A Real-Time Communication Framework for Wireless Sensor Networks

A thesis submitted to

The School of Information Technologies

The University of Sydney

In fulfilment of the requirements for the degree of

Doctor of Philosophy

Mohammed Y. Aal Salem

March 2009

DECLARATION

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

M. Aal Salem

Mar. 2009 Sydney

ABSTRACT

Recent advances in miniaturization and low power design have led to a flurry of activity in wireless sensor networks. Sensor networks have different constraints than traditional wired networks. A wireless sensor network is a special network with large numbers of nodes equipped with embedded processors, sensors, and radios. These nodes collaborate to accomplish a common task such as environment monitoring or asset tracking. In many applications, sensor nodes will be deployed in an ad-hoc fashion without careful planning. They must organize themselves to form a multi-hop, wireless communication network.

In sensor network environments, much research has been conducted in areas such as power consumption, self-organisation techniques, routing between the sensors, and the communication between the sensor and the sink. On the other hand, real-time communication with the Quality of Service (QoS) concept in wireless sensor networks is still an open research field. Most protocols either ignore real time or simply attempt to process as fast as possible and hope that this speed is sufficient to meet the deadline. However, the introduction of real-time communication has created additional challenges in this area. The sensor node spends most of its life routing packets from one node to another until the packet reaches the sink; therefore, the node functions as a small router most of the time. Since sensor networks deal with time-critical applications, it is often necessary for communication to meet real time constraints. However, research that deals with providing QoS guarantees for real-time traffic in sensor networks is still in its infancy.

This thesis presents a real-time communication framework to provide quality of service in sensor networks environments. The proposed framework consists of four components:

First, present an analytical model for implementing Priority Queuing (PQ) in a sensor node to calculate the queuing delay. The exact packet delay for corresponding classes is calculated. Further, the analytical results are validated through an extensive simulation study.

Second, report on a novel analytical model based on a limited service polling discipline. The model is based on an M/D/1 queuing system (a special class of M/G/1 queuing systems), which takes into account two different classes of traffic in a sensor node. The proposed model implements two queues in a sensor node that are served in a round robin fashion. The exact queuing delay in a sensor node for corresponding classes is calculated. Then, the analytical results are validated through an extensive simulation study.

Third, exhibit a novel packet delivery mechanism, namely the Multiple Level Stateless Protocol (MLSP), as a real-time protocol for sensor networks to guarantee the traffic in wireless sensor networks. MLSP improves the packet loss rate and the handling of holes in sensor network much better than its counterpart, MMSPEED. It also introduces the k-limited polling model for the first time. In addition, the whole sending packets dropped significantly compared to MMSPEED, which it leads to decrease the consumption power.

Fourth, explain a new framework for moving data from the sink to the user, at a low cost and low power, using the Universal Mobile Telecommunication System (UMTS), which is standard for the Third Generation Mobile System (3G). The integration of sensor networks with the 3G mobile network infrastructure will reduce the cost of building new infrastructures and enable the large-scale deployment of sensor networks.

IN MEMORY

OF

MY FATHER

ACKNOWLEDGEMENTS

To My Parents

For providing me with the opportunity to be where I am. Without them, none of this would even be possible. Special thanks for my best brother “Omer”

To My Wife

For her constant love, encouragement, and unconditional sacrifices to support my study and career.

To My Supervisor – Albert Zomaya

I thank you for your uncanny guidance, valuable advice, constructive comments, and moral support. Thank you for allowing me to learn from my own mistakes, for giving useful ideas and comments that helped me to complete my work, and for encouraging me to explore myself.

To Javid Taheri

I extend my deepest gratitude to Javid Taheri, for his insightful discussions and suggestions on various issues. Working with him has taught me to approach research with rigour, and to present research with clarity.

To My Lab Mates

I would especially like to thank my special friend Mohsin Iftikhar for his collaboration. He has been incredibly professional and a pleasure to work with. A big thank you to all my lab mates, Ali Elghirani, Khalid Almiani, and all others, for making life around the lab both more efficient and fun!

To the School of Information Technologies (USYD)

My thanks also go to the staff, post docs, and my fellow students at the Advanced Networks Research Group and at the School of Information Technologies, for creating such an enjoyable work environment. It has been a benefit interacting with these wonderful, bright, and talented people whose advice, feedback, and friendship have made my PhD experience both pleasurable and educational.

PUBLICATIONS FROM THE THESIS

The results contained in this dissertation have been published in the following forums (international journals and conferences).

CONFERENCES

- Aalsalem, M. Y., Taheri, J., and Zomaya, A., "Multiple level stateless protocol to guarantee the traffic in wireless sensor networks" in proc of IEEE Wireless Communication and Networking Conference (WCNC). Budapest, Hungary 5-8 April 2009. (Submitted)

- Aalsalem, M. Y., Taheri, J., Iftikhar, M., and Zomaya, A., "Providing QoS guarantees to multiple classes of traffic in wireless sensor networks" in proc. of IEEE/ACS AICCSA 2008.

- Aalsalem, M. Y., Iftikhar, M., Taheri, J., and Zomaya, A., "On the provisioning of guaranteed QoS in wireless sensor network through limited service polling models" in proc. of IEEE WOCN 2008.

- Aalsalem, M. Y., Everitt, D., and Zomaya, A., "An interoperability framework for sensor and UMTS networks" Third International Conference on Wireless and Mobile Communications (ICWMC'07). IEEE Computer Society and IEEE Xplore. March 4-9, 2007.

JOURNALS

- Aalsalem, M. Y., Zomaya, A., Iftikhar, M., and Taheri, J., “A QoS framework for wireless sensor networks,” *Journal of IEEE Transactions on Mobile Computing*, 2008. (Submitted)

- Aalsalem, M. Y., Zomaya, A., and Taheri, J., “A real time communication framework for wireless sensor networks,” *Journal of ACM Transactions on Sensor Networks*, 2008. (Submitted)

- Aalsalem, M. Y., Zomaya, A., and Taheri, J., “A-to-Z QoS and Routing survey in Wireless Sensor Networks” Elsevier International Journal on performance Evaluation of Wireless Ad hoc, Sensor, and Ubiquitous Networks. Dec 2008. (Submitted)

- Aalsalem, M. Y., Zomaya, A., and Taheri, J.,” End-to-End Real time communication framework for wireless sensor networks”, IEEE Journal on Selected Areas in Communication (J-SAC), Dec 2008. (Submitted)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

Over the past few years, wireless sensor networks have received a great deal of attention. This technology has changed the way we live, work, and interact with the physical environment – without boundaries. This chapter discusses the motivations behind this work and the goals of the research. A structural overview of the remainder of this thesis is also presented.

1.1 Introduction

Sensor networks have different constraints than traditional wired networks. First, the nodes in sensor networks are likely to be battery powered, and it is often very difficult to change the batteries for all of the nodes, as energy-conserving forms of communication and computation are essential to wireless sensor networks. Second, since sensors have limited computing power, they may not be able to run sophisticated network protocols. Third, since the bandwidth of the wireless links connecting sensor nodes is often limited, inter-sensor communication is further constrained. Finally, because sensor networks are often deployed by a single organization with inexpensive hardware, there is typically less need for interoperability with existing standards.

The differences between sensor networks and ad-hoc networks are summarized below [1]:

- The number of nodes in a sensor network can be vastly higher than the nodes in an ad-hoc network.
- Sensor nodes are heavily deployed in the field of interest. This dense deployment can be leveraged by the application.
- Sensor nodes are prone to collapses. This is due to several reasons, such as hardware failure, depleted batteries, and environmental factors etc. As a result, application needs a level of inherent fault tolerance and ability to reconfigure themselves.
- The topology of a sensor network changes regularly.
- Most ad-hoc networks are based on point-to-point communication, whereas many sensor networks use the broadcasting communication concept.
- Sensor nodes are limited in power, computational capacity, and memory.
- Sensor networks may not have universal identification (ID), because of the large number of sensors.

1.2 Key definition of sensor networks

Sensor networking is a challenging research area that draws on contributions from signal processing, networking and protocols, databases and information management, distributed algorithms, embedded systems, and architecture and QoS.

A number of key terms and concepts have been used throughout this thesis:

- Sensor: A transducer that converts a physical event into electrical or other signals that can be read by an observer or by a device.
- Sensor node: Also known as a mote, this is the basic unit in a sensor network and is capable of performing some processing, gathering sensory information, and communicating with other connected nodes in the network.
- Sink: Also known as a gateway, this is a special device with more power and memory than a node.
- Network topology: The study of the arrangement or mapping of the network elements, such as links, nodes, etc. it can also be a connection graph, in which nodes are sensor nodes and edges are communication links.
- Routing: The process of determining a network path from a packet source node to it is destination.
- Date-Centric: The term used when the sink sends queries to certain regions and waits for data from the sensors located in the selected regions.
- Geographic routing: Routing of data based on geographical attributes, such as locations or regions.
- Task: High-level system assignments, which include sensing, communication, processing, and resource allocation, or application tasks which may include detection, classification, localization, or tracking.
- Detection: The process of discovering the existence of a physical phenomenon.

- Classification: The assignment of class labels to a set of physical phenomena being observed.
- Resource: Elements including sensors, communication links, processors, onboard memory, and node energy reserves.
- Node service: Services, such as time synchronization and node localization that enable applications to discover properties of a node and organize themselves into a useful network.

1.3 Real-time communication

Real-time communication is the communication in which information is received at or nearly at the moment it is sent. Real-time applications demand the concept of QoS, in which there may be a scale of performance that is acceptable, and the boundary between success and failure of the system may be blurred or varied. The traditional QoS metrics do not apply in sensor network environments. However, wireless sensor networks deal with real-world environments. In many cases, sensor data must be delivered within time constraints so that appropriate observations can be made or actions can be taken. Very few results exist to date regarding meeting real-time requirements in wireless sensor networks.

1.4 Motivation

In the field of sensor network environments, much research has been conducted on topics such as power consumption, self-organisation techniques, routing between the sensors, and the communication between the sensor and the sink. On the other hand, real-time communication with the QoS concept in wireless sensor networks is still an open research field. Most protocols either ignore real time or simply attempt to process as fast as possible and hope that this speed is sufficient to meet the deadline.

1.5 Project goals

The fundamental aim of this project is to increase the accuracy of real-time communication in wireless sensor networks. In the pursuit of increased accuracy, this research will consider G/M/1 queuing systems. For this work to be relevant to real-time communication delivery, we must consider the QoS of the system with a variety of queuing schemes within suitable protocol. These aims may be realized through the following goals:

- To theoretically analyse system behaviour under widely used QoS queuing schemes, including priority queuing and polling queuing.
- To implement a simulator for the same queuing system, to allow comparison with theoretical results.
- To understand the accuracy of numerical analysis through comparison with simulation results.
- To implement the queuing schemes within our real-time framework, to verify the analytical models and numerical and simulation results.
- To deliver data from the sink to the user in the demand time.

Comprehensive analytical and numerical analysis, along with simulation results, are presented according to these goals.

1.6 Thesis contributions

This thesis provides a real-time communication framework of wireless sensor networks. It has three primary components: a suitable queuing system, a real-time communication framework, and an ability to gather the data in demand time. These components create a framework to address the real-time communication needs of wireless sensor networks. The contribution of this thesis can be summarized as follows:

1.6.1 Queuing system

- This thesis presents the closed-form expressions of the queuing delay for multiple (real-time and non-real-time) classes of traffic in a senor node through the implementation of priority queuing based on the M/G/1 queuing system. The measurement data can be useful as input to a simulation study of sensor networks (see Chapter 4).
- This thesis reports on a novel analytical model that is based on the M/G/1 queuing system. The model has been analysed on the basis of the limited-service polling discipline (both 1-limited and k-limited) to provide differential treatment to multiple classes of traffic in wireless sensor networks. The closed-form expressions of the queuing delay for two different classes of traffic have been derived (see Chapter 5).
- The measurement data can be useful as input to a simulation study of sensor networks. Also, the analytical modelling technique, and its verification through numerical and simulation results, is the first step towards finding the most appropriate queuing-scheme implementation for wireless sensor networks.

1.6.2 A real-time communication framework

- This thesis proposes a real-time framework that enables the sensor data to be delivered within time constraints that make suitable real-time actions possible.
- The framework improves the packet loss rate and the handling of holes in sensor networks.
- This framework supports multiple, dynamic routes with little or no state information.
- This framework provides built-in robustness to route failure.
- One challenge is to guarantee that the localized forwarding decisions result in the best end-to-end, real-time communication compared to others protocols.
- The framework includes a MAC-layer anycasting scheme that assists routing. This has the additional advantage of reducing the number of back-offs and, thus, the waiting times for data transmission.
- The framework is introduced by the k-limited polling model for the first time. It then uses this model in a sensor network by implementing two queues served according to a 2-limited polling model in a sensor node.
- This framework chooses the 2-limited polling model as the queuing model with the shortest elapsed (TTL) packet time first, as scheduling work gives the best possible average waiting time and minimizes the dropping packets sharply (see Chapter 6).

1.6.3 Gathering data in demand time

- This thesis presents a new approach of using the characteristics of sensor networks and mobile network infrastructure to deliver the sensor network signals.
- Communicating dynamically and intelligently between these two systems can reduce the cost and increase the lifetime of sensor networks. This approach is suitable for all organizations and for gathering data on demand.
- The feasibility and viability of the proposed method have been proven with the initial experimental work.
- Selecting the correct number of sink nodes can maximize the network lifetime as much as possible with the most economical investment. (see Chapter 7).

1.7 Thesis organization

While this chapter has given an overview of the motivation and scope of this research on wireless sensor networks, as well as a brief introduction, the remainder of this thesis is organized as follows:

- Chapter 2 presents a thorough overview of wireless sensor networks.
- Chapter 3 explains the background and related work. This chapter also presents a thorough overview of QoS frameworks, polling models, and M/G/1.
- Chapter 4 describes the procedure of formulating the queuing delay for multiple (real-time and non-real-time) classes of traffic in a senor node

through the implementation of priority queuing based on the M/G/1 queuing system.

- Chapter 5 presents and analyses the model on the basis of the limited-service polling discipline (both 1-limited and k-limited) to present differential treatment to multiple classes of traffic in wireless sensor networks
- Chapter 6 reports on a novel idea of real-time framework that enables the sensor data to be delivered within time constraints that make suitable real-time actions possible. This idea improves the packet loss rate and the handling of holes in sensor networks.
- Chapter 7 presents a new approach of using the characteristics of sensor networks and mobile network infrastructure (UMTS) to deliver the sensor network signals.
- Chapter 8 concludes by summarizing the contributions and evaluating their relevance to real-time delivery in wireless sensor networks. The direction of future research based on this thesis is also discussed.

Chapter 2

WIRELESS SENSOR NETWORKS – AN OVERVIEW

This chapter provides an intensive overview of wireless sensor networks. A wireless sensor network is a special network with large numbers of nodes equipped with embedded processors, sensors, and radios. These nodes collaborate to accomplish a common task, such as environment monitoring or asset tracking. In many applications, sensor nodes will be deployed in an ad-hoc fashion, without careful planning. In these cases, the nodes must organize themselves to form a multi-hop, wireless communication network. Advances in hardware and wireless network technologies have created low-cost, low-power, multifunctional miniature sensor devices.

2.1 Introduction

The first research in the area of wireless sensor networks was conducted in the early 1970s [2]. Since then, this technology has changed the way we live, work, and interact with the physical environment – without boundaries. The micro-electro-mechanical system (MEMS) [3], digital electronics, and wireless communications have enabled the development of a new generation of large-scale sensor nodes that are small in size and communicate unjoined over short distances with low-cost, low-power services. These new sensor nodes are suitable for a range of commercial and industrial applications [4-6]. These devices build a self-organizing ad-hoc network to forward data packets using multi-hop connections towards sink nodes [6].

The self-organization technique of sensor networks renders it possible to set up nodes randomly over a wide area being monitored, such as by dropping them from an aircraft. The self-organization feature of sensor networks includes both communications self-organization and positioning self-organization [7]. In this manner, a large number of sensor nodes are spread over the environment, without having prior information about the placement of each individual sensor [8].

Sensor nodes have a short transmission range due to their limited radio capabilities. Therefore, the data must be relayed using intermediate nodes towards the sink. In addition, it may be advantageous to use a multi-hop path, consisting of shorter links rather than a single long connection, to the sink node.

The area of sensor network quality of service (QoS) remains largely open. This is a rich area for research, because sensor deaths and sensor replenishments make it difficult to specify the optimum number of sensors that should be sending information at any given time [9].

The available energy of the sensor nodes is the most critical resource in the sensor network. The limitation of the energy source is the battery, but this is the only source of power for the nodes, which can supply the sensor with energy. The sensors cannot operate with exhausted batteries. Moreover, since sensor nodes behave as relay nodes for data propagation of other sensors to sink nodes, network connectivity decreases gradually. This may result in disconnected sub-networks of sensors. Therefore, the level of power consumption must be considered at each stage of a wireless sensor network's design. Future military applications will increasingly feature communication scenarios involving a data-gathering or intelligence-gathering wireless sensor network. Several survey papers are presented with intense background research on sensor networks [1, 5, 10-27].

2.2 Applications of sensor networks

Sensor networks are functional in a wide range of areas, such as military applications, public safety, medical, surveillance, environmental monitoring, commercial applications, habitat, and tracking [24, 26, 28, 29]. Sensor network may consist of many different types of sensors, which able to monitor a wide variety of ambient conditions that include the some of the following categories [10]:

- Measuring conditions such as temperature, humidity, pressure and noise level
- Analysis of items like soil characteristics and chemical tracking agents
- Monitoring in fields such as the environment (e.g., habitat, security), health (e.g. tracking and monitoring doctors and patients), the military (e.g., battle field, targeting), smart transportation, industrial sensing and diagnostics (e.g., factory, supply chains), and infrastructure protection (e.g., power grid, water distribution)

Deploying sensor nodes in unattended environments allows many possibilities for the exploration of new applications in the real world. Here are just some of the applications in this context [4, 5, 24, 30-38]:

- Military applications:
  - Battle damage assessment
  - Battlefield surveillance

    - Monitoring the enemies
    - Attack detection (nuclear, biological, and chemical)
- Environmental applications:
    - Forest fire detection
    - Flood detection
    - Microclimates
    - Precision agriculture
- Health applications:
    - Tracking and monitoring patients with certain diseases
    - Tracking and monitoring doctors and patients inside the hospital
    - Elderly assistance
    - Residential applications:
    - Home automation
    - Instrumented environments
    - Automated meter reading
- Commercial applications:
    - Environmental control in industrial and office buildings
    - Inventory control
    - Vehicle tracking and detection
    - Traffic flow surveillance

Some of these applications need real-time communication to deliver the data at the demanded time.

2.3 Wireless sensor network architecture

Wireless sensor networks consist of numerous sensor nodes with sensing, wireless communications, and computation abilities. These sensor nodes are dispersed in unattended environments located far from the user.

2.3.1 Sensor network units

The most important units that build the sensor node architecture are:

- Sensor node (mote): A hardware device that generates a measurable response to a change in physical or environmental conditions at different locations. Sensor nodes form a wireless network by communicating over a wireless medium. They are responsible for collecting and routing data back to the sink. Sensor nodes are very low cost devices.
- Sink (Gateway): A special device that has more power than a sensor node and is responsible for sending collected data to the user. The sink is located near or inside the sensor field. It can be stationary or moving within the sensor field.
- Sensor channels: The communication among the sensor nodes and sink.
- Network channels: The transmission from the sensor network to other networks, or between different sensor networks.
- Phenomenon: A unit of concern to the user. The phenomenon is measured, monitored, and analysed by the sensor nodes.
- User: The person interested in obtaining information about a specific phenomenon.

The first step towards understanding wireless sensor networks is to comprehend the ideas and concepts of sensor node architecture in detail.

2.3.2 Sensor node hardware architecture

Sensor nodes should be small in size, should consume extremely low energy, and should operate in high volumetric densities. Sensor nodes are autonomous and operate unattended.

Sensor node hardware can be classified into three types, each of which requires a set of trade-offs in the design choices [10]:

- Augmented general purpose computers: The examples in this class include personal digital assistants (PDAs), which comprise low-power PCs and embedded PCs (e.g., PC104). These nodes typically run on operating systems such as Windows CE, Linux, or real-time operating systems. They use standard wireless communication protocols like Bluetooth or IEEE 802.11.
- Dedicated embedded sensor nodes: The Berkeley mote family, Ember mote, and UCLA Medusa family are examples of this class [39, 40]. These nodes are generally used for commercial applications.
- System on Chip (SoC) nodes: Smart dust [4, 41], BWRC Picoradio nodes [42], and PASTA nodes are examples of this class. The goal is to find new ways of integrating complementary metal-oxide-semiconductor (CMOS), MEMS, and RF technologies to build exceedingly low-power and small-footprint sensor nodes that provide certain sensing, computation, and communication capabilities.

The common architecture of a sensor node is shown in Figure 2-1. The major components are the sensing unit, processing unit, transceiver, and power unit. The environmental information is retrieved using the sensor and converted with an analog-to-digital converter (ADC) to digital data. These data are forwarded to the processing unit to become a data packet that is then sent to the sink node for further examination. The communication between the sensor nodes is carried out with the transceiver. The power unit feeds all of these components with the necessary operational power.

The optional units, such as the location finding system, mobilizer, and power generator may be implanted to the node based on the application. Most of the applications have a need for some location information for the sensed data when they reach the sink node. Mobility may also be an application-specific requirement. Although most monitoring applications operate only static sensor nodes, for some tracking scenarios, mobility may be a major design consideration. Finally, in order to extend the lifetime of a sensor node, a power-rummaging tool, such as solar cells, can be attached to the node.

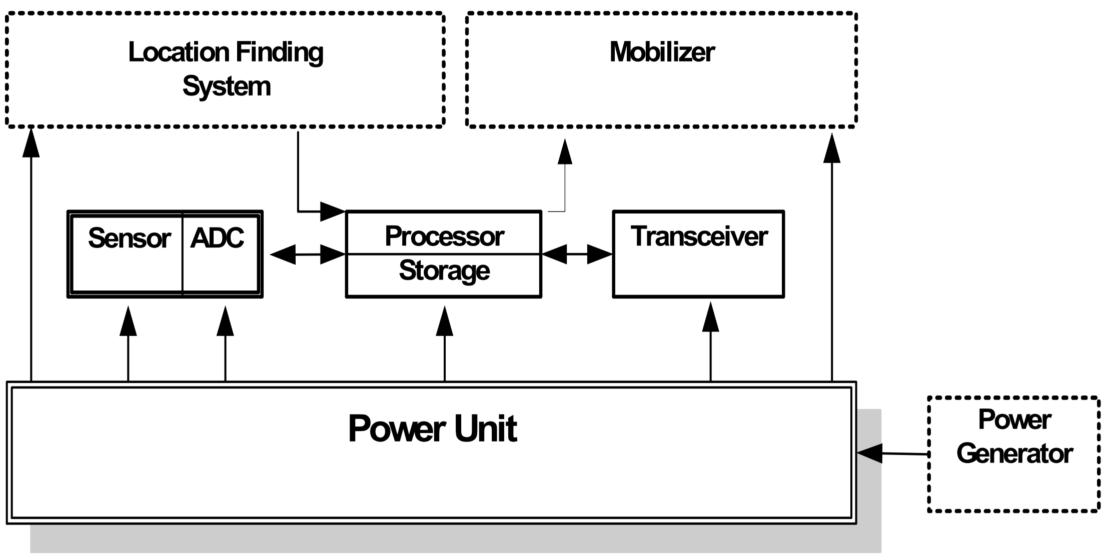

Figure 2-1: Architecture of a sensor node

2.4 Routing

Many new routing and MAC-layer protocols have been proposed for wireless sensor networks, in the hopes of tackling the issues raised by resource-constrained, unattended sensor nodes in large-scale deployments. The majority of these protocols consider energy efficiency as the main objective and assume data traffic with unconstrained delivery requirements. Table 2-1 summarizes recent research results on data routing in wireless sensor networks.

Table 2-1 : Classification and Comparison of routing protocols

	Classification	Mobility	Position Awareness	Power Usage	Negotiation Based	Data Aggregation	Localization	QoS	Delay (Latency)	Data Delivery Model	State Complexity	Scalability	Multipath	Query Based
SPIN [43]	Flat	P	No	Li	Yes	Yes	No	No	Mod	Event driven	Low	Li	Yes	Yes
Directed Diffusion [44-46]	Flat	Li	No	Li	Yes	Yes	Yes	No	Mod	Demand driven	Low	Li	Yes	Yes
Rumor Routing [47]	Flat	VL	No	N/A	No	Yes	No	No	Mod	Demand driven	Low	G	No	Yes
GBR [48]	Flat	Li	No	N/A	No	Yes	No	No	Low	Gradient	Low	Li	No	Yes
MCFA [49]	Flat	No	No	N/A	No	No	No	No	Mod	Routing table	Low	G	No	No
CADR [50]	Flat	No	No	Li	No	Yes	No	No	Low	Continuously	Low	Li	No	No
COUGAR [51]	Flat	No	No	Li	No	Yes	No	No	Mod	Leader, Query	Low	Li	No	Yes
ACQUIRE [52]	Flat	Li	No	N/A	No	Yes	No	No	Mod	Complex query	Low	Li	No	Yes

EAR [53]	Flat	Li	No	N/A	No	No	No	No	Mod	Demand driven	Low	Li	No	Yes
LEACH [54]	H	FB	No	Max	No	Yes	Yes	No	Low	Cluster head	CHs	G	No	No
TEEN & APTEEN [55, 56]	H	FB	No	Max	No	Yes	Yes	No	Mod	Event driven	CHs	G	No	No
PEGASIS [57]	H	FB	No	Max	No	No	Yes	No	High	Chains	Low	G	No	No
MECN & SMECN [58, 59]	H	No	No	Max	No	No	No	No	Low	Geographic	Low	Low	No	No
SOP [60]	H	No	No	N/A	No	No	No	No	Low	Routing table	Low	Low	No	No
HPAR [61]	H	No	No	N/A	No	No	No	No	Mod	Adaptive routing	Low	G	No	No
VGA [62]	H	No	No	N/A	Yes	Yes	Yes	No	High	Virtual grid	CHs	G	Yes	No
Sensor Aggregate [63]	H	Li	No	N/A	No	Yes	No	No	Mod	Cluster head	Low	G	No	P
TTDD [64]	H	Yes	Yes	Li	No	No	No	No	High	Query driven	Mod	Low	P	P
GAF[65]	L	Li	No	Li	No	No	No	No	Mod	Virtual grid	Low	G	No	No
GEAR[66]	L	Li	No	Li	No	No	No	No	Mod	Demand driven	Low	Li	No	No
SPAN[67]	L	Li	No	N/A	Yes	No	No	No	Mod	Continuously	Low	Li	No	No
MFR, GEDIR [68]	L	No	No	N/A	No	No	No	No	Mod	Geographic	Low	Li	No	No

GOAFR [69]	L	No	No	N/A	No	No	No	No	Mod	Geographic	Low	G	No	No
SAR [22]	QoS	No	No	N/A	Yes	Yes	No	Yes	Low	Table driven	Mod	Li	Yes	Yes
RAP [70]	QoS	No	No	N/A	No	No	No	Yes	Mod	Geographic	Mod	Li	No	Yes
SPEED [71]	QoS	No	No	N/A	No	No	No	Yes	Mod	Geographic	Mod	Li	No	Yes
MMSPEED [72]	QoS	No	No	N/A	No	No	No	Yes	Mod	Geographic	Hi	Li	Yes	Yes
RPAR [73]	QoS	No	No	N/A	No	No	No	Yes	Mod	Geographic	Hi	Li	No	Yes

F= Flat, H= Hierarchical, L= Location, P= Possible, Li= Limited, VL = Very Limited, FB= Fixed BS, Mod = Moderate, G= Good, CH= Cluster Head

2.5 Routing challenges and design issues in wireless sensor networks

The growing interest in applications that require assured end-to-end performance guarantees, along with the introduction of real-time communication, have posed additional challenges to wireless sensor networks. Depending on the application, different design purposes and constraints have been considered for sensor networks. Since the performance of a routing protocol is closely related to the architectural model, this section highlights the constraints and challenges of wireless sensor networks.

Transmission of data in such cases requires both energy and QoS-aware network management in order to ensure efficient usage of the sensor resources and effective access to the gathered measurements. However, the design of routing protocols in wireless sensor networks is influenced by many challenging factors. These factors must be overcome before efficient communication can be achieved in wireless sensor networks [74].

2.5.1 Resource constraints

Sensor nodes are usually low-cost, low-power, small devices that are equipped with only limited data-processing capabilities, transmission rates, battery power, and memory. Due to the limitations on transmission power, the available bandwidth and radio range of the wireless channel are often limited. However, energy conservation is critically essential for extending the lifetime of the network, because it is often infeasible or undesirable to recharge or replace the batteries attached to sensor nodes once they are deployed. In the existence of resource constraints, the network QoS may experience unavailability of computing and/or communication resources. For illustration, consider a number of nodes that desire to spread messages over the same wireless sensor network. These nodes must compete for the limited bandwidth that the network is able to provide. As a result, some data transmissions will perhaps experience large delays, resulting in low levels of QoS. Due to the limited memory size, data packets may be dropped before the nodes successfully send them to their destinations. Thus, it is of critical importance to use the available resources in wireless sensor networks in a very efficient way. the recent commercial sensor nodes, along with their limitations has been listed in Table 2-2 also Table 2-3 shows the recent gateways sensor nodes [75-78] .

Table 2-2: List of Sensor Nodes

Sensor Node Name	Microcontroller	Transceiver	Program + Data Memory	External Memory	Program	Support
BEAN [76]	MSP430F169	CC1000 (300-1000 MHz) with 78.6 kbit/s	4 Mbit			YATOS
BTnode [79]	Atmel ATmega 128L (8 MHz @ 8 MIPS)	Chipcon CC1000 (433-915 MHz) and Bluetooth (2.4 GHz)	64+180 K RAM	128K FLASH ROM, 4K EEPROM	C and nes C	BTnut and Tiny OS
COTS [80]	ATMEL Microcontroller 916 MHz					
Dot [81]	ATMEGA163	315-433 MHz	1K RAM	8-16K Flash	weC	
Eyes [82]	MSP430F149	TR1001		8 Mbit		Peer OS
EyesIFX vl [82]	MSP430F149	TDA5250 (868 MHz) FSK	RAM 2KB	8 Mbit		Tiny OS
EyesIFX v2 [82]	MSP430F1611	TDA5250 (868 MHz) FSK	RAM 10 KB	8 Mbit		Tiny OS
GWnode [83]	PIC18LF8722	BiM (173 MHz) FSK	64k RAM	128k Flash	C	Custom OS
IMote [84]	ARM core 12 MHz	Bluetooth with the range of 30 m	64K SRAM	512K Flash		Tiny OS
IMote 1.0 [84]	ARM 7TDMI 12-48 MHz	Bluetooth with the range of 30 m	64K SRAM	512K Flash		Tiny OS
IMote 2.0 [84]	Marvel PXA271 ARM 11-400 MHz	TI CC2420 802.15.4/ZigBee compliant radio	32MB SRAM	32MB Flash		Microsof t .NET Micro Linux TinyOS
Iris [84]	ATmega1281	Atmel AT86RF230 802.15.4/ZigBee compliant radio	8K RAM	128K Flash	Nes C	Tiny OS, Mote Works

KMote [84]	TI MSP430 microcontroller	250 kbit/s 2.4 GHz IEEE 802.15.4 Chipcon Wireless Transceiver	10k RAM	48k Flash		Tiny OS and SOS
Mica [84]	Atmel ATMEGA103 4 MHz 8-bit CPU	RFM TR1000 radio 50 kbit/s	128+4K RAM	512K Flash	nesC	Tiny OS Support
Mica2[84]	ATMEGA 128L	Chipcon 868/916 MHz	4K RAM	128K Flash		Tiny OS SOS and Mantis OS
Mica 2Dot [84]	ATMEGA 128		4K RAM	128K Flash		
MicaZ [84]	ATMEGA 128	TI CC2420 802.15.4/ZigBee compliant radio	4K RAM	128K Flash	nesC	TinyOS, SOS, MantisO S and Nano-RK
Mulle [85]	Renesas M16C	Bluetooth 2.0	31K RAM	384K+4K Flash, 2 MB EEPROM	C	TCP/IP and Bluetoot h Profiles: LAP, DUN, PAN and SPP
Nymph [86]	ATMEGA128L	CC1000		64 kB EEPROM		Mantis OS
Rene [84]	ATMEL8535	916 MHz radio with bandwidth of 10 kbit/s	512 bytes RAM	8K Flash		TinyOS
SenseNode	MSP430F1611	Chipcon CC2420	10K RAM	48K Flash	C and NesC	GenOS and TinyOS
Sun SPOT [87]	ARM 920T	802.15.4	512K RAM	4 MB Flash	Java	Squawk J2ME Virtual Machine
Telos [88]	Motorola HCS08		4K RAM			
TelosB [88]	Texas Instruments MSP430 microcontroller	250 kbit/s 2.4 GHz IEEE 802.15.4 Chipcon Wireless Transceiver	10k RAM	48k Flash		Contiki, TinyOS, SOS and MantisO S

T-Mote Sky [89]	Texas Instruments MSP430 microcontroller	250 kbit/s 2.4 GHz IEEE 802.15.4 Chipcon Wireless Transceiver	10k RAM	48k Flash		Contiki, TinyOS, SOS and MantisOS Support
weC [4]	Atmel AVR AT90S2313	RFM TR1000 RF				
Flecks [90]	Atmel ATmega 128	915 MHZ Radio Nordic nRF905	8KB RAM 4K EEPROM	8 MB Flash	C	TinyOS and FOS
FireFly [91]	Atmel ATmega 1281	Chipcon CC2420	8K RAM	128K FLASH ROM, 4K EEPROM	C	Nano-RK RTOS

Table 2-3: List of Gateway Sensor Nodes

Gateway	Microcontroller	Transceiver	Program + Data Memory	External Memory	Program	Support
Stargat	IntelPXA255	802.11	serial connection to WSN	64 MB SDRAM	32 MB Flash	IntelPXA255

2.5.2 Node deployment

Node deployment in wireless sensor networks is application dependent and affects the performance of the routing protocols. Deployment can be either deterministic or randomized. In deterministic deployment, the sensors are manually placed, and data is routed through predetermined paths. In addition, collision can be minimized through the pre-scheduling of media access. In random node deployment, the sensor nodes are scattered randomly, creating an ad-hoc routing infrastructure [22, 54, 55, 92, 93]. While the distribution of nodes is not uniform, optimal clustering becomes necessary to allow connectivity and enable energy-efficient network operation.

2.5.3 Platform heterogeneity

Sensor nodes can exhibit different functions or capabilities. The existence of a heterogeneous set of sensors raises many technical issues related to data routing. Data sensing and reporting can be produced from these sensors at different rates, subject to various QoS constraints, and can follow multiple date-reporting models.

2.5.4 Node communications

When building a clustered network, communication between sensor nodes is greatly influenced by energy considerations. Since the transmission power of wireless radio is proportional to the distance squared, or an even higher order in the existence of obstacles, multi-hop routing will consume less energy than direct communication [54]. Most of the time, sensors are scattered randomly over an area of interest, and multi-hop routing becomes unavoidable.

2.5.5 Node capabilities

Sensor nodes have limited computing power and, therefore, may not be able to run sophisticated network protocols. However, a sensor network can consist of different sensor nodes with equal [44, 46, 94] or unequal [54, 57] capacity, in terms of computation, communication, and power.

2.5.6 Production costs

Because a sensor network consists of a large number of sensor nodes, the cost of a single node is vital to justifying the overall cost of the network. Along with different functionalities, such as sensing or processing, sensors have different price points. As a result, the appropriate amount to spend on a sensor node is a challenging issue.

2.5.7 Energy consumption without losing accuracy

Computing and transmitting information in a wireless sensor network environment can use up the sensor nodes' limited supplies of energy. In addition, efficient communication and computation are essential to conserve the energy. A sensor node's useful life is heavily dependent on battery life [54].

2.5.8 Data delivery method

Depending on the time criticality of the data and also on the application of the wireless sensor network, data reporting can be categorized as continuous, event driven, query driven, or a hybrid of all of these methods. The continuous model, sensors send their data continuously to the sink at a pre-specified rate. This method is suitable for applications that need periodic data monitoring. In event-driven: most event-driven applications in wireless sensor networks are interactive, delay intolerant (real-time), mission critical and non-end-to-end applications. It means that the events sensors are expected to observe are very important to the success of the application. Query-driven methods, this data delivery model is similar to the event-driven model except that the data is pulled by the sink while the data is pushed to the sink in the event driven model. To save energy, queries can be sent on demand. A combination of the previous methods is also possible and is known as the hybrid model. The routing protocol is highly influenced by the data delivery method in terms of energy consumption and route calculations[95].

2.5.9 Fault tolerance or reliability

Sensors may fail or be blocked due to surrounding physical conditions, environmental interference, or a depletion of energy. It may be difficult to replace existing sensors, so the failure of sensor nodes should not affect the overall task of the sensor network [96]. It must be fault tolerant so that actual network conditions are transparent to the given application. Therefore, multiple levels of redundancy may be needed in a fault-tolerant sensor network.

2.5.10 Density and network size/scalability

The number of sensor nodes deployed in studying a phenomenon may be on the order of hundreds or thousands. Depending on the application, the number may even reach an extreme value of millions. The density of these nodes affects the quantity of coverage in the area of interest. The network's size affects reliability, accuracy, and data processing algorithms [25, 96]. Any routing scheme must be capable of working with the specific number of sensor nodes. In addition, sensor network routing protocols should be scalable enough to respond to events in the environment. Until an event occurs, most of the sensors can stay in the sleep state, with data from the few remaining sensors providing coarse quality.

2.5.11 Network dynamic

There are three main components in a sensor network. These are the sensor nodes, the sink, and monitored events. In many studies, sensor nodes are stationary. However, in some applications, both the sink and sensor nodes can be mobile. Nonetheless, routing messages to or from moving nodes is more challenging, since route and topology stability become important optimization factors, in addition to energy, bandwidth, and so forth. The sensed event can be either dynamic or static depending on the application [64].

2.5.12 Sensor network topology

The topology of a network affects many of its characteristics, such as latency, capacity, and robustness. The complexity of data routing and processing also depends on the network topology. Densely deploying thousands of sensor nodes in an area requires careful handling of network topology maintenance [5, 25]. There are three phases related to topology maintenance and changes (e.g., malfunctioning of some sensor nodes): the pre-deployment and deployment phase, the post-deployment phase, and the re-deployment of additional nodes phase [5]. Dealing with the inherent dynamics of wireless sensor networks requires QoS mechanisms to work in dynamic, and even changeable, environments.

2.5.13 Transmission media

In a multi-hop sensor network, communicating nodes are linked by a wireless medium. These links can be formed by radio (e.g., Bluetooth compatible 2.4 GHz and IEEE 802.11 transceivers), infrared, and optical media. Infrared is license free and robust to interference from electrical devices [97, 98].

2.5.14 Connectivity

High node density in sensor networks prevents nodes from being completely isolated from each other. Therefore, sensor nodes are expected to be well connected. This, however, may not keep the network topology from being variable or the network size from shrinking as a result of the death or failure of some sensor nodes.

2.5.15 Coverage

Each sensor node achieves a certain view of the environment that it is positioned in, and it is limited both in range and in accuracy. It can only cover a limited physical area of the environment. Hence, area coverage is also an important design parameter in wireless sensor networks.

2.5.16 Control overhead

When collisions, latency, and energy consumption increase, the number of re-transmissions in a wireless medium will also increase. Therefore, there is an increase in the number of control packets overhead to form linearly with node density relation. As a result, trade-offs among energy conservation, self-configuration, pre-node fairness, and latency may exist.

2.5.17 Data aggregation

Since sensor nodes may generate significant amounts of redundant data, similar packets from multiple nodes can be aggregated so that the number of transmissions is reduced. Data aggregation is the combination of data from different sources by using functions such as suppression (eliminating duplicates), min, max, and average [99]. This technique has been used to achieve energy efficiency and data transfer optimization in a number of routing protocols. Signal processing methods can also be used for data aggregation [60]. This scenario is referred to as data fusion. Data fusion occurs when a node is capable of producing a more accurate output signal by using techniques such as beamforming to combine the incoming signals and reduce the noise in these signals [54].

2.5.18 Security

Due to inherent constraints in wireless sensor networks, security is a vital issue. Setting security goals for sensor networks will depend on knowing what it is that needs to be protected. Sensor networks share some of the features of mobile ad-hoc networks but also have unique challenges. Therefore, security objectives should include both those of traditional networks and those suitable to the unique limitations of sensor networks. The four security goals for sensor networks are Confidentiality, Integrity, Authentication, and Availability (CIAA)[100].

Security features in wireless sensor networks focus on centralized communications approaches. Some of the threats to a wireless sensor network are described in [100-102] and categorized as follows: passive information gathering, false node, node outage, supervision of a node, node malfunction, message corruption, denial of service, and traffic analysis form details [100, 101]. There is a need to develop distributed security approaches for wireless sensor networks. In [103], details are provided regarding routing attacks in sensor networks.

2.5.19 Self-configuration

Self-configuration is critical for wireless sensor networks, since the densely deployed sensor nodes in a sensor field may fail due to such reasons as lack of energy, physical destruction, environmental interference, communications problems, or new nodes joining the network. Moreover, sensor nodes work unattended in a dynamic environment, so they need to be self-configured to begin a topology that supports communications under severe energy constraints. It is worthwhile to mention that self-configuration in a wireless sensor network is an essential factor to maintain functions properly and serve the networks' purpose [104, 105].

2.5.20 Environment

Sensor nodes are densely deployed either very close to or directly within the phenomenon to be researched. They are designed to work inside of huge equipment, at the bottom of the ocean, in a biologically or chemically infected field, in a battlefield beyond the enemy lines, and in a home or large building.

2.6 QoS requirements in wireless sensor networks

Several protocols and algorithms have been proposed for routing QoS in wire-based networks [106-108]. However, they cannot be directly applied to wireless networks due to the inherent characteristics that distinguish the two types of networks [109]. In addition, the nature of sensor networks poses unique challenges compared to general wireless networks and, thus, requires special attention. The type of target application can play an important role for QoS in wireless sensor networks. QoS in wireless sensor networks can be characterized by reliability, timeliness, robustness, availability, and security between all others [110]. The throughput, delay, jitter, and packet loss rate are the most fundamental parameters [111, 112] and may be used to measure the degree of satisfaction of these services.

2.6.1 Throughput

Throughput is the average rate of successful data deliveries over a communication channel within a certain period of time. In some situations, throughput is also called bandwidth [113]. In general, there is a relationship between throughput and performance – a larger throughput of the network leads to better system performance.

2.6.2 Delay

Delay is the time elapsed from the departure of a data packet from the source node to the arrival at the destination node. This includes the processing delay, queuing delay, transmission delay, and propagation delay.

The processing delay is the time sensors take to process the packet header, while the queuing delay is the time the packet sits in sensor queues. The transmission delay is the time it takes to push the packet's bits onto the link. The propagation delay is the time it takes for the signal to propagate through the transmission medium. Delay-sensitive applications usually require the delivery of data packets in real time. Notice that real time does not necessarily mean fast computation or communication.

2.6.3 Jitter

Jitter is referred to as variations in delay, despite many other definitions. Jitter is often caused by the difference in queuing delays experienced by successive packets. Random and deterministic are two types of jitter. Random jitter, also called Gaussian jitter, is random electronic timing noise. Deterministic jitter is a type of clock-timing jitter or data-signal jitter that is predictable and reproducible.

2.6.4 Packet loss rate

The packet loss rate is the proportion of data packets that are lost during the process of transmission. This rate can be used to represent the probability of packets being lost. A lost packet can be caused by a number of factors, including congestion, bit error, or bad connectivity. This parameter is closely related to the reliability of the network.

2.7 QoS challenges in wireless sensor networks

The information in this section is a summary of design considerations for the treatment of QoS traffic in wireless sensor networks.

2.7.1 Bandwidth constraint

Bandwidth is the rate of data transfer, or bit rate, measured in bits per second. The competition for bandwidth represents a classic problem in networking. If the presented load into the wireless sensor network exceeds the available bandwidth, the network must respond by either discarding packets or queuing them in memory and waiting for resources to become available. So, while bandwidth in wireless sensor networks is a constrained resource for which network-attached devices compete, other side effects – delay, jitter, and packet loss rate – occur as a result.

2.7.2 Buffer size constraint

Buffer size, in general, plays an important function in holding the data before forwarding it to the next node. Multi-hop routing of QoS data would typically need long sessions and buffering of even larger data, especially when the delay jitter is of concern. The buffer size constraint will increase the delay difference that packets incur while travelling on different routes, or even on the same route.

2.7.3 Queuing constraint

There are different types of queuing tools available [114, 115], such as FIFO queuing [116], Priority Queuing (PQ) [116], Weighted Fair Queuing (WFQ) [117], Class-Based Weighted Fair Queuing (CBWFQ) [118], and Low Latency Queuing (LLQ) [119]. In addition, all queuing tools are classified on the basis of six items: (1) classification, (2) drop policy, (3) maximum number of queues at an output interface, (4) maximum queue length, (5) scheduling inside each queue, and (6) output scheduler logic (scheduling among different queues). Depending on the hardware constraints of the sensor node, only some of these tools can be implemented.

2.7.4 Unbalanced traffic

Wireless sensor networks consist of hundreds or thousands of nodes, with only limited number of sinks. The traffic essentially floods from a large number of sensor nodes to a small subset of sink nodes. The QoS mechanism should be designed for unbalanced traffic.

2.7.5 Data redundancy

Data redundancies in sensor networks are characterized in the general data. Data redundancy leads to a loss of the reliability/robustness requirement of data delivery; it unnecessarily spends a great amount of valuable energy.

2.7.6 Energy balance

The energy load must be evenly distributed among all sensor nodes so that no node is drained out of energy faster than any others. An energy balance increases the life of the network.

2.7.7 Energy and delay trade-off

The transmission power of radio is proportional to the distance squared, or an even higher order in noisy or non-flat environments. Multi-hop routing is a pattern design in wireless sensor networks. While the increase in the number of hops significantly reduces the energy consumed for data collection, the accumulative packet delay expands. Since packet queuing delays control propagation delays, the increase in the number of hops can not only slow down packet delivery but also make the analysis and the handling of delay-constrained traffic difficult. Therefore, it is expected that QoS routing of sensor data would need to sacrifice energy efficiency to meet delivery requirements. In addition, redundant routing of data may be unavoidable to cope with the typical high error rate in wireless communication, further complicating the trade-off between energy consumption and delay of packet delivery.

2.7.8 Multiple sinks

The existence of multiple sinks leads to different requirements on the network. Sensor networks should be capable of supporting different QoS levels associated with multiple sinks. However, finding the optimal number of sink nodes, which will increase the lifetime of the network, is a big QoS challenge.

2.7.9 Multiple traffic types

The existence of heterogeneous sets of sensors creates challenges for multiple-traffic QoS support. For example, several applications may require a diverse combination of sensors for monitoring temperature, pressure, and humidity of the surrounding environment; detecting motion via sound signatures; and capturing the images of moving objects.

2.7.10 Packet criticality

There are two kinds of packets, high priority and low priority packets. QoS mechanisms may be required to distinguish packet importance and set up a priority structure.

2.8 Layered communication protocols

The layered communication protocols in sensor nodes and sinks consist of the physical layer, data-link layer, network layer, transport layer, application layer, power management plane, mobility management plane, and task management plane.

Table 2-4: summaries of the layered communication protocols and open Research issues

Layer	Description	Open Research Issues
Physical	Responsible for frequency selection, carrier frequency generation, signal detection modulation, and data encryption.	▪ Hardware design ▪ Strategies to overcome signal propagation effects ▪ Modulation schemes ▪ Methods of improving the transmission rate
Data link	Responsible for the multiplexing of data streams, data frame detection, medium access, and error control. Ensures reliable point-to-point and point-to-multipoint connections in a communication network.	▪ Design-scalable MAC ▪ MAC/Physical Cross Layer Design ▪ Design MAC for mobile sensor network ▪ Explore the possibility of other error control coding schemes ▪ Power saving modes of operation

Network	Takes care of routing the data supplied from the transport layer.	▪ Improvements to existing protocols to address higher topology, scalability, and real-time communication challenges
Transport	Helps to maintain the flow of data if the sensor network application requires it.	▪ Improve the existing transport protocols ▪ Sink can play important role in transport protocols ▪
Application	Different types; depends on the sensing tasks.	▪ Improve existing application-layer protocols ▪ Different application needs
Power management plane	Manages the way in which a sensor node uses its power.	▪ Introducing the rules related to data aggregation ▪ Turning sensor nodes on and off ▪ Optimizing power consumption and connectivity
Mobility management plane	Detects and registers the movement of sensor nodes.	▪ Exchange the data related to location finding algorithms. ▪ Location services
Task management plane	It balances and schedules the sensing tasks given to a specific region	▪ Optimize the way a sensor node participates in a sensing task ▪ Control node communication activities

Chapter 3

BACKGROUND -RELATED WORK

Due to inherent limitations in wireless sensor networks, real-time communication is a crucial issue. However, selecting the appropriate traffic model can lead to the successful design of wireless sensor networks. This chapter introduces the queuing problems that sensor networks face due to node resource limitations and other sensor network constraints. The network simulation plays an important role in achieving the framework goals, and an overview is followed by a brief survey of related work.

3.1 Introduction to queuing systems

Communication systems have been studied as a network of queues over the past years. D. G. Kendall introduced a standard notation for classifying queuing systems into different types [120]. Systems are described by the notation

A / B / C / D / E, where:

- A: Distribution of inter-arrival times of customers (packets)
- B: Distribution of service times
- C: Number of Servers
- D: Maximum total number of customers (packets) that can be accommodated in the system
- E: Calling population size

A and B can take any of the following distribution types:

- M: Exponential distribution (Markovian)
- D: Deterministic distribution
- E: Erlang distribution
- G: General (arbitrary) distribution

For example, D/M/n would describe a queuing system with a deterministic distribution for the inter-arrival times of customers, an exponential distribution for service times of customers, and n servers [121, 122].

3.1.1 G/M/1 queuing system

G/M/1 is a single-server queuing system, which is dual of M/G/1. In G/M/1, the arrival process is general, and service times are exponentially distributed. The authors in [123] gives detailed discussion of queuing systems.

3.1.2 Polling models

The basic polling model is a queuing model composed of a set of queues and a single server that serves the queues in cyclical order [124]. Polling models have been used in a variety of contexts since the 1960s. The advent of computer-communication networks and digital communication opened up new applications for polling models. For example, over the last two decades, polling models have been studied extensively to analyse the performance of Local Area Networks (LANs) employing different forms of token passing [125]. Generally, polling models can be classified as exhaustive, gated, or limited service. The exact details of the systems are beyond the scope of this thesis. Instead, the readers are referred to [126] for a detailed discussion of polling systems. The case, in which k = 1 results in simpler models, is referred to as 1-limited polling model (also called an alternating service polling model), in which the server serves one packet from each queue in an alternating fashion during each cycle.

3.2 Universal Mobile Telecommunications System (UMTS)

The Universal Mobile Telecommunications System (UMTS) is envisioned as the successor to the Global System for Mobile Communications (GSM). UMTS signals the move into the third generation (3G) of mobile networks. UMTS also addresses the growing demand of mobile and Internet applications for new capacity in the overcrowded mobile communications sky. This new network increases transmission speed to 2 Mbps per mobile user and establishes a global roaming standard [127].

UMTS, also referred to as Wideband-Code Division Multiple Access (W–CDMA), is one of the most significant advancements in the evolution of telecommunications into 3G networks. UMTS allows many more applications to be introduced to a worldwide base of users and provides a vital link between today's multiple GSM systems and the ultimate single worldwide standard for all mobile telecommunications, International Mobile Telecommunications–2000 (IMT–2000).

UMTS was developed mainly for countries with GSM networks [128], because these countries have agreed to free new frequency ranges for UMTS networks. Because it is a new technology that exists in a new frequency band, a whole new radio access network has to be built [129]. The advantage is that a new frequency range gives plenty of new capacity for operators. 3rd Generation Partnership Project (3GPP) is overseeing the standard development and has wisely kept the core network as similar to the GSM core network as possible. UMTS phones are not meant to be backward compatible with GSM systems. However, subscriptions (SIM cards) can be, and dual-mode phones will hopefully solve the compatibility problems. UMTS has two flavours: Frequency Division Duplex (FDD) (which will be implemented first) and Time Division Duplex (TDD).

After providing millions of customers with mobile access, the next primary needs that need to be satisfied are mobility and higher data-rate transmissions. There are several systems that are candidates for 3G [130, 131]. They can be grouped, based on their basic technology, as wideband CDMA, advanced TDMA, hybrid CDMA/TDMA, and Orthogonal Frequency Division Multiplexing (OFDM)

3.3 3rd Generation Partnership Project (3GPP)

The 3rd Generation Partnership Project (3GPP) develops specifications for a 3G system based on the UTRA (Universal Terrestrial Radio Access) radio interface and the enhanced GSM core network [132, 133]. The main objectives are to provide GSM with higher bit-rates, add different QoS classes for packet data, and contribute simultaneous usage of both circuit- and packet-switched services.

3GPP plans to provide backward compatibility with GSM and General Packet Radio System (GPRS) [130]. With backward compatibility at the signalling level and radio interface is next the expectations of packet-switched services to change more towards IP (Internet Protocol) communications, 3G networks have to evolve to meet the challenges. Furthermore, it is anticipated that media consumption via mobile networks will become a significant contributor to the traffic of the networks. The new usage patterns of mobile communications lead to an always-on society, in which most, if not all, users are continuously online to access their favourite media at all times without any delay. 3GPP incorporates two modes, FDD and TDD.

In the FDD mode, the uplink and the downlink use separate frequency bands. A bandwidth of 5 MHz is divided into 10 ms radio frames, and each frame is further divided into 15 time slots. The chip rate of UTRAN is 3.84 Mcps. Each user has a unique sequence of chips, called the spreading code, which modulates the data signal. The ratio of the chip rate and the data rate is called the spreading factor. The spreading factor used in UTRAN can vary from 4 to 512.

In the TDD mode, the uplink and the downlink use the same frequency carrier. The 15 time slots in a frame can be dynamically allocated between uplink and downlink directions; thus, the channel capacity of these links can be different.

Because GPRS is a packet-switched service, it is viable to have both technologies interact separately on the radio interface. This enables the service provider to incorporate both systems into a common network, without much change in hardware. Since 3G UTRAN uses a dual-system protocol stack, the main protocols like RLC/MAC in GSM are not the same as those used in GPRS. Conversely, core network protocols like MM (Mobility Management) and CM (Connection Management) are similar and can be reused.

The logical architecture shown in Figure 3-1 is designed to separate the CS (Circuit Switched) system from the PS (Packet Switched) system, along with providing interconnection between them through the RAN (Radio Access Network). The RAN is comprised of several node Bs that have one-to-many mapping, which means each node B can be connected to only one RNC (Radio Network Controller). Meanwhile, one RNC can manage various node Bs. To maintain communication during a soft handover, when the UE (User Equipment) moves from the coverage of a node B of one RNC to the node B of a different RNC, a vertical connection between RNCs is needed. The RAN for both CS (Circuit Switched) and PS (Packet Switched) systems is interconnected through interfaces.

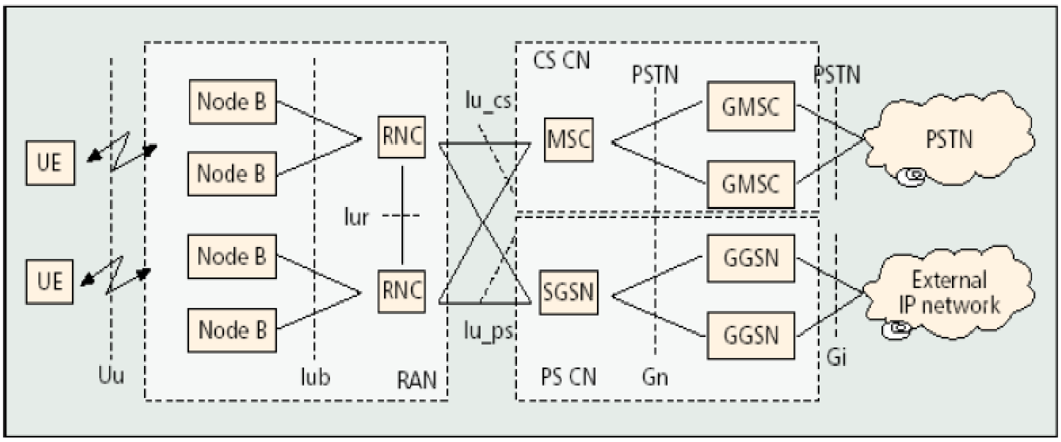

Figure 3-1: 3GPP logical architecture

3.4 Network simulation

Simulation modelling is becoming an increasingly popular method for network performance analysis. Generally, there are two forms of network simulation: analytical modelling and computer simulation. Analytical modelling is conducted by a mathematical analysis that characterizes a network as a set of equations. The main disadvantage is its overly simplistic view of the network and inability to simulate the dynamic nature of a network. Thus, the study of a complex system always requires a discrete-event simulation package, which can compute the time that would be associated with real events in a real-life situation. A software simulator is a valuable tool, especially for today's network with complex architectures and topologies. Designers can test their new ideas and carry out performance-related studies, thus freeing themselves from the burden of "trial and error" hardware implementations.

A typical network simulator can provide the programmer with the abstraction of multiple threads of control and inter-thread communication. Functions and protocols are described either by finite-state machine, native programming code, or a combination of the two. A simulator typically comes with a set of predefined modules and a user-friendly GUI. Some network simulators even provide extensive support for visualization and animation.

There are also emulators such as the NIST Network Emulation Tool (NIST Net) [134]. By operating at the IP level, it can emulate the critical end-to-end performance characteristics imposed by various wide-area network situations or underlying sub-network technologies in a lab test-bed environment (NIST NET Homepage). The academic simulators used in this thesis are J-sim and OPNET.

3.4.1 J-sim simulator features

J-sim is a free-licence simulator. It has advantages and disadvantages like any other simulator, but J-sim was chosen for these main reasons:

- J-sim has been known as JavaSim and is an open-source, component-based, compositional network simulation environment.
- J-sim is written purely in Java. At the moment, Java is one of the most widespread and well-known programming languages. Its runtime environments and compilers are available free of charge for most widely used platforms. Java is easy to learn and easy to use.
- Java pre-compiled code is interpreted in the target environment; therefore, both source texts and pre-compiled code are portable. In case of any problems, the source texts provided with J-Sim can be used to

generate new code compiled in the target environment and, therefore, completely compatible with JVM (Java Virtual Machine).

- Java provides a class called Thread, whose instances run parallel with other such instances. Thread support is built directly into the language. Therefore, no additional library is necessary, unlike with the C programming language. Moreover, an efficient method of thread synchronization is provided directly in the language. Every object has its own lock, which can temporarily suspend a currently running thread and reactivate it when a wake-up signal is received from another thread.
- Java is a fully object-oriented language, providing the concepts of classes, instances, encapsulation, inheritance, and polymorphism. Unlike in C++, the use of object principles is strictly mandatory in Java.
- The Java framework is built upon the Autonomous Component Architecture (ACA); the basic entity in J-Sim is components. Ports are the only interfaces of a component to send and receive data. When data arrives at a port, an execution context (a Java thread) is created for the component to process the data.
- Components are asynchronous, in the sense that two components may process different data at the same time without synchronizing between each other. These components can be hierarchically structured. A component may be a container mechanism and consist of subcomponents. This facilitates the hierarchical modelling of complex systems [135, 136].
- J-Sim provides basic classes for simulation, process, and queue. These classes can be either directly used or extended according to a specific user's requirements.

- There are no special actions required in order to passivate or temporarily passivate a process. Two methods are provided in the process class, whose use is intuitive and easy. The user need not know any implementation details concerning suspension and reactivation.
- J-Sim provides two possibilities of running a simulation that can be dynamically switched. The first is the batch mode, sending output to the console. The second is the interactive mode, using a graphics window to control the simulation and to display simulation output. Both modes use only standard Java services, thus rendering them fully portable. However, the possibilities of the target environment may limit their uses.

3.4.1.1 Loosely coupled, autonomous-component programming model

The performance of a component is specified in terms of contracts. A contract is bound to a specific port or group of ports, and it defines the causality of data sent and received between the component that owns the port(s) and either component that is linked to the port(s). In exacting, it does not specify the components that join in the communication. Component binding is postponed until system integration time. Autonomous refers to the capability of components to handle data in independent-execution contexts [136].

3.4.1.2 Dynamic thread execution framework for real-time process-driven

In J-Sim, the simulation engine expands the WorkerPool class and monitors the activities of all WorkerThreads. The engine maintains a globally observed, virtual system time that is proportional to the real time (e.g., 1 second in virtual time = 1000 seconds in real time). When no WorkerThread is currently active, the simulation engine adjusts the virtual system time to the nearest future so that at least one WorkerThread may become active.

3.4.1.3 Implementation of a complete suite of Internet integrated/ differentiated / best effort services protocols

J-sim supports and implements a complete suite of Internet best efforts, integrated services, and differentiated services protocols.

3.4.1.4 Dual-language environment that allows auto-configuration and no-line monitoring

It is a dual-language environment in which Java produces components and a script language that become the glue or control language to integrate components at run time and to provide high-level, dynamic control. This environment facilitates fast configuration of customized simulation scenarios and online monitoring and data collection. In the current release, it has fully integrated J-Sim with a Java implementation of the Tcl interpreter (with the Tcl/Java API extension).

3.4.2 OPNET simulator

OPNET is the abbreviation of Optimized Network Engineering Tool. This tool provides a comprehensive development environment for the specification, simulation, and performance analysis of communication networks. A large range of communication systems, from a single LAN to global satellite networks, can be supported. Discrete event simulations are used as the means of analysing system performance and their behaviour [137]. The key features of OPNET are summarized in this section.

3.4.2.1 Modelling and simulation cycle

OPNET provides powerful tools to assist the user in completing three of the five phases in a design circle (i.e., the building of models, the execution of a simulation, and the analysis of the output data).

3.4.2.2 Hierarchical modelling

OPNET employs a hierarchical structure to modelling. Each level of the hierarchy describes different aspects of the complete model being simulated.

3.4.2.3 Specialized in communication networks

Detailed library models provide support for existing protocols and allow researchers and developers to either modify these existing models or develop new models of their own.

3.4.2.4 Automatic simulation generation

OPNET models can be compiled into executable code. An executable discrete-event simulation can be debugged or simply executed, resulting in output data.

3.5 Related work

Here is an overview of the prior work that has been done in different areas relevant to traffic modelling.

3.5.1 Previous work on related queuing system

Polling models have been extensively used as a performance evaluation tool for a variety of demand-based [138-142]. A good discussion about the G/M/1 queuing system and the variety of ways in which it can be analysed is given in [143-145]. In many applications, sensor data must be delivered with time constraints to make appropriate real-time actions possible [146]. Most of the current QoS provisioning protocols [95, 147-149] in wireless sensor and ad-hoc networks are based simply on end-to-end path discovery and path recovery. Likewise, most of the existing research is focused only on reliability and lacks the ability to differentiate multiple classes of traffic that have different time constraints [71, 150, 151]. Real-time applications that require strict bounds on factors such as data rate, delay, and jitter will reveal sub-optimal performance as the network load increases. Several studies have focused on finding the maximum attainable throughput and characterizing capacity delay in wireless ad-hoc networks [152-154]. Other studies have anticipated queuing models for performance evaluation of the 802.11 MAC. The authors in [155] evaluate the packet-blocking probability and MAC-queuing delay in a basic service set with N nodes by using a finite queuing model.

The M/MMGI/1/K queuing model has been used in [156] for delay analysis over a single hop in a network. The service times of the node are modelled as a Markov-Modulated, general-arrival process. The difficulty of this approach in finding an accurate parameter description lies in the Phase-Type service. In [157], the M/MMGI/1/K queuing model has been used to analyse IEEE 802.11 DFC. This work uses single-hop criterion, an extensible and flexible approach of queuing found mainly in routing and admission control applications. The authors in [158] focus on characterizing the average end-to-end delay and maximum achievable per-node throughput in random-access, MAC multi-hop wireless ad-hoc networks with stationary nodes for hierarchical networks. The authors present an analytical model that takes into account the random-packet arrival process, the extent of locality of traffic, and the back-off and collision-avoidance mechanisms of random-access MAC. They also model random-access, multi-hop wireless networks as an open G/G/1 queuing network and use the diffusion approximation to evaluate closed form expressions for the average end-to-end delay. The authors in [159] have proposed different queues for the two different types of traffic with classifiers and schedulers. Both classes can have access to bandwidth from each other. This approach is based on cost and end-to-end constraints. This work is focused on discovering a least-cost, delay-constrained path for real-time data.

3.5.2 Previous work on the real-time communication protocols

Real-time communication is a critical service for future sensor networks to provide distributed micro-sensing in physical environments. Sensor networks need novel communication protocols to support higher-level services and should also be adaptive to avoid unpredictable congestion and holes in sensor networks.

RAP [70] is a multi-layer, real-time communication architecture for sensor networks. It provides a set of convenient, high-level query and event services. It is based on novel location-addressed communication models supported by a lightweight network stake, which integrates a transport layer location addressed protocol (LAP), a geographic routing protocol, a velocity monotonic scheduling (VMS) layer, and a contention-based MAC that supports prioritization. VMS is a concept of novel packet-requested velocity that reproduces both distance and timing constraints of sensor networks. Two versions of this algorithm are implemented. The static VMS computes a fixed, requested velocity at the sender of each packet. The requested velocity is $V = dis\ (x_0, y_0, x_d, y_d)\ /\ D$, where $dis\ (x_0, y_0, x_d, y_d)$ is the geographic distance between a sender and a destination, and D is an end-to-end deadline. The requested velocity of a packet is fixed in the networks. The dynamic VMS recalculates the requested velocity of a packet upon its arrival at each intermediate node. The requested velocity is then set to $V_i = dis\ (x_i, y_i\ , x_0, y_d)/(D\text{-}T_i)$. The requested velocity of a packet will be adjusted based on it is actual velocity.

SPEED [71] is an adaptive, location-based real-time routing protocol that aims to reduce the end-to-end, deadline-miss ratio in a sensor network. SPEED is a real-time communication protocol for sensor networks. It supports soft communication based on feedback control and stateless algorithms. It also provides three types of real time communication services: uncast, multicast, and anycast. SPEED utilizes geographic locations to make localized routing decisions. In addition, it is capable of handling congestion and provides soft real-time communication, which location-based protocols do not offer. Route discovery broadcasts in reactive routing algorithms can lead to significant delays in sensor networks. SPEED maintains only immediate neighbour information. It requires neither a routing table, as in DSDV, nor per-destination states, as in AODV. SPEED does not use any information related to deadlines. However, it provides real-time guarantees by providing a uniform packet-delivery speed across the sensor network, so that the end-to-end delay of the packet is proportional to the distance between the source and destination. SPEED does not require specialized MAC support and can work with existing best MAC protocols due to the feedback control scheme that it employs. All distributed operations in SPEED are highly localized, meaning that any action invoked by a node will not affect the whole system.

MMSPEED [72] is a novel packet-delivery mechanism for wireless sensor networks to grant service differentiation and probabilistic QoS guarantees in timeliness and reliability domains. For the timeliness domain, MMSPEED provides multiple network-wide speed options so that various traffic types can dynamically choose the proper speed options. Both SPEED and MMSPEED use fixed transmission power.

RPAR [73] varies from the previously mentioned protocols in several ways. First, RPAR is the only protocol that combines power control and real-time routing to support energy-efficient, real-time communication. Moreover, RPAR allows the application to control the trade-off between energy utilization and communication delay by specifying packet deadlines. In addition, RPAR is designed to handle faulty links. RPAR also utilizes a novel neighbourhood management mechanism that is more efficient than the periodic beacons scheme adopted by LAPC, SPEED, and MMSPEED. The main aspect of this protocol is a dynamic transmission power adjustment and routing decision in order to minimize miss ratios. The transmission power has a large impact on the delivery ratio, as it improves wireless link quality and decreases the required number of transmissions to deliver a packet. However, transmitting a packet at a high power level has a side effect of decreasing throughput due to increased channel contention and interference.

Chapter 4

PROVIDING QOS GUARANTEES TO MULTIPLE CLASSES OF TRAFFIC

Recent advances in miniaturization and low power design have led to a flurry of activity in wireless sensor networks, but the introduction of real-time communication has created additional challenges in this area. The sensor node functions as a small router most of the time, spending much of its life routing packets from one node to another until the packets reach the sink. Since sensor networks support time-critical applications, it is often necessary for communication to meet real time constraints. However, research dealing with providing QoS guarantees for real-time traffic in sensor networks is still in its infancy. In this chapter, an analytical model for implementing Priority Queuing (PQ) in a sensor node to calculate the queuing delay is presented. The model is based on the M/D/1 queuing system (a special class of M/G/1 queuing systems). Here, two different classes of traffic are considered. The exact packet delay for corresponding classes is calculated. Further, the analytical results are validated through an extensive simulation study.

4.1 Introduction

Recently, the design of sensor networks has become very important, due to several civil and military applications. Emerging sensor applications include habitat monitoring, pollution detection, weather forecasting, and monitoring disasters such as earthquakes, fires, and floods. In these new areas, similar to normal IP networks, there is real-time and non-real-time traffic in the sensor networks. Each type of traffic requires a different treatment from the network to meet the objective QoS (Quality of Service).

Queuing and scheduling have a direct impact on QoS characteristics. Different types of queuing tools, such as Priority Queuing (PQ), Custom Queuing (CQ) [115], Weighted Fair Queuing (WFQ), Class-Based Weighted Fair Queuing (CBWFQ) [118] and Low Latency Queuing (LLQ), have been developed to provide different services to heterogeneous traffic classes [119].

The communication between sensor nodes in a sensor network that builds a cluster depends on a number of factors, such as communication range, number and type of sensors, and geographical location. The efficiency of the network itself depends on the sink location, which directly affects the lifetime of the sensor network. Every cluster has a sink node that is responsible for managing the sensors in the cluster. However, the sensors within a cluster communicate with the sink via short-range wireless communication links, as illustrated in Figure 4-1. The sensor nodes need elegant and uncomplicated queuing techniques, since they usually work as small routers.

This chapter presents the implementation of Priority Queuing (PQ) in a sensor node. The M/G/1 queuing system is exploited to calculate the queuing delay for two different kinds of traffic in a sensor node. This chapter also provides a numerical solution and data from extensive simulations to verify the analytical results in order to provide guaranteed QoS to different kinds of traffic in sensor networks.

The rest of this chapter is organized as follows: Section 4.2 the queuing model and expressions of expected waiting times for two different classes are presented.

Details of numerical solution and experimental setup are provided in Section 4.3, followed by conclusions in Section 4.4.

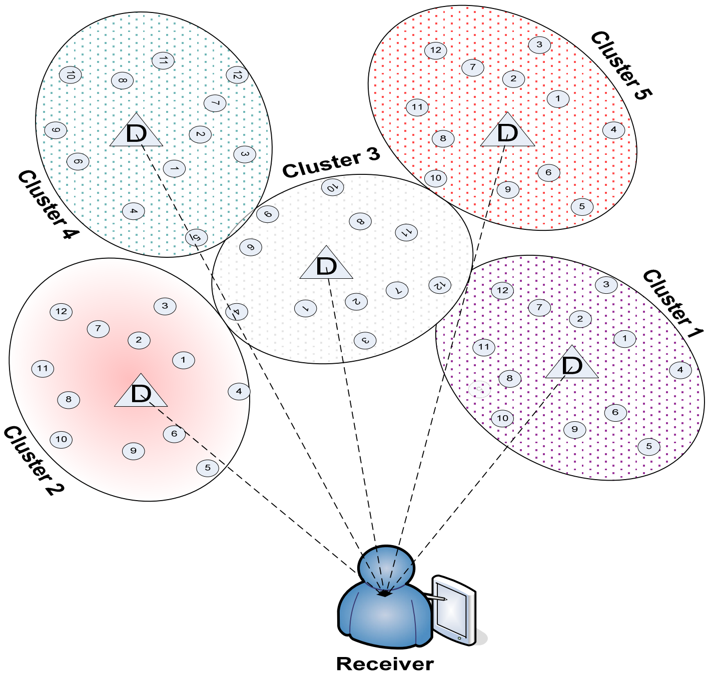

Figure 4-1 :Multi-link clustered network sensors

4.2 Queuing model

Regardless of its numerous limitations, First In, First Out (FIFO) is the default queuing algorithm in several topologies that require no configuration. Most importantly, FIFO queuing makes no decision about packet priority. FIFO queuing involves storing packets and forwarding them in order of arrival. Explode sources can cause extended delays in delivering time-sensitive application traffic and, potentially, in networking control and signalling messages. Although FIFO queuing was an effective network traffic controller before, more recent intellectual networks need more sophisticated algorithms. Furthermore, in FIFO, a full queue will cause dropping of packets – even high-priority packets. In fact, the sensor node cannot prevent this undesirable packet dropping, as it has no room in its queue.

Furthermore, FIFO cannot differentiate between a high-priority and low-priority packet.

4.2.1 Implementation of priority queuing

To overcome the limitations of the FIFO queuing discipline, Priority Queuing (PQ) is suggested as one of the applicable solutions to meet the desired QoS for real-time traffic. In this work, two queues in a sensor node are considered: high-priority and low-priority, as demonstrated in Figure 4-2. Here, the scheduler uses strict-priority logic. That is, it always serves the high-priority queue first. If there is no packet waiting in the high-priority queue, it will serve the low-priority queue. In this technique, the scheduler of the sensor node is serving different output queues

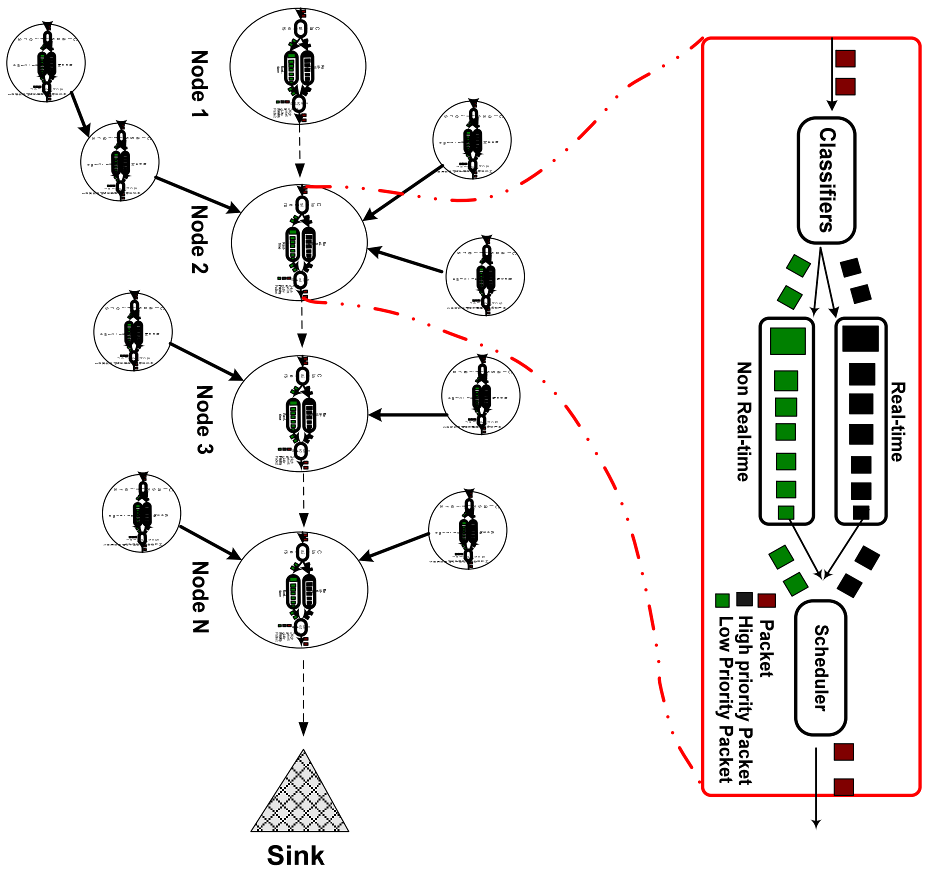

Figure 4-2 : The queuing model in sensor networks simultaneously and, hence, is behaving similarly to a multiple-queue/single-server system.

This chapter exploits the M/G/1 queuing system to model this multiple-queue/single-server system. Before explaining the formulation and notations, it is worthwhile to mention the following assumptions.

The packets that are related to high-priority queue one (Q_1) and low-priority queue two (Q_2) are called Class-1 (C_1) and Class-2 (C_2) packets with the average length of L_1 and L_2, respectively. Both C_1 and C_2 packets are travelling according to the Poisson process, with arrival rates of λ_1 and λ_2, respectively. The service times are generally distributed, and the sensor nodes and sink are all assumed to be stationary.

4.2.1.1 High-priority queue (Packets C_1)

Starting from queue one, it is assumed that the average service requirement for a C_1 packet is $E[S_1] = \dfrac{1}{\mu_1}$. The second moment of service requirement for a C_1 packet is $E[S_1^{\ 2}]$. Here, the aim is to calculate the queuing delay for each C_1 packet. Because a packet is randomly selected, its arrival time can be analysed using the PASTA property of Poisson arrival streams [160]. Here, the queuing delay is defined as the expected waiting time $E[W_1]$ in Q_1 for a C_1 packet before its being serviced.

Because Q_1 is the high-priority queue, according to strict-priority scheduler logic, the expected waiting time of C_1 packet consists of two components:

- The remaining service time of a packet in service, and,

- The time needed to serve all of the packets with the same priority (C_1) that are presented in the system at the arrival of the new, randomly selected packet.

In equation form, the expected waiting time can be written as:

$$E[W_1] = E[T_R] + E[T_1] \qquad (1)$$

where $E[T_R]$ is the expected remaining time for a packet in service when the scheduler is busy. The probability that the scheduler (server) is busy is ρ. A packet of C_1 is in service with probability $\rho_1 = \lambda_1 E[S_1]$, which is the utilization of Class-1 packets. Since the arrival time is randomly selected, the remaining service time can be viewed as that obtained for a renewal sequence consisting of generic random variables S [160]. Thus, the remaining processing time of a C_1 packet is equal to $\frac{E[S_1^{\,2}]}{2E[S_1]}$. However, because, at the arrival time of the randomly selected packet, the class (either C_1 or C_2) that is already being served is unknown, the final equation should be modified to:

$$E\left[T_R\right] = \sum_{k=1}^{2} \rho_k \frac{E\left[S_k^2\right]}{2E\left[S_k\right]} \qquad (2)$$

The second term in equation (1), $E[T_1]$, is the expected total time to serve all C_1 packets that are already waiting in Q_1 upon arrival of the randomly selected packet. Assume that the expected number of packets already waiting in queue one is $E[N_1]$. Due to the PASTA property and Little's law, on average, there are $E[N_1] = \lambda_1 E[W_1]$ C_1 packets upon arrival of this randomly selected packet [161]. Since the packets already waiting in Q_1 each require, on average, $\frac{1}{\mu_1}$ service time, $E[T_1]$ can be written as:

$$E[T_1] = \frac{E[N_1]}{\mu_1} = \frac{\lambda_1 E[W_1]}{\mu_1} = \rho_1 E[W_1] \tag{3}$$

Substituting $E[T_R]$ and $E[T_1]$ in equation (1), $E[W_1]$ can be calculated as follows a very similar expression has been given in [162] for M/G/1 with priority, as well:

$$E[W_1] = \frac{\sum_{k=1}^{2} \rho_k \frac{E[S_k^2]}{2E[S_k]}}{1-\rho_1} \tag{4}$$

The property of deterministic service time of the scheduler in a sensor node is used. Because (1) the scheduler in the sensor node needs $\frac{L_k}{R}$ time units to serve a C_k packet with the transmission rate of R, and (2) the average service requirement for a C_k packet is $E[S_k] = \frac{1}{\mu_k} = \frac{L_k}{R}$, the second moment of service requirement of a C_k packet can be expressed as follows:

$$E[S_k{}^2] = Var[S_k] + (E[S_k])^2 \tag{5}$$

In this approach, because the scheduler has a fixed service time to serve its packets and, hence, functions similarly to an M/D/1 queuing system, the service time would be deterministic with zero variance, i.e., $Var[S_k] = 0$ [163]. Thus, (5) can be simplified as:

$$E[S_k{}^2] = (E[S_k])^2 = (\frac{L_k}{R})^2$$

Based on the above assumption, the remaining service time of a C_1 packet will be $\frac{L_1}{2R}$ on average, with the probability of ρ_1 when a C_1 packet is in service. However, because, upon the arrival of the randomly selected packet, it is not clear which packet (either C_1 or C_2) is in service, equation (4) is modified as:

$$E[W_1] = \frac{\sum_{k=1}^{2} \rho_k \frac{L_k}{2R}}{1-\rho_1} = \frac{\rho_1 \frac{L_1}{2R} + \rho_2 \frac{L_2}{2R}}{1-\rho_1} \qquad (6)$$

where ρ_1 and ρ_2 are the utilizations caused by C_1 or C_2 packets with average lengths of L_1 and L_2, respectively.

4.2.1.2 Low -priority queue (Packets C_2)

We obtain the expected waiting time for a randomly selected C_2 packet arrives to the low-priority queue by analysing the events that constitute this delay. The amount of work in the system at any time is defined as the (random) sum of all service times that will be required by the packets in the system at that instant. The waiting time of a C_2 packet (which is the low-priority queue) can be written as:

$$E[W_2] = E[Z_1] + E[Z_2] + E[Z_3] + ... \qquad (7)$$

where $E[Z_1]$ is the expected amount of work seen by the arriving C_2 packet in Q_1 and Q_2 (i.e., higher priority and equal priority), plus the work needed to finish the service of a packet, which is already in service (if any). $E[Z_1]$ can be further written as:

$$E[Z_1] = E[T_2] + E[T_R] \qquad (8)$$

$E[T_R]$ is the remaining service time of the packet in service (if any), which can be calculated in the same way as it was for Q1, and $E[T_2]$ is the time needed to serve all of the packets of the higher priority class C_1 and equal priority class C_2 upon the arrival of the randomly selected C_2 packet. $E[T_2]$ is related to the number of packets per class in both queues (Q_1 and Q_2) upon arrival of the C_2 packet. Referring to the PASTA property and Little's law, there are $E[N_k] = \lambda_k E[W_k]\, C_k$ packets, on average, upon arrival of a new C_2 packet. Since each requires $\frac{1}{\mu_k}$ service time, on average, $E[T_2]$ can be calculated as follows:

$$E[T_2] = \sum_{k=1}^{2} \frac{E[N_k]}{\mu_k} = \sum_{k=1}^{2} \frac{\lambda_k E[W_k]}{\mu_k} = \sum_{k=1}^{2} \rho_k E[W_k]$$

$$= \rho_1 E[W_1] + \rho_2 E[W_2]$$

By combining the values of $E[T_R]$ and $E[T_2]$, we can write equation (8) as follows:

$$E[Z_1] = \sum_{k=1}^{2} \rho_k \frac{L_k}{2R} + \rho_1 E[W_1] + \rho_2 E[W_2]$$

Now, $E[Z_2]$ is the expected amount of work associated with higher-priority C_1 packets arriving during $E[Z_1]$, $E[Z_3]$ is the expected amount of work associated with C_1 packets arriving during $E[Z_2]$, and so on. As illustrated in Figure 4-3 , the waiting time of an arriving packet of C_2 is indeed given by the total workload building in front of it [164]. The arrows in the figure denote the arrival times of C_1 packets, and all of the oblique lines have 45 degrees angle with the time axis. In this figure, the expected waiting time is:

$$E[W_2] = E[Z_1] + E[Z_2] + E[Z_3] + E[Z_4]$$

For example, let M_j denote the number of type j arrivals over Z_i, j=1, 2, etc. Then:

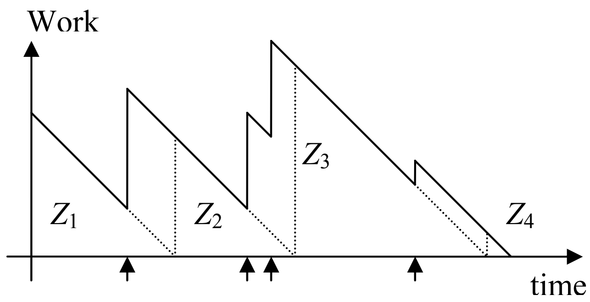

Figure 4-3 : Waiting time of a type-2 packet in terms of Z_j's.

$$W_2 = Z_1 + S_1^{M_1} + S_1^{M_2} + \cdots$$

where $S_1^{M_j}$ denotes the random sum of M_j independent service times of C_1 packets. Then,

$$E[W_2] = E[Z_1] + E[S_1]E[M_1] + E[S_1]E[M_2] + \cdots$$

since the service times and the arrival process are independent. For a stationary packet arrival process, this can be:

$$E[M_j] = E[E[M_j \mid Z_j]] = E[c_1 Z_j] = c_1 E[Z_j]$$

due to the aforementioned independence, where $c_1 > 0$ is a constant, particular to the arrival process. That is, the expectation of the number of arrivals in any period of time is proportional to the length of that period, because of it being stationary in time and linear in expectation. In our stationary Poisson traffic input process, C_1 is the expected number of arrivals per unit of time of C_1 packet (which can be called the arrival rate), each requiring $\frac{1}{\mu_1}$ service time. Hence, the expected waiting time reduces to:

$$E[W_2] = E[Z_1] + E[S_1]c_1E[Z_1] + E[S_1]c_1E[Z_2] + \cdots$$

$$= E[Z_1] + \frac{c_1}{\mu_1}(E[Z_1] + E[Z_2] + \cdots)$$

$$= E[Z_1] + \frac{c_1}{\mu_1}E[W_2]$$

In other words, during $E[W_2]$ time units, the C_2 packet has to wait; $\lambda_1 E[W_2]$ packets of C_1 arrive, on average, requiring $\frac{1}{\mu_1}$ service time. Hence, $\frac{c_1}{\mu_1}E[W_2]$ can be written:

$$\frac{\lambda_1 E[W_2]}{\mu_1} = \rho_1 E[W_2]$$

Substituting all of the values in equation (7), $E[W_2]$ can be calculated as follows:

$$E[W_2] = \sum_{k=1}^{2} \rho_k \frac{L_k}{2R} + \rho_1 E[W_1] + \rho_2 E[W_2] + \rho_1 E[W_2]$$

Bringing $E[W_2]$ to one side and simplifying allows it to be written as:

$$E[W_2] = \frac{\sum_{k=1}^{2} \rho_k \frac{L_k}{2R} + \rho_1 E[W_1]}{1 - \rho_1 - \rho_2} \qquad (9)$$

4.3 Experimentation and evaluation

A sample network was generated to show the performance of the presented model. Then, its analytical results were validated through simulation.

4.3.1 Environmental setup

A square-shaped network (200m, 200m) was selected as the benchmark to validate the accuracy of the presented formulations. One hundred nodes with 40m radio ranges were generated to operate in this network. For each node, a free-space propagation channel model is assumed, with a transmission speed of 250kbps and total packet length of 40 bytes for both C_1 and C_2 with a 15-packet capacity as their buffer sizes. Also, for each node in the sensing state, packets are generated at a constant rate of 1 packet/sec. The real-time packet generation rate is 3 packets/sec to prevent the nodes from being congested and/or overloaded [165-167].

4.3.2 Results

The results have been validated in this chapter by comparing the result obtained from the equations (numerical result) and the results obtained from the simulation (simulation results). The mathematical delay calculation has been explained in the numerical result section. The simulation technique has been clarified in the simulation result section. The result has been further broken down in the analysis section.

4.3.2.1 Numerical results, two-queues priority model

The total arrival rate for each sensor node is assumed to be 78 packets per second, with a 32Kbit packet length. The scheduler of the sensor can serve 780 packets per second. Therefore, the total utilization of the sensor node will be 0.1. Next, the utilization must be calculated for both of the queues (Q_1 and Q_2). For Q_1, real-time packets are arriving at a rate of four times as many as the non–real-time packets, which are arriving to Q_2. If the arrival rate to the real-time queue is four times more than to the non-real-time queue, the utilization of Q_1 will be 0.75 and the utilization of Q_2 will be 0.25. Using these values in equations (6) and (9) and solving them simultaneously, the expected delays for Q_1 and Q_2 would be 0.069189 ms and 0.076876 ms, respectively. Next, the arrival rate to a sensor node has been increased from 78pps to 156pps, which causes the utilization of the system to become 0.2 and the expected delay for Q_1 and Q_2 to become 0.150588 ms and 0.188235 ms, respectively. Similarly, the gradual increase in the arrival rate to a sensor node to 234pps and then to 312pps causes the utilization of the system to become 0.3 and 0.4, respectively, as shown in Table 4-1.

Table 4-1: Expected Delay for Q_1 and Q_2 packets

Utilization	Delay (Q1)	Delay (Q2)
0.1	0.069189 ms	0.076876 ms
0.2	0.150588 ms	0.188235 ms
0.3	0.247700 ms	0.353800 ms
0.4	0.365700 ms	0.914200 ms

A sample network has been generated to show the performance of the presented model. Then, its analytical results have been validated through simulation.

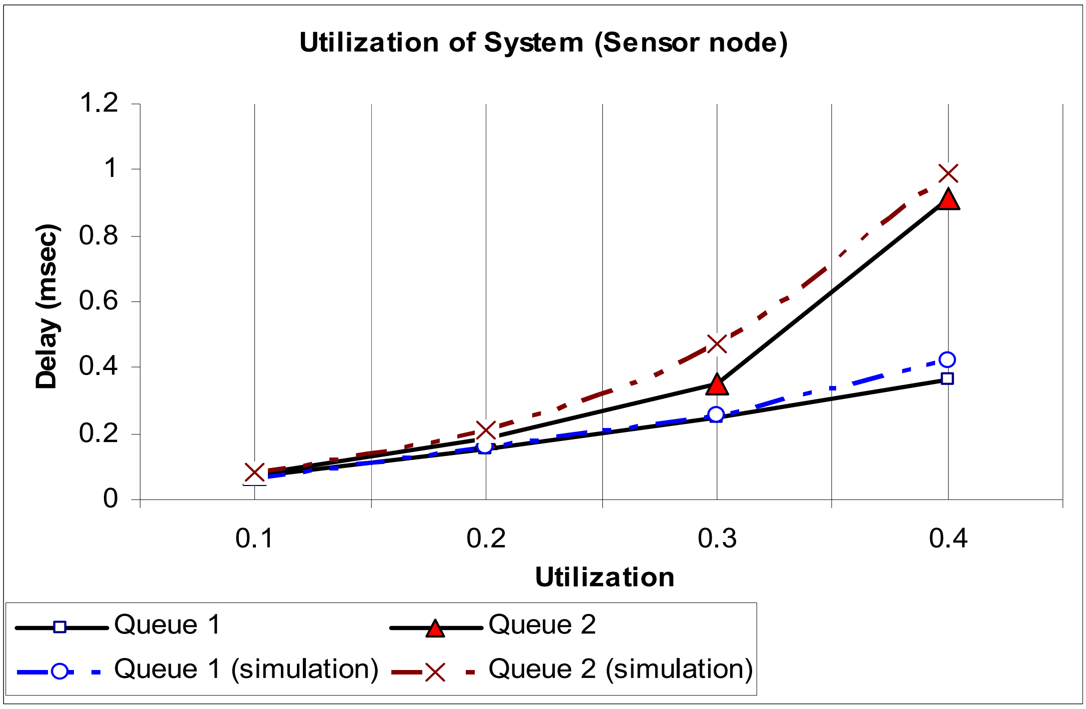

Figure 4-4: Utilization of system (sensor node)

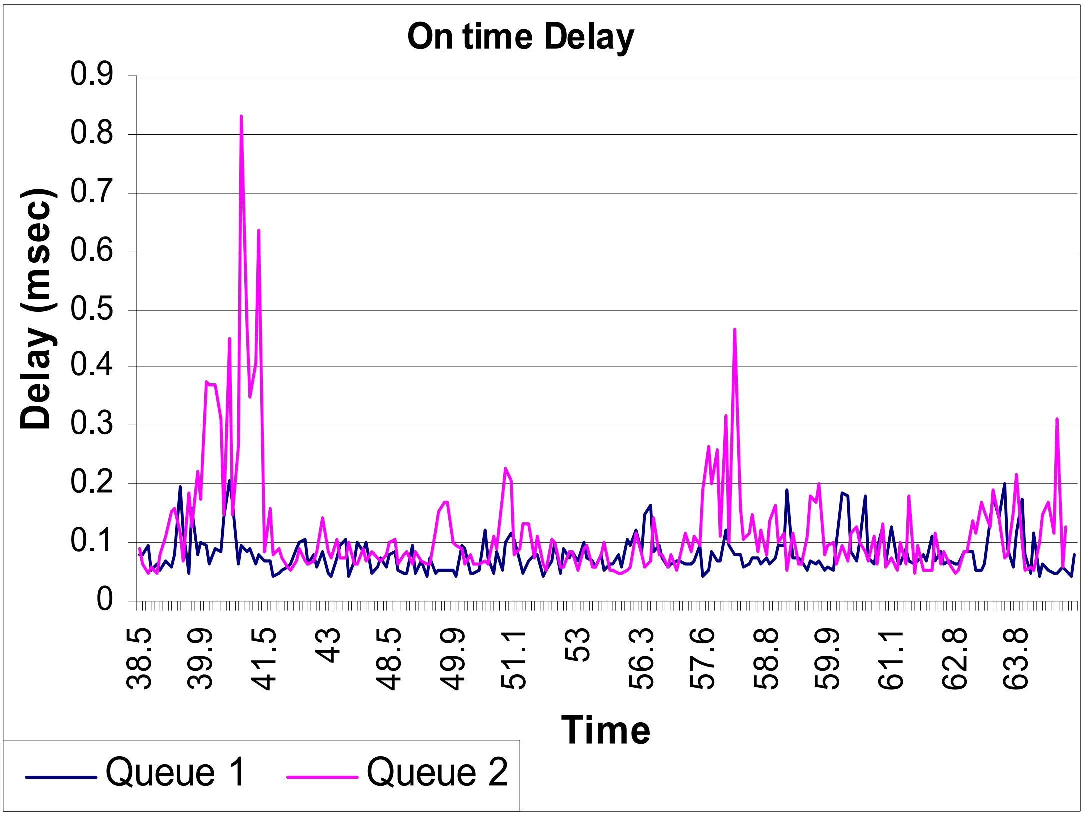

Figure 4-5: On-time delay

4.3.3 Simulation results

J-Sim has been used to simulate 100 nodes; the software provides a high-fidelity simulation for wireless communication with detailed propagation, radio, and MAC layers. The delay has been calculated for real-time traffic (Q_1) and non-real-time traffic (Q_2) on node 5, which has been selected at random out of 100 nodes during the 15 randomized runs of the simulation and shows the result and the delay characteristics of sensor nodes on the simulation.

4.3.4 Discussion and analysis

Table 4-1 shows the differences in delays for Q_1 and Q_2 for several utilizations of the sensor nodes. As can be seen, Q_2 delays exceed Q_1 delays in all positions. It is clear that Q_2 delays at the utilization of the sensor node system significantly increase, which causes an increase in expected delays. However, it must be noted that there are significant increases that started from utilization 0.3.

The horizontal axis in Figure 4-4 shows the variable of arrival rates in units of utilization, while the vertical axis shows the range of the variable of delay in units of milliseconds. Thus, the graph shows the change in delays over the arrival rates of the packets for both the numerical and simulation results. The straight lines are used to show the numerical results of Q_1 and Q_2, and the dotted lines are used to illustrate the simulation results.

The slight differences between the numerical and simulation results in both Q_1 and Q_2 are because of factors such as the retransmission of some packets and the time it takes to find a new path to forward the packet.

Figure 4-5 shows the on-time delay. The horizontal axis shows the variable of arrival rates in units of seconds, and the vertical axis shows the range of the variable of delay in units of milliseconds. Thus, the graph shows the change in delays over the arrival rates of the packets for node 5 on the simulation.

The on-time delay from 38.5s to 63.8s in Q_1 and Q_2 shows that the delay has increased and dropped periodically. The reason for the difference between the delays in Q_1 and Q_2 is because some packets have been dropped, and other packets chose different paths and did not pass though node 5. However, because each sensor node has a small buffer and the delay inside the node is gradually increasing, as shown in Figure 4-4 and Figure 4-5, the probability of dropping the packets would increase accordingly.

4.4 Conclusion

This chapter presented the closed-form expressions of the queuing delay for multiple (real-time and non-real-time) classes of traffic in a senor node through the implementation of priority queuing based on the M/G/1 queuing system. The analytical results have been verified through numerical and simulation studies. The results provide a way to analyse the performance of priority-queuing implementation in a sensor node. This measurement data can be useful as inputs to a simulation study of sensor networks. Also, the analytical modelling technique and its verification through numerical and simulation results is the first step towards finding out the most appropriate queuing-scheme implementation for wireless sensor networks.

Chapter 5

GUARANTEED QOS THROUGH LIMITED-SERVICE POLLING MODELS

Data gathering in a timely and reliable fashion has been a key concern in wireless sensor networks, particularly as related to military applications. The introduction of real-time communication has created additional challenges in this area, due to different communication constraints. Since sensor networks represent a new generation of time-critical applications, it is often necessary for communication to meet real time constraints. However, research dealing with providing QoS guarantees for real-time traffic in sensor networks is still immature. To provide guaranteed QoS in wireless sensor networks, this chapter presents a novel analytical model based on a limited-service polling discipline. The proposed model implements two queues in a sensor node that are being served according to round-robin service. The model is based on the M/D/1 queuing system (a special class of M/G/1 queuing systems), which takes into account two different classes of traffic in a sensor node. The exact queuing delay in a sensor node for corresponding classes is calculated. Further, the analytical results are validated through an extensive simulation study.

5.1 Introduction

Micro electro-mechanical systems (MEMS), digital electronics, and wireless communication have enabled the development of a new generation of large-scale sensor networks, in which small nodes communicate with each other over short distances with low-power consumption. These networks are suitable for a wide range of applications [4-6, 168].

Recently, the design of sensor networks has become more important, due to several civil and military applications. Emerging sensor applications include habitat monitoring, pollution detection, weather forecasting, and monitoring disasters such as earthquakes, fires, and floods. Just like a normal IP network, a wireless sensor networks also has real-time and non-real-time applications, each requiring a different kind of treatment from the network in terms of QoS. The communication between sensor nodes in a sensor network that builds a cluster depends on a number of factors, such as communication range, number and type of sensors, and geographic location. The efficiency of the network itself depends on the sink location, which directly affects the lifetime of the sensor network. Every cluster has a sink node that is responsible for managing the sensors in the cluster. The sensors within a cluster communicate with the sink via short-range wireless communication. The sensor nodes need elegant and uncomplicated queuing techniques, as they usually work as small routers. In many applications, sensor data must be delivered with time constraints to make appropriate real-time actions possible [146].

The rest of this chapter is organized as follows. In Section 5.2, the background of queuing models has been given, along with this chapter's proposed analytical model. A brief introduction of polling models and devoted to the equations of expected waiting time for different classes of traffic in 1-limited and k-limited polling models, is offered in Section 5.3. The numerical analysis is presented in Section 5.4. The simulation results are given in section 5.5. the discussion and analysis are given in Sections 5.6. Finally, a conclusion and future work is provided in Section 5.7.

5.2 Queuing models

Queuing and scheduling have a direct impact on QoS characteristics. There are different types of queuing tools, which have been developed to provide differential service to heterogeneous traffic classes, such as Priority Queuing (PQ), Custom Queuing (CQ), Class-Based Weighted Fair Queuing (CBWFQ) and Low Latency Queuing (LLQ) [169].

Regardless of its numerous limitations, First In, First Out (FIFO) is the default queuing algorithm in several topologies that require no configuration. Most importantly, FIFO queuing makes no decision about packet priority. FIFO queuing involves storing packets and forwarding them in order of arrival. Explode sources can cause extended delays in delivering the packets of time-sensitive applications. Although FIFO queuing seems to be an effective network traffic controller, more recent intellectual networks need more sophisticated algorithms. Furthermore, in FIFO, a full queue will cause dropping of packets – even high-priority packets, because FIFO cannot differentiate between a high-priority and a low-priority packet. In fact, the sensor node cannot even prevent this undesirable packet dropping, as it has no room for them in its queue.

This chapter presents a model of two queues in a sensor node. The model is based on the M/G/1 queuing system, which takes into account two different classes of traffic – real time and non-real time. The system is analysed on the basis of a limited-service polling model, first according to a 1-limited polling service, and then on the basis of a k-limited polling service, and derives closed-form expressions of packet delay for corresponding traffic classes. The study of polling models is important, since it gives very good insight into the qualitative behaviour of many proposed and implemented queuing disciplines and forms the basis to derive closed-form expressions of different QoS parameters, such as delay, jitter, and throughput.

Custom Queuing (CQ) is a queuing tool in which the output scheduler serves output queues in a round-robin fashion by taking a specified number of bytes from each queue [170]. CQ schedulers take the packets from each queue until either the queue is emptied or a specified number of bytes (packets) are served, whichever occurs first. The performance of a CQ scheduler can be analysed through the limited-service polling model. In this technique, the scheduler of the sensor node is serving different output queues simultaneously, and hence, behaves similarly to a multiple-queue/single-server system.

5.3 Polling model for sensor network

Over the last two decades, the polling model has been studied extensively to analyse the performance of Local Area Networks (LANs) employing different forms of token passing [125]. The basic polling model is a queuing model composed of a set of queues and a single server that serves the queues in a cyclical order [171]. Generally, polling models can be classified as exhaustive, gated, or limited service. The exact details of the systems are beyond the scope of this chapter. Instead, readers are referred to [126, 138] for a detailed discussion of polling systems. This chapter explains only the limited-service polling model, because it is related to our system. In the limited-service system, a queue is served until either:

- The buffer is emptied, or

- A specified number of packets are served, whichever occurs first. If, at most, k packets are served in one cycle, it is referred to as a k-limited polling model.

The case in which k = 1 results in a simpler model, referred to as a 1-limited polling model, in which the server serves one packet from each queue in an alternating fashion during each cycle.

Before explaining the formulation and notations, it is worthwhile to mention the following assumptions. The packets that are related to queue one (Q_1) and queue two (Q_2) are called Class 1 (C_1) and Class 2 (C_2) packets, with an average length of L_1 and L_2, respectively. Both C_1 and C_2 packets are travelling according to the Poisson process, with arrival rates of λ_1 and λ_2, respectively. The service times are generally distributed, and the sensor nodes and the sink are all assumed to be stationary.

5.3.1 1-limited polling model for sensor networks

This chapter considers a model of two queues (Q1 and Q2) in a sensor node that are being served according to a round-robin scheduling discipline. The system is analysed on the basis of a 1-limited polling model. In this type of model, since the scheduler serves one packet from Q1 and one packet from Q2 during each cycle, the expected delay for C_1 and C_2 packets will be the same due to the symmetry of service. Q1 only is studied in detail.

Starting from Q_1, we assume that the average service requirement for a C_1 packet is $E[S_1] = \frac{1}{\mu_1}$. The second moment of service requirement for a C_1 packet is $E[S_1^{\ 2}]$.

The interest is in finding out the queuing delay for a randomly selected packet arriving to queue 1. Because a packet is randomly selected, its arrival time can be analysed using the PASTA property of Poisson arrival streams [160]. Here, the queuing delay is defined as the expected waiting time $E[W_1]$ in Q_1 for a C_1 packet before its being serviced. An arriving packet of C_1 will wait for the completion of the packet already in service, plus the service times of packets in Q_1 and Q_2, according to round-robin scheduling based on 1-limited polling. If $E[N_1]$ is the expected number of packets already waiting in Q_1, then the new arriving packet of C_1 waits $R_1 + S_1^{N_1-1} + S_2^{N_1}$ time units in the queue if m =1, and $R_2 + S_1^{N_1} + S_2^{N_1}$ time units if m=2; where R_m denotes the remaining service time of a packet in service. In other words, the expected waiting time of C_1 packet consists of three components:

The remaining service time of a packet in service, if any

The time it takes to serve all C_1 packets, i.e., $E[N_1]$, that are present in Q_1 at the arrival of this randomly selected packet

The time it takes to serve $(E[N_1])$ number of packets from Q_2

In equation form, this information can be written:

$$E[W_1] = E[R_m] + E[T_1] + E[T_2] \qquad (1)$$

where $E[R_m]$ is the expected remaining time for a packet in service when the scheduler is busy. The probability that the scheduler (server) is busy is ρ. A packet of C_1 is in service with probability $\rho_1 = \lambda_1 E[S_1]$, which is the utilization of C_1 packets. Since the arrival time is randomly selected, the remaining service time can be viewed as that obtained for a renewal sequence consisting of generic random variables S [160]. Thus, the remaining processing time of a C_1 packet is equal to $\frac{E[S_1{}^2]}{2E[S_1]}$. However, it is not known, upon the arrival of this randomly selected packet, whether the packet already in service belongs to C_1 or C_2. Thus, the first term in equation (1) can be written as:

$$E\left[R_m\right] = \sum_{k=1}^{2} \rho_k \frac{E\left[S_k^2\right]}{2E\left[S_k\right]} \qquad (2)$$

The second term in equation (1), $E[T_1]$, is the expected total time to serve all C_1 packets that are already waiting in Q_1 upon arrival of the randomly selected packet. Assume that the expected number of packets already waiting in Q_1 is $E[N_1]$. Due to the PASTA property and Little's law, on average there are $E[N_1] = \lambda_1 E[W_1]$ class 1 (C_1) packets upon arrival of this randomly selected packet [161]. Since the packets already waiting in Q_1 each require, on average, $\frac{1}{\mu_1}$ service time, $E[T_1]$ can be written as:

$$E[T_1] = \frac{E[N_1]}{\mu_1} = \frac{\lambda_1 E[W_1]}{\mu_1} = \rho_1 E[W_1] \qquad (3)$$

It is known that, in the 1-limited polling model, the scheduler serves one packet from queue 1 and one packet from queue 2 in an alternating fashion during each cycle. Hence, the third term in equation (1), $E[T_2]$, is the expected total time to serve $(E[N_1])$ number of packets from Q_2 that are already waiting in Q_2 upon arrival of the randomly selected packet. It is assumed that the expected number of packets already waiting in Q_1 is $E[N_1]$. Due to the PASTA property and Little's law, on average, there are $E[N_1] = \lambda_1 E[W_1]$ class 1 (C_1) packets upon arrival of this randomly selected packet [161]. Since the packets already waiting in Q_2 each require, on average, $\frac{1}{\mu_2}$ service time, $E[T_2]$ can be written as:

$$E[T_2] = \frac{E[N_1]}{\mu_2} = \frac{\lambda_1 E[W_1]}{\mu_2}$$

We further assume that $\mu_1 = \mu_2 = \mu$, so the above equation can be written as:

$$E[T_2] = \frac{\lambda_1 E[W_1]}{\mu_1}$$

Substituting $E[R_m]$, $E[T_1]$, and $E[T_2]$ in equation (1), $E[W_1]$ can be calculated as follows:

$$E[W_1] = \sum_{k=1}^{2} \rho_k \frac{E[S_k^2]}{E[S_k]} + \frac{\lambda_1 E[W_1]}{\mu_1} + \frac{\lambda_1 E[W_1]}{\mu_1}$$

$$E[W_1] = \sum_{k=1}^{2} \rho_k \frac{E[S_k^2]}{E[S_k]} + 2\frac{\lambda_1 E[W_1]}{\mu_1}$$

$$E[W_1] = \sum_{k=1}^{2} \rho_k \frac{E[S_k^2]}{E[S_k]} + 2\rho_1 E[W_1] \qquad (4)$$

Next, the property of the deterministic service time of scheduler in a sensor node is used. Since (1) the scheduler in the sensor node needs $\frac{L_k}{R}$ time units to serve a C_k packet with the transmission rate of R, and (2) the average service requirement for a C_k packet is $E[S_k] = \frac{1}{\mu_k} = \frac{L_k}{R}$, the second moment of service requirement of a C_k packet can be expressed as follows:

$$E[S_k{}^2] = Var[S_k] + (E[S_k])^2 \qquad (5)$$

In this approach, because the scheduler has a fixed service time to serve its packets and, hence, functions similarly to an M/D/1 queuing system, the service time would be deterministic with zero variance, i.e., $Var[S_k] = 0$ [172]. Thus, (5) can be simplified as: $E[S_k{}^2] = (E[S_k])^2 = (\frac{L_k}{R})^2$. Based on the above assumption, the remaining service time of a C_1 packet will be $\frac{L_1}{2R}$, on average, with the probability of ρ_1 when a C_1 packet is in service. However, because, upon arrival of the randomly selected packet, it is not clear which packet (either C_1 or C_2) is in service, equation (4) is modified as:

$$E[W_1] = \frac{\sum_{k=1}^{2} \rho_k \frac{L_k}{2R}}{1 - 2\rho_1} = \frac{\rho_1 \frac{L_1}{2R} + \rho_2 \frac{L_2}{2R}}{1 - 2\rho_1} \qquad (6)$$

where ρ_1 and ρ_2 are the utilizations caused by class 1 and class 2 packets with average lengths of L_1 and L_2, respectively.

5.3.2 K-limited polling model for sensor networks

This research examines a model of two queues (Q_1 and Q_2) in a sensor node by considering two different classes of traffic input, i.e., real time and non-real time. The two queues are being served according to a round-robin scheduling discipline.

The system is analysed on the basis of a k-limited polling model. The scheduler logic is specified in such a way that the scheduler serves two packets from Q_1 and one packet from Q_2 during each cycle. Therefore, Q_1 can be called a 2-limited queue

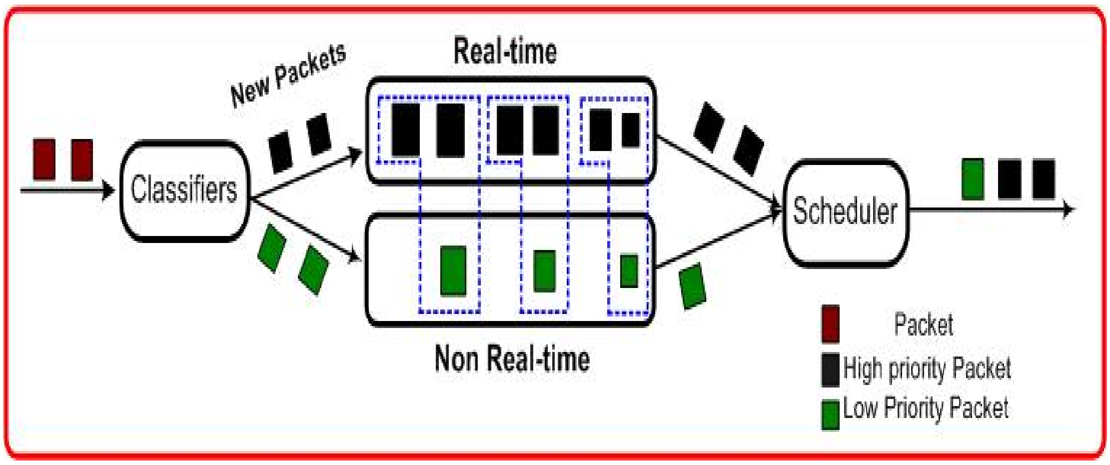

Figure 5-1: k-limited polling model in a sensor node

And Q_2 can be called a 1-limited queue. The implementation of a k-limited polling model in a sensor node is shown in Figure 5-1. Here, the aim is to calculate the queuing delay for C1 and C2 packets. Because a packet is randomly selected, its arrival time can be analysed using the PASTA property of Poisson arrival streams [160]. In this example, the queuing delay is defined as the expected waiting time $E[W_1]$ in Q1 for a C1 packet before it is serviced. An arriving packet of C1 will wait for the completion of the packet already in service, plus the service times of packets in Q1 and Q2, according to round-robin scheduling based on k-limited polling. If $E[N_1]$ is the expected number of packets already waiting in Q1, then the newly arriving packet of C1 waits $R_m + S_1^{N_1-1} + S_2^{\frac{N_1}{2}}$ time units in the queue if m =1, and $R_m + S_1^{N_1} + S_2^{\frac{N_1}{2}}$ units if m=2, where R_m denotes the remaining service time of a packet in service. In other words, the expected waiting time of a C1 packet consists of three components:

- The remaining service time of a packet in service, if any

- The time it takes to serve all C1 packets, i.e., $E[N_1]$, that are present in Q1 upon the arrival of this randomly selected packet
- The time it takes to serve $(E[\frac{N_1}{2}])$ number of packets from Q2

The equation can be written as:

$$E[W_1] = E[R_m] + E[T_1] + E[T_2] \qquad (7)$$

Where $E[R_m]$ is the expected remaining time for a packet in service when the scheduler is busy. The second term in equation (1), $E[T_1]$, is the expected total time to serve all C_1 packets that are already waiting in Q_1 upon arrival of the randomly selected packet. The first two terms, $E[R_m]$ and $E[T_1]$, can be derived the same way they were for the 1-limited polling model. Skipping the details, this chapter will proceed to derive $E[T_2]$, which is the expected total time to serve $(E[\frac{N_1}{2}])$ number of packets from Q_2 that are already waiting in Q_2 upon arrival of the randomly selected packet. Assume that the expected number of packets already waiting in Q_1 is $E[N_1]$. Due to the PASTA property and Little's law, on average, there are $E[N_1] = \lambda_1 E[W_1]$ class 1 (C_1) packets upon arrival of this randomly selected packet [161]. Since the packets already waiting in Q_2 each require, on average, $\frac{1}{\mu_2}$ service time, $E[T_2]$ can be written as:

$$E[T_2] = \frac{E[N_1]/2}{\mu_2} = \frac{(\lambda_1 E[W_1])/2}{\mu_2}$$

We further assume that $\mu_1 = \mu_2 = \mu$, so the above equation can be written as:

$$E[T_2] = \frac{(\lambda_1 E[W_1])/2}{\mu_1}$$

Substituting $E[R_m], E[T_1]$, and $E[T_2]$ in equation (7), $E[W_1]$ can be calculated as follows:

$$E[W_1] = \sum_{k=1}^{2} \rho_k \frac{E[S_k^2]}{E[S_k]} + \frac{\lambda_1 E[W_1]}{\mu_1} + \frac{\lambda_1 E[W_1]}{2\mu_1} \qquad (8)$$

Next, the property of deterministic service time of a scheduler in a sensor node is again used. Skipping the details, the above equation can be reduced as follows:

$$E[W_1] = \frac{\sum_{k=1}^{2} \rho_k \frac{L_k}{2R}}{(2-3\rho_1)/2} = \frac{\rho_1 \frac{L_1}{2R} + \rho_2 \frac{L_2}{2R}}{(2-3\rho_1)/2} \qquad (9)$$

where ρ_1 and ρ_2 are the utilizations caused by $C_1 and C_2$ packets with average lengths of L_1 and L_2, respectively.

The next step is to determine the queuing delay for a C_2 packet that is arriving to Q_2 (1-limited queue).An arriving packet of C_2 will wait for the completion of the packet already in service plus the service times of packets in Q_1 and Q_2 according to round-robin scheduling based on k-limited polling. If $E[N_2]$ is the expected number of packets already waiting in Q_2, then the newly arriving packet of C_2 waits $R_m + S_2^{N_2-1} + S_2^{2N_2}$ time units in the queue if m =2, and $R_m + S_2^{N_2} + S_2^{2N_2+1}$ units if m=1, where R_m denotes the remaining service time of a packet in service. In other words, the expected waiting time of a C_1 packet consists of three components:

The remaining service time of a packet in service, if any

The time it takes to serve all C_2 packets, i.e., $E[N_2]$, that are present in Q_2 at the arrival of this randomly selected packet

The time it takes to serve $(E[2N_2])$ number of packets from Q_1

The derivation procedure is very similar to that of a 2-limited queue. Again omitting the details, the final expression for the expected waiting time of a randomly selected C_2 packet arriving to Q_2, which is 1-limited queue, is as follows:

$$E[W_2] = \frac{\sum_{k=1}^{2} \rho_k \frac{L_k}{2R}}{(1-3\rho_2)} = \frac{\rho_1 \frac{L_1}{2R} + \rho_2 \frac{L_2}{2R}}{(1-3\rho_2)} \qquad (10)$$

5.4 Numerical analysis, polling model

The system is analysed based on the 1-limited service polling model with zero switch-over time. In the 1-limited polling model, the scheduler (server) serves one packet from each queue in an alternating fashion during each cycle. For this reason, the model is also called an alternating service model.

The system is also analysed based on the k-limited service polling model with zero switch-over time. In the k-limited polling model, the scheduler (server) serves k packets from queue 1 and one packet from queue 2 during each cycle. For the present case, the value of $k=2$ is adjusted, meaning that, during each cycle, the scheduler serves two packets from queue 1 and one packet from queue 2.

5.4.1 Numerical analysis, two queues 1-limited polling model

To extract the numerical solution, some assumptions are made. The total arrival rate for each sensor node is assumed to be 78 packets per second, with a 40-byte packet length. The transmission rate of the link is R = 250 kbps; hence, the scheduler will take $\frac{L}{R}$ sec. to serve one packet. By using the assumed values, the sensor can serve approximately 780 packets per second. Therefore, the total utilization of the sensor node's scheduler will be 0.1. Now the utilization of both queues (Q_1 and Q_2) has been calculated. It is known that the expected delay for C_1 and C_2 packets is the same in the 1-limited polling model, due to the symmetry of service. Using these values in equation (6), the expected delay for a randomly selected Q_1 and Q_2 packet would be 0.07111 ms. Next, the arrival rate to a sensor node has been increased from 78pps to 156pps, which causes the utilization of the system to become 0.2 and the expected delay for Q_1 and Q_2 to be 0.16 ms. Similarly the gradual increase in the arrival rate to a sensor node to 234 pps and then to 312 pps causes the utilization of the system to become 0.3 and 0.4, respectively. This eventually increases the queuing delay for C_1 and C_2 packets. The numerical results are shown in Table 5-1.

Table 5-1: Expected Delay for Q_1 and Q_2 packets in 1-limited polling model

Utilization	Delay for (Q_1) and (Q_2) in ms
0.1	0.0711
0.2	0.1600
0.3	0.2743
0.4	0.4270

5.4.2 Numerical analysis, two queues K-limited polling model

Similarly, by using the same values in equations (9) and (10), the expected delay for C_1 and C_2 packets in k-limited polling models is obtained. The numerical results are shown in Table 5-2.

Table 5-2: Expected Delay for Q1 and Q2 packets in k-limited Polling model

Utilization	Delay (Q_1) in ms	Delay (Q_2) in ms
0.1	0.0692	0.0759
0.2	0.151	0.1829
0.3	0.25	0.3500
0.4	0.366	0.6400

5.5 Experimentation and evaluation

A sample network has been generated to show the performance of the presented model. Its analytical results have been validated through simulation.

5.5.1 Environmental setup

A square-shaped network (200m, 200m) was selected as the benchmark to validate the accuracy of the presented formulations. One hundred nodes with a 40m radio range were generated to operate in this network. For each node, a free space propagation channel model is assumed, with a transmission speed of 250kbps and a total packet length of 40 bytes for both C_1 and C_2, with a 15-packet capacity as their buffer sizes. Also, for each node in the sensing state, packets are generated at a constant rate of 1 packet/sec. The real-time packet generation rate is 3 packets/sec to prevent the nodes from being congested and/or overloaded [165-167].

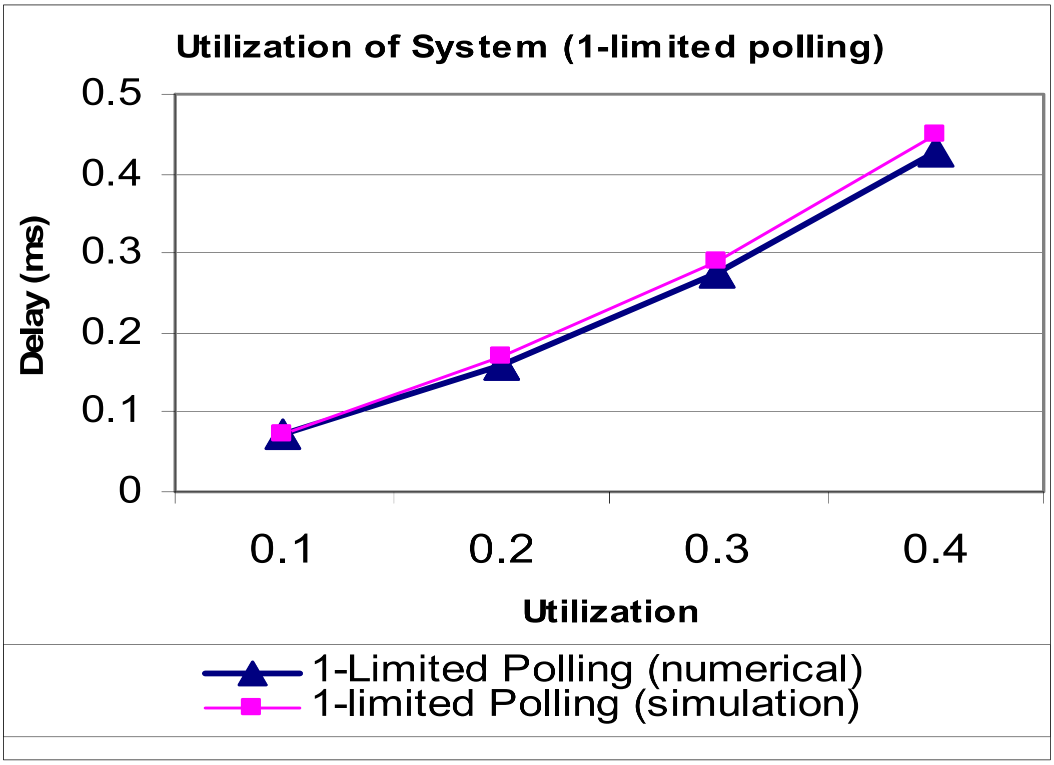

Figure 5-2: Delay vs. utilization in 1-limited polling

5.5.2 Results

The analytical results have been validated in this chapter by comparing the results obtained from the equations (numerical results) and the results obtained from the simulation (simulation results). J-sim, the software that provides a high-fidelity simulation for wireless communication with detailed propagation, radio, and MAC layers, has been used to simulate 100 nodes. The delay has been calculated for real-time traffic (Q_1) and non-real-time traffic (Q_2) on node 5, which was selected at random out of 100 nodes during the 15 randomized runs of the simulation. Figures 5-2 and 5-3 show the simulation results for 1-limited and k-limited models, respectively, which have been plotted against numerical results.

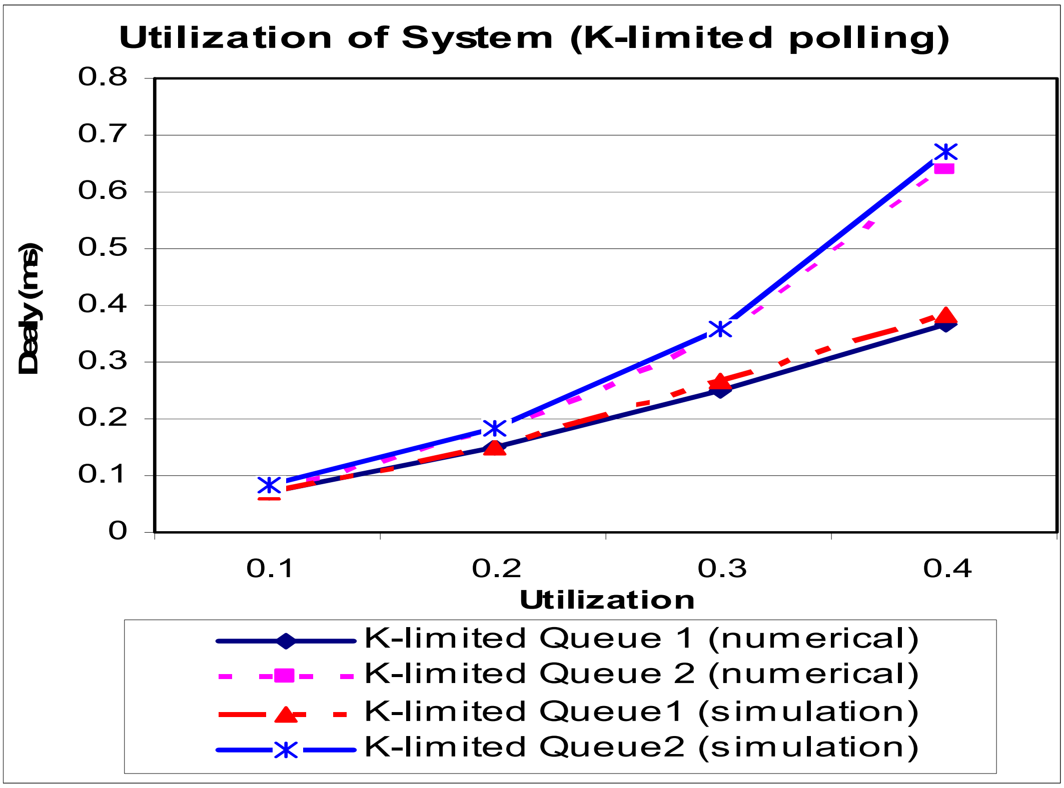

Figure 5-3: Delay vs. utilization in k-limited polling

5.6 Discussion and analysis

Table 5-1 shows the delay of Q_1 and Q_2 packets for the 1-limited polling model against the server utilization in a sensor node. The numerical and simulation results have been plotted in Figure 5-2. The simulation results are well aligned with the numerical results, thus validating the analytical model. Table 5-2 shows the delay of Q_1 and Q_2 packets in the k-limited polling model against the server utilization in a sensor node. The numerical and simulation results are shown in Figure 5-3. It can be noted that the delay for Class 2 packets exceeds those of Class 1 packets in all positions. The characteristics of k-limited polling model can be observed: as the utilization increases, there is a sharp increase in the queuing delay of Class 2 packets. Again, the simulation results are quite similar to the numerical results.

The horizontal axis in Figure 5-2 and Figure 5-1shows the variable of arrival rates in units of utilization, while the vertical axis shows the range of the variable of delay in units of milliseconds. Thus, the graph shows the change in delays over the arrival rates of the packets for both the numerical and the simulation results.

In both Figure 5-2 and Figure 5-3, it is apparent that the utilization increases because of the higher arrival rates, and the queuing delay increases sharply in a sensor node, particularly for Class 2 packets in the k-limited polling model. The slight difference between the numerical and simulation results is due to factors such as the retransmission of some packets and the time it takes to find a new path to forward the packets. Likewise, some packets chose other paths and did not pass through this node. However, because each sensor node has a small buffer and the delay inside the node is gradually increasing, as shown in Figure 5-2 and Figure 5-3, the probability of dropping the packets would increase accordingly.

5.7 Conclusion

This chapter has presented a novel analytical model that is based on the M/G/1 queuing system. The model has been analysed on the basis of a limited service polling discipline (both 1-limited and k-limited) to provide differential treatment to multiple classes of traffic in a wireless sensor network. The closed-form expressions of the queuing delay for two different classes of traffic have been derived. The analytical results have been verified through extensive simulation studies. The results presented provide a way to analyse the performance of a round-robin scheduling implementation in a sensor node. This measurement data can be useful as an input to a simulation study of sensor networks. Also, the analytical modelling technique and its verification through numerical and simulation results is the first step towards finding the most appropriate queuing-scheme implementation for wireless sensor networks. This work has provided an analytical model for implementing priority queuing in a sensor node to calculate the queuing delay [173]. Future work will focus on determining the end-to-end delay for multiple classes of traffic in sensor networks.

Chapter 6

MULTIPLE-LEVEL STATELESS PROTOCOLS

The introduction of real-time communication has created additional challenges in the field of wireless networks, due to different constraints. Sensor nodes spend most of their lifetimes routing packets from one node to another, until the packet reaches the sink. Therefore, sensor nodes primarily function as small routers. Since sensor networks represent a new generation of time-critical applications, it is often necessary for communication to meet real time constraints. However, the research that examines providing QoS guarantees for real-time traffic in sensor networks is still in its infancy. This chapter present a novel packet-delivery mechanism, the Multiple Level Stateless Protocol (MLSP), as a real-time way for sensor networks to guarantee the traffic in wireless sensor networks. The MLSP improves the packet-loss rate and the handling of holes in sensor networks much better than MMSPEED. This chapter also introduces the k-limited polling model for the first time in the routing protocols. This model is used in sensor networks by implementing two queues served according to a 2-limited polling model in a sensor node. Here, two different classes of traffic are considered, and the exact packet delay for corresponding classes is calculated/ estimated. The analytical results are validated through an extensive simulation study.

6.1 Introduction

Wireless communication, applications and/or underlying technologies, are among today's most dynamic areas of technology development. Sensor networks can be assumed as distributed computing platforms with many severe constraints, including limited CPU speed, small memory size, low power constraints, and narrow bandwidth. They are suitable for a wide range of civil and military applications [4-6, 168]. Sensor networks offer new challenges from two perspectives: (1) building communication protocols and (2) developing appropriate queuing and scheduling models. These challenges occur due to their large scale, independent operations, and extraordinarily parallel connections with a spatially distributed physical environment as well as a more strict set of resource constraints. Data gathering in a timely and reliable fashion has been a key concern here.

Wireless sensor networks particularly related to military applications and time-critical applications. Since sensor networks represent a new generation of time-critical applications include habitat monitoring, pollution detection, weather forecasting, and monitoring disasters, it is often necessary for communication to meet real time constraints. However, researches dealing with providing QoS guarantees for real time traffic in sensor networks are still very immature.

Many excellent protocols have already been developed for ad-hoc networks. They can be categorized into two groups: (1) flat routing and (2) hierarchical routing. In flat routing, all routes have equal responsibility for maintaining the routing information. Routing algorithms in this category can be further classified into three groups: (1) Proactive, (2) Reactive, and (3) Geographical [174]. Proactive routing algorithms maintain routes continuously for all reachable nodes. They usually require periodic dissemination of routing updates. Reactive routing algorithms establish and maintain nodes only if they are needed for communication. New routes are acquired when a new connection is set up and is to be maintained throughout the lifetime of connection regardless of topology changes. Geographical routing protocols utilize location for routing decisions. Hierarchical or cluster routing is the procedure of arranging nodes in a hierarchical manner.

Sensor networks however have additional requirements that were not specifically addressed here. For example, providing end-to-end real-time guarantees is a challenging problem in sensor networks. Nevertheless, communication protocols for sensor networks must supply real-time assurances. While ensuring the appropriate timing behaviour of a system has been a topic of research for decades, sensor network applications present physical space, in addition to time, as a new dimension for interaction with the environment. To provide real-time guarantees, the protocol must be specifically designed with these types of guarantees in mind.

The communication protocols among sensor nodes in a sensor network must provide real-time assurance and builds a cluster that depends on a number of factors, including communication ranges, number and types of a sensors, and geographical locations. Ensuring accurate timing behaviour of systems has been a topic of real time research for decades. The sink location directly affects the lifetime of the sensor network, and the sensor nodes need elegant and uncomplicated real-time protocols. In many applications, sensor data must be delivered within a time constraint to make appropriate real-time actions possible [146]. Most of the current QoS provisioning protocols [95, 147-149] in wireless sensor and ad-hoc networks are based only on end-to-end path discovery and path recovery. Also, much of the existing research is focused solely on reliability and lacks the ability to differentiate multiple classes of traffic that have different time constraints [71, 150, 151].

The best-effort behaviour of stander-forwarding systems does not support many categories of applications well. Real-time applications that require strict bounds on factors such as data rate, delay, and jitter will reveal sub-optimal performance as the network load increases.

The rest of the chapter is organized as follows. The design details of MLSP are presented in Section 6.2. Simulation information and results are detailed in Section 6.3. A discussion and analysis is presented in Section 6.4, followed by a conclusion in Section 6.5.

6.2 Design of MLSP

The MLSP framework consists of four elements: (1) a self-organization mechanism that assigns labels to nodes, (2) a forwarding policy, (3) a queuing model, and (4) the handling of holes.

6.2.1 Self-organization mechanism

The MLSP self-organization mechanism consists of labels assigned to the nodes to indicate certain information. Each node maintains a simple counter that indicates the sink number, level, and area it belongs to, as well as its power and buffer status, as illustrated in Figure 6-1.

The counter is a binary-bit register and is included on every sensor node. For instance, with a four-, five- , or six-bit demonstration of a level, counter could have up to 16, 32, or 64 levels from any particular sink, respectively. In this scheme, each node in the network belongs to a certain level with respect to a sink. Maintaining sink information is necessary because of the possible existence of multiple sinks. The level represents the distance (in terms of hop count) from the sink. In addition, the area is divided into four areas (which are depicted in Figure 6.2):

001	0011	010	1100100	1100100
Sink Number	Level	Area	Power	Buffer

Figure 6-1: Simple counter assigned to each node

The red area is closest to the sink. It is a very busy area and uses a special load-balancing technique to minimize congestion. From the network information, the nodes closest to the sink are the fastest to die in the entire network. Therefore, they require special treatment.

The orange area is second-closest to the sink. Most of the congestion and dropping of packets occur in this area. The orange area needs to improve the congestion control and should avoid dropping the real-time packets as much as possible, with respect to load balance between the nodes.

The green area lies farther from the sink than the orange area. Most of the action takes place in the green area, and the most important tasks in this area are to find the path and to handle holes in the path, if they exist.

The free area is the rest of the network, which covers all of the nodes not in the other areas.

In the above architecture, each area consists of a different number of hops, depending on the sink configuration. It is assumed that the number of hops in each area is static and is configured during the sink configuration period. Each sensor node is responsible for updating its power and buffer information in the specific field of its counter. This updated information does not lead to the transmission of any messages between nodes.

6.2.1.1 Initial setup (tagging process)

When the self-organization process starts, nodes are unaware of their distance to any given sink. Therefore, after a node assigns itself to a particular level/ area, it will calculate how many hops it is from the sink. Only one sink is assumed in this framework, and it is located in the lower-bottom corner.

The tagging process, as shown in Figure 6-2, is always started by a sink. This message contains the sink number, and level/area and is reported to the sensor node. Each node receiving this message will assign itself to the level/area which it belongs, with respect to the sink number (the sink number is ignored here, because only one sink exists in this framework). After receiving, each node broadcasts a message to report that it belongs to level one. All other nodes that do not yet have a level and listen to this message will increase the value of the received level by one, assign themselves to this level, and check their area before they broadcast this new level. This procedure continues until all nodes belong to a level and are assigned to an area. Once a node has assigned itself to a level and an area, it ignores all future broadcasts with level and area information. This tagging technique is designed to handle holes and to route the packets in case the routing technique fails to find a path. This eliminates the overhead due to route queries or updating.

6.2.2 Neighbourhood manager

MLSP features a novel neighbourhood manager that dynamically discovers eligible forwarding choices, with the ability to receive and forward real-time packets, as well as manage the neighbourhood table. The neighbourhood manager is invoked whenever no eligible forwarding choice exists in the neighbour table. To overcome the worst-case scenario in geographic forwards routing, in which finding a path may fail even if one exists, MLSP consists of three parts: (1) neighbourhood table management, (2) neighbour discovery, and (3) power control.

6.2.3 Neighbour discovery

When the self-organization process has finished, the neighbour discovery is directly invoked according to the following rules.

The neighbour discovery periodically broadcasts a beacon packet to its neighbours. This periodic beaconing is only used to exchange location information, power level, and available buffers between neighbours, without any extra overhead (because it uses the same packet). It obtains the power and buffer information from the counter. Comparing this to other protocols, the MLSP uses the power level and free buffer space as two important factors for choosing the target nodes. It confirms the node level and the area it belongs to in the counter.

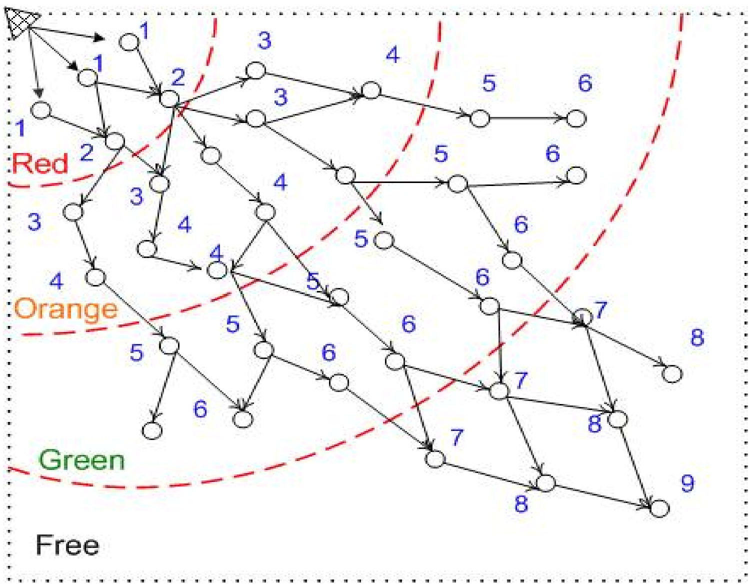

Figure 6-2: Hierarchical level-based scheme

In the worst-case scenario, if a node does not belong to any level or area for any reason, the neighbour discovery of this specific node will broadcast a beacon packet with normal power to find a neighbour for itself. If still no response is received from any neighbour, this means there is a hole in the network. In this case, this node will broadcast another beacon with higher power and follow the same steps until it receives a response from a neighbour. The level and the area of this node can then be determined based on its received/transmitted power level. From the location information of the neighbour node, its distance from the sink can also be determined. If this node is closer than its assigned level/area information, it can update its network/location information.

6.2.4 Power control

Power control is invoked when a node needs to (1) send a packet one or more hops away, (2) find a path throughout the holes by increasing the power level, or (3) handle a real-time packet that is reaching its deadline. When congestion occurs and there are real-time packets, the node will increase its power to avoid the congestion.

6.2.5 Neighbourhood table management

Neighbourhood table management is similar to a greedy geographic forward-routing algorithm, in which each node has only one table to store the location of immediate neighbours. This procedure is responsible for checking the information in each node's routing table. It then reorders the forwarding nodes in the table based on the Choosing Factors (CF), as described in the forwarding policy. If the neighbourhood table is empty and the node has a message to be sent, it will not wait for the routing queries and table updating. It can directly invoke the functionality of any casting at the MAC layer, as described in the MAC layer support section.

6.2.6 Forwarding policy

A forwarding policy is critical to guaranteeing the packet delivery. It is also very important to reduce the congestion from the side of the traffic flow in order to balance the load throughout the whole network. MLSP makes forwarding choices on a packet-by-packet basis. MLSP forwards the packets to the most-forward node that meets the packets delivery requirements:

$$CF_i = \frac{dis(x_i, y_i, x_d, y_d)}{delay} \times \left(\frac{Buffer}{MaxBuffer} \times \frac{Power}{MaxPower}\right) \times \frac{1}{Area}$$

The choosing factor involves (1) the power level of the source node, (2) the area, and (3) the buffer of the destination node. The choosing factor uses the power level and the available buffer in the destination to make the best decision. The available buffer is a critical factor because, if a node sends a packet to another node without enough available memory, there is a high possibility that the transmitted packet will be dropped. Therefore, these factors can improve the load balancing, as well as minimize the congestion, as shown in the result.

6.2.7 Traffic-load balancing and congestion manager

MLSP uses the power level and the available buffer of the destination nodes as important factors for choosing the appropriate nodes in the forwarding table. This technique also allows the node to forward the packets to different suitable nodes to minimize congestion. If congestion occurs for any reason, the node will not send the new packet, because the buffer will be full.

6.2.8 MAC-layer support

MLSP does not need real-time or QoS-aware MAC support. The MAC layer anycast has been used as a technique to find new paths to overcome the GF route failure and to pass the packet throughout the holes without any overhead, due to routing queries and updates. This technique is offered with little or no condition information.

When a node has a message that needs to be sent to the next nodes, it first broadcasts an RTS message to the node with a reduction of the level. If the sender node does not receive any response with a CTS, it will resend the RTS message again with more power to reach the nodes that are two hops or two levels away. If one or several nodes reply with a CTS, the sender node will choose one of these as the destination node and then send the information message directly to it. The sender node chooses the first-received CTS message and ignores the others. The receiver node is the node that is awake. Thus, the method essentially grants robustness and reduces the possibility of back-offs at each link. Followed by this handshaking, the sender node identifies the selected destination node and sends the information packet to it. The MAC address has been chosen as the node identifier in this scenario.

6.2.9 Queuing model

Queuing and scheduling have a direct impact on QoS characteristics. Regardless of its numerous limitations, the SPEED protocol has used First In, First Out (FIFO) as the queue system. FIFO is the default queuing algorithm in several topologies that requires no configuration. Most importantly, FIFO queuing makes no decision about packet priority. FIFO queuing involves storing packets and forwarding them in their order of arrival. Explode sources can also cause extended delays in delivering the packets of time-sensitive applications. Although FIFO queuing seems to be an effective network traffic controller, more recent intellectual networks need more sophisticated algorithms. Furthermore, in FIFO, a full queue will cause dropping of packets – even high-priority packets, as FIFO cannot differentiate between a high-priority and a low-priority packet. In fact, the sensor node cannot prevent this undesirable packet dropping, as it has no room for extra packets in its queue.

To overcome the limitations of FIFO queuing discipline, MMSPEED uses Priority Queuing (PQ) with FIFO scheduling in each queue. PQ is suggested as one of the applicable solutions to meet the desired QoS for real-time traffic. Here, two queues in a sensor node are considered – high-priority and low-priority. The scheduler uses strict-priority logic. In other words, it always serves the high-priority queue first. If there is no packet waiting in the high-priority queue, it will serve the low-priority queue. In this technique, the scheduler of the sensor node is serving different output queues simultaneously; hence, it behaves similarly to a multiple-queue/single-server system [173]. The limitation of PQ is that, any time there are packets in the high-priority queue, they will be extracted first. In this situation, the other queues could fill to capacity, resulting in packets to those queues being dropped.

The study of polling models is important, because it gives very good insight into the qualitative behaviour of many proposed and implemented queuing disciplines and forms the basis to derive closed-form expressions of different QoS parameters such as delay, jitter, and throughput. The basic polling model is a queuing model composed of a set of queues and a single server that serves the queues in cyclical order [175]. Polling models can be classified as exhaustive, gated, and limited service.

This chapter introduces the k-limited polling model for the first time and uses it in sensor networks. Two different classes of traffic, real time and non-real time, are considered to overcome the already-mentioned limitations of different queuing disciplines. This framework chose the 2-limited polling model as the queuing model with the shortest-elapsed (TTL) packet time, as this scheduling technique gives the best possible average waiting time and minimizes the dropping packets. This will be further discussed in the results.

The limited-service polling model is explained here only because it is related to this research's framework. The mathematical model for polling used here has been published in [176]. In the limited-service system, a queue is served until either the buffer is emptied or a specified number of packets are served, whichever occurs first. If, at most, k packets are served in one cycle, it is referred to as a k-limited polling model. The case in which k = 1 results in a simpler model, referred to as a 1-limited polling model. In this case, the server serves one packet from each queue in an alternating fashion during each cycle.

6.2.10 Handling holes

Greedy geographic forwards have several advantages over the traditional MANET routing algorithms for real-time sensor network applications. They do not suffer from route discovery delay and tend to select the shortest path to the destination. However, a known problem with greedy geographic forwards is that they may fail to discover a route in the presence of holes in the network. In the worst-case scenario, they may fail to find a path, even though one does exist. When there are large holes, the Euclidean distance becomes a poor approximation of the actual path length. As a result, the MLPS neighbourhood manager is responsible for finding the neighbour nodes and choosing one of them appropriately.

6.3 Experimentation and evaluation

For additional understanding and analysing of our schemes, the J-sim has been chosen to perform our simulations. MMSPEED has been used as a benchmark. The MMSPEED and MLPS results have been generated and compared to show the performance of the presented protocol.

J-sim, the software that provides a high-fidelity simulation for wireless communication with detailed propagation, radio, and MAC layers, has been used to simulate 100 nodes. It shows the result and delay characteristics of sensor nodes in the simulation. The simulation technique has been clarified in the environmental setup section. The results are explained in the discussion and analysis section.

6.3.1 Environmental setup

A square-shaped network (200m, 200m) was selected as the standard to validate the accuracy of the presented protocol. One hundred nodes with a 40m radio range were generated to operate in this network. For each node, a free-space propagation channel model was assumed with a transmission speed of 250kbps and a total packet length of 40 bytes for both real time and non-real time, with a 15-packet capacity as their buffer sizes [165] [166] [167]. For each node in the sensing state, packets are generated at a constant rate of 1 packet/sec. The real-time packet generation rate is 3 packets/sec.

6.3.2 Simulation results

The results have been validated in this chapter by comparing the result obtained from the simulation of MMSPEED and the result obtained from the simulation of the protocol. Figure 6-3 shows the total sending packets in MMSPEED and MLSP. The power consumption has been plotted in Figure 6-4. The total number of missing packets of MMSPEED and MLSP has been plotted in Figure 6-5. Figure 6-6 and Figure 6-7 show the delay calculated for real-time and non–real-time traffic on node 5, which was selected at random during the 15 randomized runs of the simulation. Finally, Figure 6-8 compares the Priority Queue and Polling Queue in MLSP.

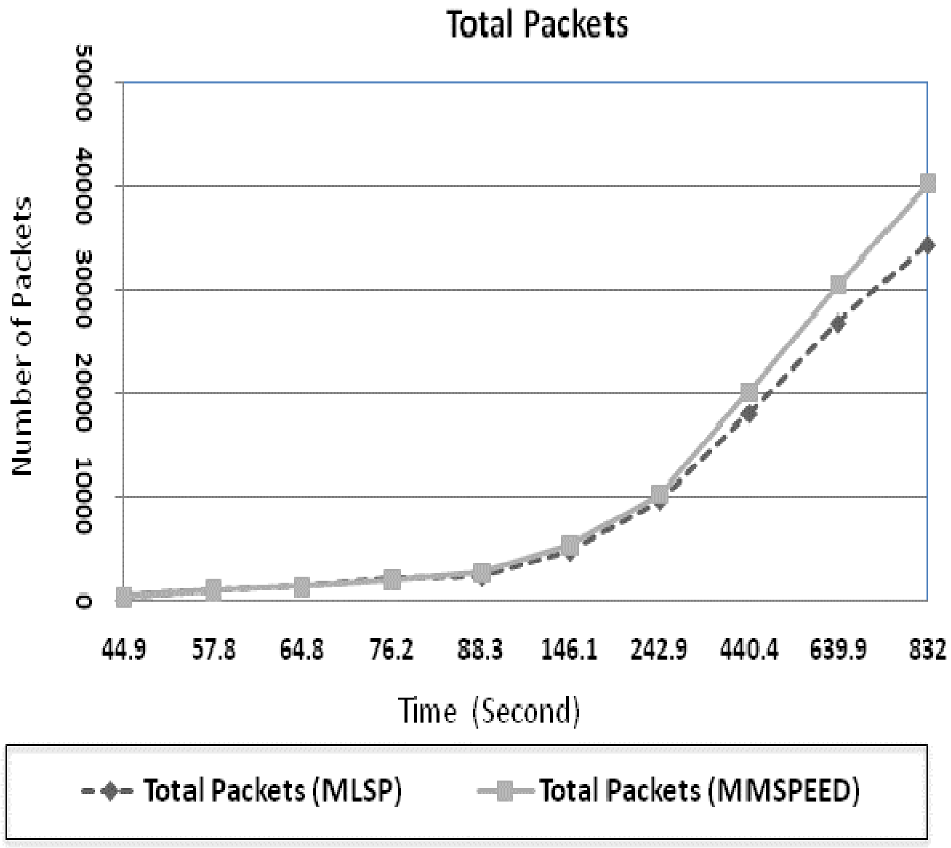

Figure 6-3 : Total sending packets in sensor network

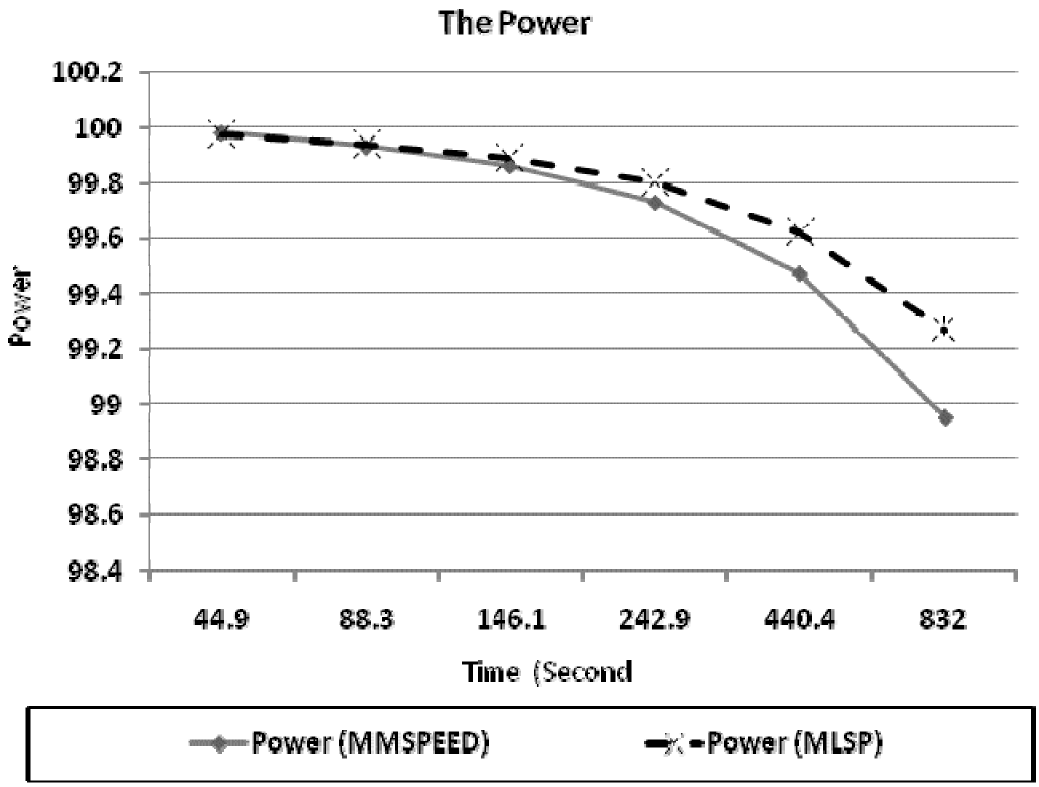

Figure 6-4: Power consumption

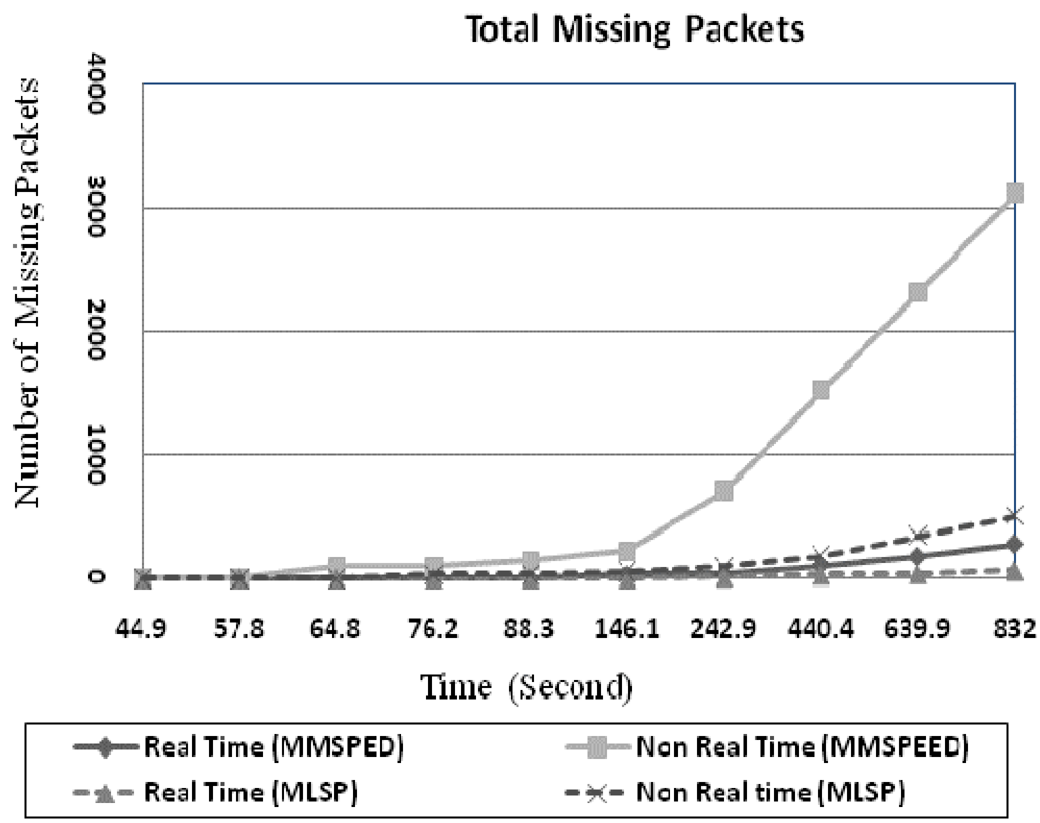

Figure 6-5: Total missing packets

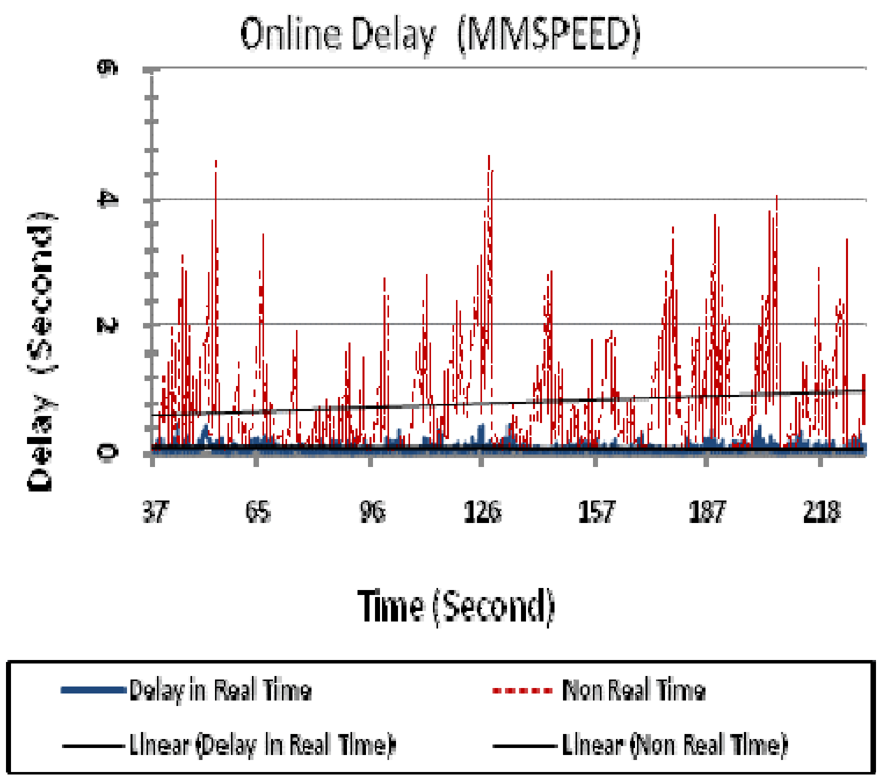

Figure 6-6: Online delay in MMSPEED

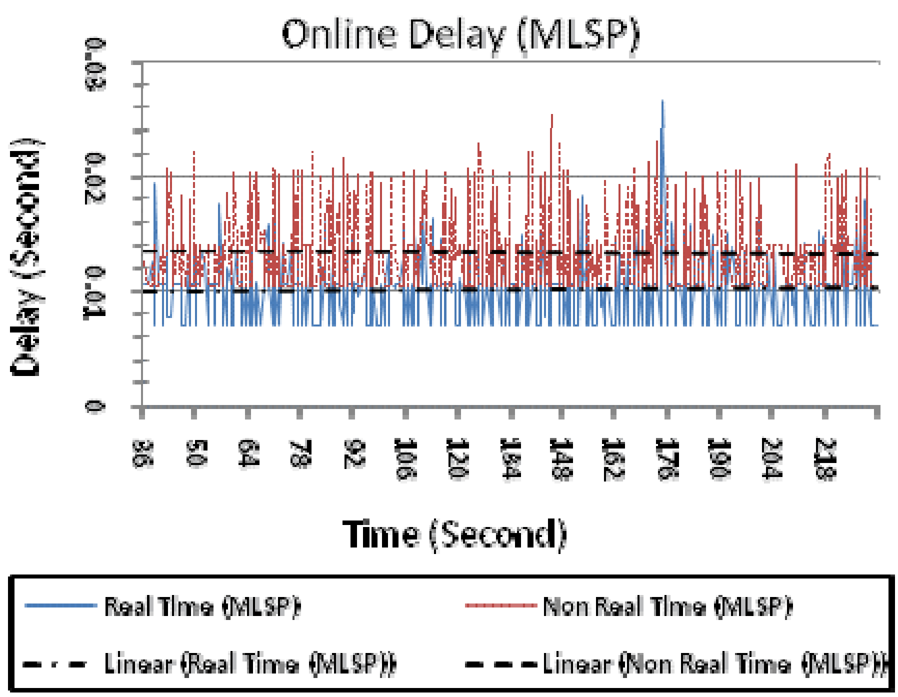

Figure 6-7: Online delay in MSLP

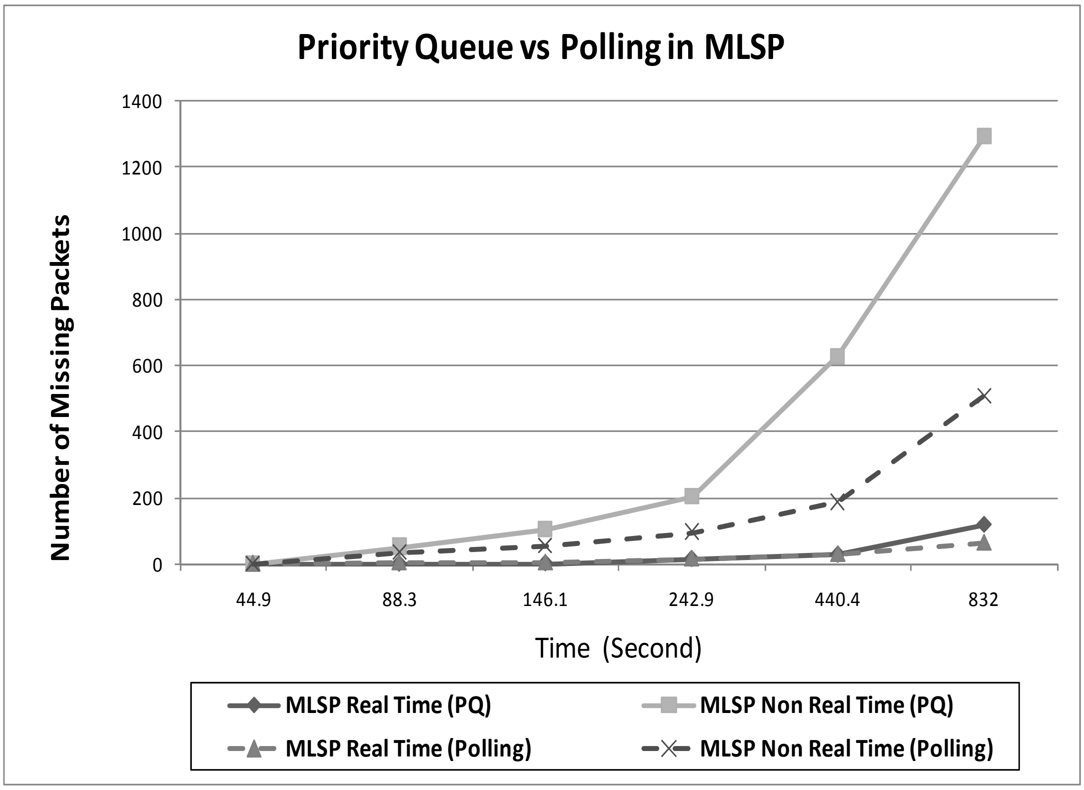

Figure 6-8: Priority queue vs. Polling queue in MLSP

6.4 Discussion and analysis

The horizontal axis in Figure 6-3 shows the time in units of seconds. The vertical axis shows the range of total sending packets. This figure shows the difference in total sending packets for the MMSPEED and MLSP protocols. All types of packets have been calculated as total sending packets. As shown in this graph, the differences are increased with time. MLSP minimizes the total sending packets, which saves more power as compared to MMSPEED. The total packets of real time and non real time in MMSPEED and MLSP are the same. MLSP, however, minimizes the number of control, update location, back pressure, and new path-finding packets. The number of missing packets plays an important role in QoS. All dropped packets are counted as missing packets.Figure 6-4 shows the power consumption in the vertical axis. The horizontal axis shows the time in units of seconds. This figure compares the power consumption between the MMSPEED and MLSP protocol. MLSP improves the power consumption of the sensor network throughout time. The first point in the shape shows that MMSPEED has a small difference in power consumption. This is attributed to the fact that MLSP has a tagging process for all nodes in the initial phase. This leads it to send extra packets. However, with time, the graph shows that MLSP decreases its power consumption sharply. The explanation: MLSP has a smart congestion control and forwarding policy that decreases the number of controlled packets and leads to power conservation.

Figure 6-5 shows the total number of missing packets in the vertical axis. The total number of real-time and non-real-time missing packets in MMSPEED are plotted against the same for MLSP. MLSP decreases the number of missing packets for real-time and non-real-time packets. The difference between the results in this figure is due to factors such as the queue model (polling), the minimized retransmission of some packets, and the time it takes to find a new path to forward the packets and to decrease the congestion so that the data can flow more smoothly. Because the cycle in the polling model allows the queue to choose the real time and non real time depends on the cycle number. The improvement is in both real time and non real time, and this gives extra advantages to MLSP.

Figure 6-6 and Figure 6-7 show the online delay for the same period of time in both MMSPEED and MLSP. In MMSPEED, the delay in the non-real-time protocol is very high and causes congestion in the nodes. The delay is different from time to time, and it increases for non-real-time traffic. The dashed line in this figure shows the delay as a constant linear for both real-time and non-real-time traffic.

Figure 6-8 shows the number of missing packets on the vertical axis. The time in units of seconds has been illustrated on the horizontal axis. This figure evaluates this chapter's algorithm with the Priority Queue and Polling Queue. The algorithm improves the miss ratio, compared to MMSPEED. MLSP with polling sharply decreases both real-time and non-real-time missing packets.

6.5 Conclusion

This chapter presented a novel algorithm to enable the sensor data to be delivered with time constraints in real-time scenarios. The k-limited polling model is introduced here for the first time. It is used in sensor networks through the implementation of two queues (real time and non-real time), served according to a 2-limited polling model in a sensor node. This improves the loss rate in the networks. The 2-limited polling model, as the queuing model with the shortest elapsed (TTL) packet time, gives the best possible average waiting time and minimizes the dropping of packets sharply, as seen in the results. This novel protocol also solved the problem of holes in sensor networks. In addition, the total sending packets dropped sharply compared to MMSPEED. This also decreased the consumption power. The analytical results have been verified through simulation studies. The results presented provide a way to analyse the performance of this protocol implementation in a sensor node. Also, this analytical modelling technique, and its verification through simulation results, is the first step towards discovering the most appropriate queuing and algorithm-scheme implementation for wireless sensor networks.

Chapter 7

INTEROPERABILITY BETWEEN THE SENSOR AND UMTS NETWORKS

This chapter introduces a new framework for moving data from a sink to the user at a low cost and power, using the Universal Mobile Telecommunication System (UMTS), which is standard in the Third Generation Mobile System (3G). Much of the research on sensor networks focuses on issues such as power consumption, self-organization, routing, and communication between sensors and the sink. On the other hand, there is little research on facilitating communication between a sink and the user. The integration of the mobile network infrastructure and sensor networks will reduce the cost of building new infrastructure and enable the large-scale deployment of sensor networks.

7.1 Introduction

Microelectromechanical systems (MEMS), digital electronics, and wireless communication have enabled the development of a new generation of large-scale sensor networks. In these networks, the nodes are small in size and communicate with other nodes across short distances with low-power consumption. This suits a wide range of applications [4-6]. These devices build a self-organizing, ad-hoc network to forward data packets, using multi-hop connections, to the sink nodes [6]. The earliest research in the area of wireless sensor networks goes back to the early seventies [2]. A number of survey papers provide a more comprehensive background on sensor networks [1, 10-13].

Through self-organization techniques, a large number of sensor nodes can be spread over an environment without a priori knowledge of where each sensor is placed [8]. Sensor nodes have a short transmission range because of their limited radio capabilities; therefore, the data must be relayed using intermediate nodes before reaching the sink. In addition, it is more advantageous to use a multi-hop path, which consists of shorter links, to the sink node, rather than using a single, long connection. A sensor network is traditionally composed of many spatially distributed sensors, which are used to observe or detect phenomena such as temperature changes or pollutant levels. Each sensor should be physically small and cost-effective, thus making it possible to deploy nodes in large numbers. A sensor is equipped with data-processing capabilities, a radio transceiver, a small microcontroller, and an energy source (usually a battery). Sensor networks can be used in different applications [10, 177, 178].

This chapter gives a detailed description of a new framework that leads to interoperability between a sensor network and UMTS networks in order to move data from a sink to a user at a low cost and low power. This chapter also analyses the requirements of and issues with such interoperability.

Section 7.2 of this chapter provides details on the sensor communication framework. Section 7.3 describes the new interoperability sensor and the UMTS approach presented in this work. Results and explanations of the interoperability framework for the sensor and UMTS networks (IFSUN) are presented in Section 7.4, followed by a conclusion and statement of future work in Section 7.5.

7.2 Sensor communication framework

Communication between sensor nodes in order to build a cluster network can be achieved in a variety of ways [177]. The efficiency of the network depends on the sink location, which directly affects the lifetime of the sensor network. However, sensors within a cluster communicate with the sink node that is allocated to them via short-range wireless communication links. Various choices are available, such as Bluetooth [178], ZigBee [10], simple RF communication, and UWB (Ultra Wide Band) [10]. There are advantages and disadvantages to each of these technologies, based on the application and the requirements of the user [179]. Users communicate with sensors through one or more committed nodes, i.e., sink nodes. These nodes are responsible for inserting sensor queries into the network, gathering responses from sensors, and forwarding them to users, as shown in Figure 7-1. The sink node in a cluster is the only node that can communicate with the outside world via a dedicated gateway node, or directly (if it is equipped with dual sink/gateway functionality).

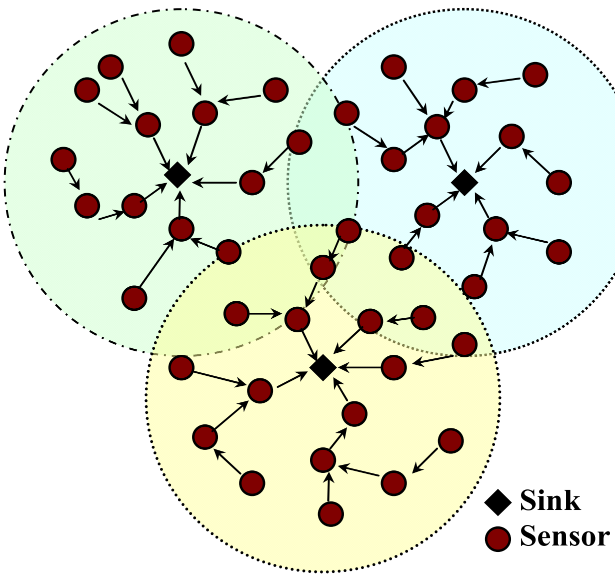

Figure 7-1 : A large-scale sensor network with three clusters

7.2.1 Sensor network frameworks

Several approaches are suggested to allow communication between the sink and the user. Three of these current methods have significant disadvantages with respect to the main constraints related to sensor networks – power consumption and cost. However, the advantages of the approach presented in this chapter [180] for the first time are available to all organizations.

7.2.1.1 Satellite communication

Satellite communication is expensive and is not available to all organizations. In addition, communicating with a satellite requires high power.

7.2.1.2 Wireless LAN networks

In this approach, a full wireless LAN network (WLAN) is installed to cover the area in which the sensor network is deployed. This technique is costly, complex, and highly dependent on the topography of the area covered by the sensor network.

7.2.1.3 A fully equipped vehicle or airplane

A vehicle or a plane is used, from time to time, to collect data from the sink nodes in a sensor network. This is a very complex and costly operation. Another drawback is that data cannot be collected on demand.

7.2.1.4 Sensor and UMTS framework network

The interoperability framework for the sensor and UMTS network (IFSUN) approach, which will be explained later in this chapter, moves data from the sink to the user at a lower cost and power by using the Universal Mobile

Telecommunication System (UMTS). UMTS is standard for the Third Generation Mobile System (3G). The integration of the mobile network infrastructure and sensor networks will reduce the cost of building new infrastructures and enable the large-scale deployment of sensor networks.

7.2.2 Sinks

Sink nodes are expensive devices that should be used economically. When the size of the network increases, the average length of the paths (or the number of hops) from the sensors to the sink node increases. As a result, energy dissipation for each packet delivery increases, as well, which will result in a shortening of the network lifetime.

One approach to overcoming this problem is to associate one sink per sensor node and to locate these sinks very close to their associated sensor nodes, as demonstrated in Figure 7-1. In this case, the transmission energy per sensor node is minimized, perhaps increasing the network's lifetime. Of course, this solution is not the most economical. The number of sink nodes is an important design criterion, which is directly dependent on the available budget. If the number of sink nodes is known, then one can estimate the number of sub-networks (clusters).

7.2.3 UMTS network

The UMTS architecture consists of a core network (CN), UTRAN (UMTS Terrestrial Radio Access Network), and user equipment (UE), as depicted in Figure 7-2. The main function of the CN is to provide switching and routing of the packet traffic. The CN is further divided into circuit-switched (CS) and packet-switched (PC) domains. The circuit-switched elements are comprised of the Mobile Services

Switching Center (MSC), Visitor Location Register (VLR) and Gateway MSC. The packet-switched elements are the Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). The SGSN is responsible for mobility management and IP-packet session management. It routes user packet traffic from the radio network to the appropriate GGSN and provides access to external packet data networks, such as the Internet and intranets. The UMTS core network has multiple SGSNs that are connected to several radio network controllers. This framework focuses on the packet-switched domain, because it provides the Multimedia Message Service (MMS) and the Wireless Application Protocol (WAP). UTRAN consists of multiple base stations (Nodes B) and Radio Network Controllers (RNC) that provide the WCDMA air interface access method for user equipment. This communication network can carry many traffic types, from real-time circuit-switched to IP-based packet-switched traffic. The RNC authorizes control functionalities for one or more Node Bs, while the Iub is a logical interface between them. The RNC and its corresponding Node Bs are called the Radio Network Subsystem (RNS). There can be more than one RNS present in a UTRAN. The term UE refers to any device that has the capability to communicate with a UMTS network. This corresponds to our proposed architecture (IFSUN) for a wireless sensor network sink or gateway [181, 182].

In a wireless environment, in which bandwidth usage is significant, short address length and simplicity of user entry on limited keypads are the distinguishing features between various systems. The international mobile subscriber identity (IMSI) is used by any system that employs the GSM standard [183]. The IMSI uses up to 15 digits – a 3-digit mobile country code (MCC), a 2-digit mobile network code (MNC), and a mobile subscriber identity number (MSIN) of up to 10 digits. The IMSI has been recognized as a better identifier than any other system.

7.2.4 Multimedia message service

At the UMTS level, which is an extension of the successful Short Messaging Service (SMS), there are Multimedia Message Services (MMS) [181, 182, 184, 185]. The insertion of the packet-switched domain into mobile networks enabled the deployment of the second generation of messaging services, which allows the exchange of multipart multimedia data, and even the composition of messages as orchestrated multimedia presentations. MMS is designed to overcome the known limitations and shortcomings of both SMS and email, making it suitable and effective for the current, evolving wireless infrastructure and fulfilling the market demands for wireless messaging.

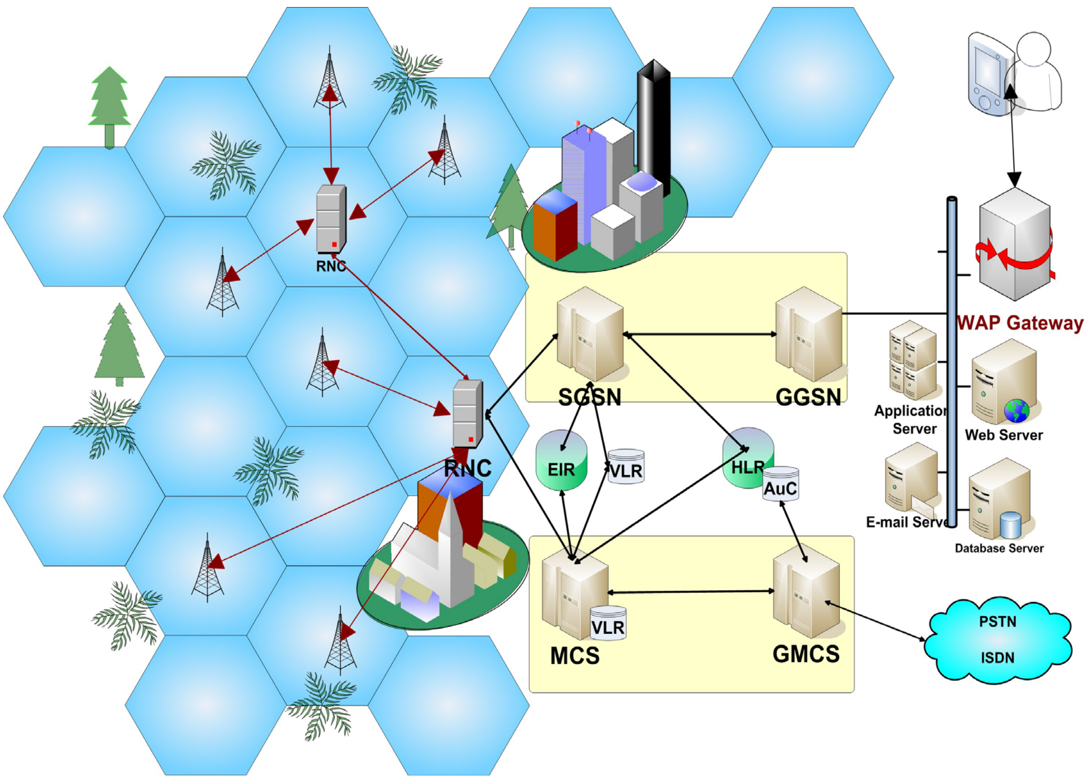

Figure 7-2 : UMTS logical architecture

MMS contains not only text, but also other elements such as voice, animated GIF images, JPG images, MIDI ringing tones, and applications, as illustrated in Figure 7-3. MMS uses the Multipurpose Internet Mail Extensions (MIME) encapsulation, which is an Internet standard that extends the specification for formatting non-ASCII messages so that they can be sent over the Internet. Therefore, any component that has its own MIME type can be sent as MMS, as well. MMS is jointly standardized by the 3GPP and WAP forum [185]. MMS was chosen for our framework for two reasons: size and addressing.

The size and features of MMS, as mentioned earlier, are very important for communication power consumption. An MMS message is not limited to any specific size; it can range from a few bytes to several kilobytes, meaning that any framework can be used, depending on the chosen application. For example, temperature measurement requires a small MMS, while taking an image of the environment requires a large MMS.

The MMS addressing model employs two types of addresses: the address of the MMS proxy and the address of the recipient. The address of the MMS Proxy-Relay, which is the Uniform Resource Identifier (URI) of the MMS Proxy-Relay, is responsible for evaluating messages and accepting the message, if it finds the message valid and resolves the recipient address.

The address of the recipient (user), which is the user-defined identifier, is given by the UMTS network and supports the address format compatible with Internet email addresses [186]. It is also expected that MMS service providers may use solutions based on static tables. In addition, it is possible to send MMS messages from a sink to an email address or a mobile phone. The following samples explain different addressing types for MMS:

To: 04022565619 / TYPE = PLMN

To: +35853132567 / TYPE = PLMN

To: Medo User < medo@user.com >

To: 192.168.0.2 / TYPE = IPv4

To: FEDC:BA98:5671:3267:FEDC:BA98:7654:3210/TYPE= IPv6.

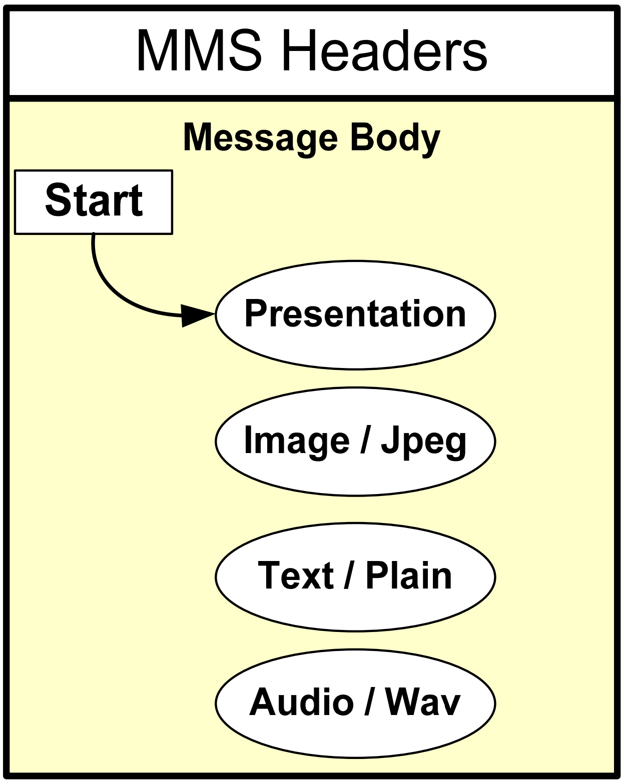

Figure 7-3: Model of MMS PDU containing a multipart message body

7.3 Interoperability sensor and UMTS network

An approach is proposed that will allow the different network standards to communicate with each other. The scope of this work is to develop a framework that will allow the integration of sensor networks into the fabric of other wireless networks.

The framework has been divided into two parts: the UMTS portion and the sink. The use of a UMTS network enables direct access to the sensor network, where the user can request data anywhere and at any time. The sink nodes are expensive devices and also have a short lifetime.

7.3.1 Framework

The proposed architecture of a wireless sensor network sink or gateway is shown in Figure 7-4. It is a flexible architecture, also known as dual sink, that supports a range of communication technologies for sensors and user applications with no or minimum modification [179].

7.3.1.1 Network interfaces

In the communication layer, two different communication technologies will be supported by the sink or gateway: the wide-range interface and the short-range communication interface. Examples of the short-range interface are Bluetooth, ZigBee, IEEE 802.16, 802.11, proprietary RF, and GPRS. The wide range is the UMTS interface.

7.3.1.2 Sensor network middleware

Sensor network middleware represents the central component of the gateway architecture. This is the layer that divides communication between the sensors and the users, encapsulates the internal organization of the sensor network, and provides API function to the users.

7.3.1.3 User application

The user application layer employs the API functions that are provided by the sensor network middleware. Depending on the application scenario, either short-range wireless communication links or wide-area networks are used for communication between users and gateways.

7.3.1.4 Security

The security layer is responsible for authentications and encryptions, to protect the data. This layer is added because the MMS does not supply its own particular security support and does not mandate any specific security solution

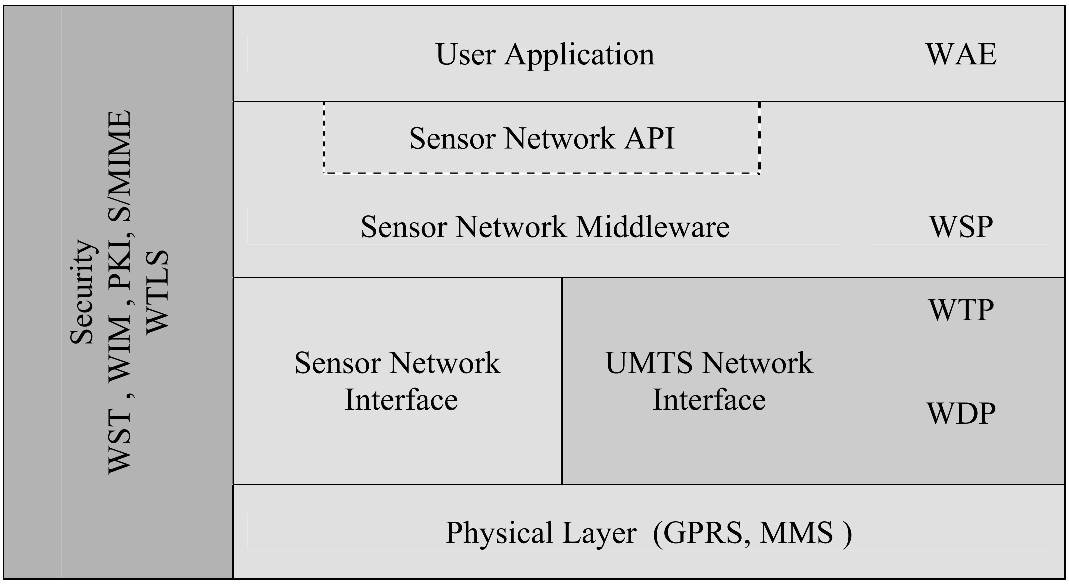

Figure 7-4: Sink or gateway architecture

7.3.2 The solutions

This section considers the possible solutions for the different components of the proposed framework. The framework solution is divided into a message solution and an internetworking solution. The aim of these solutions is to receive data on demand, or at any time and anywhere throughout the Country, and to share these data with other organizations worldwide by using Internet or email technologies.

7.3.2.1 Message solution

This solution allows the framework to implement the MMS infrastructure in a sensor network using the Wireless Access Protocol (WAP), which was chosen for its features as described in [187, 188]. This solution allows the sink to send the data that is gathered as MMS, which can easily be received by any mobile phone in any location.

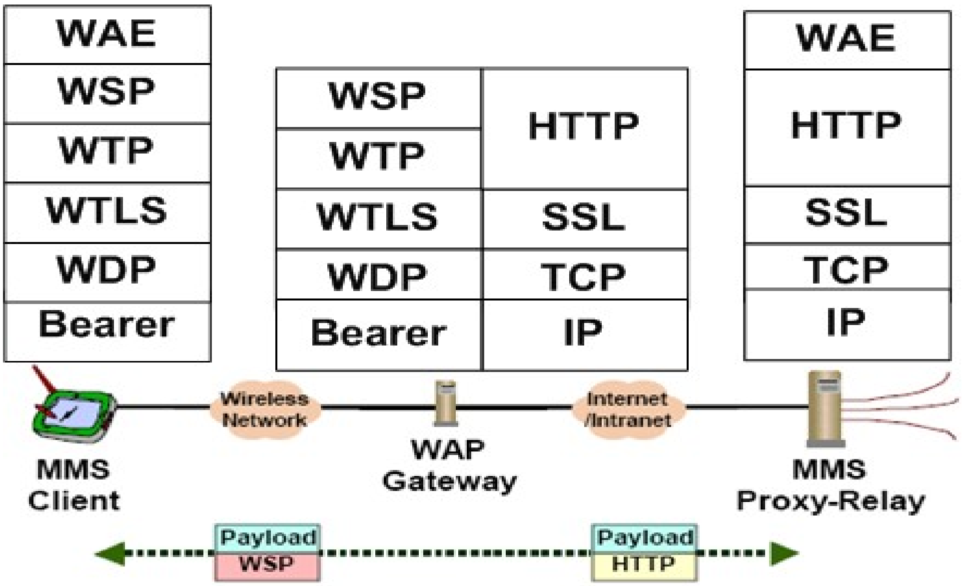

Figure 7-5: WAP implementation of MMS interface with protocol stacks

The Wireless Session Protocol (WSP) is used to transport the messages from the sink to the MMSC, and from the MMSC to the sink. The WSP is responsible for the general requirements, architecture, and functionality, as explained in [182, 185]. Figure 7-5 shows a payload carried by the WSP and HTTP. This payload represents the MMS application layer PDU (Protocol Data Unit). Below is an example of PDU encoding and decoding:

Hexadecimal PDU Message:

07911356131313F311000A9240226565910000AA27547419D40CE341D47
21B0EEA81643850D84D0651D1655033ED0651CB6D38A8078AD500.

7 Bit PDU Message (readable) is:

SMSC# +31653131313

Sender:0422565619

TP_PID:00 TP_DCS:00

TP_DCS-popis: Uncompressed Text class:0

Alphabet: Default

Max Temp = 28 and Min Temp = 15

Length: 39

Figure 7-6 shows the process for sending an MMS from an MMS client (Sink in Sensor network) to another MMS client (data receiver) as follows:

- Each sink registers itself as a mobile unit in the Home Location Register (HLR) on the UMTS network. The sink address is based on the Mobile Station ISDN Number (MSISDN) that is operated by the device. In many paging systems, users are assigned PINs that authorize a caller to deposit a message. This addressing problem can be solved by adding this number to the sink memory, or by changing the sink design to allow the USIM card to cooperate with it. After this, the sink is ready for the next step.
- Sink sends MMS to the MMS server.
- MMS server sends notification to recipient client.
- Client fetches MMS from the MMS server.

- MMS server sends delivery report to client

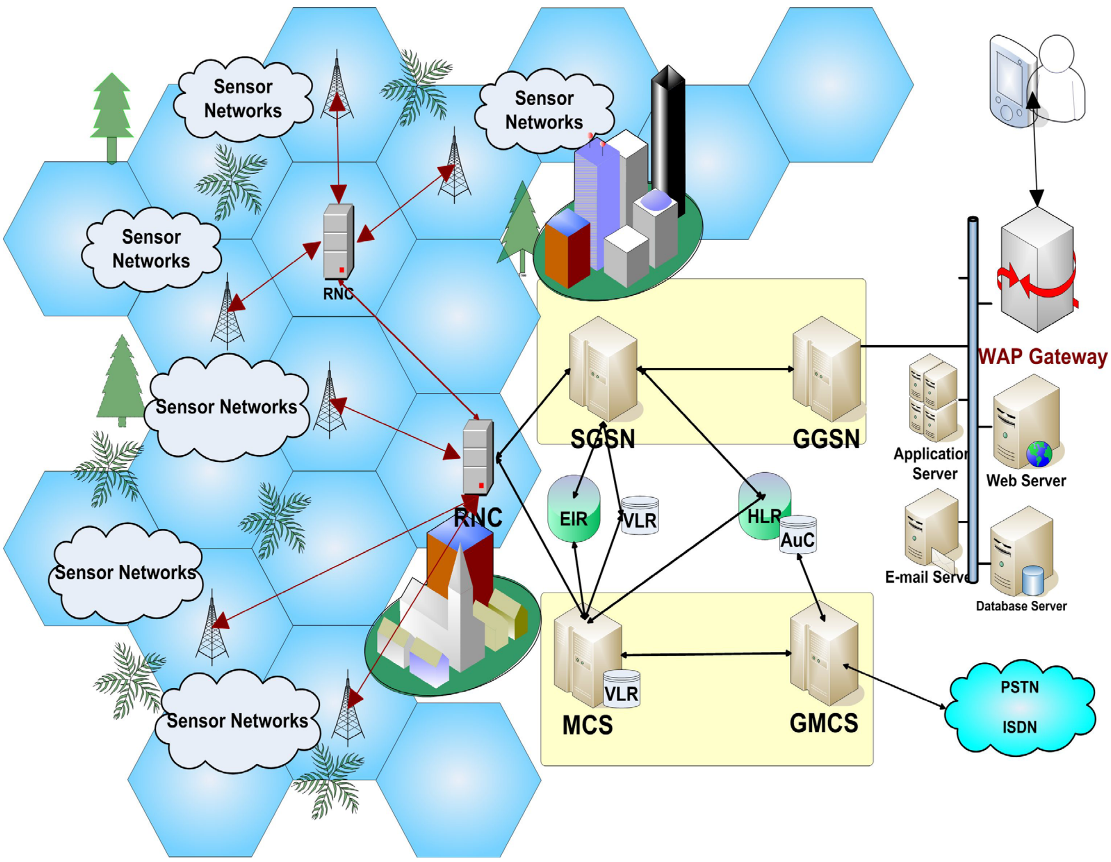

Figure 7-6: Proposed system architecture (messages solution)

7.3.2.2 Internetworking solution

The internetworking solution of this framework has been illustrated Figure 7 - 7: connecting the sensor network to the Internet via the UMTS network. The idea adopted in this solution is to obtain the data from any local network anywhere in the world. As mentioned earlier, in the MMS addressing model, the MMS address format is compatible with Internet email addresses, and the WAP gateway is responsible for converting the PDU or MMS message to HTTP format.

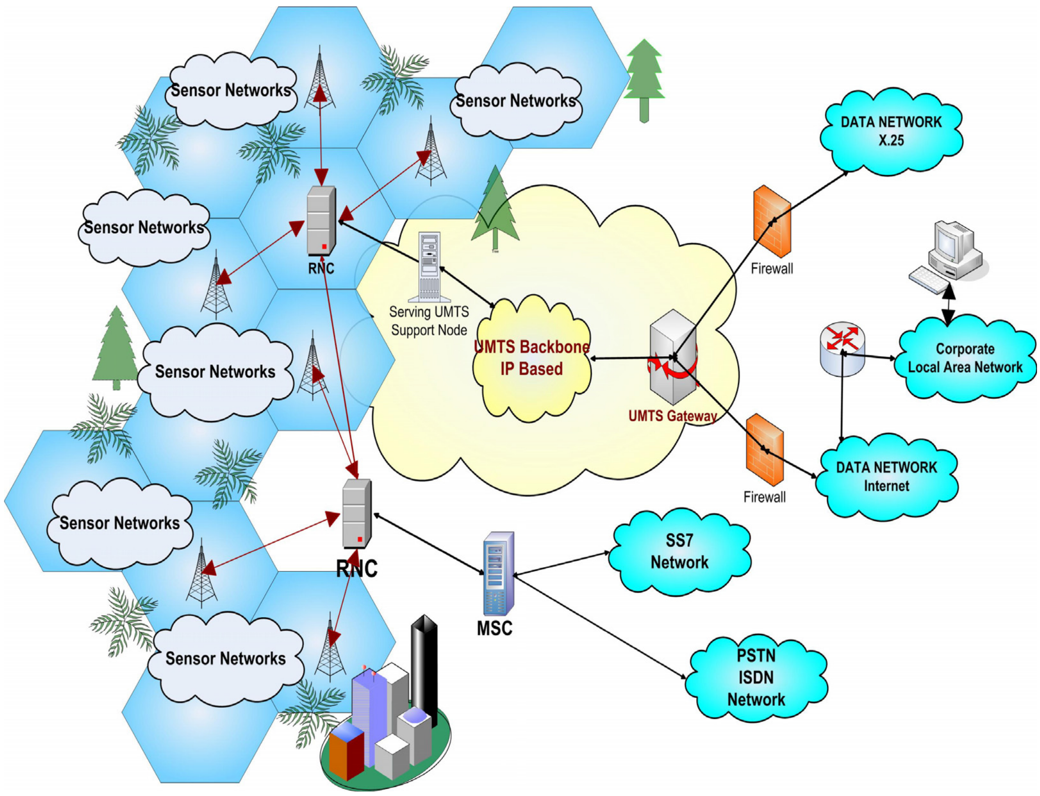

Figure 7 - 7 : Wireless data networks (Internetworking solution)

7.3.2.3 Power consumption

Power consumption is highly critical in a sensor network. Location updating enables the HLR to keep track of the subscriber's current location. The sensor network is registered in the UMTS network, and the location is then fixed. Here, the power consumption in the message and internetworking solutions can be minimized by preventing the location-updating signals. Another solution is to use solar energy to supply power to the sinks.

7.4 Experimentation and evaluation

For additional understanding and analysis of our schemes, OPNET Release 11.5 was chosen to design and evaluate IFSUN.

OPNET simulations have been used to generate a number of different forms of output. The advantage of the analysis tool in OPNET is that it displays information in the form of graphs. Graphs are presented within rectangular areas called analysis panels. An analysis panel consists of a plotting area with two numbered axes, generally referred to as the abscissa axis (horizontal) and the ordinate axis (vertical). Figures 7-6 and 7-7 are used to show the efficiency of the IFSUN framework.

7.4.1 Simulation results

The results in this chapter were validated from the simulation of the proposed system architecture (INFUS). Figure 7 - 8 shows email traffic when the sink sends emails to the server email. Figure 7-9 also explains the HTTP traffic sent through the UMTS network. Figure 7-10 demonstrates the response times authorized to each SGSN in the UMTS core network.

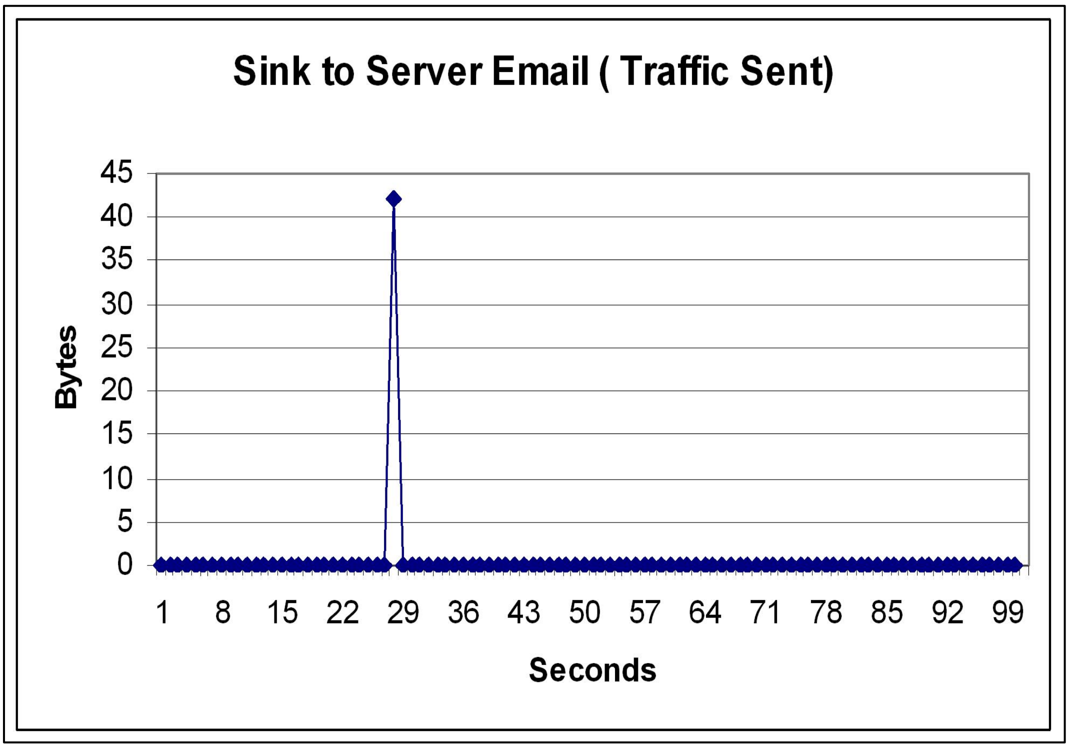

Figure 7 - 8 : Traffic sent from sink to email server via UMTS

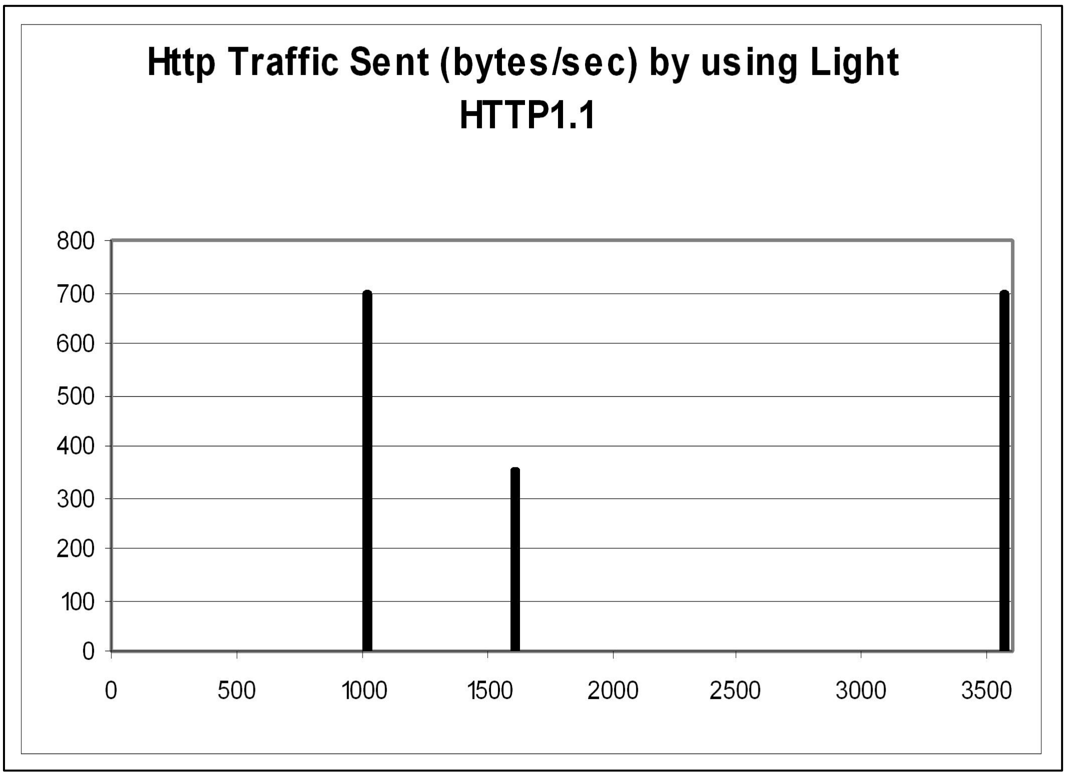

Figure 7-9 : HTTP traffic sent over the UMTS

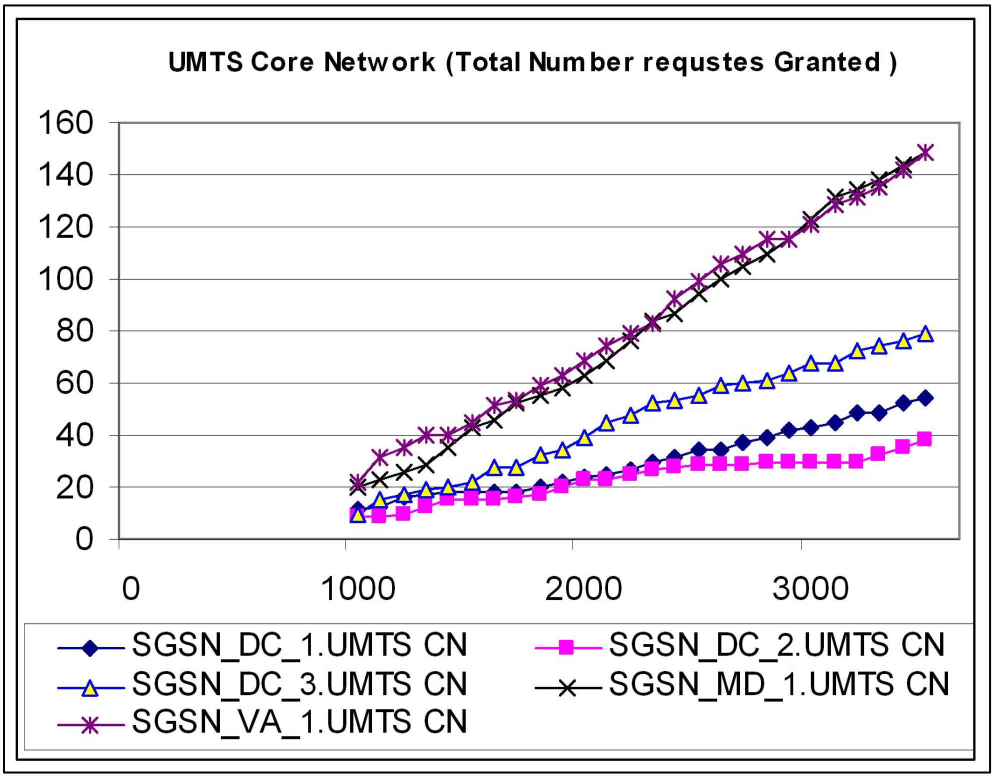

Figure 7-10: Total number of requests granted in UMTS

7.4.2 Results and explanation

The efficiency of the IFSUN framework has been shown in Figure 7-6 and Figure 7 - 7 respectively. Figure 7 - 8 shows email traffic when the sink sends emails to the server email and confirms that bandwidth is not critical in this approach. The reason for this is the small email size, as the sink can only send specific data, depending on the application it uses and the sink characteristics. Figure 7-9 explains the HTTP traffic sent through the UMTS network while the message commitment is performed through HTTP POST/GET transactions. If clients do not have HTTP-based stack capability, then these transactions are initially performed over WSP by the terminal and then transformed into HTTP by their local WAP gateway, which is efficient at using the radio resources with superior connectivity. Figure 7-10 demonstrates the response times authorized to each SGSN in the UMTS core network. This figure proves that the UMTS core network can handle delivery of the requests and gives the maximum guarantee that it will do so, which is important for a sensor network. SGSN_VA_1 and SGSN_MD_1 are depicted higher in the figure in comparison to the three other normal mobile-user SGSNs. As can be seen, they fulfill both the total number of sink requests and calling requests from the sensor network and normal mobile users.

7.5 Conclusion and future work

This chapter presented a new approach – using the characteristics of sensor networks and mobile network infrastructure to deliver sensor network signals. Communicating between these two systems dynamically and intelligently can reduce the cost and increase the lifetime of sensor networks. It was also shown that this approach is suitable for all organizations, and for gathering data on demand. The feasibility and viability of the proposed method has been proven through initial experimental work.

As shown in this chapter, delivering sensor network signals is much more complex than the mere translation of message encoding and protocols. However, work is currently underway on building a new framework to achieve the goals of enabling low-cost, low-power communication between mobile networks and sensor networks, suitable for a range of commercial and other applications and with a short transmission range. Moreover, the network lifetime must be maximized, with the least economical investment, by selecting the correct number of sink nodes.

Chapter 8

CONCLUSION AND FUTURE WORK

Sensor networks have different constraints than traditional wired networks. First, because sensors have a limited supply of energy, energy-conserving forms of communication and computation are essential to wireless sensor networks. Second, since sensors have limited computing power, they may not be able to run sophisticated network protocols. Third, since the bandwidth of wireless links connecting sensor nodes is often limited, inter-sensor communication is further constrained. Finally, because sensor networks are often deployed with inexpensive hardware by a single organization, there is less need for interoperability with existing standards. This chapter focuses to recapitulate and summarize the contribution of this thesis. The important issue of future work is also presented.

8.1 Conclusion

Many new routing and MAC-layer protocols have been proposed for wireless sensor networks, thereby tackling the issues raised by the resource-constrained unattended sensor nodes in large-scale deployments. The majority of these protocols consider energy efficiency as the main objective and assume data traffic with unconstrained delivery requirements. However, the growing interest in applications that demand certain end-to-end performance guarantees, along with the introduction of imaging and video sensors, has posed additional challenges. Transmission of data in such cases requires both energy- and QoS-aware network management in order to ensure efficient usage of the sensor resources and effective access to the gathered measurements.

The fundamental aim of this project is to increase the accuracy of real-time communication in wireless sensor networks by:

- Examining the closed-form expressions of the queuing delay for multiple (real time and non-real time) classes of traffic in a senor node through the implementation of priority queuing based on the M/G/1 queuing system.
- Presenting a novel analytical model that is based on the M/G/1 queuing system. We have analysed the model on the basis of the limited service polling discipline (both 1-limited and k-limited) to provide differential treatment to multiple classes of traffic in wireless sensor networks. The closed-form expressions of the queuing delay for two different classes of traffic have been derived. The analytical results have been verified through extensive simulation studies. The results presented provide a way to analyse the performance of round-robin scheduling implementation in a sensor node.
- Compiling measurement data that can be useful as an input to a simulation study of sensor networks. Also, this analytical modelling technique, and its verification through numerical and simulation results, is the first step towards finding the most appropriate queuing-scheme implementation for wireless sensor networks.
- Developing a novel algorithm to enable the sensor data to be delivered with time constraints and to make real-time scenarios possible. The k-limited polling model is introduced for the first time in this chapter. It is

used in sensor networks through the implementation of two queues (real time and non-real time) served according to a 2-limited polling model in a sensor node. This improves the loss rate in the networks. This novel protocol also solved the problem of holes in sensor networks. In addition, the total sending packets has dropped sharply, compared to MMSPEED, which also decreases the consumption power

- Explaining a new approach – using the characteristics of sensor networks and mobile network infrastructure to deliver sensor network signals. Communicating between these two systems dynamically and intelligently can reduce the cost and increase the lifetime of sensor networks. It was also shown that this approach is suitable for all organizations and for gathering data on demand. The feasibility and viability of the proposed method has been proven through the initial experimental work.

8.1.1 Summary of contributions

To review, this thesis had three goals: developing a suitable queuing system, creating a real-time communication framework, and gathering the data in demand time. The contributions of this thesis are further explained in the following sections.

8.1.1.1 Queuing system

This thesis presented the closed-form expressions of the queuing delay for multiple (real time and non-real time) classes of traffic in a sensor node through the implementation of priority queuing. This thesis also presented a novel analytical model that is based on the M/G/1 queuing system. The model has been analysed on the basis of the limited-service polling discipline (both 1-limited and k-limited) to provide differential treatment to multiple classes of traffic in wireless sensor networks. Based on the M/G/1 queuing system, the measurement data can be useful as an input to a simulation study of sensor networks.

8.1.1.2 A real-time communication framework

This thesis proposed a novel real-time framework that enables sensor data to be delivered within time constraints that make suitable real-time actions possible. This framework improves the packet-loss rate and the handling of holes in sensor networks. This framework also supports multiple dynamic routs with little or no state information. This novel protocol also solved the problem of holes in sensor networks. In addition, the total sending packets dropped sharply, compared to MMSPEED, which decreases the consumption power.

8.1.1.3 Gathering data in demand time

This thesis presents a new approach of using the characteristics of sensor networks and mobile network infrastructure to deliver the sensor network signals. Communicating dynamically and intelligently between these two systems can reduce the cost and increase the lifetime of sensor networks. It was also shown that this approach is suitable for all organization and for gathering the data on demand. The feasibility and viability of the proposed method has been proven through initial experimental work.

8.2 Future Work

Based on the research goals and challenges that exist in this field, the following are identified as open research issues in QoS support in wireless sensor networks.

- Node mobility. Most of the current protocols assume that the sensor nodes or/and sink are stationary.
- The optimal number of sinks to support QoS networks.
- Novel and simple QoS models, as required to guarantee real-time communication.
- A suitable queuing system for sensor networks.
- Integration of sensor networks with IP-based networks by using the Universal Mobile Telecommunication System (UMTS), and allowing the requests from the user to be made to the sink through the Internet or MMS.

Bibliography

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaiam, and E. Cayirci, "Wireless Sensor Networks: a Survey," *ELSEVIER Computer Networks,* vol. 38, pp. 393-422, 2002.

[2] U. Bilstrup, K. Sjoberg, B. Svensson, and P.-A. Wiberg, "Capacity Limitation in Wireless Sensor Networks," in *Emerging Technologies and Factory Automation, Proceedings ETFA 03 , IEEE , Vol 1, Pages 529-536*, Sept 2003, pp. 529-536.

[3] B.Warneke and K. S. J. Pister, "MEMS for Distributed Wireless Sensor Networks," in *9th International Conference on Electronics, Circuits and Systems*, Dubrovnik, Croatia, September, 2002

[4] Kahn, J. M., R. H. Katz, and K. S. J. Pister, "Next Century Challenges: Mobile Networking for Smart Dust," in *Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM '99)*, Seattle , Washington , USA, August 1999, pp. 271-278.

[5] I. F. Akyildiz, W. Su, Y. Sankarasubramaiam, and E. Cayirci, "A Survey on Sensor Networks," in *IEEE Communication Magazine*. vol. 40, no 8, August 2002, pp. 102-114.

[6] Pottie, G. J, and W. J. Kaiser, "Wireless Integrated Network Sensors," *Communication of the ACM,* vol. 43, pp. 51-58, May 2000.

[7] D. Niculescu, "Positioning in ad hoc sensor networks," *IEEE Network,* vol. 18 (4), pp. 24-29, 2004.

[8] Yuen, W. H., and C. W. Sung, "On Energy Efficiency and Netowk Connectivity of Mobile Ad Hoc Network," in *23rd International Conference on Distributed Computing System (ICDCS 2003)*, Providence, Rhode Island, USA, May 2003.

[9] R. Iyer and L. Kleinrock, "QoS Control For Sensor Networks," in *Communication, 2003 ICC03, IEEE Intrnational*, Los Angeles, 11-15 May 2003, pp. 517-521.

[10] F. Zhao and L. J. Guibas, *Wireless sensor networks : an information processing approach*. Amsterdam ; San Francisco: Morgan Kaufmann, 2004.

[11] Qi, H., S. S. Iyengar, and K. Chakrabarty, "Distributed Sensor Netowrks- a Review of Recent Research," *Journal of the Franklin Institute,* vol. 338, pp. 655-668, 2001.

[12] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, "Highly resilient, energy-efficient multipath routing in Wireless Sensor Networks," *ACM Mobile Computing and Communication Review,* vol. 5 (4), October 2001.

[13] S. S. Iyengar, L. Prasad, and H. Min, Advances in Distributed Sensor Integration: Application And Theory: Prentice-Hall, 1995.

[14] M. Tubaishat and S. Madria, "Sensor networks: an overview," *Potentials, IEEE,* vol. 22, pp. 20 - 23, Apr-May 2003.

[15] e. a. K. Sohrabi, "Protocols for Self-organization of A Wireless Sensor Network," *IEEE Personal Communications,* vol. 7, pp. 16-27, October, 2000.

[16] R. C. Shah and J. M. Rabaey, "Energy Aware Routing for Low Energy Ad Hoc Senso Networks," in *IEEE Wireless Communications and Networking Conference (WCNC)* Orland, FL, 2002.

[17] S. Tilak, N. Abu-Ghazaleh, and W. Heinzelman, "A Taxonomy of Wireless Micro-Sensor Network Models," *ACM Mobile Computing and Communication Review (MC2R),* vol. 6, April 2002.

[18] J. N. Al-Kamal and A. E. Kamal, "Routing Techniques in Wireless Sensor Networks, A survey," *Wireless Communications, IEEE,* vol. 11, pp. 6-28, 2004.

[19] K. Akkaya and M. Younis, "A survey on Routing Protocols for Wireless Sensor Networks," *Elsevier Ad Hoc network Journal,* vol. 3, pp. 325-349, 2005.

[20] K. Holger and W. Andreas, "A short Survey of Wireless Sensor Networks," in *Technical Report [TKN Technical Report TKN-03-018]* Berlin, October, 2003.

[21] A.A. Ahmed, H. Shi, and Y. Shang, "A Survey on Network Protocols for Wireless Sensor Networks," in *International Conference on Information Technology: Research and Education (ITRE'03)*, 11-13 Aug. 2003, pp. 301 - 305.

[22] K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie, "Protocols for Self-Organization of a Wireless Sensor Network," *IEEE Personal Communications,* vol. 7, pp. 16-27, October 2000.

[23] C. Kandagatla, "Survey and Taxonomy of Grid Resource Management Systems."

[24] Ning Xu, "A Survey of Sensor Network Applications," available form the World Wide Web: http://enl.usc.edu/ningxu/papers/survey.pdf.

[25] E. Yoneki and J. Bacon, "A survey of Wireless Sensor Network technologies: research trends and middleware's role," Technical Report, Ed.: [Available from the World Wide Web http://www.cl.cam.ac.uk/TechReports/UCAM-CL-TR-646.pdf], UCAM-CL-TR-646.

[26] I. Khemapech, I. Duncan, and A. Miller, "A Survey of Wireless Sensor Networks Technology," http://www.dcs.st-and.ac.uk/pst/resources/publications/khemapach05survey.pdf , 2005.

[27] K. Krauter, R. Buyya, and M. Maheswaran, "A Taxonomy and Survey of Grid Resource Management Systems for distributed Computing," *International Journal of Software: Practice and Experience (SPE), Wiley Press, New York, USA,* May 2002.

[28] V. Rajaravivarma, Y. Yang, and T. Yang, "An Overview of Wireless Sensor Network and Applications," in *35th Southeastern Symposium on System Theory*, 16-18 March, 2003, pp. 432- 436.

[29] et al and R. Min, "Low Power Wireless Sensor Networks," in *International Conference on VLSI Design* Bangalore, India, January 2001.

[30] A. Cerpa, J. Elson, M. Hamilton, and J. Zhao, "Habitat monitoring: application driver for wireless communications technology," *ACM SIGCOMM Workshop on Data Communications in Latin America and the Caribbean Cost Rica,* April 2001.

[31] D. Braginsky and D. Estrin, "Rumor Routing Algorithm for Sensor Networks," in *First ACM International Workshop on Wireless Sensor Networks and Applications (ACM WSNA 2002) pp. 22-31*, Atlanta, USA, September 2002.

[32] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *1st ACM International*

Workshop on Wireless Sensor Networks and Applications, Pages 88-97, ACM Press, 2002.

[33] Schwiebert, L., S. K. S. Gupta, and J. Weinmann, "Research Challenges in Wireless Networks of Biomedical Sensors," in *the 7th Annual International Conference on Mobile Computing and Networking (MOBICOM '01), pp. 151-165*, Rome , Italy, July 2001, pp. 151-165.

[34] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the World with Wireless Sensor Networks," in *Conf. on Acoustics, Speech and Signal Processing (ICASSP 2001)*, Salt Lake City, Utah, , May 2001.

[35] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed Compression in a Dense Microsensor Network. ," in IEEE Signal Processing Magazine, *http://www.mit.edu/people/kusuma/Papers/spmag_final.pdf*, March 2002.

[36] M. Srivastava, R. Muntz, and M. Potkonjak, "Smart Kindergarten: Sensor-based Wireless Networks for Smart Developmental Problem-solving Environments(Challenge Paper)," in *7th Ann. Intl. Conf. on Mobile Computing and Networking ACM*, Rome, Italy, July 2001. http://www.acm.org/pubs/articles/proceedings/comm/381677/p132-srivastav%a/p132-srivastava.pdf., pp. 132–138.

[37] F. Zhao, J. Shin, and J. Reich, "Information-Driven Dynamic Sensor Collaboration for Tracking Applications," in *IEEE Signal Processing Magazine*, March 2002. http://www.parc.com/cosense/pub/ieee_spm.pdf.

[38] N. F. Timmons and W. G. Scanlon, " Analysis of the performance of IEEE 802.15.4 for medical sensor body area networking," *IEEE SECON 2004,* October 2004.

[39] J. Hill, R. Szewcyk, A. Woo, D. Culler, S. Hollar, and K. Pister, "System architecture directions for networked sensors," in *8th International*

Conference on Architectural Support for Programming Languages and Operating System (ASPLOS IV), Cambridge, MA, 2000, pp. 93-104.

[40] A. Savvides and M. B. Srivastava., "A distributed computation platform for wireless embedded sensing.," in *International Conferance on Computer Design (ICCD02), Pages 220-225*, Freiburg, Germany, September 2002.

[41] J. M. Kahn, R. H. Katz, and K. Pister, "Next century challenges: Mobile networking for "smart dust"," in *5th International Conference on mobile Computing and Networking (MobiCom 1999), pages 271-278*, seattle, WA, ACM Press, August 1999.

[42] J. Rabaey, J. Ammer, J. d. Silva, D. Pater, and S. Roundy, "Picoradio supports ad-hoc ultra-low power wireless networking," in *IEEE Computer Magazine*, July 2002, pp. 42-48.

[43] W. R. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive protocols for information dissemination in wireless sensor networks," in *In proceeding of the 5th Annual ACM/IEEE international Conference on Mobile Computing and Networks (MobiCom'99)*, Seattle, Washington, Aug, 1999, pp. 174–185.

[44] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *6th Annual ACM/IEEE International Conference on Mobile Computing and Networking(MobiCom'00)*, Boston, MA, August 2000.

[45] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Direct diffusion for wireless sensor networking," *IEEE / ACM Transactions on Networking,* vol. 11 (1), Feb 2003.

[46] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," in *Proceeding of the 5th annual ACM/IEEE International Conference on Mobile Computing and Networking MobiCom 99*, Seattle, WA , USA, 1999, pp. 263-270.

[47] D. Braginsky and D. Estrin, "Rumor Routing Algorithm for Sensor Networks," in *First ACM International Workshop on Wireless Sensor Networks and Applications (ACM WSNA 2002) pp. 22-31*, Atlanta, GA, USA, October 2002.

[48] C. Schurgers and M. B. Srivastava, "Energy efficient routing in wireless sensor networks," in *MILCOM Proceedings on Communications for Network-Centric Operations: Creating the Information Force*, McLean, VA, 2001.

[49] F. Ye, A. Chen, S. Liu, and L. Zhang, "A scalable solution to minimum cost forwarding in large sensor networks," in *Proceedings of the tenth International Conference on Computer Communications and Networks (ICCCN)*, 2001, pp. 304-309.

[50] M. Chu, H. Haussecker, and F. Zhao, "Scalable information driven sensor querying and routing for ad hoc heterogeneous sensor networks," *The International Journal of High Performance Computing Applications* vol. 16, pp. 293-313, 2002.

[51] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," in *SIGMOD Record*, September 2002.

[52] N. Sadagopan and e. al, "the ACQURIRE mechanism for efficient querying in sensor networks," in *Proceedings of the First International Workshop on Sensor Network Protocol and Application*, Anchorage, AK, May 2003.

[53] R. C. Shah and J. Rabaey, "Energy aware routing for low energy and ad hoc sensor networks," in *IEEE Wireless Communications and Networking Conference (WCNC 2002)*, Orlando, Florida, Mar 2002.

[54] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless sensor networks," in *Proc. of the 33rd Hawaii International Conference System Sciences*, Hawaii, January 2000.

[55] A. Manjeshwar and D. P. Agrawal, "TEEN: A Protocol for Enhanced Efficiency in Wireless Sensor Networks," in *Proc. of the 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing*, San Francisco, CA, April 2001.

[56] A. Manjeshwar and D. P. Agrawal, "APTEEN: ahybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks," in *Proceedings of 2nd International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing*, Ft. Lauderdale, FL, April 2002.

[57] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power Efficient GAthering in Sensor Information Systems," in *IEEE Aerospace Conference, Big Sky*, Montana, March 2002.

[58] V. Rodoplu and T. H. Ming, "Minimum energy mobile wireless networks," *IEEE Journal of Selected Area in Communications,* vol. 17, pp. 1333-1344, 1999.

[59] L. Li and J. Y. Halpern, "Minimum energy mobile wireless networks revisited," in *Proceedings of IEEE International Conference on Communications (ICC'01)*, Helsinki, Finland, June 2001.

[60] L. Subramanian and R. H. Katz, "An Architecture for Building Self Configurable Systems," in *IEEE/ACM Workshop on Mobile Ad Hoc Networking and Computing*, Boston, MA, August 2000.

[61] Q. Li, J. Aslam, and D. Rus, "Hierarchical Power-aware Routing in Sensor Networks," in *In Proceedings of the DIMACS Workshop on Pervasive Networking*, May, 2001.

[62] J. N. Al-Karaki, R. Ul-Mustafa, and A. E. Kamal, "Data Aggregation in Wireless Sensor Networks - Exact and Approximate Algorithms," in

Proceedings of IEEE Workshop on High Performance Switching and Routing (HPSR'04), Phoenix, Arizona, USA, April 18-21, 2004.

[63] Q. Fang, F. Zhao, and L. Guibas, "Lightweight Sensing and Communication Protocols for Target Enumeration and Aggregation," in *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking and computing(MOBIHOC)*, 2003, pp. 165-176.

[64] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, "A two-tier data dissemination model for large-scale wireless sensor networks.," in *8th Annual International Conference on Mobile Computing and Netwotking MOBICOM*, Pages 148-159, ACM / IEEE Press, 2002, pp. 148-159.

[65] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad hoc routing," in *Proceedings of the 7th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'01)*, Rome, Italy July 2001.

[66] Y. Yu, D. Estrin, and R. Govindan, "Geographical and energy-aware routing: a recursive data dissemination protocol for wireless sensor networks," UCLA Computer Science Department Technical Report May 2001.

[67] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, "Span: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks," in *Proceedings of the 7th ACM International Conference on Mobile Computing and Networking*, July 2001.

[68] I. Stojmenovic and X. Lin, "GEDIR: Loop-Free Location Based Routing in Wireless Networks," in *In International Conference on Parallel and Distributed Computing and Systems* Boston, MA, USA, Nov. 3-6, 1999.

[69] F. Kuhn, R. Wattenhofer, and A. Zollinger, "Worst-Case optimal and average-case e±cient geometric ad-hoc routing," in *Proceedings of the 4th*

ACM International Conference on Mobile Computing and Networking, 2003, pp. 267-278.

[70] C. Lu, B. Blum, T. Abdelzaher, J. Stankovic, and T. He, "RAP: A Real-Time Communication Architecture for Large-Scale Wireless Sensor Networks," in *Eighth IEEE Real-Time andEmbedded Technology and Applications Symposum (RTAS '02)*, Sept 2002, pp. 55-66.

[71] T. He, J. Stankovic, C. Lu, and T. Abdelzaher, "SPEED: AStateless Protocol for Real-Time Communication in Sensor Networks," in *Proceedings of IEEE International Conference on Distributed Computing Systems*, 2003, pp. 46 – 55.

[72] E. Felemban, C.-G. Lee, E. Ekici, R. Boder, and S. Vural, "MMSPEED: multipath Multi-SPEED protocol for QoS guarantee of reliability and. Timeliness in wireless," in *IEEE transactions on mobile computing (IEEE trans. mob. comput.) , vol. 5, no6, pp. 738-754* 2006.

[73] O. Chipara, Z. He, G. Xing, Q. Chen, X. Wang, C. Lu, J. Stankovic, and T. Abdelzaher, "Real-time Power Aware Routing in Wireless Sensor Networks " in *Technical Report WUSEAS-2005-31, Washington University in St. Louis,* . July 2005.

[74] M. Younis, K. Akkaya, M. Eltoweissy, and A. Wadaa, "On Handling QoS Traffic in Wireless Sensor Networks," in *the 37th Hawaii International Conference on System Sciences* 2004.

[75] Wikipedia, "Sensor node," http://en.wikipedia.org/wiki/Sensor_node#List_of_Sensor_Nodes, 2008.

[76] B. Horizonte, "MSc Thesis: BEAN: Uma Plataforma Computacional para Rede de Sensores Sem Fio (BEAN: A Computer Platform for Wireless Sensor Network)," in *Computer Science Department, MG, Brazil*. vol. Master

Universidade Federal de Minas Gerais: Universidade Federal de Minas Gerais, 2004.

[77] K. Pister, "Smart Dust, http://robotics.eecs.berkeley.edu/~pister/SmartDust/," 2003.

[78] B. Warneke, B. Liebowitz, and K. S. J. Pister, "Smart Dast: communicating with a cubic-millimeter computer," *IEEE Computer,* pp. 2-9, January 2001.

[79] ETH Zurich, "BTnodes - A Distributed Environment for Prototyping Ad Hoc Networks," http://www.btnode.ethz.ch/, 28 Oct 2008.

[80] S. Hollar, "Cots dust. Master's thesis." vol. Master: University of California at Berkeley, 2000, p. 105.

[81] Axcess International INC, "http://www.axcessinc.com/products/dot.html."

[82] By Richard Zurawski and Contributor Richard Zurawski, *Embedded Systems Handbook: handbook*: CRC Press, 2006.

[83] "http://envisense.org/glacsweb/tech/probe/probe.html."

[84] http://www.xbow.com/Home/wHomePage.aspx.

[85] http://www.csee.ltu.se/~jench/mulle.html.

[86] University of Colorado at Boulder Mantis, "Multimodal networks of in-situ sensors. http://mantis.cs.colorado.ed."

[87] S. M. Laboratories, "http://www.sunspotworld.com/."

[88] Telos., "Telos wireless sensor node, http://www.moteiv.com/info.php," March 2004.

[89] Moteiv Corp, "Tmote Sky, Ultra low power IEEE 802.15.4 compliant wireless sensor module," Moteiv Corporation, 55 Hawthorne St., Suite 550 San Francisco, CA 94105 11/13/2006.

[90] CSIRO, "CSIRO Sensor and Sensor Networks Research Progress Report www.sensornets.csiro.au," July 2006 to June 2007.

[91] Carnegie Mellon University, "Real-Time Wireless Sensor Network Platform, http://www.ece.cmu.edu/firefly/resources.html," Real-Time & Multimedia Systems Lab.

[92] M. Younis, M. Youssef, and K. Arisha, "Energy- Aware Routing in Cluster-Based Sensor Networks," in 10th IEEE/ACM International Symposium on Modeling , Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 02), Texas, USA, October 2002.

[93] C. E. Perkins, *Ad hoc networking*. Boston: Addison-Wesley, 2001.

[94] J. M. Rabaey, M. J. Ammer, J. L. d. Silva, D. Patel, and S. Roundy, "PicoRadio Supports Ad Hoc Ultra-Low Power Wireless Networking," *Computer,* vol. 33, pp. 42-48, July 2000.

[95] T. Chen, J. Tsai, and M.Gerla, "QoS Routing Performance in Multihop Multimedia Wireless Networks," in *IEEE Sixth International Conference of Universal Personal Communications*, 1997, pp. 557–561.

[96] K. G. Hoblos, M. Staroswiecki, and A. Aitouche, "Optimal Design of Fault Tolerant Sensor Networks," in *EEE Int'l.Conf. Cont. Apps* Anchorage, AK, Sep 2000.

[97] P. Favre, "A 2V, 600 $_{m}$A, 1 GHz BiCMOS super regenerative receiver for ISM applications," *IEEE Journal of Solid State Circuits 33 2186-2196,* 1998.

[98] T. Melly, A. Porret, C. C. Enz, and E. A. Vittoz, "A 1.2 V, 430 MHz, 4dBm power amplifier and a 250 $_{m}$W Frontend, using a standard digital

CMOSprocess," *IEEE International Symposium on Low Power Electronics and Design Conference, San Diego, USA,* pp. 233-237, August 1999.

[99] B. Krishnamachari, D. Estrin, and S. Wicker, "Modeling Data Centric Routing in Wireless Sensor Networks," in *IEEE INFOCOM*, New York, June 2002.

[100] A. Perrig, J. Stankovic, and D. Wagner, "Security in wireless sensor networks," *Communications of the ACM (CACM), Wireless sensor networks, Special Issue: Wireless sensor networks,* vol. 47, pp. 53-57, June, 2004.

[101] Avancha S. et al, "Wireless Sensor Networks," *Kluwer Academic/Springer , Verlag Publishers,* 2003.

[102] E. Shi and A. Perrig, "Designing Secure Sensor Networks," *IEEE Wireless Communications,* pp. 38- 43, December, 2004.

[103] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks:attacks and countermeasures," *Elsevier's Ad Hoc Networks Journal, Special Issue on Sensor Network Applications and Protocols,* vol. 1, pp. 293-315, 2003.

[104] N. Bulusu, D. Estrin, L. Girod, and J. Heidemann, "Scalable coordinstion for wireless sensor networks: self-configuring localization system," in *International Symposium on Comunication theory and application (ISCTA 2001)*, Ambleside , UK, July 2001.

[105] B. Krishnamachari, "On the Complexity of Distributed Self-Configuration in Wireless Networks," *Kluwer Academic Springer Publishers, Telecommunication Systems,* vol. 22, pp. 33-59, 2003.

[106] Z. Wang and J. Crowcraft, "QoS-based Routing for Supporting Resource Reservation," *IEEE Journal on Selected Area of Communications,* Sept 1996.

[107] Z. Wang, Internet QoS : architectures and mechanisms for quality of service. San Francisco: Morgan Kaufmann, 2001.

[108] Q. Ma and P. Steenkiste, "Quality-of-Service routing with Performance Guarantees," in *Proc. of the 4th IFIP Workshop on Quality of Service*, May 1997.

[109] Y. Ge and J. Hou, "An analytical model for service differentiation in IEEE 802.11," in *ICC '03*, 2003.

[110] F. Xia, "Review: QoS Challenges and Opportunities in Wireless Sensor/Actuator Networks " *Sensors,* vol. 8, pp. 1099-1110 21 February 2008.

[111] El-Gendy, M. A. Bose, A. Shin, and K. G, "Evolution of the Internet QoS and support for soft real-time applications," *the IEEE,* vol. 91, pp. 1086-1104, 2003.

[112] Y. J. Li, C. S. Chen, Y.-Q. Song, and Z. Wang, "Real-time QoS support in wireless sensor networks: a survey," in *7th IFAC Int Conf on Fieldbuses & Networks in Industrial & Embedded Systems (FeT'07)*, Toulouse, France, Nov. 2007.

[113] By Larry L. Peterson, Bruce S. Davie, and Contributor Bruce S. Davie, *Computer Networks: A Systems Approach*: Morgan Kaufmann, 2003.

[114] C. System, "Cisco system,," vol. 2004, pp. 1-19, 2001.

[115] "Configuring Custom Queuing," http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t5/qcfcq.htm, May, 2005.

[116] Syngress Media Inc, Staff, Wayne Lawson Inc, S. Media, and R. Benn, *Configuring Cisco AVVID: Architecture for Voice, Video and Integrated Data*: Elsevier, 2000.

[117] Anthony T. Velte and Toby Velte, *Cisco: A Beginner's Guide*: McGraw-Hill Professional, 2004.

[118] "CBWFQ (Cisco IOS Software Release 12.0.5T)," in *http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t5/cbwfq.htm*, March, 2005.

[119] "LLQ (Cisco IOS Software Release 12.0.7T)," http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t7/pqcbwfq.htm, Ed., Jan, 2003.

[120] E. Glenbe, G. Pujolle, and J. C. C. Nelson, *Introduction to Queueing Networks*: John Wiley and Sons, 1987.

[121] M. Tanner, *Practical Queueing Analysis*: The IBM Mcgraw Hill Series, 1995.

[122] M. Tanner, "Practical Queueing Analysis," in *The IBM Mcgraw Hill Series*, 1995, pp. 337-339.

[123] Bose Sunjay K, "An Introduction to Queueing Systems," *Springer,* 2002.

[124] A. J. Weststrate, "Analysis and Optimization of Polling Models," PhD Thesis, Feb. 1992.

[125] W. P. Gronenenkijk, "Conservation Lawsin Polling Systems," PhD Thesis, University of Utrecht, 1990.

[126] H. Takagi, "Analysis of Polling Systems.," Research Reports and Notes Computer System Series The MIT press, 1986.

[127] Mobile Streams, "www.mobilegprs.com," Newbury, Berkshire, ENGLAND, 2002.

[128] GSM Association, " www.GSMWorld.com," the world wide web site of the GSM Association, 2005.

[129] H. Holma and A. Toskala, *WCDMA for UMTS : radio access for third generation mobile communications*, 3rd ed. Hoboken, N.J.: John Wiley & Sons, 2004.

[130] J. Korhonen, *Introduction to 3G mobile communications*, 2nd ed. Boston, MA: Artech House, 2003.

[131] C. Smith and D. Collins, *3G wireless networks*. New York: McGraw-Hill, 2002.

[132] Nokia, "A History of Third Generation Mobile 3G , www.nokia.com," 2004.

[133] 3GPP, "www.3GPP.org," 2005.

[134] "NIST http://snad.ncsl.nist.gov/nistnet/."

[135] "J-Sim. http://www.j-sim.org/."

[136] "J-Sim wireless extension tutorial. http://www.j-sim.org/v1.3/wireless/wirelesstutorial.htm."

[137] "OPNET. http://www.opnet.com."

[138] O. J. Boxma and B. W. Meister, "Waiting-Time Approximations in Multi-Queue System with Cyclic Service," *Performance Evaluation 7,* pp. 59-70, 1987.

[139] Winands, A. E.M.M., I.J.B.F., and G. J. v. Houtum, "A Two-Queue Model with Alternating Limited Service and State-Dependent Setups," *Analysis of Manufacturing System-Production Management,* pp. 200-208, 2005.

[140] O. J. Boxma and W. P. Groenendijk, "Two Queues with Alternating Service and Switching Times."

[141] O. J. Boxma, "Two Symmetric Queues with Alternating Service and Switching Times," *E. Gelenbe (ed.) Performance 84, North Holland, Amsterdam,,* pp. 409-431.

[142] W. Chang and D. G. Down, "Asymptotics for Polling Models with Limited Service Policies," *IEEE Conference on Decision and Control, Sydney, Australia,,* pp. 1427-1432, Dec. 2000.

[143] Leonard Kleinrock, " Queueing Systems," *John Wiley and Sons,* vol. 1, 1975.

[144] A. A. Rizvi and A. Hussain, "Analysis of Single Server Queueing System with Heavy Tailed Distributions," in *IEEE INMIC*, 2003, pp. 176-181.

[145] C. Yunn-Kuang and J.-C. Ke, "Interval Estimation of Mean Response Time for a G/M/1 Queueing System," *Empirical Laplace Function Approach Methods in Applied Sciences,* pp. 707-715, 2007.

[146] J. A. Stankovic, "Research Challenges for Wireless Sensor Networks," SIGBED Review: Special Issue on Embedded Sensor Networks and Wireless Computing, 1(2), July 2004.

[147] R. Sivakumar, P. Sinha, and V. Bharghavan, "CEDAR: Core Extraction Distributed Ad Hoc Routing Algorithm," *IEEE Journal on Selected Areas of Communications,* vol. 17, pp. 1454–1465, 1999.

[148] S. Chen and K. Nahrstedt, "Distributed Quality-of-Service Routing in Ad Hoc Networks," *IEEE Journal on Selected Areas of Communications,* vol. 17, pp. 1488–1505, 1999.

[149] B. Hughesand and V. Cahill, "Achievingreal-timeguaranteesinmobileadhocwirelessnetworks," in *Proceedings of the Work-*

in-Progress session of the 24th IEEE Real-Time Systems Symposium, December 2003.

[150] B. Deb, S. Bhatnagar, and B. Nath, "ReInForm: Reliable Information Forwarding Using Multiple Paths in Sensor Networks," in *Proceedings of IEEE International Conferenceon Local Computer Networks.*, pp. 406 – 415.

[151] F. Stann and J. Heidemann, "RMST: Reliable Data Transport in Sensor Networks," in *Proceedings of IEEE International Workshop on Sensor Network Protocols and Applications*, 2003, pp. 102–112.

[152] A. E. Gamal, J. Mammen, B. Prabhaker, and D. Shah, "Throughput-delay trade-off in wireless networks," in *IEEE INFOCOM*, March 2004.

[153] J. Li, C. Blake, D. S. D. Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *MobiCom 01, Pages 61-69* New York, USA: ACM Press, 2001.

[154] P. Gupta and P. R. Kumar, "Capacity of wireless networks," *IEEE Trans. on Information Theory,* pp. 388-404, March 2000.

[155] G. zeng, H. Zhu, and I. Chlamtec, "A Novel Queueing model for 802.11 Wireless LANs," in *WNCG Wireless Network Symposium*, 2003.

[156] O. Tickoo and B. Sikdar, "A Queueing model for finite load IEEE 802.11 random access MAC," in *IEEE ICC*, Paris, France, 2004.

[157] M. Ozdemir and A. B. McDonald, "An M/MMG1/1/K queuing model for IEEE 802.11 ad hoc networks," in *Proceeding of PE-WASUN'05*, 2004, pp. 107-111.

[158] N. Bisnik and A. Abouzeid, "Queuing Network Models for Delay Analysis of Multihop Wireless Ad Hoc Networks," in *The 2006 International Wireless Communications and Mobile Computing Conference (IWCMC 2006)* Vancouver, Canada: IEEE Communication Society and ACM, July, 2006.

[159] K. Akkaya and M. Younis, "An Energy-Aware QoS Routing Protocol for Wireless Sensor Network," in *Proceedings of the 23 rd International Conference on Distributed Computing Systems Workshops (ICDCSW'03)*, 2003.

[160] R. Nelson, "Probabililty, Stochastic Processes, And Queueing Theory, The Mathematics of Computer Science Modeling," *Springer-Verlag New York,Inc. ,* pp. 283-323, 1995.

[161] J.D. Little, "A Proof of Queueing Formula L=λW," *Operations Research, 9, pp 383-387,* 1961.

[162] B. R. Haverkort, "Performance of Computer Communication System," John Wiley and Sons, 1998, pp. 115-132.

[163] D. Gross and C. M. Harris, *Fundamentals of Queueing theory* vol. Third Edition: John Wiley and Sons Inc. , 1998.

[164] M. Iftikhar, M. Caglar, and B. Landfeldt, "An Analytical model based on G/M/1 with Self-Similar input to provide end-to-end QoS in 3G networks " in *IEEE/ACM Mobiwac 06 (MSWIM)*, Torremolinos, Malage, Spain, 2nd Oct 2006.

[165] J. B. Andresen, T. S. Rappaport, and S. Yoshida, "Propagation Measurements and Models for Wireless Communications Channels," in *IEEE Communications Magazine*. vol. 33, January 1995.

[166] M. Gerla, G. Pei, and S.-J. Lee, "Wireless, Mobile Ad-Hoc Network Routing," in *Proceedings of IEEE/ACM FOCUS'99*, New Brunswick, NJ, , May 1999.

[167] "Data sheet for the Acoustic Ballistic Module," SenTech Inc., http://www.sentechacoustic.com.

[168] G. J. Pottie and W. J. Kaiser, "Wireless Integrated Network Sensors," *Communication of the ACM,* vol. 43, pp. 51-58, May 2000.

[169] W. Odom and M. J. Cavanaugh, "IP Telephony Self-Study Cisco DQoS, Exam Certification Guide," *Cisco Press,,* pp. 3-314, 2004.

[170] "Configuring Custom Queuing," http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t5/cbwfq.htm, March, 2005.

[171] J. A. Weststrate, "Analysis and Optimization of Polling Models," PhD Thesis, Feb. 1992.

[172] D. Gross and C. M. Harris, "Fundamentals of Queueing theory," vol. Third Edition, pp. 211-213, 1998.

[173] M. Y. Aalsalem, J. Taheri, M. Iftikhar, and A. Y. Zomaya, "Providing QoS Guarantees to Multiple Classes of Traffic in Wireless Sensor Networks," in *The sixth ACS/IEEE International Conference on Computer Systems and Applications (AICCSA-08)*, Doha, Qatar, March 31 - April 4, 2008

[174] J. Stankovic, T. Abdelzaher, L. Chenyang, L. Sha, and J. C. Hou, "Real Time Communication and Coordination in Embedded Sensor Networks," *IEEE,* vol. 91 No.7, July 2003.

[175] A. Jan, "Analysis and Optimization of Polling Models," PhD Thesis, Feb. 1992.

[176] M. Y. Aalsalem, M. Iftikhar, J. Taheri, and A. Y. Zomaya, "On the Provisioning of Guaranteed QoS in Wireless Sensor Networks through Limited Service Polling Models," in *The fifth IEEE and IFIP International Conference on wireless and Optical communications Networks (WOCN2008)* Surabaya, East Java Indonesia: IEEE, 2008.

[177] J. J. F. Hair, R. E. Anderson, R. L. Tatham, and W. C. Black, *Multivariate Data Analysis With Reading*: Prentice-Hall, Inc, 1995.

[178] E. H. Callaway, *Wireless sensor networks : architectures and protocols*. Boca Raton, Fla. ; London: Auerbach Publications, 2004.

[179] S. Krco, D. Cleary, and D. Parker, "P2P mobile sensor networks," in *Proceedings of the 38th Annual Hawaii International Conference*, Jan, 2005.

[180] M. Y. Aalsalem, D. Everitt, and A. Y. Zomaya, "An Interoperability Framework for Sensor and UMTS Networks," in *Third International Conference on Wireless and Mobile Communications (ICWMC'07)* Guadeloupe: IEEE Computer Society and IEEE Xplore, March 2007.

[181] 3GPP, "Multimedia Messaging Service(3GPP TS 22.140)," 2001.

[182] 3GPP, "Multimedia Messaging Service Architecture and functionality(3GPP TS 23.140)," 2001.

[183] 3GPP, "TS 23.003 V4.3.0 "Numbering, Addressing and Identification."

[184] 3GPP, "TS 23.140 V6.3.0 "Multimedia Messaging Service (MMS), Functional description, Stage 2 "," 2003-09.

[185] 3GPP, "MMS ArchOverview," 2001.

[186] P. Resnick, "Internet Message Format," in RFC2822. URL:*http://www.ietf.org/rfc/rfc2822.txt*, April 2001.

[187] T. S. Rappaport, *Wireless Communication Principles and Practice*: , Prentice-Hall, 2002.

[188] O. M. A. (OMA), "Wireless Application Protocol Architecture Specification," in *WAP-210-WAPArch-20010712-a*. vol. 2006: OMA, 2006.