fewer resources, rather than first come first serve in general IEEE 802.15.4.

5.3.3 Bandwidth allocation module

The BAN Priority is set to indicate the priority of the services provided to the BAN of the sender. The algorithm is designed based on IEEE 802.15.4 protocol due to its widely application in WBAN. The proposed mechanism would provide the maximum reliability as much as possible to reduce the energy consumption and the throughput of the network. Emergency data node, issued by the MAC frame, shall be tagged as EN04; normal non real time data nodes are tagged as NNRN01; normal real time data nodes are tagged as NRN02; abnormal data nodes are tagged as AN03.The number indicates each node’s priority. Normal non real time data nodes use slotted CAP and real time data nodes use GTS to transmit messages. When abnormalities and emergency are detected, GTS in CFP will be allocated to meet requirements of reliability. Then coordinator sends commands to sensors that they can alter speed according to the new configuration.

5.4 A reliable medium access mechanism

IEEE 802.15.4 is a communication standard for low data rate, low power consumption and low cost Wireless Personal Area Networks for a wide range of fields. The proposed RMAM is based on IEEE 802.15.4 MAC standards with small changes. Our system operates on 2.4 GHz RF band with a star topology. A slotted CSMA/CA mechanism is used to access the channel for non-real time data and GTS requests during the CAP. In the CFP, the dedicated bandwidth is used for time critical data. Each node is a type of sensor, which requests and sends one type of message with dynamical rate in run time. All non-real time messages should be periodically dispatched in CAP period. Other real time information transmitted in GTSs allocated by coordinator with priority value. From a point of view, the GTS allocation is similar to a time division multiple access mechanism. A reserved bandwidth is previously granted for given data. Adaptive parameters of the super frame will be assigned by the determined amount of time slots. Therefore scheduling slots is feasible to ensure data reliability in our proposed method from two aspects: GTS allocation and CAP back off mechanism.

5.4.1 GTS allocation

To solve restrictions of IEEE 802.15.4, in our proposed scheme Superframe slots are not fixed to 16 and vary accordingly to priority assign module. The Superframe Duration (SD) of this protocol is adjusted by changing the SO. On the other hand, the proposed protocol changes the number of slots according to priority requirement, so that it can achieve better utilization, save energy and better reliability. The equations of modified SD and BI are listed as:

$$BI = aBaseSlotDuration \times (BO + NumSlot) \tag{5.3}$$

$$SD = aBaseSlotDuration \times NumSlot \tag{5.4}$$

NumSlot is determined by the sum of each node's required slots according to the data rate and priorities. BO should be determined by the following equation:

$$16 \times aBaseSlotDuration \le aBaseSlotDuration(BO + NumSlot) \le 16 \times aBaseSlotDuration \times 2^{14} \tag{5.5}$$

When emergency data are detected, coordinator set BO=0 to cancel inactive time until all emergency data are successfully received. If required slots exceed the maximum we allocate bandwidth by nodes' weight β:

$$W_i = W_u \frac{\beta_i}{\Sigma_{i=1}^{n} \beta_i} \tag{5.6}$$

W_i is the bandwidth allocated for the i_{th} node, W_u is the bandwidth utilized for GTS .

$$Wu = 16 \times aBaseSlotDuration \times 2^{14} - k * aBaseSlotDuration \tag{5.7}$$

where we leave k slots for CAP and $2 \le k \le 16$.

5.4.2 CAP access mechanism

The CAP allocation shall be executed when the sensor nodes which want to send normal non real time data. Nodes that operate in beacon enable mode must utilize the slotted CSMA-CA access mechanism. In this algorithm, each node with non-real time data uses different binary exponential back-off (BE) by given weight. The higher the weight is, the shorter the back off time is. Therefore high-weight nodes can access the medium with higher probability in CSMA.

The pseudo code is showed to execute the optimization process:

Step1 Packet arrives

Set NB, maxBE, minBE

wait (channel_becomes_free)

send(frame)

wait_until (end_of_frame) or (collision)

Step2 If collision detected,

Check node weight β

Set BE= Floor (minBE/ β)

Step3 r = 2BE * slotTime

Wait(r*slotTime)

perform CCA

if channel idle go to step1

else set NB=NB+1,BE=BE+1

if BE>maxBE report failure

else go to step3

else :

wait (inter-frame_delay)

5.4.3 A case analysis

Now we use an example to explain the allocation scheme in detail: there are 5 nodes assumed to transfer data, node 1 forwards normal real time data, node 2 forwards emergency data and node 3 forwards abnormal real time data and node 4 and 5 forward non real time data. Therefore node 1, 2 and 3 use GTSs in CAP by coordinator while node 4 and 5 use CAP. If the rate required for nodes are: Node 1= 50kbps, Node 2=150kbps, Node 3=100kbps, Node 4=25kbps, Node 5=25kbps. When devices work on 2.4 GHz band and the size of a symbol is 4 bits. If a 50 frames/second sampling frequency is chosen, the number of symbols per frame required by devices should be: Node1= 250symbols, Node2= 750 symbols, Node3= 500 symbols, Node4= 125symbols and Node5= 125symbols. When the size of the base slot is set to 60 symbols and now the required slots for Node1, Node2, Node3, Node4 and Node5 are 5, 13, 9, 3 and 3 respectively. Now total slots should be allocated to all nodes are NumSlot = 5+13+9+3+3= 33 and for GTS are

27. In standard IEEE 802.15.4 the number of GTS is limited to 7 and it cannot satisfy the requirement of nodes. Nodes have to wait for another cycle to complete transmission, which will cause latency of critical information and even bring life-threaten result. While in our proposed mechanism the increase and decrease of slots number can be dynamically adjusted in superframe configuration so it would achieve maximum WBAN-width utilization and minimum time latency to guarantee high reliability.

5.5 Performance analysis and simulation

5.5.1 Latency

Latency plays a very important role in real-time applications. Transmitting critical data in short time constraint is extremely critical that data should be reliably received within minimum latency. For a general star networks based on IEEE 802.15.4, if there are N nodes with real time data and one coordinator allocate 1 GTS per superframe for each node with the maximum 7 slots in CFP. It will cause a worst case queuing latency of $16 * \mathrm{Slot}_{\mathrm{Duration}} * \left\lceil \frac{\mathrm{N}}{7} \right\rceil$. Hence, the worst-case uplink latency is given by equation:

$$Latency = QLantency + Tx_{Time} = 16 \times Slot_{duration} \times \left\lceil \frac{N}{7} \right\rceil + \frac{f}{R} \quad (5.8)$$

Where f is the data frame size in bits and R is the data bit rate. Therefore latency grows linearly with the number of nodes in regular allocation and it is terrible for emergency data with waiting a long queuing. According to our proposed algorithm, if emergency data are allocated enough slots in a super frame and can be transferred without queuing latency to guarantee transmission in shortest duration.

5.5.2 Energy consume

The energy of signals consumed in wireless channel to carry information is an essential performance which depends on two parameters: transmission power and duration. The energy expenditures are different with various communication techniques. The transmission power is set by a transmitter and the energy in our system is similar to the model in reference [68] and can be defined by

$$E(R) = P_l t_l + P_{tx} \frac{f}{R} \quad (5.9)$$

Where P_l is consumed power in the period listening, t_l is the time in the period of listening and P_{tx} is the power in the period of transmitting. So it is simple to count the network's energy consumption in various rates.

5.5.3 Packet breakdown

The packet breakdown means a packet is broken after either it gets the maximum number of transmission attempts or a new frame gets to the node while it was still waiting for its transmission. The packet breakdown rate is a function of packet arrival rate λ and beacon interval BI. These parameters decide how long, on the average, a frame will have to wait for its opportunity for transmission and be possible to be dropped as new frames arrive. According to reference[69] the breakdown packets are related with the successful probability as in (3.10)

$$P_{breakdown} = P_{d0} + (1 - P_{d0})\sum_{i=1}^{n} P_{di} \qquad (5.10)$$

$$P_{d0} = 1 - \frac{\lambda BIe^{-\lambda BI}}{1 - e^{-\lambda BI}} \;, \qquad P_{di} = (1 - P_{d0})P_e^i(1 - e^{-\lambda BI})\, e^{-(i-1)\lambda BI}$$

Where $e^{-(i-1)\lambda BI}$ represents no arrivals in the previous i-1 superframes and $1 - e^{-\lambda BI}$ represents at least one arrival in superframe i. Also$(1 - P_{d0})$ is the probability that the packet is not fail in the first superframe.

5.5.4 Throughput

At the same time the coordinator detects the system throughput S (t) of the superframe, t is given by the ratio of the average length of successfully allocated payload in a GTS time slot to the average length of a GTS time slot, namely as S(t) in (3.11):

$$S(t) = \frac{P_{s(t)} L_{pl} \tau_n}{\theta_{min} T_{ss}} \qquad (5.11)$$

where Ps (t) = 1 − Pbreak (t), Lpl is the length of payload of each data packet, τ_n is the number of data packets, T_{ss} is the length of superframe slot, and θ_{min} is the minimum number of superframe slots [70].

5.6 Simulationand results analysis

5.6.1 Simulation Platform

We tested RMAM by Castalia Simulator. Castalia is a simulator for Wireless Sensor Networks (WSN), Body Area Networks (BAN) and generally networks of low-power embedded devices. It is based on the OMNeT++ platform and can be used by researchers and developers who want to test their distributed algorithms and/or protocols in realistic wireless channel and radio models, with a realistic node behaviour especially relating to access of the radio. Castalia can also be used to evaluate different platform characteristics for specific applications, since it is highly parametric, and can simulate a wide range of platforms. The main features of Castalia are:

（1）Advanced channel model based on empirically measured data：

- Model defines a map of path loss, not simply connections between nodes
- Complex model for temporal variation of path loss
- Fully supports mobility of the nodes
- Interference is handled as received signal strength, not as separate feature

（2）Advanced radio model based on real radios for low-power communication.

- Probability of reception based on SINR, packet size, modulation type.
- PSK FSK supported, custom modulation allowed by defining SNR-BER curve.
- Multiple TX power levels with individual node variations allowed
- States with different power consumption and delays switching between them
- Realistic modeling of RSSI and carrier sensing

（3）Extended sensing modelling provisions

- Highly flexible physical process model.
- Sensing device noise, bias, and power consumption.

（4）MAC and routing protocols available.

（5）Designed for adaptation and expansion.

5.6.2 Body Area Network simulation in Castalia

In Castalia, go to Simulations /BAN test and open omnetpp.ini, a lot of configuration is defined and can be adjusted to simulate a variety of scenarios to evaluate MAC in Body Area Networks. All scenarios use the throughput Test application, where all nodes send packets to a sink/hub node as a star topology is used.

For the simulation, 6 nodes are used in a system, 1 Coordinator and 5 sensor nodes on a star topology. The sensor nodes adopt CC2420 radios for a simulation time of 600 seconds with 3 repetitions. Table 5.2 explains the specifications of nodes. By this simulation model, we compare general IEEE 802.15.4 protocol with RMAM and present the performance parameters like latency, energy consumption, packet breakdown and throughput with the bandwidth allocation for priority devices.

Table 5.2 Specification of nodes

Nodes	Node1	Node2	Node3	Node4	Node5
Priority scope	2,3,4	2,3,4	2,3,4	1,3,4	1,3,4
Initial priority	2	2	2	1	1
Basic rate(Packets/s)	50	50	50	25	25
Maximum rate (Packets/s)	150	150	150	150	150
Weight	0.3	0.3	0.2	0.1	0.1

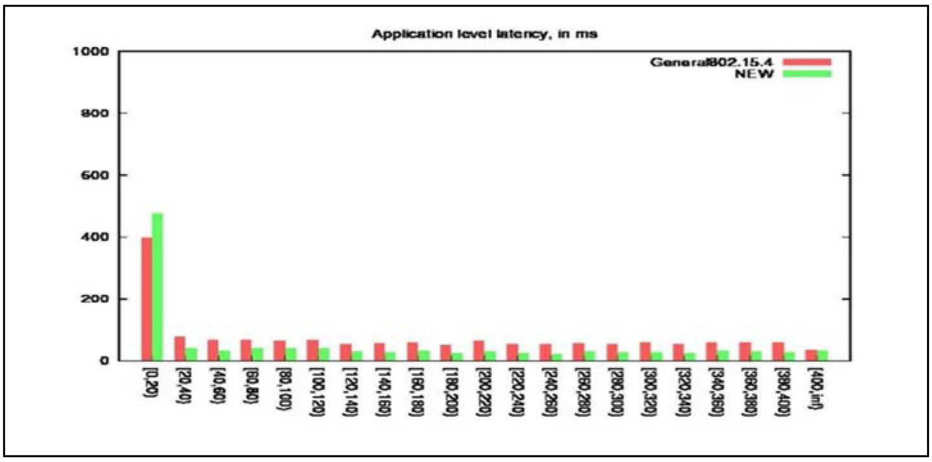


Figure 5. 6 The latency comparison

Figure 5.6 shows the comparison of latency. We can see that new mechanism is performing better in the whole process although the latency is longer in initial period due to the beacon requirement and response. The energy expenditure of each node is depicted in figure 5.7. We also see that RMAM is a more energy efficient while node 2 expends higher energy according to the traffic needs.

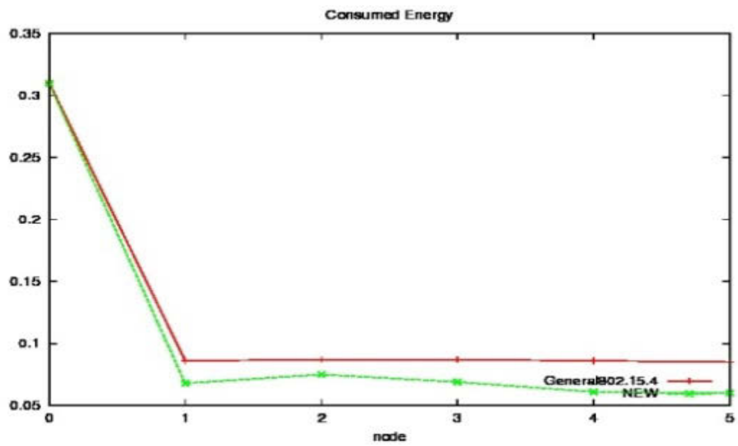


Figure 5. 7 Energy expenditure of each node

Figure 5.8 and 5.9 show the packets breakdown as fraction of 1 in two different scenarios. Looking at the figure it is easy to get an impression of how better the application is. The portion of failed packets because of no Ack or busy channel is higher especially when node is in the maximum rate while in new method the ratio decreased sharply because it is allocated enough bandwidth. The quality of service for node 2 with highest priority is guaranteed. Then we notice that the performance throughput of RMAM is better in figure 5.10 when the channel is allocated well. Finally it can be noticed that when the traffic rate reaches the larger volumes there are early indications of saturation, especially for the general IEEE802.15.4 model.

Generally the RMAM performs better compared with the general mode because it makes a more efficient use of the wireless medium and reduces collisions. This method proposes a priority based MAC protocol that meets the requirements of WBAN such as: real-time transmission, collision Avoidance, Adaptively, minimal end to end delays, high Throughput by considerable utilization of bandwidth.

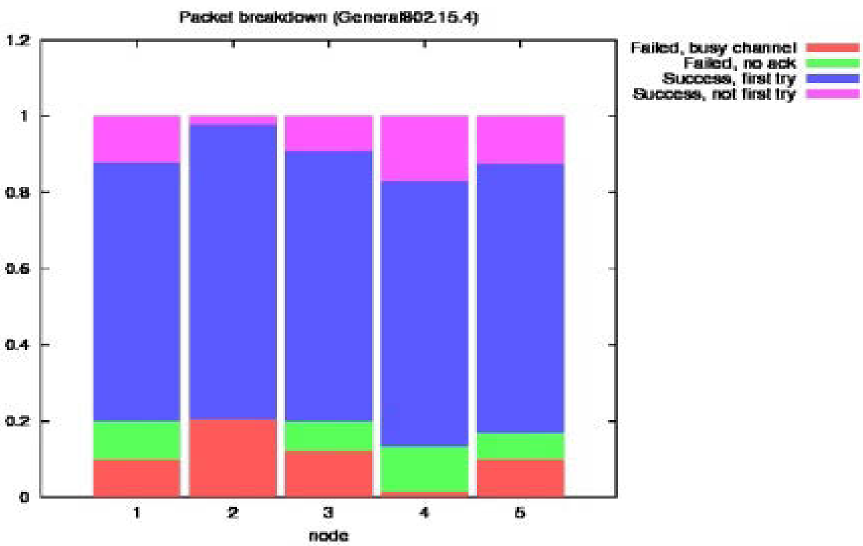


Figure 5. 8 Packet breakdowns in 802.15.4

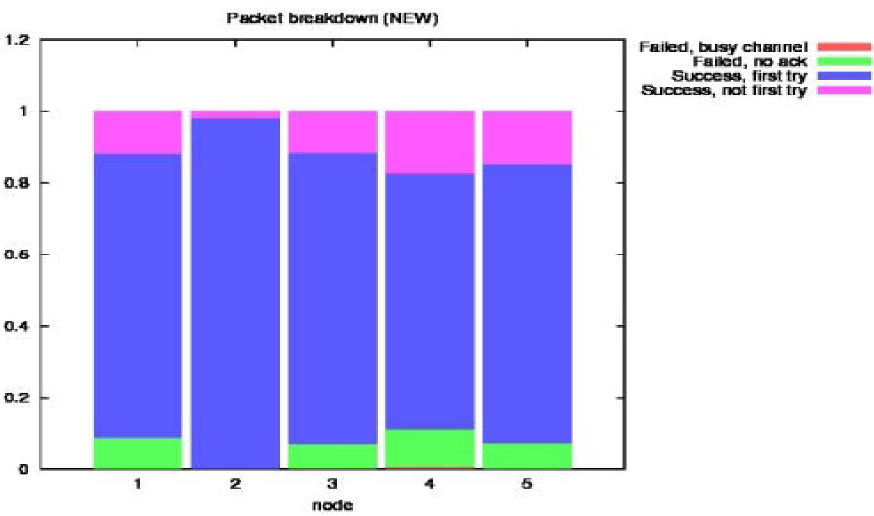


Figure 5.9 Packet breakdowns in NEW mechanism

This mechanism introduces dynamic Superframe reconfiguration in bandwidth allocation by priority to transfer signals. The validity is demonstrated by Castalia and results show it significantly decreases delay and reaches a considerably high packet delivery ratio dropped in the entire network as compared to the general IEEE 802.15.4 model. In the future characteristics of physical and application layers such as RF band, trade off and environment interference should be considered for better quality of services.

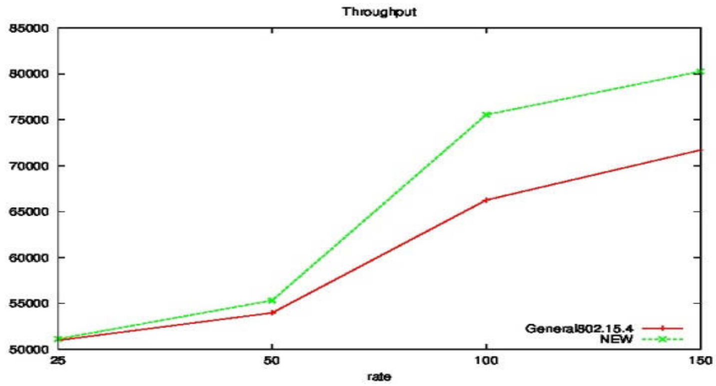


Figure 5. 10 Throughput Comparisons

5.7 Conclusions and Future Work

Reliability is a key requirement in WBANS, especially for healthcare and related areas. Recently, a lot of interest has increased in the development of a reliable system for wide range of applications. However there are many challenges in designing such a network, with consideration of various constraints including energy efficiency, low latency, high throughput and other metrics.

An automatically slots allocation model is proposed to avoid collision and prevent faults. The object of this chapter is to propose a priority based MAC protocol that meets the requirements of WBAN such as: real-time transmission, collision avoidance, adaptation, minimal end to end delays, and low packet breakdown by considerable utilization of bandwidth. This mechanism introduces dynamic Superframe reconfiguration in bandwidth allocation with priorities to transfer signals. The validity is demonstrated by Castalia and results show it significantly decreases delay and reaches a considerably high packet delivery ratio in the entire network as compared to the general IEEE 802.15.4 model. In the future more characteristics of physical and application layers should be considered for better reliability.

Chapter 6 An adaptive fuzzy control medium access in Wireless Body Area Networks

6.1 Introduction

A wireless body area network (WBAN) is a novel system which uses low power, tiny wireless sensors to continuously collect physiological, behavioural or other health related parameters and forward them to professionals for analysis or make decision [19, 71]. WBANs can monitor vital body signs such as heart-rate, temperature, blood pressure, ECG, EEG and pH level of patients. By replacing cables with wireless links, WBANs can support less invasive and more comfortable and efficient services both in hospital and outside the hospital. WBANs create particular interest to the healthcare sector to provide efficient healthcare services and ongoing clinical management. Examples of implantable devices that can be equipped with wireless transceivers including cochlear implants, retina implants, glaucoma sensors, intracranial pressure sensors, capsule endoscopes, glucose sensors, insulin pumps, heart pacemakers, cardiac monitors and so on.

WBAN is accepted in some innovative and significant utilization such as healthcare, sports, entertainment and military because it can provide convenient and economical services with invasive or non-invasive mini gadgets [6, 72-74]. However special concerns and requirements in some applications hinder the wide spread usage. For example data traffic rates may range from very low to very high for different sensors in a network, which requires adaptive communication protocols. Particularly for users in some critical or emergency medical status, highest reliability, shortest access and immediate response are very high demanded. The dropping of critical signals could eventually lead to life threatening consequences in urgent treatments. These special features distinguish WBAN to conventional wireless sensor networks (WSN) in capability and induce many new issues in utilization [72]. Therefore effective and feasible solutions or standards should be considered in different network situations, such as: flexible data rate, various nodes priority and capacity of channel, which may be constrains in existed protocols. Generally there are two main wireless medium access mechanisms: one is based on slots contention such as CSAM/CA and the other is contention free based on slots assignment. For contention mechanism the collision probability goes up exponential along with the number of node and causes the reducing of throughput. While for fixed slots assignment mechanism it is difficult to make a balance for diverse traffic load and priorities. So appropriate medium access schemes are very significant to solve some limitations and obtain better performance.

MAC is the key part of all communication protocol stacks and it provides the basic performance to achieve Quality of Service (QoS) in wireless networks. An adaptive MAC should support various applications and different kinds of data including periodic, no periodic, burst and continuous data to meet high QoS. The aim of MAC protocol is to prevent collisions and to avoid simultaneous transmissions while pursuing maximum throughput, least latency, highest communication reliability and best energy efficiency. Therefore a wealth of research papers are inspired to propose coordination or optimization methods on MAC layers for a more reliable and efficient communication. Pangun Park derived an adaptive algorithm to formalize the power consumption while guaranteeing reliability and delay constraints in the packet transmission[63]. Yan Zhang used application specific control channels in MAC to provide priority Guarantee [64]. H. Ozgur Sanli and Hasan Çam proposed a practical coverage and rate allocation protocol to exploit this dependency in realistic environments[65]. G. Wu considered resource reservation based on priority queue but it serves sensors firstly when they require less bandwidth reservation. Those protocols consider more of the power consumption and efficiency of transmission while the priority is paid less attention. Or the priority is considered but cannot guarantee efficient resources allocation and prior services for emergency services.

In this chapter a novel adaptive medium access scheme based on fuzzy logic system is proposed. With several input parameters an adaptive scheme is designed to improve throughput, reduce collision and guarantee high priority services. The performance of the scheme is evaluated with simulations in Castalia-3.2.

For the remaining part of this chapter, section 2 explains the fuzzy logic system and basic components. Section 3 briefly introduces PHY layer and MAC Sublayer for wireless communication. Section 4 introduces the major medium access mechanisms. Section 5 explains the structure of adaptive control model for medium access. Section 6 presents the implementation of the module. Section 7 describes the performance and section 8 evaluates the simulation results. Finally section 9 makes a conclusion.

6.2 Fuzzy Logic control

Generally, a Fuzzy Logic System (FLS) is a nonlinear mapping of an input data vector into a scalar output. Its ultimate goal is to provide foundations for approximate reasoning with imprecise propositions based on fuzzy set theory, which is similar to the classical reasoning using precise propositions based on the classical set theory[75].

6.2.1 A basic Fuzzy Logic System

A FLS is widely used in fuzzy logic controller, which is a control and decision system approach that mimics human control logic, in the same way a human would make decisions. Fuzzy logic is an extension of classic logic to provide a simple way for a definite conclusion based on vague, ambiguous or imprecise input information. A fuzzy system handles input linguistic variables to produce output linguistic variables and can be constructed of basic components: Fuzzification, Rule Based, inference Engine and Defuzzification as in Figure 6.1[76].

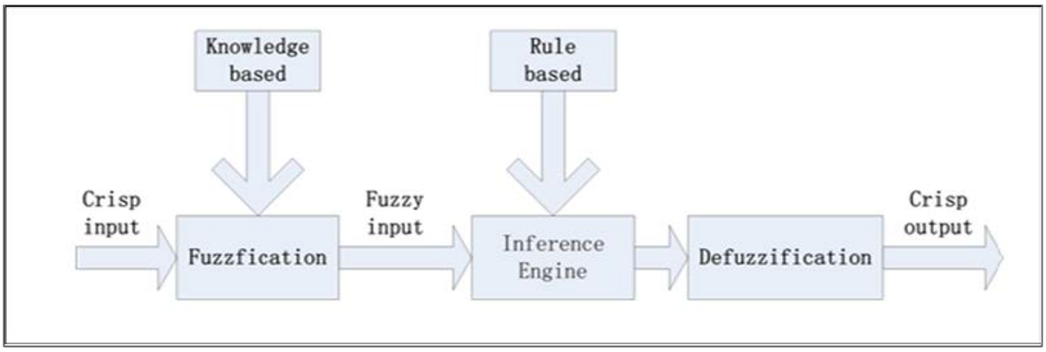


Figure 6. 1 Basic fuzzy system

Fuzzification transfers system inputs (crisp inputs) space to fuzzy sets in a defined universe in terms of the membership function, of the fuzzy sets. The membership function is a graphical representation of the magnitude of participation of each input. The membership function is associated to each system input: $F = fuzzifier(x_i)$, where x_i represents a crisp input value, F is a fuzzy set and fuzzifier represents a fuzzification operator.

In a fuzzy set we name all the elements of the Universe and assign them a number between 0 and 1. This number demonstrates to what degree this element belongs to the defined fuzzy set. The fuzzy logic controller triangular membership functions are defined over the range of the fuzzy input values and linguistically describe the variable's universe of discourse.

The Fuzzification Module transforms the physical values (position, voltage, degree, etc.) of the process signal into a normalized fuzzy subset consisting of a subset for the range of the input values and a normalized membership function describing the degree of confidence of the input belonging to this range.

Rule Base:

A fuzzy rule is a conditional statement with the following form: if x is A then y is B, where x and y are linguistic variables, A and B are linguistic values represented by fuzzy sets defined on the universe X and Y. A fuzzy rule have multiple parts, which can be combined by "and" or "or" operators. IF-THEN rule base performing the inference is in the general form:

R1: IF controller input e_1 is E_{11} AND ... AND controller input e_n is E_{1n}, THEN controller output u_1 is U_1.

…

Rm: IF controller input e_1 is E_{m1} AND ... AND controller input e_n is E_{mn}, THEN controller output um is U_m.

In general, m rules produce m controller outputs, u_1, ... , u_m, belonging to m fuzzy subsets. The establishment of this rule base depends heavily on the designer's work experience, knowledge about the physical plant, analysis and design skills and so on.

The Fuzzy Logic Rule Base: Designing a good fuzzy logic rule base is critical to obtain a satisfactory controller for a particular application. Classical analysis and control strategies should be incorporated in the establishment of a rule base. A general procedure in designing a fuzzy logic rule base includes the following: Determining the process states and control variables; Determining input variables to the controller.

Inference Engine: The Inference Engine uses the knowledge and the fuzzy inputs to make inference by a reasoning method. All fuzzy rules will be activated and are combined together. The aim of the "engine" is to create control actions, in fuzzy terms, according to the information provided by the fuzzification module and the set-point tracking requirement.

Defuzzification: The defuzzification module is the connection between the control rule base and the physical process to be controlled, which plays the role of a transformer mapping the controller outputs (generated by the control rule base in fuzzy terms) back to the crisp values that the process can accept. Hence, in a sense the defuzzification module is the inverse of the fuzzification module. The controller outputs u_1,…, u_m, generated by the rule base above, are fuzzy signals belonging to the fuzzy subsets U_1, ..., U_m, respectively. It reconverts the fuzzy output values, deriving from the inference mechanism, into crisp values. Between the defuzzification step and the physical plant, a post-processing unit mapping the point wise signal u to a physical signal may be needed, depending again on the nature of the underlying process. What has to be determined in this stage is essentially a defuzzification formula. There are several commonly used, logically meaningful, and practically effective defuzzification formulas available, which are by nature weighted average formulas in various forms.

The Defuzzification Module: The defuzzification module is in a sense the reverse of the fuzzification module: it converts all the fuzzy terms, which are created by the rule base of the controller to crisp terms (numerical values) and then sends them to the physical system (plant, process), so as to execute the control of the system.

The defuzzification module performs the following functions: It creates a crisp, overall control signal, u, by combining all possible control outputs from the rule base into a weighted average formula.

6.2.2 Application of Fuzzy logic system

Fuzzy logic system has been applied in many areas such automobile systems, digital image processing, pattern recognition and so on [76-79]. A. Boulkroune proposed an adaptive fuzzy tracking control for a class of MIMO non affine uncertain systems [80]. Huanqing Wang proposed a fuzzy decentralized control approach for a class of uncertain stochastic nonlinear large-scale systems [81]. Shaocheng Tong proposed Observer-based fuzzy adaptive robust control of nonlinear systems with time delays and un-modeled dynamics [82]. H. B. Kazemian presented Neuro-Fuzzy applications to Moving Picture Expert Group (MPEG-4) video transmission in ZigBee [83]. Also there are some researchers use fuzzy logic in wireless medium access control. For example: Hwang proposed the EDCA Performance Improvement Protocol [84]. Liu and Hsu proposed a Protocol for multi channels WLANs with a simple fuzzy controller to tune the size of the medium access control backoff window[85]. B. Otal applied Fuzzy-logic Scheduling for Highly Reliable and Energy-efficient Medical Body Sensor Networks[86]. John Wallace proposed Fuzzy Logic Optimisation of MAC Parameters and Sleeping Duty-Cycles in Wireless Sensor Networks[87].

Using fuzzy logic model in medium access is proposed by the above papers but they mostly only focus on modifying back window to reduce the possibility of collisions without considering nodes' working status. It is not enough for both CAP and CFP accesses schemes in WBAN.

6.3 PHY and MAC

In the OSI networking model, the physical layer (PHY) is a fundamental layer below the logical data structures of the higher level. It defines the means of transmitting raw bit, which will be converted to a physical signal and transmitted over a transmission medium. The physical layer provides an electrical, mechanical, and procedural interface to the transmission medium. The shapes and properties of the electrical connectors, the frequencies to broadcast on, the modulation scheme to use and similar low-level parameters are specified here.

Media access control (MAC) is a sublayer of the data link layer sublayer. It acts as an interface between the logical link control (LLC) sublayer and the network's physical layer. It provides addressing and channel access control mechanisms so that several terminals or network nodes can communicate within a multiple access network with a shared medium.

6.3.1 PHY layer

The PHY provides two services: the PHY data service and the PHY management service. The PHY data service enables the transmission and reception of PHY protocol data units (PPDUs) across the physical radio channel.

For 802.15.4 standard, Physical layers (PHYs) are defined for:

(1) Devices operating in the license-free 868–868.6 MHz, 902–928 MHz, and 2400–2483.5 MHz bands

(2) Devices with precision ranging, extended range, and enhanced robustness and mobility

The standard provides for ultra-low complexity, ultra-low cost, ultra-low power consumption, and low data rate wireless connectivity among inexpensive devices. The raw data rate is high enough (250 kb/s) to satisfy a set of applications but is also scalable down to the needs of sensor and automation needs (20 kb/s or below) for wireless communications.

The PHY provides an interface between the MAC sublayer and the physical radio channel, via the RF firmware and the RF hardware. The PHY provides two services, accessed through two SAPs: the PHY data service, accessed through the PHY data SAP (PD-SAP), and the PHY management service, accessed through the PLME-SAP. The maximum PSDU size (in octets) the PHY shall be able to receive is 127 bytes.

6.3.2 MAC sublayer

The MAC sublayer provides two services: the MAC data service and the MAC management service. The MAC data service enables the transmission and reception of MAC protocol data units (MPDUs) across the PHY data service.

The features of the MAC sublayer are beacon management, channel access, GTS management, frame validation, acknowledged frame delivery, association, and disassociation. In addition, the MAC sublayer provides hooks for implementing application-appropriate security mechanisms.

The MAC sublayer provides an interface between the next higher layer and the PHY. The MAC sublayer conceptually includes a management entity called the MLME. This entity provides the service interfaces through which the invoking layer management. The MLME is also responsible for maintaining a database of managed objects pertaining to the MAC sublayer.

MAC in IEEE 802.15.4

IEEE 802.15.4 standard defines a beacon-enabled modality, which uses a slotted Carrier sense multiple accesses with collision avoidance (CSMA/CA) and the optional Guaranteed Time Slot (GTS) allocation mechanism. A coordinator regularly forwards the beacon frames in each beacon interval and communicate with sampling sensors in duplex during active period, named the superframe duration SD. After this period system goes into the low-power status within the inactive period. The structure of the superframe is determined by two parameters, the beacon order BO and the superframe order SO, which determine the length of the superframe and its active period, respectively.

The superframe consists of three components: a beacon, a contention access period (CAP), and a contention free period (CFP). The beacon is transmitted in Slot 0 without CSMA/CA, and the CAP starts immediately after the beacon. A slotted CSMA/CA mechanism is used to access the channel of non time critical data frames and GTS requests during the CAP. The CFP, if presents, commences after the CAP to the end of the active period. Any allocated GTSs should be located within the CFP. In the CFP, the dedicated bandwidth is used for time critical data frames. All frames transmitted in the CAP will use a slotted CSMA/CA mechanism to access the channel. MAC command frames shall always be transmitted in the CAP.

The CFP will start on a slot boundary immediately following the CAP, and it will finish before the end of the active period. If any GTSs have been allocated by the coordinator, they will be located within the CFP and occupy contiguous slots. The CFP duration will increase or decrease depending on the total length of all calculated GTSs.

6.4 Major Medium access Mechanisms for CAP and CFP

6.4.1 CSMA/CA Mechanism for CAP

CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) is a protocol for carrier transmission in computer networking, where carrier sensing is used while nodes try to avoid collisions by transmitting only when the channel is sensed to be "idle". For example: if a node wants to transmit a data packet, firstly the MAC sublayer initializes four variables: th enumber of backoffs (NB=0), contention window (CW=2), backoff exponent (BE=macMinBE), and retransmission times (RT=0). Then the MAC sublayerdelays for a random number of complete backoff periods are in the range $\left[\mathbf{0},\mathbf{2}^{BE}-\mathbf{1}\right]$. If the number of backoff periods is greater than the remaining number of backoff periods, the MAC sublayer will stop the backoff counts down at the end of the CAP and resumes it at the start of the next superframe.

Otherwise, the MAC sublayer counts its backoff delay. When the backoff period is zero, the node needs to perform the first clear channel assessment (CCA). The MAC sublayer proceeds in the remaining CSMA/CA algorithm steps (i.e., two CCAs). The frame transmission and any acknowledgment can be completed before the end of the CAP. If the MAC sublayer cannot proceed, it shall wait until the start of the CAP in the next superframe and apply a further random backoff delay in the range $\left[\mathbf{0},\mathbf{2}^{BE}-\mathbf{1}\right]$ before evaluating whether it can proceed again. Otherwise the MAC sublayer proceeds the CCA in the current superframe. If two consecutive CCAs are idle, then the node commences the packet transmission. If either of the CCA fails due to busy channel, the MAC sublayer increases the value of both NB and BE by one, up to a maximum value macMaxCSMABackoffs and macMaxBE, respectively. Hence, the values of NB and BE depend on the number of CCA failures of a packet. Once BE reaches macMaxBE, it remains at the value macMaxBE until it is reset. If NB exceeds mac Max CSMA Backoffs, then the packet is discarded due to channel access failure. Otherwise, the CSMA/CA algorithm generates a random number of complete backoff periods and repeats the process. Here, the variable macMaxCSMA Backoffs represents the maximum number of times the CSMA/CA algorithm is required to backoff. If channel access is successful, the node starts transmitting packets and waits for ACK. The reception of the corresponding ACK is interpreted as successful packet transmission. If the node fails to receive ACK due to collision or ACK timeout, the variable RT is increased by one up to macMaxFrameRetries. If RT is less than macMaxFrameRetries, the MAC sublayer initializes two variables CW=0, BE=macMinBE and follows the CSMA/CA mechanism to re-access the channel. Otherwise, the packet is discarded due to the retry limit.

6.4.2 GTS Allocation in CFP

The coordinator is responsible for the GTS allocation and determines the length of the CFP in a superframe. To request the allocation of a new GTS, the devices end the GTS request command to the coordinator. The coordinator confirms its receipt by sending an acknowledgment frame within CAP. Upon receiving a GTS allocation request, the coordinator checks whether there are sufficient resources and, if possible, allocates the requested GTS. The GTS capacity in a superframe satisfies the following requirements:

1. The maximum number of GTSs to be allocated to devices is seven, provided there is sufficient capacity in the superframe.

2. The minimum length of a CAP is aMinCAPLength.

Therefore, the CFP length depends on the GTS requests and the currently available capacity in the superframe. If there is sufficient bandwidth in the next superframe, the coordinator determines a device list for GTS allocation, based on a first-come-first-served (FCFS) policy. Then the coordinator includes the GTS descriptor, which is the device list that obtains GTSs in the following beacon to announce the allocation information. The coordinator makes this decision within a GTS Desc Persistence Timesuperframes. Note that on receipt of the acknowledgment to the GTS request command, the device continues to track beacons and wait for at most a GTS Desc Persistence Timesuperframes.

A device uses the dedicated bandwidth to transmit the packet within the CFP. In addition, a transmitting device ensures that its transaction is completed in one interframe spacing (IFS) period before the end of its GTS.

6.5 The structure of fuzzy control medium access (FCMA)

According to the specific circumstances in WBAN the implementation of medium access is operated by the coordinator. First of all a device or a node, such as a PAD, a laptop or computer, is defined as a coordinator which is in charge of receiving data from other nodes. The coordinator analyses, sets priorities, allocates slots and transfers to the service centre. When abnormal or emergency data are detected, the coordinator changes initial priorities and allocate slots again for new requirements.

This chapter proposes a novel adaptive control of medium access based on IEEE 802.15.4 for reliable and effective communication. The protocol is adjusted dynamically by a fuzzy logical optimization control.

For this optimal fuzzy control medium access the architecture of the system is divided into three logical function modules: Data acquisition module, fuzzy logic control module and implementation module. Figure 6.2 illustrates the structure of a FCMA system. All the modules are introduced in the following sections.

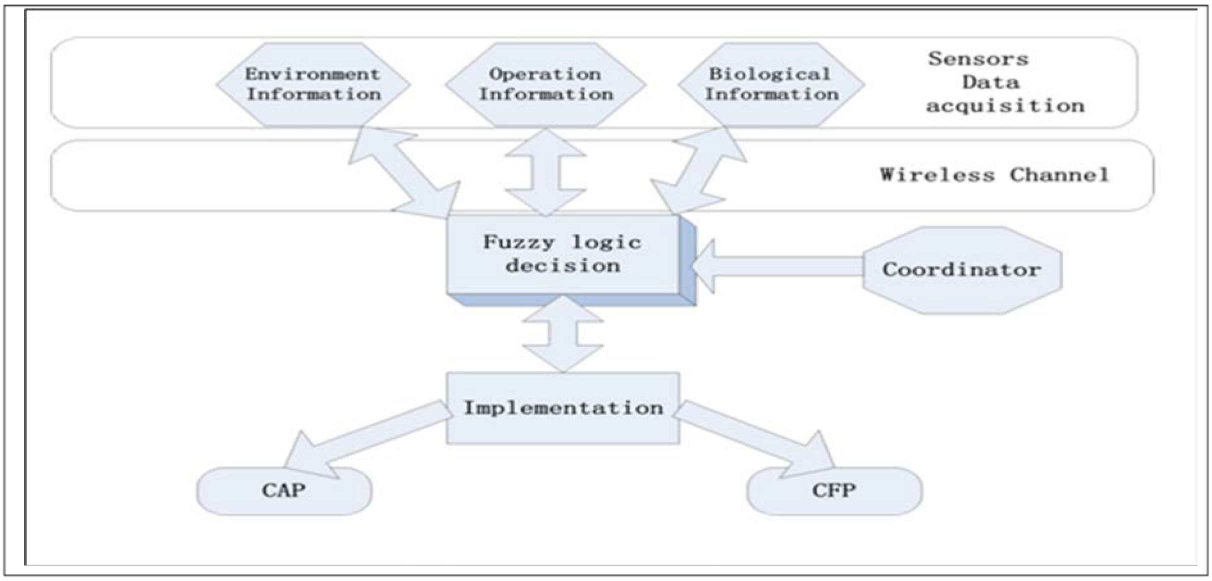


Figure 6. 2 Structure of FCMA

6.5.1 Data acquisition module

In the data acquisition module sensors collect environmental, biological and runtime information, such as blood pressure, heart rate, SpO2, environment temperature and so on. The sampling frequencies of biomedical sensors usually are no more than 300 Hz since biological signals oscillate at low frequencies. When a large array of sensors are used , e.g., EEG 192-channel recording, the data rate will increase up to hundreds of kbps[66]. Table 6.1 shows the data rate of several biomedical measurements. If environment sensors send video information high data rate is required. Thus these sensors should be classified by critical or non-critical nodes with priorities in different applications. Data acquired by all sensors go through wireless channel to the coordinator in WBAN.

Table 6. 1 Sampling frequency of biomedical sensors

Sensors	Sampling frequency (Hz)	Basic Data Rate
ECG 1-point	250	6 kbit/s
ECG 12-point	250	72 kbit/s
EEG 1-channel	200	4.8 kbit/s
EEG 192-channel	200	921.6 kbit/s
Blood pressure	60	1.44 kbit/s
SpO2	300	7.2 kbit
Body temperature	0.1	2.4 bit/s

6.5.2 Fuzzy logic control module

The fuzzy logic control is executed by the coordinator. The data are classified into different types according to collected parameters. Some data are very critical under certain situations. For example: heart rate, ECG, oxygen and other indicators demonstrate the important real-time physiological status. Those vital data should be ensured with high end-to-end transmission reliability and low delivery delay. However non critical data are not time urgent and real time transmission is not necessary for them. In our model data are classified into four types: non real-time data, normal real-time data, abnormal data, and emergency data. The coordinator can analyse and judge abnormality of information with pre-defined activation threshold or principal components analysis[67]. If abnormal or emergency data are detected, the involved nodes should send information as soon as possible. So there are three possible priorities of each node: normal, abnormal and emergency. When sensors with transient abnormal data return to normal status, the priority decreases to its original value. High priority nodes will be guaranteed more resources or low probability of collisions while low priority will be allocated less resources, rather than first come first serve and random of collisions probability in standard IEEE 802.15.4.

A slotted (CSMA/CA) mechanism is used in CAP to access the channel for non real time data and GTS requests are used in CAP for real time data. In the CFP, the dedicated bandwidth is guaranteed for time critical data. Each node is a type of sensor, which requests and sends one type of message dynamically. All non-real time messages should be periodically dispatched in CAP period with fuzzy control of contention window. Other real time information transmitted in GTS is allocated by coordinator with fuzzy control output.

The coordinator builds a fuzzy control model to manage the access schemes. The input parameters are fed into the fuzzy logic system and different fuzzy rule are used to acquire an optimal solution to two access mechanisms: contention access period (CAP) and contention free period (CFP). In CAP coordinator modifies the size of the contention window base on fuzzy control factor to allow different possibility of access for nodes and meet certain criteria. In CFP, the coordinator allocates slots according fuzzy control factor. The Fuzzy Logic control model is a nonlinear system, which is realized by two variables. The input variables are defined as: $priority \subset \{normal, abnormal, high\}$, $rate \subset \{low, medium, high\}$. The output decision is defined as contention window associated with a fuzzy set {short, long, longer, longest} for CAP or allocation slots as {wait, effort, best effort, urgency} CFP. Therefore scheduling of backoff periods or slots is feasible to ensure data reliability, increase throughput and reduce latency.

6.6 Implementation module

6.6 .1 Fuzzy logic control in CAP

As defined in 802.15.4 CAP applies CSMA/CA which uses the Random Back-off time algorithm to reduce the probability of collisions. Each node on the network can sense the channel before transmitting the data packet. If the channel is busy, the node waits for a randomly period of time, and then checks again to see if the channel is clear. This waiting period is called backoff time and counted down by a backoff counter. Once the backoff counter reaches zero and the channel is clear the node sends data. A node sets its backoff counter to a random integer number uniformly distributed over the interval $[1, CW]$ where CW is called contention window and $CW \in [CW_{min}, CW_{max}]$. In varying the size of contention window, certain nodes can more quickly access the channel with shorter backoff time. The contention window is decided by the fuzzy rules with two input parameters: priority and data rate. As discussed in previous part there are usually three priorities for each nodes and the data rate is divided as low, medium and high. A convenient method to define all rules is using a decision table called as a fuzzy association memory bank matrix. The decision table for CAP is showed in Table1. Since the input two variables have three states respectively, the total possible output is 9 values.

The CAP allocation shall be executed when the sensor nodes want to send normal non real time data. Nodes in beacon enable mode utilize the slotted CSMA/CA access mechanism. The contention window is adjusted based on the fuzzy control output. If input values are low rate and normal priority, the contention window is big. On the contrary, if input values are high rate and emergency priority the contention window shall be as small as possible.

Table 6. 2 Fuzzy decision table for CAP

Rate	Priority		
	normal	abnormal	emergency
Low	slow	ordinary	immediate
Medium	slow	fast	immediate
High	ordinary	fast	immediate

6.6.2 Fuzzy logic control in CFP

In CFP allocation a reserved bandwidth is previously granted for given data. Adaptive parameters of the super frame will be assigned by the determined amount of time slots. Now we use an example to explain the allocation scheme in detail: 5 nodes are assumed to transfer data: node 1 sends normal real time data, node 2 sends emergency data and node 3 sends abnormal real time data, node 4 and 5 forward non real time data. Therefore node 1, 2 and 3 use GTSs in CFP by coordinator while node 4 and 5 use CSMA/CA in CAP. If the rate required for nodes are: Node1= 50kbps, Node2=100kbps, Node 3=100kbps, Node4=20kbps, Node5=20kbps. When the size of a symbol is 4 bits ,100 frames/second sampling frequency is chosen and the size of a base slot is 60 symbols, then required slots for Node1, Node2, Node3, Node4 and Node5 are 3, 5, 5, 1 and 1 respectively. The total slots should be allocated to all nodes are NumSlot = 3+5+5+1+1= 15 and for GTS are 13. In standard IEEE 802.15.4 the number of GTS is limited to 7 and it cannot satisfy the requirement of nodes, which will cause latency of critical information and even bring life-threaten results. While in FCMA mechanism the increase and decrease of slots number can be dynamically adjusted in superframe configuration so it would achieve maximum bandwidth utilization and minimum latency to guarantee high reliability in emergency.

To solve restrictions of IEEE 802.15.4, superframe slots can vary accordingly to fuzzy logic control output. The proposed protocol changes the number of slots according to requirement, so that it can achieve better utilization, save energy and better reliability. The equations of modified SD and BI can be defined as:

$$BI = aBaseSlotDuration \times (BO + NumSlot) \tag{6.1}$$

$$SD = aBaseSlotDuration \times NumSlot \tag{6.2}$$

NumSlot is determined by the sum of each node’s required slots according to the data rate and priorities. BO should be determined by the following equation:

$$16 \times aBaseSlotDuration \leq aBaseSlotDuration(BO + NumSlot) \leq 16 \times aBaseSlotDuration \times 2^{14} \quad (6.3)$$

If the GTS capacity is available for all nodes, which means there are enough slots for total data, each node can be allocate slots as they need. If the GTS capacity is not enough we allocate slots according to the fuzzy control output. For example if the priority is the highest and the rate is also highest this node's service is urgent and its slots should be ensured firstly. If the priority is lowest and rate is lowest, this node may have to wait. Scheduling slots with the control factor can ensure emergency nodes with high level service, which improve the reliability for critical information. The decision table for CFP is showed in Table 6.3:

Table 6. 3 Fuzzy decision table for CFP

Rate	Priority			Capacity
	normal	abnormal	emergency	
Low	waiting	effort	urgency	No
Medium	waiting	best effort	urgency	No
High	effort	Best effort	urgency	No

6.7 Performance analysis

An effective MAC mechanism may improve the network performance, which can be evaluated by key factor such as latency, packet breakdown and throughput.

6.7.1 Latency

Latency plays a very important role in real-time applications. In WBANs critical data should be reliably received within minimum latency. In CSMA/CA mechanism the latency is associated with the contention window and back off times as in equation 6.4 [88] :

$$Delay = \overline{CW} + T_{data} + T_{I-ACK} + \mathbf{2}T_{pSTFS} + \mathbf{2}\tau \quad (6.4)$$

$\overline{CW}$ is the average backoff time decide by contention window and times of retry, T_{data} is the transmission delay , T_{I-ACK} is the immediate acknowledgement period . T_{pSTFS} is a short inter frame spacing time and τ is the propagation delay. For CFP model in a star topology of IEEE802.15.4, if there are N nodes with real time data and the coordinator allocate 1 GTS per superframe for each node, it will cause a worst case queuing latency of $16 * \mathrm{Slot_{Duration}} * \left\lceil \frac{\mathrm{N}}{\mathrm{maxGTS}} \right\rceil$[89]. Hence, the uplink latency in GTS is given by equation 6.5 :

$$Latency = QLantency + T_{Transimission} = \mathbf{16} * Slot_{duration} * \left[\frac{N}{\mathbf{max}\, GTS} \right] + T_{data} + T_{I-ACK} + \mathbf{2}\tau \tag{6.5}$$

Therefore if latency grows linearly with the number of nodes in regular allocation and it is serious for emergency data with waiting a long queuing. But in FCMA algorithm, if emergency data are allocated enough slots in a super frame and can be transferred without queuing latency (QLatency) to guarantee immediate transmission.

6.7.2 Packets breakdown

The packet breakdown means a packet is broken after either it gets the maximum number of transmission attempts or a new frame gets to the node while it was still waiting for its transmission. The packet breakdown rate is a function of packet arrival rate λ and beacon interval BI. These parameters decide how long, on the average, a frame will have to wait for its opportunity for transmission and be possible to be dropped as new frames arrive. According to reference[69] ,the packet breakdown can be defined in equation 6.6:

$$P_{breakdown} = P_{d0} + (1 - P_{do}) \tag{6.6}$$

$$P_{d0} = 1 - \frac{\lambda B_I e^{-\lambda B_I}}{1-e^{-\lambda B_I}}, \quad P_{di} = (1 - P_{d0}) P_e^i (1- e^{-\lambda B_I}) e^{-(i-1)\lambda B_I}$$

Where $e^{-(i-1)\lambda B_I}$ represents no arrivals in the previous i-1 superframes and 1-$e^{-\lambda B_I}$ represents at least one arrival in superframe i. Also$(1 - P_{d0})$ is the probability that the packet is not fail in the first superframe. In the general IEEE 802.15.4 the probability failure is associated with the waiting time and collision while in FCMA this probability can be reduced due to the efficient usage of bandwidth.

6.7.3 Throughput

For CAP Throughput (S) can be defined as the ratio of the payload size (in bytes) as x , to the total transmission latency per payload size and is given by equation 6.7:

$$S(t) = \frac{\mathbf{8} * x}{Latency} \tag{6.7}$$

For CFP the system throughput S (t) of the superframe t is given by the ratio of the average length of successfully payload in a GTS time slot to the average length of a GTS time slot, namely:

$$S(t) = \frac{P_{s(t)} L_{pl} \tau_n}{\theta_{min} T_{ss}} \tag{6.8}$$

where Ps (t) = 1 − Pbreak (t), Lpl is the length of payload of each data packet, τ_n is the number of data packets, T_{ss} is the length of superframe slot, and θ_{min} is the number of superframe slots [70].According to the discussion before, when the probability of packet breakdown (Pbreak (t)) is reduced compared with IEEE 802.15.4, the throughput in FCMA increases consequently.

6.8 Simulation results

In this section, we present our experimental analysis to evaluate the performance of proposed protocol in CAP and CFP. Firstly we build the fuzzy control system with Matlab. Figure 6. 3 shows the control model in CAP. There are two input values: priority, rate and one output of contention window decided by fuzzy rules. The input function is a triangular-shaped membership function and the output function is Gaussian distribution curve. Figure 6.4 shows a 3D surface of the system output with different input values. From it we can decide contention based on the node's input priority and rate.

Then FCMA is evaluated by Castalia Simulator, which is an open source discrete event-drive simulator developed on OMNet++. For the simulation, there are 6 nodes in a network with 1 Coordinator and 5 nodes of a star topology.

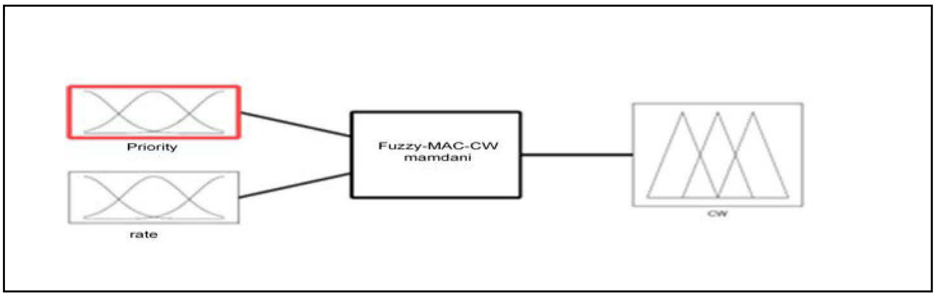


Figure 6. 3 Fuzzy logic control of CAP

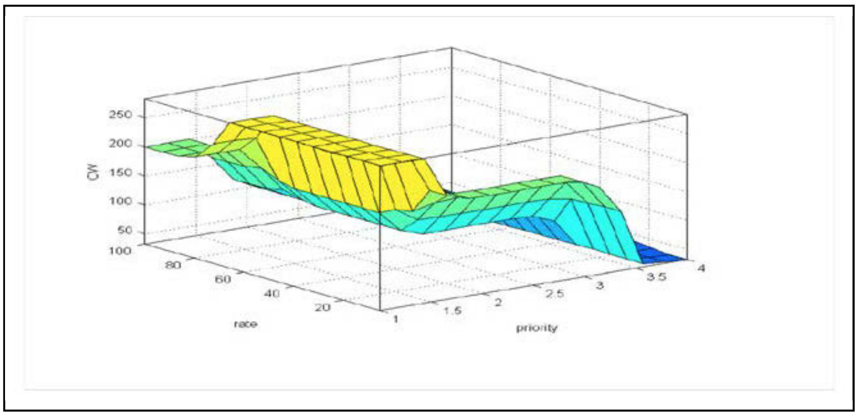


Figure 6. 4 Output of system

Table 6.4 explains the specifications of nodes. By this simulation model, the standard IEEE802.15.4 protocol and FCMA are compared with performance parameters such as latency, packet breakdown and throughput.

Table 6. 4 Specification of nodes

Nodes	Node1	Node2	Node3	Node4	Node5
Priority scope	[2,4]	[2,4]	[2,4]	[1,3]	[2,3]
Initial priority	normal	normal	normal	normal	normal
Basic rate(Packets/s)	50	50	50	20	20
Maximum rate (Packets/s)	100	100	100	100	100

Figure 6.5 shows the comparison of application-level latency histograms with many time buckets in CAP. The optimal mechanism performs better in whole process due to the control of contention window and reduction of collisions. Then we investigate the throughput in Figure 6.6. In this figure the y axis shows the average packets received by the coordinator and the x axis is the rate for each node in packets/sec. It shows when data rate is low, the throughput is monotonic increasing. But the throughput almost does not change when data rate is high in standard scheme, which means there is an indication of saturation. While using the optimal scheme the throughput still goes up because the packets dropping is low of less collisions and this is in accordance with figure 6.7. It demonstrates there are more packets received from each node in new method than in standard one. Especially in high rate and high priority the successful packets of node 5 reach the maximum in new mechanism. This observation explains the fuzzy control of back off period in CAP at the MAC level reduce collisions and latency, both of which consequently upgrade the throughput.

Then we implement the adaptive control in CFP and acquire results as shown in figure 6.8 and 6.9. In figure 6.8, when data rate is not high the throughput is same. While the traffic rate reaches the larger volumes, it is obvious to see the increasing throughput is higher in FCMA. Also there is huge difference of packets breakdown portion when data rate is high. When data rate is 100 packets/second the breakdown ratio is close to 40% with standard protocol because it only allocates slots evenly and causes long queue for some nodes, which is the main reason for overflow. When using FCMA packets breakdown drops sharply because it makes a more efficient usage of the wireless medium according to nodes' priority and rate. When the number of nodes is set as 11 with two high priority nodes 8 and 9 the performance gain is more clearly as shown in figure 6.10.These two nodes are allocated enough slot and acquired more successful transmission.

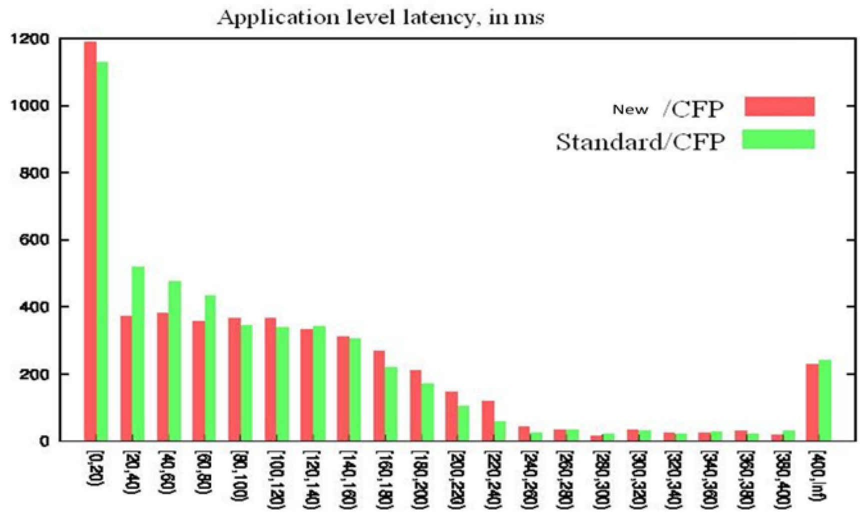


Figure 6. 5 The latency of CAP

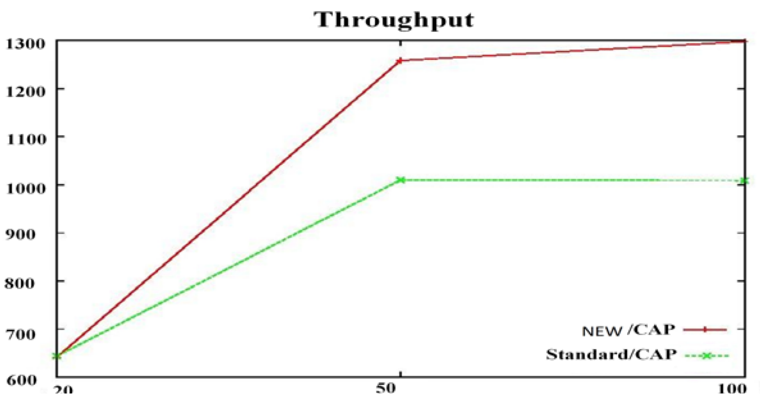


Figure 6. 6 Throughput of CAP

From the analysis it proves that FCMA strengthens WBAN performance not only in the whole throughput but also guarantees the QoS for nodes with prior requirements, which is very important for health related applications.

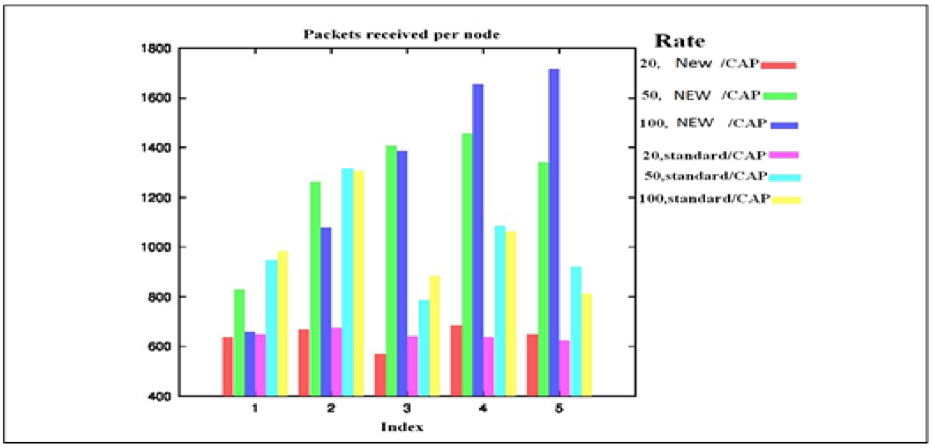


Figure 6. 7 Packets received (5 nodes+1coordinator)

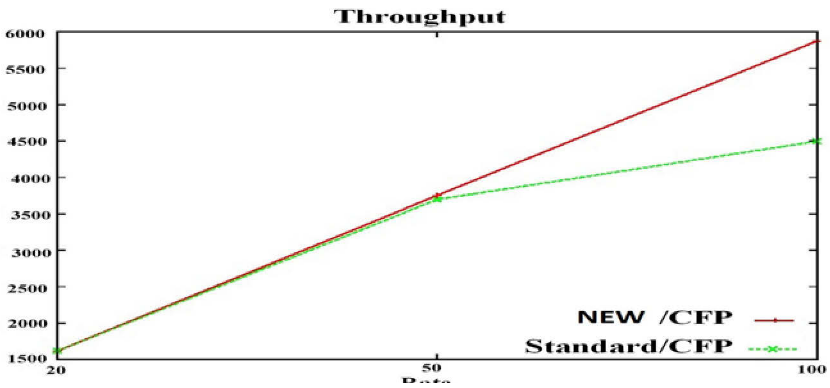


Figure 6. 8 Throughput in CFP

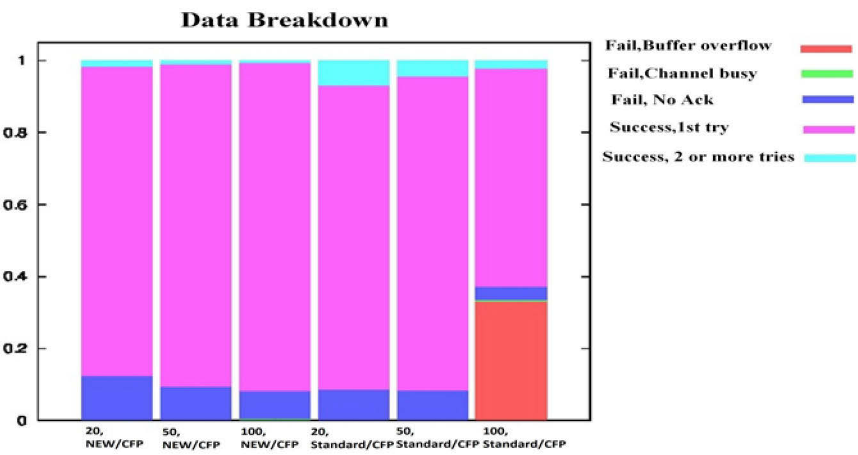


Figure 6. 9 Packets breakdown in CFP

6.9 Conclusions

This chapter proposes an adaptive MAC scheme based on fuzzy logic system to improve performance of WBAN such as: collision avoidance, latency, high throughput by guaranteeing multiple services of different conditions. This new mechanism introduces dynamic Superframe reconfiguration and utilizes fuzzy control scheme for contention window in CAP and slots allocation in CFP. The validity is evaluated by Castalia with specified nodes and obtained results show it significantly decreases latency and gains a high packet delivery ratio as compared to the standard IEEE 802.15.4 protocol. In the future research of other characteristics in physical and application layers such as hidden nodes, trade off and environment interference may be considered for better QoS.

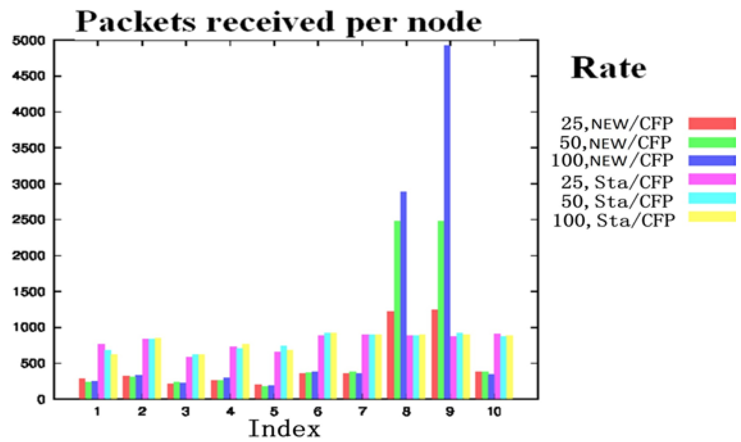


Figure 6. 10 Packets received (10 nodes +1coordinator)

Chapter 7 Intelligent Management of medium resource access in Wireless Body Area Network

7.1 Introduction

A. Background

Healthcare is a critical issue in the current world due to the rapid growth of demands in health and medical related services including baby care, diet monitor, sports assists, diseases, people caring and monitoring[90, 91]. Simultaneously with the growing population of old people, new technologies and applications are required to provide auxiliary methods for elderly independence and reduce healthcare costing [73]. Medical organizations and institutions always make effort to generate efficient, convenient and safe services for people who are in need at any time and any places. But because of the limited financial support, insufficient devices and professional human resources it is not easy to solve some problems and guarantee reliable services with traditional methods. Wireless body area network (WBAN) is a new technique to challenge the general model and bring benefits in health and medical applications [19,92,93]. WBAN uses sensors on, around or inside the human body to collect biological and environment parameters for further analysis or provide monitoring. Compared to traditional wireless sensor networks the human body environment and QoS in medical services bring new challenges in WBAN application, such as extreme energy efficiency, unique characteristics of the wireless channel, low latency and high data reliability[71, 94]. Therefore it requires appropriate protocols to support various functions and qualifications in health and medical related scenarios.

There are several wireless local area network protocols such as Bluetooth, Zigbee and WiFi for short distance services but they are not available for WBAN specially targeting health and medical application because specific working conditions are not considered in these standards. IEEE 802.15.6 Task Group (TG) developed an optimized communication standard to serve a variety of applications including medical, health and personal entertainment [95]. It outlines the basic elements, such as packet formats and message exchange protocols in medium access control (MAC) and physical layers to support a reliable wireless communication. But it still leaves open questions for further development, such as: how to manage the various access schemes including contention-based, scheduled, or improvised access? Is it available to increase the resource utilization and performance with intelligent control [71]? So there are still a lot of work to be done and the medium access control is a key factor [96], which has been studied by many researchers from many aspects.

B. Related works

Recently a wealth research of MAC has been done for enhancing the performance of WBAN. Sana Ullah derived numerical formulas to determine the maximum throughput and minimum delay of IEEE 802.15.6-based CSMA/CA Protocol [88]. They proved and evaluated the throughput and delay bounds for different data rates and frequency bands with one sender and one receiver. Shih Kuei-Ping proposed a virtual carrier sense MAC protocol with power controlling to avoid collisions [97]. It effectively reduces the energy consumption by a collection of stations and prevents collisions due to Point issue. Pangun Park designed an adaptive MAC protocol in control and monitoring applications [63]. A model is proposed based on Markov chain to minimize the power consumption and guarantee delay limitation and reliability. A. M. Asadi discussed a framework and surveyed proposed algorithms as opportunistic scheduling. In the application a user attempts to maximize the network throughput to meet certain constraints [98] . Mario Eagalj used a game-theoretic scheme to investigate the selfish behavior of nodes in CSMA/CA for a simple, localized and distributed protocol which successfully supervises multiple selfish nodes to reach a Pareto-optimal Nash equilibrium[99]. Boulis studied the trade-offs caused by a mix of two MAC techniques: contention-based access, and polling-based access. This paper also presented some MAC techniques including dynamic allocation of slots, retransmissions and power control based on studies of the BAN channel. The performance was compared by simulations [100].Tian Jun proposed a Game-theory Model based on CSMA Protocol in Wireless Network for more throughputs[101] . The analytical can get the transmit power and optimal carrier sensing threshold with maximizing the total network throughput.

These research concerned TDMA or CSMA/CA respectively but did not consider multiple access schemes together in the application. Usually contention-free access and contention access coded are simultaneously deployed in the application. The scheduling problem might be quite hard under various channel conditions and error-prone wireless channels. Therefore taking into account the requirement of applications and channel status, an intelligent control strategy is necessary for flexible and ideal resource allocation and better performance. So we establish an optimal strategy to manage various access schemes under different situations including services and link states. To the best of our knowledge it is the first time to propose a comprehensive management model of different access schemes for WBAN. The proposed scheme can be applied to generate an efficient and smart MAC protocol and the performance is evaluated by a set of simulations.

The contributions of this chapter contains: 1) A series of utility functions are defined with considering various situations in data transmission. These specific definitions can facilitate network to adapt to communication requirement and channel states. 2) An intelligent medium access scheme is proposed to obtain flexible and efficient allocation of resource and bring benefits for WBAN in different applications.

The rest of the chapter is organized as follows: Section 2 introduces the main access mechanisms and their features respectively. Section 3 analyses the throughput and energy consumption under different mechanisms. Section 4 proposes the intelligent management model. Section 5 illustrates simulations and results. Section 6 makes conclusions.

7.2 Wireless Body Area Network Standard: IEEE 802.15.6

IEEE 802.15.6 standard provides an international standard for a short-range (i.e., about human body range), low power, and highly reliable wireless communication for use in close proximity to, or inside, a human body (but not limited to humans)[95]. It uses existing industrial scientific medical (ISM) bands as well as frequency bands approved by national medical and/or regulatory authorities. Support for quality of service (QoS), extremely low power, and data rates up to 10 Mbps is required while simultaneously complying with strict non-interference guidelines where needed. This standard considers effects on portable antennas due to the presence of a person (varying with male, female, skinny, heavy, etc.), radiation pattern shaping to minimize specific absorption rate (SAR) into the body, and changes in characteristics as a result of the user motions.

7.2.1 Network topology

All nodes and hubs are to be organized into logical sets, referred to as body area networks (BANs) in this specification, and coordinated by their respective hubs for medium access and power management. There is to be one and only one hub in a BAN, whereas the number of nodes in a BAN is allocated from zero to mMaxBANSize. In a one-hop star BAN, frame exchanges occur directly between nodes and the hub of the BAN. In a two-hop extended star BAN, the hub and a node exchange frames optionally via a relay-capable node.

Access coordination at the MAC sublayer between BANs is not specified in this standard. Improved mechanisms for coexistence and interference mitigating between adjacent or overlapping BANs are provided. Nodes referenced in this standard are in the context of a given BAN, unless noted otherwise.

All nodes and hubs are internally partitioned into a physical (PHY) layer and a medium access control (MAC) sublayer, in accordance with the IEEE 802 reference model, as shown in Figure 1.5. Direct communications between a node and a hub transpire at the PHY layer and MAC sublayer as specified in this standard; the PHY layer and MAC sublayer of a node or a hub use only one operating channel at any given time. Message security services occur at the MAC sublayer, and security key generations take place inside and/or outside the MAC sublayer. Within a node or a hub, the MAC provides its service to the MAC client (higher layer) through the MAC service access point (SAP) located immediately above the MAC sublayer, while the PHY provides its service to the MAC through the PHY SAP located between them. On transmission, the MAC client passes MAC service data units (MSDUs) to the MAC sublayer via the MAC SAP, and the MAC sublayer passes MAC frames (also known as MAC protocol data units or MPDUs) to the PHY layer via the PHY SAP. On reception, the PHY layer passes MAC frames to the MAC sublayer via the PHY SAP, and the MAC sublayer passes MSDUs to the MAC client via the MAC SAP. Both MAC SAP and PHY SAP are not exposed and their specifications are beyond the scope of this standard.

There may be a logical node management entity (NME) or hub management entity (HME) that exchanges network management information with the PHY and MAC as well as with other layers. The HME is a superset of the NME in terms of the management functionality they each support. However, the presence of the NME or HME and the partitioning between the NME or HME and the MAC or the PHY is not mandated, nor is the behaviour of the NME or HME specified, in this standard.

7.2.2 Time base

All nodes and hubs establish a time reference base, as shown in Figure 3, if their medium access is to be scheduled in time, where the time axis is divided into beacon periods (superframes) of equal length and each beacon period (superframe) is composed of allocation slots of equal length and numbered from 0, 1,..., s, where $s \leq 255$. An allocation interval may be referenced in terms of the numbered allocation slot comprising it, and a point of time may be referenced in terms of the numbered allocation slots preceding or following it as appropriate.

If time reference is needed for access scheduling in its BAN, the hub is required to choose the boundaries of beacon periods (superframes) and hence of the allocation slots therein. In beacon mode operation for which beacons are transmitted, the hub needs to communicate such boundaries by transmitting beacons at the start or other specified locations of beacon periods (superframes), and optionally timed frames (T-Poll frames) containing their transmit time relative to the start time of current beacon period (superframe). In non-beacon mode operation for which beacons are not transmitted but time reference is needed, the hub is required to communicate such boundaries by transmitting timed frames (T-Poll frames) also containing their transmit time relative to the start time of current superframe.

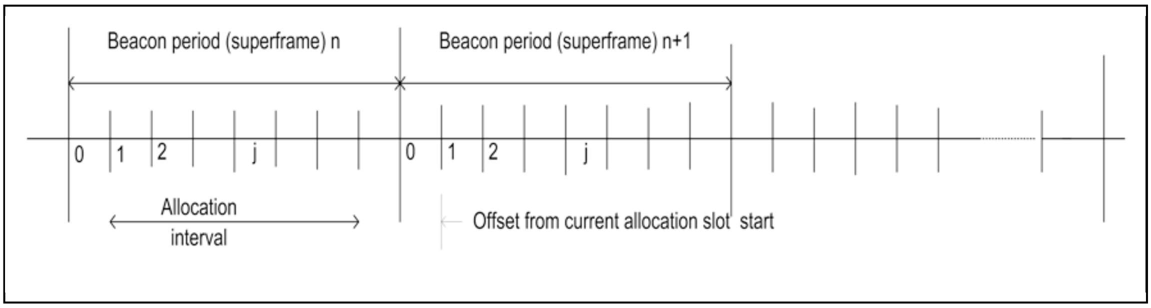


Figure 7. 1 Time reference base

A node requiring a time reference in the BAN needs to derive and recalibrate the boundaries of beacon periods (superframes) and allocation slots from reception of beacons or/and timed frames (T-Poll frames).

A frame transmission may span more than one allocation slot, starting or ending not necessarily on an allocation slot boundary.

The IEEE 802.15.6 standard defines a MAC layer that provides services for Physical layers including Narrowband, Ultra-wideband and Human Body Communications (HBC) layers[102]. According to the standard the WBAN is consisted of one hub/coordinator and some nodes (sensors), ranging from zero to maximum WBAN Size. A single coordinator or hub controls the entire operation in data transmission. The time axis or channel is usually divided into beacon periods or superframes of equal length. Each superframe contains a number of allocation slots and these slots have equal time duration. The coordinator transmits beacons to determine the superframe boundaries and allocate slots.

The MAC frame structure contains three parts: Beacon, Downlink and Uplink. Beacon is used to describe frame structure and guarantee synchronization. It also contains information which should be periodically broadcasted to nodes. The Downlink is set for data transmission from the coordinator to nodes and the function of Uplink part is vice versa. The Uplink part usually has two sub-parts: Contention Access Part (CAP) and Contention Free Part (CFP). CAP is based on CSMA/CA or slotted ALOHA and it is also used to transmit MAC control information. In CFP Guaranteed Time Slot (GTS) is designated to nodes for scheduled data transmission.

7.2.3 Beacon

A beacon frame contains a Frame Payload that is formatted as shown in Figure 7.2. It is locally broadcast by a hub in every beacon period (superframe).

2) Non beacon Mode with Superframe Boundaries. In this mode, the hub operates during the MAP period only.

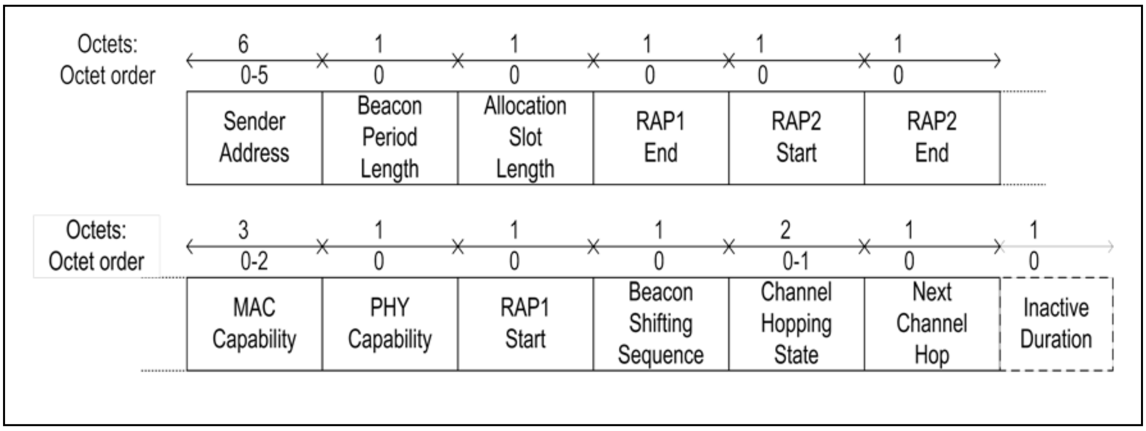


Figure 7. 2 Frame Payload format for Beacon frames

3) Non beacon Mode without Superframe Boundaries. In this mode, the hub provides unscheduled polled or posted allocations or a combination of both.

Usually the beacon mode with Superframe Boundaries is widely accepted in application so this model is applied in our study. Although it defines several access mechanisms, such as random, improvised, unscheduled and scheduled access, it doesn't explicitly illustrate how to apply these mechanisms and obtain the efficient usage of medium. The duration of different schemes is not adaptively allocated based on the traffic characteristics. Therefore a game theory scheme is proposed and applied to obtain flexible and efficient allocation of resource in WBAN.

The MAC protocol of WBAN aims to manage the usage of the medium access mechanism for a short range, low cost, low power and highly reliable wireless communication. Generally there are three main medium access mechanisms for WBAN including: contention access of CSMA/CA, scheduled access of guaranteed time slots and improvised access of polling [95].

A. CSMA/CA

CSMA/CA is a connect-less service to transfer asynchronous message with no guarantee of bandwidth and latency [88, 103]. It adapts quite well in the variable condition of traffic and is quite robust against interferences. In CSMA/CA scheme a node begins with listening the status of channel before sending. If the channel is found to be idle, the node sends the first packet in the transmit queue. If the channel is busy (either another node transmission or interference), the node waits the end of the current transmission and a random period of time, which means starting the contention). The CSMA/CA applies collision avoidance instead of collision detection in wired Ethernet because it is difficult to detect collision in wireless communication. But it can't acquire high efficiency for connection oriented services.

B. Guaranteed time slots

Guaranteed time slots is a time division multiplex access (TDMA) scheme[70]. The base station or the coordinator is responsible for nodes regulation in a network. The channel is divided into time slots, which are normally of fixed size. Every node in the network can be allocated one or more slots where it can transmit data. Slots are usually constituted in a frame and the coordinator determines the frame structure in a beacon. Each node follows the instruction in the beacon. This mechanism is suitable for predictable needs and is available to achieve low latency and guarantee bandwidth. Due to the inflexibility and connection oriented it is not appropriate for variable size packets and the burst traffic. This mechanism heavily depends on the quality of the channel.

C. Polling

In polling access the coordinator totally controls over the channel, but the frame structure is no more fixed and variable size packets are allowed to transmit. With a specific poll packet, the coordinator triggers the transmission scheme by the node. When the node receives a poll packet, and then sends what it wants to transmit.

Polling is a scheme between a connection oriented service (TDMA but with flexible packet size) and connection less-service (asynchronous transmission). The coordinator can either totally use polling of the network to check if all nodes have something to send ,which is available for limited nodes, or use reservation slots where each node can request a connection to transmit data.

7.3 WBAN Performance analyses

The purpose of WBAN is to offer reliable, flexible, high channel efficiency and cost saving services for professionals and patients. In order to evaluate the performance of a WBAN, latency, energy consumption, successfully received packets and breakdown packets are key factors to be measured with simulation.

A. Latency

In a network, latency is an expression of how much time it takes for a packet of data from one designated point to another. For transmitting critical data in WBAN time constraint is extremely critical so that data should be received within minimum latency. Usually latency can be given by (7.1):

$$\text{Latency} = O_{\text{waiting}} + Tx_{\text{Time}} = O_{\text{waiting}} + \frac{F}{R} \tag{7.1}$$

Where waiting is the overall waiting time, which contains queuing latency, acknowledgement time and contention window period in CSMA/CA. Tx Time is the transfer time for data, which is determined by F and R, F is the data frame size and R is data bit rate. For example: as defined in [104] the total duration of the frame:

$$T_{X\,Time} = 8 * \{\frac{L_{PHY} + L_{MACHDR} + L_{address} + x + L_{MACFIR}}{R_{data}}\} \tag{7.2}$$

L_{PHY} is the Length of the PHY header with 6 bytes, L_{MACHDR} is the Length of the MAC header with 3 bytes, $L_{ADDRESS}$ is the Length of the MAC address info field and the length of 1 PAN-identifier is 2 bytes, x is the number of bytes that are received from the upper layer, L_{MACFTR} is the Length of the MAC footer with 2 bytes and R_{data} is the Raw data rate.

Therefore latency increases linearly with the amount of nodes and it is terrible for emergency data with waiting a long queuing. If emergency data use reserved slots and are transferred without queuing latency then the system can fulfil transmission in the shortest period.

B. Energy consumption

The energy consumption is different from various communication techniques. For example energy is consumed for sensing operation. A large amount of energy is wasted in " idle listening" period. When two nodes want to communicate with each other, both of them should be active during this time. If there is not proper scheduling, the receiving node does not know the purpose of the transmitting node. The receiving node has to continuously listen to the channel.

The transmission power is set by a transmitter and the energy in our system is the model in reference [68] and can be defined by (7.3):

$$E(R) = P_l t_l + P_{tx} \times (t_{tone} + \frac{F}{R}) + P_{rx}(t_{tone} + \frac{F}{R}) \tag{7.3}$$

where P_l is the consumed power in the state of listening, t_l is the time of listening, t_l,t_{tone} are the channel listening and tone sending time respectively, P_{tx} is the consumed power in the state of transmitting, P_{rx} is the consumed power consumption in the state of receiving state. So it is available to obtain the network's energy expenditure at various rates.

C. Throughput

The throughput is related to the data rate and the probability of successful packets during the transmission. It represents the real successful data in specific period and can be defined as the average amount of packets which are transmitted in a unit time.

For CSMA/CA the throughput can be defined as the ratio of the payload size to the total deliver time for each payload[88] in (7.4):

$$S_C = \frac{B * x}{Deliverytime(x)} \tag{7.4}$$

where x is the payload, B is the bits in one packet payload, *Deliverytime(x)* is associated with average contention windows, data transmission time, immediate acknowledgement time, short inter frame spacing time and propagation delay.

While for GTS the throughput can be defined as (7.5) [70]:

$$S_G = \frac{P_S(t)L}{\theta T_{SS}} \tag{7.5}$$

$P_S(t)$ is the probability of a successful GTA allocation, L is the length of payload, θ is the number of superframe slots, T_{SS} is a superframe slot length.

For a whole network when CSMA and GTS are both used in the MAC, the total throughput can be defined in (7.6):

$$S = \frac{S_C \times N_C + S_G \times N_G}{N_C + N_G} \tag{7.6}$$

where N_C is the number of slots for CSMA/CA, which is the period for RAP in 802.15.6. N_G is the number of slots for GTS. So the normalized system throughput is associated to the access schemes, which affects the successful probability, the effective slots and latency for data transmission.

D. Successful received packets

Usually the system throughput of the superframe is decided by the successfully payload in the delivery time. But sometimes it is not easy to accurately measure or calculate the time and the throughput is not the primary factor in special occasions so successful received packets in equation 7.7 can also be used to evaluate the performance to ensure the maximum packets are received:

$$S_i = P_{si} \times L \times A \tag{7.7}$$

where A is a normalization constant to convert payload to packets in designated period, L is the payload, P_{si} is the probability that any node successfully transmit data in designated period. The probability can be affected by different access schemes and channel states.

7.4 Intelligent medium access model and simulations

As previously described, either GTS or Polling usually requires a service slot or reservation slot. If the MAC is connection oriented and the rate of new connection is low, maybe a single service slot is enough. If the MAC is packet oriented and the requested rate is high, then the protocol needs to offer more reservation slots. On the other side for some variable data rate, CSMA/CA is the preferred scheme due to the flexibility. Therefore intelligent management of various access schemes is in need and can effectively improve the utilization of bandwidth and obtain better performance. All nodes in a network act like competitors and they share the same channel to transmit data. Each node's action will affect other nodes' access. This competition is consistent with the principle of Game theory. So Game theory can be used as an effective and smart method to implement the resource allocation in MAC for WABN. Hence we propose an intelligent manage access mechanisms with game theory for various conditions including traffic requirements and channel states.

7.4.1 Introduction of Game theory

Game theory is a theoretical framework that attempts to mathematically acquire both human and non-human (e.g. computer, animal, plant) behaviour during a strategic situation[105]. A strategic situation is a situation that involves the interaction of two or more entities where the individual's success depends on the choice of actions by others. A logical behaviour would try to find equilibria between all entities (named as players), for example: sets of actions called as strategies that players will not want to change, since if they will most probably benefit less. Therefore, game theory can be used to model situations of interaction and can offer solutions so that mutually agreeable sequences of actions can be employed by the players. Game theoretic models make the assumption that the entities make rational choices, which are profitable according to each entity's own interpretation of profit.

A strategic situation, where the actions of a participant may alter another's outcome, is primarily characterized by the players' strategies. In addition, a strategic situation contains other elements that must be taken into consideration when modelling such a situation as a game, e.g. chance and skill (elements that are not easily controlled or modified). Game models, i.e., models of specific strategic situations, may be categorized in various ways due to the several elements that they contain.

Figure 7.3 illustrates all of the possible actions in game trees that can be taken by all of the players and indicating all of the possible outcomes of the game.

Each node represents a point of choice for a player. One player is illustrated as a node. The links between the nodes represent a possible action for that player. The payoffs are specified at the bottom of the tree.

In Figure 7.3 there are two players. Player 1 moves first and chooses either A or B. Player 2 knows the move chosen by Player 1 and hence chooses C or D. Each sequence of choices from the players results in a different set of payoffs for the two players. Depending on the outcome of a game, a payoff is provided for each player, which reflects their gain at the end of the game (or each round of the game in the case of an iterative game).

The normal form game (a.k.a. strategic form) is usually represented by a table which shows the players, strategies, and payoffs, as the example in Table 7.1. In the example there are two players; one chooses the row and the other chooses the column. Each player has two strategies, which are specified by the number of the row and the number of the column. The payoffs are provided in the interior. Two numbers are provided, one for each player. The first number is the payoff received by the row player (Player 1); the second is the payoff for the column player (Player 2). Suppose that Player 1 plays A and that Player 2 plays C. Then Player 1 gets a payoff of 4, and Player 2gets 3. When a game is presented in normal form, it is presumed that each player acts simultaneously or, at least, without knowing the actions of the other. Every extensive game form has an equivalent normal game form.

Furthermore, an interaction may happen only once or repeatedly; in the first situation we are faced with one-shot game models, while the second situation requires repeated game models. Many of the real-world problems modelled employ game players that interact infinitely many moves (one or more). However, one class of games studied is the one of infinitely repeated games, a model used also for games whose horizon (i.e., the number of moves) is not known.

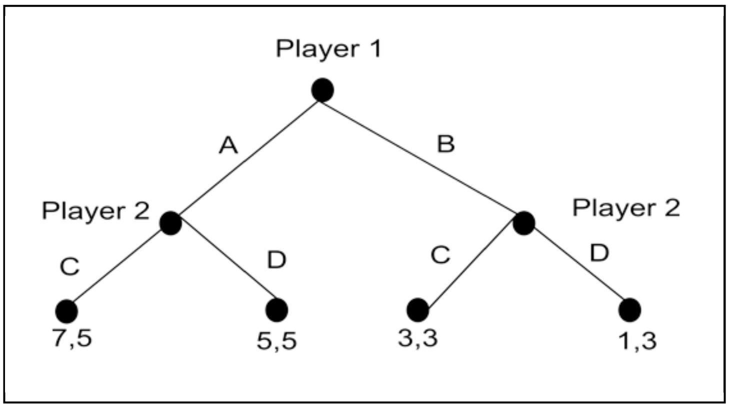


Figure 7. 3 An example tree of a sequential form game

The focus of attention is usually not so much on what is the best way to play such a game, but rather on whether one or the other player has a winning strategy.

Table 7. 1 An example table of a normal game form

	Player 2 plays C	Player 2 plays D
Player 1 plays A	4,3	2,1
Player 1 plays B	0,0	-1,-2

An additional dichotomy is whether the players are in complete conflict, where the model employed is a non-cooperative one, or they have some commonality, where a more cooperative game model may be more appropriate. Such commonality could be, for instance, that the players by cooperating increase their individual payoff or even that the players participate in groups and the payoff is given to a group and not to an individual player. Therefore, in a cooperative game the players are able to form binding commitments. In non-cooperative games this is not possible. Of the two types of games, non-cooperative games, and by non-cooperative we refer to the case where the players in the game are antagonistic and the individual players are modelled in detail, are able to model situations to the finest details, producing more accurate results on the individual level, whereas cooperative games, situations where the group payoff is studied instead of the individual, focus on the game at large. Hybrid games contain cooperative and non-cooperative elements. For instance, coalitions of players are formed in a cooperative game, but these play in a non-cooperative way.

Games are motivated by profitable outcomes that await the players once the actions are taken. These outcomes are referred to as payoffs. Payoffs for a particular player capture everything in the outcomes that the particular player cares about. If a player faces a random prospect of outcomes, then the number associated with this prospect is the average of the payoffs associated with each component outcome, weighted by their probabilities.

The solution to a strategic game is derived by establishing equilibria. Equilibria may be reached during the interaction of players' strategies when each player is using the strategy that is the best response to the strategies of the other players (i.e., given the strategies of the other players, the selected strategy results in the highest payoffs for each player participating in the game).

The idea of equilibrium is a useful descriptive tool and furthermore, an effective organizing concept for analysing a game theoretic model. For normal form games the Nash Equilibrium is used as a solution concept, where every player's action is the best response to the actions of all the other players.

For sequential-moves games, e.g., repeated games, the equilibrium used is known as the subgame perfect equilibrium or the rollback equilibrium (for finite repeated games). In such games the players must use a particular type of interactive thinking; players plan their current moves based on future consequences considering also opponents' moves. Therefore, the equilibrium in such a game must satisfy this kind of interactive thinking, and subgame perfect equilibrium does exactly that, by planning the best responses for every possible subgame or interaction.

Game Theory provides appropriate models and tools to handle multiple, interacting entities attempting to make a decision and seeking a solution state that maximizes each entity's utility. Game Theory has been extensively used in networking research as a theoretical decision-making framework, e.g., for routing, congestion control, resource sharing and heterogeneous networks. These interactions may benefit from such a theoretical framework that considers decision-making, interacting entities. We advocate that the game theoretical framework is suitable for generating profitable behaviours/strategies for interacting entities in conflicting situations, and we explore its application upon seemingly conflicting interactions, as for example those occurring in converged heterogeneous communication networks.

A game model comprises a set of players, which choose their actions in each period of the game to maximize that period's expected payoff. Payoff usually represents profit, quantity, utility, or other continuous measurement or simply illustrates the desirable outcomes[106]. Game theory has been used for medium access control. For example: Tao cui provided a general game theoretic framework for contention based medium access control; Li jun used a game-theoretic approach to control contention [107, 108].

A static non–cooperative game model is used for access control in this study. The non-cooperation game reflects a competitive situation where every player needs to take its decision independently of other players, giving the possible choices of other players and their effect on the players' utilities.

Definition 1: A general random access game G is defined as a triple $G = \{N, (S_i, i \in N), (U_i, i \in N)\}$, where $N = (1, \ldots, n)$ is the finite set of players, S_i is the available strategies and U_i is the utility function to evaluate the payoff. In our analysis N represents wireless nodes in a WBAN. Choosing the access scheme is the strategy that every node uses.

In game theory, Nash Equilibrium (NE) is a solution concept of a game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing only its own strategy unilaterally [106]. We use the best response concept to calculate the Nash Equilibrium. If a player A has a dominant strategy s_A, then there exists a Nash equilibrium where A plays s_A. For two players A and B, there exists a Nash equilibrium in which A plays s_A and B plays a best response to s_A. If both A and B have strictly dominant strategies, there exists a unique Nash equilibrium in which each plays his strictly dominant strategy.

The performance of a network can be analysed by various utility functions for MAC. These specific definitions are useful for different conditions and strategies to approach the equilibrium of the game, which is a stable state that nodes choose the best action for a given utility. Therefore a series of utility functions is proposed for various conditions.

7.4.2 Intelligent medium access model

We formulize the system model into a static game theory model, where all players make decisions rationally either simultaneously or asynchronously without knowing the decisions of other players. We design the utility function to analyse the behaviour of sensors and acquire the best response.

A. Normal status

For normal network status we use game theory to obtain the maximum payoff with defined utility function. Nodes (sensors) are players in this game and all nodes make their decision simultaneously.

Since we assume that a player's object is to maximize the utility and the coordinator determines strategies to acquire maximum payoff in normal status.

Because packets, latency and energy are key parameters to evaluate the system performance we use these factors to define the utility in equation (7.8) to represent the payoff,

$$U_N = \sum_{i=1}^{n} U(i) = \Sigma \frac{U(S_i)}{K \times U(L_i) \times U(E_i)} \tag{7.8}$$

where U(S) is the successful received packets in designated period, U(L) is the average latency for packets, U(E) is the consumed energy by all nodes and K is a coefficient to adjust the utility according to network requirement and conditions. When the utility is energy-oriented, the K is set as the maximum value to emphasize the influence of consumed energy. The constraint for the function is that the total allocated slots in the control process cannot exceed maximum available slots B. The utility goes up with the increasing successful packets and the decreasing latency and consumed energy. By varying the access scheme the system gets the maximum U_N as the best response as the equilibrium.

B. Packets oriented

For some packets oriented services the maximum amount of successful packets is a major demanded factor, so nodes should be allocated enough slots to guarantee the requirement. For example if the data rate is 1024Kbps and the slot allocation length is chosen to be 10ms. A 128 byte data packet needs 1ms to be transmitted with the BAN radio so the total period for a packet (radio transition times+ TX + ACK) is 1.16ms. It means each allocation slot can hold 8 packets. Considering the packet rate the required slots are calculated by (7.9):

$$Slots = \frac{PR \times BP}{NP} \tag{7.9}$$

PR is the packets rate, BP is the time of a beacon period and NP is the number of packets in a slot. Therefore the utility for packets oriented services is decided by (7.10):

$$U_p = \sum_{i=1}^{n} U(Si) \tag{7.10}$$

For a node of the highest rate 100 packets/second, calculating a theoretical required number of slots is 4. At this condition the node gets the highest successful packets and the system obtains the best value of the successful received packets from all nodes as shown in figure based on the simulation. When the allocated slots are 4 the amount of successful received packets reaches the maximum and more slots can't improve the performance. Also for the rate of 50 packets/second and 25 packets/second, two scheduled slots are enough for data transmission. So choosing suitable slots in GTS can get the best response of the utility.

C. Emergency

If the coordinator detects emergency information, such as: a person's falling or abnormal signals of critical specific nodes by data analysis[67], it turns into emergency state and triggers new allocation at the next frame. At this stage, some nodes may increase their sampling rate to guarantee reliable monitoring and the polling mechanism is used for high data flow without considering energy consumption.

As defined in 802.15.6, if the coordinator gets a packet of one node whose more Data field is 1, it means this node has one extra packet to transmit. After the coordinator counts the number of such packets, it sends a poll message enclosed in the ACK packets. Therefore nodes obtain required polling-based slots.

These nodes are treated as selfish players[99], which try their best to occupy the time of channel for maximum transmission probability.

At this condition the successful packets (7.7) of specific nodes or one node is used to define the utility function.

$$U = \sum_{j=p}^{q} U(S_j), p, q \in (1,2,...n) \tag{7.11}$$

Where "p, q " is the serial number of specific nodes.

These selfish players can increase their expected payoff at the expense of other nodes by using polling strategy. We focus on the individual utility of each sensor in this condition.

D. Bad link status

The link status comes from the recently information of the channel state. If the amount of consecutive breakdown packets increases heavily and then reaches a threshold, it means the channel condition is worse. The access scheme should be modified and the increasing of random access period can decrease the possibility of breakdown packets. Because GTS is highly depends on the channel status while CSMA/CA is flexible in slots using. So using more slots of CSMA/CA is preferred for this condition and the throughput in (7.3) is used as the utility. The allocation of large random access slots can bring benefits to the value of the payoff for this condition.

7.5 Simulation and analysis

We consider a body area system with N wireless sensors to one designated receiver (coordinator) and we assume all nodes are working in the same communication range (i.e., every node can hear other nodes to avoid complicated problems caused by the hidden terminal). The simulation is executed based on the protocol of IEEE 802.15.6 because it already defines the MAC frame format and communication modes. Some simulation parameters are presented in table 7.2.

The duration for each simulation operation is 50 seconds and the results are averaged measurement over 5 repeated runs.

Simulation 1: for normal status: We set simulation for 6 nodes (one coordinator and five nodes for data acquisition) in a wireless body area network, in which the data rates of sending nodes are 25 packets/second, 50 packets/second and 100 packets/second respectively.

The utility is calculated to evaluate the access scheme as shown in figure 7.5. The y axis is the utility and the x axis represents different strategies. “No-polling&no GTS” means only CSMA is deployed in the MAC, “no-plooing&GTS” means CSMA and GTS are deployed in the MAC, “polling&GTS” means polling and GTS are both used in the MAC.For low data rate (25 packets/second with blue bar), three different strategies get the similar results. While for high data rate (100 packets/second with pink bar) polling&GTS obtain the maximum payoff. At the same situation CSMA (no-polling&noGTS) the payoff of Utility is the lowest. Therefore we can use the function, which is defined by successful received packets, latency and consumed energy, to get a balanced evaluation and choose the appropriate mechanism for all nodes to make a more efficient use of the wireless medium. It shows that for low rate using CSMA/CA and GTS with small slots can reach the high utility while for high rate GTS with more slots is better. TDMA schemes of scheduled slots can make a more efficient use of the wireless medium and high utility for high data rate.

For normal wireless protocol of CSMA/CA, when a device wants to transmit data, the device waits for a random period of back off before trying to access the channel. The back off time is randomly generated in the interval [0, 2BE-1]. This will cause more collisions and then reduce the network throughput, increase latency especially for high data rate. With intelligent manage of the access, the resource can be dynamically allocated according to rate and acquire better utility as shown in Figure 7.4. Simulation 2: For packets-oriented condition we acquire the payoff of all successful packets with different allocated slots and the result is shown in figure 7.5. The y axis is the average packets received from all nodes and the x axis is the sending rate for each node measured in packets/second.

Nodes are sending packets with 50 packets/second, 100 packets/second and 200 packets/second. According to (5), adequate slots assignment is the primary factor to ensure the best response. When the data rate is low the requested GTS is small. Otherwise nodes demand more slots for maximum payoff. When the rate is 100 packets/second and the access scheme of 4 slots is allocated in GTS the coordinator can receive the most packets from every node. While the rate increases to 200 packets/second, the overall received packets do not increase and keep the same value as shown in the green line, which means the slots allocation can't meet the requirement for high data rate. So when the rate is 200 packets/second 6 slots for GTS is the best strategy because it acquires more packets. Therefore dynamical adjust slots allocation according to operation requirement.

Table 7. 2 Simulation Parameters

Parameters	Values
Physical Data Rate	1024kbps
Network topology	Star
Packet size	128 bytes
Simulation time	50 seconds
Start up Delay	1 second
MAC Buffer Size	48 bytes
Tx Output Power	-10dBm
Receiver Sensitivity	-87dBm
Time slot size	10ms

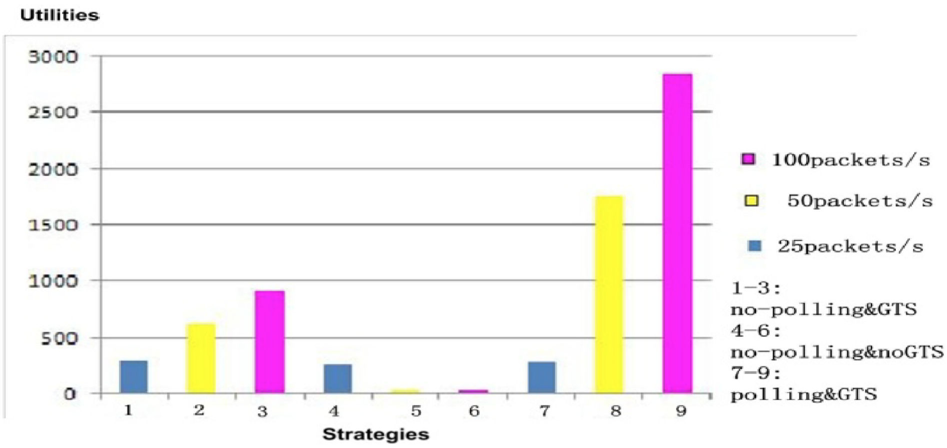


Figure 7. 4 Utilities of various accesses

Simulation 3: For emergency situation we assume a network, in which a node requires urgent services if abnormal signals are detected. Then the scheme of polling is triggered and specific slots are allocated for it. This node is regarded as a selfish player to maximize its payoff without considering other nodes' utility. In figure 7.6 we notice that in this situation the ratio of breakdown packets is the smallest when the node 2 uses the polling scheme. It means the ratio of successful packets, which is shown in pink colour, is the highest in all nodes. At this situation this player (node 2) obtains its best utility with smart decision due to the requirement.

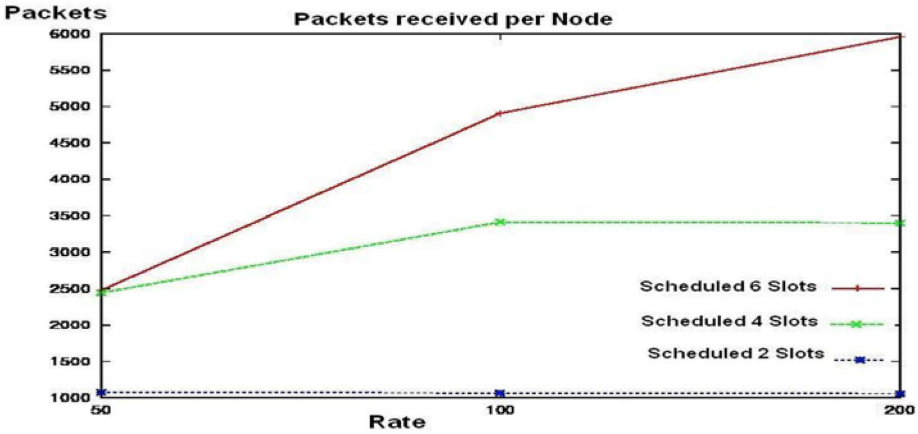


Figure 7. 5 Packets-oriented Access

Simulation 4: For bad link status we use a channel with high interference to test the access scheme. The period of contention access of 2 slots, 6 slots and 16 slots is defined as short CAP, medium CAP and long CAP respectively in figure 7.7. As defined before, the utility is determined by throughput. For low data rate (25 packets/second) the increase of contention access period does not improve while for high data rate (100 packets/ second) the raising of utility is obvious. Therefore for high data rate in bad link status using long contention period is beneficial to the system performance.

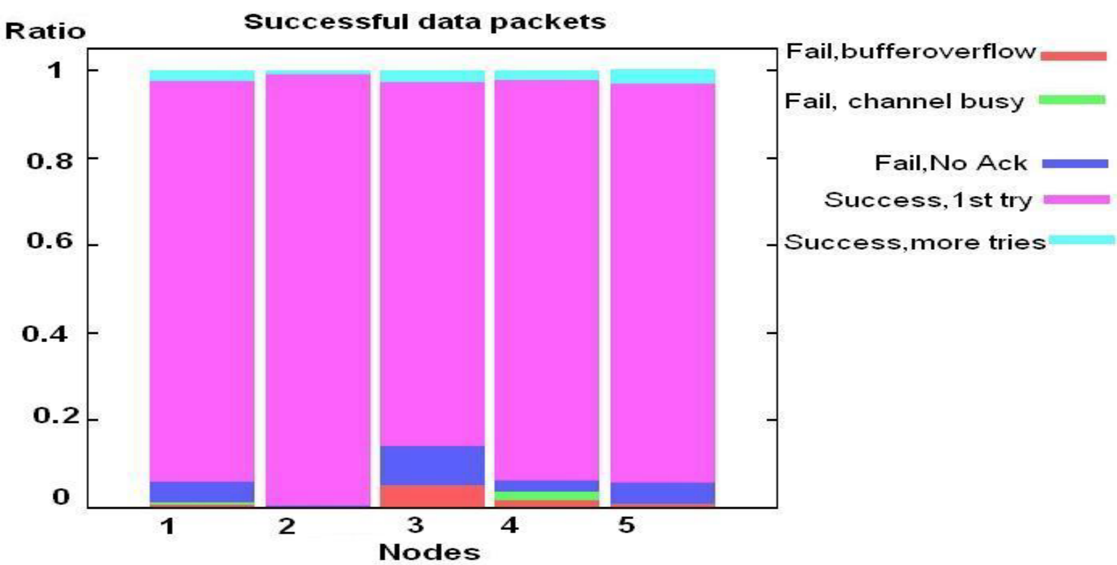


Figure 7. 6 Successful data packets

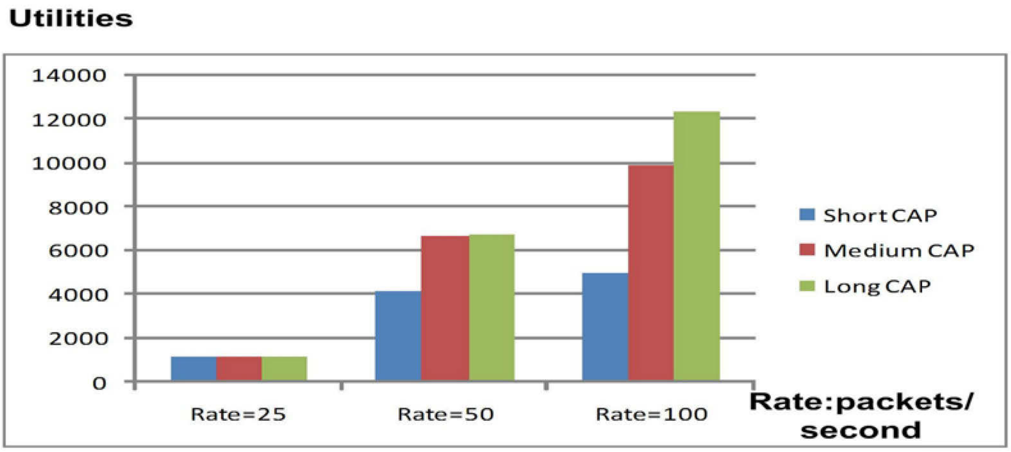


Figure 7. 7 Bad link statuses

7.6 Conclusions

For WBAN medium access control for wireless communication is a key factor to improve the system performance. Using game theory to intelligently manage multiple access schemes is proposed in this paper for better utility. Players in the game model choose various strategies in different situations to maximize the expected payoff. The coordinator is responsible for arranging access schemes to obtain better resource allocation or meet specific requirement. All nodes observe the demanding and channel state to ensure the dynamic and intelligent model to be implemented. The proposed model is evaluated in the simulation to illustrate the benefits. In the future research more cognitive knowledge in the network will be considered to ensure high quality of services.

Chapter 8 Conclusions and Future Work

8.1 Development of WBAN in healthcare

In recent years, Wireless Body Area Networks provide ubiquitous real-time monitoring and / or actuation and attract attention in medical applications not only for prevention and early detection of medical problems but also for assisting surgeons in minimally invasive surgeries. It is possible to see in the future healthcare related sensing will be a convergence in many fields—health, environment, activity, and lifestyle. The expense of health care grows up every year so people will look for affordable, portable and comfortable medical devices, especially non-prescribed monitoring gadgets. Also many people want to record their continued health status with easy to use, comfortable and non-intrusive devices. Now patients have more awareness of the technologies and services that have appeared in the health and fitness areas and they are well educated. So they will have more capabilities in the decision making process with healthcare therapies and have questioning and demanding more of their healthcare providers.

The use of personalized medicine will accelerate in the future and it will become a standard practice for healthcare. With new technologies genomic information will be used to choose the most effective treatment for an individual person rather than the treatment based on population-centred statistics. The individual treatment will be tailored and continuously modified with the real time biological data. Sensing will play an important role in this process, including determining effective drug and monitoring the efficacy. With the evolution of MEMS and the increased integration technology, the sensors' size will shrink and meet the expectations and requirements of patients. Novel epidermal sensing technologies, for example "smart skin," in which sensors can be tattooed onto the skin, will enable vital signs are monitored by clinicians.

Another exciting body sensor network platform is the using of smart phones. Due to the pervasiveness and familiarity they can obtain better adoption for sensing and management devices. Some major smart phone companies, such as Samsung and Apple, are delivering future devices and services based on a combination of existing sensors, innovative smart phones, and cloud-based software.

Innovative products will provide continuous, reliable sensing in any places and any times according to an individual's required lifestyle. This will strengthen the cloud services and related requirements including reliability, security, privacy, configurable access and integrated electronic technologies and can offer many potential benefits to society. WBAN is expected to avoid failure and enhance throughput for various application. Therefore new health communication standards and approaches must be developed to meet stringent technical requirements especially for monitoring of life related signals, such as indicators of a heart attack.

8.2 Conclusions

New approaches to sensing, such as the development of biosensors and microelectromechanical systems (MEMS) technology, have significantly increased the availability of WBAN applications. But the challenge of reliability, which is affected by special characteristics including remote deployment, energy consumption and resource constraints, will increase. Further, the requirement for reliability of WBAN is higher than other WSN due to the sensitive data, health care and medical services. The reliability continues to be an active research direction for further applications.

The aim of the research illustrated in this thesis is to explore new approaches to improve the reliability of Wireless Body Area Networks. The technologies that we use to improve the performance of wireless sensor and wireless body area networks are based on fault management and fault prevention with consideration of requirements, limitations and characteristics of WBAN.

This thesis firstly provides a brief overview of WBAN development and application and it looks at many factors, which affect applications for health related services. Then it puts forward solutions to enhance the reliability with fault detection and identification. After the description of major types of MAC protocols and the illustration of a number of challenges for the operation of the IEEE 802.15.4 standard and IEEE 802.15.6 draft standard this thesis proposed novel mechanisms to control the medium access and manage wireless resources , which increase the performance including: throughput, latency, fairness and successful delivery packets. With these techniques WBAN achieves better reliability with adaptive controlling of access processes and intelligent management based on operation conditions.

The main contributions are outlined as following:

(1) A simple and effective algorithm is proposed for abnormalities detection of IMU in motion monitoring based on PCA. By reducing dimensions with PCA and calculate upper control limits this algorithm compares normal and abnormal samples to find abnormalities. The effectiveness of algorithm has been illustrated through the simulation and real measured IMU data from the laboratory.

(2) In order to guarantee the accuracy and availability of COSMED K4b^2, especially when it performs in various statuses or activities, we design a comprehensive fault detection and

identification procedure based on statistical techniques of PCA, clustering of K-means and BP neural network in multiple processes. It can be a generic detection model and easily implemented in sensors with significant features such as dimension reducing for less energy consumption in transmission, clustering for fault detect in different processes and identification of crashed sensors.

(3) MAC protocols should be designed to accommodate changes in network topology and traffic characteristics. The choice of a medium access protocol is a substantial part of fault prevention for high reliability. We deployed an automatically slots allocation model based on priority to avoid collision and prevent faults. This MAC mechanism is an effective way to achieve a consensus decision according to priorities while minimizing data latency for emergency messages.

(4) An optimal MAC scheme is proposed based on fuzzy logic system to improve performance of WBAN such as: collision avoidance, latency, throughput by guaranteeing multiple services of different conditions. This new mechanism introduces dynamic Superframe reconfiguration and utilizes fuzzy control scheme for contention window in CAP and slots allocation in CFP. The validity is evaluated by Castalia with specified nodes and obtained results show it significantly decreases latency and gains a high packet delivery ratio as compared to the standard IEEE 802.15.4 protocol.

(5) The using of wireless medium is a key factor to improve the system performance. Using game theory to intelligently manage multiple access schemes is proposed in this thesis for better performance. Players in the game model choose various strategies in different conditions to maximize the expected payoff. The coordinator is responsible for arrange access schemes to obtain better resource allocation or meet specific requirement. All nodes observe the demand and channel state to ensure the dynamic and intelligent model to be implemented. The proposed model is evaluated in the simulation to illustrate the benefits.

8.3 Future Work

Combined with supporting technologies, sensors such as accelerometers, ambient light sensors and magneto meters are becoming common in modern mobile devices. Sensor technologies will continue to advance, both in terms of improvements to existing approaches and the development of new sensing modalities. The crossover will range from health sensing into sports and wellness sensing and new demands are needed for sensor technologies and communication protocols.

There are many techniques, which may be driven by specific application requirements or other demands, can enhance the reliability. It is clear to see there are many exciting prospects for the further development of reliability in WBAN.

For health and medical application a more detailed and thorough research should be done to consider specific pattern. Future research may investigate the cross-layer implementation, coexistence of various network, interference mitigation, and new communication protocols to fulfil the ubiquitous healthcare. Cognitive knowledge and contextual sensing information may be taken into consideration and dynamic strategies can be performed for extremely high reliability. Self-healing mechanism should be proposed to avoid failure when node or network does not operate properly.

Reference

[1] I. P. C. I. US Census Bureau, "Global Population Profile: 2002," ed, 2004, p. 49.
[2] D. o. E. a. S. A. United Nations, Popolation Division. (2013). *World Popoulatuin Ageing*.
[3] M. Company. (2010). *mHealth: A new vision for healthcare*. Available: http://www.mckinsey.com/Search.aspx?q=mHealth&l=Insights%20%26%20Publications,
[4] J. G. Hilke Brockmann, "The cost of population aging: Forecasting future hospital expenses in Germany," *MPIDR Working paper,* p. 13, 2005.
[5] B. A. and B. M.D. (2008). *Advances in Heart Failure*.
[6] S. Heiermann, K. K. Hedayati, M. J. Muller, and M. Dittmar, "Accuracy of a Portable Multisensor Body Monitor for Predicting Resting Energy Expenditure in Older People: A Comparison with Indirect Calorimetry," *Gerontology,* vol. 57, pp. 473-479, 2011.
[7] C. P. Waltenegus Dargie, *Fundamentals of wireless sensor networks : theory and practice*: John Wiley &Sons Ltd, 2010.
[8] F. Jacob, *Handbook of Modern Sensors*, 4th ed ed. New York: Springer, 2010.
[9] K. Y. Chen, K. F. Janz, W. Zhu, and R. J. Brychta, "Redefining the roles of sensors in objective physical activity monitoring," *Medicine and Science in Sports and Exercise,* vol. 44, pp. 13-23, 2012.
[10] C. N. S. Michael J. McGrath (2013). *Sensor Technologies:Healthcare, Wellness and Environmental Applications*. Available: http://it-ebooks.info/book/3253/
[11] W. M, " The computer for the 21st century," *Scientific American,* vol. 265, pp. 94-104, 1991.
[12] S. S. Garg SK, Edelman SV, " Improved glucose excursions using an implantable real-time continuous glucose sensor in adults with type 1 diabetes," *Diabetes Care,* vol. 27, pp. 734-738, 2004.
[13] P. A. Steil GM, Rebrin K. , "Closed-loop insulin delivery-the path to physiological glucose control. ," *Advanced Drug Delivery Reviews* vol. 56, pp. 125-144, 2004.
[14] M. Yuwono, B. D. Moulton, S. W. Su, B. G. Celler, and H. T. Nguyen, "Unsupervised machine-learning method for improving the performance of ambulatory fall-detection systems," *Biomedical Engineering Online,* vol. 11, Feb 16 2012.
[15] D.-J. Kim and B. Prabhakaran, "Motion fault detection and isolation in Body Sensor Networks," *Pervasive and Mobile Computing,* vol. 7, pp. 727-745, Dec 2011.
[16] P. Zappi, T. Stiefmeier, E. Farella, D. Roggen, L. Benini, and G. Troester, *Activity recognition from on-body sensors by classifier fusion: sensor scalability and robustness*, 2007.
[17] "<Embedded self-healing layer for detecting and.pdf>."
[18] G. Wu, J. Ren, F. Xia, and Z. Xu, "An Adaptive Fault-Tolerant Communication Scheme for Body Sensor Networks," *Sensors,* vol. 10, pp. 9590-9608, Nov 2010.
[19] O. Aziz, B. Lo, A. Darzi, and G. Yang, *Body sensor network*: Springer, 2011.
[20] I. Koren, *Fault tolerant systems*.
[21] S. B. Lo, R. K. Thiemjarus, and a. G. Yang, " Body sensor network–a wireless sensor platform for pervasive healthcare monitoring," presented at the the 3rd International Conference on Pervasive Computing, 2005.
[22] J. Puentes, Julien Montagner, Laurent Lecornu, and Jaakko Lahteenmaki,, "Quality Analysis of Sensors Data for Personal Health Records on Mobile Devices," in *Pervasive Health Knowledge Management*, R. K. Bali, Indrit Troshani, and Steve Goldberg,, Ed., ed New York: Springer, 2013, pp. 103-134.
[23] E. Sobahni-Tehrani, "Fault detection, isolation, and identification for nonlinear systems using a hybrid approach.," Master, Applied Science in Electrical and Computer Engineering, Concordia University, Montreal, Quebec, Canada, 2008.
[24] D. Törnqvist, "Statistical Fault Detection with Applications to IMU Disturbances," Division of Automatic Control Department of Electrical Engineering, Linköpings universitet, . , Linköping, Sweden, 2006.

[25] É. J. d. Oliveira, I. M. d. Fonseca, and H. K. Kuga, "Fault Detection and Isolation in Inertial Measurement Units Based on -CUSUM and Wavelet Packet," *Mathematical Problems in Engineering,* vol. 2013, pp. 1-11, 2013.

[26] Shuguang Sun and R. Liu:, "Error Calibration and FDI Technology of Gyros in Redundant IMU. ," in *International Workshop on Database Technology and Applications*, Hubei, 2009, pp. 398-401.

[27] W.Haerdle and L. Simar., *Applied Multivariate Statistical Analysis,*, Second Edition ed. Heidelberg: Springer, 2007.

[28] A. Ibaida, I. Khalil, and F. Sufi, "Cardiac abnormalities detection from compressed ECG in wireless telemonitoring using principal components analysis (PCA)," presented at the Intelligent Sensors, Sensor Networks and Information Processing 2009.

[29] W. L. S. Valle, and S. Qin, , "Selection of the number of principal components: A new criterion with comparison to existing methods," *Ind. Eng. Chem. Res,* vol. 38, pp. 4389-4401, 1999.

[30] W.Haerdle and L. Simar, *Applied Multivariate Statistical Analysis, Second Edition*. Heidelberg: Springer, 2007.

[31] J.C.Eisenmann, N.Brisko, D.Shadrick, and S.Welsh, "Comparative analysis of the Cosmed Quark b super(2) and K4b super (2) gas analysis systems during submaximal exercise," *Journal of Sports Medicine and Physical Fitness,* vol. 43, pp. 150-155, 2003.

[32] A.I.d. Silva, L.C. Fernandes, and R. Fernandez, " Energy expenditure and intensity of physical activity in soccer referees during match-play " *Journal of Sports Science and Medicine,* vol. 7, pp. 327-334, 2008.

[33] A.Lay-Ekuakille, P.Vergallo, A.Trotta, A.Trabacca, M. D. Rinaldis, F.Angelillo, and L.Petrara, "Ergospirometric Signal Processing for Human Breath and Force Measurements," in *IEEE International Workshop on Medical Measurements and Applications Proceedings*, 2010, pp. 113-116.

[34] E. E. Kuffel, S. E. Crouter, J. D. Haas, E. A. Frongillo, and D. R. Bassett, Jr., "Validity of estimating minute-by-minute energy expenditure of continuous walking bouts by accelerometry," *Int J Behav Nutr Phys Act,* vol. 8, p. 92, 2011.

[35] D. K. Spierer, M. Hagins, A. Rundle, and E. Pappas, "A comparison of energy expenditure estimates from the Actiheart and Actical physical activity monitors during low intensity activities, walking, and jogging," *Eur J Appl Physiol,* vol. 111, pp. 659-67, Apr 2011.

[36] C. Srl, "K4b2 User manual, XVIII Edition," ed. Rome, Italy: http://www.cosmed.it, 2008, pp. 1-162.

[37] J. C. Eisenmann, N. Brisko, D. Shadrick, and S. Welsh, "Comparative analysis of the Cosmed Quark b(2) and K4b(2) gas analysis systems during submaximal exercise," *Journal of Sports Medicine and Physical Fitness,* vol. 43, pp. 150-155, Jun 2003.

[38] L. R. M. Naughton, R.Sherman, S.Roberts, and D.J.Bentley, "Portable gas analyser Cosmed K4b2 compared to a laboratory based mass spectrometer system," *Journal of Sports Medicine and Physical Fitness,* vol. 45, pp. 315-323, 2005.

[39] M. Gayda, L. Bosquet, M. Juneau, T. Guiraud, J. Lambert, and A. Nigam, "Comparison of gas exchange data using the Aquatrainer system and the facemask with Cosmed K4b2 during exercise in healthy subjects," *Eur J Appl Physiol,* vol. 109, pp. 191-9, May 2010.

[40] R. Duffield, B. Dawson, H. Pinnington, and P. Wong, "Accuracy and reliabillity of a Cosmed K4b 2 portable gas," *Sci Med Sport* vol. 7, pp. 11-22, 2004 2004.

[41] COSMED. K4b2 Cardio Pulmonary Exercise Testing [Online]. Available: http://www.cosmed.com/images/pdf/productliterature/K4b2_Brochure_EN_C09052-02-93_A4_web.pdf

[42] COSMED. (2012). *K4b2Cardio Pulmonary Exercise Testing*.

[43] I. T. Jolliffe, *Principal Component Analysis*. Secaucus, NJ, USA Springer 2002.

[44] L. H. Xie Tingfeng, Wu Jianjun, "Fault Detection and Diagnosis for Sensors of LRE Based on PCA," *Journal of Astronautics,* vol. 28, pp. 1668-1672, 1703, 2007.

[45] L. Yu, J.-h. Zhu, and L.-j. Chen, "Parametric Study on PCA-based Algorithm for Structural Health," presented at the Prognostics & System Health Management Conference, Macao, 2010.

[46] M. P. S. Chawla, "PCA and ICA processing methods for removal of artifacts and noise in electrocardiograms: A survey and comparison," *Applied Soft Computing,* vol. 11, pp. 2216-2226, 2011.
[47] Y. L. Borgne and G. Bontempi, " Unsupervised and supervised compression with principal component analysis in wireless sensor networks," in *13th ACM International Conference on Knowledge Discovery and Data Mining*, 2007, pp. 94–103.
[48] M. A. Rassam, A. Zainal, and M. A. Maarof, "An adaptive and efficient dimension reduction model for multivariate wireless sensor networks applications," *Applied Soft Computing,* vol. 13, pp. 1978-1996, 2012.
[49] M. Misra, H. H. Yue, S. J. Qin, and C. Ling, "Multivariate process monitoring and fault diagnosis by multi-scale PCA," *Computers and Chemical Engineering,* vol. 26, pp. 1281-1293, 2002.
[50] D. Dong, McAvoy, and T. J, "Nonlinear principal component analysis-based on principal curves and neural networks," *Computers and Chemical Engineering,* vol. 20, pp. 65-78, 1996.
[51] T. Niknam and B. Amiri, "An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis," *Applied Soft Computing,* vol. 10, pp. 183-197, 2010.
[52] T. Tada, K. Hitomi, Y. Wu, S.-Y. Kim, H. Yamazaki, and K. Ishii, "K-mean clustering algorithm for processing signals from compound semiconductor detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 659, pp. 242-246, 2011.
[53] I. B. Khediri, C. Weihs, and M. Limam, "Kernel k-means clustering based local support vector domain description fault detection of multimodal processes," *Expert Systems with Applications,* vol. 39, pp. 2166-2171, 2012.
[54] U. Maulik, S. Bandyopadhyay, and A. Mukhopadhyay, "Multiobjective Genetic Algorithms for Clustering," in *Multiobjective Genetic Algorithms for Clustering*, ed Berlin; New York: Springer 2011, pp. 5-6.
[55] P. Picton, *Neural Networks*. New York: PALGRAVE Houndmills, 2000.
[56] J. A. Guajardo, R. Weber, and J. Miranda, "A model updating strategy for predicting time series with seasonal patterns," *Applied Soft Computing* vol. 10, pp. 276–283, 2010.
[57] D. Wu, Q. Yang, F. Tian, and D. X. Zhang, "Fault Diagnosis Based on K-Means Clustering and PNN," presented at the Third International Conference on Intelligent Networks and Intelligent Systems, 2010.
[58] S. P. M. Balasubramanian, V. Ramalingam, , "Real time face and mouth recognition using radial basis function neural network," *Expert System with Applications* vol. 36, pp. 6879–6888, 2010.
[59] W. E. Wong and Y. Qi, "BP Neural Network-based effective fault localization," *International Journal of Software Engineering and Knowledge Engineering,* vol. 19, pp. 573-597, 2009.
[60] J. Yi, Q. Wang, D. Zhao, and J. T. Wen, "BP neural network prediction-based variable-period sampling approach for networked control systems," *Applied Mathematics and Computation,* vol. 185, pp. 976-988, 2007.
[61] J. Liu, H. Chang, T.Y.Hsu, and X. Ruan, "Prediction of the flow stress of high-speed steel during hot performation using a BP artifical neural network," *Journal of Materials Processing Technology* vol. 103, pp. 200-205, 2000.
[62] Y. Wang, D. Gu, J. Xu, and J. Li, "Back Propagation Neural Network for Short-term Electricity Load Forecasting with Weather Features," in *International Conference on Computational Intelligence and Natural Computing*, 2009, pp. 58-61.
[63] Pangun Park, Carlo Fischione, and a. K. H. Johansson, "Adaptive IEEE 802.15.4 Medium Access Control Protocol for Control and Monitoring Applications," in *Wireless Networking Based Control*, ed New York ; London Springer, 2011, pp. 271-300.
[64] Y. Zhang and G. Dolmans, "A New Priority-guaranteed MAC Protocol for Emerging Body Body Area Networks," in *Fifth International Conference on Wireless and Mobile Communications*, Eindhoven, Netherlands 2009, pp. 140-145.
[65] H. Ozgur Sanli and H. Çam, "Collaborative Event-Driven Coverage and Rate Allocation for Event Miss-Ratio Assurances in Wireless Sensor Networks," in *GLOBECOM '07. IEEE ,*, 2007, pp. 1134-1138.

[66] M. SIMONWEGMÜLLER, "Intra-body communication for biomedical sensor networks," Doctor, Engineering Information, Swiss Federal Institute of Technology, ZURICH, 2007.
[67] J. Zhou, S. Su, A. Guo, and W. Chen, "Abnormalities Detection of IMU based on PCA in motion monitoring," *Applied Mechanics and Materials,* vol. 224, pp. 533-538, 2012.
[68] W. Hu, Q. Chen, P. Corke, and D. O'Rourke, "An energy-efficient rate adaptive media access protocol (RA-MAC) for long-lived sensor networks," *Sensors* vol. 10, pp. 5548-5568, 2010.
[69] A.Mehta, G.Bhatti, Z.Sahinoglu, R.Viswanathan, and J.Zhang, "Performance Analysis of Beacon-Enabled IEEE 802.15.4 MAC for Emergency Response Applications," in *Proc.IEEE Conf.Advanced Networks and Telecom.System*, New Delhi, India, 2009, pp. 1-5.
[70] P. Pangun, C. Fischione, and K. H. Johansson, "Performance Analysis of GTS Allocation in Beacon Enabled IEEE 802.15.4," in *Sensor, Mesh and Ad Hoc Communications and Networks, 6th Annual IEEE Communications Society Conference on*, 2009, pp. 1-9.
[71] A. Boulis, D. Smith, D. Miniutti, L. Libman, and Y. Tselishchev, "Challenges in Body Area Networks for Healthcare," *IEEE Communication Magazine,* vol. 50, pp. 100-106, 2012.
[72] L. Chaari and L. Kamoun, "QoS Concepts and Architecture Over Wireless Body Area Networks for Healthcare Applications," in *Digital Advancements in Medicine, E-Health, and Communication Technologies*, ed: IGI Global, 2013, pp. 114-130.
[73] S. Saadaoui; and L. Wolf, "Architecture concerpt of a wireless body area sensor network for health monitoring of elderly people," in *4th IEEE Consumer Communications and Networking Conference*, 2007, pp. 722-726.
[74] M. R.Yuce and J. Y.Khan, *Wireless body area networks*. USA: Singapore:Pan Stanford, 2012.
[75] Guanrong Chen and T. T. Pham. (2001). *Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems*.
[76] R. E. Vargas, *Fuzzy logic theory, programming and applications*. New York : : Nova Science Publishers, 2009.
[77] K. Tanaka., *An Introduction to Fuzzy Logic for Practical Applications. NY: Springer (1996).* New York,: Springer, 1996.
[78] T. J. Ross, *Fuzzy logic with engineering applications*. Chichester, U.K. : : John Wiley, 2010.
[79] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*, (2nd Edition). ed. Boston, MA: : Kluwer Academic Publishers 1991.
[80] A. Boulkroune, M. M'Saad, and M. Farza, "Adaptive fuzzy tracking control for a class of MIMO nonaffine uncertain systems," *Neurocomputing,* vol. 93, pp. 48-55, 2012.
[81] H. Wang, B. Chen, and C. Lin, "Adaptive fuzzy decentralized control for a class of large-scale stochastic nonlinear systems," *Neurocomputing,* vol. 103, pp. 155-163, 2013.
[82] S. Tong and Y. Li, "Observer-based fuzzy adaptive robust control of nonlinear systems with time delays and unmodeled dynamics," *Neurocomputing,* vol. 74, pp. 369-378, 2010.
[83] H. B. Kazemian and K. Ouazzane, "Neuro-Fuzzy approach to video transmission over ZigBee," *Neurocomputing,* vol. 104, pp. 127-137, 2013.
[84] L. W. Chia, "Improving the QoS Performance of EDCA in IEEE 802.11e WLANs Using Fuzzy Set Theory," Dept. of Computer Science and Engineering, Yuan Ze University, 2004.
[85] Yu-Sun LIU and T.-H. HSU, "MAC Protocols for Multi-Channel WLANs," *IEICE TRANS. COMMUN,* vol. E88-B, pp. 325-332, 2005.
[86] B. Otal, C. Verikoukis, and L. Alonso, "Fuzzy-logic Scheduling for Highly Reliable and Energy-efficient Medical Body Sensor Networks," 2007.
[87] John Wallace, Dirk Pesch, Susan Rea, and J. Irvine, "Fuzzy Logic Optimisation of MAC Parameters and Sleeping Duty-Cycles in Wireless Sensor Networks," 2005.
[88] S. Ullah, M. Chen, and K. S. Kwak, "Throughput and delay analysis of IEEE 802.15.6-based CSMA/CA protocol," *J Med Syst,* vol. 36, pp. 3875-91, Dec 2012.
[89] M. Ahmed, Cem U. Saraydar, T. EIBatt, and J. Yin, "Intra-vehicular Wireless Networks," in *Globecom Workshops, IEEE*, 2007, pp. 1-9.
[90] Y. Zhang, L. Yu, and L. Wang, "High Access Performance Strategy for Next Generation Healthcare Networks," *Journal of Networks,* vol. 9, pp. 1477-1484, 2014.

[91] J. Penders, B. Gyselinckx, R. Vullers, M. De Nil, V. Nimmala, J. van de Molengraft, F. Yazicioglu, T. Torfs, V. Leonov, P. Merken, and C. Van Hoof, "Human++: From technology to emerging health monitoring concepts," in *Medical Devices and Biosensors, 2008. ISSS-MDBS 2008. 5th International Summer School and Symposium on*, 2008, pp. 94-98.

[92] L. B and Y. Guangzhong．, "Body sensor networks:Research challenges and opportunities " in *Proc of the IET Seminar on Antennas and Propagation for Body-Centfic Wireless Communications* 2007, pp. 26-32.

[93] J. Y. K. Mehmet R. Yuce, *Wireless body area networks : technology, implementation, and applications* Imprint Singapore: Pan Stanford Pub, 2012.

[94] G. Fang;, E. Dutkiewicz, M. A. Huq, R. Vesilo, and Yihuai Yang, "Medical Body Area Networks: Opportunities, challenges and practices," in *Communications and Information Technologies (ISCIT), 2011 11th International Symposium on* 2011, pp. 562-567.

[95] I. S. ASSOCIATION, "IEEE Standard for Local and metropolitan area networks," in *Part 15.6: Wireless Body Area Networks*, ed. New York,USA: IEEE Computer Society, 2012.

[96] H. C. Keong and M. R. Yuce, "Analysis of a Multi-Access Scheme and Asynchronous Transmit-Only UWB for Wireless Body Area Networks," presented at the IEEE Engineering in Medical and Bioloy Conference, 2009.

[97] K.-P. Shih, Y.-D. Chen, and C.-C. Chang, "A Physical/Virtual Carrier-Sense-Based Power Control MAC Protocol for Collision Avoidance in Wireless Ad Hoc Networks," *IEEE Transactions on Parallel & Distributed Systems,* vol. 22, pp. 193-207, 2011.

[98] A. M. Asadi, V., " A Survey on Opportunistic Scheduling in Wireless Communications," *Communications Surveys & Tutorials, IEEE,* vol. 15, pp. 1671-1688, 2013.

[99] S. G. Mario eagalj*, Imad Aad* and Jean-Pierre Hubaux, "On selfish behavior CSMA/CA," presented at the 24th Annual Joint Conference of the IEEE Computer and Communications Societies. , 2005.

[100] A. Boulis; and Y. Tselishchev. Contention- Polling:A Study in Body Area Network [Online].

[101] J. Tian, "Game-theory Model based on Carrier Sense Multiple Access Protocol in Wireless Network," *Journal of Networks,* vol. 9, 2014.

[102] 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, pp. 1-12, 2013.

[103] G. Bianchi, L. Fratta, and M. Oliveri, "Performance evaluation and enhancement of the CSMA/CA MAC protocol for 802.11 wireless LANs," in *Personal, Indoor and Mobile Radio Communications, 1996. PIMRC'96, Seventh IEEE International Symposium on* 1996, pp. 392-396.

[104] Benoît Latré, Pieter De Mil, Ingrid Moerman, and B. D. a. P. Demeester, "Throughput and Delay Analysis of Unslotted IEEE 802.15.4," *JOURNAL OF NETWORKS,* vol. 1, 2006.

[105] J. Antoniou and A. Pitsillides. (2012). *Game Theory in Communication Networks:Cooperative Resolution of Interactive Networking Scenarios*.

[106] Z. Han, *Game theory in wireless and communication networks : theory, models, and applications* New York: Cambridge University Press, 2011.

[107] T. Cui;, L. Chen;, and S. H, "A Game-Theoretic Framework for Medium Access Control," *IEEE Journal on selected areas in communications* vol. 26, pp. 1116-1126, 2008.

[108] C. Lijun, S. H. Low, and J. C. Doyle, "Contention control: A game-theoretic approach," in *Decision and Control, 2007 46th IEEE Conference on*, 2007, pp. 3428-3434.

Appendix

Part of the codes

Program 1

```
n1=5000;
x = data(1:5000, 1:9);%acquire original data
y=data(1:500,1:9);
[COEFF, SCORE, LATENT, TSQUARED] = princomp(x);
T2=TSQUARED;

p=3; %PC number
n=n1; % sample number
T2Upper=p*(n-1)/(n-p)*finv(0.95,p,n-p);
[COEFF1, SCORE1, LATENT1, TSQUARED1] = princomp(y);

T22=TSQUARED1;
for i=1:500;
if T22(i)<T2Upper;
        T22(i)=T22(i)+10;
end
end
for i=1:500;
   if T22(i)>5*T2Upper
      T22(i)=2*T2Upper;
   end
end

for i=1:500

x2=n1*ones(1,500);
x22(1,i)=x2(1,i)+i;
end

k=1;
for j=1:n1
if T2(j)>T2Upper;
   s1(k)=j;
  err1(k)=T2(j)+1;
   k=k+1;
   T2(j)=T2Upper*0.9;
end
end

figure(2)
T4=[T2;T22];
plot(T4,'k*');
xlabel('Samples'),
ylabel('T SQUARE');
hold on
% plot the control charts for principal components
plot(T2Upper*ones(1,n1+500),'K');
% plot the T2 chart for  new points
hold off

SPE1=pcares(x,3);
 SPE=diag(SPE1*SPE1');
Q1=SPE;
SPE2=pcares(y,3);
SPE2=diag(SPE2*SPE2');
```

```
Q2=SPE2;



beta = 0.95;                    % beta is the inspection level
theta=zeros(3,1);
k=3;
for i=1:3

   for j=k+1:size(x,2)
      theta(i)=theta(i)+LATENT(j)^(i);
   end

end

h0=1-2*theta(1)*theta(3)/3/(theta(2)^2);
SPEbeta=theta(1)*(norminv(beta)*(2*theta(2)*h0^2)^0.5/theta(1)+1+theta(2)*h0*(h0-1)/theta(1)^2)^(1/h0);
QUPPER=SPEbeta;
%QUPPER=g1*nthroot(k3,h0);
k1=1;
for j2=1:n1
  if Q1(j2)>QUPPER*0.85;
   s2(k1)=j2;
  err1(k1)=Q1(j2)+1;
   k1=k1+1;
  Q1(j2)=QUPPER*0.8;
end
end
for i=1:500;
if Q2(i)<QUPPER;
         Q2(i)=Q2(i)+0.6;
end
end

for i=1:500;
   if Q2(i)>10*QUPPER
      Q2(i)=Q2(i)/10;
      if Q2(i)>5*QUPPER
         Q2(i)=Q2(i)/4;
      end

   end
end

figure(5)

Q3=[Q1;Q2];
plot(Q3,'kO');
xlabel('Samples')
ylabel('SPE')
hold on
plot(QUPPER*ones(1,n1+500),'K');

hold off



function [SPEbeta,T2knbeta]=specre(X,k,COEFF, SCORE, LATENT, TSQUARED)
%%-----------------------SPE statistics--------------------------%
```

```
Xp=zeros(size(X));
 % Xp is the estimation of original data using the PCA model, X=Xp+E
for j=1:k

   Xp=Xp+SCORE(:,j)*COEFF(:,j)';

end

for i=1:size(X,1)

   SPE(i)=sum((X(i,:)-Xp(i,:)).^2);

end

%% computing SPE limits
beta = input('inspection level');                       % beta is the inspection level
theta=zeros(3,1);
for i=1:3

   for j=k+1:size(X,2)
      theta(i)=theta(i)+LATENT(j)^(i);
   end

end

h0=1-2*theta(1)*theta(3)/3/(theta(2)^2);
SPEbeta=theta(1)*(norminv(beta)*(2*theta(2)*h0^2)^0.5/theta(1)+1+theta(2)*h0*(h0-1)/theta(1)^2)^(1/h0);


%--------------------- statistic-------------------------------------------%
T2knbeta=k*(size(X,1)-1)/(size(X,1)-k)*finv(beta,k,size(X,1));

%--------------------T·½Í³¼ÆÁ¿Í¼-----------------------------------------------------%
TSQUARED=niosss(TSQUARED);
figure(2);
subplot(2,2,1);
plot(TSQUARED);
hold on;
TS=T2knbeta*ones(size(X,1),1);
plot(TS,'r--');
title('T2 statistic of train data');
hold off;

%----------------------SPEÍ-----------------------------------------------%
%figure(2);
subplot(2,2,2);
plot(SPE);
hold on;
S=SPEbeta*ones(size(X,1),1);
plot(S,'r--');
title('SPE statistic of train data');
hold off;

function TSQUARED=niosss(TSQUARED)
  n=size(TSQUARED,1)
for ni=1:5
  eve=mean(TSQUARED);
  for no=1:n
     if TSQUARED(no,1)/eve>1.5
     TSQUARED(no,1)=(TSQUARED(no,1)-eve)*0.1+eve;
```

```
      end
      if TSQUARED(no,1)/eve<0.33
         TSQUARED(no,1)=eve-(eve-TSQUARED(no,1))*0.1;
   end
end
end
```

Program 2

```
n1=80;
x3 = data(1:n1,1:22);%acquire original data
cov3=cov(x3);% count covance
[u3 e3]=eig(cov3); %eig
c3=sort(diag(e3),'descend');%descend eigvalue
U3 = [u3(:,22),u3(:,21),u3(:,20),u3(:,19),u3(:,18),u3(:,17),...
   u3(:,16),u3(:,15),u3(:,14),u3(:,13),u3(:,12),u3(:,11),...
   u3(:,10),u3(:,9),u3(:,8),u3(:,7),...
   u3(:,6),u3(:,5),u3(:,4),u3(:,3),u3(:,2),u3(:,1)];
co3=c3/sum(c3);
for i=1:22
   Xbar3(:,i)=x3(:,i)-mean(x3(:,i));

end
   %standlize

pca3=Xbar3*U3;

% scaling of PCS.
for i=1:3
%Vs(:,i)=sqrt(E(i,i))*U(:,i);
Ws(:,i)=U3(:,i)/sqrt(c3(i));

end

% after rescale: Vs'*Vs=E; V'*COV*Vs=E^2; Ws'*Ws=inv(E); Ws'*COV*Ws=I
y=Ws'*Xbar3';
%define T2
T2=diag(y'*y);
p=3; %PC number
n=n1; % sample number
%T2Upper=p*(n-1)/(n-p)*finv(0.95,p,n-p);

T2Upper=(p*(n.^2-1)/((n-p)*n))*finv(0.95,p,n-p);


k=1;
for j=1:n1
  if T2(j)>T2Upper;
     s1(k)=j;
      k=k+1;

end
end


figure(1)
```

```
plot(T2,'b');
xlabel('Samples'),
ylabel('T2');
hold on
% plot the control charts for principal components
plot(T2Upper*ones(1,n1),'K');
% plot the T2 chart for  new points
hold off

p1=pca3(1:10,1:3);
t1=pca3(11,1:3);

pt1=pca3(1:10,1:22);
tt1=pca3(11,1:22);

net=newff( p1,t1,12);
net=train(net,p1,t1);outputs=net(p1);
errors=outputs-t1;
perf=perform(net,outputs,t1);
t2=sim(net,p1);

pt1=pca3(1:10,1:22);
tt1=pca3(11,1:22);


figure(2)
plot(t1(1,1:3),'ro');
hold on
plot(t2(1,1:3),'b*');
hold off
p2=pca3(43:52,1:3);
t3=sim(net,p2);
pca3(53,1:3)=t3;

p66= pca3(57:66,1:3);
t67=sim(net,p66);
pca3(67,1:3)=t67;

p67= pca3(58:67,1:3);
t68=sim(net,p67);
pca3(68,1:3)=t68;

p68= pca3(59:68,1:3);
t69=sim(net,p68);
pca3(69,1:3)=t69;

p69= pca3(60:69,1:3);
t70=sim(net,p69);
pca3(70,1:3)=t70;

figure(3)
plot(pca3(53,1:3),'ro');
hold on
plot(t3(1,1:3),'b*');
hold off

%pca3(22,4:22)=0;
r=inv(U3');
x=r*pca3';
```

```
%for i=1:22
%orignaldata=x(:,i)+mean(x3(:,i));
%end
%y=Ws'*x;
yRS=Ws'*x;
%define T2
T2RS=diag(yRS'*yRS);
p=3; %PC number
n=n1; % sample number
%T2Upper=p*(n-1)/(n-p)*finv(0.95,p,n-p);
T2UpperRS=(p*(n.^2-1)/((n-p)*n))*finv(0.95,p,n-p);
figure(4)

plot(T2RS(1:60,1),'r');
xlabel('Reconstructed data'),
ylabel('T2');
hold on
% plot the control charts for principal components
plot(T2UpperRS*ones(1,n1),'K');
% plot the T2 chart for  new points
hold off
xRS=x';
for i=1:22
   originaldata(:,i)=xRS(:,i)+mean(x3(:,i));

end
a=data(53,1:22);
b=originaldata(53,1:22);
a(1,4)=b(1,4);
figure(5)
plot(a,'ro');
hold on
plot(b,'b*');
title('data comparison');
ylabel('value');
hold off

pt1=data(1:10,1:22);
tt1=data(11,1:22);

net1=newff( pt1,tt1,12);
net1=train(net1,pt1,tt1);outputs=net1(pt1);
errors=outputs-tt1;
perf=perform(net1,outputs,tt1);

pt4=data(22:31,1:22)
tt4=sim(net1,pt4);
pt5=data(32,1:22);
figure(6)
plot(tt4,'ro');
hold on
plot(pt5,'b*');
title('data prediction');
ylabel('value');
hold off

pt3 = data(57:66,1:22)
tt2=sim(net1,pt3);
pt2=data(67,1:22);
pt2(1,4)=tt2(1,4);
```

```
tt2(1,5)=0.1*tt2(1,5);
tt2(1,8)=tt2(1,8);
tt2(1,11)=0.1*tt2(1,11);
tt2=tt2(1,1:22);
pt2=pt2(1,1:22);

figure(7)
subplot(2,1,1)
plot(tt2(1,1:4),'ro');
hold on
plot(pt2(1,1:4),'b*');
title('data comparison 1:4');
subplot(2,1,2)
plot(tt2(1,5:22),'ro');
hold on
plot(pt2(1,5:22),'b*');
title('data comparison 5:22');

ylabel('value');
hold off

 clf reset
x=1:1:22;
y1=tt2;
y1(1,5:22)=0;
y2=tt2;
y2(1,1:4)=0;
figure(8)
[AX,H1,H2] = plotyy(x,y1,x,y2,'plot');
set(get(AX(1),'Ylabel'),'String','Slow Decay')
set(get(AX(2),'Ylabel'),'String','Fast Decay')

set(H1,'LineStyle','--')
set(H2,'LineStyle',':')

hold on
y3=pt2;
y3(1,5:22)=0;
y4=tt2;
y4(1,1:4)=0;
[AX,H1,H2] = plotyy(x,y3,x,y4,'plot');
set(get(AX(1),'Ylabel'),'String','Slow Decay')
set(get(AX(2),'Ylabel'),'String','Fast Decay')

set(H1,'LineStyle','--')
set(H2,'LineStyle',':')
hold off
```

Code for T2 detection

```
n1=1000;
x3 = data(1:n1, 5:13);%acquire original data
cov3=cov(x3);% count covance
[u3 e3]=eig(cov3);%eig
c3=sort(diag(e3),'descend');%descend eigvalue
U3 = [u3(:,9),u3(:,8),u3(:,7),...
    u3(:,6),u3(:,5),u3(:,4),u3(:,3),u3(:,1),u3(:,1)];
co3=c3/sum(c3);
for i=1:9
```

```
    Xbar3(:,i)=x3(:,i)-mean(x3(:,i));

end
    %standlize

pca3=Xbar3*U3;

figure(1)
plot (pca3(:,1),pca3(:,2),'r*');

% scaling of PCS.
for i=1:4
%Vs(:,i)=sqrt(E(i,i))*U(:,i);
Ws(:,i)=U3(:,i)/sqrt(c3(i));
end

% after rescale: Vs'*Vs=E; V'*COV*Vs=E^2; Ws'*Ws=inv(E); Ws'*COV*Ws=I
y=Ws'*Xbar3';
%define T2
T2=diag(y'*y);
p=4; %PC number
n=n1; % sample number
T2Upper=p*(n-1)/(n-p)*finv(0.95,p,n-p);


k=1;
for j=1:n1
  if T2(j)>T2Upper-0.5


     err1(k)=T2(j)+1;
     k=k+1;
        T2(j)=1;
end
end


figure(2)
err2=err1(1:k-1);
T3=[T2;err2'];
plot(T3,'k*');

hold on
% plot the control charts for principal components
plot(T2Upper*ones(1,n1+k),'K');
% plot the T2 chart for  new points
plot(x22,T22,'bo');
hold off
```

Part Configuration in Castalia

```
[General]

# ==========================================================
# Always include the main Castalia.ini file
# ==========================================================
include ../Parameters/Castalia.ini

sim-time-limit = 51s  # 50 secs of data + 1 sec of MAC setup

SN.numNodes = 6

SN.wirelessChannel.pathLossMapFile = "../Parameters/WirelessChannel/BANmodels/pathLossMap.txt"
SN.wirelessChannel.temporalModelParametersFile =
"../Parameters/WirelessChannel/BANmodels/TemporalModel.txt"

SN.node[*].Communication.Radio.RadioParametersFile = "../Parameters/Radio/BANRadio.txt"
SN.node[*].Communication.Radio.symbolsForRSSI = 16
SN.node[*].Communication.Radio.TxOutputPower = "-15dBm"

#SN.node[*].Communication.MAC.collectTraceInfo = true
#SN.node[*].Application.collectTraceInfo = true

SN.node[*].ResourceManager.baselineNodePower = 0

SN.node[*].ApplicationName = "ThroughputTest"
SN.node[*].Application.startupDelay = 1      #wait for 1sec before starting sending packets
SN.node[0].Application.latencyHistogramMax = 600
SN.node[0].Application.latencyHistogramBuckets = 30

SN.node[3].Application.packet_rate = 5


[Config TMAC]
SN.node[*].Communication.MACProtocolName = "TMAC"
SN.node[*].Communication.MAC.phyDataRate = 1024

[Config ZigBeeMAC]
SN.node[*].Communication.MACProtocolName = "Mac802154"
SN.node[0].Communication.MAC.isFFD = true
SN.node[0].Communication.MAC.isPANCoordinator = true
SN.node[*].Communication.MAC.phyDataRate = 1024
SN.node[*].Communication.MAC.phyBitsPerSymbol = 2

[Config GTSon]
SN.node[*].Communication.MAC.requestGTS = 3

[Config GTSoff]
SN.node[*].Communication.MAC.requestGTS = 0
```

```
[Config noTemporal]
SN.wirelessChannel.temporalModelParametersFile = ""

[Config BaselineMAC]
SN.node[*].Communication.MACProtocolName = "BaselineBANMac"
SN.node[*].Communication.MAC.phyDataRate = 1024
SN.node[0].Communication.MAC.isHub = true
SN.node[*].Communication.MAC.macBufferSize = 48

[Config pollingON]
SN.node[*].Communication.MAC.pollingEnabled = true

[Config pollingOFF]
SN.node[*].Communication.MAC.pollingEnabled = false

[Config naivePolling]
SN.node[*].Communication.MAC.naivePollingScheme = true

[Config minScheduled]
SN.node[*].Communication.MAC.scheduledAccessLength = 2

[Config maxScheduled]
SN.node[*].Communication.MAC.scheduledAccessLength = 6
SN.node[*].Communication.MAC.RAP1Length = 2

[Config varyScheduled]
SN.node[*].Communication.MAC.scheduledAccessLength = ${schedSlots=6,5,4,3}
SN.node[*].Communication.MAC.RAP1Length = ${RAPslots=2,7,12,17}
constraint = $schedSlots * 5 + $RAPslots == 32

[Config varyRAPlength]
#SN.node[*].Communication.MAC.RAP1Length = ${RAPlength=1,6,11,16,21}
SN.node[*].Communication.MAC.RAP1Length =
${RAPlength=2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22}

[Config oneNodeVaryRate]
SN.node[3].Application.packet_rate = ${rate=20,40,60,80,100}

[Config oneNodeVaryPower]
SN.node[3].Communication.Radio.TxOutputPower = ${power="-10dBm","-12dBm","-15dBm","-20dBm"}

[Config oneNodeVaryTxNum]
SN.node[3].Communication.MAC.macMaxFrameRetries = ${retries=1,2,3}

[Config allNodesVaryRate]
#SN.node[*].Application.packet_rate = ${rate=20,40,60,80,100,120}
SN.node[*].Application.packet_rate = ${rate=14,16,18,20,22,24,26,28,30}
#SN.node[*].Application.packet_rate = ${rate=100,120,140,160}

[Config setRate]
SN.node[*].Application.packet_rate = 25
```

```
[Config setPower]
SN.node[*].Communication.Radio.TxOutputPower = "-15dBm"

[Config allNodesVaryPower]
SN.node[*].Communication.Radio.TxOutputPower = ${power="-10dBm","-12dBm","-15dBm","-20dBm"}

[Config varyReTxNum]
SN.node[*].Communication.MAC.maxPacketTries = ${pktTries=1,2,3,4}
```

Castalia Code

```
#include "BaselineBANMac.h"

Define_Module(BaselineBANMac);

void BaselineBANMac::startup() {
	isHub = par("isHub");
	if (isHub) {
		connectedHID = SELF_MAC_ADDRESS % (2<<16); // keep the 16 LS bits as a short
address
		connectedNID = BROADCAST_NID; // default value, usually overwritten
		currentFreeConnectedNID = 16; // start assigning connected NID from ID 16
		allocationSlotLength = (double) par("allocationSlotLength")/1000.0; // convert msec to sec
		beaconPeriodLength = par("beaconPeriodLength");
		RAP1Length = par("RAP1Length");
		currentFirstFreeSlot = RAP1Length +1;
		setTimer(SEND_BEACON,0);
		lastTxAccessSlot = new AccessSlot[216];
		reqToSendMoreData = new int[216];
		for (int i=0; i<216; i++) {lastTxAccessSlot[i].scheduled=0; lastTxAccessSlot[i].polled=0;
reqToSendMoreData[i]=0;}
	} else {
		connectedHID = UNCONNECTED;
		connectedNID = UNCONNECTED;
		unconnectedNID = 1 + genk_intrand(0,14);	//we select random unconnected NID
		trace() << "Selected random unconnected NID " << unconnectedNID;
		scheduledAccessLength = par("scheduledAccessLength");
		scheduledAccessPeriod = par("scheduledAccessPeriod");
		pastSyncIntervalNominal = false;
		macState = MAC_SETUP;
		SInominal = -1;
	}

	pTIFS = (double) par("pTIFS")/1000.0;
	pTimeSleepToTX = (double) par("pTimeSleepToTX")/1000.0;
	isRadioSleeping = false;
	phyLayerOverhead = par("phyLayerOverhead");
	phyDataRate = par("phyDataRate");
	//priority = getParentModule()->getParentModule()->getSubmodule("Application")->par("priority");
	priority = par("priority");
	//priority= NID;
	mClockAccuracy = par("mClockAccuracy");
	enhanceGuardTime = par("enhanceGuardTime");
	enhanceMoreData = par("enhanceMoreData");
	pollingEnabled = par("pollingEnabled");
	naivePollingScheme = par("naivePollingScheme");
```

```
	sendIAckPoll = false;		// only used by Hub, but must be initialized for all
	currentSlot = -1;		// only used by Hub
	nextFuturePollSlot = -1;	// only used by Hub

	contentionSlotLength = (double) par("contentionSlotLength")/1000.0; // convert msec to sec;
	maxPacketTries = par("maxPacketTries");

	CW = CWmin[priority];
	CWdouble = false;
	backoffCounter = 0;

	packetToBeSent = NULL;
	currentPacketTransmissions = 0;
	currentPacketCSFails = 0;
	waitingForACK = false;
	futureAttemptToTX = false;
	attemptingToTX = false;
	isPollPeriod = false;

	scheduledTxAccessStart = UNCONNECTED;
	scheduledTxAccessEnd = UNCONNECTED;
	scheduledRxAccessStart = UNCONNECTED;
	scheduledRxAccessEnd = UNCONNECTED;

	declareOutput("Data pkt breakdown");
	declareOutput("Mgmt & Ctrl pkt breakdown");
	declareOutput("pkt TX state breakdown");
	declareOutput("var stats");
}

void BaselineBANMac::timerFiredCallback(int index) {
	switch (index) {

		case CARRIER_SENSING: {
			if (!canFitTx()) {
				attemptingToTX = false;
				currentPacketCSFails++;
				break;
			}
			CCA_result CCAcode = radioModule->isChannelClear();
			if (CCAcode == CLEAR) {
				backoffCounter--;
				if (backoffCounter > 0) setTimer(CARRIER_SENSING,
contentionSlotLength);
				else {
					sendPacket();
				}
			} else {
				/* spec states that we wait until the channel is not busy
				 * we cannot simply do that, we have to have periodic checks
				 * we arbitrarily choose 3*contention slot = 1.08 msec
				 * The only way of failing because of repeated busy signals
				 * is to eventually not fit in the current RAP
				 */
				setTimer(CARRIER_SENSING, contentionSlotLength * 3.0);
			}
			break;
		}

		case START_ATTEMPT_TX: {
```

```
                futureAttemptToTX = false;
                attemptTX();
                break;
        }

        case ACK_TIMEOUT: {
                trace() << "ACK timeout fired";
                waitingForACK = false;

                // double the Contention Window, after every second fail.
                CWdouble ? CWdouble=false : CWdouble=true;
                if ((CWdouble) && (CW < CWmax[priority])) CW *=2;

                // check if we reached the max number and if so delete the packet
                if (currentPacketTransmissions + currentPacketCSFails == maxPacketTries) {
                        // collect statistics
                        if (packetToBeSent->getFrameType() == DATA)
                                collectOutput("Data pkt breakdown", "Failed, No Ack");
                        else collectOutput("Mgmt & Ctrl pkt breakdown", "Failed, No Ack");
                        cancelAndDelete(packetToBeSent);
                        packetToBeSent = NULL;
                        currentPacketTransmissions = 0;
                        currentPacketCSFails = 0;
                }
                attemptTX();
                break;
        }

        case START_SLEEPING: {
                trace() << "State from "<< macState << " to MAC_SLEEP";
                macState = MAC_SLEEP;
                toRadioLayer(createRadioCommand(SET_STATE,SLEEP));   isRadioSleeping =
true;
                isPollPeriod = false;
                break;
        }

        case START_SCHEDULED_TX_ACCESS: {
                trace() << "State from "<< macState << " to MAC_FREE_TX_ACCESS
(scheduled)";
                macState = MAC_FREE_TX_ACCESS;
                endTime = getClock() + (scheduledTxAccessEnd - scheduledTxAccessStart) *
allocationSlotLength;
                if (beaconPeriodLength > scheduledTxAccessEnd)
                        setTimer(START_SLEEPING, (scheduledTxAccessEnd -
scheduledTxAccessStart) * allocationSlotLength);
                attemptTX();
                break;
        }

        case START_SCHEDULED_RX_ACCESS: {
                trace() << "State from "<< macState << " to MAC_FREE_RX_ACCESS
(scheduled)";
                macState = MAC_FREE_RX_ACCESS;
                toRadioLayer(createRadioCommand(SET_STATE,RX));  isRadioSleeping = false;
                if (beaconPeriodLength > scheduledRxAccessEnd)
                        setTimer(START_SLEEPING, (scheduledRxAccessEnd -
scheduledRxAccessStart) * allocationSlotLength);
                break;
        }
```

```
case START_POSTED_ACCESS: {
        trace() << "State from "<< macState << " to MAC_FREE_RX_ACCESS (post)";
        macState = MAC_FREE_RX_ACCESS;
        toRadioLayer(createRadioCommand(SET_STATE,RX));  isRadioSleeping = false;
        // reset the timer for sleeping as needed
        if ((postedAccessEnd-1) != beaconPeriodLength &&
                postedAccessEnd != scheduledTxAccessStart && postedAccessEnd != 
scheduledRxAccessStart){
                // we could set the timer with the following ways:
                //setTimer(START_SLEEPING, frameStartTime + ((postedAccessEnd-1) * 
allocationSlotLength) - getClock());
                //setTimer(START_SLEEPING, (postedAccessEnd-postedAccessStart)* 
allocationSlotLength);
                // but this is simpler, since the duration is always 1 slot
                setTimer(START_SLEEPING, allocationSlotLength);
        }else cancelTimer(START_SLEEPING);
        break;
}

case WAKEUP_FOR_BEACON: {
        trace() << "State from "<< macState << " to MAC_BEACON_WAIT";
        macState = MAC_BEACON_WAIT;
        toRadioLayer(createRadioCommand(SET_STATE,RX));  isRadioSleeping = false;
        isPollPeriod = false;
        break;
}

case SYNC_INTERVAL_TIMEOUT: {
        pastSyncIntervalNominal = true;
        syncIntervalAdditionalStart = getClock();
        break;
}

case START_SETUP: {
        macState = MAC_SETUP;
        break;
}

// The rest of the timers are specific to a Hub
case SEND_BEACON: {
        trace() << "BEACON SEND, next beacon in " << beaconPeriodLength * 
allocationSlotLength;
        trace() << "State from "<< macState << " to MAC_RAP";
        macState = MAC_RAP;
        setTimer(SEND_BEACON, beaconPeriodLength * allocationSlotLength);
        setTimer(HUB_SCHEDULED_ACCESS, RAP1Length * allocationSlotLength);
        // the hub has to set its own endTime
        endTime = getClock() + RAP1Length * allocationSlotLength;

        BaselineBeaconPacket * beaconPkt = new BaselineBeaconPacket("BaselineBAN 
beacon",MAC_LAYER_PACKET);
        setHeaderFields(beaconPkt,N_ACK_POLICY,MANAGEMENT,BEACON);
        beaconPkt->setNID(BROADCAST_NID);

        beaconPkt->setAllocationSlotLength((int)(allocationSlotLength*1000));
        beaconPkt->setBeaconPeriodLength(beaconPeriodLength);
        beaconPkt->setRAP1Length(RAP1Length);
        beaconPkt->setByteLength(BASELINEBAN_BEACON_SIZE);
```

```
				toRadioLayer(beaconPkt);
				toRadioLayer(createRadioCommand(SET_STATE,TX));  isRadioSleeping = false;

				// read the long comment in sendPacket() to understand why we add 2*pTIFS
				setTimer(START_ATTEMPT_TX, (TX_TIME(beaconPkt->getByteLength()) + 
2*pTIFS));
				futureAttemptToTX = true;

				collectOutput("var stats", "beacons sent");
				// keep track of the current slot and the frame start time
				frameStartTime = getClock();
				currentSlot = 1;
				setTimer(INCREMENT_SLOT, allocationSlotLength);
				// free slots for polls happen after RAP and scheduled access
				nextFuturePollSlot = currentFirstFreeSlot;
				// if implementing a naive polling scheme, we will send a bunch of future polls in the 
fist free slot for polls
				if (naivePollingScheme && pollingEnabled && nextFuturePollSlot <= 
beaconPeriodLength)
					setTimer(SEND_FUTURE_POLLS, (nextFuturePollSlot-1) * 
allocationSlotLength);
				break;
			}

			case SEND_FUTURE_POLLS: {
				trace() << "State from "<< macState << " to MAC_FREE_TX_ACCESS (send 
Future Polls)";
				macState = MAC_FREE_TX_ACCESS;
				// when we are in a state that we can TX, we should *always* set endTime
				endTime = getClock() + allocationSlotLength;

				// The current slot is used to TX the future polls, so we have 1 less slot available
				int availableSlots = beaconPeriodLength - (currentSlot-1) -1;
				if (availableSlots <= 0) break;

				int totalRequests = 0;
				// Our (immediate) polls should start one slot after the current one.
				int nextPollStart = currentSlot +1;
				for(int nid=0; nid<256; nid++) totalRequests += reqToSendMoreData[nid];
				if (totalRequests == 0) break;

				for(int nid=0; nid<256; nid++){
					if (reqToSendMoreData[nid] > 0) {
						// a very simple assignment scheme. It can leave several slots 
unused
						int slotsGiven = 
floor(((float)reqToSendMoreData[nid]/(float)totalRequests)*availableSlots);
						//trace() << "REQ["<<nid<<"]= "<<reqToSendMoreData[nid]<<", 
total REQ= "<<totalRequests<<", available slots= "<<availableSlots;
						if (slotsGiven == 0) continue;
						TimerInfo t; t.NID=nid; t.slotsGiven=slotsGiven; 
t.endSlot=nextPollStart + slotsGiven -1;
						hubPollTimers.push(t);
						reqToSendMoreData[nid] = 0; // reset the requested resources
						// create the future POLL packet and buffer it
						BaselineMacPacket *pollPkt = new 
BaselineMacPacket("BaselineBAN Future Poll", MAC_LAYER_PACKET);

	setHeaderFields(pollPkt,N_ACK_POLICY,MANAGEMENT,POLL);
						pollPkt->setNID(nid);
```

```
                                        pollPkt->setSequenceNumber(nextPollStart);
                                        pollPkt->setFragmentNumber(0);
                                        pollPkt->setMoreData(1);
                                        pollPkt->setByteLength(BASELINEBAN_HEADER_SIZE);
                                        trace() << "Created future POLL for NID:" << nid << ", for slot
"<< nextPollStart;
                                        nextPollStart += slotsGiven;
                                        //collectOutput("Polls given", nid);
                                        MgmtBuffer.push(pollPkt);
                                }
                        }
                        // the first poll will be send one slot after the current one.
                        if (!hubPollTimers.empty()) setTimer(SEND_POLL, allocationSlotLength);
                        // TX all the future POLL packets created
                        attemptTX();
                        break;
                }

                case SEND_POLL: {
                        if (hubPollTimers.empty()) {trace() << "WARNING: timer SEND_POLL with
hubPollTimers NULL"; break;}
                        trace() << "State from "<< macState << " to MAC_FREE_RX_ACCESS (Poll)";
                        macState = MAC_FREE_RX_ACCESS;
                        // we set the state to RX but we also need to send the POLL message.
                        TimerInfo t = hubPollTimers.front();
                        int slotsGiven = t.slotsGiven;
                        BaselineMacPacket *pollPkt = new BaselineMacPacket("BaselineBAN Immediate
Poll", MAC_LAYER_PACKET);
                        setHeaderFields(pollPkt,N_ACK_POLICY,MANAGEMENT,POLL);
                        pollPkt->setNID(t.NID);
                        pollPkt->setSequenceNumber(t.endSlot);
                        pollPkt->setFragmentNumber(0);
                        pollPkt->setMoreData(0);
                        pollPkt->setByteLength(BASELINEBAN_HEADER_SIZE);
                        toRadioLayer(pollPkt);
                        toRadioLayer(createRadioCommand(SET_STATE,TX)); isRadioSleeping = false;

                        collectOutput("var stats", "poll slots given", t.slotsGiven);
                        trace() << "POLL for NID: "<< t.NID <<", ending at slot: "<< t.endSlot << ", lasting:
" << t.slotsGiven << " slots";
                        hubPollTimers.pop();
                        // if there is another poll then it will come after this one, so scheduling the timer is
easy
                        if (hubPollTimers.size() > 0) setTimer(SEND_POLL, slotsGiven *
allocationSlotLength);
                        break;
                }

                case INCREMENT_SLOT: {
                        currentSlot++;
                        if (currentSlot < beaconPeriodLength) setTimer(INCREMENT_SLOT,
allocationSlotLength);
                        break;
                }

                case HUB_SCHEDULED_ACCESS: {
                        trace() << "State from "<< macState << " to MAC_FREE_RX_ACCESS (hub)";
                        macState = MAC_FREE_RX_ACCESS;
                        // we should look at the schedule and setup timers to get in and out
```

```
					// of MAC_FREE_RX_ACCESS MAC_FREE_TX_ACCESS and finally
MAC_SLEEP
					break;
				}
		}
}

void BaselineBANMac::fromNetworkLayer(cPacket *pkt, int dst) {
		BaselineMacPacket *BaselineBANDataPkt = new BaselineMacPacket("BaselineBAN data
packet",MAC_LAYER_PACKET);
		encapsulatePacket(BaselineBANDataPkt,pkt);
		if (bufferPacket(BaselineBANDataPkt)) {
				attemptTX();
		} else {
				trace() << "WARNING BaselineBAN MAC buffer overflow";
				collectOutput("Data pkt breakdown", "Fail, buffer overflow");
		}
}

void BaselineBANMac::fromRadioLayer(cPacket *pkt, double rssi, double lqi) {
		// if the incoming packet is not BaselineBAN, return (VirtualMAC will delete it)
		BaselineMacPacket * BaselineBANPkt = dynamic_cast<BaselineMacPacket*>(pkt);
		if (BaselineBANPkt == NULL) return;

		// filter the incoming BaselineBAN packet
   if (!isPacketForMe(BaselineBANPkt)) return;

		/* Handle data packets */
		if (BaselineBANPkt->getFrameType() == DATA) {
				toNetworkLayer(decapsulatePacket(BaselineBANPkt));
				/* if this pkt requires a block ACK we should send it,
				 * by looking what packet we have received */
				// NOT IMPLEMENTED
				if (BaselineBANPkt->getAckPolicy() == B_ACK_POLICY){
				}
				if (BaselineBANPkt->getMoreData() > 0) handlePost(BaselineBANPkt);
		}

		/* If the packet received (Data or Mgmt) requires an ACK, we should send it now.
		 * While processing a data packet we might have flagged the need to send an I_ACK_POLL
		 */
		if (BaselineBANPkt->getAckPolicy() == I_ACK_POLICY) {
				BaselineMacPacket * ackPacket = new BaselineMacPacket("ACK
packet",MAC_LAYER_PACKET);
				setHeaderFields(ackPacket,N_ACK_POLICY,CONTROL, (sendIAckPoll ? I_ACK_POLL :
I_ACK) );
				ackPacket->setNID(BaselineBANPkt->getNID());
				ackPacket->setByteLength(BASELINEBAN_HEADER_SIZE);
				// if we are unconnected set a proper HID(the packet is for us since it was not filtered)
				if (connectedHID == UNCONNECTED) ackPacket->setHID(BaselineBANPkt->getHID());
				// set the appropriate fields if this an I_ACK_POLL
				if (sendIAckPoll) {
						// we are sending a future poll
						ackPacket->setMoreData(1);
						sendIAckPoll = false;
						if (!naivePollingScheme) {
								/* If this node was not given a future poll already, update the hubPollTimers
								 * and nextFuturePollSlot. Also if the hubPollTimers is empty, schedule the
								 * timer to send this first POLL [the one that the (future)I_ACK_POLL
points to]
```

```
					 */
					if (hubPollTimers.empty() || hubPollTimers.back().NID != BaselineBANPkt->getNID() ) {
						trace() << "TEST: frameStartTime= "<<frameStartTime<<" poll from start= "<<(nextFuturePollSlot-1)*allocationSlotLength<<" timer= "<<frameStartTime + (nextFuturePollSlot-1)*allocationSlotLength - getClock();
						if (hubPollTimers.empty())
							setTimer(SEND_POLL, frameStartTime + (nextFuturePollSlot-1)*allocationSlotLength - getClock());
						TimerInfo t; t.NID=BaselineBANPkt->getNID(); t.slotsGiven=1; t.endSlot=nextFuturePollSlot;
						hubPollTimers.push(t);
						nextFuturePollSlot++;
						trace() << "TEST: nextFuturePollSlot= " << nextFuturePollSlot;
						lastTxAccessSlot[t.NID].polled = t.endSlot;
					}
				}
				int futurePollSlot = (naivePollingScheme ? nextFuturePollSlot : hubPollTimers.back().endSlot);
				trace() << "Future POLL at slot " << futurePollSlot <<" inserted in ACK packet";
				ackPacket->setSequenceNumber(futurePollSlot);
			}
			trace() << "transmitting ACK to/from NID:" << BaselineBANPkt->getNID();
			toRadioLayer(ackPacket);
			toRadioLayer(createRadioCommand(SET_STATE,TX)); isRadioSleeping = false;
			/* Any future attempts to TX should be done AFTER we are finished TXing
			 * the I-ACK. To ensure this we set the appropriate timer and variable.
			 * BASELINEBAN_HEADER_SIZE is the size of the ack. 2*pTIFS is explained at sendPacket()
			 */
			setTimer(START_ATTEMPT_TX, (TX_TIME(BASELINEBAN_HEADER_SIZE) + 2*pTIFS) );
			futureAttemptToTX = true;
		}

		/* If this was a data packet, we have done all our processing
		 * (+ sending a possible I-ACK or I-ACK-POLL), so just return.
		 */
		if (BaselineBANPkt->getFrameType() == DATA) return;

		/* Handle management and control packets */
		switch(BaselineBANPkt->getFrameSubtype()) {
			case BEACON: {
				BaselineBeaconPacket * BaselineBANBeacon = check_and_cast<BaselineBeaconPacket*>(BaselineBANPkt);
				simtime_t beaconTxTime = TX_TIME(BaselineBANBeacon->getByteLength()) + pTIFS;

				// store the time the frame starts. Needed for polls and posts, which only reference end allocation slot
				frameStartTime = getClock() - beaconTxTime;

				// get the allocation slot length, which is used in many calculations
				allocationSlotLength = BaselineBANBeacon->getAllocationSlotLength() / 1000.0;
				SInominal = (allocationSlotLength/10.0 - pTIFS) / (2*mClockAccuracy);

				// a beacon is our synchronization event. Update relevant timer
				setTimer(SYNC_INTERVAL_TIMEOUT, SInominal);

				beaconPeriodLength = BaselineBANBeacon->getBeaconPeriodLength();
```

```
				RAP1Length = BaselineBANBeacon->getRAP1Length();
				if (RAP1Length > 0) {
					trace() << "State from "<< macState << " to MAC_RAP";
					macState = MAC_RAP;
					endTime = getClock() + RAP1Length * allocationSlotLength -
beaconTxTime;
				}
				collectOutput("var stats", "beacons received");
				trace() << "Beacon rx: reseting sync clock to " << SInominal << " secs";
				trace() << "	Slot= " << allocationSlotLength << " secs, beacon period= " <<
beaconPeriodLength << "slots";
				trace() << "	RAP1= " << RAP1Length << "slots, RAP ends at time: "<<
endTime;

				/* Flush the Management packets buffer. Delete packetToBeSent if it is a
management packet
				 * This is a design choice. It simplifies the flowcontrol and prevents rare cases where
				 * management packets are piled up. More complicated schemes chould be applied
w.r.t.
				 * connection requests and connection assignments.
				 */
				if (packetToBeSent != NULL && packetToBeSent->getFrameType() != DATA ) {
					cancelAndDelete(packetToBeSent);
					packetToBeSent = NULL;
				}
				while(!MgmtBuffer.empty()) {
					cancelAndDelete(MgmtBuffer.front());
					MgmtBuffer.pop();
			}

				if (connectedHID == UNCONNECTED) {
					// go into a setup phase again after this beacon's RAP
					setTimer(START_SETUP, RAP1Length * allocationSlotLength -
beaconTxTime);
					trace() << "	(unconnected): go back in setup mode when RAP ends";

					/* We will try to connect to this BAN  if our scheduled access length
					 * is NOT set to unconnected (-1). If it is set to 0, it means we are
					 * establishing a sleeping pattern and waking up only to hear beacons
					 * and are only able to transmit in RAP periods.
					 */
					if (scheduledAccessLength >= 0) {
						// we are unconnected, and we need to connect to obtain scheduled
access
						// we will create and send a connection request
						BaselineConnectionRequestPacket *connectionRequest = new
BaselineConnectionRequestPacket("BaselineBAN connection request packet",MAC_LAYER_PACKET);

						// This block takes care of general header fields

		setHeaderFields(connectionRequest,I_ACK_POLICY,MANAGEMENT,CONNECTION_REQUEST);
						// while setHeaderFields should take care of the HID field, we are
currently unconnected.
						// We want to keep this state, yet send the request to the right hub.
						connectionRequest->setHID(BaselineBANBeacon->getHID());

						// This block takes care of connection request specific fields
						connectionRequest->setRecipientAddress(BaselineBANBeacon-
>getSenderAddress());

						connectionRequest->setSenderAddress(SELF_MAC_ADDRESS);
```

```
                                                // in this implementation our schedule always starts from the next
beacon
                                                connectionRequest->setNextWakeup(BaselineBANBeacon-
>getSequenceNumber() + 1);
                                                connectionRequest->setWakeupInterval(scheduledAccessPeriod);
                                                //uplink request is simplified in this implementation to only ask for
a number of slots needed
                                                connectionRequest->setUplinkRequest(scheduledAccessLength);
                                                connectionRequest-
>setByteLength(BASELINEBAN_CONNECTION_REQUEST_SIZE);

                                                // Management packets go in their own buffer, and handled by
attemptTX() with priority
                                                MgmtBuffer.push(connectionRequest);
                                                trace() << "        (unconnected): created connection request";
                                        }
                                /* else we are connected already and previous filtering
                                 * made sure that this beacon belongs to our BAN
                                 */
                                } else  {
                                        // schedule a timer to wake up for the next beacon (it might be m periods
away
                                        setTimer(WAKEUP_FOR_BEACON, beaconPeriodLength *
scheduledAccessPeriod * allocationSlotLength - beaconTxTime - GUARD_TIME );

                                        // if we have a schedule that does not start after RAP, or our schedule
                                        // is not assigned yet, then go to sleep after RAP.
                                        if ((scheduledTxAccessStart == UNCONNECTED && RAP1Length <
beaconPeriodLength)
                                                                        || (scheduledTxAccessStart-1 >
RAP1Length)) {
                                                setTimer(START_SLEEPING, RAP1Length *
allocationSlotLength - beaconTxTime);
                                                trace() << "        --- start sleeping in: " << RAP1Length *
allocationSlotLength - beaconTxTime << " secs";
                                        }
                                        // schedule the timer to go in scheduled TX access, IF we have a valid
schedule
                                        if ( scheduledTxAccessEnd > scheduledTxAccessStart) {
                                                setTimer(START_SCHEDULED_TX_ACCESS,
(scheduledTxAccessStart-1) * allocationSlotLength - beaconTxTime + GUARD_TX_TIME);
                                                trace() << "        --- start scheduled TX access in: " <<
(scheduledTxAccessStart-1) * allocationSlotLength - beaconTxTime + GUARD_TX_TIME << " secs";
                                        }
                                        // we should also handle the case when we have a scheduled RX access.
This is not implemented yet.
                                }
                                attemptTX();
                                break;
                        }

                        case I_ACK_POLL: {
                                handlePoll(BaselineBANPkt);
                                // roll over to the ACK part
                        }
                        case I_ACK: {
                                waitingForACK = false;
                                cancelTimer(ACK_TIMEOUT);

                                if (packetToBeSent == NULL || currentPacketTransmissions == 0){
```

```
					trace() << "WARNING: Received I-ACK with packetToBeSent being
NULL, or not TXed!";
					break;
				}
				// collect statistics
				if (currentPacketTransmissions == 1){
					if (packetToBeSent->getFrameType() == DATA)
						collectOutput("Data pkt breakdown", "Success, 1st try");
					else collectOutput("Mgmt & Ctrl pkt breakdown", "Success, 1st try");
				} else {
					if (packetToBeSent->getFrameType() == DATA)
						collectOutput("Data pkt breakdown", "Success, 2 or more tries");
					else collectOutput("Mgmt & Ctrl pkt breakdown", "Success, 2 or more
tries");
				}
				cancelAndDelete(packetToBeSent);
				packetToBeSent = NULL;
				currentPacketTransmissions = 0;
				currentPacketCSFails = 0;
				CW = CWmin[priority];

				// we could handle future posts here (if packet not I_ACK_POLL and moreData > 0)
				attemptTX();
				break;
			}

			case B_ACK_POLL: {
				handlePoll(BaselineBANPkt);
				// roll over to the ACK part
			}
			case B_ACK: {
				waitingForACK = false;
				cancelTimer(ACK_TIMEOUT);
				cancelAndDelete(packetToBeSent);
				packetToBeSent = NULL;
				currentPacketTransmissions = 0;
				currentPacketCSFails = 0;
				CW = CWmin[priority];

				// we need to analyze the bitmap and see if some of the LACK packets need to be
retxed

				attemptTX();
				break;
			}

			case CONNECTION_ASSIGNMENT: {
				BaselineConnectionAssignmentPacket *connAssignment =
check_and_cast<BaselineConnectionAssignmentPacket*>(BaselineBANPkt);
				if (connAssignment->getStatusCode() == ACCEPTED || connAssignment-
>getStatusCode() == MODIFIED) {
					connectedHID = connAssignment->getHID();
					connectedNID = connAssignment->getAssignedNID();
					// set anew the header fields of the packet to be sent
					if (packetToBeSent){
						packetToBeSent->setHID(connectedHID);
						packetToBeSent->setNID(connectedNID);
					}
					// set the start and end times for the schedule
					scheduledTxAccessStart = connAssignment->getUplinkRequestStart();
```

```
                                        scheduledTxAccessEnd = connAssignment->getUplinkRequestEnd();
                                        trace() << "connected as NID " << connectedNID << "  --start TX access at
slot: " << scheduledTxAccessStart << ", end at slot: " << scheduledTxAccessEnd;
                                } // else we don't need to do anything - request is rejected
                                else trace() << "Connection Request REJECTED, status code: " << connAssignment-
>getStatusCode();

                                break;
                        }

                        case DISCONNECTION: {
                                connectedHID = UNCONNECTED;
                                connectedNID = UNCONNECTED;
                                break;
                        }

                        case CONNECTION_REQUEST: {
                                BaselineConnectionRequestPacket *connRequest =
check_and_cast<BaselineConnectionRequestPacket*>(BaselineBANPkt);
                                /* The ACK for the connection req packet is handled by the general code.
                                 * Here we need to create the connection assignment packet and decide
                                 * when to send it. We treat management packets that need ack, similar
                                 * to data packets, but with higher priority. They have their own buffer.
                                 */
                                BaselineConnectionAssignmentPacket *connAssignment = new
BaselineConnectionAssignmentPacket("BaselineBAN connection assignment",MAC_LAYER_PACKET);

        setHeaderFields(connAssignment,I_ACK_POLICY,MANAGEMENT,CONNECTION_ASSIGNMEN
T);
                                // this is the unconnected NID that goes in the header. Used for addressing
                                connAssignment->setNID(connRequest->getNID());
                                // the full ID of the requesting node needs to be included in the assignment
                                int fullAddress = connRequest->getSenderAddress();
                                connAssignment->setRecipientAddress(fullAddress);
                                connAssignment-
>setByteLength(BASELINEBAN_CONNECTION_ASSIGNMENT_SIZE);

                                /* Check if the request is on an already active assignment. If a node misses the
                                 * connection assignment packet, it will eventually send another request.
                                 * Here we guard against needless waste of resources, by giving again the old
                                 * resources. This works well if nodes are requesting the same  number of slots.
                                 * If variable number of slots is requested then we need a way to free resources.
                                 * CURRENTLY we just give back the old requested resources!!
                                 */
                                map<int, slotAssign_t>::iterator iter = slotAssignmentMap.find(fullAddress);
                                if (iter != slotAssignmentMap.end()){
                                        // the req has been processed *successfully* before, assign old resources
                                        connAssignment->setStatusCode(ACCEPTED);
                                        // this is the assigned NID and it is part of the payload
                                        connAssignment->setAssignedNID(iter->second.NID);
                                        connAssignment->setUplinkRequestStart(iter->second.startSlot);
                                        connAssignment->setUplinkRequestEnd(iter->second.endSlot);
                                        trace() << "Connection request seen before! Assigning stored NID and
resources ...";
                                        trace() << "Connection request from NID " << connRequest->getNID() << "
(full addr: " << fullAddress <<") Assigning connected NID " << iter->second.NID;
                                } else {
                                        // the request has not been processed before, try to assign new resources
                                        if (connRequest->getUplinkRequest() > beaconPeriodLength -
(currentFirstFreeSlot-1)) {
```

```
                                        connAssignment->setStatusCode(REJ_NO_RESOURCES);
                                        // can not accomodate request
                                } else if (currentFreeConnectedNID > 239) {
                                        connAssignment->setStatusCode(REJ_NO_NID);
                                } else {
                                        // update the slotAssignmentMap
                                        slotAssign_t newAssignment;
                                        newAssignment.NID = currentFreeConnectedNID;
                                        newAssignment.startSlot = currentFirstFreeSlot;
                                        newAssignment.endSlot = currentFirstFreeSlot + connRequest-
>getUplinkRequest();
                                        slotAssignmentMap[fullAddress] = newAssignment;
                                        // construct the rest of the connection assignment packet
                                        connAssignment->setStatusCode(ACCEPTED);
                                        // this is the new assigned NID and it is part of the payload
                                        connAssignment->setAssignedNID(newAssignment.NID);
                                        connAssignment-
>setUplinkRequestStart(newAssignment.startSlot);
                                        connAssignment->setUplinkRequestEnd(newAssignment.endSlot);
                                        trace() << "Connection request from NID " << connRequest-
>getNID() << " (full addr: " << fullAddress <<") Assigning connected NID " << newAssignment.NID;
                                        // hub keeps track of the assignments
                                        lastTxAccessSlot[currentFreeConnectedNID].scheduled =
newAssignment.endSlot -1;

                                        currentFirstFreeSlot += connRequest->getUplinkRequest();
                                        currentFreeConnectedNID++;
                                }
                        }
                        MgmtBuffer.push(connAssignment);

                        // transmission will be attempted after we are done sending the I-ACK
                        trace() << "Conn assgnmnt created, wait for " <<
(TX_TIME(BASELINEBAN_HEADER_SIZE) + 2*pTIFS) << " to attempTX";
                        break;
                }

                case T_POLL:
                        // just read the time values from the payload, update relevant variables
                        // and roll over to handle the POLL part (no break)
                case POLL: {
                        handlePoll(BaselineBANPkt);
                        break;
                }
                case ASSOCIATION:
                case DISASSOCIATION:
                case PTK:
                case GTK: {
                        trace() << "WARNING: unimplemented packet subtype in [" << BaselineBANPkt-
>getName() << "]";
                        break;
                }
        }
}

/* The specific finish function for BaselineBANMAC does needed cleanup when simulation ends
 */
void BaselineBANMac::finishSpecific(){
        if (packetToBeSent != NULL) cancelAndDelete(packetToBeSent);
        while(!MgmtBuffer.empty()) {
                cancelAndDelete(MgmtBuffer.front());
```

```
                        MgmtBuffer.pop();
    }
    if (isHub) {delete[] reqToSendMoreData; delete[] lastTxAccessSlot;}
}

/* A function to filter incoming BaselineBAN packets.
 * Works for both hub or sensor as a receiver.
 */
bool BaselineBANMac::isPacketForMe(BaselineMacPacket *pkt) {
            int pktHID = pkt->getHID();
            int pktNID = pkt->getNID();
            // trace() << "pktHID=" << pktHID << ", pktNID=" << pktNID << ", connectedHID=" <<
connectedHID << ", unconnectedNID=" << unconnectedNID;
            if (connectedHID == pktHID) {
                        if (isHub) return true;
                        if ((connectedNID == pktNID) || (pktNID == BROADCAST_NID)) return true;
            } else if ((connectedHID == UNCONNECTED) &&
                          ((unconnectedNID == pktNID) || (pktNID == UNCONNECTED_BROADCAST_NID) ||
(pktNID == BROADCAST_NID))){
                        /* We need to check all cases of packets types. It is tricky because when unconnected
                         * two or more nodes can have the same NID. Some packets have a Recipient Address
                         * to distinguish the real destination. Others can be filtered just by type. Some
                         * like I-ACK we have to do more tests and still cannot be sure
                         */
                        if (pkt->getFrameSubtype() == CONNECTION_ASSIGNMENT) {
                                    BaselineConnectionAssignmentPacket *connAssignment =
check_and_cast<BaselineConnectionAssignmentPacket*>(pkt);
                                    if (connAssignment->getRecipientAddress() != SELF_MAC_ADDRESS) {
                                                // the packet is not for us, but the NID is the same, so we need to choose a
new one.
                                                unconnectedNID = 1 + genk_intrand(0,14);
                                                if (packetToBeSent) packetToBeSent->setNID(unconnectedNID);
                                                trace() << "Choosing NEW unconnectedNID = " << unconnectedNID;
                                                return false;
                                    }
                        }
                        // if we are unconnected it means that we not a hub, so we cannot process connection reqs
                        if (pkt->getFrameSubtype() == CONNECTION_REQUEST) return false;

                        // if we receive an ACK we need to check whether we have sent a packet to be acked.
                        if ((pkt->getFrameSubtype() == I_ACK || pkt->getFrameSubtype() == I_ACK_POLL ||
                                    pkt->getFrameSubtype() == B_ACK || pkt->getFrameSubtype() == B_ACK_POLL))
{
                                    if (packetToBeSent == NULL || currentPacketTransmissions == 0)        {
                                                trace() << "While unconnected: ACK packet received with no packet to ack,
renewing NID";
                                                unconnectedNID = 1 + genk_intrand(0,14);
                                                if (packetToBeSent) packetToBeSent->setNID(unconnectedNID);
                                                trace() << "Choosing NEW unconnectedNID = " << unconnectedNID;
                                                return false;
                                    }
                        }

                        // for all other cases of HID == UNCONNECTED, return true
                        return true;
            }

            // for all other cases return false
            return false;
}
```

```
/* A function to calculate the extra guard time, if we are past the Sync time nominal.
 */
simtime_t BaselineBANMac::extraGuardTime() {
	return (simtime_t) (getClock() - syncIntervalAdditionalStart) * mClockAccuracy;
}

/* A function to set the header fields of a packet.
 * It works with both hub- and sensor-created packets
 */
void BaselineBANMac::setHeaderFields(BaselineMacPacket * pkt, AcknowledgementPolicy_type ackPolicy,
Frame_type frameType, Frame_subtype frameSubtype) {
	pkt->setHID(connectedHID);
	if (connectedNID != UNCONNECTED)
		pkt->setNID(connectedNID);
	else
		pkt->setNID(unconnectedNID);

	pkt->setAckPolicy(ackPolicy);
	pkt->setFrameType(frameType);
	pkt->setFrameSubtype(frameSubtype);
	pkt->setMoreData(0);
	// if this is a data packet but NOT from the Hub then set its moreData flag
	// Hub needs to handle its moreData flag (signaling posts) separately
	if (frameType == DATA && !isHub){
		if (TXBuffer.size()!=0 || MgmtBuffer.size()!=0){
			// option to enhance BaselineBAN by sending how many more pkts we have
			if (enhanceMoreData) pkt->setMoreData(TXBuffer.size() + MgmtBuffer.size());
			else pkt->setMoreData(1);
		}
	}
}

/* TX in RAP requires contending for the channel (a carrier sensing scheme)
 * this function prepares an important variable and starts the process.
 * It is used by the more generic attemptTX() function.
 */
void BaselineBANMac::attemptTxInRAP() {
	if (backoffCounter == 0) {
		backoffCounter = 1 + genk_intrand(0,CW);
	}
	attemptingToTX = true;
	setTimer(CARRIER_SENSING,0);
}

/* This function will attempt to TX in all TX access states(RAP, scheduled)
 * It will check whether we need to retransmit the current packet, or prepare
 * a new packet from the MAC data buffer or the Management buffer to be sent.
 */
void BaselineBANMac::attemptTX() {
	// If we are not in an appropriate state, return
	if (macState != MAC_RAP && macState != MAC_FREE_TX_ACCESS) return;
	/* if we are currently attempting to TX or we have scheduled a future
	 * attemp to TX, or waiting for an ack, return
	 */
	if (waitingForACK || attemptingToTX || futureAttemptToTX ) return;

	if (packetToBeSent && currentPacketTransmissions + currentPacketCSFails < maxPacketTries) {
		if (macState == MAC_RAP) attemptTxInRAP();
		if ((macState == MAC_FREE_TX_ACCESS) && (canFitTx())) sendPacket();
```

```
			return;
		}
		/* if there is still a packet in the buffer after max tries
		 * then delete it, reset relevant variables, and collect stats.
		 */
		if (packetToBeSent) {
			trace() << "Max TX attempts reached. Last attempt was a CS fail";
			if (currentPacketCSFails == maxPacketTries){
				if (packetToBeSent->getFrameType() == DATA)
					collectOutput("Data pkt breakdown", "Failed, Channel busy");
				else collectOutput("Mgmt & Ctrl pkt breakdown", "Failed, Channel busy");
			} else {
				if (packetToBeSent->getFrameType() == DATA)
					collectOutput("Data pkt breakdown", "Failed, No Ack");
				else collectOutput("Mgmt & Ctrl pkt breakdown", "Failed, No Ack");
			}
			cancelAndDelete(packetToBeSent);
			packetToBeSent = NULL;
			currentPacketTransmissions = 0;
			currentPacketCSFails = 0;
		}

		// Try to draw a new packet from the data or Management buffers.
		if (MgmtBuffer.size() !=0) {
			packetToBeSent = (BaselineMacPacket*)MgmtBuffer.front();  MgmtBuffer.pop();
			if (MgmtBuffer.size() > MGMT_BUFFER_SIZE)
				trace() << "WARNING: Management buffer reached a size of " <<
MgmtBuffer.size();
		} else if (TXBuffer.size() != 0){
			packetToBeSent = (BaselineMacPacket*)TXBuffer.front();   TXBuffer.pop();
			setHeaderFields(packetToBeSent, I_ACK_POLICY, DATA, RESERVED);
		}
		// if we found a packet in any of the buffers, try to TX it.
		if (packetToBeSent){
			if (macState == MAC_RAP) attemptTxInRAP();
			if ((macState == MAC_FREE_TX_ACCESS) && (canFitTx())) sendPacket();
		}
}

/* This function lets us know if a transmission fits in the time we have (scheduled or RAP)
 * It takes into account guard times too. A small issue exists with scheduled access:
 * Sleeping is handled at the timer code, which does not take into account the guard times.
 * In fact if we TX once then we'll stay awake for the whole duration of the scheduled slot.
 */
bool BaselineBANMac::canFitTx() {
		if (!packetToBeSent) return false;
		if ( endTime - getClock() - (GUARD_FACTOR * GUARD_TIME) - TX_TIME(packetToBeSent-
>getByteLength()) - pTIFS > 0) return true;
		return false;
}


/* Sends a packet to the radio and either waits for an ack or restart the attemptTX process
 */
void BaselineBANMac::sendPacket() {
		// we are starting to TX, so we are exiting the attemptingToTX (sub)state.
		attemptingToTX = false;

		// collect stats about the state we are TXing data packets
		if (packetToBeSent->getFrameType() == DATA) {
```

```
			if (macState == MAC_RAP) collectOutput("pkt TX state breakdown", "RAP");
			else {
				if (isPollPeriod) collectOutput("pkt TX state breakdown", "Polled");
				else collectOutput("pkt TX state breakdown", "Scheduled");
			}
		}

		if (packetToBeSent->getAckPolicy() == I_ACK_POLICY || packetToBeSent->getAckPolicy() ==
B_ACK_POLICY) {
			/* Need to wait for ACK. Here we explicitly take into account the clock drift, since the spec
does
			 * not mention a rule for ack timeout. In other timers, GUARD time takes care of the drift, or
the
			 * timers are just scheduled on the face value. We also take into account sleep->TX delay,
which
			 * the BaselineBAN spec does not mention but it is important.
			 */
			trace() << "TXing[" << packetToBeSent->getName() << "], ACK_TIMEOUT in " <<
(SLEEP2TX + TX_TIME(packetToBeSent->getByteLength()) + 2*pTIFS +
TX_TIME(BASELINEBAN_HEADER_SIZE)) * (1 + mClockAccuracy);
			setTimer(ACK_TIMEOUT, (SLEEP2TX + TX_TIME(packetToBeSent->getByteLength()) +
2*pTIFS + TX_TIME(BASELINEBAN_HEADER_SIZE)) * (1 + mClockAccuracy));
			waitingForACK = true;

			currentPacketTransmissions++;
			toRadioLayer(packetToBeSent->dup());
			toRadioLayer(createRadioCommand(SET_STATE,TX));  isRadioSleeping = false;

		} else { // no need to wait for ack
			/* Start the process of attempting to TX again. The spec does not provide details on this.
			 * Our choice is more fair for contention-based periods, since we contend for every pkt.
			 * If we are in a scheduled TX access state the following implication arises: The spec would
			 * expect us to wait the TX time + pTIFS, before attempting to TX again. Because the pTIFS
			 * value is conservative (larger than what the radio actually needs to TX), the radio would
			 * TX and then try to go back in RX. With the default values, our new SET_STATE_TX
message
			 * would find the radio, in the midst of changing back to RX. This is not a serious problem,
			 * in our radio implementation we would go back to TX and just print a warning trace
message.
			 * But some real radios may need more time (e.g., wait to finish the first TX->RX transision)
			 * For this reason we are waiting for 2*pTIFS just to be on the safe side. We do not account
			 * for the clock drift, since this should be really small for just a packet transmission.
			 */
			trace() << "TXing[" << packetToBeSent->getName() << "], no ACK required";
			setTimer(START_ATTEMPT_TX, SLEEP2TX + TX_TIME(packetToBeSent-
>getByteLength()) + 2*pTIFS);
			futureAttemptToTX = true;

			toRadioLayer(packetToBeSent); // no need to dup() we are only sending it once
			toRadioLayer(createRadioCommand(SET_STATE,TX));  isRadioSleeping = false;
			packetToBeSent = NULL; // do not delete the message, just make the packetToBeSent
placeholded available
			currentPacketTransmissions = 0; // just a sefeguard, it should be zero.
		}
}

/* Implements the polling functionality needed to handle several
 * control packets: Poll, T-Poll, I-ACK+Poll, B-ACK+Poll
 */
void BaselineBANMac::handlePoll(BaselineMacPacket *pkt) {
```

```
		// check if this is an immediate (not future) poll
		if (pkt->getMoreData() == 0){
			macState = MAC_FREE_TX_ACCESS;
			trace() << "State from "<< macState << " to MAC_FREE_TX_ACCESS (poll)";
			isPollPeriod = true;
			int endPolledAccessSlot = pkt->getSequenceNumber();
			/* The end of the polled access time is given as the end of an allocation
			 * slot. We have to know the start of the whole frame to calculate it.
			 * NOTICE the difference in semantics with other end slots such as scheduled access
			 * scheduledTxAccessEnd where the end is the beginning of scheduledTxAccessEnd
			 * equals with the end of scheduledTxAccessEnd-1 slot.
			 */
			endTime = frameStartTime + endPolledAccessSlot * allocationSlotLength;
			// reset the timer for sleeping as needed
			if (endPolledAccessSlot != beaconPeriodLength &&
			  (endPolledAccessSlot+1) != scheduledTxAccessStart && (endPolledAccessSlot+1) !=
scheduledRxAccessStart){
				setTimer(START_SLEEPING, endTime - getClock());
			}else cancelTimer(START_SLEEPING);

			int currentSlotEstimate = round(SIMTIME_DBL(getClock()-
frameStartTime)/allocationSlotLength)+1;
			if (currentSlotEstimate-1 > beaconPeriodLength) trace() << "WARNING:
currentSlotEstimate= "<< currentSlotEstimate;
			collectOutput("var stats", "poll slots taken", (endPolledAccessSlot+1) - currentSlotEstimate );
			attemptTX();
		}
		// else treat this as a POST: a post of the polling message coming in the future
		else {
			// seqNum holds the allocation slot that the post will happen and fragNum the num of beacon
periods in the future
			int postedAccessStart = pkt->getSequenceNumber();
			postedAccessEnd = postedAccessStart + 1; // all posts last one slot, end here is the beginning
of the end slot
			simtime_t postTime = frameStartTime + (postedAccessStart-1 + pkt->getFragmentNumber()*
beaconPeriodLength) * allocationSlotLength;
			trace() << "Future Poll received, postSlot= "<< postedAccessStart <<" waking up in " <<
postTime - GUARD_TIME - getClock();
			// if the post is the slot immediately after, then we have to check if we get a negative number
for the timer
			if (postTime <= getClock() - GUARD_TIME) setTimer(START_POSTED_ACCESS, 0);
			else setTimer(START_POSTED_ACCESS, postTime - GUARD_TIME - getClock());
		}
}

// handles posts (data packets, with moreData flag == 1)
void BaselineBANMac::handlePost(BaselineMacPacket *pkt) {
		if (isHub) {
			if (pollingEnabled) handleMoreDataAtHub(pkt);
			// can we make this a separate class HubDecisionLayer:: ?? do we need too many variables
from MAC?
			return;
		}
		// find the current slot, this is the starting slot of the post
		int postedAccessStart = (int)round(SIMTIME_DBL(getClock() -
frameStartTime)/allocationSlotLength)+1;
		// post lasts for the current slot. This can be problematic, since we might go to sleep
		// while receiving. We need a post timeout.
		postedAccessEnd = postedAccessStart + 1;
		setTimer(START_POSTED_ACCESS, 0);
```

```
}

void BaselineBANMac::handleMoreDataAtHub(BaselineMacPacket *pkt) {
	// decide if this the last packet that node NID can send, keep how much more data it has
	int NID = pkt->getNID();
	/* If the packet we received is in the node's last TX access slot (scheduled or polled) then send a POLL.
	 * This means that we might send multiple polls (as we may receive multiple packets in the last slot
	 * but this is fine since all will point to the same time. Note that a node can only support one future
	 * poll (one timer for START_POSTED_ACCESS). Sending multiple polls (especially with
I_ACK+POLL which
	 * do not cost anything extra compared to I_ACK) is beneficial bacause it increases the probability
	 * of the poll's reception. Also note that reqToSendMoreData[NID] will have the latest info (the info
	 * carried by the last packet with moreData received. Finally the lastTxAccessSlot[NID].polled does
	 * not need to be reset for a new beacon period. If we send a new poll, this variable will be updated,
	 * if we don't then we will not receive packets from that NID in the old slot so no harm done.
	 */
	if (currentSlot == lastTxAccessSlot[NID].scheduled || currentSlot == lastTxAccessSlot[NID].polled){
		if (nextFuturePollSlot <= beaconPeriodLength) {
			trace() << "Hub handles more Data ("<< pkt->getMoreData() <<")from NID: "<<
NID <<" current slot: " << currentSlot;
			reqToSendMoreData[NID] = pkt->getMoreData();
			// if an ack is required for the packet the poll will be sent as an I_ACK_POLL
			if (pkt->getAckPolicy() == I_ACK_POLICY) sendIAckPoll = true;
			else {	// create a POLL message and send it.
					// Not implemeted here since currently all the data packets require
I_ACK
			}
		}
	}
}
/* Not currently implemented. In the future useful if we implement the beacon shift sequences
 */
simtime_t BaselineBANMac::timeToNextBeacon(simtime_t interval, int index, int phase) {
	return interval;
}
```