DESIGN OF REAL-TIME MEDIUM ACCESS CONTROL AND THE OPTIMIZATION OF END-TO-END DELAY USING GENETIC ALGORITHM

BY

JUNE SEUNG YOON

A dissertation submitted to the Graduate Faculty in Electrical engineering in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
The City University of New York

2012

This manuscript has been read and accepted for the Graduate
Faculty in Engineering in satisfaction of the dissertation
requirement for the degree of Doctor of philosophy.

	Myung J. Lee
Date	Chair of Examining Committee
	Mumtaz Kassir
Date	Executive Officer

Jizhong Xiao

Sang-Woo Seo

Yi Sun

Truong-Thao Nguyen

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

DESIGN OF REAL-TIME MEDIUM ACCESS CONTROL AND THE OPTIMIZATION OF END-TO-END DELAY USING GENETIC ALGORITHM

By
June Seung Yoon

Advisor: Professor Myung Jong Lee

The Quality of Service (QoS) of wireless ad hoc and sensor networks is identified as one of the most important research issues due to the difficulty of overcoming the unstable and unforeseeable characteristics of wireless medium. The difficulty is amplified for the networks that handle real-time data such as audio/video stream, vital sign monitoring, and factory process control. These are so delay-sensitive that usually become inefficacious if belated. Therefore, the most prevailing measure for the real-time data applications is the end-to-end delay. This dissertation addresses two medium access control protocols that can manage the QoS required for wireless body area networks and wireless industrial sensor networks, and a metaheuristic scheduling algorithm for multi-channel, multi-hop wireless sensor networks. First, the preemptive slot allocation and non-preemptive access medium access control (PNP-MAC) protocol, designed for wireless body area networks is introduced in chapter 1. Wireless body area network applications, especially medical applications are very exacting in delay; thus, PNP-MAC equips diverse QoS features to meet the QoS requirements of the applications. Second, chapter 2 introduces Enhanced Guaranteed Time Slot (EGTS), developed for industrial applications that require capabilities to support large scale multi-hop networks and deterministic delay. EGTS's multi-channel over the Time Division Multiple Access (TDMA) enables multiple simultaneous data transmissions in a single time slot, and the three-way-hand shaking allocation mechanism allows nodes to use non-overlapped slots and channels, leading the data forwarding with minimal end-to-end delay. Third, chapter 3 introduces Genetic Algorithm Assisted Scheduling (GAS), a scheduling algorithm designed based on the genetic algorithm, one of the most popular metaheuristic algorithms. Genetic algorithm is shown to provide a near-optimal solution at a reasonable computation time for our slot and channel allocation problem given an end-to-end delay bound. Newly designed features such as scheduling order of flow, pre-channel assignment, and sifting along with enhanced genetic algorithm operations expedite search process and lead to the optimal solution.

Acknowledgements

I would like to express my sincere appreciation to my advisor Professor Myung Jong Lee. He has provided me with insightful suggestions and motivations that enabled me to keep moving forward. I am deeply thankful for his help and guidance during my doctoral study.

I would like to thank Professor Truong-Thao Nguyen, Professor Yi Sun, Professor Jizhong Xiao, and Professor Sang-Woo Seo for providing useful comments and serving on my dissertation committee. I also want to acknowledge the companionship and support of all my colleagues, Yang, Rui, Taerim, Ghangseop, Yanchen and Junsun in Advanced Wireless Networks Lab.

I would like to give my special thanks to my parents who give me unreserved love and encourage me to explore the new world.

Finally I would like to give my greatest gratitude to my wife Hyunchong who gave me patient love that enabled me to complete this work. I want to dedicate this dissertation to my wife Hyunchong.

Table of Contents

List of Figures

List of Tables

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Over the last few decades wireless sensor network (WSN) technology has been drawing tremendous attentions and research focuses due to its practicalities in wide spectrum of applications such as environmental monitoring, surveillance, collaborative computing, social networking, and vehicular networking. Along with the tremendous attentions, rapid advances in Micro-Electric Mechanical System (MEMS) and VLSI technologies realize miniaturized, low-power, short-range wireless communication devices. Thanks to the tiny form factor, devices equipped with microprocessor, memory, sensor, transceiver, antenna as well as battery can be placed in one tiny wireless sensor device. Emergence of wireless sensor network technology, in addition to currently existing wireless technologies offers considerable flexibility. Wireless sensor network is autonomous system that does not rely on infrastructure but self-organizes itself in ad hoc fashion. Hence, wireless sensor networks allow users to deploy them easily even in large scale networks.

Recently, wireless sensor networks broaden their effectiveness in myriad applications. Among other things, the chief beneficiary of wireless sensor network technology is the applications in monitoring and control areas including HVAC control, inventory tracking, data gathering as well as home and factory automation. Actually, Zigbee alliance [46] has already provided various profiles for those applications together with a network protocol. Consequently, the wireless sensor network becomes a highly attractive solution to alternate current wired networks which are bulky, expensive, complex and difficult to deploy and maintain. Many organizations and standards such as wireless HART, BACnet over Zigbee, ISA100, IEEE802.15 TG4e, and IEEE802.15 TG6 [45][43][44][32][14] have exerted themselves to design appropriate wireless solutions for real-time monitoring, tracking, control, and diagnosis applications. As other wireless technologies did before, wireless sensor networks will bring out remarkable impacts on the paradigm of real-time data treatment.

The most recent emerging areas that welcome wireless sensor network technology most are wireless body area networks (BAN) and wireless industrial sensor networks. The sensors attached surface or inside of the human body, and machines, conveyor belt, or utility pipelines form a tree-based network and report sampled data to a central controller periodically. The most critical performance measure in this kind of application is latency or end-to-end delay. This dissertation introduces two application-specific real-time medium access control protocols: PNP-MAC for body area networks and EGTS for wireless industrial sensor networks, and an end-to-end delay optimization algorithm using genetic algorithm for multi-hop, multi-channel networks.

1.2 Outlines

1.2.1 PNP-MAC

A recent strong demand in wireless sensor network technology stems from the need for untethered communication around human body and it sparks emergence of new wireless technology, Body Area Network (BAN). The need and the basic concept of BAN technology were derived from the medical area in order for appropriate medical treatment at the right time by monitoring and recording the patients' physiological phenomenon and vital signs for example electroencephalogram (EEG), electrocardiogram (ECG), electromyogram (EMG), and pulsation.

A challenging issue in Body Area Networks is to provide guaranteed and diverse levels of Quality of Service (QoS). Since BAN supports a wide range of medical and entertainment applications, it has to support various types of traffics such as low and high rate, periodic and non-periodic, and continuous. BAN technology should serve all these applications and fulfill distinctive requirements of them even when they operate simultaneously. In chapter 2, we introduce preemptive slot allocation and non-preemptive access medium access control (PNP-MAC), a novel medium access protocol for BAN. It includes preemptive slot allocation method and non-preemptive channel access that support priority and guaranteed QoS. It also addresses a flexible superframe structure that enables fast slot allocation and prompt adaptation to changes in time slot allocation. Moreover PNP-MAC provides access schemes not only for narrowband PHY but also Ultra-wide band (UWB) PHY which is one of the promising radio technology candidates for BAN due to its substantially low energy consumption.

As BAN technology has been drawing enormous needs and attentions, IEEE 802.15 task group 6 (TG6) was established to develop a standard for the physical layer (PHY) and the medium access control (MAC) layer protocols for BAN [14]. Needless to say, QoS is the common feature to be achieved in all kinds of wireless networks. However, QoS in BAN is incomparably critical than others because BAN mainly handles health and even life related real-time physiological information, which becomes meaningless if belated.

In light of the distinct characteristics, the first consideration in designing a MAC protocol for

BAN is to provide guaranteed QoS. PNP-MAC's various QoS features are able to cope with the requirements. Hence, we could contribute PNP-MAC to IEEE802.15.6 as a part of the MAC protocol.

1.2.2 EGTS

In chapter 3, we introduce a novel wireless sensor network MAC protocol, the Enhanced Guaranteed Time Slot (EGTS) that is cut out for industrial application by enhancing GTS mechnism of IEEE802.15.4 MAC protocol [31][32] . It provides robust and timely data delivery in industrial wireless sensor network applications. To this end, we enhance and add new functionalities to current IEEE802.15.4 such as multi-superframe extension, muti-channel extension, and multi-hop extension, providing deterministic service for large scale industrial applications. However, its utilization is not confined only for industry but can be applied to the applications requiring high level of QoS such as reliability, jitter, latency and robustness.

The important requirement for industrial application-specific MAC is deterministic QoS such as jitter and latency. Sensor readings from the factory process and equipment should be delivered in a timely manner. Belated information may become meaningless and could make problems worse in controlling process and equipment, and may push users to crisis.

In order to satisfy the requirements, first, EGTS introduces a fortified superframe structure which extends GTS in space, time, and channels. EGTS supports multi-hop mesh connectivity by extending GTS in space. Moreover, EGTS is provided in multi-channel while GTS of IEEE802.15.4 is provided in a single channel. For time extension, we propose Superframe Reiteration (SR), Multi-superframe Order (MO) and CAP Reduction (CR). Superframe reiteration may not go with an energy saving policy but is suitable for the applications under timing constraints, where providing deterministic guarantees is the foremost important requirement to consider than everything else. Secondly, EGTS performs enhanced time slot allocation mechanism in conjunction with channel allocation through the collaborative three-way-handshaking which results in simultaneous transmissions with less interference and deterministic low end-to-end delay. Thirdly, EGTS provides channel adaptation to enhance robustness to cope with dynamic channel conditions. A factor we should not overlook is the adaptivity to and robustness against industrial environment. Indoor is not a friendly environment for wireless communication because of the severe multi-path effect caused by static and mobile obstacles. A thorough investigation of channel stability in a factory [35] shows it clearly.

IEEE802.15.4e standard is the MAC amendment of IEEE802.15.4 for the industrial applications where deterministic end-to-end delay should be supported [42]. Our EGTS protocol is contributed to IEEE802.15.4e MAC in the name of DSME (Deterministic Synchronous Multi-channel Extension) mode.

1.2.3 GAS

DSME (Deterministic Synchronous Multi-channel Extension) mode in the IEEE802.15.4e standard [42] exploits multi-channel over TDMA [34][42]. Thanks to the multi-channel extension, slot utilization in DSME is amplified up to as many as the number of channels thus, end-to-end delay can be reduced substantially. However, although DSME guarantees deterministic end-to-end-delay, it is not the optimal because scheduling in DSME is done by a first-come, first-served fashion. In order to optimize end-to-end-delay we have to investigate all possible node-channel-slot combinations of all the nodes in a network; but, it is not an easy task in practice, especially in large scale multi-hop networks.

The transmission scheduling in multi-hop networks is a typical combinatorial NP-complete problem [51] [52] that has excessively large candidate solutions but no specific given solution is provided. A popular method to attain the optimal solution in this kind of combinatorial problem is metaheuristics [58]. Metaheuristc is a computational method that optimizes a problem by iteratively trying to improve a candidate solution satisfying a given objective function. Comparing to exhaustive search, metaheuristics can offer near-optimal solution very fast. In chapter 4, we propose GAS, Genetic algorithm-assisted transmission scheduling method for multi-hop, multi-channel networks. Genetic Algorithm (GA) is a metaheuristic tool inspired from the evolutionary process of nature and the rule 'The survival of the fittest'. The creatures in nature have chromosomes which in turn consist of strings of genes. The genes in a chromosome determine characteristics of the creature and fitness to the surrounding environment. The ones who have better fitness than others survive; otherwise they become extinct. The survived ones mate with each other and have chances to turn genes over their descendants. As the evolutionary process iterated, the genes of surviving creatures tend to be similar each other since sharing of some dominant genes does not make any improvement later on and eventually they all resemble each other, then, evolutionary process ends; in other words the process reaches the optimal or near-optimal solution.

GA is a proper mataheuristic method that can be applied to a combinatorial problem having multiple of variables such as nodes, slots, and channels in our scheduling problem. In chapter 4, we introduce enhanced GA operations and additional supportive features including flow-based allocation, pre-channel assignment that can promote GA to find a near-optimal node-slot-channel combination that minimizes the maximum end-to-end delay.

Chapter 2

PNP-MAC

2.1 Introduction

Rapid advances in Micro-Electric Mechanical System (MEMS) and VLSI technologies realize miniaturized, low-power, short-range wireless communication devices. Thanks to the tiny form factor, devices equipped with processor, memory, sensor, transceiver, as well as battery can be attached on human body and even be implanted under the skin without discomforting everyday life. Along with the advances, the new needs for untethered communication around human body sparks emergence of new wireless technology, Body Area Network (BAN).

BAN is a network of devices used in, on, or around the human body. Two major application domains driving the development of BAN are medical applications (e.g., monitoring of respiratory rate, glucose level, brain waves) and entertainment applications (e.g., wireless video goggle, interactive game) [13]. The needs and the basic concept of BAN technology were initiated from the medical area in order for appropriate medical treatment at the right time by monitoring and recording the patients' physiological phenomenon and vital signs for example

Electroencephalogram (EEG), Electrocardiogram (ECG), Electromyogram (EMG), pulsation, and more, and then forward the collected data to remote medical staffs or doctors to take steps for diagnosis, prescription or other types of treatments in time.

A challenging issue in Body Area Network is to provide guaranteed and diverse levels of Quality of Service (QoS). Since BAN supports a wide range of medical and entertainment applications, it has to support various types of traffics such as low rand high rate, periodic and non-periodic, and continuous. A BAN should serve all these applications and fulfill distinctive requirements of them even when they operate simultaneously. In this chapter, we introduce preemptive time slot allocation and non-preemptive access medium access control (PNP-MAC), a novel medium access protocol for BAN. It includes preemptive slot allocation method and non-preemptive channel access that support priority and guaranteed QoS. It also addresses a flexible superframe structure that enables fast slot allocation and prompt adaptation to changes in time slot allocation. Moreover PNP-MAC provides access schemes not only for narrowband PHY but also Ultra-wide band (UWB) PHY which is one of the promising radio technology candidates for BAN due to its substantially low energy consumption. While low energy consumption is definitely beneficial for battery powered wireless sensor devices, too low transmission power level of UWB PHY makes wireless devices difficult to perform CSMA (carrier sense multiple access), the most popular wireless carrier detection technique. Therefore MAC should have special schemes in order for exploiting UWB.

In conjunction with currently existing infrastructures such as wireless LAN and cellular networks, BAN technology can bring revolutionary impact on the paradigm of healthcare. While users lead their daily lives unmindful of sensor devices’ operation, the devices seamlessly monitor users’ physiological activities and report them to a coordinator or a server device.

Irrespective of time and location, the data collected at a coordinator are automatically forwarded to a remote medical clinic via wireless LAN or cellular network in real time [20]. If needed, based on the diagnosis from a medical clinic, users can take appropriate feedbacks right away in the forms of prescription, treatment appointment, ambulance service. Instead of burdensome visit to medical clinics and occasional checkups, users can always take advantages of wireless BAN technology, facilitating anytime and anywhere healthcare.

In addition to medical services, a variety of other areas for example, entertainment, sports, fitness and social networking are being envisioned as BAN applications. Among others, fitness is the one that welcomes highly the BAN technology. Health and fitness are everyone’s interest for the whole life and the interest keeps on rising along with the world’s general trend ‘Well-Being’. Like in medical applications, fitness devices survey users’ physiological activities and body status (e.g., exercise time, quality of sleeping, pulsation rate), and transmit them to a data collection server such as PC, cell phone or portable media player. Utilizing the collected data, users can improve their body performance, and manage their health. Fitness devices applying BAN technologies are already found in the market [17][18][19] and innumerable products are expected before long.

As BAN technology has been drawing tremendous needs and attentions, IEEE 802.15 task group 6 (TG6) was established to develop a standard for the physical layer (PHY) and the media access control (MAC) layer protocols for BAN [14]. Needless to say, Quality of Service (QoS) is the common feature to achieve in every wireless network. However, QoS in BAN is incomparably critical than others because BAN mainly handles real-time, health, and even life related information, which becomes meaningless if belated.

In light of the distinct characteristics, the first consideration in designing a MAC protocol for

BAN is to provide guaranteed QoS. The BAN applications have diverse QoS requirements in terms of latency and reliability [11]. Most of the monitoring applications for physiological data are periodic and low rate, while entertainment applications are high rate and streaming in nature. Emergency alarm, the most time-critical data in BAN, is non-periodic and is not easily predictable and it requires instantaneous delivery with the highest reliability. BAN should be able to cope with these various latency and reliability requirements even when all those applications are used simultaneously.

The remaining chapters are instructed as follows: Section 2.2 reviews related work. Multi-hop spanning tree protocol, energy efficient MAC protocols including UWB PHY applied MAC protocols, IEEE802.15.4 MAC protocol and its variants are reviewed. Detail design of PNP-MAC including its structure, operations and QoS features are described in section 2.3. Section 2.4 shows performance evaluation of PNP-MAC in various scenarios using the OPNET network simulator [26]. Section 2.5 addresses MicaZ [27] test-bed setup deployed on the body of a volunteer and measures packet reception rate in motionless and in motion scenarios. It backs up the results of OPNET simulation as well as provides practicality of PNP-MAC. The Section 2.6 presents some concluding remarks.

2.2 Literature Review

WASP (Wireless Autonomous Spanning tree Protocol), a spanning tree protocol for a multi-hop BAN was presented in [21]. It models a multi-hop tree topology BAN network and introduces an algorithm for data transmission scheduling and silence period to minimize end-to-end delay. According to the technical requirements of IEEE 802.15 TG6 [1], the transmission range of the BAN devices is up to 5 meters which is large enough to cover a patient room. Thus, a multi-hop tree may not be a realistic scenario but a single hop star topology is more practical. The MicaZ test-bed experiments based on the on-body scenario in section 2.5 shows it clearly. A coordinator device placed on the waist, about a center of body, captures data from head and ankle, the farthest point of a body at -10dBm transmission power.

Energy efficient protocols for BAN have been proposed in [9][22][23]. Low energy consumption in BAN is as critical as the issue of QoS, especially for implant devices. Although some wireless charging techniques, for instance inductive charging [27] are emerging, battery replacement or recharging of the implant device is inconvenient, costly and often painful. Another concern on the energy efficiency is that there exists a tradeoff between the energy saving and QoS. Thus, one should balance among energy efficiency, latency, and reliability when design protocols for BAN.

Regarding energy efficiency, [8][15] suggest that Ultra-wideband (UWB) is a suitable radio technology for BAN due to its ultra-low power radiation. Besides, its wide bandwidth supports high data rate for real-time multi-media services. For both medical and entertainment applications, UWB is one of the most desirable radio technology but its ultra-low power radiation make CSMA infeasible. With energy efficient UWB, it is time to consider a new MAC protocol design exploiting its advantages.

Much of earlier works, [5][8][10][16] adopt IEEE 802.15.4 MAC protocol [31][32] or its variants as a MAC protocol for BAN. Although IEEE 802.15.4 was originally designed for low-power and low-rate applications in wireless personal area networks (WPAN), it offers some desirable features for BAN. A flexible superframe, supporting CSMA and TDMA is well-fitted to digest periodic and non-periodic applications. Its Guaranteed Time Slot (GTS) in contention-free TDMA access period guarantees deterministic delay through periodic exclusive slot allocation. The GTSs are managed by a PAN coordinator in accordance with the request from its immediate neighbors. Typically the communications in BAN take place between many tiny BAN sensor devices and a more capable data collection device (e.g., handset, PDA) or a base station. The handset or the base station can take a role as a BAN coordinator that can manage time slot allocation. The features of IEEE 802.15.4 GTS, however, are not sufficient to satisfy QoS requirements of BAN. First of all, it allows only up to seven GTSs, which is not enough for the BAN applications. For example, a BAN may consist of tens of Electroencephalogram (EEG) nodes generating continuous waveform data that need to be delivered with small delay, for instance less than 125 ms in medical applications [1]. Secondly, the high latency associated with the GTS allocation may pose a significant problem in BAN. In order to use a GTS, a device should request allocation of a GTS and receive a beacon confirming the successful allocation at the next beacon interval. Thus, at least one beacon interval should pass before using a GTS. Such allocation latency may not be acceptable for time-critical medical applications such as emergency alarms. Thirdly, GTS allocation is first-come, first-served basis. Since it does not consider priority, high priority data can be blocked if all the GTSs are already allocated to lower priority data. In this case, time slots for the high priority data can be allocated only after de-allocation of low priority data. However, the process of de-allocation also takes at least one beacon interval. Therefore, allocation of high priority data after de-allocation of low priority data takes at least two beacon intervals, causing severe delay.

BodyQoS [4] separates QoS scheduler from the underlying MAC implementation, thus, unlike IEEE802.15.4, it does not suffer from the limited number of time slots. However, BodyQoS adopts non-preemptive slot allocation schemes that high priority data can be blocked by low priority data. Moreover, the polling scheme may incur additional overhead, burdening restricted

BAN resources.

Our proposal to TG6 [2] and previous work [24] have addressed the concerns on QoS for BAN. However, it lacks solutions to support UWB radio and practical experiments with real devices in practical scenarios. This time, we suggest a new access schemes for UWB exploiting its low energy and wide bandwidth and made practical experiment using MicaZ test-bed with on-body BAN scenarios.

2.3 PNP-MAC Design

The biggest concern in designing a MAC protocol for BAN is how to provide QoS to the applications with various requirements. In PNP-MAC, priority takes on an important role to resolve this. Use of priority in MAC enables differentiated services to applications with different requirements. This section first defines priorities of various traffics and explains flexible structure of superframe, fast slot allocation and adaptation, prioritized back-off in contention access period, preemptive allocation and non-preemptive access in data transmission TDMA period, random and prioritized CCA in ETS, and backoff-and-reserve and backoff-and-go for UWB radio.

2.3.1 Priority Classification

Priority is determined in accordance with the reliability and the timeliness of data. Generally, emergency alarm is placed on the top of the priority, and then comes medical data which is followed by non-medical entertainment data. It is because health or life related physiological data is obviously far more critical than entertainment data. TG6 technical requirements [1] also

Table 2.1 Priority and back-off class classification

Back-off Class	Priority	Data Types	Examples
0	**0**	Emergency alarm	Abnormal vital sign Battery depletion
1	**1**	Medical continuous	EEG/ECG/EMG
	2	Medical routine	Temperature Blood pressure Oxygen saturation
2	**3**	Non-medical continuous	Real time audio and video
	4	Others	File transfer

set delay bound for medical data (125ms) that is much stricter than entertainment data (250ms).

Classification in terms of timeliness of data is in the order of continuous, periodic, and then followed by non-periodic data. Back-off class is for the differentiation of back-off duration while contention and it consists of three classes, alarm, medical, and non-medical data. Considering all these, we summarize the classification of back-off class and priority in Table 2.1 with some examples.

Priority is exploited in preemptive data transmission slot allocation, prioritized back-off in contention access, random and prioritized CCA (RP-CCA) and non-preemptive access in emergency alarm transmission slot (ETS). Note, however, that in practice the priority classification should accommodate applications requirements.

2.3.2 BAN Superframe

PNP-MAC inherits the best breeds of contention-based and contention-free medium access techniques; thereby, it is able to support various types of traffics: continuous streaming, periodic, non-periodic, and bursty as well. As indicated in Fig.2.1, BAN superframe consists of six

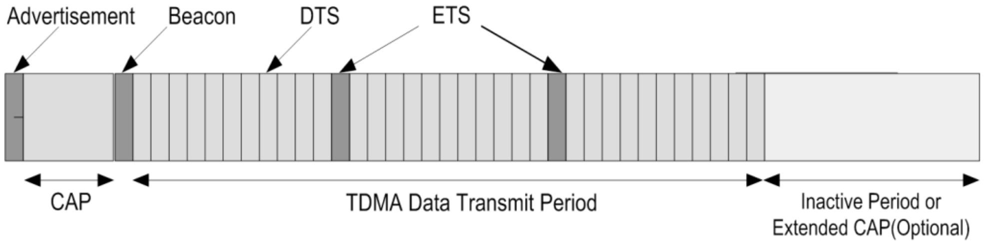


Fig. 2.1 BAN Superframe structure

periods: advertisement, contention access period (CAP), beacon, data transmit slot (DTS), emergency data transmit slot (ETS), and optional inactive or extended CAP (ECAP) period. The BAN superframe admits flexible configuration in accordance with the demands from devices. Initially, BAN coordinator begins BAN superframe with only advertisement and CAP. Advertisement informs newly joining devices and member devices of information regarding the BAN coordinator and superframe such as BAN identifier, superframe interval, and the length of CAP period. Once taking all the BAN related information from the advertisement, member devices can transmit control frames such as network association request, DTS allocation request or transmit non-periodic data during CAP period using prioritized back-off.

DTS period is not initiated until allocation requests from member devices are received. Once received, the BAN coordinator allocates requested number of DTSs for the DTS requestors. After DTS allocation is completed and has nothing to transmit during CAP, a device can remain asleep during advertisement and CAP period for energy saving.

Beacon broadcast right after CAP period confirms allocation and announces changes in allocation status. All the devices that reserved DTSs have to listen to beacon to check its allocation status before data transmission. The devices may wake up once every multiple BAN superframe interval if data transmission interval is longer than a BAN superframe period. The beacon also synchronizes devices in a network.

ETS is dedicated to emergency alarms for non-preemptive access. Configurable number of ETSs is placed during TDMA data transmit period to meet delay tolerance and support reliability. The number of emergency alarm sources and delay tolerance determine the number of ETSs and the analysis is of it is shown in section 2.5.2.

At the end of DTS period, the BAN superframe proceeds with inactive period or ECAP until next advertisement. The BAN coordinator can dynamically adjust the size of the inactive period or ECAP considering its energy status and total number of requested DTSs from devices.

2.3.3 Fast Allocation and Adaptation

For real-time data, especially in medical applications, data delivery in timely manner is a critical concern because belated delivery of data may not be effective and leads erroneous diagnosis. For example, life critical situations that may be observed frequently in emergency rooms, ambulances or battle fields, the most urgent task is to diagnose the patient's vital status as fast as possible. Therefore, data transmission should be expedited even when a device is installed on the patient's body.

PNP-MAC's fast allocation mechanism accommodates this urgency. As shown in Fig. 2.1, advertisement, CAP, and beacon are concatenated at the beginning part of the BAN superframe. The concatenation of the periods enables fast allocation and adaptation as follows: A device requests DTS allocation during CAP after recognizing existence of the BAN coordinator from the advertisement. Next, the BAN coordinator allocates DTS for the device and notifies it through the following beacon immediately. After confirming its DTS allocation status from the beacon, the device can transmits data at its designated slot. Fig. 2.2 shows the allocation process

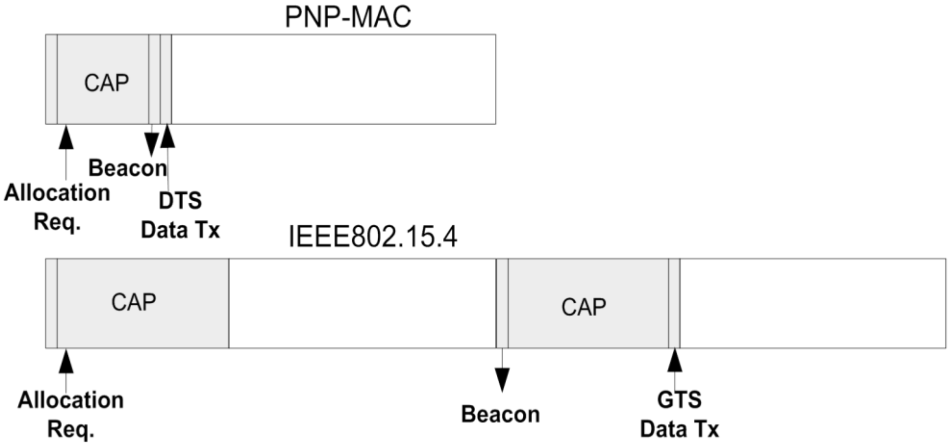


Fig. 2.2 Comparison of allocation process
between PNP-MAC and IEEE802.15.4 MAC

of PNP-MAC and compares it with IEEE802.15.4. Unlike IEEE802.15.4, a consecutive process of DTS allocation and data transmission is completed within current superframe duration. Furthermore, while IEEE802.15.4 allocates GTS from the last one, PNP-MAC allocates from the first one, resulting in substantially reduced allocation delay and enhanced QoS support. In the same way, allocation failure can be detected instantly. If the allocation request fails by any reason, the device that has requested DTS allocation will know it from the following beacon that does not include its allocation status. Then the device requests DTS allocation again at the next CAP period.

2.3.4 Prioritized back-off in CAP

In a contention-based access scheme such as CSMA/CA, the devices intending to transmit data have to perform random back-off and carrier sensing before data transmission for collision avoidance and fairness (e.g., IEEE802.15.4). However, in order to utilize priority based service, PNP-MAC adopts prioritized back-off in CAP period so as to differentiate back-off duration for data that have different levels of QoS requirement.

As shown in Table 2.1, three back-off classes, 0 for emergency alarm, 1 for medical data, and 2 for non-medical data are grouped based on the priorities. The back-off duration is chosen randomly from the range [0, $2^{BE} \cdot$ (back-off class+1) - 1] where BE denotes back-off exponent. Hence, the higher the priorities, the higher the opportunity to seize the channel against lower priority data. It assures services for high priority data even when traffics with different priorities contend heavily at the same time.

2.3.5 Preemptive DTS Allocation

Similar to the prioritized back-off in CAP period, high priority data take precedence over low priority data in DTS allocation as well. To this end, a privilege, named preemptive allocation, is conceded to high priority data. Fig. 2.3 (A) explains DTS preemption when high and low priority data contend for DTS allocation simultaneously.

The BAN coordinator sorts all the DTS allocation requests collected during CAP period in the order of their priorities and allocates DTS slots from the first to the DTS requestors orderly. It is to provide DTS to high priority data first otherwise data even they have higher priority than other

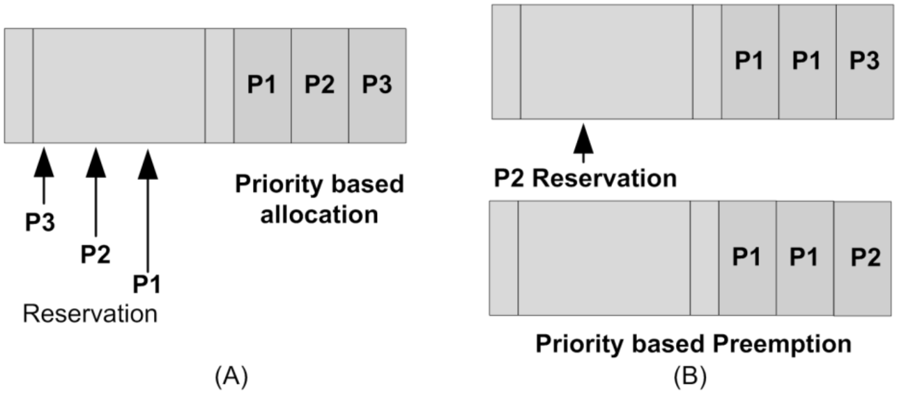


Fig. 2.3 Preemptive DTS allocation

competitor may fail to reserve DTS slots.

A DTS allocation request, even when DTS allocation is already done and remaining DTSs are insufficient to admit it, can preempt DTSs if its priority is higher than that of the current DTSs assignees. First, the BAN coordinator de-allocates DTSs allocated to the devices with lower priority comparing to the newly joining DTS requestor to make enough room for it and then allocates the DTSs to the new device as in Fig. 2.3(B). The original assignee driven out from the allocation may suffer extended delay, but high priority data, in view of QoS, should go ahead of low priority data. The expelled assignee may transmit its data during CAP period or ECAP period, if possible. All the changes in DTS allocation caused by the preemptive allocation become effective promptly through the allocation status announcement of upcoming beacon. The preemptive allocation guarantees services for higher priority.

2.3.6 Non-Preemptive Access

BAN should always be ready to perceive an emergency situation such as abnormality of organs or battery depletion of an implant device to report it to a coordinator in time. Considering significant result of missing a notification of emergency situation, failure of reporting it in time may worsen the situation.

However, prompt and reliable reporting of the emergency situation is by no means easy task because emergency situation is abrupt and unforeseeable. First of all, transmission of emergency alarm using prioritized back-off during CAP period is possible but is not satisfactory in view of reliability. Successful alarming depends on number of contenders and their priorities at the moment. Besides, due to the nature of emergency, reserving a DTS in advance is not feasible and slot preemption may be hard to satisfy delay tolerance, if BAN superframe interval is longer than the delay tolerance (e.g., one second in [1]). Therefore, considering its primary position in reliability and urgency, an emergency alarm deserves a special treatment, namely, non-preemptive access.

As the name, non-preemptive access represents, an emergency alarm does not reserve and preempt a slot in advance but transmits during ETS period, a time period dedicated only for emergency alarms. Configurable number of ETSs is pre-arranged which is sufficiently less than the delay tolerance across the data transmission period. The number of ETSs is determined in consideration of superframe size, delay tolerance and number of alarm sources. Thanks to the ETS period, emergency alarms can be transmitted without interruption to and from other data transmissions.

However, in fact, the best way to minimize delay and enhance reliability for the emergency alarm is slot preemption with the aid of DTS assignees. An emergency alarm can be transmitted at any DTS if current DTS assignees can perform short clear channel assessment (CCA) at the beginning of reserved DTS. The assignee remises its occupancy if it detects emergency alarms during CCA. However, it is a waste of precious resource especially for the battery operated devices to perform CCA at every reserved slot for emergency alarms that may not occur frequently. The energy consumption of CCA is non-trivial for energy-constrained BAN devices. Hence, we argue that our non-preemptive approach to support emergency alarms assures successful transmission of both DTS assignee and emergency alarm.

2.3.7 Access schemes in ETS

ETS enables emergency alarms to achieve QoS but on the other hand it entails two problems, under-utilization of DTS and collision among multiple emergency alarms in ETS period. Firstly, in practice, the situations triggering emergency alarm messages may not frequent in daily life. Reserving ETSs for rarely occurring emergency alarm is obviously waste of resource and it could lead inefficient and under-utilized usage of DTS. Secondly, failure of emergency alarm delivery in ETS due to contention among emergency alarms that arise simultaneously. For example, abnormal functioning of heart may trigger all the sensors related to heartbeat consecutively such as pulsation, blood pressure, and ECG. Simultaneous occurrence of multiple emergency alarms may cause collision in ETS because all alarms try to transmit as soon as upcoming ETS period begins. To resolve this, we propose three ETS access schemes, RP-CCA (Random and Prioritized CCA) for narrow band PHY and BR (Backoff-and-Reserve) and BG (Backoff-and-Go) for UWB PHY.

2.3.7.1 Random-and- Prioritized CCA

In order to overcome under-utilization of ETS, we allow non-emergency data to access ETS. When the number of available DTSs is not enough to meet request, a coordinator can assign the requestor device to transmit non-emergency alarm data during ETS. However, the device assigned to an ETS must perform CCA first for the duration specified in (2.1) to avoid collision with an emergency alarm and to remise ETS to the emergency alarm. An integer N in (2.1) is fixed number and determines total CCA duration and must be small enough to not to disturb data transmission. If an emergency alarm is detected during CCA, transmission is postponed and emergency alarm takes ETS.

$$\text{CCA Duration of non-emergency alarm} = \text{unitCCAduration} \times N \quad (2.1)$$

$$\text{CCA Duration of emergency alarm} = [0, \text{unitCCAdurtion} \times (N-1)] \quad (2.2)$$

To prevent collision among emergency alarms, availability of the ETS has to be secured before emergency alarm transmission. Similarly to the previous case, an emergency alarm must perform Random and Prioritized CCA (RP-CCA) for the duration dictated in (2.2). Since RP-CCA duration is chosen randomly and shorter than (2.1), emergency alarm takes precedence over the assignee and avoids collision with other emergency alarms as illustrated in Fig. 2.4.

Only the one who chooses the smallest RP-CCA duration transmits an alarm in current ETS but all the others defer their transmission until next ETS or CAP period whichever comes earlier.

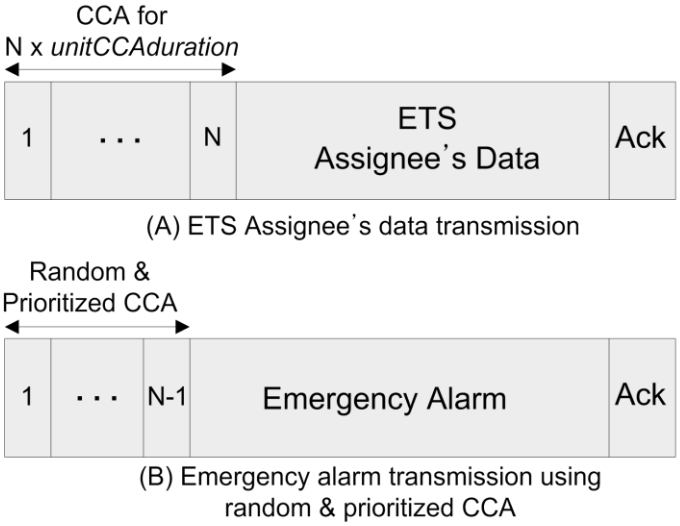
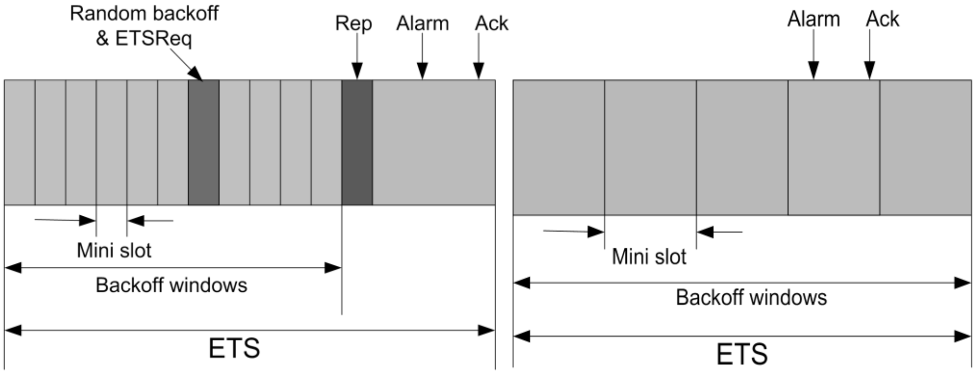


Fig. 2.4 ETS sharing and RP-CCA access

2.3.7.2 Backoff-and-Reserve and Backoff-and-Go

BAN devices should be tiny enough to be attachable or even implantable inside the body. This necessity entails severe restriction on battery size and capacity. Moreover, radio radiation power around human body is strictly mandated by Specific Absorption Rate (SAR) rule, e.g., 1.6W/Kg in US, to prevent harmful affection to human tissue [1][28][29].

In light of these restrictions, Ultra-wideband (UWB) has been spotlighted as an appropriate



(A) Backoff-and-Reserve (B) Backoff-and-Go

Fig. 2.5 ETS Access schemes with UWB PHY

radio technology for BAN. UWB is a radio technology that can be used at very low energy levels for short-range, high-speed communications with a large bandwidth (>500 MHz). Since the transmission power of UWB is much lower than the other WPAN standards, such as Bluetooth or IEEE802.15.4, it is appropriate to cope with the restrictions of radio radiation in BAN [8][9].

However, adoption of UWB in PNP-MAC brings about a problem in exploiting CCA. Because of its low radiation energy level, CCA of UWB signal is very hard. Thus, we need another access scheme that does not rely on CCA.

In the case of CAP period, access scheme can be simply replaced with Aloha protocol which does not require carrier sensing for example IEEE802.15.4a [33]. Therefore, a new access scheme in CAP is to be a combination of slotted Aloha as in IEEE802.15.4a and prioritized back-off. The prioritized back-off in CAP not only supports priority based access but also helps to reduce collision probability which is a critical drawback of Aloha protocol. Therefore we can increase success probability of slotted Aloha which is at most 36% in theory.

Since RP-CCA developed for ETS access of narrowband radio is not compatible with UWB either, a new access scheme is required for UWB to access ETS. Therefore, we propose two new random back-off based access schemes for ETS access, BR (Backoff-and-Reserve) and BG (Backoff-and-Go). As shown in Fig. 2.5 (A), BR exploits random back-off and slot reservation scheme. In BR, an ETS consists of a group of mini slots and an emergency alarm transmission slot. The size of a mini slot is used as unit back-off duration in ETS and is large enough to accommodate ETS allocation request or reply frame.

The detail operations of BR are as follows. When transmitting an emergency data, a device picks a random number k, ranging [1 to br_backoff window]. The device back-off for (k-1) mini slot duration and transmits an ETS request frame at the last k^{th} mini slot. The BAN coordinator replies during reply period only to the device that requests first. Only the device receiving reply can exclusively occupy remaining emergency alarm transmission slot of ETS. Similarly, BG as in Fig. 2.5(B) adopts random back-off and follows the same back-off rule as the BR but it does not rely on a reservation scheme. In BG, the ETS is divided into several mini slots that are large enough to carry an emergency alarm and an acknowledge frames. Just like BR, when transmitting an emergency data, a sensor device picks a random number k, ranging [1 to bg_backoff window] and back-off for (k-1) mini slot duration and transmits the emergency alarm at the last k^{th} mini slot. Since lacks of reservation and small range of back-off window (small number of mini slots) comparing to BR, BG may have difficulty when large number of competitors is contending at the same time. However, an advantage of BG is that it can convey multiple alarms as many as the number of mini slots while the maximum number of emergency alarm that BR can carry is only one all the time.

2.4 Performance Evaluation

We have implemented PNP-MAC protocol using OPNET network simulator which is the most popular network simulation tool and evaluated its QoS performances and also compared them with IEEE 802.15.4 MAC, one of the most popularly used protocols for wireless sensor networks. In consideration of human height and on-body scenarios, the topology designed for simulation is one hop star topology. In practice, a BAN coordinator may be placed around waist or inside of a pocket since the coordinator devices are small enough such as handset and media player. Then, variable number of sensor devices which have different functions can be placed around 1m range from the coordinator like the test-bed experiment scenario in Fig. 2.16.

For the purpose of fair comparison with IEEE802.15.4, we applied BO=6 and SO=3, the superframe configuration parameters of IEEE802.15.4, to PNP-MAC to make the superframe duration and slot size equal. According to IEEE802.15.4 MAC specification, superframe duration is determined by aBaseSuperframeDuration$\times 2^{BO}$ where aBaseSuperframeDuration = 960 symbols = 15.36ms. A symbol size is 4 bit and takes 16us at 250kbps. Slot size is obtained from aBaseSlotDutation $\times 2^{SO}$ where aBaseSlotDutation = 60 symbols = 0.96ms.

Therefore, applying parameters BO=6 and SO=3, BAN superframe duration becomes 0.983sec including the same size of 1 advertisement slot, 8 CAP slots, 1 beacon slot, and 118 DTS slots. Some of the DTS will transform to ETS. In this simulation, inactive period or ECAP is not applied. Transmission power is 0dBm and channel capacity is 250Kb/sec. All these simulation parameters are applied to MicaZ test-bed experiments as well. The body channel model adopted in this simulation is CM3 [3] developed for on-body communication scenario simulations by TG6. Simulation time of all the scenarios is 5 minutes.

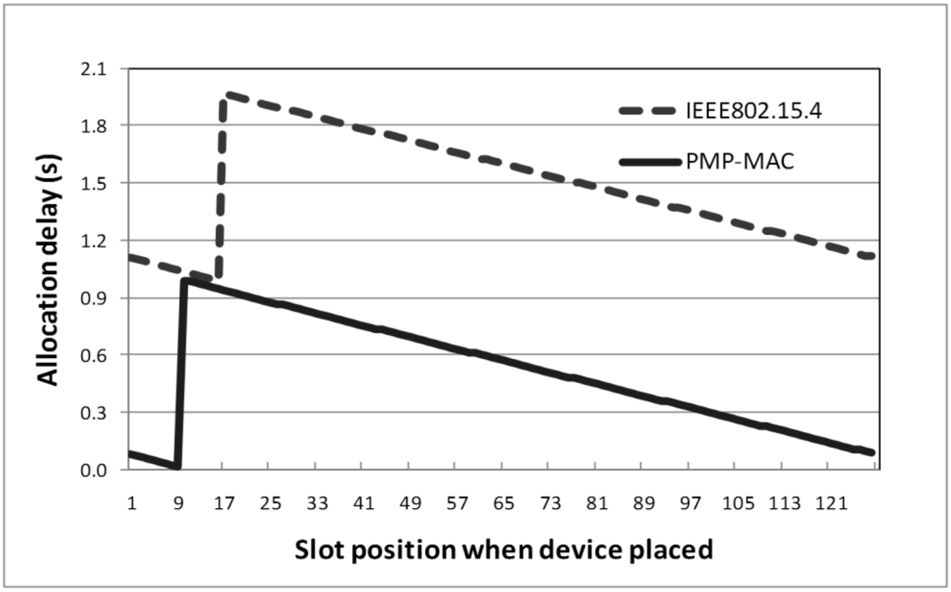


Fig. 2.6 Allocation delays in PNP-MAC and IEEE802.15.4

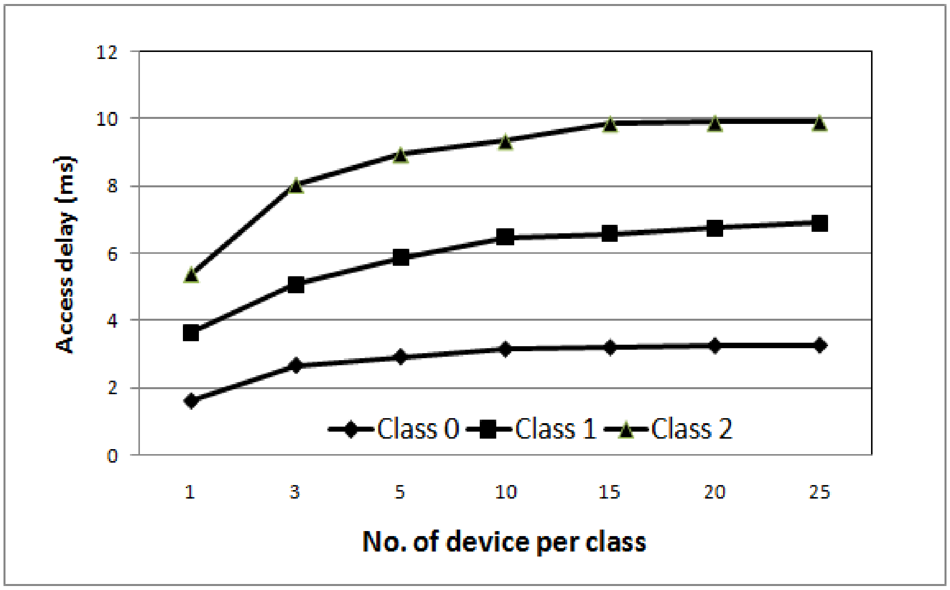


Fig. 2.7 Access delays of classes with prioritized back-off in CAP

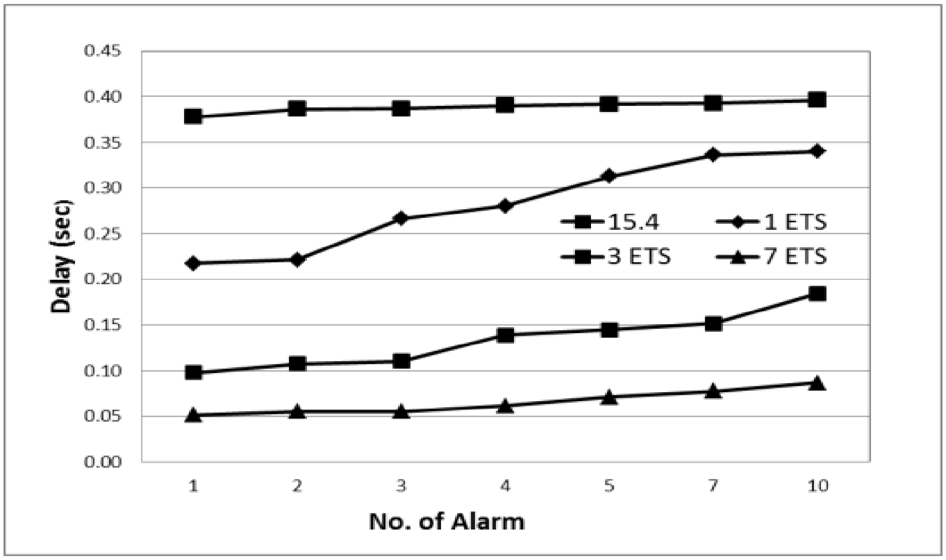


Fig. 2.8 Emergency alarm delays with RP-CCA and IEEE802.15.4

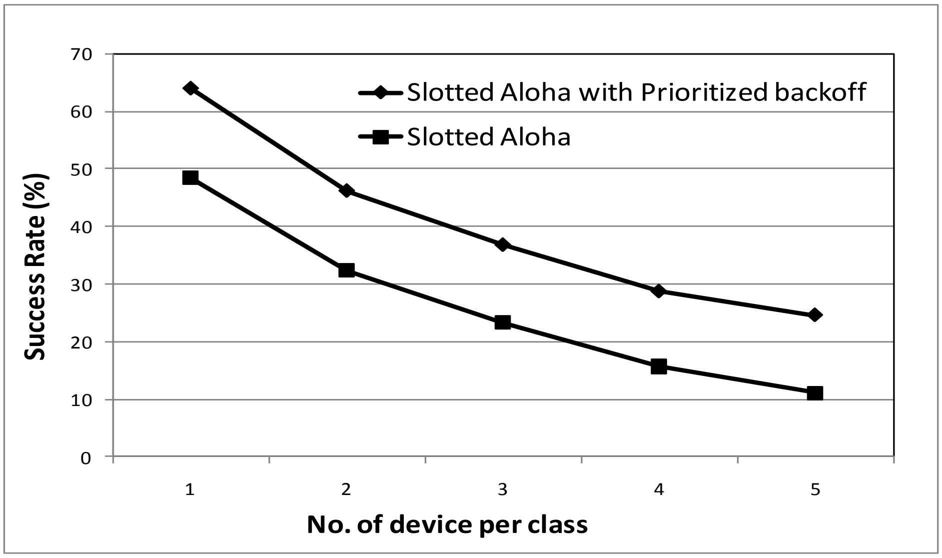


Fig. 2.9 Ratio of acknowledged packet in Aloha
with/without prioritized backoff

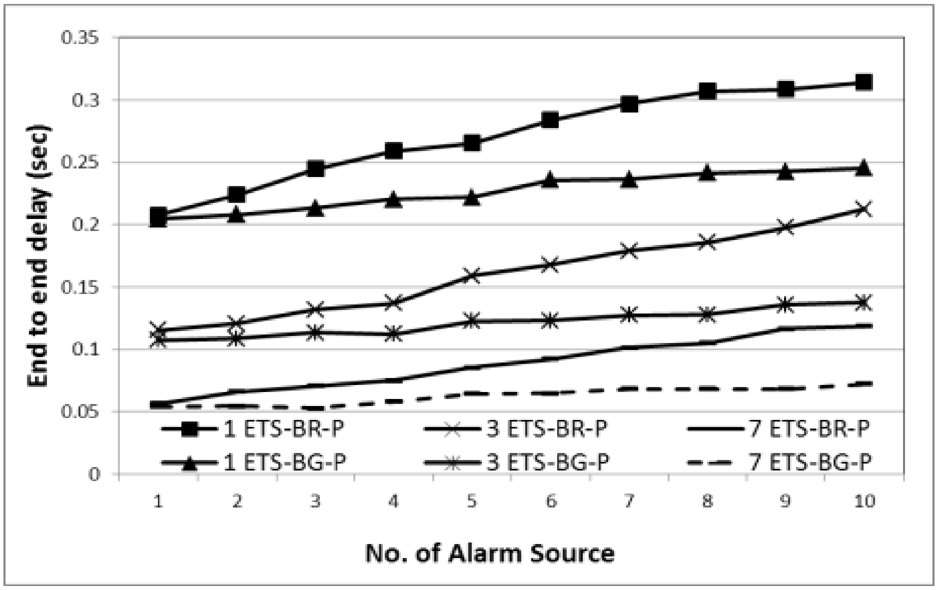


Fig. 2.10 Emergency alarm delays of BR and BG with Poisson traffic

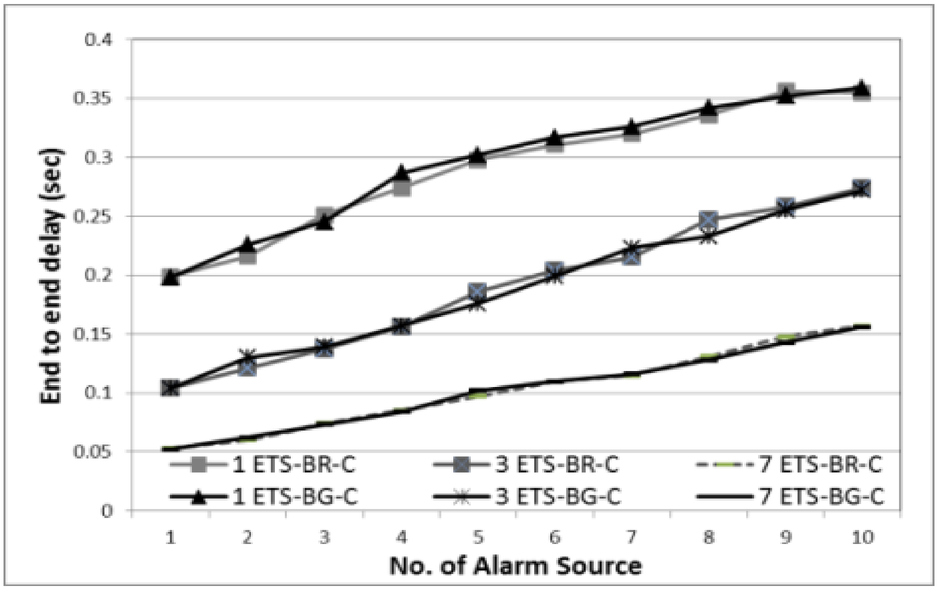


Fig. 2.11 Emergency alarm delays of BR and BG with constant traffic

The first evaluation is the allocation delay comparison between PNP-MAC and IEEE802.15.4. The allocation delay is time duration from device deployment to data transmission at the allocated slot as in slot allocation process in section 2.3.3. As soon as a device is placed on the body of patience with switched on, the device begin to find advertisement. Once receiving advertisement, it reserves one or more DTS during CAP and confirms the reservation status from beacon. If beacon confirms DTS reservation, a device transmits data at the designated DTS. In IEEE802.15.4, a device waits for beacon and reserves GTS right after receiving beacon but the reservation is not confirmed until next beacon period, causing unnecessary delay. It becomes more severe as the superframe duration increases. As observed in Fig. 2.6, PNP-MAC's fast allocation completes entire allocation process within current BAN superframe and it takes only about 1/3 of allocation delay of IEEE802.15.4.

Fig. 2.7 is the measurement of access delay of prioritized back-off in CAP. Access delay is the time duration from first frame transmission to successful frame reception at a coordinator. Devices with three different back-off classes as indicated in Table 2.1 are co-located in a BAN network. The number of devices in each class varies from 1 to 25, in other words, the total number of devices varies from 3 to 75. As indicated in section 2.3.4, the maximum back-off duration is determined by $[2^{BE} \cdot (\text{backoff class}+1) - 1]$ where BE is in the range from 3 to 5. The BE increases when transmission fails. Unit back-off duration used is 320usec. Since low backoff class value, a high priority node has a high probability to choose short back-off duration, leading low access delay. The effect of prioritized backoff becomes distinct when contention is heavy. The delay difference among three classes becomes larger as total number of node increase.

Average emergency alarm delay of RP-CCA and IEEE802.15.4 are shown in Fig. 2.8. Emergency alarm delay is the time taken from alarm generation till its transmission. Emergency alarm frame size is 18 Byte, and average Poisson inter-arrival rate is 0.1 frame/sec. A parameter N explained in (2.3) for RP-CCA is 3 and unitCCAduration = 128usec. The number of emergency alarm sources varies from 1 to 10 and three different number of ETS, 1, 3, and 7 are tested with each number of emergency alarm source. In case of IEEE 802.15.4, CAP is the only period that alarms can be transmitted thus, regardless of the time an emergency situation occurs, the an alarm message should wait until next CAP period. Therefore, average delays in IEEE802.15.4 do not change that much. However, thanks to the ETS, delay in PNP-MAC is dramatically reduced and improvement becomes evident as the number of ETS increases. Moreover, RP-CCA prevents collision among emergency alarms thus, helps reliable delivery. Considering the length of superframe, number of alarm sources, and delay requirement, we can adjust number of ETS.

We designed Aloha-like access schemes for CAP and ETS for UWB PHY. Fig. 2.9 shows success rate of slotted Aloha in CAP with and without prioritized back-off. The number of devices per each back-off class varies from 1 to 5 thus, the total number of devices is 3 to 15. Prioritized back-off lessens the collision a critical problem of slotted Aloha and raises success rate across the entire number of device range.

Another Aloha-like access scheme for ETS are BR and BR. Fig. 2.10 and Fig. 2.11 show performance comparison of BR and BG in ETS with two different traffic types, Poisson and constant traffic at 0.5 frame/sec. The number of mini slots is 10 for BR and 5 for BG. When the traffic is Poisson as in Fig. 10, BG always maintains lower delay than BR. As mentioned earlier, it is because that BG can digest more alarms with less time. The advantage of BG amplifies as the number of alarm sources increases.

Constant traffic simulation is for the case when multiple emergency alarms occur at the same time. As observed in Fig.11, when the traffic is constant, there is almost no delay difference between BR and BG and it induces longer delay compared to Poisson traffic. Since all the alarm sources generate alarms constantly, contention in ETS is very severe for both BR and BG. Though BG can carry more alarms than BR in Poisson traffic, it is insufficient when traffic is constant and heavy. Undelivered alarms during ETSs attempt to transmit during CAP period. Overall, we can say BG is more beneficial than BR regardless of traffic types.

2.5 Test-bed Experiment

2.5.1 MicaZ Test-bed Setup

The objective of the performance evaluation using test-bed is to ensure feasibility of PNP-MAC not only in a network simulator but also in real-world with real devices. For the test-bed setup, we used Crossbow MicaZ mote, one of the most popular wireless platforms for low power, short range wireless sensor networks. It is equipped with AtMega128 micro controller, 128Kbyte programmable memory, CC2420 RF transceiver chip and supports up to 250Kbps data rate at 2.4GHz ISM band. Though MicaZ is not an exclusive BAN device, its small form factor and capabilities such as low power, short range, and low data rate are appropriate enough to emulate BAN devices thus, to validate PNP-MAC's functionality. Fig. 2.12 shows the architecture of implementation and interfaces among layers. The lowest layer represents MicaZ hardware which equivalent to a physical layer. The layer in the middle is MAC and OS layer in which PNP-MAC and operating system reside. Usually TinyOS is popularly used as an operating system in MicaZ test-bed but this time we designed our own OS which is much simpler and lighter than TinyOS. This OS can go with not only PNP-MAC but also other MAC protocols as

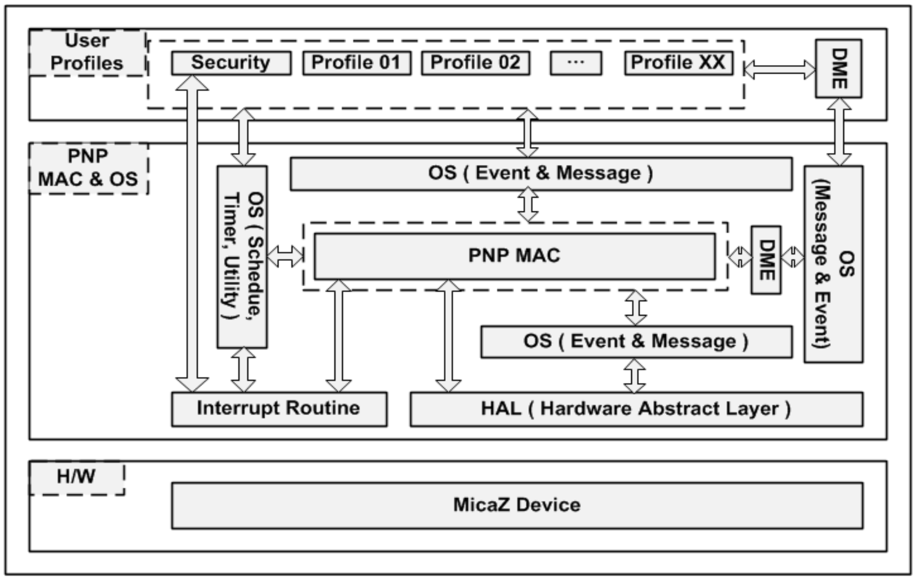


Fig. 2.12 Implementation architecture in MicaZ

well. OS is a central manager, coordinating MAC and physical layer to make them operate properly. It handles various events, messages, interruptions, schedules and informs them to PNP-MAC layer and HAL (Hardware abstraction layer) using primitives. The top layer can be interpreted as an application layer. Though this application layer is not involved in this experiment, it would contain security feature and multiple profiles for different applications as Bluetooth or Zigbee does. BAN has two different scenarios, on-body and in-body. In-body scenario is for the communication of between a coordinator and implant devices but implant devices are not readily available, we do test-bed experiment with only on-body scenario by attaching MicaZ devices on the human body.

2.5.2 Experiment and Analysis

First experiment is to observe emergency alarm delays. Since simulation shows BR is not as good as BG, we test only RP-CCA and BG. For fairness, the experiments are done under the identical parameters and scenarios as those of simulations. First, we put a MicaZ device as

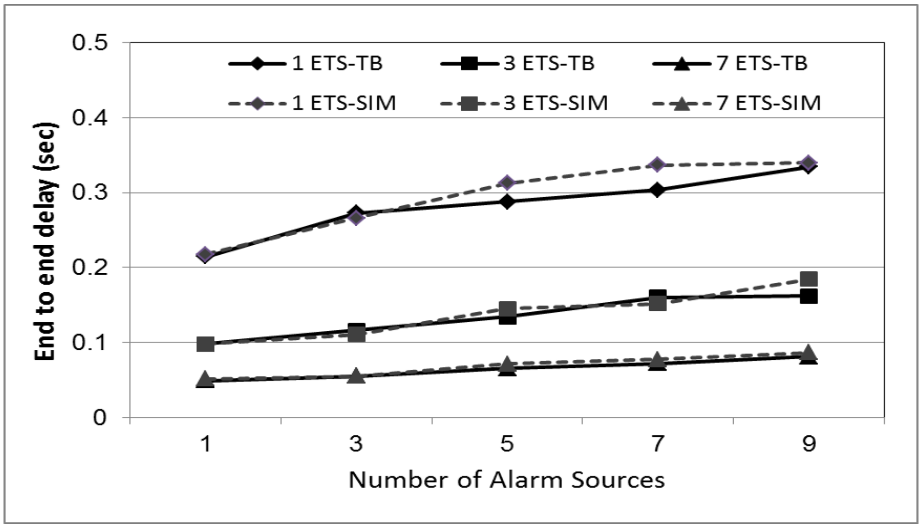


Fig. 2.13 Comparison of emergency alarm delays of RP-CCA in simulation and testbed

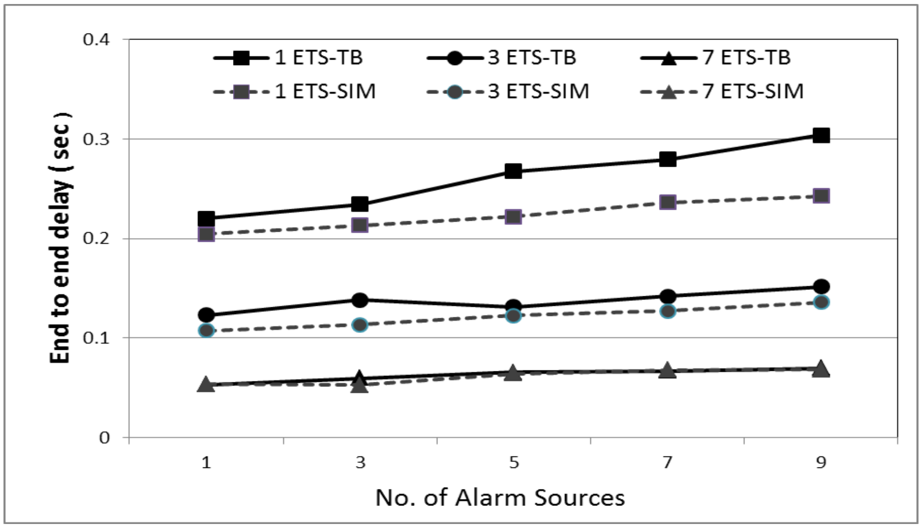


Fig. 2.14 Comparison of emergency alarm delays of BG in simulation and testbed

a coordinator at the center of a table in an office and placed various numbers of children MicaZ devices around the coordinator in accordance with the identical topologies used in simulation of Fig. 2.8 and Fig. 2.10.

Comparisons of emergency alarm delays of RP-CCA and BG between simulation and test-bed are illustrated in Fig.2.13 and Fig 2.14. Taking into consideration of random nature in data generation, CCA duration, and back-off duration, the experiment results show consistency with the simulation results. Through this performance comparison, we claim that PNP-MAC performs properly with real device as well and it also supports credibility of PNP-MAC.

A critical concern to meet the delay requirement of emergency alarm is determination of the number of ETSs. In determination of the number of ETSs, the worst case situation should be considered in which all the alarm sources generate emergency alarm frames simultaneously right after CAP. Although CAP helps achieving low emergency alarm delay as in Fig. 2.13 and Fig. 2.14, it is not be helpful when a large number of alarms are generated simultaneously after CAP and remaining time to next superframe is longer than delay requirement.

In light of the worst case scenario, we need to analyze the required number of ETSs before deployment. We first estimate minimum number of ETSs for BG as follows. First, p, the probability that at least one alarm is successfully transmitted is given by (2.3) where m≥2 is the number of mini slots and n≥1 is the number of alarm sources. i represents i^{th} transmission attempt, in other words i^{th} ETS. Thus, n_i is the number of contending alarms at i^{th} ETS and p_i is the success probability at the ETS.

$$p_i = \frac{{}_{m}\mathrm{C}_{1} \cdot {}_{(m-1)}\prod_{(n_i-1)}}{{}_{m}\prod_{n_i}} = \frac{m(m-1)^{n_i-1}}{m^{n_i}} = \left(\frac{m-1}{m}\right)^{n_i-1} \qquad (2.3)$$

Considering successfully transmitted alarms, n decreases at each ETS and p varies accordingly as well. Therefore, the number of alarms n contending for the next ETS is given by (2.4).

$$n_{i+1} = n_i - (p_i \times n_i) \qquad (2.4)$$

(2.3) and (2.4) feed each other repeatedly until i satisfying $n_i \leq 1$; that is, no more alarms are left to transmit after ith ETS. After finding i, the total delay for the successful transmission of all alarms can be calculated as (2.5). T_{wait} is waiting time for the first ETS, T_{int} is ETS interval, T_{tx} is transmission time including back-off, and D_{req} is delay requirement.

$$T_{delay} = T_{wait} + (i-1)T_{int} + T_{tx} \leq D_{req} \qquad (2.5)$$

However, in the worst case, waiting time could be the same as the ETS interval and transmission time could be the same as the ETS duration. Thus, (2.5) is rewritten as (2.6) and ETSs are placed in TDMA period using the interval obtained from (2.7).

$$iT_{int} + T_{ETS} \leq D_{req} \;\; (T_{wait} = T_{int},\, T_{tx} = T_{ETS}) \qquad (2.6)$$

$$T_{int} \leq \frac{D_{req} - T_{ETS}}{i} \qquad (2.7)$$

One can estimate the required T_{int} by using (2.3)-(2.7). As an example, for the same parameters used in the simulations, i.e., T_{ETS}=7.68 ms, m=5, n=7, and D_{req} =1sec, T_{int} must be less than or equal to 198.5ms. In other words, the maximum ETS interval will be 25 slots.

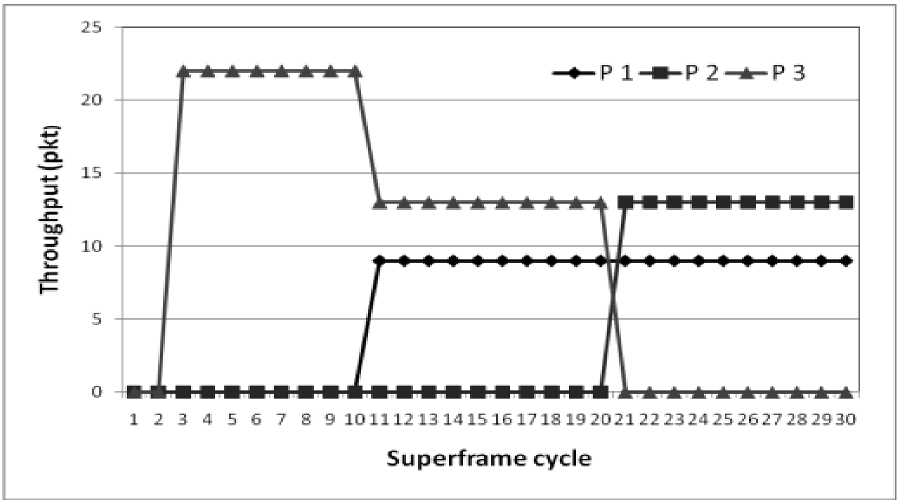
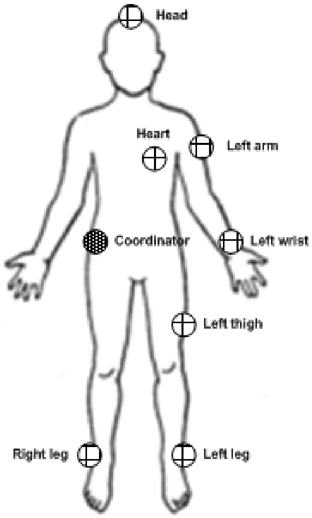


Fig. 2.15 Priority based Preemptive DTS allocation in testbed

Second experiment is to verify the behavior of slot preemption according to the priorities of data. As explained earlier in section 2.3.5, even when DTSs are exhausted, higher priority data can preempt a DTS that is already allocated to low priority data. As shown in Fig. 15, at the beginning of the experiment, priority 3 for example real-time audio occupies whole slots. Priority 1, medical continuous data for example ECG or EMG joins network and preempt some part of the DTSs from the 10th cycle. Priority 2, medical routine data for example blood pressure preempts entire remaining slots from the 20th cycle. Although priority 3 data is the original slot assignee, it remises all slots to higher priority data. It may use CAP and ECAP for its data transmission. Next two experiments are to observe performance of PNP-MAC in a practical situation. Contrary to the previous table-top experiment, we attached MicaZ devices on a human body as shown in Fig. 2.16. We placed 8 MicaZ devices including a coordinator on the volunteer’s body. A coordinator device is placed on the right side of waist, about the center of the body and the others are placed on the head, heart, arm, wrist, thigh, and ankles respectively. Each of them can be regarded as sensor devices monitoring EEG, ECG, EMG, blood pressure, and pulsation. All sensor devices send frames periodically to the coordinator at very BAN superframe.



Sensor location	Emulation
Head	EEG
Heart	ECG
Thigh	EMG
Arm	Blood Pressure
Wrist Ankles	Pulsation

Fig. 2.16 On-body experiment scenario setup

Another purpose of this experiment is to observe how location of a device, transmission power, and movement of the human body impact the performance. To this end, we set 3 different levels of transmission powers, -10dBm, -15dBm, and -25dBm and measured packet reception rate (PRR) in stationary scenario and mobile scenario.

Fig. 2.17 is the PRR measurement in the stationary scenario. A volunteer with 8 MicaZ devices tied on his body as in Fig. 2.16 sits on a chair in an office for 5 minutes while devices are transmitting and receiving data. Same experiments were repeated with 2 other transmission powers levels. The result shows that there is no frame loss when transmission power is at -10dBm and a very little degradation when the power is -15dBm. The overall performance gets the worst when the power is -25dBm as expected. One can find that the performance depends heavily on the location of devices. PRR of the right ankle that is nearly on the line of sight toward the coordinator maintains the same level regardless of transmission power variation but

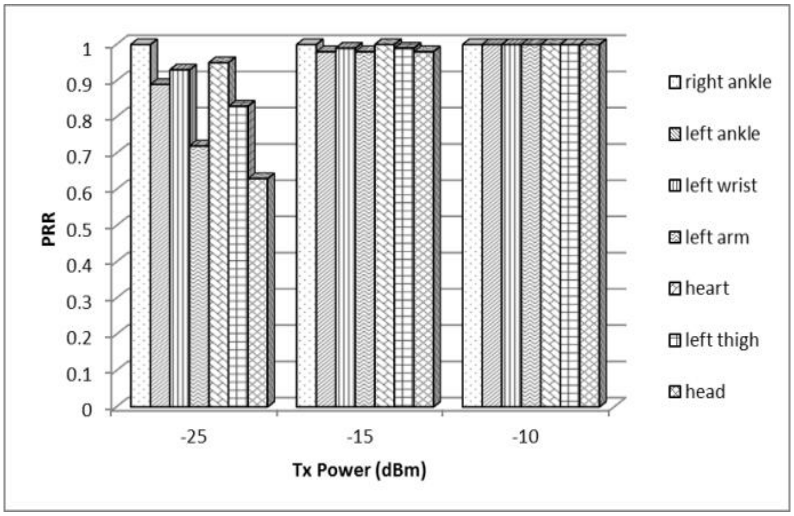


Fig. 2.17 PRR vs. transmission power of
the stationary scenario in testbed

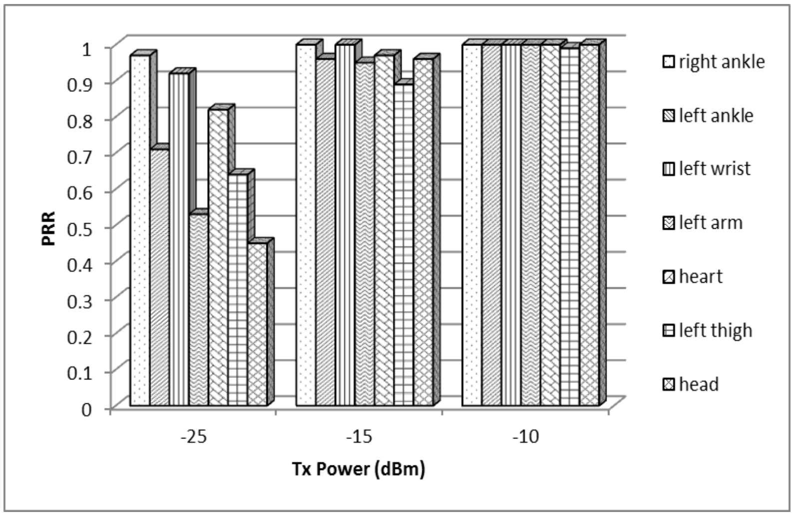


Fig. 2.18 PRR vs. transmission power of
the body movement scenario in testbed

PRRs of the head and the left arm decrease steeply -25dBm. Though the distances from them to the coordinator are not longer than the distance from the right ankle to the coordinator, diffraction and absorption caused by torso obstructs frame delivery as transmission power becomes weaker. The PRRs of the left ankle and the left thigh that lack of line of sight show similar pattern as well. Fig. 2.18 presents the PRR in the presence of the body movement. Instead of sitting on a chair motionlessly, the volunteer walks around the office slowly at about 1m/sec.

Although user is in motion, the PRRs when transmission power is -10dBm show almost same result as those of previous motionless scenario. However, the PRR degradation is distinct when transmission power is -15dBm. Aside from the left wrist that is close to the coordinator, the PRRs of the all the devices located left side are worse than those of the devices located at right side of the body and the performance degradation is more severe than the stationary scenario. But still it shows at least 88% of success rate. The PRRs are worst when transmission power is -25dBm. Average PRR of the devices located at the left side is about 60% but the PRR of head is barely over 40%.

These experiments reveal that transmission power, at least -10dBm is required to achieve reliable PRR performance regardless a user is in motion or motionless state. The fact is close to the transmission power requirement of TG6 that a transmitter shall be capable of transmitting at least –10dBm EIRP (Effective isotropic radiated power).

2.6 Conclusion

BAN is far more fastidious about QoS than other wireless networks. Supporting a wide range of applications and data types, besides strict QoS requirements are hard challenges in designing a

MAC protocol. PNP-MAC is provided to meet the challenges. Its versatile superframe, prioritized services, and preemptive nature support QoS satisfactorily. Its fast allocation substantially reduces initial delay and engages prompt transmission. PNP-MAC also provides such reliable solutions as ETS, RP-CCA, BG, and BR for emergency alarm for narrowband and UWB radios. Intensive simulations and test-bed experiments demonstrate that PNP-MAC operates properly and is capable of supporting a variety of QoS.

The PNP-MAC protocol in this chapter was accepted as a part of IEEE802.15.6, Wireless Body Area Network Standard.

Chapter 3

EGTS

3.1 Introduction

Over the last decade wireless technologies have made revolutionary advancement and provided users quick and convenient connectivity regardless of time and space limitation. Wireless technologies not only enable users to set up a network quickly and conveniently, but also enable them to set up a network where it is inconvenient or impossible to wire cables.

Wireless sensor network (WSN), one of the most recent wireless technologies has been drawing tremendous attentions and research focuses as well because of its practicalities in a wide range of applications. Emergence of wireless sensor network technology, in addition to currently existing infrastructure-based wireless technologies realizes more flexibility in device deployment and network formation. Low cost, low power and small form factor of devices in wireless sensor networks enable users to deploy them easily regardless of the scales of networks. Moreover, it broadens practicality in myriad applications. One of the most recent application is the BAN addressed in the prevous chapter. Another application that welcomes wireless sensor network technology most is the applications in industry such as factory process monitoring, control, and automation.

Among other things, the major beneficiary of wireless sensor network technology is the applications of monitoring and control including environment monitoring, HVAC control, inventory tracking, as well as home and factory automation. Actually, Zigbee alliance [46] has already provided various profiles for those applications together with a network protocol. In light of the facts, wireless sensor network is highly attractive solution to alternate current wired networks in industry which are bulky, expensive, complex and difficult to deploy and maintain. Many researchers and organizations such as wireless HART, BACnet over Zigbee, ISA100 and IEEE802.15 TG4e [45][43][44][32] have exerted themselves to design appropriate wireless solutions for industrial applications. Like other wireless technologies did, they can provide remarkable impacts on control, tracking, monitoring, and diagnostics of the manufacturing processes and equipments [41].

While numerous MAC protocols for wireless sensor network have been designed, IEEE802.15.4 [31][32] could be the most suitable candidate for wireless sensor network MAC protocol. Although it was designed for low-rate applications in short range of wireless personal area networks, its mechanisms are well-fitted for other wireless sensor networks and sensor devices. Moreover, its flexible superframe structure can cope with various requirements of a wide range of applications. However, adoption of IEEE802.15.4 as it is to a certain application is admittedly challenging issue because networks with different scenarios and applications have distinctively different requirements and level of QoS. It may not guarantee same level of performance as it does in other applications.

CHAPTER 3 EGTS

In this chapter, we introduce a novel wireless sensor network MAC protocol, EGTS (Enhanced Guaranteed Time Slot) that is cut out for industrial applications by adopting IEEE802.15.4 MAC protocol as a basis to provide robust and timely data delivery in industrial wireless sensor networks. To this end, we enhanced and added new functionalities to current IEEE802.15.4 providing required QoS support of industrial applications. However, its utilization is not confined only for industry but can be extended to the applications requiring high level of QoS. Among the things to be considered first in designing an industry-specific MAC is the deterministic QoS such as jitter, latency, reliability, and scalability. Predictable deterministic latency is the foremost requirement in real-time monitoring and control systems. Sensor readings from the factory process and equipment should be delivered in timely manner. Belated information may meaningless and belated response could make problems worse and may push users to crisis.

Although IEEE802.15.4 provides synchronization mechanism and TDMA guaranteed time slots to provide QoS, its advantage is quite limited. As shown in Fig. 3.1, a well scheduled beacon synchronizes all the devices in a network and helps on time data transmission but relatively long inactive period in the superframe causes inevitable long end-to-end delay. It is because, not only IEEE802.15.4 but most wireless sensor network MAC protocol focuse on lifetime extension by reducing active period in order to overcome energy constraint. Fig. 3.1 is an example of synchronized 3-hop network and a 0.983sec long superframe consists of the 8 same length of the time duration. In this example, data transmission is possible only when ISD (Incomming Superframe Duration) of child and OSD (Outgoing Superframe Duration) of parent are aligned because of the inactive duration. Therefore, data transmission from node 4 to node 0 takes almost 2sec, twice of the superframe interval.

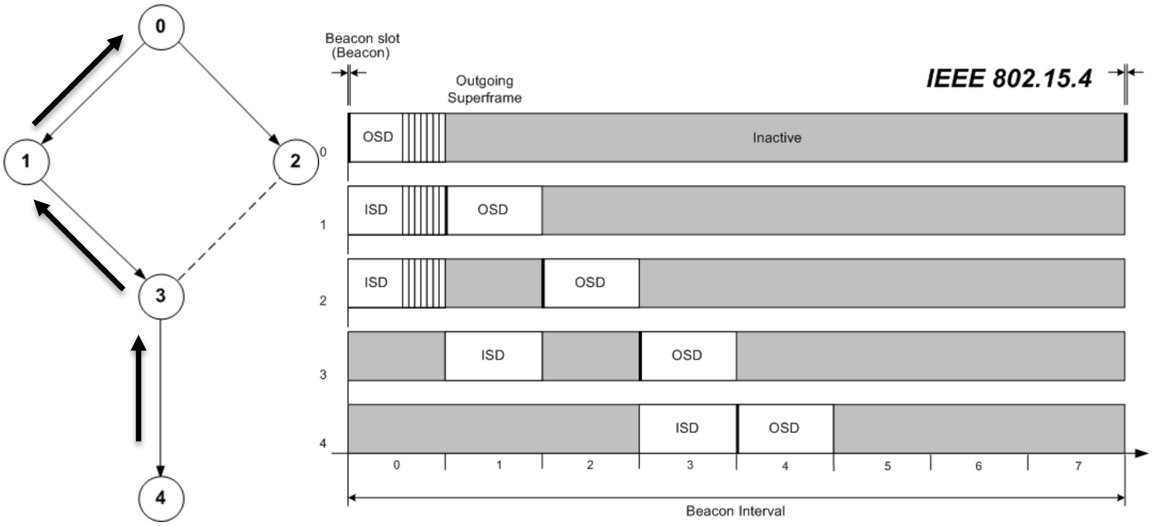


Fig. 3.1 Delay of 3 hop data forwarding in IEEE802.15.4 (BO=6)

End-to-end delay from node 4 to node 1 is (3/8 + 6/8 + 7/8)*0.983 = 1.966

Futhermore, IEEE802.15.4 can support only up to seven guaranteed TDMA slots. It is far less for densy large scale industry monitoring networks besides, TDMA slots are not supported more than one hop as shown in Fig. 3.1, causing long end-to-end delay which is fatal to real-time applications.

In this chapter we propose EGTS (Enhanced Guaranteed Time Slot) protocol that enhances existing IEEE 802.15.4 GTS to provide robust and timely data delivery for multi-hop, multi-channel industrial wireless sensor networks. First, EGTS introduces a fortified superframe structure which extends GTS in space, time, and channel. In IEEE 802.15.4, GTS is supported only within one hop from the coordinator node. However, EGTS supports multi-hop mesh connectivity by extending GTS in space. Moreover, EGTS supports multi-channels while GTS of IEEE802.15.4 supports only a single channel. For time extension, we propose Superframe Reiteration (SR), Multi-superframe Order (MO) and CAP Reduction (CR). Superframe reiteration may not go with energy saving policies of traditional wireless sensor networks MAC protocols but for the applications under timing constraints, providing deterministic guarantees is the foremost requirement to consider than everything else. Multi-superframe order is a tunable parameter similar to beacon order (BO) and superframe order (SO) of IEEE802.15.4 and is to combine multiple superframes in one beacon interval. CAP reduction, together with multi-superframe order assures seamless and reliable data forwarding over multi-hop networks.

Secondly, EGTS performs enhanced time slot allocation mechanism in conjunction with channel allocation through the collaborative three-way-handshaking. The collaboration between a source and a destination node results in fast and reliable data delivery and simultaneous transmissions with deterministic low end-to-end delay.

Third, EGTS provides channel adaptation to enhance robustness to dynamic channel conditions. A factor we should not overlook is the adaptivity to and robustness against industrial environment. Indoor is not a friendly environment for wireless communication because of the severe multi-path effect caused by obstacles. A thorough investigation of channel stability in a factory [35] shows it clearly. In harsh factory environment, there are lots of obstacles such as machinery, pillars and walls that have different reflectivity and permissibility. Moreover, unexpectedly moving-in and out stuffs and people readily affect channel quality instanteneously, leading to high probability of packet loss. In this chapter, we also exploit channel adaptation in TDMA guaranteed time slot so as to prevent packet loss caused by channel quality degradation. It ensures not only reliable data delivery in harsh industrial environment but parallel transmissions that help to reduce end-to-end delay.

The rest of this paper is organized as follows. Section 3.2 presents related work. In section 3.3, the detail designs of EGTS are introduced. Section 3.4 analyzes the end-to-end delay bound of EGTS services using Network Calculus theory [48][49]. Section 3.5 presents the simulation study using OPNET network simulator. Simulations with various metrics and scenarios show that EGTS attains dramatic improvement in delay, throughput, jitter and traffic capacity. Finally, section 3.6 concludes this chapter.

3.2 Literature Review

In order to acheive high throughput with less interference in wireless communication, multi-channel MAC is considered as one of the popular countermeasures, and significant number of multi-channel MACs for ad hoc networks have been developed. However, most of them are developed for IEEE802.11-based as hoc networks which are not suitable for wireless sensor networks. They are designed for the devices that have relatively higher computation power and less concerns for energy cosumption. Recently, several multi-channel MAC protocols designed exclusively for wireless sensor networks are presented in [37][38][39][40]. MMSN (Multi-Frequency Media Access Control for Wireless Sensor Networks) [37] introduced multi-channel MAC for slotted CSMA in wireless sensor networks. MMSN assigns an unique channel to every node and maximizes parallel transmissions, resulting in increased throughput. Although it can achieve high throughput, it is difficult to use in latency constraint network. Moreover, fixed channel assignment may lose connections if channel quality changes frequently. Channel assignment should be flexible and adaptive to the surrounding environment.

TMCP (Tree-based Multi-Channel Protocol) [38] proposed centralized channel allocation scheme in tree-based wireless sensor networks. TMCP allocates channels to disjoint sub-trees and realizes parallel transmissions with less interference. Although it provides intra-tree interference avoidance mechanism from neighboring nodes on the same tree but cannot eliminate intra-tree interference completely. Aforementioned, channel quality in factories is very capricious and can be affected heavily by obstacles nearby. A path that uses same channel can be broken readily in harsh and lossy factory environments.

Y-MAC [39] is an energy efficient schedule based multi-channel MAC protocol for dense wireless sensor networks. It provides synchronization and time slot assignment mechnism between a source and a destination pair as IEEE802.15.4 GTS does. However, it cannot guarantee deterministic latency. When multiple nodes intend to transmit for the same destination, the source nodes have to contend the medium at the beginning of time slot and only one node can have opportunity to transmit. The others that fail to occupy the slot try again at the next time slot. As number of hops or source nodes increase, contention for slot reservation becomes heavier, resulting in highly increased latency and waste of energy.

MCMAC (Multi-Channel MAC protocol) [40] introduces a dynamic centralized channel allocation mechanism for clustered networks. In MCMAC, a cluster head collects channel allocation requests from member nodes and assigns currently available channels to both a transmitter and a receiver. Through the coordination of a cluster head, MCMAC effectively allocates available channels to member nodes, therefore, increases throughput and saves energy. On the other hand, the cluster head has to hold big burden to manage cluster and its members. Especially, severe power consumption is inevitable.

A on-demand channel assignment mechnism, similar to virtual carrier sensing of IEEE802.11 was introduced in [36]. The RTS/CTS/Notification three-way-handshaking is used in channel negotiation between a transmitter and a receiver pair. Together with overhearing, this simple handshaking mechanism can avoid channel overlapping effortlessly. We adopt this handshaking approach in channel allocation of EGTS. In general, this approach is costly for wireless sensor networks but if it is integrated with guaranteed time slot allocation, channel and time slot allocation can be completed at once by exchanging small ABT (Allocation Bit map Table). GTS handshaking is the distributed approach thus, nodes can select a best available channel independently according to their location and channel condition. Superframe enhancement including SR, MO and CR and dynamic multi-channel allocation provide quantitative QoS improvement in end-to-end delay.

3.3 Enhanced Guaranteed Time Slot

As we mentioned, IEEE 802.15.4 has many useful features to be used for wireless mesh sensor networks that require guaranteed QoS in terms of latency, throughput, and reliability. In this section, we introduce EGTS that enhances the IEEE 802.15.4 GTS to satisfy QoS requirement of industrial applications.

3.3.1 EGTS Multi-Hop, Multi-Superframe, and Multi-Channel Extension

The GTS service in IEEE 802.15.4 provides only for single hop communication in star topology. However, EGTS enhances GTS service and support multi-hop communication in mesh topology. EGTS multi-hop extension complicates the time synchronization mechanism because beacon collisions may occur in multi-hop environment. Once a beacon collides with another beacon, the collision persists because of the constant beacon interval without a backoff period for the beacon. The mechanism for beacon scheduling is explored in [45].

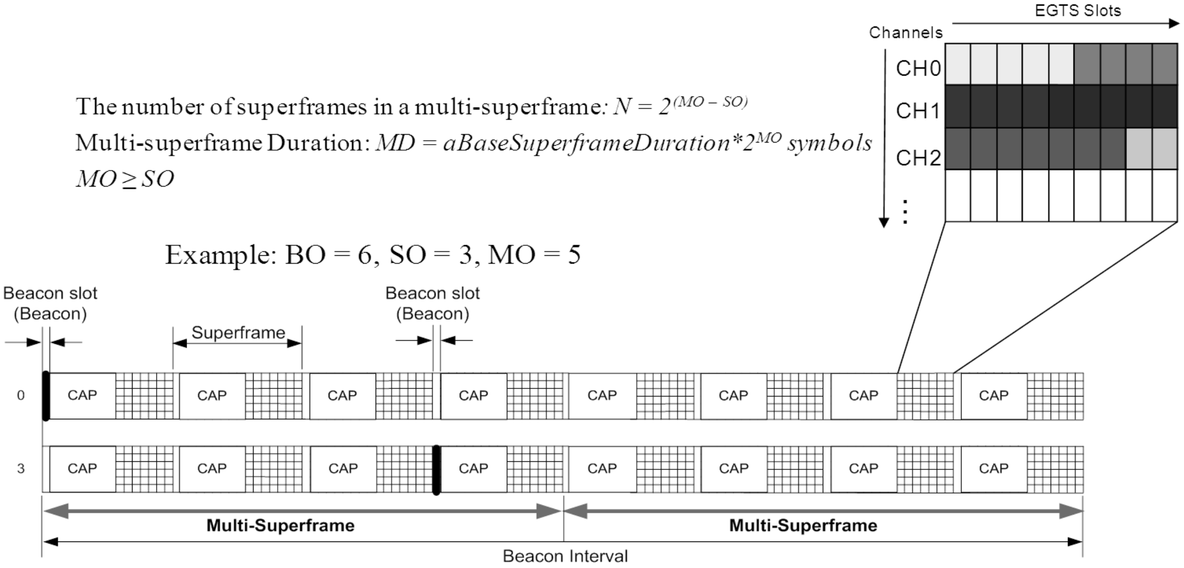
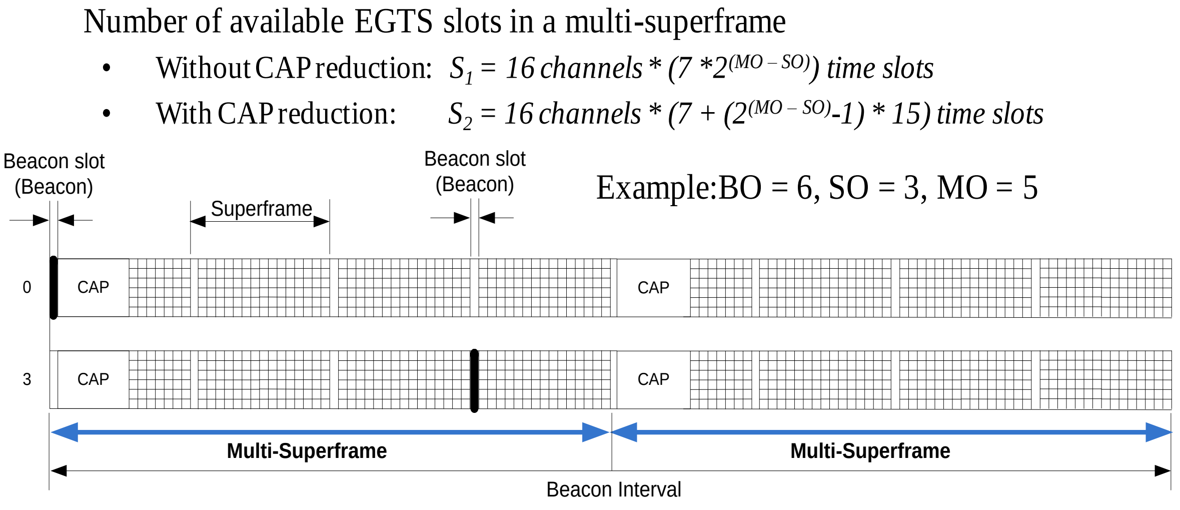


Fig. 3.2. Structure of Multi-superframe and Multi-channel

In order to increase flexibility and availability of guaranteed time services, the limitation of seven GTS slots should be released. Therefore, we expanded single superframe into a multi-superframe structure as in Fig. 3.2. In IEEE 802.15.4, the superframe stucture consists of beacon, CAP (contention access period) and CFP (contention free period) periods in sequence. In order to support guaranteed services, if a node requests all seven GTS slots, services for the other nodes would not be guaranteed. A solution for this limitation of GTS slots would be an extension of superframe to increase available number of GTS slots. A multi-superframe is a cycle of repeated superframes, each of which consists of a beacon, a CAP and a CFP period. The structure of multi-superframe is descrided by the values of macBeaconOrder (BO), macSuperframeOrder (SO), and macMult-superFrameOrder (MO). The macMult-superFrameOrder (MO) represents the number of a multi-superframe in a beacon interval. The multi-superframe duration, MD, is defined as aBaseSuperframeDuration*2^{MO} symbols, for 0 ≤

SO ≤ MO ≤ BO. In EGTS multi-channel extension, this mechnism fully utilizes all sixteen available channels supported by the existing IEEE 802.15.4 compliant devices, and there are seven GTS slots in each superframe. Thus, for example if 16 channels are avaiable, we can have total 112 EGTS slots for guaranteed services in each superframe instead of 7 GTS. An example of a mulit-superframe and of multi-channel structure is shown in Fig. 3.2. In this example, the beacon interval, BI, is eight times as long as the superframe duration, SD, and twice as long as the multi-superframe duration, MD.


Number of available EGTS slots in a multi-superframe

- Without CAP reduction: $S_1 = 16\ channels * (7 * 2^{(MO-SO)})\ time\ slots$
- With CAP reduction: $S_2 = 16\ channels * (7 + (2^{(MO-SO)}-1) * 15)\ time\ slots$

Example:BO = 6, SO = 3, MO = 5

Fig. 3.3. Structure of CAP reduction

In order to accommodate increment in slot allocation requests, the EGTS can further increase number of EGTS slots by reducing CAP period. As shown in Fig. 3.3, the CAP period reduction is achieved starting from the second superframe by replacing CAP period with eight EGTS slots while the very first CAP remains as it is in the first superframe to accommodate contention access traffics and control frames such as EGTS request and reply. In Fig. 3.3, the number of available EGTS slots without CAP reduction is 448 slots while the number of available EGTS slots with the CAP reduction becomes 832. Therefore, multi-superframe and multi-channel extensions can secure large number of time slots and are capable to achieve QoS requirements and reliability in data transmissions, especially for the very high level of reliability required real-time applications, such as a stream of multimedia, real-time target tracking, and emergency alarm.

3.3.2 Three-Way-Handshaking (TWH) Channel and Slot Assignment

Depending on the time-sensitivity of data, a node can exploit deterministic EGTS service using the three-way-handshaking allocation or transmits data during CAP using contention-based transmission such as CAMS/CA and random backoff. While EGTS data frames are transmitted on one of the multiple channels, other types of frames such as beacon, command, and CAP data frames are transmitted on the same single channel which is referred to as the common channel. The common channel selection is at the discretion of the network layer. Every node should recognize this common channel, and is able to listen to all the neighbors’ transactions during the CAP period.

EGTS adopts the three-way-hand shaking for the channel and the time slot assignment. To this end, we first define slot as a tuple of (time slot id, channel id). A source node exchanges the tuple with its intended receiver to negotiate best available time slot and channel for both of them. In order to exchange the tuple, each node maintains a local EGTS Allocation Bitmap Table (ABT) that contains slot and channel usage information at each instant. Initially, the ABT table is set to zero in all rows and columns. A single bit is used to indicate whether a slot is assigned or not. Set the single bit means a specified pair of a time slot in column and a channel in row is already assigned to its neighbors or itself. Otherwise, the time slot and channel pair is availiable to use.

If the number of multi-superframes and channels increase, the size of ABT table is expanded accordingly. Therefore, the size of the entire ABT table might be too large to be transmited by a EGTS command frame. Althougit it can be transmitted, transmission of large frame could be a big overhead especially for the resouce limited sensor devices. For example, if MO = 7 and SO = 3, the size of the ABT will be 224 bytes. Hence, we propose transmission of an sub-block of ABT which contains a portion of ABT along with the index of the beginning and the ending points of a block which is just 28 bytes in length.

The three-way-hanshaking allocation algorithm requires three command frames: EGTS request, EGTS reply, and EGTS notify. A node having time-critical data in its buffer is able to request EGTS slots for the desired QoS service instead of the contention-based transmission during CAP period. First, a node unicastss EGTS request command frame containing its ABT and the number of slots it will use to the intended destination. Upon receipt of the EGTS request command frame, the destination node performs an OR operation with the requestor’s ABT and its own ABT to see if there are unused time slots and channels. If so, the destination node assigns slot and channel to the requesting node and updates its ABT table. Afterward, the destination node broadcasts EGTS reply command frame that contains newly updated ABT. On the other hand, if there are no available time slots, the node broadcasts EGTS reply command frame to the requesting node but this time it contains a slot unavailability message. Upon receipt of ETS reply command frame from the intended destination node, the requesting node also updates its ABT by OR operation, and broadcasts EGTS notify command frame with the updated ABT. Then, the requesting node waits for the allocated time slot, and switches its channel to the allocated one to transmit data during the current superframe. The destination node also waits until the allocated time slot and switches its channel for data receiving. The switching delay should be considered when a node changes its channel for either transmission or reception. The neighbor nodes around the sender and the destination that overheard either broadcested EGTS reply or EGTS notify command frame update their ABT as well. Similar to the purpose of RTS/CTS virtual carrier sensing in IEEE802.11 that avoids hidden terminal problem, the three-way-hand-shaking also avoids collisions caused by accidental slot and channel overlapping.

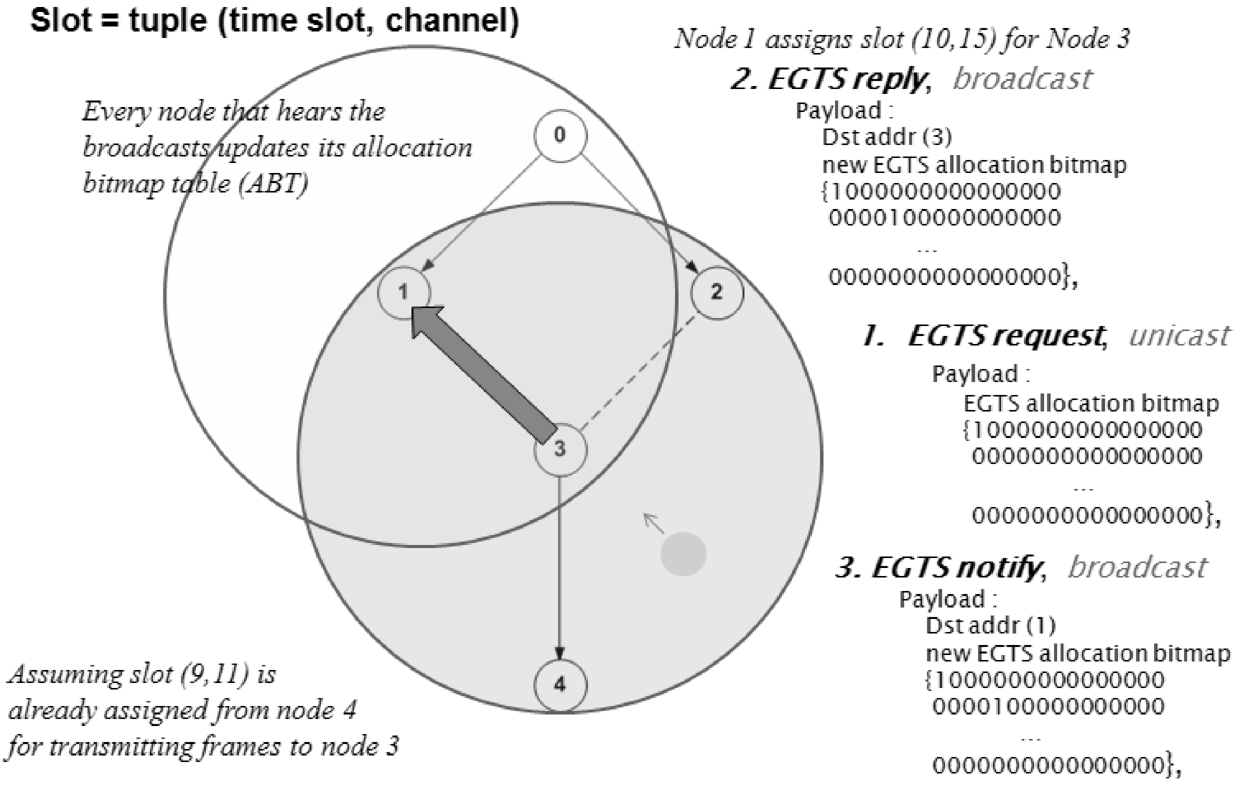


Fig. 3.4. TWH-EGTS algorithm

Fig. 3.4 illustrates the three-way-handshaking allocation algorithm. As indicated in the figure, we assume that the slot number 9 and the channel number 11 are already assigned for the transmission from node 4 to node 3. Since the communication range of node 3 includes node 4, node 2, and node 1, they should know the slot and channel that are assigned to node 3 and 4. If node 3 has time-critical data to transmit to node 1, it transmits EGTS request command frame containing its ABT that is marked 1 (9,11) by unicast. The rows and the columns of ABT represent time slots and channels, respectively. After receiving an EGTS request, node 1 updates its ABT by performing an OR operation with ABT of node 3 and its own, and then selects a new anailable slot and channel pair ((10,15) in the the Fig. 3.4) that is marked in the second row for the time slot and the fifth column for the channel. Node 1 replies EGTS reply frame with newly updated ABT (10,15) via broadcast to inform the update to its neighbor nodes. After receiving the ABT from node 1, node 3 updates its own ABT by performing an OR operation with its current ABT (9,11) prior to broacast the EGTS notify frame. By broadcasting of updated ABT, the neighbors of node 3 are informed reservation of slot 10 and channel 15 by node 3 and node 1, and update their ABT to prevent collision caused by slot and channel overlapping in the future. EGTS slot deallocation can also use same control frames and procedure. In order to deallocate the reserved EGTS slot, EGTS request command frame encapsulates the deallocation subfield in EGTS characteristics field and EGTS request/reply/notify frames are exchanged between a source and a destination node.

The slot allocation scheme can be one of the following two methods, namely, the random slot allocation and the staggered slot allocation. In the random slot allocation, to ensure collision and interference avoidance, a allocated slot should be unique in two-hop away distance. It is feasible because of the broadcating of EGTS replay and notify.

For the random allocation, a node chooses first available time slot that has at least one vacant channel in ABT to avoid a null selection, and then chooses a channel that is available in that same time slot. When selecting a channel, the first available one is chosen.

The staggered allocation is similar to the purpose of DMAC [50] that minimizes hop delay for the QoS of highly time-critical data. Therefore, in the staggered slot allocation, a node that has time-critical data sets a flag of the staggered slot allocation field when the node tansmits EGTS request frame. After receving EGTS request frame, the destination node selects the first available time slot that minimizes frame residing time before choosing a channel, and then it specifies its allocated slot in the EGTS request frame, which is relayed to the next node. Likewise, the next node also chooses a time slot that can relay a frame with the lowest residing time and then select an available channel. This pattern is keep repeated until the final destination node is reached. Comapring to the random slot allocation, the staggered allocation is more complex and may take more time but the staggering mechnism realizes consecutive hop-to-hop forwading, resulting in short end-to-end delay. It is evident as hop distance from a sender to the coordinator in tree increases. The simulation results of end-to-end delay comparison between random and staggered allocation clarify it. The staggered slot allocation is more effective way for high level of QoS.

In EGTS allocation, there might be a possiblity of allocation conflict, especially when mobile node is involed in the middle of the three-way-handshaking. Therefore, we thoroughly investigate about EGTS slot allocation conflict and explain three possible conflicts that may arise. First, as in the Fig. 3.4, once association is done, the blue-shaded mobile node is moving into the communication range of node 3. It has time-critical data, and transmits EGTS request frame to node 3. Assume that at this instance, node 3 is waiting for EGTS reply frame from the node 1. While waiting for EGTS reply, node 3 has to executes EGTS request frame for the newly joined blue-shaded node and replies EGTS reply frame indicating a slot and channel assignment to the bule-shaded node. Right after transmitting of EGTS reply to the blue-shaded node, node 3 receives EGTS reply frame from node 1. Although the possiblity is very rare, the worst case scenario is that slot allocation from node 1 to 3 and slot allocation from node 3 to blue-shaded node are identical. If this coincident happened, the blue-shaded node transmits data at the exactly same time when node 3 transmits date to node 1. Therefore, the transmission attempt of blue-shaded node fails all the time unless it trys to reserve another slot. This is similar to the well-known hidden terminal problem in a ad hoc network with single transceiver.

Second conflict is when a node misses EGTS reply or notify frame because of the characteristics of unreliable broadcasting in capricious factory environment. The missing broadcasting of EGTS replay or notify leads incorrect update of ATB.

Third, new joining node does not realize the recent slot allocation information at a certain area. A newly joining node can know slot usage information only when it overhear on-going EGTS reservation control frames. Otherwise it may exchange reservation control frames several times until it finds proper slots.

If slot conflict happens, the original ower of the EGTS slot can detect it by overhearing neighbors' EGTS reply and notify frames. Once a confliction is detected, the original owner of the EGTS slot transmits an EGTS duplicated allocation notification to the node that is using same duplicated slot by unicast, including duplicated slot ID (time slot,channel) during CAP period. The EGTS duplicated allocation notification forces the nodes having duplicated slot to retry EGTS reservation process.

The EGTS slot and channel negotiation by the three-way-handshaking guarantees that the same EGTS slot and channel pair not to be used by neighbors within two-hop distance but they can be reused beyond two-hop away. Furthermore, the three-way-handshaking is performed by a fully distributed fashion and by on the fly that is very efficient and suitable for the resource-constraint sensor devices.

3.3.3 Channel Adaption

One of the significant problems of wireless communications is the unpredictable variation of the channel quality. Especially in factory-like indoor environment having various obstacles and moving objects, channel quality degradation arised by fading, multipath and shadowing is a severe problem that obstructs successful data reception. The unsuccessful data delivery lowers not only throughput and packet reception rate but increases end-to-end delay and energy consumption. Therefore, EGTS aiming at applications of industry has to have countermeasure for the channel quality variation.

One of the solution for channel quality variation is adaptive channel allocation. Adaptive channel allocation is performed when the link quality of the designated channel falls down below a certain level of received power threshold which is managed by the application layer. The channel quality can be measured by monitoring RSSI (Received Power Strength Indicator) or LQI (Link Quality Indicator) of its own data or of other on-going data from neighbors. If a node detects degradation of channel quality, it transmits EGTS request command frame with a bad flag set to its destination node. Upon receiving, the destination node selects a new channel that is the farthest from the current channel because the channels located close to the current channel may suffer similar quality degradation. Once channel selection is done, the node chooses a new EGTS slotin which the newly selected channel is available, and replies the slot and channel update with EGTS reply command frame.

Multi-channel adaptation can be be used for data that requires higher reliability in a harsh environment where channel quality varies frequently.

3.4 Delay Bound Analysis

Deterministic delay bound is the most important requirement of industrial wireless sensor networks. The delay bound of IEEE 802.15.4 GTS in one-hop star topology has been analyzed in [48] using Network Calculus theory [49][50]. Now, we analyze the delay bound of EGTS in multi-hop mesh topology using this theory.

In industrial wireless sensor networks, the sensor devices typically generate periodic data flow with very small jitter. Following the arrival model in [48], let us consider a data flow represented as a cumulative arrival function R(t) that is upper bounded by the linear arrival curve

$$A(t) = b + r \cdot t \tag{3.1}$$

The arrival curve A(t) represents the maximum number of bits that can be generated by a sensor device in interval [0,t] as illustrated in Fig. 3.5.

Note that only the first hop allows bursty arrival while the next consecutive hops do not. Hence, the latency of the consecutive hops are deterministic. We will first derive the delay bound of the first hop transmission, and then add the deterministic delay of remaining hops to derive entire end-to-end delay.

The delay bound of the first hop can be derived by comparing the arrival curve of the data flow and the service curve of EGTS. The arrival curve is presented as (3.1) and the service curve is the minimum number of bits that are assured to be transmitted in time interval [0, t].

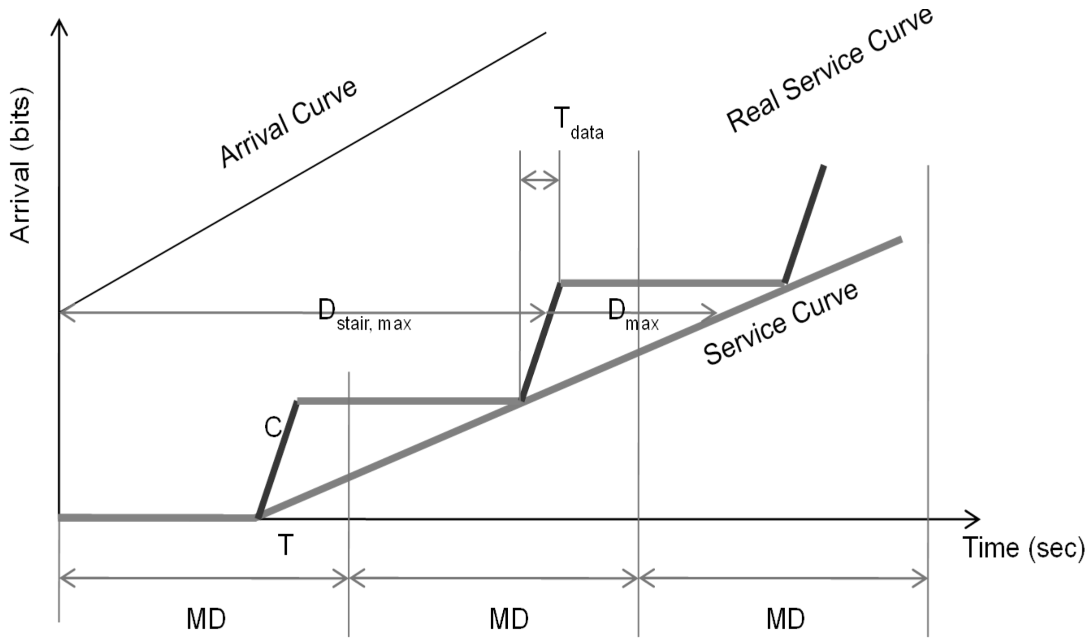


Fig. 3.5. Graphical interpretation of arrival curve and service curve of EGTS

First, we need to define the actual service time during a multi-superframe. More than one data frame can be transmitted in each EGTS duration. If the remaining time is not enough to complete the transmission of the next data frame, the transmission is postponed until the next allocated EGTS in the next multi-superframe. We denote the duration of a time slot as T_{slot} and the actual service time in an EGTS as T_{data}. Note that T_{data} is always smaller than T_{slot}.

Second, we define the maximum latency T waiting time of data frame before the transmission. The maximum latency happens when the data frame arrives just after the end of the assigned EGTS. Hence, the maximum latency is calculated as

$$T = MD - T_{slot} \tag{3.2}$$

where MD is the time duration of a multi-superframe.

Let us define the k-th multi-superframe as the time interval $[(k-1)MD, k\cdot MD]$. Then the service curve in the first multi-superframe is:

$$\begin{aligned} S^1(t) &= \sup[C\{t-(MD-T_{slot})\}] && \text{if } 0 \leq t \leq MD - (T_{slot} - T_{data}) \\ &= 0 && \text{otherwise} \end{aligned}$$

The service curve in the second multi-superframe is:

$$\begin{aligned} S^2(t) &= C\cdot T_{data} + \sup[C\{t-(2\cdot MD-T_{slot})\}] && \text{if } MD \leq t \leq 2\cdot MD - (T_{slot} - T_{data}) \\ &= 0 && \text{otherwise} \end{aligned}$$

The service curve in k-th superframe can be derived by recurrence as:

$$\begin{aligned} S^k(t) &= (k-1)\cdot C\cdot T_{data} + \sup[C\{t-(k\cdot MD-T_{slot})\}] \\ & \qquad \text{if } (k-1)\cdot MD \leq t \leq k\cdot MD - (T_{slot} - T_{data}) \qquad (3.3) \\ &= 0 \qquad \text{otherwise} \end{aligned}$$

Then, the overall service curve $S_{stair}(t)$ is derived as:

$$S_{stair}(t) = \sum S^k(t) \qquad \text{if } (k-1)\cdot MD \leq t \leq k\cdot MD - (T_{slot} - T_{data}) \qquad (3.4)$$

The overall service curve $S_{stair}(t)$ is a stair function as shown in Fig. 3.5.

For simplicity, the real service curve $S_{stair}(t)$ can be approximated as a linear function $S_{approx}(t)$ with a slope parameter R which is expressed as:

$$R = T_{data}\cdot C / MD \qquad (3.5)$$

The slope parameter R represents the amount of bits that are transmitted in each multi-superframe.

According to Network Caculus theory, the delay bound D_{max} for a data flow with the arrival curve A(t) and the service curve $S_{approx}(t)$ is the maximum horizontal distance between A(t) and $S_{approx}(t)$:

$$D_{max} = \sup[\inf\{\tau \geq 0: A(s) \leq S_{approx}(S+\tau)\}] \tag{3.6}$$

as illustrated in Fig. 3.5.

The first hop delay bound can be computed as:

$$D^1_{max} = b / R + T = b \cdot MD / T_{data} \cdot C + T \tag{3.7}$$

Finally, the end-to-end delay bound can be derived simply by adding the constant D^C which represents the total delay of the consecutive hops. Hence the end-to-end delay is bounded by:

$$D^E_{max} = b \cdot MD / T_{data} \cdot C + T + D^C \tag{3.8}$$

3.5 Performance Evaluation

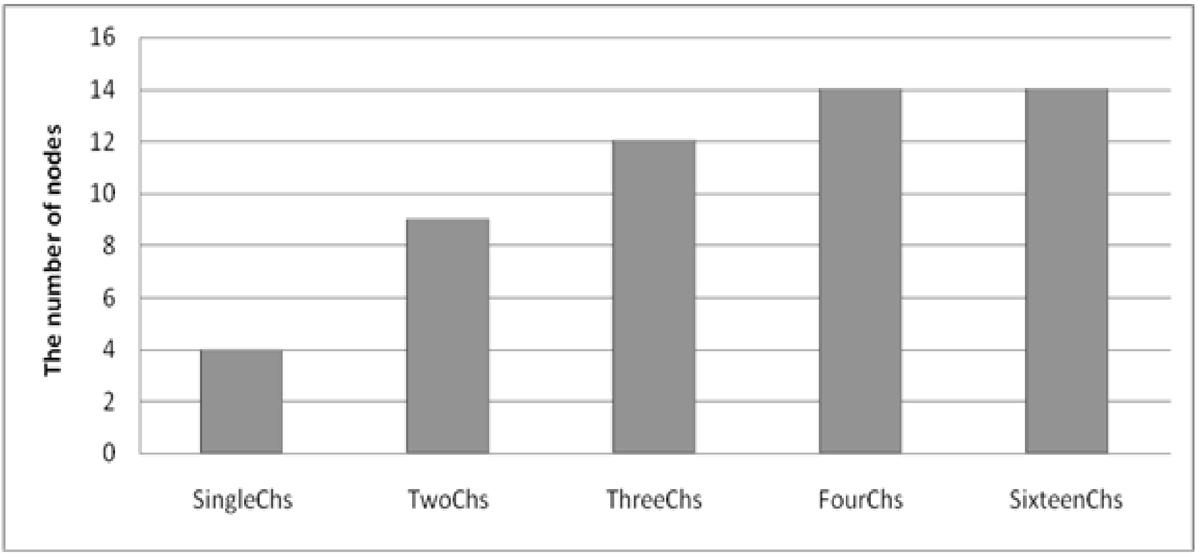
We evaluate the performance of an EGTS algorithm in terms of number of connections, end-to-end delay, standard deviation, and the cumulative distribution function (CDF) in binary traffic models, such as Poisson and CBR in diverse topologies.

The simulation is carried out on a 100m x 100m square-type field and fifty nodes are distributed randomly in it, and their communication range is 13m. In order to ensure reliable data transaction, a buffer-overflow issue is disregarded. A single packet for EGTS data approximately corresponds with one time slot considering transmission time, propagation time, and channel switching time. The actual packet size is 35 bytes and the length of a time slot is 1.92ms. The values of BO and SO are 4 and 1, respectively. All the simulation time is 900 sec.

In the simulation, source nodes are selected randomely and data transmitted from the source node converge into a coordinator, which is the typical scenario for data collection in wireless sensor networks. The formulation of topology and beacon scheduling are performed according to [47] in advance and the EGTS allocation algorithm is implemented according to the procedure in section 3.3.2.

3.5.1 Supported Number of Connections

In this simulation, we can observe the inherent limitation of IEEE 802.15.4, which uses a single channel and supports only seven GTS slots in a superframe. In single hop communications, each connection can be supported with one GTS slot. However, for multi-hop communications, each intermediate node requires at least two slots per each connection, one for receiving and the other for forwarding. Depending on the topology, especially in mesh topologies, a network may be able to support more than seven connections with seven GTS slots because of spatial reuse. However, in a tree-based topology, which is a typical topology in industrial sensor networks, a network cannot support more than seven connections with seven GTS slots because of the bottleneck effect at the coordinator.



Fig, 3.6. The number of supported EGTS connections, MO=1

First we implemented GTS and EGTS without multi-superframe extention. IEEE802.15.4 GTS is tested with a single channel and EGTS is tested two, three, four and sixteen channels. As Fig. 3.6 shows, IEEE802.15.4 with single channel support only four end-to-end connections. However, EGTS using multiple-channels, instead of a single channel, achieves up to 3.5 times more connection. The advantage of using multi-channel, however, terminates when number of channel is four, which is a saturation point. The maximum number of source nodes that EGTS can support remains same from four channels to sixteen channels. It is because of the bottleneck effect at the coordinator. This limitation basically stems from seven fixed number of GTS slots. The limitation can be solved using multi-superframe extension. Consequently, the overall performance of the current scenario depends more on the number of available slot, instead of the number of available channels.

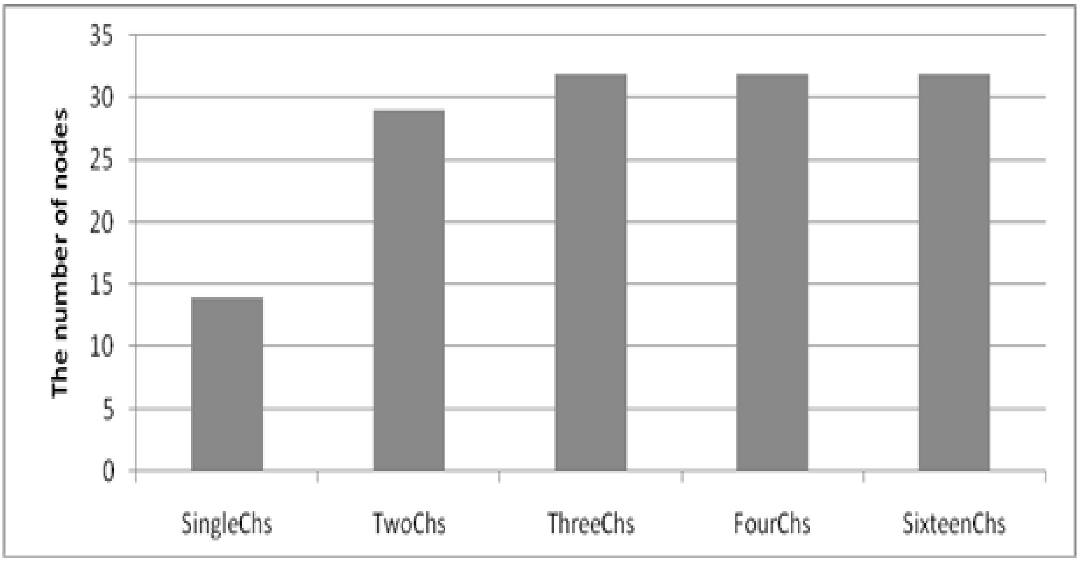


Fig. 3.7 The number of supported EGTS connections, MO = 2

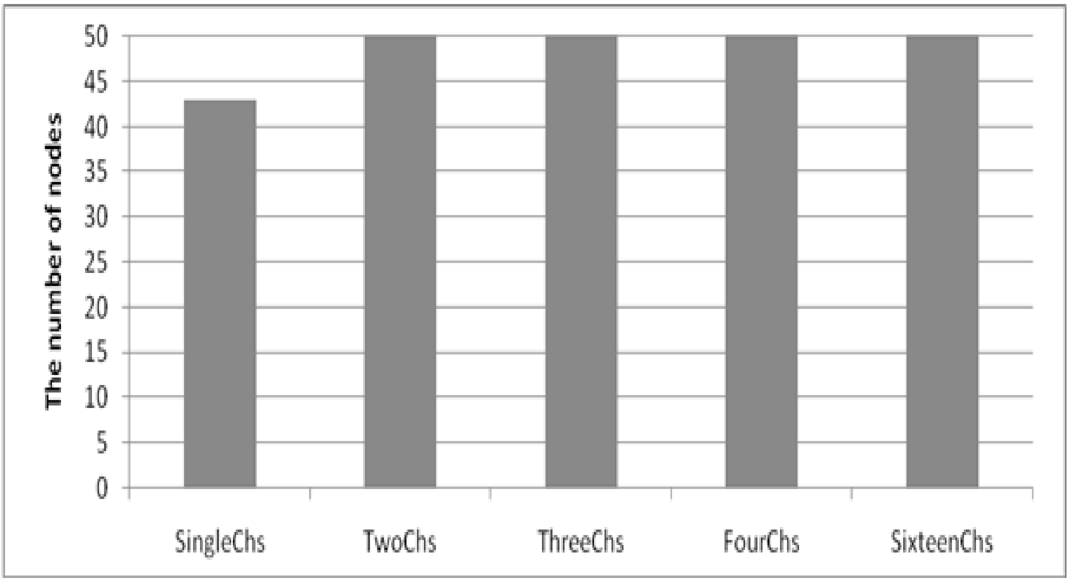


Fig. 3.8 The number of supported EGTS connections, MO = 3

In order to overcome the deficiencies of IEEE802.15.4 GTS, we proposed MO which allows multiple superframes in one beacon interval thus, the multiple superframes can provide more than seven EGTS slots. When MO=2 and SO=1, we can have two superframes in a beacon interval and they provides up to fourteen EGTS slots. As in the Fig. 3.7, the number of supported EGTS connections reaches around 30 out of total 50 source nodes. However, even with sixteen channels, the network could not support more than 35 connections. All 50 connections are supported when MO=3 and more than two channels are used.

3.5.2 End-to-End Delay

For evaluation of end-to-end delay, we implemented random slot allocation, staggered slot allocation, and non-beacon mode of IEEE802.15.4. The results of this simulation is the average of 36 simulations in 6 different topologies with variable number of source nodes. The total number of source node varies from 3 to 50 in each topology. The IEEE 802.15.4 non-beacon mode does not support TDMA GTS but it just provides contention access period without complex beaconing, synchronization, and TDMA GTS management. As shown in Fig. 3.9, when the number of source nodes is less than 30, the end-to-end delay in IEEE802.15.4 non-beacon mode is comparable to TDMA random allocation and TDMA staggered EGTS allocation mechanism. Beacon, GTS, EGTS, and multi-superframe of IEEE802.15.4 and EGTS are costly mechnisms for guaranteed predictable deterministic delay and providing high level of QoS for the time-critical data. Fig. 3.9 shows that if a network size is small, IEEE802.15.4 non-beacon mode, one of the simplest access techniques achieves end-to-end- delay as low as those of TDMA access techniques that are complex and costly. However, as the size of network increases end-to-end delay of IEEE802.15.4 non-beacon mode rapidly increases. As a result, although it is complex and costly, TDMA access of GTS or EGTS is better choice for a large scale industrial wireless sensor networks.

As explained before, random slot allocation is does not consider hop delay and there are many occasions that the next available EGTS slot does not exist. However, staggered slot allocation utilizes consecutive data forwarding by selecting slots that minimize data residing time at the forwarder nodes. Therefore, data receiving and forwarding can be completed within current multi-superframe duration and eventually, total end-to-end delay from a source node to the coordinator can be minimized. Hence, as the comparison in Fig. 3.9 shows, the end-to-end delay of the random slot allocation is greater than the staggered slot allocation. However, still, random slot allocation is significantly better than IEEE802.15.4b non-beacon mode in terms of end-to-end delay.

Fig. 3.10 also shows the average end-to-end delay measurement of the simulations with six different topologies but this time fourteen source nodes are randomly chosen and traffic load varies from 0.1kbps to 0.5kbps in two traffic pattern, CBR (Constant Bit Rate) and Poisson. IEEE802.15.4 model in this simulation contains inactive period, a popular model for energy constraint wireless sensor networks. As shown in Fig. 3.10, end-to-end delay remains constant in staggered slot allocation with CBR data traffic. Since traffic pattern is CBR, the data generation rate in a superframe duration is also constant. In this setting, the number of connections between sources and the coordinator exceed traffic generation rate. However, the delay for Poisson traffic pattern in the staggered slot allocation slightly increases as traffic load increases because forwarding capacity is not enough to handle traffic generation rate and more than one data may accumulate in buffer, causing gradual increase in end-to-end delay.

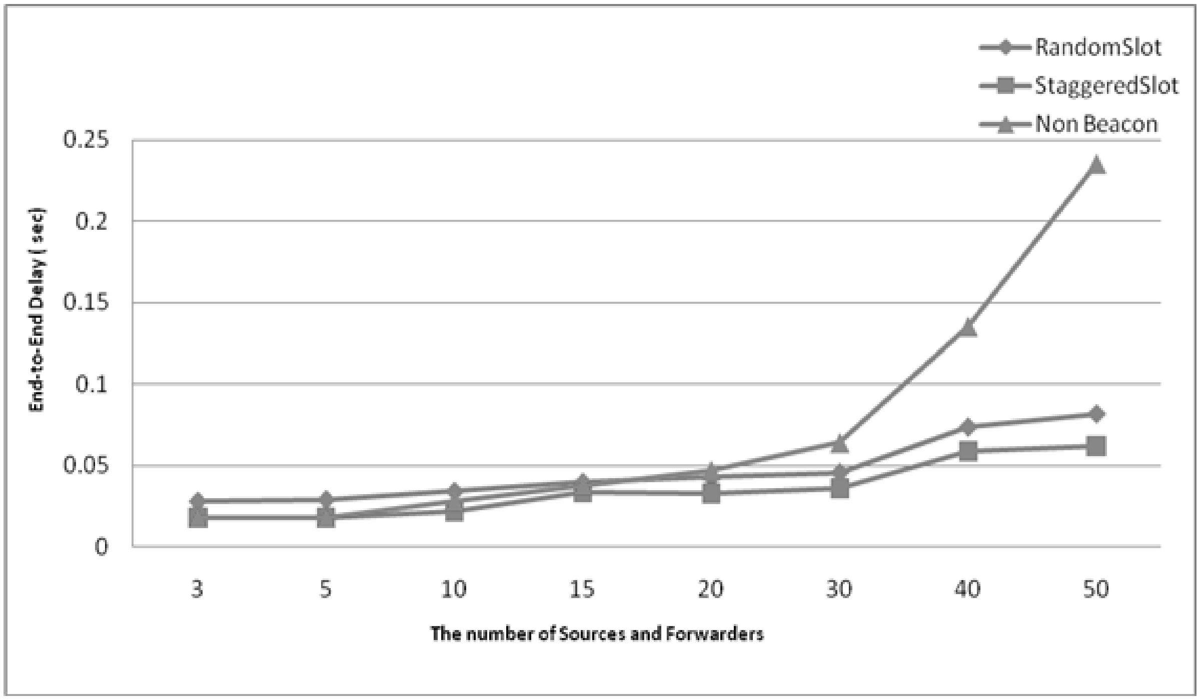


Fig. 3.9 End-to-End Delay vs. the number of sources

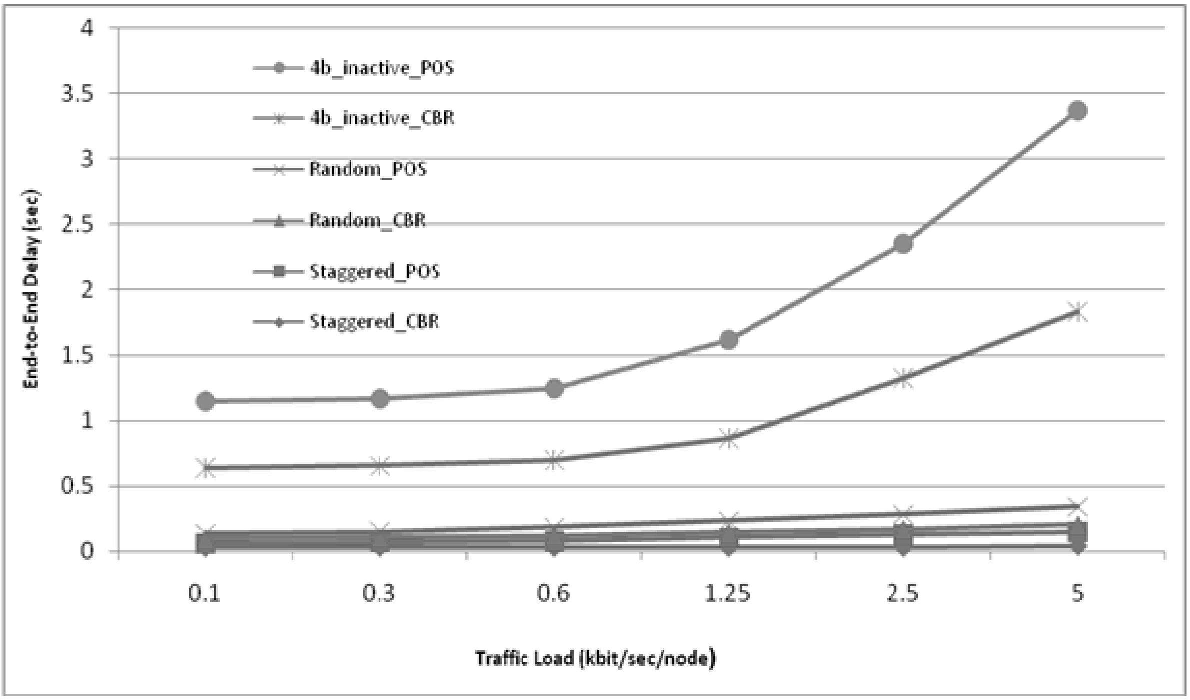


Fig. 3.10 End-to-End Delay vs. Traffic Load

However, comparing to other allocation mechanisms, the increasement is ignorably small. The same phenomena also occur in the random slot allocation with both CBR and Poisson traffic pattern. End-to-end delay increases as traffic load increases regardless of traffic pattern and end-to-end delay in Poisson traffic is little longer than in CBR. It is also due the random nature of Poisson traffic. End-to-end delay of IEEE802.15.4b inactive mode is far greater than the previous ones. End-to-end delay in IEEE 802.15.4b inactive mode with CBR traffic is substatially longer than previous random and staggered allocations. The main reason for this is that data transmission is possible only in active period, which is just 1/8 of a beacon interval plus, although it is not significant, backoff delay contributes to end-to-end delay increament as well. In terms of energy efficiency, inactive mode is one of the best way that lowers energy consumption but the inactive period is a significant barrier for the time-sensitive data and is not appropriate for industrial applications.

Fig. 3.11 compares standard deviation of previous end-to-end delay measurement in Fig. 3.10 when traffic pattern is CBR. The reason to see the standard deviation is that standard deviation of end-to-end delay can be considered as a jitter, one of the QoS measurement in real-time applications. Investigating jitters, we can assure stability of the mechanism for multi-hop data delivery. The standard deviation of staggered slot allocation stay constant across entire traffic load range. It is because that the end-to-end delay in entire traffic load range was constant as we saw in Fig. 3.10. However, the jitter for CBR traffic in the random slot allocation increases as the traffic load increases. It confirms again that staggered allocation also performs batter than random allocation in terms of jitter. As the traffic load becomes heavier, the standard deviation in IEEE 802.15.4b also increases substantially as expected. The jitter in IEEE 802.15.4b is about

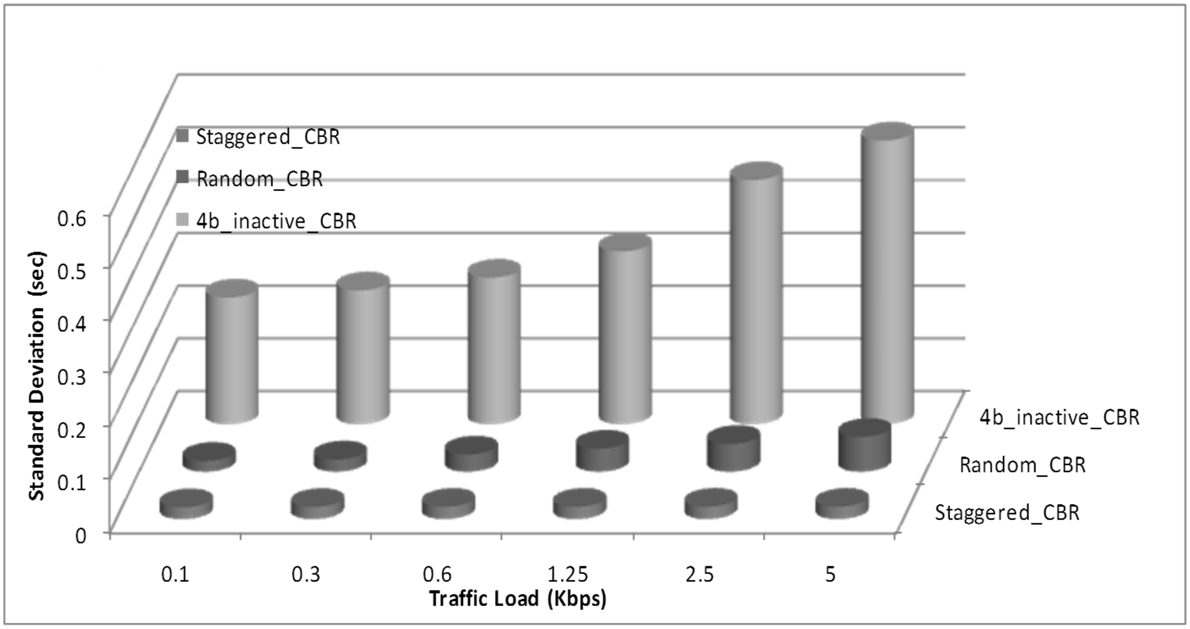


Figure 3.11. Standards Deviation vs. CBR Traffic Load

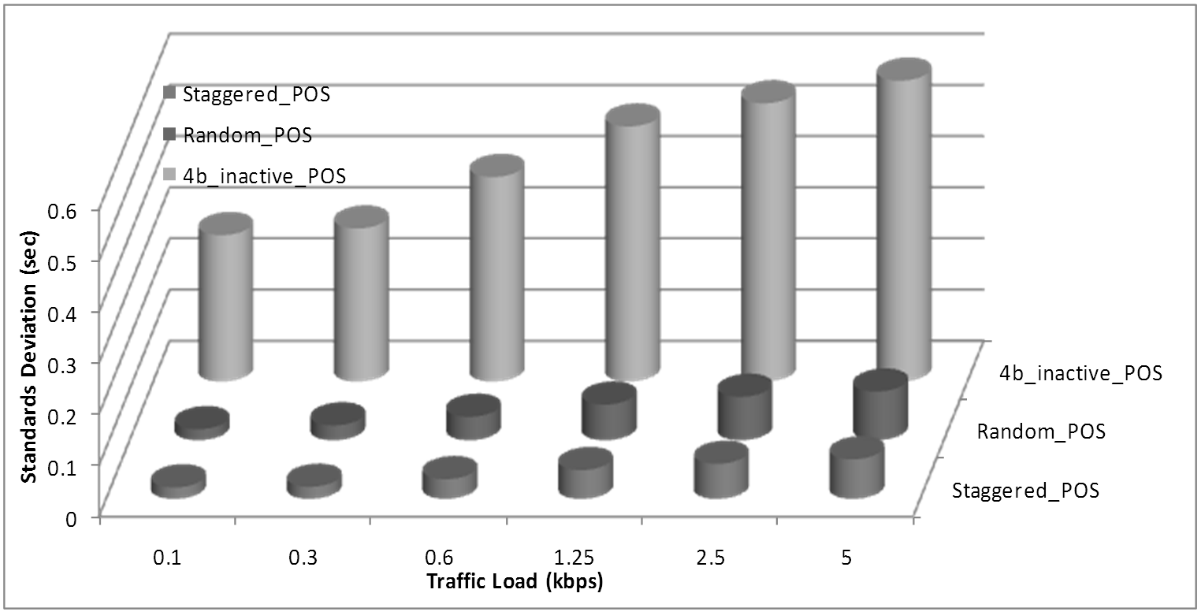


Figure 3.12. Standards Deviation vs. Poisson Traffic Load

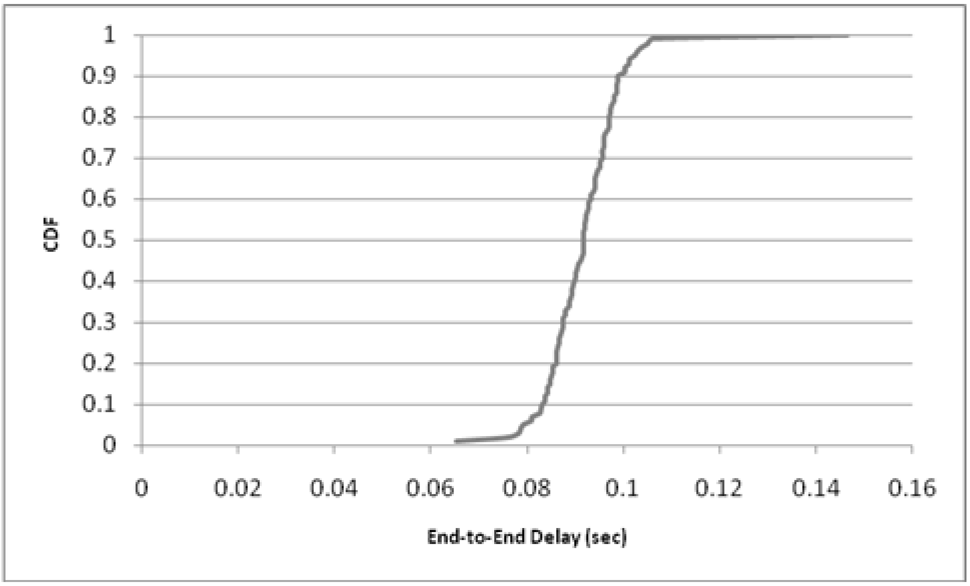


Figure 3.13. CDF of Poisson traffic

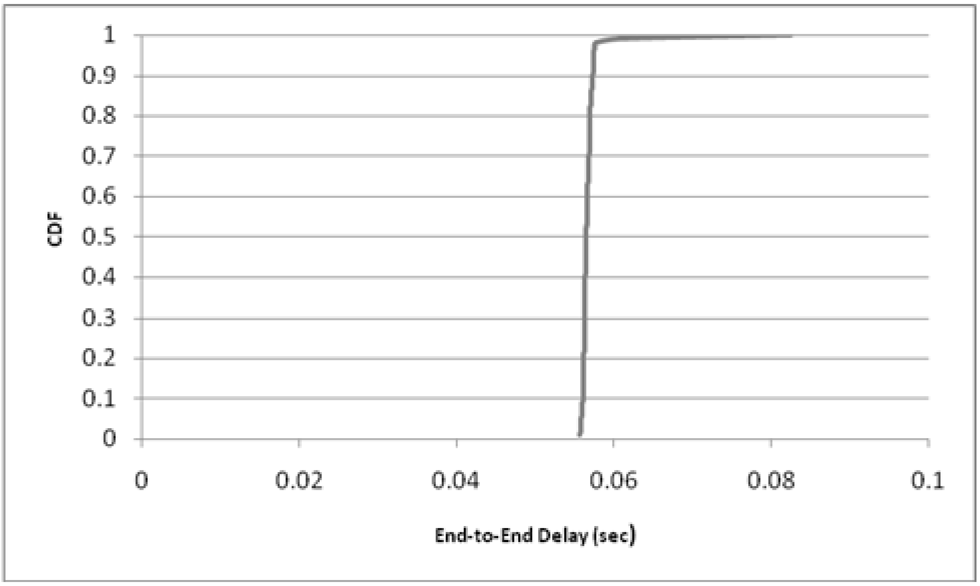


Figure 3.14. CDF in CBR traffic

ten times higher than the other two allocation mechanisms. Therefore, IEEE 802.15.4b inactive mode which supports only contention-based access scheme during the CAP period is not suitable for the applications that require deterministic delay.

Fig. 3.12 is the jitter measurment with Poisson traffic. Jitter in all three mechnisms are worse than in CBR. It is also because of the randomness in data generation regardless that makes extra buffering delay.

Fig. 3.13 and Fig. 3.14 show CDF (Comulative Distribution Function) of staggered allocatoin. CDF when traffic pattern is Poisson in Fig 3.13 shows the end-to-end delay distribution spreads from 0.065sec to around 0.1sec, implying 99% of data is delivered in 0.1sec. However, CDF when traffic pattern is CBR, delay distribution spread only from 0.057sec to 0.06sec. Hence, in this simulation setup, QoS can be guaranteed if the end-to-end delay requirement is greater than 0.1 seconds.

3.6 Conclusion

In this chapter we presented EGTS protocol that enhances the current GTS mechanism of IEEE 802.15.4 to meet the QoS requirement of wireless sensor networks for industrial applications. EGTS equips several QoS features, multi-hop extension, multi-channel extension, and multi-superframe extension. Multi-hop extention realizes data delivery among the nodes apart multiple hop away which is not supported by IEEE802.15.4. Two distributed scheduling methods, random allocation and staggered allocation using three-way-handshking allow maxium slot utilization and minimum end-to-end delay. Multi-channel extention enables simulaneous transmissions by aptopting multiple channels in a single EGTS slot. A EGTS slot integrated with

16 channels can allow up to 112 transmission in a superframe. It is suitable for industrial applications that require large periodic data delivery. Multi-superframe extention overcomes time slot limitation. 7 GTS of IEEE805.15.4 is not enough for the large industrial networks delaing with time-sensitive data. Multi-superframe extention reiterate superframes with controlable parameter MO. It can support large amount time-sensitive traffics and provide high levels of QoS. CAP reduction is also helpful when the number of EGTS slots in multi-superframes is not enough. Except the first CAP, all the others can transform to 8 EGTS slots and provide more room for time-critical data delivery.

EGTS protocol in this chapter was accepted as one of the MAC modes, named DSME (Deterministic Synchronous Multi-channel Extension) of IEEE802.15.4e standard, amendment of IEEE802.15.4 for industrial applications.

Chapter 4

GAS

4.1 Introduction

Wireless sensor network becomes a de facto wireless communications technology for the real-time monitoring and control applications such as factory automation and process monitoring. In factory application, as explained in chapter 3, sensor nodes attached to the machines, conveyor belt, or utility pipelines form a tree-based network and report sampled data to a central controller periodically through multiple relay nodes on the route. The most critical performance measure in this kind of application is end-to-end delay.

IEEE802.15.4e standard is aiming to the industrial wireless sensor network applications dealing with the real-time data for which a deterministic end-to-end delay should be supported [42]. DSME (Deterministic Synchronous Multi-channel Extension) mode in the standard exploits multi-channel over TDMA [34][42]. Thanks to the multi-channel extension, slot utilization in DSME is multiplied up to as many as the number of channels thus, end-to-end delay can be reduced substantially. However, although DSME guarantees deterministic end-to-end-delay, it is far from the end-to-end delay optimization because scheduling in DSME is done first-come, first-served fashion. Staggered allocation, comparing to random allocation offers smaller end-to-end-delay. In ideal case data from a source node flows continuously along the route until its intended destination but in practice, continuous slot allocation may not possible all the time because slot allocation is affected by allocation status of other neighboring data flow. Therefore, the end-to-end delay in DSME, though it is deterministic, may not optimal. In order to optimize the end-to-end-delay, we have to investigate all possible combinations of slots and channels for the all nodes in a network but in practice, it is infeasible.

The transmission scheduling in multi-hop networks is a NP-complete problem [51] [52] and, what is more, the adoption of multiple channels enlarges search space and makes the scheduling problem more complex. It is a typical combinatorial problem. A popular method to attain optimization in this kind of combinatorial problems is metaheuristic [58].

Metaheuristc is a popular optimization tool especially for the optimization problems that have large search space where all possible candidate solutions are in. In fact the best method to find the most appropriate solution is exhaustive search. However, exhaustive search is not practical if the size of candidate solutions is large. A famous example for it is the travelling salesman problem that has large candidate solution space and the size of the space grows more than exponentially as the size of the problem, the number of cities, increases.

Metaheuristic is a computational method that optimizes a problem by iteratively trying to improve a candidate solution that satisfying a given objective function. Comparing to exhaustive search, metaheuristic can offer the optimal solution very fast. However, metaheuristics do not guarantee the optimal solution and usually they offer near-optimal solution and it lacks repeatability.

The most popular metaheuristics being used are simulated annealing [60], genetic algorithms [61], tabu search [62], particle swarm optimization [63]. They are applied to various combinatorial problem optimizations, including routing in networks [64], job-shop scheduling [65], and electricity distribution system design [66].

In this chapter, we propose GAS, GA (Genetic Algorithm)-assisted transmission scheduling method for multi-hop, multi-channel networks. GA is the one of the most popular metaheuristics that inspired from the evolutionary process of nature. Evolution in nature is a process of the survival of the fittest. The creatures in nature have chromosomes that are consists of strings of genes. The genes in a chromosome determine characteristics and fitness to the surrounding environment. The ones who have superior fitness to others can survive, and mate with others who have different genes. This crossover of the genes, a major operation in genetic algorithm occurs every generation. The new born creatures that have 50% of genes from each parent may survive and turn the genes over to the next generation if newly recombined genes are better than others otherwise they become extinct. These iterative crossover operations for a very long generation eventually lead the creatures to evolve to adapt to its environment, in other words, reaching the optimal solution in optimization problems. Another major operation of genetic algorithm is mutation, unexpected gene recombination. Although it is not frequently occurs, mutation affects to the fitness and it may suggest new directions of evolution. As evolutionary process iterated, the genes of creatures tend to be similar each other since sharing of dominant genes does not make any improvement at a certain point and eventually they all resembles each other, a local minima problem in optimization. However, mutation is the gene recombination that does not follow crossover rule, and only small portion of variations in genes can make substantial difference from the parents. It may direct creatures evolving sluggishly to evolve faster or evolve to other new directions. In consequence, mutation supplies diversity in evolutionary process.

GA is a proper mataheuristic method that can be applied to a combinatorial problem having multiple of variables such as nodes, slots, and channels in our scheduling problem. In this chapter, we will introduce GA with enhanced GA operations such as sifting and adaptive mutation, and additional supportive features including flow-based allocation, pre-channel assignment that can assist GA to find the optimal node-slot-channel combinations, minimizing the maximum end-to-end delay.

Flow-based scheduling realizes consecutive data forwarding and helps minimizing end-to-end delay. Pre-channel assignment mitigates computational load in searching for the optimal combinations and reduces solution search time considerably. Performance evaluation show that GAS effectively searches the optimal schedule satisfying the objective, minimization of the maximum end-to-end delay, and comparisons with other scheduling schemes show GAS outperforms all the others.

4.2 Network Model and Problem Formulation

4.2.1 Network Model

The network model we consider here is a static multi-hop, multi-channel tree networks the same as the networks of EGTS in chapter 3. Therefore, the most important QoS measure is end-to-end delay as well. Each sensor node in a network transmits data periodically to a coordinator every TDMA cycle at designated time slot. All the slots are of the same size. A shortest path tree rout from each node to a coordinator is established in advance. In this model, we assume that the coordinator has all the required information for scheduling such as number of nodes, nodes’ ID on the path, and connectivity among the nodes.

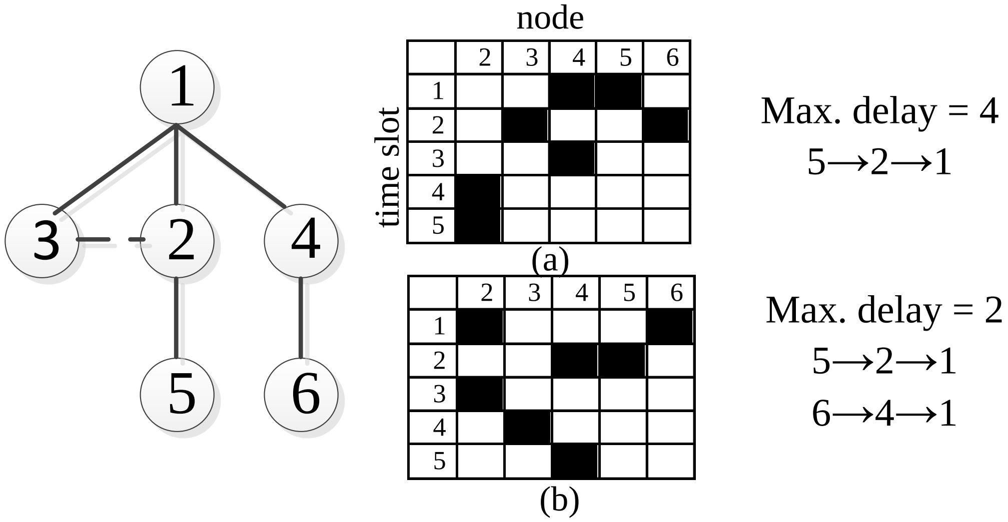


Fig. 4.1 Impact of scheduling on TDMA frame length and end-to-end delay

In practice, this assumption raises extra overhead but it is an inevitable tradeoff for the higher level of QoS. The assumption can be realized by exchanging ABT-like table of IEEE802.15.4e [42] and child number report of IEEE802.15.5 [54].

4.2.2 Problem Formulation

As mentioned, genetic algorithm is widely used in various areas for optimization problems. Genetic algorithm is also utilized for TDMA broadcast scheduling to minimize TDMA frame length and maximize channel utilization [52][53]. They show GA has superior capability in searching for the optimal schedule that satisfies the objective, TDMA frame length minimization. Shortening a TDMA frame length as much as possible is the typical way for delay minimization in networks. However, TDMA frame length minimization may not a proper objective in view of the end-to-end delay optimization for multi-hop networks. It is clearly shown in the example of

Fig. 4.1. The example network in Fig. 4.1 consists of 1 coordinator, node 1, and 5 child nodes, node 2 to 6. Next to the figure of the network are two available slot allocation schedules for the network. Using a schedule either (a) or (b), all 5 child node can transmit their own data in 5 time slots. In terms the objective of TDMA frame length minimization in [52][53], both schedule (a) and (b) are the best solution because they have the same 5 slots of TDMA frame length and they are the minimum frame length. However, in terms of the end-to-end delay, the schedule (a) and (b) have big difference. In the schedule (a), data flow from node 5 to node 1 via node 2 takes 4 time slots, leaving at slot 1 and arriving at slot 4. In the schedule (b), same data flow from node 5 to node 1 via node 2 takes only 2 slots, leaving at slot 2 and arriving at slot 3. Although both schedules (a) and (b) have the same length of the shortest frame length, the maximum end-to-end delay difference is twofold. The difference must be severe as the number of nodes in a network increases. Consequently, frame length minimization does not necessarily mean minimized end-to-end delay.

To be sure, exhaustive search is the best approach in searching for the optimal solution, but it is almost infeasible for our scheduling problem. Assume we have a network with N nodes and the number of time slots and channels given for the network are C and T respectively. The number of hops of each node n toward a coordinator is represented by h_n. Then, the total number of hops H that is equivalent to the total number of transmission in one TDMA cycle, for instance H=7 in Fig. 4.1, is expressed as (4.1) thus, the size of the search space can be determined by (4.2).

$$H=\sum_{n} h_n \tag{4.1}$$

$$|\text{Search space}| = (C \cdot T)^{H} \tag{4.2}$$

The objective function to find the optimal schedule S_{opt} that minimizes the maximum end-to-delay is described as:

$$S_{opt} = \underset{S}{\operatorname{argmin}}\,[\,\mathrm{MaxD}(s_m)\,] \tag{4.3}$$

where $D(S) = \{d_1(s_1), d_2(s_1), \ldots, d_{n-1}(s_m), d_n(s_m)\}$, a set of end-to-end delays of each node n with schedule S_m. S_m is a subset of the search space $S = \bigcup_{i,j,k}\{(n_i, c_j, t_k) \mid i = [1, N], j = [1, C], k = [1, T]\}$.

As shown in (4.2), size of the search space grows exponentially with increment of any component, C, T, and H, thus, substantial computational load and long runtime are inevitable for exhaustive search. Table 4.1 shows it distinctly. The number of nodes used for evaluation is 5, 6, and 7, and the maximum end-to-end delays are measured in 5 different random topologies. Same sets of experiments are repeated with our GAS algorithm for comparison. Details of GAS algorithm and parameters will be described later in the following two sections.

First experiment is the frame length comparison. In all the three cases, exhaustive search performs better than GAS. However, the difference is not that much significant. The biggest difference is 1 slot when the number of nodes is 6 and differences of the other two cases are less than 1, 0.4 when the number of nodes is 5 and 0.8 when the number of nodes is 7. Second experiment is the maximum end-to-end delay comparison. Although, exhaustive search performs better than GAS in all cases, the biggest difference is just 0.76 slots. However, while frame length and the maximum end-to-end delay have very small differences, runtimes of the two

TABLE 4.1

Performance Comparison: Exhaustive Search vs. GAS (C=3)

No. of Nodes	Max Delay (slot)		Frame Length (slot)		Runtime (min)	(Sec)
	EXH	GAS	EXH	GAS	EXH	GAS
5	3.8	4.2	5.8	6.2	8.2	0.64
6	4	4.61	6.4	7.4	16.09	1.2
7	4.25	5.01	8	8.8	34.97	1.71

The results are average of 5 different topologies using
Intel Core2 Duo 1.8 GHz processors with 2GB RAM

approaches are significantly different. The maximum difference is more than 1200 times when the number of nodes is 7. Although exhaustive searches find little better solutions than GAS, the runtimes are excessively long even with small number of nodes. Therefore, exhaustive search is not a practical approach for the optimization problem and this is the reason that we use a metaheuristic method.

4.3 GAS: GA-Assisted Scheduling Algorithm

4.3.1 Flow-Based Scheduling

GAS adopts flow-based scheduling scheme. In this scheme, not only a source node but the relay nodes to a coordinator, namely flow, are scheduled at once. Scheduling inside of a flow is done by staggered fashion like the staggered allocation of EGTS in the chapter 3 to minimize hop delay in data forwarding. Flow ID is the same as the node ID. For example, scheduling of flow 5 in Fig.4.1 is to assign best available time slots that minimize hop delay from node 5 and node 2.

However, as mentioned before, considering the relationships such as connectivity and interference with neighbor nodes, effectiveness of the staggered allocation relies on the schedules of neighbors, of course, its schedule affects to other neighbors as well. Depending on the schedules of neighbor flows, data residing time at the forwarder nodes varies, affecting final end-to-end delay. Therefore, in order to satisfy our objective function, minimizing the maximum end-to-end-delay, we suggest optimization of scheduling order of each flow (SF) in a network. Flow-level comparison, instead of node level is also advantages in reducing runtime. Fig. 4.1 confirms how SF affects to end-to-end delay. The SF applied to the schedule of Fig.4.1 (a) is SF_A (4 3 6 5 2) and SF applied to Fig.4.1 (b) is SF_B (2 5 6 3 4). SF_B results in just half of the maximum end-to-end delay comparing to that of SF_A. The numbers in parenthesis represent the source node ID which is same as flow ID for example scheduling order of SF_A is (node **4**→node 1, node **3**→node 1, node **6**→node 4→node 1, node **5**→node 2→node 1, node **2**→node 1). Scheduling from the each source node to the coordinator is completed at once.

4.3.2 Pre-Channel Assignment

The chromosomes in the creatures consist of the strings of the genes and they indicate the characteristics of the creature. Likewise, GA also have chromosomes that consist of artificially encoded multiple genes and they are one of the elements in an entire search space. In GAS algorithm, an individual chromosome represents a candidate SF and an individual gene represents a candidate combination of a node and a channel.

In GAS, a i^{th} chromosome $Pi = [g_1, g_2, g_3, \ldots g_{H-1}, g_H]$ is consist of H integer genes. The reason that the number of genes is H instead of N is scheduling should be done based on the total number of the transmission not total number of the nodes. As mentioned earlier, the H of the example in Fig. 4.1 is 7 because node 2 and node 4 transmit not only their own data but forward data from their child node 5 and node 6. Therefore, H genes in a chromosome reflect node and channel pair for entire transmission in a network.

The gene g_h is encoded by $g_h = 10^{d} \cdot n_i + c_j$ where n_i is the ID of node i and d is the highest digit of channel C_j. The range of i and j are [1, H] and [1, C] respectively. For example, if channel 15 is assigned to node 3, the corresponding gene will be an integer number 315. The series of n_i in each chromosome is obtained by random permutation of the nodes and channel c_j is also chosen randomly from the given C channels. The order of the series n_i also represents SF thus, time slot allocation for each flow follows the order and the channels that are already encoded in genes are assigned to each node in a flow. Therefore, we can reduce the effort for available channel scanning, interference avoidance, and negotiation with parent or destination node.

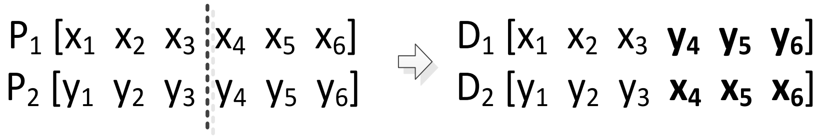
One may say the randomness involved in SF determination and channel assignment is not an efficient method because well designed chromosomes having high fitness can make evolution process fast, resulting in a better solution. But the problem in it is biased and pre-matured evolution. The diversity in genes provides more chances to have various chromosomes in other words, wider explore in search space is possible. It may add overhead to reach ultimate level of evolution but there are more chances to try various gene combinations, leading more efficient solution searches. The creatures in an isolated area for example Galapagos Island, lack diversity in genes thus, they have less chance to meet creatures with different types of genes, retarding or stopping evolution process. Not only that, their evolution is biased only to adapt to the environment they are locating where isolated from the outer world. Although the chromosome setup looks like chaotic because of the randomly permuted SFs and randomly selected channels, it provides diverse bloods and there is a high probability of appearing the offspring having more improved genes than their parent generation.

In conclusion, randomly permuted SF and randomly selected channel provide more diversity than intentionally manipulated ones and result in improved solution search, and pre-channel assignment significantly lowers computational overhead for channel assignment, reducing runtime considerably. Details on the chromosome and gene encoding will be explained in the following section and the quantity investigation will be shown in performance evaluation.

4.3.3 Encoding

The first part of the GA operation is the encoding of candidate solutions. A candidate solution, which is called a chromosome, represents one of the possible solutions in a search space. A chromosome consists of several genes. In our case, genes are combinations of two integer numbers, indicating a node and a channel assigned to it. Putting all the individual genes together, a candidate solution, representing an SF and node-channel match-ups are formed. As mentioned in the previous section, the encoding is a random process and they are done without any assumption or given condition.

First step of the encoding is the formation of a node list (NL) and a channel list (CL). List all nodes' IDs where each ID is an integer number, in ascending order as many times as the number of its required transmissions which is the same as the H in (4.1). Hence, the total number of elements in an NL can be re-written as H = $\sum_{L=1}^{L_{max}}(L \cdot N_L)$ where L is the level of a tree and N_L is the number of nodes at each level. Note the level L also represents the number of hops (or required number of transmissions) from a source node at level L to the coordinator on top of the tree. Thus, H also represents the upper bound of required time slots within a TDMA cycle. Once



(a) Crossover at crossover point 4

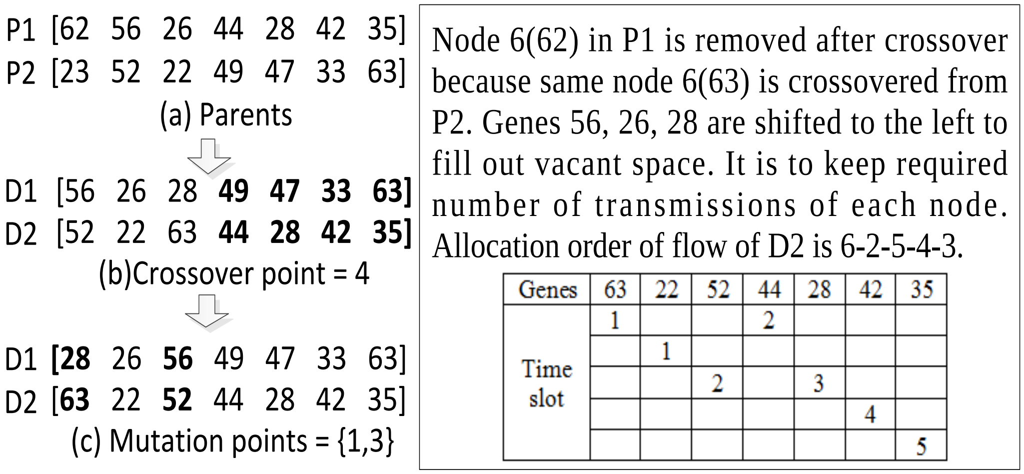
(b) Mutation at mutation points {1,3}

Fig. 4.2 Pairing: Crossover and Mutation

node list NL = $\{N_1, N_2, \dots N_{H-1}, N_H\}$ is prepared, next is the formation of a channel list, CL = $\{C_1, C_2, \dots C_{j-1}-1, C_j\}$ where C_j is a channel number represented by an integer and j is the given number of channels for example 16 in IEEE802.15.4. Last step is the random permutation for SF and random channel selection as explained in previous section. The size of the initial population pool is configurable depending on the complexity of a given problem.

4.3.4 Paring

Paring is a strategy to probe more search space and evolve toward the optimal solution by recombining genes of parent chromosomes, namely crossover and mutation. Crossover interchanges arbitrarily selected portions of genes of parents and produces two new descendants. Fig. 4.2 (a) describes single point crossover operation with crossover point = 4. The genes from the 4th to the 7th of the both parents P1 and P2 are interchanged and produce descendants D1 and D2 which have new order and combinations of genes. In GAS, it represents a new SF and new nodes and channels match-ups. As GA operations are iterated for a long period of generations, the chromosomes in a population pool tend to be alike each other because of the repeated genes interchange with a limited size of population pool, making evolution speed slower and stopping



P1 [62 56 26 44 28 42 35]
P2 [23 52 22 49 47 33 63]
(a) Parents

D1 [56 26 28 **49 47 33 63]**
D2 [52 22 63 **44 28 42 35]**
(b)Crossover point = 4

D1 **[28** 26 **56** 49 47 33 63]
D2 **[63** 22 **52** 44 28 42 35]
(c) Mutation points = {1,3}

Node 6(62) in P1 is removed after crossover because same node 6(63) is crossovered from P2. Genes 56, 26, 28 are shifted to the left to fill out vacant space. It is to keep required number of transmissions of each node. Allocation order of flow of D2 is 6-2-5-4-3.

Genes	63	22	52	44	28	42	35
Time slot	1			2			
		1					
			2		3		
						4	
							5

Fig. 4.3 Example of paring of the network in Fig.4.1

evolution before reaching the optimal point. Mutation is to prevent this pre-matured completion, the local minima in the optimization problem.

An example of two points swapping mutation is shown in Fig. 4.2 (b). As shown in Fig. 4.2 (b), the two genes of randomly selected mutation point = {1, 3} in P1 are swapped and produce mutated descendant D1 that has different gene order from the parent. When evolutionary process stalls for a long time, we can assume that the process falls in local minima. Mutation helps to escape from local minima by jumping to a different point of a search space and provides diversity.

In GAS, mutation provides new SF by swapping the positions of two genes, we can produce very new SF and it may impact much on the maximum end-to-end delay because of the reciprocal relationships with other neighboring flows' schedule. The rates of crossover and mutation are controlled by configurable probabilities. Comparing to crossover, probability of mutation is very low because high rate of mutation obstructs convergence toward the optimal solution. Fig. 4.3 is an example of pairing. Since the fixed number of appearance of each node in a chromosome which is equivalent to the number of transmissions, redundant nodes caused by crossover are removed.

4.3.5 Sifting of the First Generation

The quality of the optimal solution at the end of evolutionary process highly relies on the quality, fitness in GA, of the first generation. Therefore, fitness proportionate selection such as elitism or tournament is the popular approach for the quality of heredity but as mentioned before, it limits diversity of genes and may lead prejudiced evolution. In order to satisfy both diversity and quality of heredity, GAS sifts chromosomes with high fitness. The procedure is as follows. First, generate candidate chromosomes more than the size of a population pool and evaluate their fitness. Then, GAS sifts only top rankers as many as the size of a population pool from the candidates. The sifting occurs before GAS operations and it adds little more time to the total runtime. However, sifting provides quality of heredity and avoids prejudiced evolution. The performance evaluations confirm it clearly.

4.3.6 Selection

Selection and pairing in the next section work just like a breeding of living things. Two randomly selected chromosomes from the initial population pool take roles as parents and produce two new descendants which may have higher fitness than their parents. In general, random selection is not as fast as elitism or tournament, a strategy for selecting superior ones to others. However, the chosen candidates, superior-to-others but not-yet-optimal may take hold off following generations to complete optimization pre-maturely due to the lack of diversity. Therefore, like the other operations, we prefer randomness in parent selection. The random selection can explore search space wider and have high probability to find the optimal solution and helps to avoid pre-matured termination of the algorithm.

4.3.7 Adaptive Mutation

Mutation provides diversity when the algorithm stalls in local minima for a long time but too frequent mutation may lead to loss of good solution and of convergence to it. In GAS, the probability of mutation P_{mut} increases additively by α_{mut} whenever the algorithm stalls, providing higher chance to escape from local minima. Once mutation is successful, the probability of mutation is reset to its initial value.

4.3.8 Fitness evaluation

The fitness is a measure of the maximum end-to-end delay of a candidate solution. Since SF and nodes and channels match-ups are already craved in genes and chromosomes, only consideration is to allocate appropriate time slots that minimize end-to-end delay. The slot allocation procedure is explained in the pseudo-code of the Fig. 4.4.

End-to-end delay of a flow f_i is expressed as $DL(f_i)=\sum DL(N_k)$ where $DL(N_k)$ accounts the delay at node k belonging to the flow i and i is the flow ID or source node ID that attempting to transmit data to the coordinator via forwarder nodes on the path. The fitness function for a candidate solution of flow i, is given in (4.4).

$$Fitness\,(DL_i) = \frac{1}{\max(DL(f_i))} \tag{4.4}$$

In accordance the rule, ‘The survival of the fittest’, the one who scores the highest fitness, in other words, having the minimum of the maximum end-to-end delay among the parents and the descendants survives as a successor for the next generation.

```
While $t \leq$ upper bound of time slot;  (initial t=1)
  if time slot $t$ is already allocated to a node(s) $k$
    if $s^p \neq \forall k^p$ ;*check if parents of $s$ and $k$ are same or not
    *check if $s$ and $k$ interrupt each other
      if $C(s, k^p) + C(s^p, k) = 0, \forall k$ when $ch_s = ch_k$
        allocate $t$ to node $s$, break while loop
      elseif $(s \neq k^p \ \Lambda \ s^p \neq k), \ \forall k$ when $ch_s \neq ch_k$
        allocate $t$ to node $s$, break while loop
      end if
    end if
  else
    allocate $t$ to node $s$, break while loop
  end if
  $t = t + 1$
end while
Note:  $ch_k$: channel assigned to node k,   $k^p$: parent of node k,
     s: current node,   k: neighbor(s) of node s
     C: connectivity matrix
     $C(s, k) = 1$  if there exist a link between node s and k
            0 otherwise
```

Fig. 4.4 Pseudo-code for slot allocation

4.4 Performance Evaluation

We first compare 5 different versions of GAS to inspect how the components of GAS affect to its performance. The parameters used for the performance evaluation are listed in Table 4.2 and performance comparisons are shown in Table 4.3. The reason that we take only 3 channels out of 16 available in IEEE802.15.4 is only 3 of them are not-overlapped with WiFi channels. From Fig. 4.5 to Fig. 4.8 are topologies used in the simulation. Solid lines are the paths from each node to the coordinator node 1 located at the center and dotted lines represent connectivity among the nodes.

TABLE 4.2
EVALUATION PARAMETERS FOR GAS

Parameters	Values
Size of the population pool	0.25N
Number of candidates for Sifting	0.5N
Method of crossover	Single point
Probability of crossover (P_{xover})	0.7
Method of mutation	Two points swapping
Probability of mutation (P_{mut})	0.3
Probability of mutation increment (α_{mut})	0.02
Stall duration limit	35 iteration
Max. iteration	100
Number of Channels	3

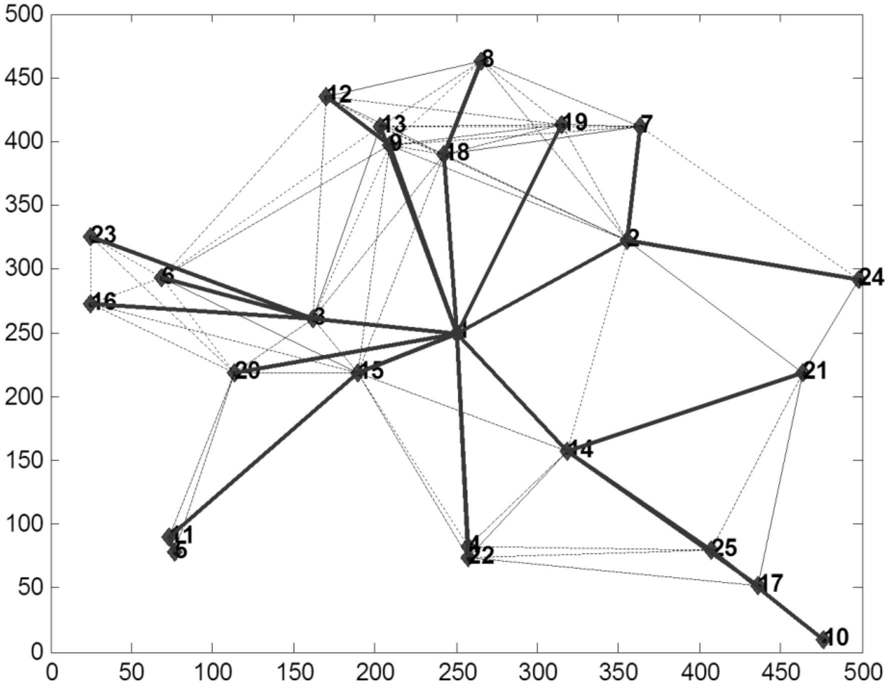


Fig. 4.5 Example topology when N=25

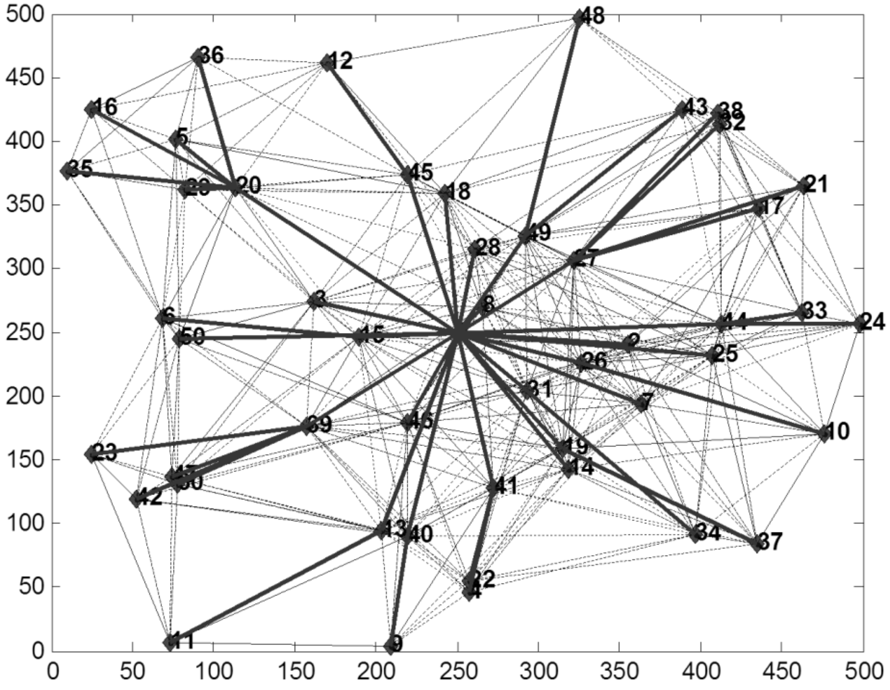


Fig. 4.6 Example topology when N=50

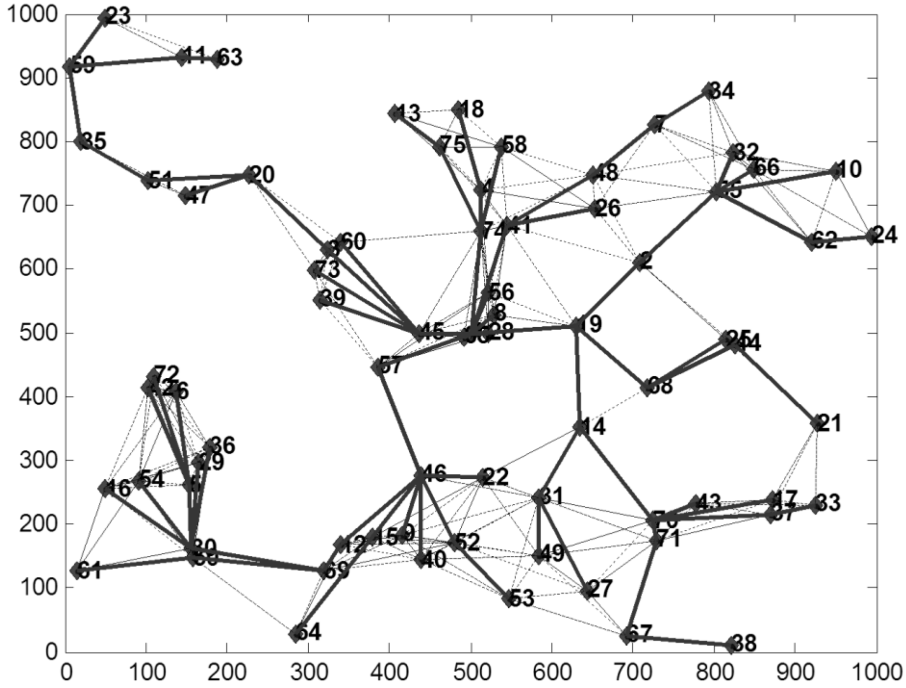


Fig. 4.7 Example topology when N=75

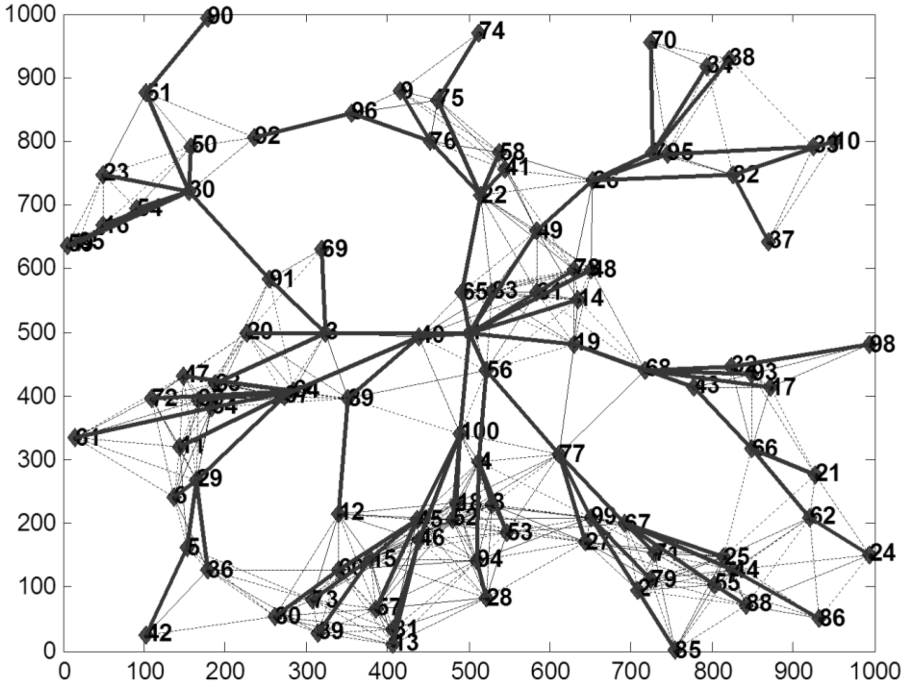


Fig. 4.8 Example topology when N=100

TABLE 4.3
PERFORMANCE COMPARISONS OF DIFFERENT VERSIONS OF GAS

NO. OF NODES	RUN TIME (SEC)			
	GAS	GAS-C	GAS-CM	GAS-CMS
25	185.8	47.8	46.5	49.6
50	832.3	342.6	342.4	344.0
75	1986.9	1299.4	1289.9	1327.1
100	5618.8	4476.8	4456.9	4909.1

NO. OF NODES	MAX. END-TO-END DELAY (SLOT)			
	GAS	GAS-C	GAS-CM	GAS-CMS
25	4.5	4.8	4.2	4.1
50	11.9	10.9	9.5	7.9
75	27.9	23.5	21.5	20.8
100	67.4	68.0	66.9	56.7

4.4.1 Comparisons of GASs

Table 4.3 lists comparisons of the runtime and the maximum en-to-end delay of 5 different versions of GAS: GAS, GAS-C, GAS-CM, and GAS-CMS. GAS is the default version that has only flow-based allocation scheme and it assigns an available channel to each node while time slot allocation. GAS-C has additional pre-channel assignment scheme, GAS-CM adds adaptive mutation to GAS-C, and finally GAS-CMS adds sifting to GAS-CM. The measurements are average of 30 sets of simulations.

Aforementioned, pre-channel assignment decreases computational load, leading substantially reduced runtime. GAC-C does not show improvement in end-to-end delay measurement. In some cases, they are worse than GAS, though it is small. It is because of the random channel selection and it may be disadvantage in staggered time slots allocation. However, the difference is very small. Another measurement is runtime. Comparing to GAS which assigns channel while time slot allocation, GAS-C reduces runtime considerably up to 74.3% when the number of nodes is 25. The light computational load derived from pre-channel assignment helps to reduce runtime greatly.

GAS-CM which applies adaptive mutation does not make distinctive improvement in runtime but it attains lower maximum end-to-end-delay than GAS and GAS-C. The additively increased mutation rate whenever evolution stalls provides more diversity and raises probability to escape from local minima. The runtime of GAS-CMS is little longer than GAS-C and GAS-CM because of the sifting process before evolutionary process but is still far lower than GAS. The advantage of the GAS-CMS is shown in end-to-end delay measurement. Though it takes little more runtime, end-to-end delay measurement excels all the others. It is distinct as number of the nodes is increase.

4.4.2 End-to-end delay comparisons of GAS, Level and Random Allocation

This time end-to-end delay performance of GAS-CMS is compared with two other SF methods: level-based and random-based SF. Same staggered allocation inside of the flow is applied for both level-based and random-based SF but the difference is how to determine SF. The rationale of level-based ordering is simple. The longer the hop distances from a source node to the coordinator node, the higher the probability to reside at the forwarder nodes because of the already completed neighbors' schedules. Therefore, in level-based allocation, SF is determined by the order of hop-distance from a source node to the coordinator. The furthest node from the coordinator is allocated first and the closest one is allocated last. If tied, the order is determined

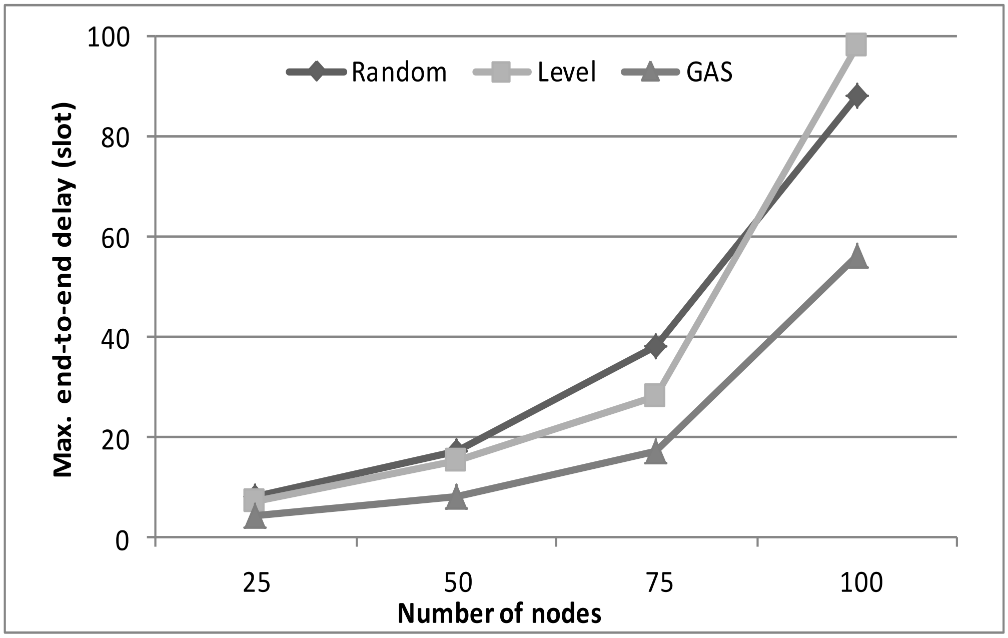


Fig. 4.9 End-to-end delay comparisons of GAS, Level, and Random allocation

arbitrarily. Random-based scheme determines SF randomly and is compatible with the first-come, first-served policy of EGTS or DSME mode of IEEE802.15.4e.

Fig. 4.9 shows end-to-end delay comparison in the four different sizes of networks. GAS is superior to the others in all topologies and the delay difference becomes greater as the number of nodes increases. The ability to evolve using crossover and mutation iteratively leads GAS to find the schedule with lower end-to-end delay. An interesting thing is that level-based SF. The simplest scheme shows almost similar to or better performance than the random-based SF up to 75 nodes. It shows that the simple level-based SF works nicely if network size is small. Then we can save computational overhead, runtime, and energy.

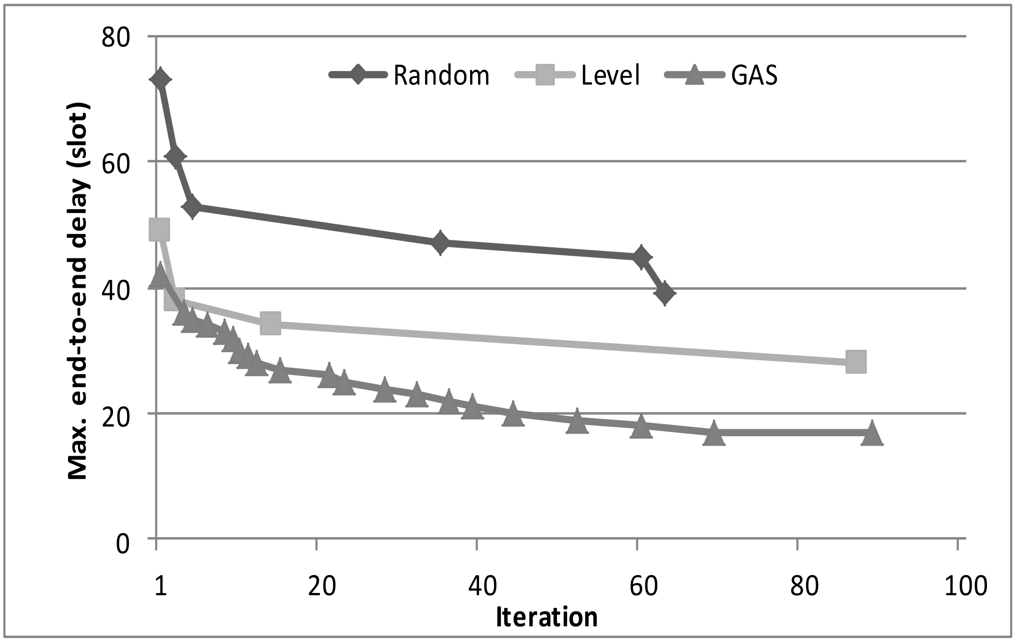


Fig. 4.10 End-to-end delay improvement trace

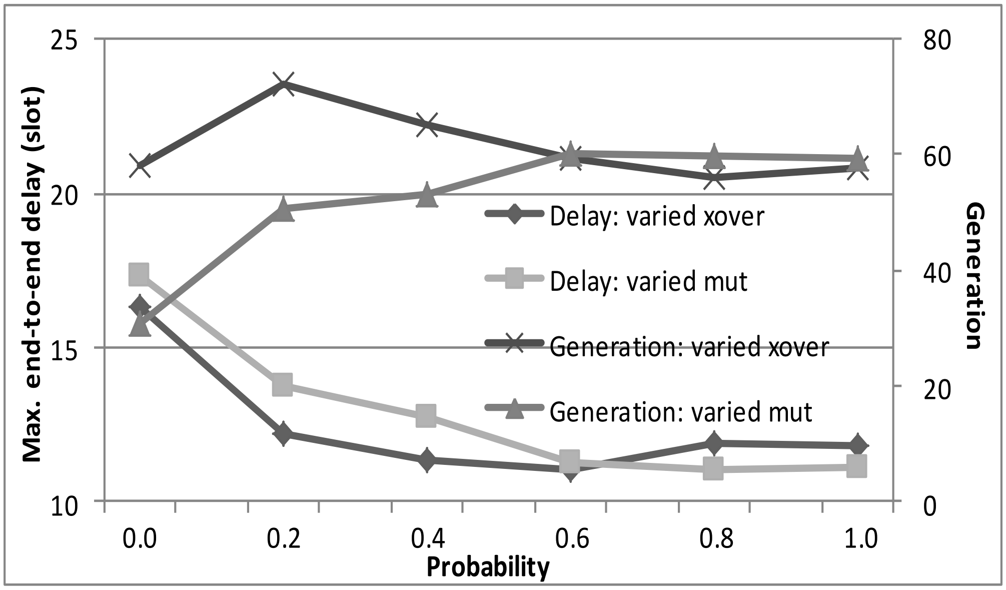


Fig. 4.11 Impact of probabilities of crossover and mutation

The trace of solution improvement in 100 iterations when the number of nodes is 75 is observed in Fig. 4.10 which affirms the GA's capability of the optimal solution search. The dots in the graphs are the points where new improved solutions are found. In case of random-based SF, it improves only 5 times and the improvement stops at around 65th iteration. This optimal scheduling problem has huge number of candidate combinations and capability of random-based SF is not enough to handle them. In level-based SF, the number of improvement is just 3 but better than random-based SF due to the inherent capability to lessen data residing time at the forwarder nodes on the long path. On the other hand, GAS finds improved solutions more than 20 times steadily and the results outperform the others in all iteration ranges. The evolution rate becomes slower as iteration is repeated because as the evolutionary operations are repeated chromosomes in a population pool tend to resemble each other. Ultimately, all of them turn to same, indicating the end of the evolution.

Fig. 4.11 shows how probabilities of crossover and mutation influence the length of iteration and end-to-end delay of GAS. The number of nodes in a network is 50. While the probability of crossover is 0.5, the probability of mutation varies from 0 to 1, and vice versa. The general trend is that when the probabilities of both crossover and mutation are low, evolution ends early with very high end-to-end delay. However, as the probabilities increase, number of iterations is also increases but end-to-end delays decrease. It is intuitive because frequent exchanges of the genes diversify descendants and more improved solutions appear as generations pass. In this network setup, the optimality is shown around probability 0.5 to 0.8. When crossover: 0.5, mutation: 0.8, the end-to-end delay is the lowest, and when crossover: 0.8, mutation: 0.5, the runtime is the shortest. It implies that the probabilities of crossover and mutation should be configured according to the problems.

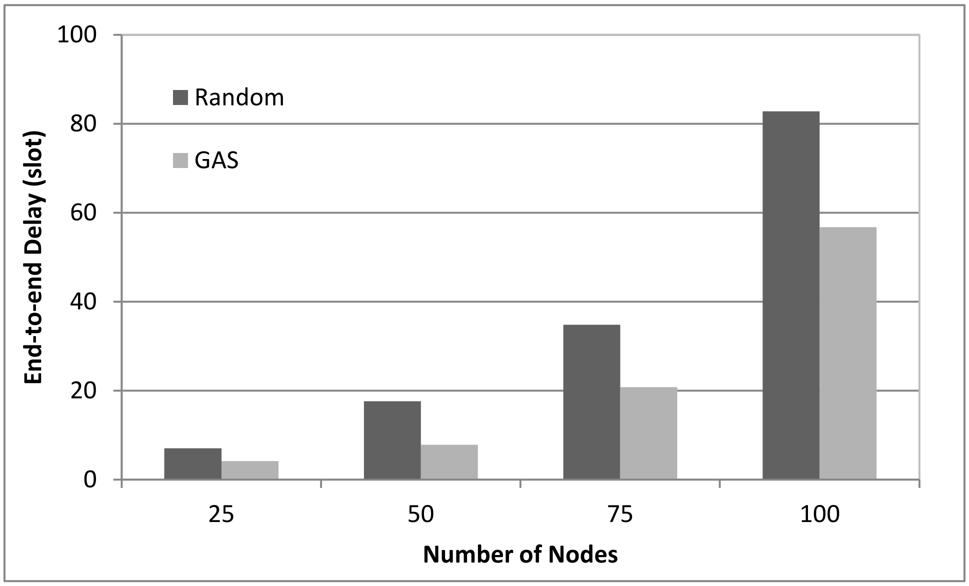


Fig.4.12 End-to-end delay comparison between random-based SF and GAS for the same time duration

Previous end-to-end delays in Fig. 4.9 were measured for the duration of the same 100 iterations but the time duration of an iteration of each scheme is different. In order to consider fairness, we simulated GAS and random-based SF for the same amount of time period. The runtime of GAS-CMS in Table 4.2 are applied for random-based SF. As the results of Fig. 4.9, the end-to-end delays of GAS are far less than that of random-based SF regardless of time and iterations.

4.4.3 End-to-end Comparison with Simulated Annealing

Simulated annealing (SA) is a probabilistic metaheuristic proposed in [60] for finding the global minimum of a cost function that may possess several local minima. It works by emulating the physical process whereby a solid is slowly cooled so that when eventually its structure is frozen, this happens at a minimum energy configuration [69]. Simulated annealing tries to avoid

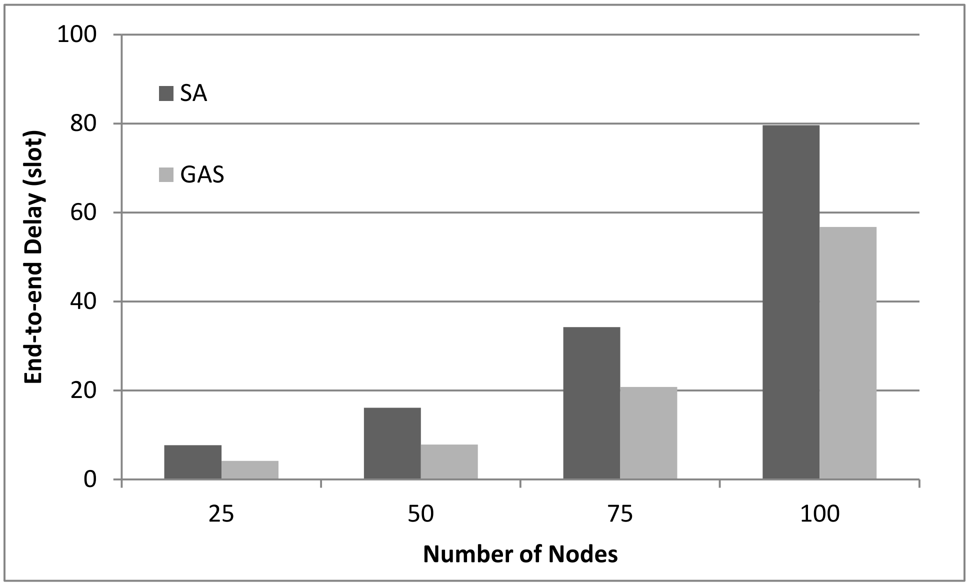


Fig.4.13 End-to-end delay comparison between simulated annealing and GAS for the same time duration

local minima by jumping out of them early in the computation. Toward the end of the computation, when the temperature, or probability of accepting a worse solution, is nearly zero, this simply seeks the bottom of the local minimum. The chance of getting a good solution can be traded off with computation time by slowing down the cooling schedule. The slower the cooling, the higher the chance of finding the optimum solution, but it makes the runtime longer. Thus, effective use of this technique depends on finding a cooling schedule that gets good enough solutions without taking too much time [70]. SA's major advantage over other metaheuristic methods is an ability to avoid local minima. The algorithm accepts not only changes that decrease objective function, but changes that increase it with a probability of (4.5).

$$P = Exp\left(-\frac{\Delta}{T}\right) \tag{4.5}$$

where Δ is the change in the value of the objective function and T is the temperature. In this simulation we set initial temperature T=100 and cool rate=0.8. In order for fair comparison, time durations of GAS-CMS in Table 4.2 are applied to SA. As shown in Fig. 4.13, the results of SA are slightly better than random-based SF in Fig. 4.12 but still worse than GAS. SA searches only one point of the search space while GAS searches and evaluates various combinations of genes using crossover and mutation operations. Although same period of time is given, the ability to search improved solutions of GA surpasses that of SA.

4.5 Conclusion

In this chapter, we proposed GAS, genetic algorithm-assisted transmission scheduling for multi-hop, multi-channel networks. GAS, as a scheduling algorithm for EGTS or for other multi-hop, multi-channel wireless networks, enhances QoS in terms of end-to-end delay by exploiting evolutionary process of genetic algorithm.

Fortified GA operations including sifting and adaptive mutation as well as QoS features such as node-channel pair encoding in genes, flow-based scheduling order, pre-channel assignment search improved optimal solutions fast and efficiently.

Through the experiments of the five different versions GAS, we affirmed the impact of the each enhancement. Furthermore, solution searching capability of GAS was observed by the end-to-end delay comparison with random-based and level-based SF. In addition, performance comparison with SA also confirmed GAS is an effective method for the optimal scheduling of for multi-channel, multi-hop wireless sensor networks.

Chapter 5

Conclusion and Future Work

This dissertation presented two real-time MAC protocols, PNP-MAC for wireless body area networks and EGTS for wireless industrial wireless sensor networks, and GAS, an optimization algorithm that minimizes the maximum end-to-end delay in multi-hop, multi-channel wireless sensor networks. Although the three topics address different contents, they share a common fundamental principle: to support high level of QoS in terms of end-to-end delay.

BAN is a small range network that has a wide range of real-time applications. Various levels of QoS from entertainment applications for example interactive game and video goggle to medical applications such as health management and physiological phenomenon monitoring are the challenges of BAN. Moreover, medical applications’ time-sensitivity that requires the highest level of reliability makes BAN very fastidious about QoS. PNP-MAC is provided to meet those challenges. Its versatile superframe offers CAP period, TDMA period, and emergency alarm exclusive ETS period; hence, PNP-MAC can manipulate diverse types of traffics and support QoS for them. Moreover, its prioritized services and preemptive nature support the required QoS satisfactorily. Consecutively arranged advertisement, CAP, and beacon realize fast allocation and adoption to changes in slot allocation. Fast allocation substantially reduces initial delay and engages DTS allocation promptly. Another critical challenge in BAN is low energy consumption. UWB is emerging as a promising radio solution for low power transmission. However, the low power transmission makes clear channel assessment (CCA) infeasible thus, PNP-MAC provides such reliable solutions as BG and BR that do not depend on CCA for emergency alarm transmission via UWB radios. Desirable performance of PNP-MAC meeting given QoS is demonstrated by OPNET network simulator and test-bed experiments using MicaZ motes in realistic scenarios.

EGTS enhanced the GTS mechanism of IEEE 802.15.4 to satisfy QoS requirements of the large scale industrial wireless sensor networks which are different from those of BAN. While PNP-MAC serves 1-hop, 5m range netowrks, EGTS serves for multi-hop long range industrial networks via multiple forwarder nodes, namely multi-hop extention. Multi-hop extention realizes data delivery among the nodes located multiple hops away which is not supported by IEEE802.15.4. Schduling is done by exchanging ABT table of each node using the three-way-handshking, EGTS reques/EGTS reply/EGTS notify. EGTS can choose one of the distributed scheduling methode, random allocation or staggered allocation while the three-way-handshaking.

In order to support guaranteed QoS, EGTS adopts not only TDMA but also multiple channels over a TDMA time slot, namely multi-channel extention. Multi-channel extention allows multiple simulaneous transmissions by adopting multiple channels in a single EGTS slot, leading to substantially increased throughput and expedited simultaneous data delivery. It is very appropriate for industrial applications that has large data traffic requiring deterministic and guaranteed QoS. Multi-superframe extention overcomes time slot limitation of GTS in IEEE802.15.4. 7 GTS slots of IEEE805.15.4 is not enough for a large scale of industrial networks which mainly deal with periodic time-sensitive data. Multi-superframe extension reiterates superframes with the configurable parameter MO. It flexibly controls the number of EGTS slots based on the demands from the member nodes. CAP reduction is also helpful scheme when the traffic is very heavy and the number of EGTS in multi-superframe is not enough. Except the first CAP, each of all the other CAP periods can transform into 8 additional EGTS slots and provides more room for time-critical data delivery. The non-time sensitive data, or best effort data still can use the CAP period in the first superframe.

In chapter 3, we proposed GAS, a genetic algorithm-assisted transmission scheduling for multi-hop, multi-channel networks. The main purpose of GAS is for the optimal scheduling that minimizes the maximum end-to-end delay in multi-hop networks for example wireless industrial sensor networks. TDMA access scheme in EGTS provides deterministic end-to-end delay but the delay may not be optimal since EGTS scheduling considers only available slots and channels at a certain time of a certain node; thus, no optimality is considered. However, the scheduling for multi-hop networks is a combinatorial NP-complete problem; in other words, no specific solutions are given for it. Hence, we take a metaheuristic approach that provides fast and efficient optimization tool for complex combinatorial problems. GAS is an optimization algorithm that takes advantages of GA, one of the most popular evolutionary metaheuristics. We improved the traditional GA by exploiting sifting and adaptive mutation. We also added supportive features for GAS such as flow-based scheduling order, pre-channel assignment, enabling faster searches for the optimal solution. Therefore, EGTS, if integrated with GAS, can produce more efficient schedules for time-sensitive real-time data. Extensive experiments affirmed the GAS performs better than first-come, first-served EGTS scheduling in terms of end-to-end delay. Furthermore, performance comparison with simulated annealing also confirmed GAS is an effective scheme for scheduling of time slot and channel assignment for multi-channel, multi-hop wireless sensor networks.

GAS is a centralized scheduling scheme. A coordinator takes charge in scheduling of an entire network which inevitably consumes non-negligible time even though it is only one time event at the initialization stage of a network. Since GAS is designed to combine with EGTS, currently it works for static tree-based networks. Therefore, as a future work, we plan to design a distributed version of GAS which takes less time and works not only in tree-based networks but in mobile mesh networks. In future works, other metaheuristic such as particle swarm optimization (PSO) and harmony search (HS) will be investigated.

References

[1] Bin Zhen, Maulin Patel, SungHyup Lee, EunTae Won, Arthur Astrin, IEEE TG6 15-08-0644-09-0006-tg6-technical-requirements-document.

[2] June S. Yoon, Gahng S. Ahn, Seong-Soon Joo, Myung J. Lee, IEEE TG6 15-09-0336-01-0006-ieee802-15-6-versatile-mac-proposal-for-wban.

[3] Kamya Yekeh Yazdandoost, Kamran Sayrafian-Pour, IEEE TG6 15-08-0780-09-0006-tg6-channel-model.

[4] Gang Zhou, Jian Lu, Chieh-Yih Wan, Mark D. Yarvis, and John A. Stankovic, “BodyQoS-Adaptive and Radio-Agnostic QoS for Body Sensor Networks”, IEEE INFOCOM 2008, April 2008, pp565-573.

[5] Nicholas F. Timmons, William G. Scanlon “Analysis of the Performance of IEEE 802.15.4 for Medical Sensor Body Area Networking”, Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004.

[6] M. Yuce, C.K Ho, “Implementation of body area networks based on MICS/WMTS medical bands for healthcare systems”, IEEE International Conference of the Engineering in Medicine and Biology Society, 2008.

[7] Emil Jovanov, Aleksandar Milenkovic, Chris Otto, Piet C de Groen, "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation", Journal of Neuro Engineering and Rehabilitation May 2005.

[8] Kenichi Takizawa, Huan-Bang Li, Kiyoshi Hamaguchi, and Ryuji Kohno, "Wireless Patient Monitoring using IEEE802.15.4a WPAN", IEEE Ultra-Wideband, 2007. ICUWB 2007, Sept. 2007, pp: 235-240.

[9] Okundu Omeni, Alan Chi Wai Wong, Alison J. Burdett, and Christofer Toumazou, “Energy Efficient Medium Access Protocol for Wireless Medical Body Area Sensor Networks”, IEEE Transactions on Biomedical Circuits and Systems, Vol. 2, No. 4, December 2008.

[10] Bin Zhen, Huan-Bang Li, and Ryuji Kohno, “IEEE Body Area Networks for Medical Applications”, IEEE ISWCS 2007.

[11] M. A. Ameen, Ahsanun Nessa, Kyung Sup Kwak, "QoS Issues with Focus on Wireless Body Area Networks", Third International Conference on Convergence and Hybrid Information Technology, 2008.

[12] David Malan, Thaddeus Fulford-Jones, Matt Welsh, Steve Moulton, “CodeBlue: An Ad Hoc Sensor Network Infrastructure for Emergency Medical Care", International Workshop on Wearable and Implantable Body Sensor Networks 2004.

[13] NIST, http://w3.antd.nist.gov/ban/

[14] IEEE 802.15 TG6 http://www.ieee802.org/15/pub/TG6.html

[15] Chengshi Zhao, Zheng Zhou, and Kyungsup Kwak, **“**Novel UWB Transceiver for WBAN Networks: A Study on AWGN Channels”**,** ETRI Journal, vol.32, no.1, Feb. 2010, pp.11-21.

[16] P. Kumar, M. Güne¸s, A. Basset A. Mamou, J.Schiller, "Real-time Bandwidth and Energy Efficient IEEE 802.15.4 for Medical Applications", 7. GI/ITG KuVS Fachgespräch "Drahtlose Sensornetze", Sep. 2008.

[17] http://www.fitbit.com/

[18] http://www.polarusa.com

[19] http://www.apple.com/ipod/nike/

[20] Sana Ullah, Pervez Khan, Niamat Ullah, Shahnaz Sallem, Henry Higgins, Kyung Sup Kwak, “ A Review of Wireless Body Area Networks for Medical Applications”, Communications, Network and System Sciences, 2009.

[21] Bart Braem, Benoıt Latre, Ingrid Moerman, Chris Blondia, Piet Demeester, “The Wireless Autonomous Spanning tree Protocol for Multihop Wireless Body Area Networks”, International Conference on Mobile and Ubiquitous Systems 2006.

[22] S. Marinkovic, C.Spagnol and E. Popovici, “Energy-Efficient TDMA-based MAC Protocol for Wireless Body Area Networks”, International Conference on Sensor Technologies and Applications 2009.

[23] Begonya Otal, Luis Alonso, Christos Verikoukis, “Highly Reliable Energy-Saving MAC for Wireless Body Sensor Networks in Healthcare Systems”, IEEE Journal on selected area in communications Vol. 27, No. 4, MAY 2009.

[24] June S. Yoon, Gahng-Seop Ahn, Seong-Soon Joo, Myung J. Lee, “PNP-MAC: Preemptive slot allocation and Non-Preemptive transmission for Providing QoS in Body Area Networks”, IEEE Consumer Communications and Networking Conference 2010.

[25] June S. Yoon, Gahng-Seop Ahn, Myung J. Lee, “Recent MAC standardization activities in IEEE 802.15 working group”, Journal of standards and technology review, Dec. 2010

[26] Opnet, http://www.opnet.com

[27] Xbow, http://www.xbow.com

[28] International Commission on Non-Ionizing Radiation Protection (ICNIRP).Statement on the "Guidelines for limiting exposure to time-varying electric, magnetic and electromagetic fields (up to 300 GHz)", 2009.

[29] IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz, IEEE Std C95.1, 2005.

[30] http://www.wirelesspowerconsortium.com/

[31] Wireless medium access controls (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs) IEEE Standard 802.15.4-2003.

[32] Wireless medium access controls (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs) IEEE Standard 802.15.4b-2006.

[33] Wireless medium access controls (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs) IEEE Standard 802.15.4a Alternate PHYs-2007

[34] G. Ahn, T. Park, Y. Kim, J. Yoon, R. Zhang, M. Lee, C. Shin, S. Joo, J. Chae, IEEE TG4e 15-08-0775-01-004e-distributed-multi-channel-mesh-extension.

[35] Doherty, L.; Lindsay, W.; Simon, J. “Channel-Specific Wireless Sensor Network Path Data”, ICCCN 2007. Proceedings of 16th International Conference Aug. 2007.

[36] S.-L. Wu, C.-Y. Lin, Y.-C. Tseng, and J.-P. Sheu, "A new multichannel MAC protocol with on-demand channel assignment for multi-hop mobile ad hoc networks", Parallel Architectures, Algorithms and Networks, 2000(l-SPAN 2000), July 2000, pp.232-237.

[37] Zhou, G. Huang, C. Yan, T. He, T. Stankovic, J. A. Abdelzaher, T. F. “MMSN: Multi-Frequency Media Access Control for Wireless Sensor Networks”, INFOCOM 2006,April 2006, pp.1-13.

[38] Yafeng Wu, Stankovic, J.A., Tian He, Shan Lin, “Realistic and Efficient Multi-Channel Communication in Wireless Sensor Networks”, INFOCOM 2008, April 2006, pp.1193 – 1201.

[39] Youngmin Kim, Hyojeong Shin, and Hojung Cha, “Y-MAC: An Energy-efficient Multi-channel MAC Protocol for Dense Wireless Sensor Networks”, IPSN, International Conference on Information Processing in Sensor Networks, 2008 pp.53 – 63.

[40] X. Chen, P. Han, Q. He, S. Tu, and Z. Chen, “A Multi-Channel MAC Protocol for Wireless Sensor Networks”, IEEE CIT, 2006.

[41] L. Q. Zhuang, K. M. Goh and J. B. Zhang “The Wireless Sensor Networks for Factory Automation - Issues and Challenges”, Emerging Technologies and Factory Automation, 2007.

[42] IEEE802.15 WPAN TG 4e, http://www.ieee802.org/15/pub/TG4e.html

[43] BACnet, http://www.bacnet.org

[44] ISA, http://www.isa.org

[45] Wireless HART, http://www.hartcomm.org

[46] Zigbee Alliance, http://www.zigbee.org

[47] A. Koubaa, M. Alves, and E. Tovar, “GTS Allocation Analysis in IEEE 802.15.4 for Real-Time Wireless Sensor Networks”, 20th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2006), Rhodes Island, Greece, April 2006.

[48] J.-Y. Le Boudec and P. Thiran, “Network Calculus: A Theory of Deterministic Queuing Systems for the Internet”, Springer, Lecture Note CS, Vol. 2050, 2001.

[49] R. L. Cruz, “A Calculus for Network Delay, Part I: Network Elements in Isolation, Part II: Network Analysis”, IEEE Transactions on Information Theory, Vol. 37, No. 1, January 1991.

[50] Gang Lu, Bhaskar Krishnamachari, Cauligi S. Raghavendra, “An Adaptive Energy-Efficient and Low-Latency MAC for Data Gathering in Wireless Sensor Networks”, Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International

[51] S.C. Ergen. P. Varaiya, “TDMA scheduling algorithm for wireless sensor networks,” The Journal of Mobile Communication, Computation and Information, Volume 16, Number 4, pp. 985-997, Nov. 2009.

[52] G. Chkrabolty, “Genetic algorithm to solve optimum TDMA transmission schedule in broadcast packet radio networks,” IEEE Transactions on communications, Volume 52, Issue 5, pp.765-777, May 2004.

[53] X. Wu, B.S. Sharif, O.R. Hinton and C.C. Tsimenidis, “Solving optimum TDMA broadcast scheduling in mobile ad hoc networks: a competent permutation genetic algorithm approach”, Communications, IEE Proceedings, Volume 152, Issue 6, pp. 780 – 788, Dec. 2005

[54] IEEE 802.15.5: Mesh Topology Capability in Wireless Personal Area Networks (WPANs), IEEE Std. 802.15.5, 2009.

[55] Cidon. I, Sidi. M, “Distributed assignment algorithms for multihop packet radio networks Computers, IEEE Transactions on Volume: 38 , Issue: 10 pp: 1353 – 1361, 1989

[56] Dimitrios J. Vergados, Maria Koutsogiannaki and Dimitrios D. Vergados, “Optimizing End-to-End TDMA Scheduling in Ad-hoc Networks on Random Topologies”, The Sixth Annual Mediterranean Ad Hoc Networking WorkShop, 2007

[57] LI Pingsheng, LI Bin, XIE Xiaoli, WANG Meng, “Improved Algorithm about the Maximum Flow in Road Network Using Auxiliary Graph Theory, Journal of Transportation Technology, Volume 9, Issue 1, February 2009

[58] Aleksandar Tsenov, “Simulated Annealing and Genetic Algorithm in Telecommunications Network Planning”, International Journal of Computational Intelligence, 2006

[59] Blum, C., Roli, A., “Metaheuristics in combinatorial optimization: Overview and conceptual comparison”, ACM Computing Surveys. pp. 268–308, 2003

[60] Kirkpatrick, S., Gelatt Jr., C.D., Vecchi, M.P., "Optimization by Simulated Annealing", Science, 1983

[61] Holland, J.H., “Adaptation in Natural and Artificial Systems”, University of Michigan Press, 1975

[62] Glover, F., "Future Paths for Integer Programming and Links to Artificial Intelligence", Computer and Operations Research, 1986

[63] Kennedy, J., Eberhart, R., "Particle Swarm Optimization". Proceedings of IEEE International Conference on Neural Networks, 1995

[64] Leonardo Badia, Alessio Botta, Luciano Lenzini, “A genetic approach to joint routing and link scheduling for wireless mesh networks”, Journal of Ad hoc Networks, 2008

[65] A. Tamilarasi and T. Anantha kumar, “An enhanced genetic algorithm with simulated annealing for job-shop scheduling”, International Journal of Engineering, Science and Technology Vol. 2, No. 1, , pp. 144-151, 2010

[66] Young-Jae Jeon, Jae-Chul Kim,, Jin-O. Kim, Joong-Rin Shin , Kwang Y. Lee, “An Efficient Simulated Annealing Algorithm for Network Reconfiguration in Large-Scale Distribution Systems”, IEEE Transactions on power delivery, Vol. 17, No. 4, 2002

[67] Chlamtac, I.; Farago, A., “Making Transmission Schedules Immune to Topology Changes in Multi-Hop Packet Radio Networks Networking”, IEEE/ACM Transactions on Volume: 2 , Issue: 1, pp: 23 – 29, 1994

[68] General Simulated Annealing Algorithm, http://www.mathworks.com

[69] Bertsimas, D. and Tsitsiklis, J., “Simulated annealing”, Statistical Science 8, 10-15, 1993

[70] NIST, “Simulated Annealing”, http://www.nist.org