Design of Implicit Routing Protocols for Large Scale Mobile Wireless Sensor Networks

by

Tsung-Ta Wu

A Thesis

presented to the University of Strathclyde

in fulfilment of the requirements for the degree of Doctor of Philosophy

in

Department of Electronic and Electrical Engineering

Centre for Intelligent Dynamic Communications

September 2011

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Tsung-Ta Wu September 2011

To

Dad, Mum

And

Chen-I Kuo

Acknowledgements

My heartfelt thanks to Dr. Kae Hsiang Kwong for his patience and teaching through example; without him, I could not have reached this stage of my career. I also thank my supervisors Prof. Ivan Andonovic and Dr. Craig Michie for their enthusiastic encouragement, guidance and sincere support throughout the PhD.

Over the years at Glasgow, I have made many friends. In particular, I want to thank Karen, Angeline, Hooi Ping and Christine, for they have always helped me one way or another especially when I need it most. My years in Glasgow were much warmer and memorable because of them. I also want to thank all the colleagues with whom I worked with and who have helped me over the years. In particular, thanks to Kostas, Swee, Gavin, Charlie, Leo, Liu. And a big thanks to Hock Guan Goh for his inspiration and encouragement to pursue the goal with determination.

Above all, I owe tremendous dept of gratitude to my parents, Wu Ching-Mao and Wu Chen Chun-Ying. Through their constant support in all respects, they guide me to excel both in life and in work. Their success and love drive me to accomplish any goal that I set.

Finally, and always, I thank my beautiful wife, Chen-I Kuo. Without her love, dedication and sacrifice, nothing is ever possible.

Abstract

Most developments in wireless sensor networks (WSNs) routing protocols address static network scenarios. Schemes developed to manage mobility in other mobile networking implementations do not translate effectively to WSNs as the system design parameters are markedly different.

Thus this research focuses on the issues of mobility and scalability in order to enable the full potential of WSNs to self-organise and co-operate and in so doing, meet the requirements of a rich mix of applications. In the goal of designing efficient, reliable routing protocols for large scale mobile WSN applications, this work lays the foundation by firstly presenting a strong case supported by extensive simulations, for the use of implicit connections. Then two novel implicit routing protocols - Virtual Grid Paging (VGP) and Virtual Zone Registration and Paging (VZRP) - that treat packet routing from node mobility and network scalability viewpoints are designed and analysed. Implicit routing exploits the connection availability and diversity in the underlying network to provide benefits such as fault tolerance, overhead control and improvement in QoS (Quality of Service) such as delay. Analysis and simulation results show that the proposed protocols guarantee significant improvement, delivering a more reliable, more efficient and better network performance compared with alternatives.

Contents

List of Figures

List of Tables

Abbreviations

ACK	Acknowledgment
AIMRP	Address-light, Integrated MAC and Routing Protocol
AODV	Ad-hoc On-demand Distance Vector
ARP	Address Resolution Protocol
BFS	Breadth-First Search
BRP	Bordercast Resolution Protocol
BS	Base Station
CBR	Continuous Bit Rate
CDMA	Code Division Multiple Access
CO	Communication Overhead
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
CTS	Clear To Send
DD	Direct Diffusion
DSDV	Destination-Sequenced Distance-Vector routing
DSR	Dynamic Source Routing
DV	Distance Vector
ECA	explicit connection approach
ECD	Explicit Connection Duration
ECI	Explicit Connection Interval
FDMA	Frequency Division Multiple Access
FIFI	First In First Out
FSR	Fisheye State Routing

GAF	Geographical Adaptive Fidelity
GloMoSim	Global Mobile Information System Simulation Library
GPS	Global Positioning System
IARP	IntrAzone Routing Protocol
ICD	Implicit Connection Duration
IC	Implicit Connectivity
ICI	Implicit Connection Interval
ICOM-TD	Implicit Cluster-based Overlay Multicast protocol exploiting Tree Division
IDR	Implicit Distance-based Routing
IERP	IntErzone Routing Protocol
IFq	InterFace queue
LL	Link Layer
MAC	Medium Access Control
MANET	Mobile Ad-hoc Network
MSC	Mobile Switch Centre
NetIF	Network InterFace
NAM	Network Animator
NDP	Neighbour Discovery Protocol
OMNeT++	Objective Modular Network Testbed in C++
OPNET	Optimized Network Engineering Tool
PAMAS	Power Aware Multi-Access protocol with Signalling for Ad Hoc Networks
PCS	Personal Communication Service

PDA	Personal Digital Assistant
QoS	Quality of Service
RAM	Random Access Memory
RPS	Radio Paging System
RREQ	Route Request
RREP	Route Reply
RRER	Route Error
RTS	Ready To Send
S-MAC	Sensor-MAC
TDMA	Time Division Multiple Access
TEEN	Threshold sensitive Energy Efficient sensor Network
TTDD	Two-Tier Data Dissemination
VGP	Virtual Grid Paging
VZRP	Virtual Zone Registration and Paging
WSN	Wireless Sensor Network
ZRP	Zone Routing Protocol

Chapter 1

Introduction

1.1 Overview

Typically, in Wireless Sensor Networks (WSNs) [Akyildiz et al. 2002; Karl and Willig 2005], a wireless node communicates with a remote sink using multi-hop communication through an explicit (predetermined and fixed) routing path. Many traditional routing protocols such as Direct Diffusion [Intanagonwiwat et al. 2003], Ad-hoc On-demand Distance Vector (AODV) [Perkins and Royer 1999], Dynamic Source Routing (DSR) [Johnson et al. 2001] focus on finding the most appropriate and reliable explicit routing path. However, as node mobility and node density increase, the network topology changes rapidly resulting in a decrease in the robustness of the explicit routing path. As a consequence, a significant amount of routing overhead is introduced to maintain such vulnerable routing paths.

Any routing overhead is undesirable and needs to be controlled to avoid unnecessary energy dissipation and decrease in network performance. In recent years, many positioning technologies [Albowicz et al. 2001; Chehri et al. 2008; Cheng et al. 2004; Doherty et al. 2001; He et al. 2003] have been proposed utilising WSNs. For any physical position of nodes, geophysical position information can be integrated and used to route the packet toward its destination via an implicit (totally dynamic, the forwarding decision is made in real-time) routing path without any prior knowledge of neighbouring connections. In contrast to explicit routing, implicit routing has distinct benefits in establishment-free and maintenance-free operation; the use of implicit routing conserves energy and bandwidth.

The thesis explores solutions that facilitate implicit routing in WSNs. Two implicit routing protocols are proposed addressing two major issues, node mobility and network scalability. Although the aforementioned issues have not hitherto been investigated together within the scope of WSNs, they do play an important role when considering any sophisticated application such as the "Smart City" [Ahn et al. 2009; Al-Hader et al. 2009; Filipponi et al. 2010] concept consisting of a massive number of nodes (information providers) and sinks (end users) scattered over a large area with all participating nodes assumed to be mobile. To the best of our knowledge, no such study has been attempted thus far.

The research first examines routing performance analysis with both explicit and implicit connection solutions under different degrees of node mobility and network scalability. In an attempt to improve routing performance, two implicit routing protocols are developed and analysed. Details of protocols operation and their performance are then given.

Chapter 2 provides the background research core to the understanding of the subsequent chapters. It starts with a survey on WSN discussing the hardware platform, architectures and protocols currently in use. It also presents the system model that captures the network scenario and assumption used in the subsequent analyses. Finally, it provides the definitions of mobility and scalability, the main performance parameters examined throughout the dissertation.

Chapter 3 starts by presenting the state-of-the-art in routing for wireless networks. Routing protocols are categorised into explicit [Intanagonwiwat et al. 2003; Johnson et al. 2001; Perkins and Royer 1999] and implicit [He et al. 2003b; Tanenbaum 1981] connection approaches, followed by a detailed definition of their design principles. A series of analysis metrics used to capture and quantify the impact of node mobility and network scalability on network performance are then presented. Comprehensive simulation based evaluations are carried out to show the potential and benefit of the implicit connections.

Chapter 4 presents two novel implicit routing protocols referred to as Virtual Grid Paging (VGP) and Virtual Zone Registration and Paging (VZRP). It starts with discussion of two well-known, established systems viz. Radio Paging Systems [Molisch 2005; Rappaport 2002] and cellular networks [Molisch 2005; Rappaport 2002] that inspire the development of VGP and VZRP, respectively. It also introduces a critical topology control algorithm - the Geographical Adaptive Fidelity (GAF) [Xu et al. 2001] - that functions in tandem with the proposed routing schemes. The rest of chapter 4 focuses on a definition of the principles underpinning VGP and VZRP.

Chapter 5 evaluates the routing performance of VGP and VZRP; the evaluation is based on both analysis and simulation. The VGP scheme seems to be promising as results show that it provides high reliability in delivering packets. Both numerical analysis and simulation results confirm that VZRP is a highly efficient solution for delivering packets.

Chapter 6 concludes by summarising the contributions, highlighting important results followed by suggestions for future work.

1.2 Contributions

The contributions of this thesis are threefold. Firstly, a series of analysis metrics are proposed that capture and quantify the impact of node mobility and network scalability on network performance. The proposed matrix allow network performance to be evaluated from the perspective of the physical layer directly. The metrics can be generally applied to any mobile network to predict optimised routing performance under specific node mobility and network scalability while isolating the influence of upper layer routing protocols; results obtained serve as design guidelines when deploying a wireless mobile network. These analysis metrics are also used to assess the potential of implicit connections, the core focus of the research.

Secondly, the thesis presents two novel implicit routing protocols viz. VGP and VZRP. They are designed to fulfil the needs of wireless routing protocols for mobility handling, routing overhead control, routing reliability and network scalability. Compared with many traditional explicit routing protocols, including

Direct Diffusion, AODV and DSR, they naturally adapt to changes in network topology even when operated in a network consisting a large number of mobile nodes. To support reliable routing in such large, scalable and mobile networks is a central contribution.

Thirdly, in order to evaluate the performance of proposed protocols and compare with general explicit connection approaches (ECAs) in depth, analysis models are developed and used to examine the simplified routing overhead for each approach. In particular the model for ECAs also captures the impact of node mobility and network scalability on routing performance, hitherto never been quantified from that perspective. Performance evaluations are conducted through extensive simulations and results confirm that the proposed implicit routing protocols can be of value in large scale mobile WSNs.

A summary of the contributions follows;

- Analysis metrics are defined that support the analysis of routing performance in highly mobile environments in general (refer to Section 3.3.1).
- Capture of the routing performance for both explicit and implicit connection scenarios under varied mobility and network scale by utilising the above analysis metrics under OMNeT++ simulator (refer to Section 3.3.2).
- Introduction of an implicit routing protocol, VGP, to achieve high reliability in large scale mobile WSNs (refer to Section 4.3).

- Introduction of a second implicit routing protocol, VZRP, to achieve high efficiency in large scale mobile WSNs (refer to Section 4.4).
- Routing overhead analysis for proposed implicit routing protocols, VGP and VZRP (refer to Section 5.2).
- Routing overhead analysis for general explicit routing protocols (refer to Section 5.2).
- A comprehensive routing performances study for VGP and VZRP using the NS-2 simulator (refer to Section 5.4).

There is a Journal paper [Kwong et al. 2011] and three conference papers [Di et al. 2010; Wu et al. 2010a; Wu et al. 2010b] are published; partial of the presented results are derived from the analysis metrics conducted in chapter 3. The main body of this research, chapter 4 and chapter 5 are looking forward to be published in the near future.

Chapter 2

Background

2.1 Introduction

This chapter provides the background and summarises related research in Wireless Sensor Networks (WSNs). A survey on WSNs is followed by a definition of mobility, scalability and a system model is introduced as a foundation to the analyses contained in subsequent chapters. The chapter concludes with a summary that highlights the pros and cons of mobility models and states the research challenges for designing a routing protocol able to support mobility and scalability.

2.2 A Survey on Wireless Sensor Network

Recent progress in wireless communication and digital electronics has enabled the development of low-cost, low-power, multi-functional sensor nodes which are small in size and capable of communicating across short distances. The combination of these developments has stimulated widespread research and deployment of wireless sensor network (WSN) concepts [Akyildiz et al. 2002; Karl and Willig 2005; Romer and Mattern 2004]. In general, a WSN consists of spatially distributed autonomous devices, so called “nodes”, using sensors to cooperatively monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants at different locations. One of the unique characteristic of WSNs are that they are based on the collaborative effort of a large number of nodes to fulfil tasks as, usually, a single node is incapable of doing so; and they use wireless connections to enable this collaboration.

The following sections describe the fundamentals of WSNs beginning with their physical components and system hierarchy. Then some traditional applications are highlighted and topologies explored. The chapter concludes by outlining the communication architecture and protocols developed for each layer are discussed.

2.2.1 WSN Components and System Hierarchy

Establishing a WSN requires the constituent nodes to meet the needs derived from the specific requirements of any given application. Generally, they are small form factor, low cost, energy efficient, equipped with sensors that fit the requirements of the application, provide the necessary computation and memory resources to execute on tasks that enable the application, comprise communication facilities to form a collaborated network over wireless channels and a power supply is needed to provide energy for all on-node operations [Altaan 2010]. The basic sensor node comprising the five functional components [Akyildiz et al. 2002], their relationship and composition is depicted in Figure 2.1.

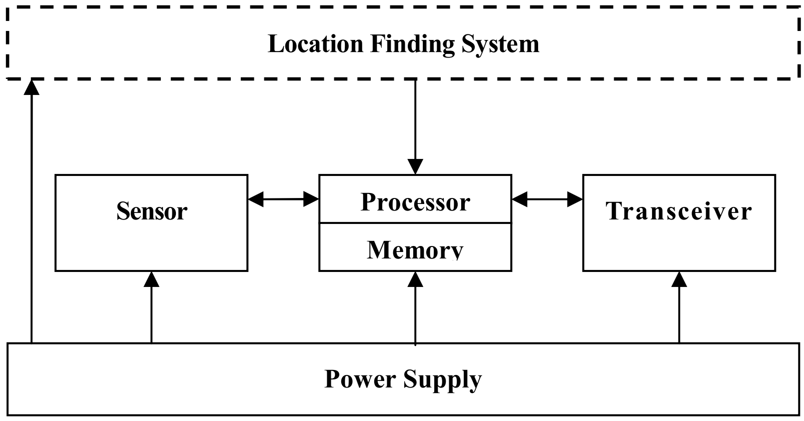

Figure 2.1 Components of a wireless sensor node.

- *Processor:* processes all relevant data, capable of executing arbitrary code.
- *Memory:* to store programs and intermediate data; usually different types of memory are used. Generally, RAM is used for program and Flash is used for data storage.
- *Sensor:* interface to the physical world. Components that can observe physical parameters of the environment or detect certain physical changes of monitoring objects.
- *Transceiver:* connecting nodes into a network require interfaces for sending and receiving information over a wireless channel.
- *Power supply:* as no tethered power supply is available, typically, batteries are used to provide energy.

The node may also have application dependent additional components such as a location finding system [Akyildiz et al. 2002]. Most sensor network routing techniques and sensing tasks require knowledge of location with high accuracy [Akyildiz et al. 2002]. Location information may be acquired through Global

Positioning System (GPS) technology [Parkinson and Spilker 1996]. If required. the sensor node board MTS420CC manufactured by Crossbow [Crossbow 1995] can be used for sensor nodes to obtain their physical location with high accuracy (error ±5m). However, the usage of GPS introduces additional cost and consumes considerable energy to operate. To avoid such a limitation, much effort has been devoted to estimating the abstract location through radio without GPS [Albowicz et al. 2001; Doherty et al. 2001; Bahl and Padmanabhan 2000; He et al. 2003a; Cheng et al. 2004].

A typical WSN consists of three hierarchical entities: sensor node, the sink and task management node (Figure 2.2). A sizeable number of sensor nodes are typically scattered around the target field to collect data and forward it back to the task management node for further computations and interaction with users [Akyildiz et al. 2002]. The ‘sink’ acts as an intermediary between sensor nodes and task management node by providing higher communications capabilities and a larger power resource [Akyildiz et al. 2002]. The sink may communicate with the task manager node via the Internet or satellite. Note that Figure 2.2 simply shows an example of how the three entities collaborate together. In practice, the formation of the network is application dependent.

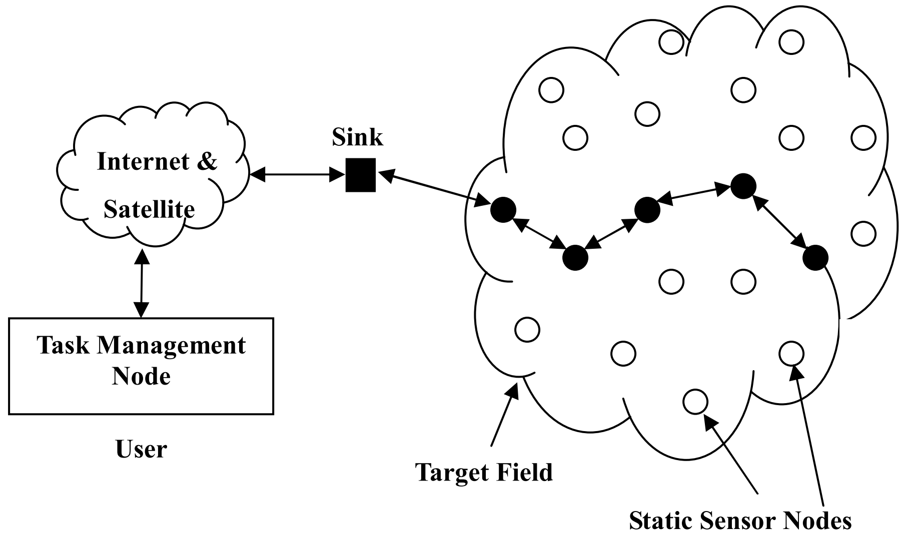

Figure 2.2 Sensor nodes scattered in a target field.

The scope of this study is focussed on the interaction between source nodes and sink. In general, source nodes are defined as the scattered sensor nodes that generate data; the sink is defined as the node where the data is delivered. The detailed definition of both source node and sink are given in the next Section.

2.2.2 Sensor Network Scenarios and Types of Applications

Section 2.2.1 has introduced a typical WSN system hierarchy and has briefly touched on the definition of source node and sink. The Section discusses WSN scenarios comprising different types of source nodes and sinks, types of communication paradigms and types of mobility.

A source is any scattered node in the network that generates data; a sink is the destination that data is transported to. There are essentially three types of sink [Karl and Willig 2005]; one of sensor nodes constituting the network [Intanagonwiwat et al.

2003]; or a device outwith the network, such as a handheld Personal Digital Assistant (PDA) used to interact with the sensor network [Chang et al. 2007]; or a gateway to another larger network such as the Internet [Chang and Tassiulas 2004; Kulkarni et al. 2006] (Figure 2.2).

In relation to WSNs, communication paradigms can be classified as; one-to-one [Karp and Kung 2000], one-to-many [Luo et al. 2003], many-to-one [Kulkarni et al. 2006], and many-to-many [Chang and Tassiulas 2004; Kweon et al. 2009]. The simplest is one-to-one comprising one source node and one sink, the data flow being always from source to sink. In many cases, there are multiple source nodes and/or multiple sinks [Karl and Willig 2005]. The most challenging case would be many-to-many; multiple source nodes sending information to multiple sinks, where either all or some of the data has to reach all or some the sinks. Figure 2.3 illustrates four communication paradigms.

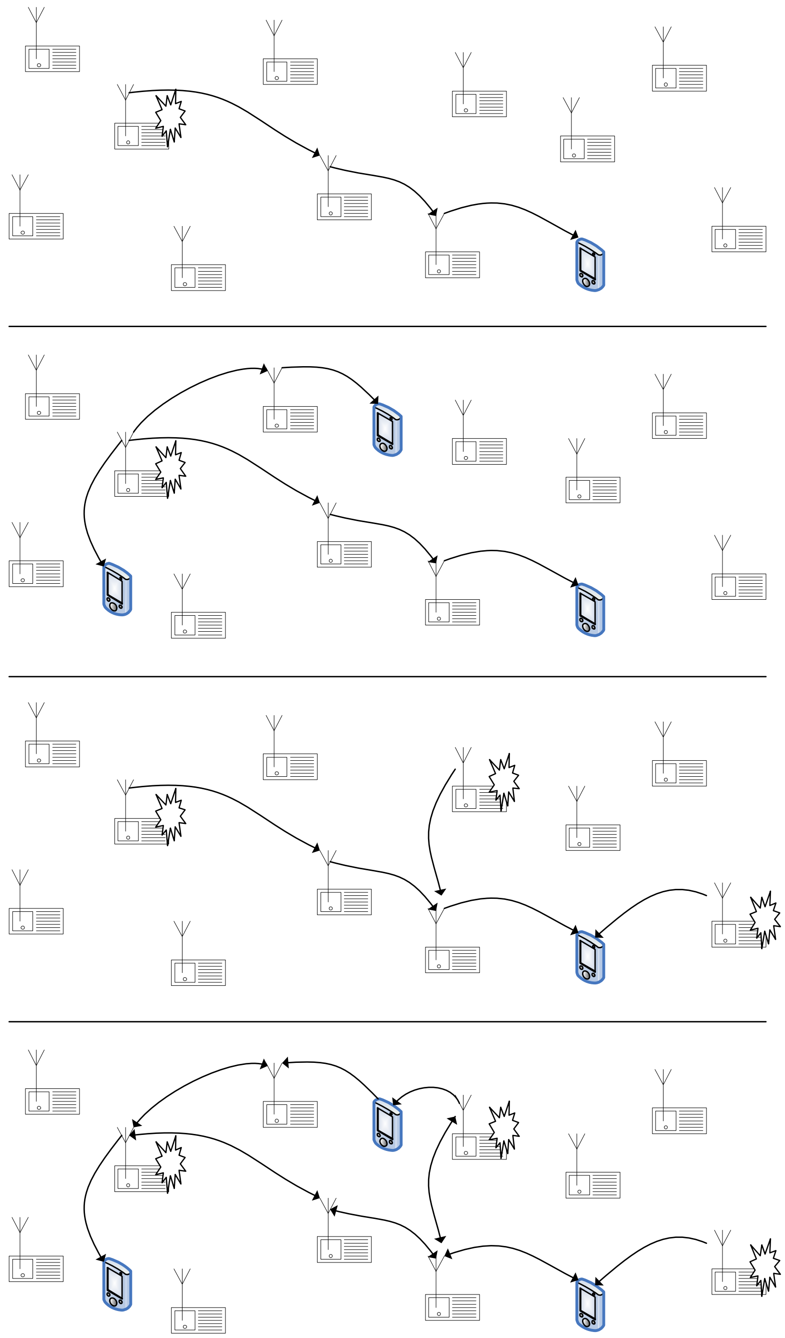

Figure 2.3 Four communication paradigms; from the top, are one-to-one, one-to-many, many-to-one and many-to-many.

In the scenarios discussed above, all participants are stationary; however one of the main features of wireless communication is the ability to support mobile participants.

In WSNs, mobility can appear in three main forms [Karl and Willig 2005];

- *Node mobility:* sensor nodes themselves can be mobile. The impact of this type of mobility is highly application dependent. In examples like environmental control e.g. sensor nodes deployed on/in the ground to measure soil moisture, the nodes are static [Sikka et al. 2006]; in livestock monitoring e.g. sensor nodes attached to cattle to record activity, mobility is one of the main design challenges [Wark et al. 2007].
- *Sink mobility:* data sinks can also be mobile. This can be viewed as a special case of node mobility, if the sinks themselves are part of the WSN [Huang et al. 2006]. The important aspect here is the mobility of a sink when it is not part of the sensor network e.g. a user requesting information via a PDA while walking in a building [Zhang et al. 2009].
- *Event mobility:* In applications like event detection [Doumit and Agrawal 2002] and particularly in tracking applications [Shen et al. 2001], the causes of the event or the objects to be tracked are usually mobile. This can also be viewed as a special case of node mobility if the sensor node is attached to the studied object i.e. the object is not a part of a network e.g. sensor network as applied in battlefield scenarios to track the location of enemy tanks [Luo et al. 2003].

Communication protocols for WSNs will have to render appropriate support for these different forms of mobility. In this thesis, the most challenging case is considered which assumes that all entities including sensor nodes and multiple sinks are mobile.

The effect of mobility under such conditions has rarely been studied in depth. Detailed definitions underpinning the analyses and description of the network scenario to be analysed are provided in Section 2.3.

The remainder of this section discusses the most relevant interaction patterns between source nodes and sinks. The types of applications representative of these interactions can be categorised as follows [Karl and Willig 2005];

- *Event detection:* sensor nodes only report to the sink once they have detected the occurrence of a predefined event [Bonnet et al. 2000]. The simplest events can be detected locally by a single sensor node in isolation e.g. when a temperature threshold is exceeded [Shen et al. 2001]; more complicated types of events require the collaboration of a cluster of nearby sensor nodes to decide whether a composite event has occurred e.g. when a temperature gradient becomes too steep [Li et al. 2003].
- *Periodic measurements:* sensor nodes tasked with periodic reporting of measured values [Goense et al. 2005; Mainwaring et al. 2002; Tolle et al. 2005]. These reports can be triggered by a detected event or queried from the sink; the reporting period and interval is application dependent.
- *Tracking:* the event of interest can be mobile e.g. an enemy in surveillance scenarios within a battlefield [Luo et al. 2003; Romer 2004]. The WSN can be used to report updates on the event's location to the sink(s), potentially with estimates about speed and direction. To achieve this, typically sensor nodes have to cooperate before updates can be reported.

2.2.3 WSN Characteristics and Applications

Applications enabled by WSNs are bounded by the unique set of characteristics presented by the platform. Therefore it is important to understand these unique characteristics before a meaningful description of possible applications can be presented;

- *Self-organising:* one of the distinct characteristics of WSNs is that sensor nodes operate in a self-organising fashion [Akyildiz et al. 2002]. Sensor network protocols allow nodes to self-configure into an appropriate network and operate unattended once they are randomly (most likely) and densely deployed within the environment. As both the environment and the WSN itself change dynamically (node mobility, failing nodes, new tasks, depleted energy), the system has to adapt to maintain functionality.
- *Fault tolerance:* since nodes may run out of energy, or since the wireless channel between two nodes could degrade, it is vital that the network is tolerant to such faults [Wang et al. 2005]. Generally to mitigate such faults, the redundant deployment of sensor nodes is considered [Karl and Willig 2005]. However, this strategy impinges on many wireless networking issues: many nodes increase the potential of interference with each other [Papadimitriou et al. 2010], many additional routes are presented [Gomez and Campbell 2007], nodes might use greater transmission power to talk to distant nodes directly resulting in a reduction in the re-use of channel bandwidth [Monks et al. 2001], and routing protocols might have to be recomputed even for small node movements [Lim et al. 2002].

- *Lifetime:* sensor nodes are designed to operate against a design lifetime e.g. in excess of a year in some applications such as environmental control or 'as long as possible' [Khan et al. 2010]. In many scenarios, this criterion is the most challenging since sensor nodes are typically battery powered. Thus, energy management protocols are routinely invoked to extend the lifetime of the network/application [Chen et al. 2001]. One striking protocol – the Geographical Adaptive Fidelity (GAF) - developed by [Xu et al. 2001] will be described in Section 4.2.3 since its structure and functionality is explored and used within the body of this thesis.
- *Scalability:* many WSNs comprise a large number of nodes, either due to the large extent of the sensing environment or to provide a certain degree of fault tolerance; thus any resultant architectures and supporting protocols must be able scale [Li et al. 2011]. Scalability is one of major design considerations in this present work. In order to provide a foundation understanding, a detailed definition of scalability is revisited in Section 2.4.
- *Wide range of densities:* the 'density', defined as the number of nodes per unit area, can vary considerably to fit the requirement of different applications. Even within a given application, the density can vary over time and space as nodes fail or move [Xu et al. 2001]. The network protocol should be designed to adapt to such variations.

In this research, the goal of the routing protocols is to ensure that the wireless sensor platform conforms to the aforementioned characteristics, in so doing allowing applications as demanding as disaster relief and military operations;

- *Disaster relief applications:* one of the most often cited applications for WSNs are disaster relief operations [Cayirci and Coplu 2007; Fujiwara et al. 2004; Pogkas et al. 2005; Saha and Matsumoto 2007]. A typical scenario is wildfire detection: sensor nodes are equipped with thermometers that can determine their own location (relative to each other or in absolute coordinates). Sensor nodes may be strategically, randomly, and density deployed over, for example, a forest. They collectively produce a "temperature map" of the area that can be accessed remotely, for example, by firemen equipped with PDAs [Li et al. 2006]. Some of these disaster relief applications have commonalities with military applications where sensor nodes detect, for example, enemy troop movements [Doumit and Agrawal 2002].
- *Environment control and biodiversity mapping:* WSNs can be used to monitor the environment, for example, with respect to chemical pollutants [Ailamaki et al. 2003; Yang et al. 2002]. Closely related to environmental control is the understanding of plant and animals that live in given habitats [Mainwaring et al. 2002; Zhang et al. 2004].
- *Intelligent buildings:* WSNs have valuable applications in real-time monitoring of the temperature, airflow, humidity, and other physical parameters that consume significant power within large, dispersed office buildings [Kappler and Riegel 2004]. The goal is to increase the comfort level of inhabitants and reduce energy consumption. In addition, such sensor nodes can also be used to monitor the mechanical strength of buildings in seismically active zones [Chintalapudi et al. 2006]. For example, the sensor node can emit warning alerts, if it detects defects in parts of buildings which

represent weaknesses in the infrastructure and are consequently susceptible during earthquakes [Kurata et al. 2004]. Other types of sensors might be geared to detect people enclosed in collapsed buildings, providing valuable information to a rescue team [Cayirci and Coplu 2007].

- *Facility management:* WSN are used in management of facilities where the site is larger than a single building [Braun 2007; Malatras et al. 2008]. A related application is tracking in (say) airports for the control of passengers' access points [Karl and Willig 2005]. Such an application is beneficial for airline companies as it avoids delays caused by passengers mistakenly entering the wrong area or wrong terminal. Furthermore it has concomitant value in improving the security, for example restricting passenger entry in highly secure areas [Liu et al. 2007a]. A wide-area WSN tracks passenger's positions and issues alerts to local authorities; this application shares many concepts military applications.
- *Precision agriculture:* applying WSN to agriculture allows precise irrigation and fertilisation [Lea-Cox and Ross 2001] through distributed humidity/soil composition sensors across fields [Sikka et al. 2006]. Also, livestock health identification can be implemented by attaching a sensor(s) to each sheep or cow, continually monitoring the health status of individual animals (by checking body temperature, step count or activity), raising alerts if given thresholds are exceeded [Radenkovic and Wietrzyk 2006; Wark et al. 2007].
- *Medicine and health care:* each patient is fitted with small, lightweight wireless sensor nodes; each sensor node has its specific task e.g. one sensor node may be detecting heart rate while another is detecting the blood pressure

[Heinzelman et al. 2004]. The advantage of untethered sensors is considerable here. Doctors may also carry a sensor node, which allows other doctors to locate them within the hospital [Virone et al. 2006]. Such patient and doctor tracking systems within hospitals can be lifesaving.

This brief list of application scenarios highlights the vast design space [Romer and Mattern 2004].

2.2.4 Protocol Stack

A protocol stack used to facilitate the design of WSN-enabled applications is depicted in Figure 2.4; the stack consists of a power management plane, a mobility management plane, a task management plane governing the physical layer, the MAC (Medium Access Control) layer, the routing layer, the transport layer and the application layer [Akyildiz et al. 2002].

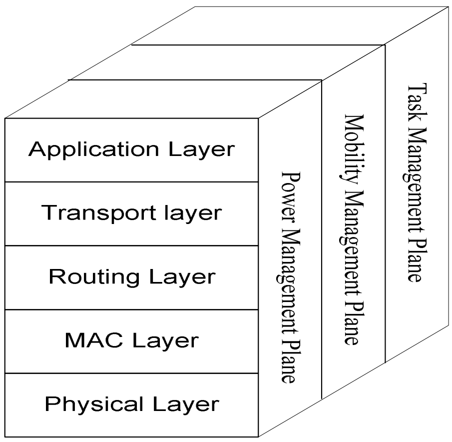

Figure 2.4 WSN protocol stack.

The power management plane manages the rule a sensor node uses its power e.g. the sensor node may turn off its receiver upon of a message from one of its neighbours [Xu et al. 2001]. This is to avoid generating duplicate messages, in so doing wasting energy. The sensor node conserves energy when its receiver is in sleep mode. Also, when the power level of a sensor node is low, it broadcasts to its neighbours that it can no longer participate [Akyildiz et al. 2002]. Sensor nodes with higher conserved energy will assume the task of routing resulting in better energy balancing over the network.

The mobility management plane detects and registers the movement of sensor nodes, so a route back to the user is always maintained, and the sensor node can keep track of neighbour nodes. By registering the position of neighbour nodes, on overall network balancing of individual node power and task usage can be achieved. The task management plane balances and schedules tasks to be executed within a specific application environment. Not all nodes are required to perform the sensing task at the same time. The task management plane also balances and schedules data forwarding. Not all sensors are required to participate in the task of transporting data under the networking protocol, avoiding unnecessary network overheads [Xu et al. 2001]. The management plane provides the strategy for the sensor nodes to collaborate to conserve power, route data with the minimum of overhead in dynamic mobile environments, and share resources between sensor nodes.

The goal of the physical layer is to provision a hardware platform flexible in supporting a range of suitable modulation schemes and transceiver architectures that are simple, low-cost, but robust enough to provide reliable connectivity [Abouei et al. 2011]. The MAC layer is responsible for resource allocation through managing functions such multiplexing of data streams, data frame detection, medium access and error control. The fundamental task of the MAC layer is to regulate the access of nodes to a shared medium in such a way that certain application-depended performance requirements (such as delay, throughput, fairness and energy conservation) are satisfied [Zhou et al. 2006a]. The routing layer governs the establishment of viable paths from source to sink and is used to implement a number of strategies, multi-hop being prominent in many applications [Rashid et al. 2008]. Whenever a source node cannot send a packet directly to its destination node viz. single hop but has to rely on the assistance of intermediate nodes to forward these packets on its behalf, a multi-hop network results. The design of routing protocols follow the core principles of power efficiency, data centric (the actual data is more valuable, rather than the identity of data sender), data aggregation, attribute-based addressing (querying an attribute of stimulus, rather than querying an individual node) [Shen et al. 2001] and location awareness. A detailed survey of existing routing protocols is provided in Section 3.2.

The transport layer is commonly assigned to provide reliability [Akan and Akyildiz 2005]. In the particular case of WSNs, the reliability refers not only to the delivery of data packets but also to the ability to detect physical phenomena. The problem of detection reliability associates highly with the coverage of sensing node network which is determined by the deployed node density [Liu and Towaley 2003; Liu and Towaley 2004; Huang and Tseng 2003]. Reliable data packet delivery requires the ability to detect and repair the loss of packets in predominately multi-hop environments. A number of algorithms have been proposed to enhance such reliability [Deb et al. 2003a; Deb et al. 2003b; Dulman et al. 2003]. Finally, the application layer is responsible for presenting all data to the application and propagating requests from the application layer down to the lower layers. The application layer software depends on the deployment and application [Alam et al. 2008].

The stack is not a function of any specific hardware or software so that an implementer is able to design their hardware and software to fit the requirements of the application. For example, the IEEE 802.15.4 [IEEE Standard 2003] standard specifies the physical and MAC layers, whereas Zigbee [Gislaso 2008] specifies the entire protocol stack. It is not in the scope of this thesis to examine various standards in depth.

2.3 System Model

Two major design issues - node mobility and network scalability - are addressed in the thesis for WSNs comprising a large number of sensor nodes and sinks. The traffic flow is many-to-many, as depicted in Figure 2.3. Both sensor nodes and sinks are assumed to be mobile and freely roam around a relatively large area; the length of any physical connection is potentially up to several hundreds or even thousands of meters. Due to the limits in resources provided by WSNs, the design of an effective routing protocol that facilitates such large scale multi-hop communication is a significant challenge.

The scenario studied can be represented as a dynamic graph $G = \{V, S, E(t)\}$, where V is the set of mobile nodes, S is the set of mobile sinks and $E(t)$ is the set of network edges at time t. Let $n = |V|$ be the number of mobile nodes and $m = |S|$ be the number of mobile sinks participating in a WSN. Node/sink $i \in V$ can hear node/sink $j \in V$ if i is within the radio range of j. Let $H(i)$ be a set of nodes/sinks which i can hear and $H(j)$ to be a set of nodes/sinks which j can hear. It is then obvious that i and j can hear each other if and only if $i \in H(j)$ and $j \in H(i)$.

Since the two identified design issues are not commonly investigated together within the scope of WSN, to gain more understanding, a further detailed definition of mobility and scalability is presented in Section 2.4 and Section 2.5, respectively.

2.4 Definition of Mobility

A typical WSN network topology is static [Romer and Mattern 2004]. However, many applications such as disaster relief (Section 2.2.2) involve mobility. This section aims to provide a better understanding of mobility which plays a key focus in the context of the thesis. The definition of mobility can be described from the perspectives of moving patterns and moving speed, discussed in Section 2.4.1 and Section 2.4.2, respectively.

2.4.1 Moving Patterns

To the best of our knowledge, no mobility models have been proposed specifically to describe the moving pattern of mobile entities within WSNs. Most early research on mobility focused on understanding the behaviour of cellular networks [Camarda et al.

1996; Xie and Goodman 1993]. Mobility patterns have been used to derive traffic and mobility prediction models in the study of various problems in cellular systems, such as handoff, location management, paging, registration, call time, and traffic load [Bansal et al. 1999; Yu and Leung 2001]. Recently, mobility models have been explored also in ad-hoc networks. The following sections provide a survey of various mobility models; each generates distinct movement profiles and possesses their own unique characteristics.

2.4.1.1 Random Walk

The Random Walk Mobility Model was first described mathematically by Einstein in 1926 [Broch 1998]. It was developed to mimic the erratic movement of entities in nature, moving in an extremely unpredictable manner akin to the movement of molecules. In this mobility model, a mobile node moves from its current location to a new location by randomly choosing a direction and speed in which to travel. The new speed and direction are both chosen from pre-defined ranges [speedmin, speedmax] and [0, 2π], respectively. Each movement in the Random Walk Model occurs in either a constant time interval t or a constant distance travelled d, at the end of which a new direction and speed are calculated. If a mobile node which moves according to this model reaches a simulation boundary, it "bounces" off the simulation border with an angle determined by the incoming direction. The mobile node then continues along this new path. This effect is called the "border effect" [Bettstetter and Wagner 2002]. The Random Walk Model is widely used to study cellular networks [Rubin and Choi 1997] and ad-hoc networks [Garcia-Luna-Aceves and Madrga 1999; McDonald et al. 1999], in which it is sometimes referred to as

“Brownian Motion”.

Note that The Random Walk is a memory-less mobility pattern because it retains no knowledge on its past locations and speed i.e. the current speed and direction of a mobile node is independent of its past speed and direction. This characteristic can generate unrealistic movements such as sudden stops and sharp turns.

Figure 2.5 shows an example movement pattern of a single mobile node under the Random Walk model within an area of 500mx500m for 20 simulated minutes. The mobile node chooses a direction between 0 and 2π and a speed between 5m/s and 10m/s. The mobile node constantly travels for 330m before changing its direction and speed.

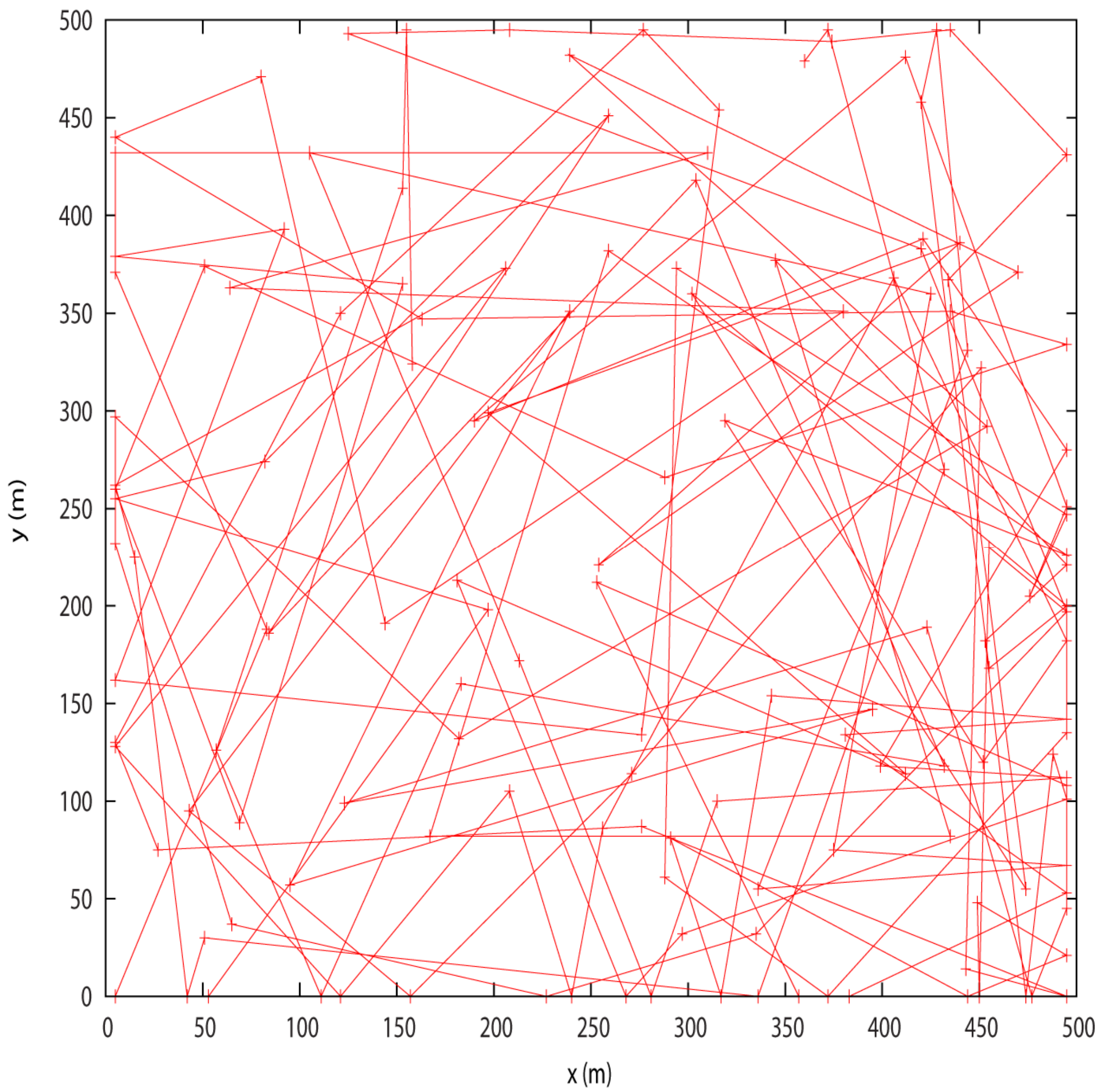

Figure 2.5 Travelling pattern of a mobile node using the Random Walk Mobility Model.

2.4.1.2 Random Waypoint

The Random Waypoint [Johnson and Maltz 1996] Mobility Model is an extension of the Random Walk, in which pause times between changes in direction and/or speed are included. A mobile node initially remains at a location for a certain time then it moves to a new randomly-chosen destination within the simulation area at a speed uniformly distributed between [speedmin, speedmax]. Upon arrival, the mobile node pauses for a specified time before starting the process again. The Random Waypoint Model is widely used in the development of well-known routing protocols for ad-hoc networks such as DSR [Johnson and Maltz 1996] and AODV [Perkins and Royer 1999].

Note that a complex relationship between node speed and pause time exists in the Random Waypoint Model [Boleng 2001]. Results show that a scenario with fast mobile nodes and long pause times produces a more stable network than a scenario with slower mobile nodes and shorter pause times. More specifically, long pause times (over 20 seconds) produce a stable network (few link changes per mobile node) even at relatively high speeds (up to 50m/s). If this model is used in a performance evaluation, appropriate parameters need to be evaluated to fit the requirement of the design task. Although the Random Waypoint Model incorporates pause times to describe a more realistic behaviour of user, it still performs unrealistic movements such as sudden stops and sharp turns due to the memory-less feature.

Figure 2.6 shows an example of movement under Random Waypoint for 60 simulated minutes. The speed of the mobile node is uniformly chosen between 5m/s and 10m/s; the pause time is set to 20s. Note that Random Waypoint does not exhibit the “border effect” as in Random Walk since node distribution is non-uniform with the node density maximum at the centre region and almost zero around the boundary of the area. Specifically, the probability of a mobile node choosing a new destination located in the centre of the simulation area, or a destination which requires travel through the middle of the simulation area, is high. This phenomenon is referred to as the 'non-uniform spatial distribution problem' [Mousavi et al. 2007].

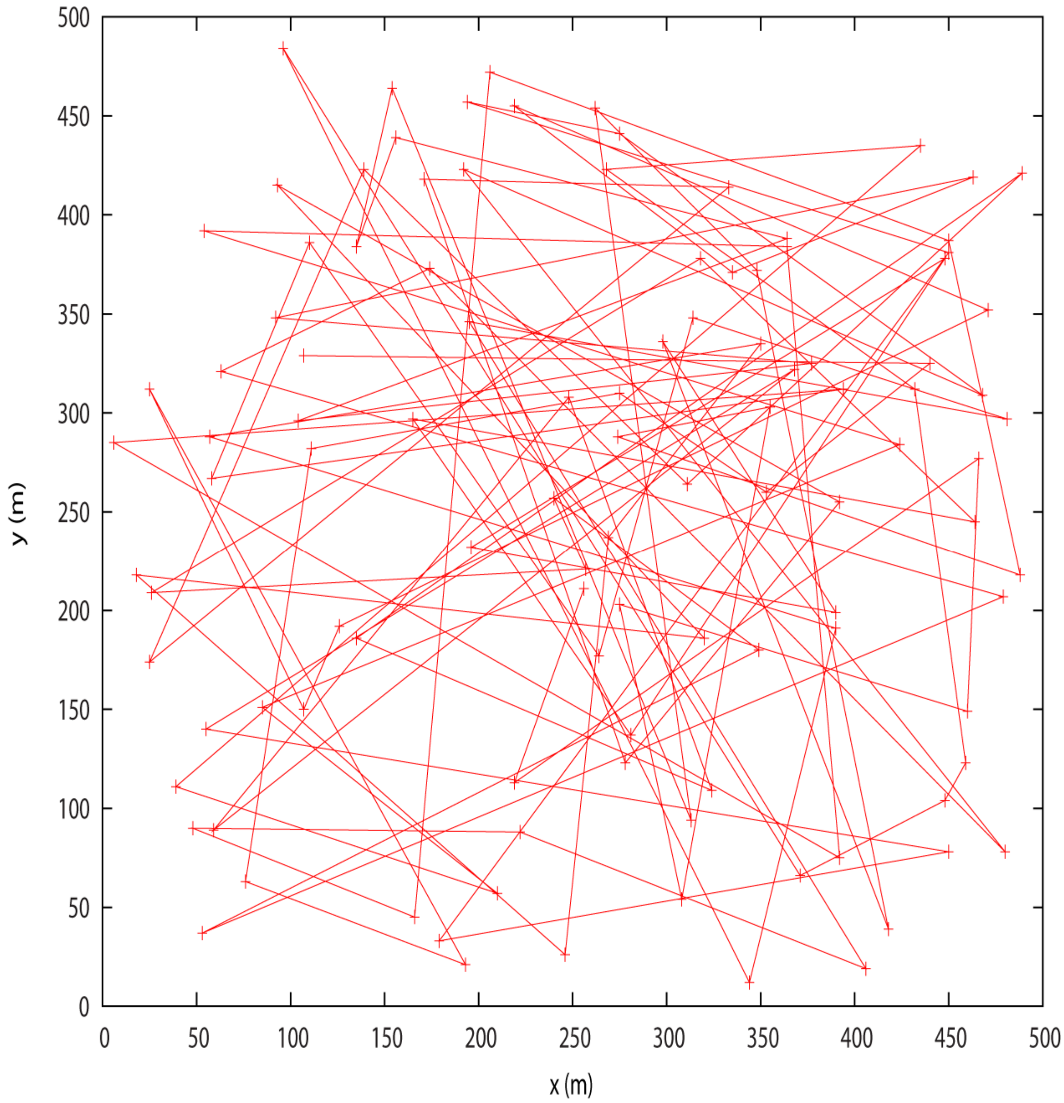

Figure 2.6 Travelling pattern of a mobile node using the Random Waypoint Mobility Model.

2.4.1.3 Random Direction

The Random Direction Mobility Model [Royer et al. 2001] was created to overcome the aforementioned non-uniform spatial distribution problem inherent in the Random

Waypoint Model. In this model, a mobile node chooses a random direction in which to travel similar to the Random Walk Model. A mobile node then travels to the border of the simulation area in that direction with a speed uniformly distributed between [speedmin, speedmax]. On reaching the simulation boundary, the mobile node pauses for a specified time, chooses another angular direction (between 0 and 180 degrees) and continues the process.

Figure 2.7 shows an example path of a mobile node under the Random Direction Model for 60 simulated minutes. The speed of the mobile node is uniformly chosen between 5m/s and 10m/s; the pause time is set to 20s. Since the mobile node always pauses at the border of the simulation area, from the perspective of multi-hop communication, the average hop count for data packets using Random Direction will be much higher than the average hop count of most other mobility models e.g. Random Waypoint. In addition, network partitioning will be more likely with Random Direction compared to other mobility models.

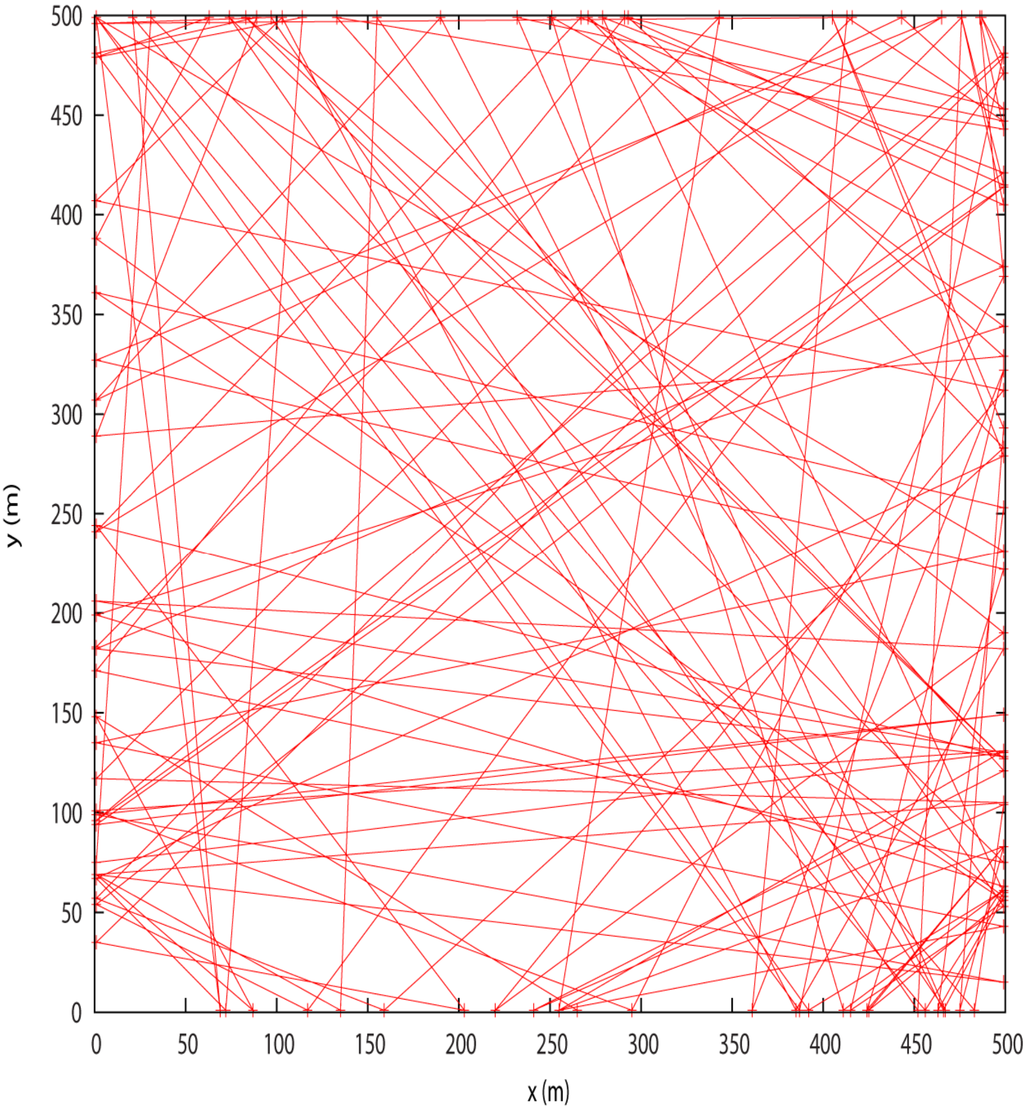

Figure 2.7 Travelling pattern of a mobile node using the Random Direction Mobility Model.

2.4.1.4 Gaussian-Markov

A mobile node e.g. pedestrian or visitor in a conference site or moving vehicle usually travels with a destination in mind. Furthermore, a change of the node's velocity within a short time is limited by physical restrictions. Therefore, a mobile node's future location and velocity are likely to be correlated with its past and current location and velocity. Thus, the memory-less nature of all the mobility models discussed above renders all unsuitable to represent such behaviour. In order to provide a more realistic mobility pattern, a 'memory' Gaussian-Markov Mobility Model is proposed in [Liang and Haas 1999], originally for the simulation of a PCS (Personal Communication Service); this model has also been used for the simulation of ad hoc network protocol performance [Tolety 1999].

The Gaussian-Markov model was designed to adapt to different levels of randomness via one tuning parameter, including as its two extreme cases, the Random Walk and the constant velocity fluid-flow model [Xie and Goodman 1993; Xie et al. 1993]. The latter is most suitable for vehicular traffic on highways; performing frequent "stop-and-go" tasks. In the Gaussian-Markov model, initially each mobile node is assigned a current speed and direction. At fixed intervals of time, n, movement occurs through updating the speed and direction of each mobile node. Specifically, the value of speed and direction at the *nth* instance is calculated based upon the value of speed and direction at the (n-1)*th* instance and a random variable using the following:

$$s_n = \alpha s_{n-1} + (1-\alpha)\bar{s} + \sqrt{(1-\alpha^2)}s_{x_{n-1}} \quad (2.1)$$

$$d_n = \alpha d_{n-1} + (1-\alpha)\bar{d} + \sqrt{(1-\alpha^2)}d_{x_{n-1}} \quad (2.2)$$

where s_n and d_n are the new speed and direction of the mobile node at time interval n; α, where $0 < \alpha < 1$, is the tuning parameter used to vary the randomness; $\bar{s}$ and $\bar{d}$ are constants representing the mean value of speed and direction as $n \to \infty$. $S_{x_{n-1}}$ and

$d_{x_{n-1}}$ are random variables from a Gaussian distribution with zero mean and unit variance and is independent of s_n and d_n, respectively.

As α approaches zero, Equation (2.1) and Equation (2.2) represent a drifting Random Walk mobility pattern with mean $\bar{s}$ and $\bar{d}$ and standard deviation $s_{x_{n-1}}$ and $d_{x_{n-1}}$, respectively. As α approaches one, Equation (2.1) and Equation (2.2) represent a constant velocity fluid flow mobility pattern with $s_n = s_0$ and $d_n = d_0$ for all n. The Gaussian-Markov model can represent a wide spectrum of mobility patterns with various degrees of memory, the Random Walk and fluid-flow models being two extreme cases.

At each time interval the next location is calculated based on the current location, speed, and direction of movement. Specifically, at time interval n, a mobile node's position is given by:

$$x_n = x_{n-1} + s_{n-1} \cos d_{n-1} \tag{2.3}$$

$$y_n = y_{n-1} + s_{n-1} \sin d_{n-1} \tag{2.4}$$

where (x_n, y_n) and (x_{n-1}, y_{n-1}) are the x and y coordinates of the node's position at the n^{th} and $(n-1)^{st}$ time intervals, respectively, s_{n-1} and d_{n-1} are the speed and direction of the node, respectively, at the $(n-1)^{st}$ time interval.

To ensure that a mobile node does not remain near an edge of the grid for a long period of time, they are forced away from an edge when they enter within a certain distance of the edge, executed by modifying the mean direction variable $\overline{d}$ in Equation (2.3) and Equation (2.4). For example, initially the mobile node is assigned a current direction, say 0 degrees (trends to the right), when the node nears the right edge of the simulation grid, the value $\overline{d}$ is changes to 180 degrees; the node's new direction is away from the right edge of the simulation grid.

As shown in Figure 2.8, the Gaussian-Markov Mobility Model eliminates sudden stops and sharp turns encountered in memory-less mobility models by calculating the new direction and speed based on the previous direction and speed.

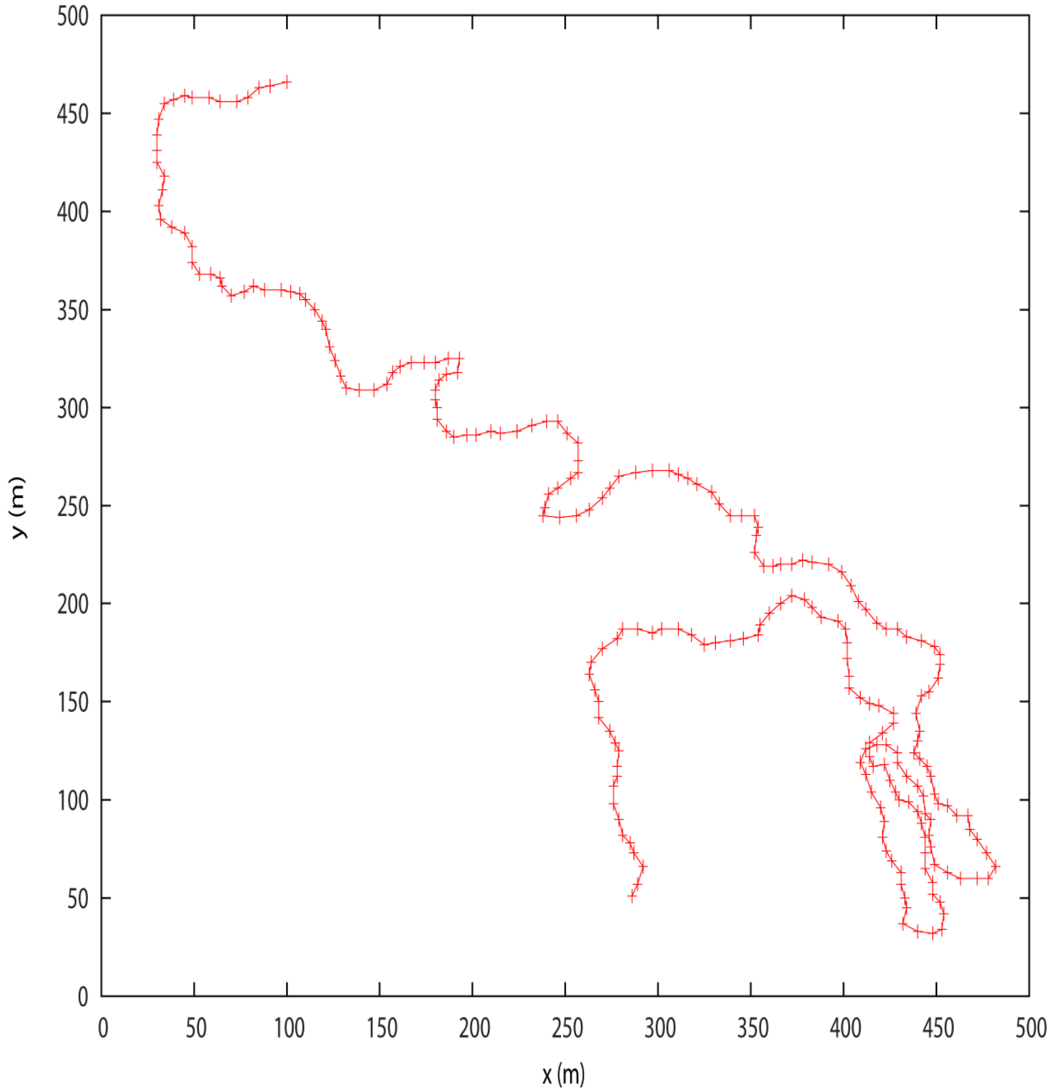

Figure 2.8 Travelling pattern of a mobile node using the Gaussian-Markov Mobility Model.

2.4.2 Moving Speed

Previous sections have discussed the most well-known mobility models which mainly describe the moving pattern of mobile entities (the manner in which they travel around the field) in a wireless communication system. This section describes the grades of mobility based on speed of moving nodes. Three different grades exist [Molisch 2005]:

- *Low Speed:* many communications devices are operated at pedestrian pace (speeds range from 4.51km/h to 4.75km/h) [Roads 1997]. Mobile phones, as well as PDAs operated by walking users are typical examples. For WSNs, this speed is appropriate when sensor nodes are attached to walking humans.
- *High Speed:* high speed mobility usually describes speed ranges from 30km/h to 150km/h, a speed relevant to sensor nodes attached to moving cars.
- *Extremely High Speed:* extremely high speed mobility represents high-speed trains and planes, which cover speeds between 300km/h and 1000km/h.

Note that an upper bound of speed must be defined i.e. the maximum speed of mobile nodes a WSN can support. Such a fundamental limitation is as a consequence of the Doppler shift [Rappaport 2002] originating at the physical layer. Due to the nature of radio propagation, it is well known that the relative motion of transmitter and receiver produces an apparent change in frequency. The magnitude of the Doppler shift is given by:

$$f = \frac{v}{\lambda}\cos\alpha \tag{2.5}$$

where v is the relative speed of transmitter and receiver, λ is the wavelength, α is the angle between the direction of movement and the direction of wave propagation. The maximum Doppler shift f_m occurs when $\alpha = 0$ or $\alpha = 180$ (transmitter and receiver move in the same direction or opposite directions). In this case;

$$f_m = \pm \frac{v}{\lambda} \tag{2.6}$$

Exceeding the Doppler shift maximum f_m renders the signal irresolvable e.g. for a wireless device operating at a frequency of 2.4GHz with a frequency tolerance of 800Hz with the receiver static and the transmitter travelling directly towards the receiver, the maximum speed limitation of transmitter is $v = f_m * \lambda = 800 * 0.125$ m = 360km/h.

2.5 Definition of Scalability

Scalability is defined as the ability to maintain the network performance characteristics such as packet received ratio, routing efficiency and latency irrespective of the 'size of the network'. Most of existing research in routing protocols such as Direct Diffusion [Intanagonwiwat et al. 2003] and AODV [Perkins and Royer 1999] treat the size of the network as the number of sensor nodes. In this thesis, the size of the network is not only considered as defined by the number of sensor nodes but also considers the physical size in the study of scalability. With WSNs potentially consisting of hundreds or up to thousands of sensor nodes, as well as delivering applications over a wide range e.g. from smart building to wildfire detection, scalability is a core requirement.

The physical size of a network presents a new challenge in terms of developing appropriate WSN routing protocols. Intuitively, the success of routing relies on intermediates between sender and receiver. As the physical size of network increases, the sender and receiver are more likely to be outwith the single hop connection regime. This implies that more intermediates will be involved in the transport of packets. A more detailed study on scalability in relation to the physical size of the network is presented in Section 3.3.2.3.

2.6 Conclusions

In this chapter, the basic concepts of Wireless Sensor Network were discussed. The discussions begun from the node level up to the system level and introduced potential applications supported by the platform. Networking scenarios are highlighted and a system model is given which will form the foundation for the analyses within the main body of the thesis.

The chapter then focuses on the definition of mobility and scalability which represent the major focus in the scope of this research. One of the main challenges that the development of effective WSN routing protocols must address are the mobility of sensor nodes and mobile sink(s) (the mobility of events was not considered in this work) and network scalability in terms of numbers of sensor nodes and physical size of network.

Several mobility models had been discussed designed to mimic the movement of mobile entities in a tractable way. ;their pros and cons are summarised in Table 2.1. The Random Waypoint Mobility Model is the most widely used when evaluating wireless networks [Johnson et al. 2001; Luo et al. 2003; Perkins and Royer 1999] and hence is adopted in the subsequent analyses in this thesis.

Table 2.1 The pros and cons of mobility models.

Mobility Model	**Pros**	**Cons**
Random Walk	● The simplest model to implement ● Generates unpredictable movements, enabling a long-running simulation to consider all locations and node interactions	● Unrealistic movement patterns ● Sharp and sudden turns ● Behaviour not observed in real applications
Random Waypoint	● The most common use mobility model, because of its simplicity ● A foundation for building other mobility models	● Exhibits speed decay ● Exhibits density wave ● Memory-less movement behaviours
Random Direction	● A variation of the random waypoint without drawback of density wave ● Uniform distribution of chosen routes	● Unrealistic movement patterns ● Average distances between mobile nodes are much higher than other models, leading to decrease the routing protocols performance
Gaussian-Markov	● The velocity of a mobile node at any time slot is a function of its previous velocity ● More realistic compared to the Random Waypoint	● Rarely used in mobile networks simulation studies, mainly due to the relatively larger complexity involved in simulating

The impact of mobility (Random Waypoint) and scalability on routing performance is further studied in depth in chapter 3.

Chapter 3

Existing Routing Protocols

3.1 Introduction

This chapter reviews of a number of existing routing protocols and examines them through a comprehensive network performance analysis. The goal is to indentify a candidate approach that gracefully handles node mobility and achieves network scalability. Section 3.2 presents a taxonomy of routing protocols in wireless networking, categorised into explicit and implicit connection approaches followed by a detailed comparison of their design principles. Section 3.3 proposes a series of metrics used to analyse, capture and quantify the impact of node mobility and network scalability on network performance. Comprehensive simulation based evaluations are carried out to show the potential and benefit of an implicit connection strategy. Section 3.4 concludes with a summary that highlights the design principles of explicit and implicit connections and draws conclusions obtained from the analysis.

3.2 Routing in Wireless Networks

A large number of routing protocols have been proposed to facilitate wireless communication driven by a wide range of requirements [Murthy and Manoj 2004]. In general, routing protocols need to perform the following:

- Determine and detect changes in network topology.
- Maintain network topology and connectivity.
- Schedule packet transmission and channel assignment.
- Route the packet to its destination.

3.2.1 Characteristics of an Ideal WSN Routing Protocol

The design of appropriate routing protocols for WSNs is one of the major challenges facing the discipline. Most existing protocols mainly address routing in fixed networks [Förster and Murphy 2010]. However many evolving WSN implementations must support mobile nodes whilst meeting a stringent scalability requirement; this section identifies the salient properties of such a routing protocol.

- *Distributed Implementation:* WSNs are autonomous, self-organising systems, which do not rely on centralised authorities. Therefore, protocols must be based on distributed routing algorithms.
- *Adaptability to Changing Topology:* routing protocols should adapt to changes in network topology. In most scenarios both sensor nodes and sink(s) are assumed to be static and the network topology is only subject to change as a consequence of sensor node failure and/or the redeployment of new participating nodes. In mobile scenarios where both sensor nodes and sink(s) are mobile, the network topology changes continually, the degree and variety

dependent on the mobility characteristics such as moving speed and pause time.

- *Adaptability to Network Scale:* the number of participating sensor nodes may be of the order of hundreds or thousands. Any routing scheme needs to handle this number of nodes and fully utilise the benefits (such as multipath connectivity) from the resultant high node density. Furthermore, network scalability in terms of physical network size may be in the region of several hundred meters or indeed kilometres in extent. For example, in bridge structure monitoring applications, the routing protocol must provide reliable connections over a large number of hops [Pakzad et al. 2008].
- *Efficient Bandwidth Utilisation:* since the bandwidth resource of wireless networks is limited, especially for a low resource platform such a WSN, any reduction in control overhead is much sought after. If as a consequence of the implementation of a routing protocol, excessive control traffic is generated, the bandwidth available for the data traffic will reduced significantly. Moreover, the larger the network, the more bandwidth is consumed to facilitate data transport resulting in a serious impact on network performance. Any routing protocol should minimise the amount of control traffic in so doing maximising the ability of the platform to transport data [Sridharan and Krishnamachari 2007].
- *Energy Conservation:* energy is always a primary concern in WSN deployments since sensor nodes are typically battery powered, and it is often impractical to change or recharge batteries. The main sources that dissipate energy have been identified in [Ye et al. 2002]. Results shows that by placing

the radio into a sleep mode, sensor nodes conserve significant energy therefore prolonging the useful lifetime of the network. However, a trade-off between energy efficiency and network performance exists. Thus it is desirable that the routing protocol is able to accommodate a sleep period without causing any adverse consequences e.g. decreased network connectivity and increasing latency.

3.2.2 Taxonomy of Routing Protocols

Taking into consideration the main design criteria has led to the development of a large number of routing protocols for WSN applications [Biradar et al. 2009; Singh et al. 2010]. Some mainly consider energy conservation and network scalability in isolation [Zhou and Shi 2006]; others have been proposed for Mobile Ad-hoc Network (MANET), although considering issues of mobility and QoS, nevertheless may also be applicable to WSN [Singh et al. 2010].

Consequently there are many ways in which protocols can be categorised. In this section, routing protocols for wireless networks are broadly classified into two main categories: explicit routing and implicit routing. Some representative examples of each class are shown in Figure 3.1. This taxonomy divides the protocols reflecting fundamental design and implementation strategies.

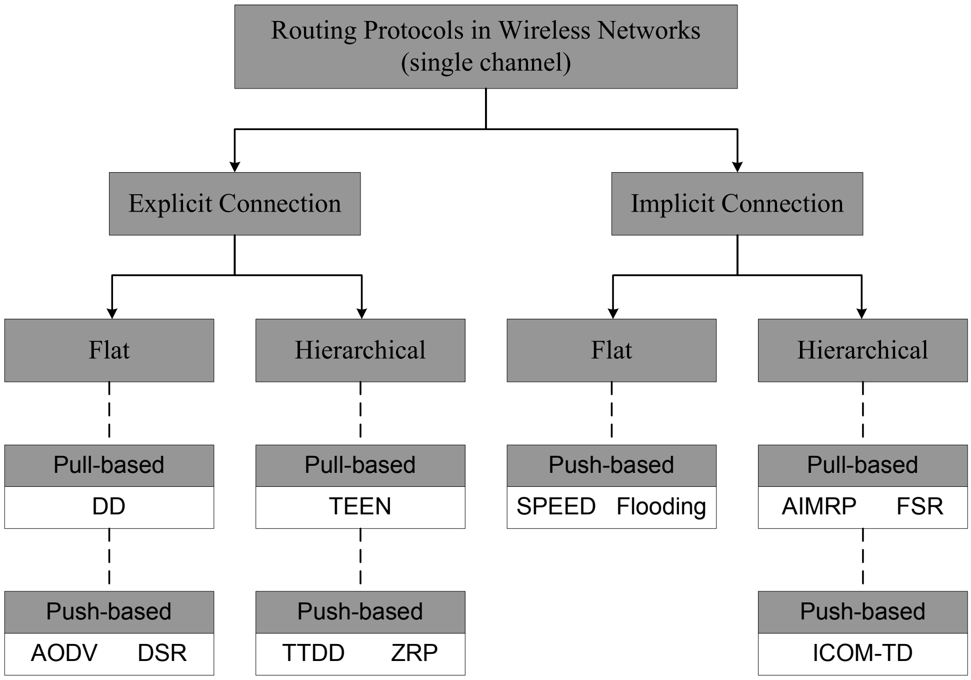

Figure 3.1 Taxonomy of routing protocols in wireless networking.

- *Communication Model:* routing protocols can be categorised according to designs for multi-channel or single channel systems [Feeney 1999]. Multi-channel protocols are generally based on Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA) or Frequency Division Multiple Access (FDMA); examples can be found in [Bahl et al. 2004; Chiang et al. 1997; Kyasanur and Vaidya 2006; Yan and Gharavi 2006]. Although multi-channel protocols have the potential to use available channel bandwidth more efficiently, this approach is difficult to use in WSNs, since the requirement for precise synchronisation for executing channel hopping introduces a considerable overhead [Wang et al. 2009]. Single-channel protocols based on Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) schemes [IEEE Standard 2003] that resolve contention over a shared media are more appropriate in ad hoc environments. A large class of protocols has been designed based on the usage of a single-channel [Singh et al. 2010].

- *Traffic Connection:* routing protocols can be categorised according to the certainty of traffic connection [Liu et al. 2004]. In explicit routing, a specific path is established and maintained for data traffic transport. The main drawback of explicit routing is that once the path is established, there is lack of flexibility to accommodate any change of network topology. Conversely, with implicit routing, data traffic is transported from upstream to downstream without taking a pre-defined path. Hence, implicit routing provides a certain degree of flexibility to changes in network topology. The strength of implicit routing is further studied in Section 3.3.2.
- *Topology:* routing protocols can be categorised according to network topology. A flat topology is defined as all nodes operational, executing the same tasks [Karl and Willig 2005]. A hierarchical topology is defined when a hierarchy is established locally characterised by some nodes executing a 'special' role e.g. controlling neighbouring nodes. Hierarchical protocols attempt to improve scalability and data transport efficiency by enabling collaboration amongst sensor nodes [Karl and Willig 2005].
- *Routing Scheme:* information on available routes can be published or subscribed to, provided by either the source or sink. Thus protocols can be classified as push-based and pull-based [Karl and Willig 2005]; the selection of the best option is application dependent. For example, in a battlefield surveillance application, a solider needs to determine the location of enemy

tanks. Nodes detecting tank movement and location can periodically push (broadcast) the information throughput the network. This push-based strategy is efficient when there are many 'soldiers' in the network constantly in need of specific information. However, there is a trade-off as the bandwidth consumed by the broadcast feature is not efficient when the demand for information falls below a certain threshold i.e. is relative low. Alternatively, a pull-based strategy can be considered, in which the soldier only broadcasts a request for information when needed.

Table 3.1 Salient feature of routing protocols in wireless networking.

Protocol	**Salient Feature**
Direct Diffusion (DD) [Intanagonwiwat et al. 2003]	DD uses the naming scheme for the data to reduce unnecessary transmission. It also uses a reinforcement technique to enhance the transmission rate for the shortest path.
Ad-hoc On-demand Distance Vector (AODV) [Perkins and Royer 1999]	AODV minimises the number of required broadcasts by creating routes on a demand basis.
Dynamic Source Routing (DSR) [Johnson et al. 2001]	DSR is based in the concept of source routing and mobile nodes are required to maintain route caches containing the source route of which the mobile node is aware.
Threshold sensitive Energy Efficient sensor Network (TEEN) [Manjeshwar and Agrawal 2001]	TEEN uses the concept of hard and soft threshold to support time critical applications.
Two-Tier Data Dissemination (TTDD) [Luo et al. 2003]	TTDD provides data delivery to multiple mobile sinks based on a grid structure data dissemination policy.
Zone Routing Protocol (ZRP) [Haas et al. 2002]	ZRP defines a routing zone and utilises degrees of clustering.
SPEED [He et al. 2003b]	SPEED uses the information from each neighbour and geographic forwarding to ensure soft real-time end-to-end communication.
Flooding [Tanenbaum 1981]	Flooding is the simplest algorithm for multi-hop routing in mobile wireless networking.

Address-light, Integrated MAC and Routing Protocol (AIMRP) [Kulkarni et al. 2006]	AIMRP organises the network into concentric tiers around the sink(s) and routes data by forwarding them from one tier to another via implicit connections, in the direction of the sink.
Fisheye State Routing (FSR) [Pei et al. 2000]	FSR is based on hazy link state (approximates a global topology) algorithm used in wired networks; the accuracy of the link state is increased as the packet approaches the destination to ensure the success of packet delivery.
Implicit Cluster-based Overlay Multicast protocol exploiting Tree Division (ICOM-TD) [Choi et al. 2007]	ICOM-TD achieves a high packet delivery ratio by utilising implicit clustering and tree division under varied network scalability.

Table 3.1 summarises the salient features of routing protocols in wireless networking (n.b. Figure 3.1 summarises the range of design choices). Table 3.2 shows how different routing protocols are categories and also compares different routing techniques according to a number of metrics. A discussion in detail of all cited protocols is not within the scope of this thesis and thus only a selection of representative routing protocols are described in the following sections.

Table 3.2 Comparison of routing protocols in wireless network.

	Proposed field	**Mobility**	**Scalability**	**QoS**	**Optimal route**	**Energy awareness**	**Position awareness**	**Traffic pattern**
DD	WSN	Ltd.	Ltd.	No	Yes	No	No	Base station pattern
AODV	MANET	Yes	Low	Poss.	Yes	No	No	Peer-to-peer pattern
DSR	MANET	Yes	Low	No	Yes	No	No	Peer-to-peer pattern
TEEN	WSN	Fixed base station	Good	Yes	No	Yes	No	Base station pattern
TTDD	WSN	Mobile base station	Low	No	No	No	Yes	Base station pattern

ZRP	MANET	Yes	Good	No	No	No	No	Peer-to-peer pattern
SPEED	WSN	No	Ltd.	Yes	Yes	Yes	Yes	Base station pattern
Flooding	Computer network	Yes	Bad	Yes	Yes	No	No	Peer-to-peer pattern
AIMRP	WSN	Poss.	Ltd.	Ltd.	No	Yes	No	Base station pattern
FSR	MANET	Yes	Ltd.	No	No	No	No	Peer-to-peer pattern
ICOM-TD	MANET	Yes	Good	Ltd.	No	No	No	Peer-to-peer pattern

3.2.3 AODV

Ad-hoc On-demand Distance Vector (AODV) routing is a well-known algorithm for Mobile Ad-hoc Networks (MANET), supporting peer-to-peer communications [Perkins and Royer 1999]. AODV adopts both a modified on-demand broadcast route discovery approach used in DSR [Johnson et al. 2001] and the concept of destination sequence number adopted from Destination-Sequenced Distance-Vector routing (DSDV) [Perkins and Bhagwat 1994].

When a source node desires to send a message to a destination node and does not already have a valid route to that destination, it initiates a ‘Path Discovery’ and sends a Route Request (RREQ) message to its neighbours, which then forward the request message to their neighbours, and so on, until either the destination or an intermediate node with a ‘fresh enough’ route to the destination is located. Figure 3.2 illustrates the propagation of the broadcast RREQs across the network. Nodes receiving the RREQs update their information about the source. They also set up a backward link to the source in their routing tables. Each RREQ contains the source node’s address and a Broadcast ID that uniquely identifies it. It also has a current sequence number that determines the freshness of the message. Thus, a message number with a higher sequence number is considered to be fresher or more recent than that with a lower sequence number. The RREQ also contains a hop count variable that keeps track of the number of hops from the source.

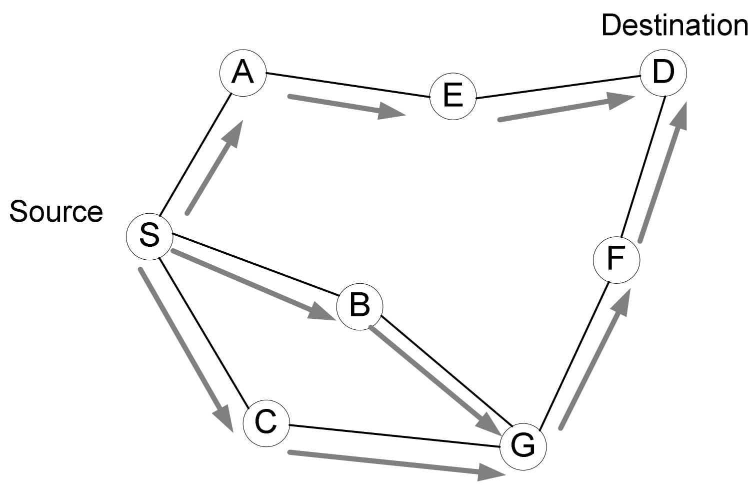

Figure 3.2 Propagation of RREQs.

Once the RREQ reaches the destination or an intermediate with a fresh route, the destination/intermediate node responds by unicasting a Route Reply (RREP) packet back to the neighbour from which it first received the RREQ (Figure 3.2). As the RREP propagates back to the source node, the intermediate nodes setup forward pointers to the actual destination (Figure 3.3).

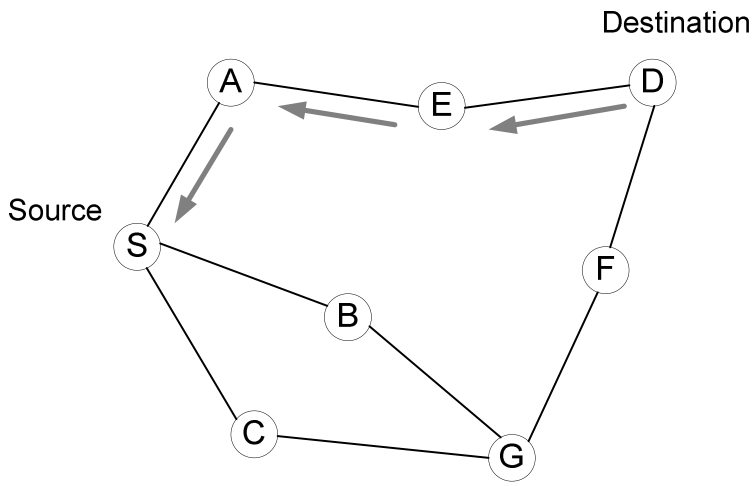

Figure 3.3 Propagation of the RREP.

When the source node receives the RREP, it checks whether it has an entry for the route. If it does not have any entry in its routing table, the node creates a new entry in the table. Otherwise it checks the sequence number of the RREP; if the RREP arrives with the same sequence number as in its tables but with a smaller hop count, or a greater sequence number (indicating fresher route), it updates its routing table and starts using this better route. Once an entry for the new route has been created, the node can start communication with the destination.

In the mobile scenario, routes are maintained as follows. If a source node moves, it is able to reinitiate the route discovery protocol to find a new route to the destination. If a node along the route or the destination moves, its upstream neighbour notices the move (using the HELLO beaconing or link layer detection) and propagates a link failure notification message (RRER) to each of its active upstream neighbours to inform them of the erasure of that part of the route. These nodes in turn propagate the

RRER to their upstream neighbours, and so on until the source node is reached. The source node may then choose to reinitiate route discovery for that destination if a route is still desired.

AODV is used in many comparative evaluation studies relating to mobility, and is used as the benchmark for other newer protocol designs (including the proposed protocols in this work). Its performance has been well-studied and investigated [Broch et al. 1998; Johansson et al. 1999; Das et al. 2000]. Particularly, Broch et al. [1998] concluded that the flat structure of explicit protocols such as AODV and DSR perform well in some cases yet have certain drawbacks in others. Specifically, DSR and AODV performed equally well, delivering over 95% of the packets regardless of mobility rate. The use of source routing in DSR increases the routing overhead information required by the protocol. AODV although eliminates the source routing overhead, still suffers from a high overhead because of its Distance Vector algorithm. For this reason, AODV has proved more expensive than DSR at high rates of mobility. The performance analysis in terms of routing overhead is further investigated in Section 5.2.2.

3.2.4 ZRP

The Zone Routing Protocol (ZRP), as the name implies, is based on the concept of zones. A routing zone is defined for each node separately, and the zones of neighbouring nodes overlap. The routing zone has a radius r expressed in hops. The zone thus includes nodes whose distance from the node in question is at most r hops. An example routing zone is shown in Figure 3.4, where the routing zone of S includes the nodes A–I, but not K. In the example, the radius is marked as a circle around the node in question. It should be noted that the zone is defined in hops, not as a physical distance.

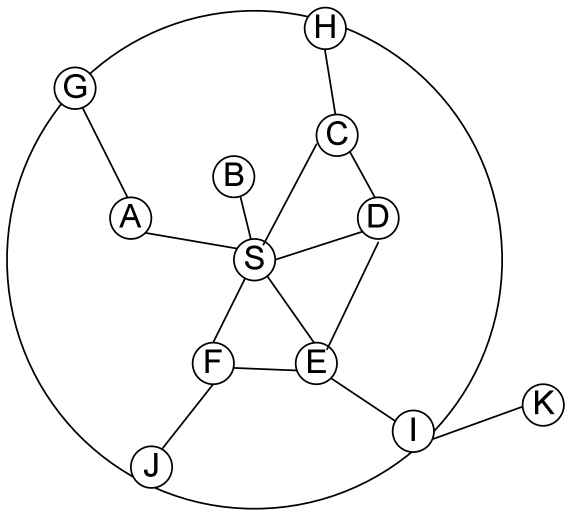

Figure 3.4 Example routing zone with *r* =2.

The nodes of a zone are divided into peripheral nodes and interior nodes. Peripheral nodes are nodes whose minimum distance to the central node is exactly equal to the zone radius *r*. The nodes whose minimum distance is less than *r* are interior nodes. In Figure 3.4, nodes A–F are interior nodes; nodes G–J are peripheral nodes and node K is outside the routing zone. Note that node H can be reached by two paths, one of length 2 and one of length 3 hops. The node is within the zone, since the shortest path is less than or equal to the zone radius.

In order to detect new neighbouring nodes and link failures, ZRP relies on a Neighbour Discovery Protocol (NDP) provisioned by the Media Access Control

(MAC) layer. NDP transmits “HELLO” beacons at regular intervals. Upon receiving a beacon, the neighbour’s table is updated. Neighbours, for which no beacon has been received within a specified time, are removed from the table.

By dividing the network into overlapping, variable-size zones, ZRP consists of several components, which only together provide the full routing benefit. Each component works independently of the other and they may use different technologies in order to maximize efficiency in their particular role. Components of ZRP are the IntrAzone Routing Protocol (IARP), the IntErzone Routing Protocol (IERP) and the Bordercast Resolution Protocol (BRP), all briefly described below.

The IntrAzone Routing Protocol (IARP) is used by a node to communicate with the interior nodes of its zone limited by the zone’s radius (the number of hops from the node to its peripheral nodes). As the network topology changes, then the local neighbourhood of a node will also change. Thus nodes need to update the routing information continuously in order to determine the peripheral nodes as well as to maintain a map of which nodes can be reached locally. IARP allows for local route optimisation through the removal of redundant routes and the shortening of routes if a route with fewer hops has been detected, as well as bypassing link-failures through multiple (local) hops.

The global reactive routing component, the IERP, takes advantage of knowledge of the local topology of a node's zone and, using a reactive approach enables communication with nodes in other zones. When there is request for a route, route queries are issued. The delay caused by route discovery (in contrast to IARP, where the route is immediately available) is minimized through ‘bordercasting’, an approach in which the node does not submit the query to all local nodes, but only to it's peripheral nodes. A node does not send a query back to the nodes the request came from, even if they are peripheral nodes. It is necessary to disable pro-active updates for local routes to convert an existing reactive routing protocol for use as the IERP. The IERP takes advantage of the local routing information provided by the IARP, as well as changing the way route discovery is handled. Instead of flooding a route request to all nodes, it uses the Bordercast Resolution Protocol (BRP) to initiate route requests with peripheral nodes only.

The BRP is used to direct route requests initiated by the IERP to peripheral nodes, removing redundant queries and so maximizing efficiency. Bordercast trees are constructed by utilising the map provided by the IARP. The BRP keeps track of which nodes a query has been delivered to, so that it can prune the bordercast tree of nodes that have already received the query. When a node receives a query packet for a node that does not lie within its local routing zone, it constructs a bordercast tree so that it can forward the packet to its neighbours. These nodes, upon receiving the packet, reconstruct the bordercast tree so that they can determine whether or not it belongs to the tree of the sending node. If it does not belong to the bordercast tree of the sending node, it continues to process the request and determines if the destination lies within its routing zone and taking the appropriate action, upon which the nodes within this zone are marked as covered.

The IARP, IERP and BRP together fulfil network-wide packet routing which is illustrated in the following example. In Figure 3.5, assume that node S has a packet to send to node X. The zone radius is $r = 2$. The node uses the routing table provided by IARP to check whether the destination is within its zone. Since it is not, a route request is issued using IERP. The request is bordercast to the peripheral nodes (gray in the picture). Each of these searches their routing table for the destination.

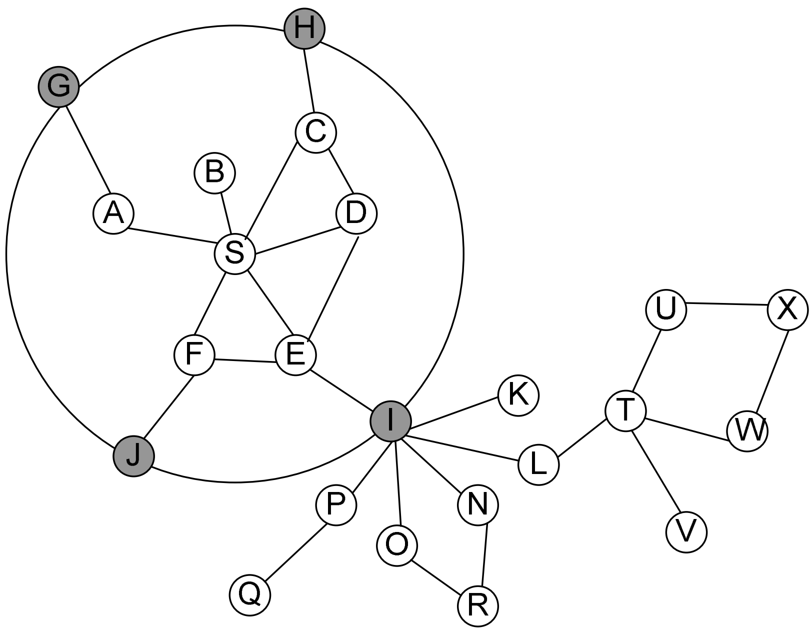

Figure 3.5 The routing zone of node S.

Node I does not find the destination in its routing table provided by IARP. Consequently, it broadcasts the request to its peripheral nodes (shown in gray in Figure 3.6). Due to query control mechanisms, the request is not passed back to nodes D, F and S.

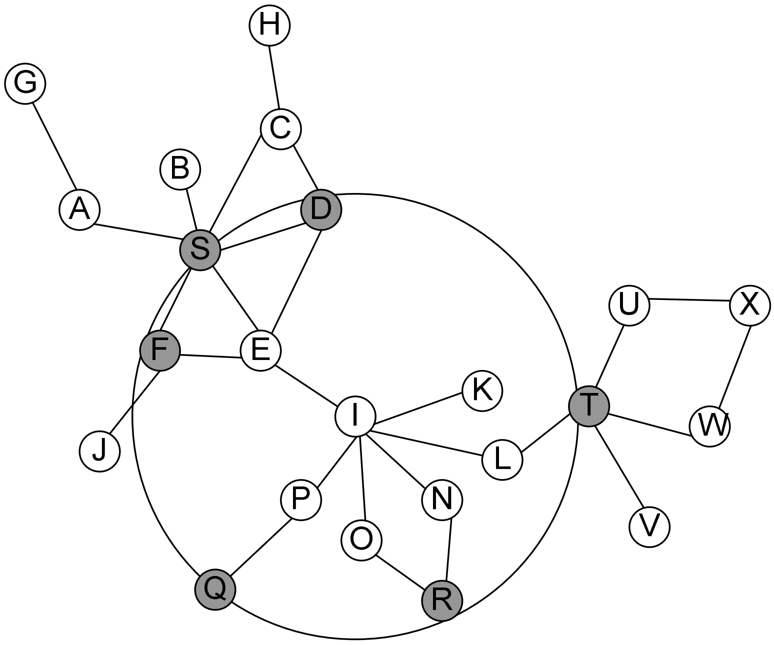

Figure 3.6 The routing zone of node I.

Finally, the route request is received by node T, which can locate the destination in its routing zone (within a 2 hop distance). Node T appends the path from itself to node X to the path in the route request. A route reply, containing the reversed path is generated and sent back to the source node. If multiple paths to the destination are available, the source would receive several replies; then the packet can be forwarded through an appropriate reverse path and arrive at node X.

The performance of ZRP has also been well-studied and compared with AODV [Mittal and Kaur 1999; Raju et al.2010]. Results show that ZRP performs poorer than AODV for different degrees of mobility [Mittal and Kaur 1999] as well as number of mobile nodes [Raju et al.2010]. Due to the use of a hierarchical structure, ZRP should provide better scalability compared with the flat structured AODV. This may be as a consequence of the high cost of maintaining the hierarchical structure, congesting the network with control packets. Raju et al. [2010] conclude that a better algorithm is needed to improve the performance of ZRP.

3.2.5 Flooding

Flooding is an established technique that can also be used for routing in wireless networking; Tanenbaum [1981] describes the Flooding algorithm as “one of the simplest routing algorithms”, the foundation of which is broadcasting an incoming packet to all neighbours. As long as source and destination nodes are in the same connected path, the packet is certain to arrive at the destination. To avoid packets circulating endlessly, a node should only forward a packet it has not yet received (piggybacking a unique source identifier and sequence numbers in the packet). Also, packets usually carry some form of expiration data (time to live, maximum number of hops) to avoid needless propagation of the packet e.g. if the destination is not reachable at all.

Flooding is an implicit connection approach which does not rely on costly topology maintenance and complex route discovery algorithms. Moreover, Flooding is most robust against mobility in wireless networking compared with other explicit connection protocols. This is due to the fact that the success of packet delivery relies on the connectivity of the “entire” network, not on the connectivity of a specific path as in AODV or DSR. A more detailed explanation on the previous assertion is provided in Section 3.3. Flooding also has the advantage that any packet lost due to environmental interference e.g. obstacle or transmission collision is quickly replaced by one of its copies. However, Flooding has several drawbacks such as [Heinzelman et al. 1999]:

- *Implosion:* a situation where duplicated messages are sent to the same node. For example, if sensor node A has N neighbour sensor nodes that are also the neighbours of sensor node B, sensor node B receives N copies of the message sent by sensor node A.
- *Overlap:* if two nodes share the same region, both of them may sense the same stimuli at the same time. As a result, neighbouring nodes receive duplicated messages.
- *Resource Blindness:* the Flooding protocol does not take into account the available energy resources. An energy resource aware protocol considers the amount of energy available to individual nodes at all times.

In order to overcome the aforementioned deficiencies, many extensions of Flooding have been proposed [Ababneh and Selvakennedy 2007; Campista et al. 2008; Jaquet et al. 2001]. For example each node forwards an incoming packet to a subset of its neighbours determined by a topology-control algorithm, equivalent to flooding on a reduced topology. Those approaches are referred to as 'controlled flooding'. Note that the first of the proposed routing protocol defined in this work – the Virtual Grid Paging (VGP) - can also be categorised as one of them.

3.2.6 AIMRP

AIMRP is a cross-layer design which integrates the MAC and routing layers to minimise the protocol overhead. It is designed particularly for sensor-to-sink WSNs applications with latency constraints. Another attractive feature is that it provides a 'power-saving mode' which requires no co-ordination between nodes; nodes are allowed to shut their radio modules down without a sleep associated latency penalty. The routing structure is flat, with all nodes treated equally by the algorithm, utilising the features of implicit connections.

AIMRP firstly organises (just after the deployment) the entire network into annular tiers concentrated at the sink and of thickness αR where R is communication range, α is a protocol parameter which effects network connectivity; its minimum value has been calculated as 0.45 [Kulkarni et al. 2006]. Tier numbering is the key to facilitate implicit connections for sensor-to-sink applications. The network is illustrated in Figure 3.7, the sink being at the centre of the region of interest. If node S in tier n intends to send data to the sink, this data is forwarded towards the sink by one of the so-called 'next-hop' nodes, shown in the hatched region. A node in tier n forwards its data to any node located in tier n − 1 (or n − 2 and so on). In turn, the node which receives this data forwards it to the next tier n − 2. In this way, at each hop, data from a node in one tier is forwarded to some other node in another tier closer to the sink. Thus if at each hop there exists at least one node with a lower tier number within the range of the sender node, then a path from a source node to the sink is guaranteed.
Note that there is no routing information exchange amongst sensor nodes. Instead of creating a routing path through a series of specific intermediated nodes like AODV and TTDD, AIMRP forwards the data from one tier to another using an implicit connection; the actual router among the next-hop nodes is not predetermined and is selected 'on the air'. The selection of router is done by using two-way handshaking;

refer to [Kulkarni et al. 2006] for detail.

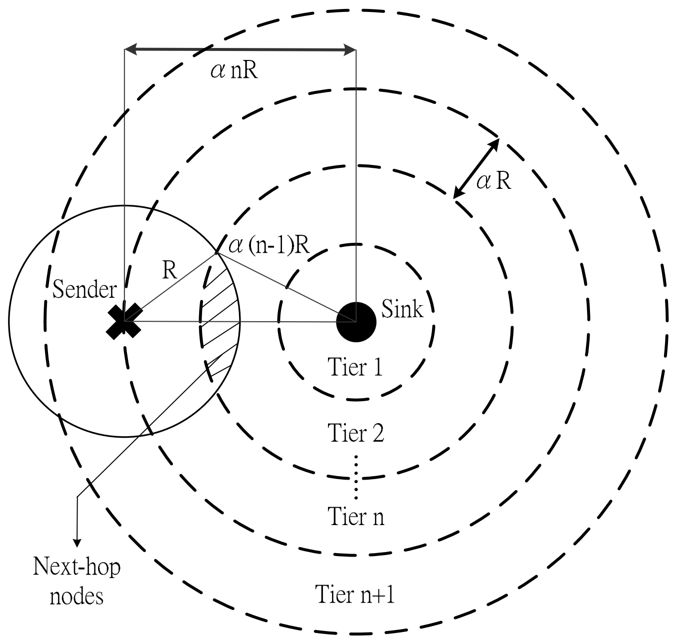

Figure 3.7 Formation of a tier structure.

The power-saving mode in AIMRP is completely asynchronous and utilises random duty-cycling, distinct from other power saving scheme including S-MAC [Ye et al. 2002], PAMAS [Singh and Raghavendra 1998] and IEEE 802.11 power-saving mode [IEEE Standard 1999]. Nodes are allowed to shut off their radio modules for a random duration t_σ and conserve energy without affecting data forwarding; the protocol inherently assumes that there are multiple next-hop nodes available for data forwarding from a higher tier to a lower tier region. It is worth noting that the sleep time duration t_σ can be set to random.

Since nodes use multi-hop relaying to send their data to the sink, they cannot put their radio modules off indefinitely. Moreover, the parameter t_σ is tuneable and determined by the end-to-end latency required by the application. All nodes should wake-up sufficiently often to ensure that for a given node trying to send a packet to the sink, there will always be nodes available to relay the data within the specified latency period. The dimensioning of t_σ has discussed in detail in [Kulkarni et al. 2006].

AIMRP has been proposed for use in WSNs where both sensor nodes and sinks are static [Kulkarni et al. 2006]. The network configuration is established just after deployment and performed only once. AIMRP is able to support mobility by reconfiguring the entire network periodically; however the concomitant overhead is significant which decreases network performance. Moreover, AIMRP is only designed to support the traffic flow for sensor-to-sink (many-to-one). In the case of multiple mobile sinks and many-to-many traffic flows, the total overhead limits scalability in terms of number of mobile sinks that can be deployed i.e. the total overhead is multiplied by the total number of mobile sinks.

3.2.7 Summary of Suitable Applications

Different routing protocols are suitable for different kinds of network configurations and node characteristics (static or mobile). Figure 3.8 depicts a chart that aids protocol evaluation based on the principles underpinning the protocol framework.

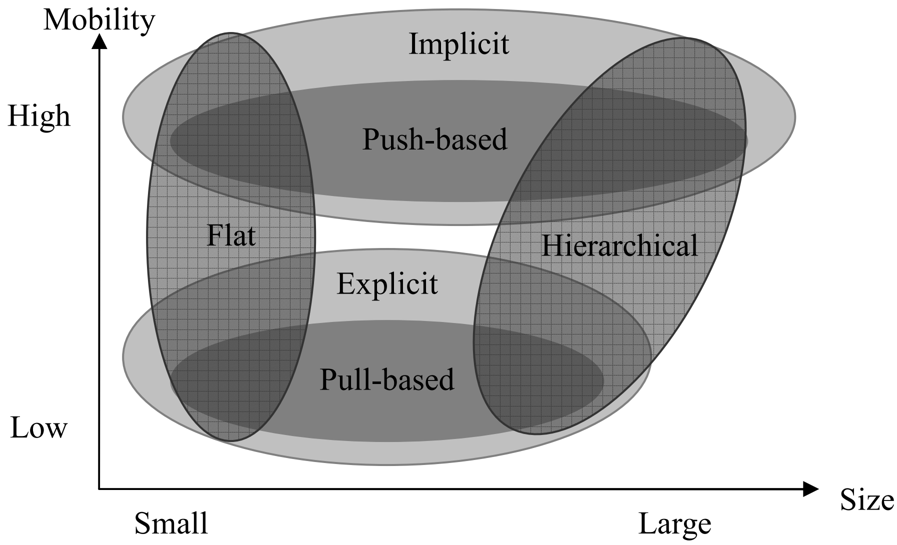

Figure 3.8 Suitability of different kind of routing protocols in terms of mobility and scale.

Firstly, flat structure protocols suffer from decreasing throughput per node as node density increases. Gupta and Kumar [2000] obtain an upper bound on the throughput of ad-hoc wireless networks, which decreases as $O(1/\sqrt{n})$ per node, as the number of nodes n increases. Their results motivate the design of hierarchical structure approaches. Moreover, hierarchical structure approaches should provide better scalability than the flat structure, since the former relies on local information exchange rather than global information flooding to perform data routing. However, a trade-off between routing overhead and protocol complexity exists. An effective method is needed to maintain the hierarchical structure with minimum influence on routing performance. Secondly, pull-based protocols may not be suitable in more dynamic networks since there are too many and too frequent update messages need to be circulated proactively, in order to maintain the correct and 'not stale' route for data forwarding. Conversely, push-based protocols reactively find the latest route to the destination only when needed and the routing path needs only to be maintained during the period of data transmission. Finally, the implicit connection based protocols inherently adapt to changes in network topology and consequently provide better network scalability in terms of node density and physical network size. In the next Section, a more detailed study is presented to show the advantages of implicit connections.

3.3 Analysis of Routing Performance for Explicit/Implicit Routing Solutions

As illustrated in Figure 3.8, the founding principles of the work presented in the thesis is that implicit connections inherently adapt better to changes in network topology compared with explicit connections. This is due to the fact that the implicit routing does not need to create nor maintain a path every time a source-destination connection request is made. Implicit connection utilises the actual/relative geographical (n.b. except Flooding which requires no network information) information to guide the forwarding packet to any or specific destination (dependent on the type of geographical information used) regardless of the level of mobility of intermediated sensor nodes integral to the path. With explicit connections, the established routing path is prone to be broken rapidly when nodes move. Consequently, additional energy and bandwidth is consumed to recover the broken path.

Several analysis metrics are proposed to quantify the routing performance of explicit/implicit connections as a function of a number of key network parameters.

3.3.1 Analysis Metrics

Increasing node mobility and node density has a strong impact on network performance in terms of connection reliability, connection stability and capacity. Node mobility is the most studied factor used to evaluate the routing performance of routing protocols [Bai et al. 2004]. Moreover, increasing node density as well as physical network size also strongly impact on network performance. This section investigates the impact on routing performance taking node mobility, node density and physical network size into account. In order to further capture and quantify those impacts, five analysis metrics are used: Explicit Connection Duration (*ECD*), Explicit Connection Interval (*ECI*), Implicit Connection Duration (*ICD*), Implicit Connection Interval (*ICI*) and Implicit Connectivity (*IC*). Those metrics are used to represent the stability of explicit connections, stability of implicit connections, and the real-time capacity of connections.

3.3.1.1 Explicit Connection Duration (ECD)

Considering an explicit connection protocol like AODV, an end-to-end routing path is established between source and destination - say node i and node j - after propagation of RREQ and RREP as discussed previously. Assuming that node i and node j are in the first instance connected by the end-to-end routing path R = $\{n_1, n_2, \ldots n_k\}$ consisting of k nodes (including the source node i and destination node j), at time t_1, the *ECD* is defined as the length of the longest time interval $\{t_1, t_2\}$. t_2 is the time when one of the $k - 1$ links (n_1, n_2), (n_2, n_3),..., (n_{k-1}, n_k) becomes broken;

consequently the source node and destination node become disconnected [Bai et al. 2004]. Formally,

$$ECD_{ij}(n_1, n_k, t_1) = \min_{1 \le z \le k-1} LD(n_z, n_{z+1}, t_1) \tag{3.1}$$

where LD (n_z, n_{z+1}, t_1) is the link duration between nodes n_z and n_{z+1}. Moreover, two nodes n_z and n_{z+1} are not within the transmission range of each other at time $t_1 - \varepsilon$ and $t_2 + \varepsilon$ where ε is any given time period and $\varepsilon > 0$. Note that ECD_{ij} (n_1, n_k, t_1) is the shortest path between node n_1 and n_k viz. node i and node j at time t_1 determined by the Breadth-First Search (BFS) [Cormen et al. 1998], a path selection rule in explicit connection protocols (the shortest path is selected when multiple paths are available). The average ECD representing the reliability of an end-to-end routing path from the perspective of the entire network, can then be obtained by:

$$\overline{ECD} = \frac{\sum_{t_1=0}^{T} \sum_{i=1}^{N} \sum_{j=i+1}^{N} ECD_{ij}(n_1, n_k, t_1)}{P} \tag{3.2}$$

where T is the evaluation time. N is the number of mobile nodes, k is the number of node consist the explicit connection, P is the number of tuples (n_1, n_k, t_1).

3.3.1.2 Explicit Connection Interval (ECI)

The explicit connection interval $ECI(i,j,t_1)$ is defined as the time interval $\{t_1, t_2\}$ between two consecutive end-to-end path connections between node i and node j. Specifically, both t_1 and t_2 are the moment the source node and the destination node become connected. Moreover, two nodes i and j are connected by an end-to-end path at $t_1 + \varepsilon$ and $t_2 + \varepsilon$ and disconnected at time $t_1 - \varepsilon$ and $t_2 - \varepsilon$ for $\varepsilon > 0$. The average ECI can be written as:

$$\overline{ECI} = \frac{\sum_{t_1=0}^{T}\sum_{i=1}^{N}\sum_{j=i+1}^{N} ECI(i,j,t_1)}{P} \tag{3.3}$$

where P is the number of tuples (i,j,t_1).

The average *ECI* represents the degree of dynamism in the change of explicit connections i.e. the connection stability. In a dense network (where network partitioning is rare), the value implies how frequently the end-to-end path needs to be re-established. More routing overhead is introduced by frequent changes in the end-to-end path, consequently decreasing the overall routing performance. Note that in a sparse network, the average *ECI* is not a suitable representation of the degree of dynamics in connection, as a large value is not indicative of nodes enjoying less changes in connection, but is more likely an indication of network partitioning [Xue and Kumar 2004].

Note that the value of *ECI* is always equal or larger than *ECD*. When source and destination is still connected, but the old path is broken and the new path is established immediately, *ECI* is equal to *ECD*. When source and destination is disconnected due to network partitioning and no available path exists, *ECI* is larger than *ECD*, since the length of connection interval is equal to the length of connection duration plus disconnection duration. The difference between *ECD* and *ECI*, gives the disconnection duration representing potential disjoint packet delays. When a packet fails to reach its destination due to the lack of available paths, it will be temporally buffered and carried by the sender or intermediated node until an end-to-end path is found between sender and destination; large delays may result (n.b. buffer delay is larger than transmission delay). As expected, the value of average *ECI* is larger than the average *ECD*.

3.3.1.3 Implicit Connection Duration (ICD)

Considering an implicit connection protocol like Flooding, the connection between source node and destination node (say node i and node j) depends on the connectivity of the entire network. As long as more than one path exists between source and destination, the two nodes are said to be connected implicitly. Assuming that the two nodes i and j are initially connected by a single routing path R_1 or multiple routing paths MR = $\{R_{1,} R_{2, \ldots} R_m\}$ consisting of m different paths (m varies from time to time during its connection), at time t_1, then the *ICD* is the length of the longest time interval $\{t_{1,}\ t_2\}$. t_2 is the time when no path exists between node i and node j or these two nodes i and j belong to a different network partition. Formally,

$$ICD_{ij}(R_1, R_m, t_1) = \max_{1 \le x \le m} PD(R_x, t_1) \tag{3.4}$$

where PD $(R_{x,}\ t_1)$ is the path duration of path R_x measured from time t_1. Path $R_x = \{n_{1,}\ n_{2, \ldots} n_k\}$ consisting of k nodes and corresponding path durations can be calculated by using Equation (3.1). The average *ICD* representing the reliability of an implicit connection from the perspective of entire network can then be obtained by:

$$\overline{ICD} = \frac{\sum_{t_1=0}^{T} \sum_{i=1}^{N} \sum_{j=i+1}^{N} ICD_{ij}(R_1, R_m, t_1)}{P} \tag{3.5}$$

where P is the number of tuples (R_1, R_m, t_1)

3.3.1.4 Implicit Connection Interval (ICI)

The implicit connection interval $ICI(i,j,t_1)$ is defined as the time interval $\{t_{1,}\ t_2\}$

between two consecutive implicit connections between node i and node j. Specifically, both t_1 and t_2 are the instants the source node and the destination nodes become connected Moreover, two nodes i and j are connected by an implicit connection at $t_1 + \varepsilon$ and $t_2 + \varepsilon$ and disconnected at time $t_1 - \varepsilon$ and $t_2 - \varepsilon$ for $\varepsilon > 0$. The average *ICI* can be written as:

$$\overline{ICI} = \frac{\sum_{t_1=0}^{T}\sum_{i=1}^{N}\sum_{j=i+1}^{N} ICI(i,j,t_1)}{P} \tag{3.6}$$

where P is the number of tuples (i,j,t_1)

The average *ICI* represents the degree of change of implicit connections i.e. connection stability. It also serves as an indicator to compare with explicit connections and reveals the advantage of implicit connections when different network parameters are considered.

Similar to the explicit connection, the value of average *ICI* is always equal or larger than the average *ICD*. The source and destination will become disconnected if and only if there is a network partition. Therefore the difference between *ICD* and *ICI* is the disconnection duration which represents potential disjoint packet delays for implicit connection solutions [Hay and Giaccone 2009]. When a packet fails to reach its destination owing to a lack of available paths (network partition occurs), it will be temporally buffered and carried by the sender or intermediated node until at lease one path is found between sender and destination; a large delay may result.

3.3.1.5 Implicit Connectivity (IC)

Implicit connectivity is defined as the fraction of time for which at least one path is available between two nodes *i* and *j* [Bai et al. 2004]. Formally,

$$IC(i,j) = \frac{\sum_{t=0}^{T} Available(i,j,t)}{T} \tag{3.7}$$

where the *Available*(*i*,*j*,*t*) is set to 1 when at least one path is available between source to destination, the value is set to 0 otherwise. Then the average *IC* can be written as:

$$\overline{IC} = \frac{\sum_{i=1}^{N} \sum_{j=i+1}^{N} IC(i,j)}{P} \tag{3.8}$$

where *P* is the number of tuples (*i*,*j*)

The average *IC* represents the potential capacity for implicit connection based routing. Packet delivery using implicit connection depends directly on the connectivity of entire network. Thus for instance, considering the simplest implicit connection based protocol Flooding, the packet delivery ratio can be predicted by calculating the average implicit connectivity (ignoring the physical layer interference such as collision and fading).

3.3.2 Results

In order to fully understand the impact of physical conditions on network performance for both explicit and implicit connection based protocols and allow a comparative analysis to be carried out, the five metrics described above are evaluated using the OMNeT++ network simulator [OMNeT++ 2003]. Firstly, this section shows how the proposed metrics differentiate between node mobility, node density and physical network size. At the same time, this section reveals the superior performance of implicit connections, providing a greater potential to provide a solution that supports WSNs with high node mobility and network scalability.

The Random Waypoint Mobility Model is used to generate different mobility patterns for the purpose of the comparison. Unless specified otherwise, the scenario contains 27 mobile nodes randomly distributed and moving in an area of 600mx600m for a period of 30 simulated minutes. The minimum speed is set to 0m/s, the maximum speed 10m/s, with a pause time of 20s. The transmission range of the nodes was assumed to be 160m (a typical distance for the low frequency wireless device, such as 433MHz or 916MHz) in each direction [Maroti et al. 2005]. For the following results, each data point represents an average of 10 runs using different seeds with a corresponding confidence interval of 95%. During the simulation, the time resolution is set to 0.1 simulated second.

3.3.2.1 Results with Varied Node Mobility

Figure 3.9 shows the average *ECD* and *ECI* as a function of moving speed. Firstly, in respect of *ECD*, as expected, the average connection duration decreases significantly as maximum speed is increased. This decreased connection duration means that the explicit connection between any two nodes tends to be broken easily; explicit connections are thus considered to be unreliable. Moreover, the value of *ECI* also decreases and this supporting result corroborates the conclusion that explicit connections change rapidly at high moving speeds indicative of unstable connections.

As highlighted in Section 3.3.1.2, the result shows that the value of average *ECI* is larger than the average *ECD* as the disconnection duration has been taken into the account. The average disconnection duration is equal to the difference between the average *ECI* and *ECD* and decreases as moving speed increases. The rapid changes in network topology at high speed stimulates an immediate reconnection (disconnection duration is zero) after a path is broken. In most cases, the disconnection duration is 0s, unless the disconnection is induced by a network partition. In the case of a network partition, a new path is re-established quicker as the network merges at a more rapid rate due to the high speed dynamics. The average disconnection duration tends to be shorter as moving speed increases [Afanasyev et al. 2009].

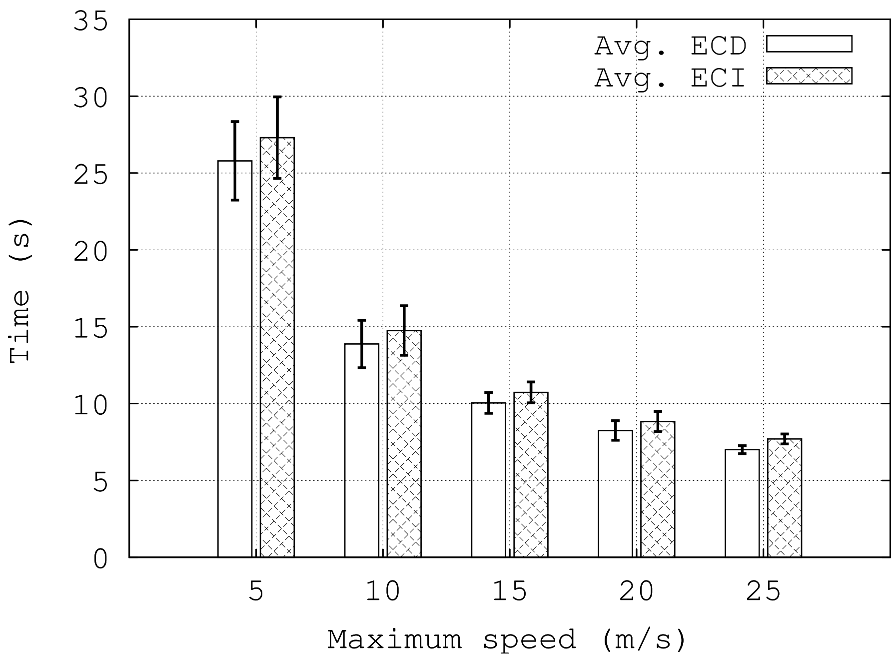

Figure 3.9 Average ECD and average ECI as a function of maximum node speed.

Figure 3.10 shows the average *ICD* and *ICI* as a function of maximum node speed and compared with the average *ECD* and *ECI*. The average *ICD* and *ICI* trend is similar to the one in Figure 3.9, in which, the value of average *ICD* and *ICI* decrease as maximum speed increases. This result implies that the reliability and stability of implicit connections are also impacted by node mobility. However, the result also shows that implicit connections are ~23 times (on average) more robust than explicit connections. As mentioned in Section 3.3.1.4, the value of average *ICI* is larger than the average *ICD* as the disconnection duration has been taken into account. The average disconnection duration equals the difference between the average *ICI* and *ICD* and decreases as the moving speed increases. This is due to the fact that when the disconnection is caused by a network partition, a new path is re-established quicker due to the relatively fast dynamics of the network.

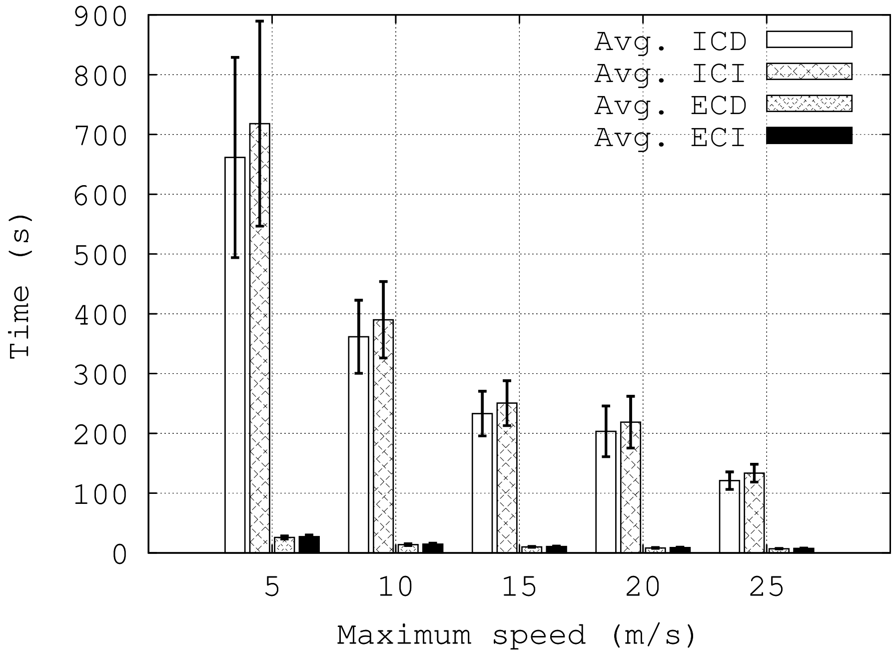

Figure 3.10 Average ICD and average ICI with varied maximum speed.

Figure 3.11 shows the average implicit connectivity as a function of maximum node speed. As the speed increases from 5m/s to 25m/s, the average implicit connectivity stabilises at a slightly decreased value from 0.94 to 0.91. This result implies that network performance in terms of packet received rate can be maintained as high as 91% (considering Poisson traffic) even at a maximum speed of 25m/s. Although, the network topology changes rapidly due to high moving speed and no network partition is evident. In other words, the routing performance is not be affected by node mobility since the routing protocol can gracefully adapt to changes in network connections.

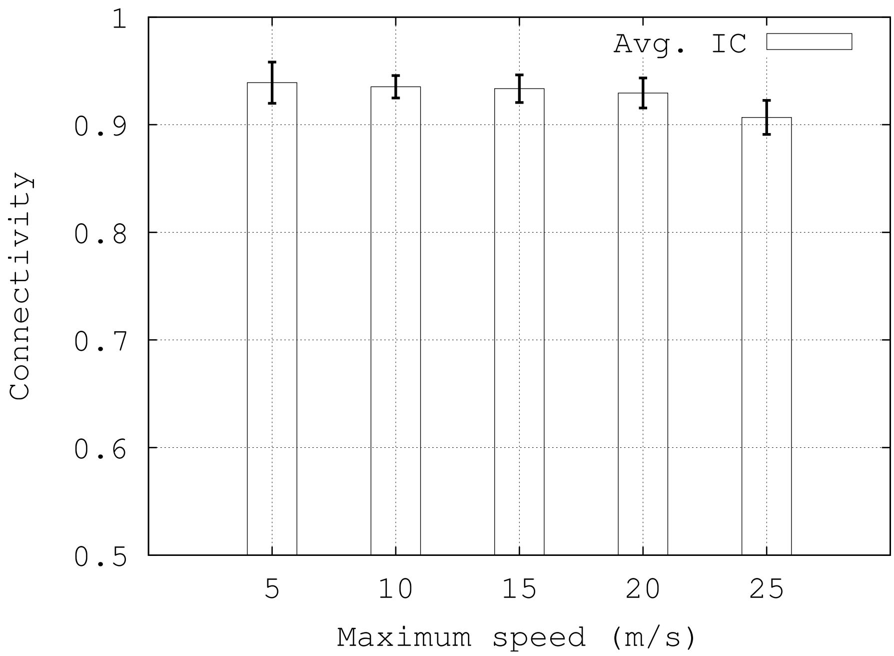

Figure 3.11 Average IC as a function of maximum speed.

3.3.2.2 Results with Varied Node Density

Figure 3.12 shows the average *ECD* and *ECI* as a function of node density. The result shows both the average *ECD* and *ECI* decreases significantly as the number of nodes increases, indictaing that explicit connections are unreliable. This phenomenon is caused by the "edge effect" reported in [Lim et al. 2002]. The authors found that the shortest routing path is very unreliable in high density mobile wireless networks since the shortest path is composed of links that connect farthest neighbours which are more likely to be located on the edge of transmission range of other nodes. In this case, a small movement of any node may easily break the path. In contrast in low node density networks, the shortest routing path is established across neighbours located far away from the edge of the transmission range resulting in higher robustness to mobility. Moreover, the duration of the disconnection tends to zero as the number of node exceeds 45. Zero disconnection duration coupled with low average *ECD* and *ECI* means that the explicit connection suffers rapid and periodic disconnections although the network remains fully connected due to the high node density.

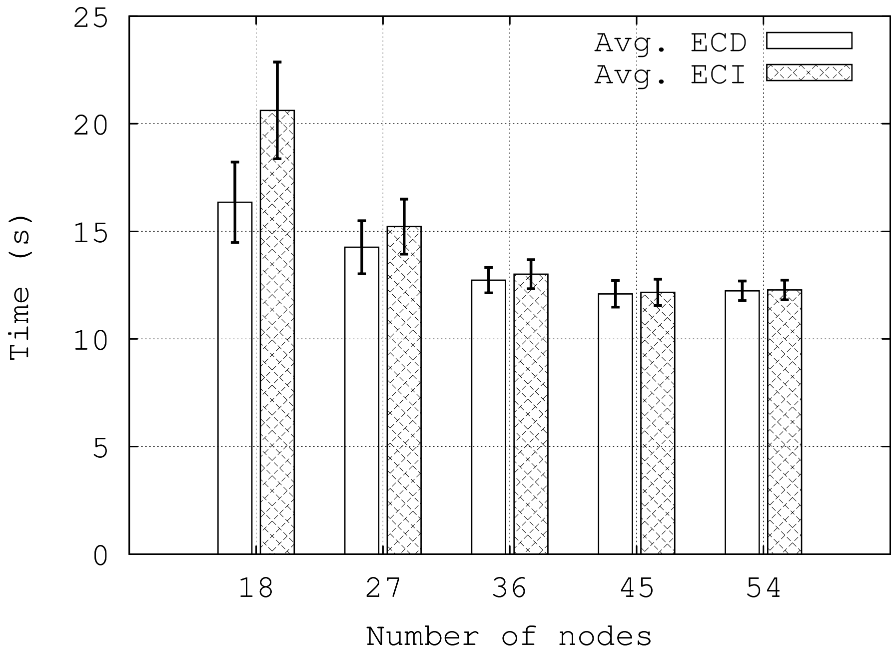

Figure 3.12 Average ECD and average ECI as a function of the number of nodes.

Implicit connections provide better r stability as the number of nodes is increased (Figure 3.13). Such significant improvement is because the entire network is deemed 'fully connected' as the number of nodes exceeds 45 (the actual degree of connectivity is shown in Figure 3.14). Moreover, the duration of the disconnection decreases marginally with an increase in the number of nodes, since the node is rarely disconnected from the main network above a certain high node density threshold [Xue and Kumar 2004].

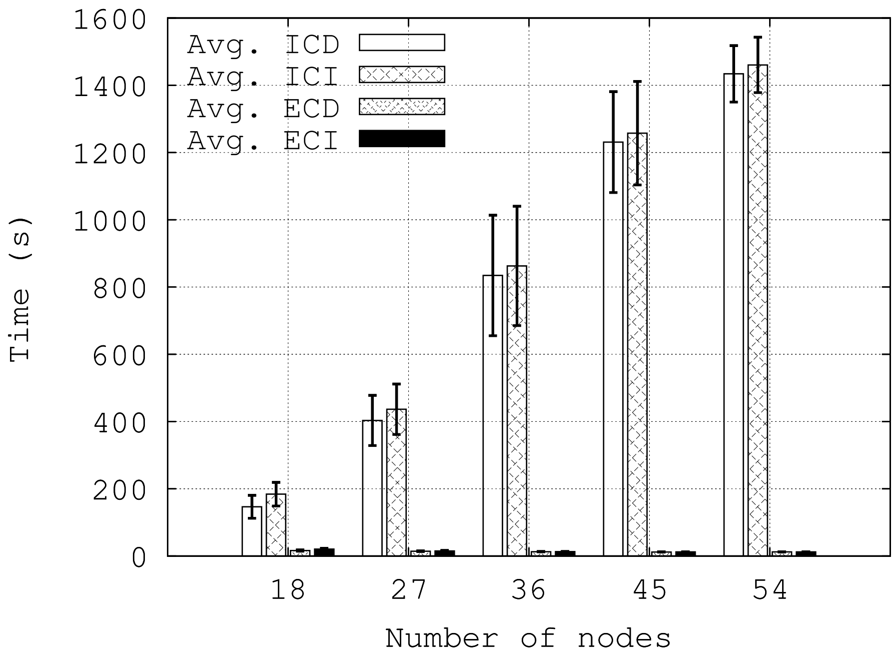

Figure 3.13 Average ICD and average ICI as a function of the number of nodes.

Figure 3.14 shows the average *IC* as a function of node density, approaching 1 (fully connected) beyond a node number of 45. This result implies that implicit connections obviate the "edge effect" and provide a potential solution for large-scale mobile wireless networking.

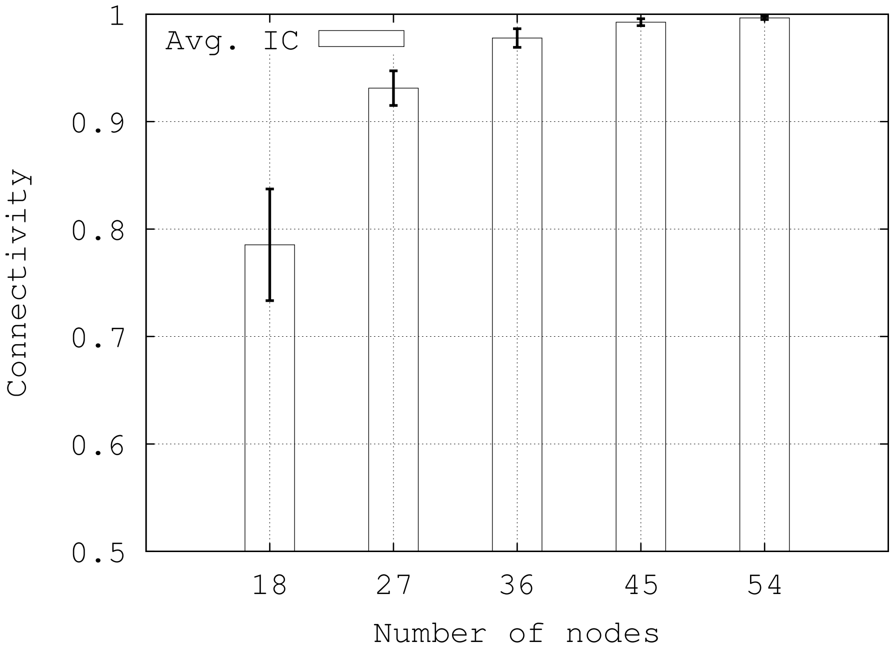

Figure 3.14 Average IC as a function of the number of nodes.

3.3.2.3 Results with Varied Physical Network Size

Figure 3.15 shows the average *ECD* and *ECI* as a function of physical network size. Note that the length of area ranges from 400m to 1000m in increments of 200m. In the previous simulation environment, 27 nodes are randomly distributed within a 600mx600m area. Other sizes are generated by scaling the square and keeping the average density of nodes constant (3 nodes per 40000m^2). As expected, results show the average connection duration decreases significantly as the physical network size is increased. Since the established explicit connection contains more intermediate nodes (*k* is increased in Equation (3.1)), the sender and receiver tend to be further apart within a large area. Paths are disconnected sooner (as expressed in Equation (3.1)) due to node mobility; therefore explicit connections are unreliable within large physical size networks. Moreover, the duration of the disconnection decreases with increases in physical size, since paths are disconnected due to changes in network topology rather than network partitioning.

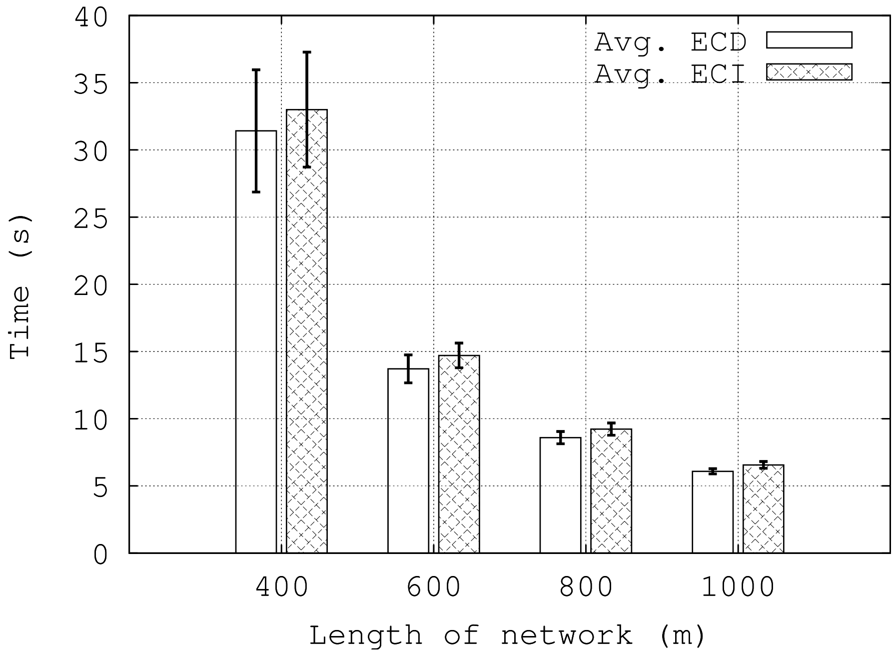

Figure 3.15 Average ECD and average ECI as a function of physical (size) length of network.

Figure 3.16 shows the average *ICD* and *ICI* as a function of physical network size and are compared with the average *ECD* and *ECI*. Both the average *ICD* and *ICI* decrease as the physical network size is increased since the probability of the network partitioning is higher when all nodes move within a larger area. Although the reliability and stability of implicit connection is impacted by increasing physical network size, it nevertheless provides a higher robustness to large scale. Therefore the implicit connection approach shows potential to be effective within large physical size networks.

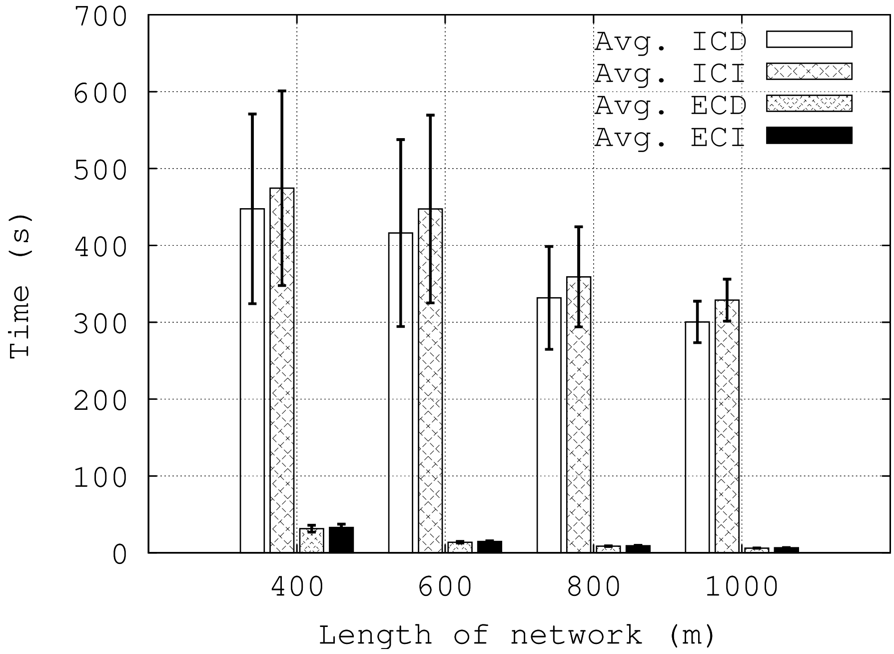

Figure 3.16 Average ICD and average ICI as a function of physical (size) length of network.

Figure 3.17 shows the average implicit connectivity as a function of physical network size, which stabilises as a function of increasing physical network size. This result shows that the potential network performance in terms of packet received rate is independent of the network size. In other words, network scalability can be achieved by utilising implicit connections.

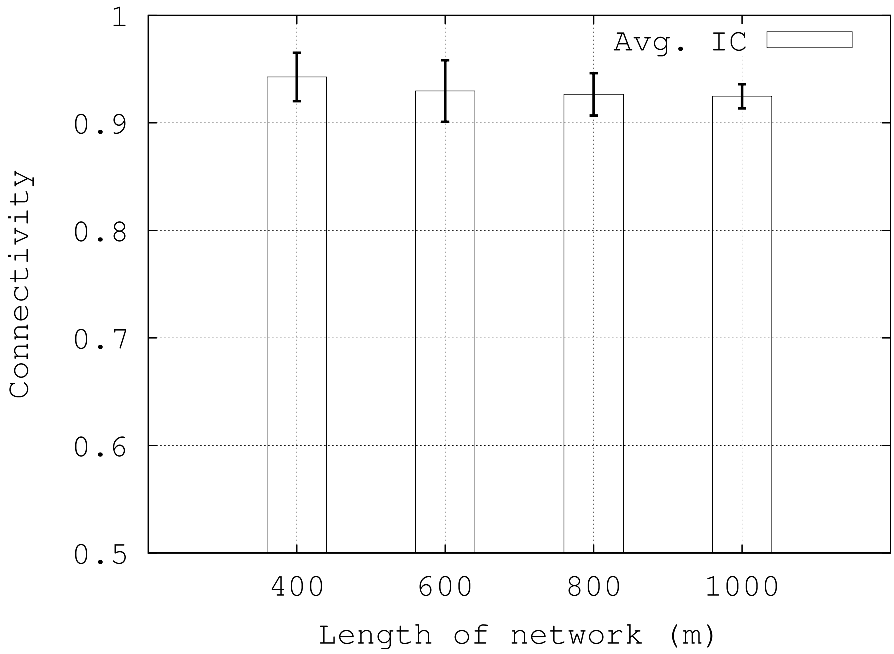

Figure 3.17 Average IC as a function of physical (size) length of network.

3.3.2.4 Discussion

Two routing algorithms - explicit connection and implicit connection - have been evaluated. The explicit connection represents the most common and widely used routing protocol, establishing end-to-end routing paths for data delivery. The implicit connection is a relatively novel approach to routing with great potential that propagates data toward its destination via a region (e.g. AIMRP) or overall (e.g. Flooding) network connections. The goal of this section is to investigate the impact of key design parameters of node mobility, node density and physical network size on routing performance of these two algorithms. Five metrics are used to capture and quantify the comparison in performance. Note that both algorithms are analyzed from the perspective of the physical layer implying that the overheads on searching and establishing the connection are not considered; these results represent the upper bound performance of the two strategies.

In summary, the analysis concludes that the explicit connection algorithm is highly impacted by increasing node mobility resulting in unstable connections. Additionally, explicit connections are also impacted by increasing node density as well as physical network size meaning that network scalability is highly constrained. In contrast, implicit connections provide a significant improvement in all cases compared with explicit connections. Specifically, implicit connections are more robust to increasing node mobility and physical network size inferring that it adapts well to rapid changes in network topology. Finally, implicit connections overcome the issue of the "edge effect" which occurs with explicit connections as node density is increased; it is able to take full advantage of dense deployments of nodes (a common characteristic in WSNs, refer to Section 2.2.3 for detail) resulting in significant increases in connection reliability and stability.

3.4 Conclusions

Different classifications of routing protocols are presented and are classified under explicit connection and implicit connection. In the explicit connection approach, a multi-hop routing path is established and maintained for data traffic transport. The main drawback of explicit routing is that the established routing path lacks flexibility to accommodate changes of network topology due to the node mobility as well as the increasing network scale. In the implicit connection approach, data traffic is transported from upstream to downstream based on geographical information and does not utilise a pre-defined path. Hence, implicit routing provides a certain degree of flexibility to adapt to changes in network topology and scale. The benefits of the latter are supported through a series of analytical analyses. Specifically, results show that implicit connections present higher stability as a function of increasing node mobility and physical network size compared with explicit connections. Results also show that the implicit connections overcome the issue of “edge effect” that occurs in explicit connections; the former takes advantage of a fully connected network resulting in a significant improvement in connection reliability and stability.

Chapter 4

Implicit Routing Schemes

4.1 Introduction

In this chapter, two novel implicit routing schemes - the Virtual Grid Paging (VGP) routing protocol and Virtual Zone Registration and Paging (VZRP) routing protocol - inspired by radio paging systems [Molisch 2005; Rappaport 2002] and cellular networks [Molisch 2005; Rappaport 2002], respectively, are proposed.

The objective of VGP is to efficiently transport data to individual or groups of mobile sinks via implicit connections maintaining the communication overhead constant regardless of the degree of node mobility and node density. To further enhance the efficiency and scalability of this routing protocol, VZRP is developed based on the concept of the 'Register and Page' strategy deployed in cellular networks [Bar-Noy and Kessler 1993]. The goal of VZRP is to diminish the influence of node mobility and achieve a highly scalable and reliable wireless connection with minimum overhead. Section 4.2 describes the concepts and inspiration behind these implicit connection protocols. Section 4.3 and Section 4.4 present VGP and VZRP, respectively, whilst Section 4.5 concludes with a summary that highlights the features of the two proposed schemes.

4.2 The Concept of Implicit Connection

One of major advantages of wireless network is untethered operation, able to provide uninterrupted communication services while the users and constituent entities are on the move. Mobility is a common feature of most wireless networks including WSNs, however, most existing studies of the latter are mostly confined to static scenarios; the issue of mobility is rarely addressed [Cugola and Migliavacca 2009]. Motivated by the potential afforded by implicit connection approaches (discussed in Section 3.3) in the scenario of varied node mobility and network scalability, the research seeks solutions using as the foundation the principles underpinning implicit connection. The two implicit connection-based wireless networks providing the inspiration behind the proposed schemes - radio paging system and cellular networking - are discussed first. Having understood the pros and cons of those two classic networking schemes, two novel and practical implicit connection-based routing protocols are developed (presented in Section 4.3 and Section 4.4). In order to leverage those concepts and render them applicable to WSNs, a topology control algorithm - the Geographic Adaptive Fidelity (GAF) [Xu et al. 2001] - is used, gracefully integrated with the proposed routing protocols to enable robust connections; further detail of GAF is also presented in this section also.

4.2.1 The Radio Paging System (RPS)

The Radio Paging System is a uni-directional wireless system very popular during the 1980s and early 1990s. The fundamental function of RPS is to deliver a short message to alert an individual, or group of individuals within the coverage area of the network typically formed by a number of base stations (BSs). Specifically, each BS sends out the same message (network-wide broadcast or so called simulcasting) initiated by the call centre which has addressed the message to an individual user or a group of users. The users are allowed to roam freely from one location to another while still able to receive the message successfully as long as they lie within the coverage of BSs regardless of the their degree of mobility. The principle of RPS is illustrated in Figure 4.1.

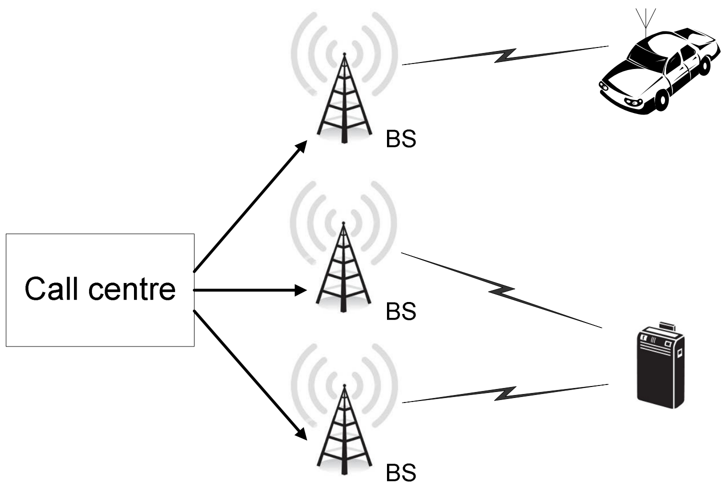

Figure 4.1 Principle of the Radio Paging System.

RPS exhibits the following features that are also desirable for WSNs;

- *Unlimited users:* due to network-wide broadcasts, the number of users (receivers) is unlimited, since they all receive the same information. In WSNs, it is likely that multiple mobile sinks are present in the network and it is desirable to be able to support any number of mobile sinks.

- *Unlimited mobility:* there is no need for a call centre to know the users locations. Connections can be considered as implicit.
- *Economical bandwidth:* RPS is efficient when the amount of transmitted data is relatively small (typically limited to about 80 characters in length); therefore the consumed bandwidth is marginal. In WSNs, due to inherent limited bandwidth resources, the packet size is also relatively small - 30bytes to 50bytes - compared to >512bytes used in mobile ad hoc networks (MANETs) [Zhou et al. 2006b].
- *Simulcasting:* information is transmitted to a conventional pager through “simulcasting” i.e. “simultaneous broadcasting” from multiple BSs. This makes paging reliable, since the signal transmitted to pagers (users) emanates from multiple locations; the chance of successfully receiving the message is thus dramatically improved. Penetration into buildings, tunnels, and underground areas is better because the signals emanate from several different directions. However, an additional communication overhead is introduced directly related to the size of network; the larger the network, more BSs are present, consequently, more overhead is consumed during simulcasting. A trade-off between reliability and overhead exists [Dulman et al. 2003].

RPS relies on an existing infrastructure (a group of base stations) that provide sufficient radio coverage. To leverage RPS concepts and make them applicable to infrastructure-less WSNs, the most critical step is to appoint a set of mobile nodes to act as the base stations mimicking the role as in RPSs. The major challenge is that the appointed sensor nodes need to be fully connected with each other through low-power transceivers (typically less than 300m in outdoor environments [Crossbow 1995]) while ensuring network-wide radio coverage. Here the solution utilises a topology control algorithm – the Geographic Adaptive Fidelity (GAF) - that provides the foundation of the implementation of the aforementioned requirement. A detailed description of GAF is presented in Section 4.2.3.

4.2.2 Cellular Networking

In a cellular network, the coverage area is divided into many small areas referred to as "cells" [Molisch 2005; Rappaport 2002]; one BS provides coverage for a cell area. The interconnection of BSs is similar to RPS viz. all BSs are fully connected and provide sufficient radio coverage over the required network area. It is important to note that, since RPS delivers short text messages to mobile users utilising simulcasting, the consequence is a severe limitation in the number of users that communicate simultaneously. Such a limitation has root in the fact that the bandwidth is shared by all users within the network, and a primitive level of interference prevails for users to communicate via broadcasting. In contrast, cellular networks deliver voice call (data hungry) quality services to typically, a group of mobile users. In other words, cellular networks require a cost (communication overhead) efficient algorithm able to handle significantly increased traffic loads especially when the number of users is large. The solution to delivering these types of services to a group of mobile users is to track or keep registration on each user(s). The two basic operation involved in tracking user(s) are Register and Page [Bar-Noy and Kessler 1993].

To illustrate the concept of tracking, two simple tracking strategies are discussed. The first is the Always-Register strategy, in which each mobile user transmits a registration message to a Mobile Switch Centre (MSC) whenever it moves into a new cell. Clearly, under this strategy the overhead due to transmissions of registration messages is very high, especially in networks with a small cell size and a large number of highly mobile users. However, since the current location of each user is always known, the overhead for MSC finding users is negligible. The second methodology is the Never-Register strategy, in which users never send registration messages to a MSC on their location. Clearly, under this strategy there is no overhead associated with the registration process. Whenever there is a need for a MSC to find a particular user, a network-wide search is initiated, the overhead of which is very high. These two simple strategies demonstrate the basic trade-off inherent to tracking. In essence, the above simple strategies are representative of the two possible extremes, the former where cost is minimized and the latter where cost is maximized [Bar-Noy and Kessler 1993]. Cellular networks use a combination of these two extremes. Many optimisation policies [Bar-Noy and Kessler 1993; Hajek et al. 2008; Madhavapeddy et al. 1995] have been proposed to minimise the total cost of tracking. Further discussion of these policies in detail lies outside the scope of thesis, but the concept of tracking inspires the development of the implicit routing (Section 4.4) protocols reported herein.

Cellular networks exhibit the following features that are also desirable for WSNs;

- *Bi-directional traffic:* in contrast to RPS (uni-directional traffic), data can originate from either the network or the user; a cellular user can be called, or initiates a call. Although most traffic in WSNs is assumed to be uni-directional (from source node to sink), on some occasions bi-directional traffic is still required, for example after a sink receives sensory data from source node, it may decide to change the sampling interval of the source node to improve data accuracy.
- *Unlimited mobility:* cellular users can be located anywhere within a network. Neither call generator nor call receiver need to know the user's location; it is the MSC that has to manage the mobility of the user, hence the connection between call generator and call receiver is considered as implicit.
- *Efficient operation:* due to the use of tracking, a call generator is able to setup a communication link directly to call receiver via a MSC resulting in a significantly reduced communication overhead. From the perspective of the entire network, available bandwidth can be reused as multiple calls can take place simultaneously.

The operation of cellular networks also relies on existing infrastructure (a group of base stations) that provides sufficient radio coverage, as well as the MSC that maintains the location of each mobile user. To leverage the concept of cellular networks and make it applicable to the infrastructure-less WSN, a set of mobile nodes need to act as base stations and particular mobile node(s) act as the MSC capable of communicating with base station-like mobile nodes and other MSCs (if more than one MSC is present). The topology control algorithm GAF once again provides the foundation for the solution.

4.2.3 Geographical Adaptive Fidelity (GAF)

As mentioned previously, to leverage the concept of RPS and cellular network and make them applicable to WSNs, there is a need to assign a group of mobile nodes to act as the base stations fully connected with each other to provide sufficient radio coverage for a functional network. In order to achieve this requirement, the following methods are considered.

Assume that there are a large number of mobile sensor nodes *N* moving within a 2D area *A*. To ensure robust network connectivity and radio coverage, the most intuitive method is to enlist each mobile sensor nodes *N* to act in the role of base station. However, the method is not efficient for message transport due to excessive overlap in radio coverage resulting in the broadcast storm problem [Tseng et al. 2002] with network-wide broadcasts (as in RPS) [Molisch 2005; Rappaport 2002].

Alternatively, a randomly selected subset of *N* can be used to act as base stations to avoid unnecessary broadcasting. However, the risk that the network will partition (not fully connected) surfaces, since it is likely that the selected base station-like nodes are not evenly distributed but are concentrated in several locations. Moreover mobility has more of an impact on network connectivity when only a small fraction of *N* are assigned in this role as demonstrated in Section 3.3.2.2.

A topology control algorithm, Geographical Adaptive Fidelity (GAF) [Xu et al. 2001], was proposed to reduce energy consumption in ad hoc wireless networks while maintaining acceptable network connectivity. GAF dynamically maintains network connectivity with only a small fraction of nodes. Therefore, GAF forms a fundamental element of the proposed protocols; VGP and VZRP are gracefully integrated with GAF to facilitate robust wireless connectivity within large scale dynamic WSNs. The protocol stack of the proposed solution is illustrated as Figure 4.2.

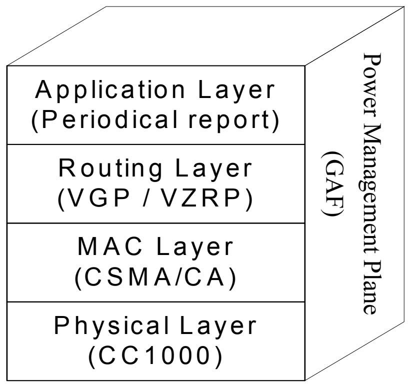

Figure 4.2 The proposed system protocol stack.

The general functionality of each layer can be found in Section 2.2.4. To aid understanding, specific protocols are assumed for each layer; the system is studied within that context. Specifically, this work considers the periodic reporting [Tilak et al. 2002] application, in which source nodes report data periodically triggered by an event. The proposed routing protocols VGP (refer to Section 4.3) and VZRP (refer to Section 4.4) manage the delivery of data to mobile sink(s). To avoid collisions, all data and routing overhead packets are sent using CSMA/CA [Bianchi 2000]. This thesis chooses the commonly available wireless device MICA2 as the basis for a test-bed platform. MICA2 features a low-power transceiver CC1000 [Crossbow 1995] which has a maximum transmission range of ~150m in an outdoor environment. Finally, use is made of GAF to improve energy efficiency by enabling collaboration amongst sensor nodes. GAF divides the network into virtual grids and within a specific sized grid, all nodes are equivalent from the perspective of routing. Then it conserves energy by indentifying redundant nodes within each single grid and powering off these nodes to limit energy consumption, whilst keeping a constant level of routing fidelity. GAF moderates this principle using system-level information together with geographical location information (in WSNs, geographical location can be estimated using [Albowicz et al. 2001; Bahl and Padmanabhan 2000; Cheng et al. 2004; Doherty et al. 2001; He et al. 2003a; Want et al. 1992]); nodes that are either source or sink remain powered on whilst the remainder monitor and balance energy use. The rest of this section provides a description of GAF.

GAF targets unnecessary energy dissipation at the MAC layer including overhearing and idle listening, two major sources of energy waste [Ye et al. 2004]. In wireless communication, packets are encoded and propagate over relative large ranges. All nodes within range receive each packet to determine if the received packet needs to be discarded or received locally. Overhearing is said to occur on those nodes that discard received packets. Idle listening is the energy consumed on monitoring the radio channel even when no traffic exists. [Stemm and Katz 1997] show that energy dissipation owing to idle listening cannot be ignored as it consumes roughly the same amount of energy as sending and receiving. In order to avoid these unnecessary energy dissipations, the GAF algorithm employs the information from the MAC layer to determine when the node can power its radio off to conserve energy lost both in overhearing, and to invoke an idle state when no traffic transport exists [Xu et al. 2001]. GAF also employs location information and active node communication to determine node density and redundancy. If there is significant node redundancy in a network, multiple paths exist between nodes, GAF powers off some intermediate nodes while still maintaining connectivity viz. it power offs redundant nodes to conserve power but not at the expense of routing performance.

With GAF, each node uses location information to associate itself with a "virtual grid", where all nodes present within the same grid are equivalent with respect to forwarding packets. The definition of a virtual grid is that for any two adjacent grids A and B, all nodes in A can communicate with all nodes in B and vice versa. Thus all nodes in each grid are equivalent in terms of routing. Figure 4.3 shows an example of a virtual grid created by GAF; here three virtual grids A, B, and C are created. According to the definition, node 1 can communicate with any of node 2, 3, and 4, node 2, 3, and 4 can all communicate with node 5. In order to satisfy the definition, the size of a virtual grid r is the upper bound. Specifically, the communication range R must be larger or equal to the possible farthest distance in any two adjacent grids. For example, node 2 of grid B must be able to reach node 5 of grid C, in which the distance is the long diagonal of grid B and grid C. Formally:

$$r^2 + (2r)^2 \le R^2 \qquad (4.1)$$

or

$$r \leq \frac{R}{\sqrt{5}} \tag{4.2}$$

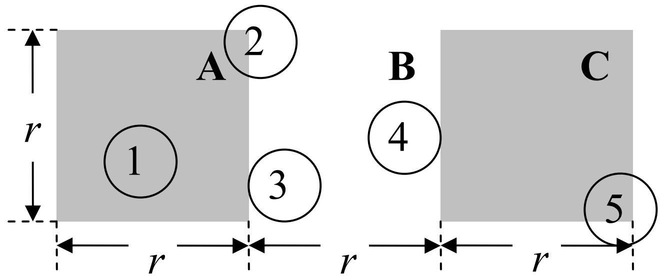

Figure 4.3 Example of a virtual grid created by GAF.

In Figure 4.3, node 2, 3, and 4 are said to be equivalent and two of them can sleep while the connectivity amongst grids A, B, and C is maintained nevertheless. Let's say node 2 and 3 are asleep while node 4 is awake. In this case, node 1, 4 and 5 are capable of serving as base stations – akin to RPS and cellular networks - since each provides sufficient radio coverage for full connectivity with neighbouring grids.

Once a virtual grid is formed, coordination is carried out among nodes in the same grid to determine which will sleep and for how long. Nodes then periodically wake up and trade places to accomplish load balancing. In GAF, nodes are in one of three states: sleeping, discovery and active. A state transition for GAF is shown in Figure 4.4.

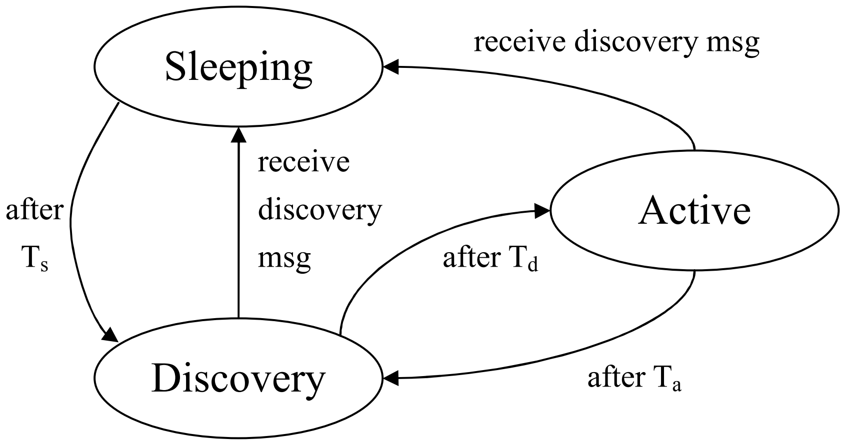

Figure 4.4 State transitions in GAF.

Initially nodes start out in the *discovery* state. In this state, a node turns on its radio and sets a timer for T_d. When the timer fires, the node decides to handle the task of routing and broadcasts its discovery message which contains node id, grid id, estimated node active time (*enat*), expected node grid time (*engh*), and node state. The discovery message is used to notify equivalent nodes how long they can sleep T_s which can then power down their radio. After the transmission of a discovery message, the node enters state *active*, stays for duration T_a before returning back to discovery state. Note that during state *active*, the node periodically re-broadcasts its discovery message at interval T_d. This re-broadcast mechanism is used to negotiate with nodes that have just joined the grid (say) due to mobility. A node in sleeping state wakes up after duration T_s before transitioning back to *discovery* state.

The estimated node active time (*enat*) is set at a value less than the time to use up all remaining energy (*enlt*). This will ensure that the node can hand off the role of active node to other nodes preventing node die off first and balancing the overall energy usage second. The expected node grid time (*engh*) is calculated based on grid size *r* and the node's current speed *s*, in which *engh* = *r/s*.

T_a is set to value *enat*. T_s is set to the smaller value of *enat* and *engt* which determines how long the node can stay in sleeping mode. The reason is that the active node may stay in the same grid until the end of *enat* or the active node may leave its grid beforehand resulting in a grid without an active node, thereby compromising routing fidelity. In either case, the node has to wake up and negotiate the role of active node.

The value of T_d is influenced by node rank in order to encourage high ranked nodes to suppress low ranked nodes, allowing the latter to enter sleeping mode as quickly as possible in so doing guaranteeing that each grid only maintains one active node - the highest ranked in the grid – maximising network lifetime [Xu et al. 2001]. Node ranking is determined by the following three rules. First, a node in *active* state has higher rank than a node in discovery state. This rule lets the prioritised *active* node continue the role for the duration of T_a. Second, for nodes in the same state, the node with longer expected lifetime (*enat*) has higher rank. This rule prioritises nodes with longer *enat* to be the *active* node in the grid. Finally, node ids are used to break ties.

[Xu et al. 2001] evaluates the performance of GAF with AODV and DSR and compares it with pure AODV and DSR. Their simulation results show that GAF saves energy by 40% to 60%. Their analysis suggests that network lifetime increases proportionally to node density, since the number of active nodes used to maintain routing fidelity remains constant; in one example, a four-fold increase in node density leads to network lifetime increase by 3 to 6 times. More importantly, GAF does not compromise data delivery. Their results also show that both GAF with AODV and unmodified AODV provide a similar data delivery ratio and average packet delay even at high node mobility (20m/s).

In the following sections, two implicit routing protocols are presented integrated with GAF, the latter not only helping to improve network lifetime at a certain node density, but also to dynamically assign the role of base station (active node).

4.3 Virtual Grid Paging Routing (VGP)

GAF partitions the entire network into a virtual grid of tetragon cells. In each grid, a highest ranked node assumes the active node role, thus driving the remainder of the participating nodes into sleeping mode to conserve energy. Discovery messages are broadcast periodically to treat node mobility.

Virtual Grid Paging (VGP) protocol is proposed for routing in large-scale mobile WSNs; VGP closely integrated with GAF forms and maintains a fully connected virtual grid network structure. The design of VPG draws from RPS principles and is a one-way data dissemination protocol, in which mobile sensor nodes are aware of their task to deliver data to the appropriate mobile sink. In the example of battlefield applications, the task (mission of the mobile node) is to record the event of interest, i.e. the abnormal physical condition of soldiers or an environmental condition of value. These events are of interest to all mobile sinks or to mobile sinks located within a certain range of a centralised source node (refer to Section 4.3.2). The remainder of this section describes the VGP protocol in detail.

As mentioned in Section 2.3, the network under consideration comprises a large number of mobile sensor nodes and mobile sinks. The traffic flow is assumed to be many-to-many (the traffic is generated from multiple mobile source nodes and delivered to multiple mobile sinks via multi-hop). The proposed VGP effectively translates the aforementioned scenario into emulating an RPS system-like network by utilising GAF. Specifically, VGP treats active nodes as temporally shifting (the active node changes dynamically according to their mobility and energy level) base stations within a RPS, since each active node is responsible for managing connections within its grid. Effectively the protocol treats the source node and mobile sink as the mobile call centre and pager user, respectively. Note that unlike the call centre in RPS which connects each static base station (infrastructure) through wired connections, the source node in VGP fully connects with each base station via multi-hop wireless links. Figure 4.5 shows an example network configuration with VGP.

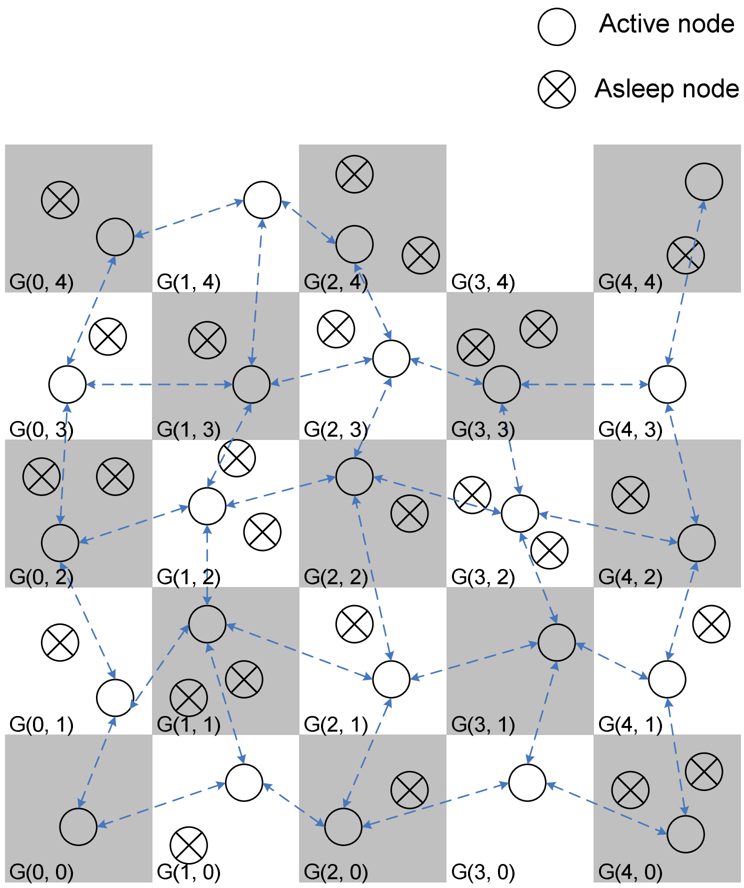

Figure 4.5 An exampled network configuration with VGP.

In Figure 4.5, grids are indentified as G(*x*, *y*) using conventional *xy*-coordinates. *x, y* represent the column and row, respectively. Circles represent mobile sensor nodes in active state and the cross circles represent those in sleeping state. There is only one (maximum) active node in each virtual grid, and each active node is fully connected with all other active nodes via multi-hop links. Note that a possibility of no mobile sensor nodes present in one virtual grid for a period of time exists, resulting in a so-called 'void grid' in the network e.g. see G(3, 4). Void grid management is discussed in the next Section.

The similarities between RPS and VGP can be characterised by the following properties:

- The mobile sink (c.f. a user in RPS) only receives data. A packet (sensory data) originates at mobile nodes (c.f. a call centre in RPS), not at mobile sinks.
- The data is intended for, and received by, a single mobile sink or a group of sinks.
- With RPS, the size of transmitted data is small - typically limited to ~80bytes in length. In WSNs, due to inherent limits in bandwidth, the packet size is also relatively small - 30bytes to 50bytes.
- With RPS, multiple base stations are deployed in a pre-determined manner to ensure coverage over an entire region. Recall that the GAF algorithm partitions the entire network into grids. Each grid is fully connected with each other via active nodes; similarly coverage of the entire network in VGP is assured.

4.3.1 Grid-based Flooding

On establishing the VGP configuration, the propagation of data from source node to multiple mobile sinks is achieved using the network-wide broadcast technique of grid-based flooding, similar to the approach used in RPS. Grid-based flooding is similar to Flooding (refer to Section 3.2.5), but with the former, only active nodes participate into the process of rebroadcast. Therefore data is propagated from the originating source node grid through neighbouring grids and so on until it has reached its destination sink or sinks. For instance, in Figure 4.5, assume that the source node in G(2, 2) has generated the data and desires to forward to all mobile sinks in the network. Regardless of the position of mobile sinks, the source node at G(2, 2) will firstly broadcast the data to the nearest active nodes. Since the source node and active node at G(2, 2) have the same connectivity to its neighbouring active nodes, the broadcasted data is guaranteed reception by active nodes located at G(1, 2), G(2, 1), G(2, 3), G(3, 2) and G(2, 2) as long as one active node exists in those grids. The remaining mobile nodes may not able to receive that data; either they may be out of transmission range of the source node or are currently in sleeping state. On receipt of new broadcasted data, all active nodes rebroadcast that data. The rebroadcasted data is then transmitted to active nodes located at G(0, 2), G(1, 1), G(2, 0), G(3, 1), G(4, 2), G(3, 3), G(2, 4) and G(1, 3). The procedure of rebroadcast will repeat until the data floods throughout the entire network. Thus, the source node is able to communicate with all mobile sinks via intermediated active nodes as long as both source node and mobile sink are within radio coverage of any active nodes.

A network may contain void grid(s), as a consequence of an uneven deployment of mobile sensor node throughout the network area, resulting in no active node within a grid. However, void grid(s) do not impact grid-based flooding. When data propagates to the edge of a void grid, it naturally detours around that grid and continues to travel beyond. It is unlikely that the network will partition because of void grid(s) unless the number of participating mobile nodes is insufficient to form a fully connected network. The issues with sparse networks are not in the scope of this thesis; some relevant studies can be found in [Pelusi et al 2006; Zhang 2006; Zhang and Wolff 2008].

The above seemingly simple grid-based flooding operation overcomes several challenges. As discussed in Section 3.2.5, two major deficiencies of Flooding - implosion and resource blindness - are eliminated through integration with GAF. Recall that implosion is an undesired phenomenon in which duplicated messages are sent to the same nodes needlessly consuming additional resource i.e. bandwidth and energy. Due to GAF in VGP, the cost of the flooding is controlled, since a constant number of active nodes promote flooding. Resource blindness occurs because Flooding does not take into account available energy resources. Recall that in GAF, the node with most energy acts as the active node; nodes with less energy resources are driven into sleeping mode. VGP harnesses the advantage from Flooding that provisions highly robust connections to changing network topologies; in fact, in highly mobile networks flooding may be the only feasible routing strategy [Goldsmith 2006]. Finally, VGP does not require any prior knowledge of global topology; each node acts based on information of its local neighbourhood only. In order to avoid data circulating endlessly within the network, the data needs to piggyback additional essential information, discussed further in the next Section.

4.3.2 Data Format

The proposed data format used in grid-based flooding is shown in Figure 4.6. The format comprises the source id, sequence number in the packet (seq num), maximum number of hops (max num hops) and payload (data). The first two fields - source id and seq num - are used to avoid packet circulating endlessly; an active node checks the recorded cache every time data is received and only rebroadcasts the data it has not seen yet, similar to ordinary flooding. Max num hops is an optional field used to indicate the maximum hop (grid) distance the packet can travel, avoiding the needless transmission of packets e.g. data targeted to mobile sink(s) located within certain range of origin. The value of max num hops decreases by 1 each time it is rebroadcasted. The value of max num hops is tuneable according to the requirement of the application. Finally, the payload field contains the data acquired by the source node.

source id 2bytes	seq num 2bytes	max num hops 2bytes	payload

Figure 4.6 Data format for grid-based flooding in VGP.

4.3.3 Collision Resistance

VGP provides some degree of resistance to collisions enhancing the success of data dissemination. When the data packet is transported, every active node in the grid will rebroadcast the received packet once. In this case, each active node (mobile sink) is guaranteed to receive at last 4 copies of the same packet from its vicinity grids as flooding propagates throughout the network. The probability that all received copies contend for the same active node (or mobile sink) is negligible. In general, the average number of received copies is equal to the average number neighbouring active nodes. Without considering the edge effect, the average number of neighbours (NH) for any active node is given by the following [Hamma et al. 2006]:

$$NH = \frac{\pi R^2}{C^2} N \tag{4.3}$$

where R is the communication range, C is the field side and N is the total number of active nodes in network area C^2. Assume that the size of the virtual grid is set to its maximum equal to $R / \sqrt{5}$ as given in Equation (4.2). Since every grid contains one active node, N equals to $C^2 / (R / \sqrt{5})^2$; then NH is a constant value of 15.7 calculated from Equation (4.3).

4.3.4 Discussion

The proposed VGP has the following three features advantageous for use in mobile WSNs. Firstly, improved scalability; the overhead for a mobile sensor node to disseminate a data packet is constant regardless of the number of mobile sensor nodes and mobile sinks in the network. Secondly, no additional mobile management is required for both mobile sensor nodes and mobile sinks. In fact, the degree of mobility is not a factor and the approach affords transparency to data dissemination. Finally, there is no need to establish and maintain an explicit routing path between mobile sensor node and mobile sink; therefore improved bandwidth and energy conservation results.

There is several drawbacks that limit the scalability of physical network size, system capacity and the routing efficiency. Specifically, the communication overhead is a direct function of the physical network scale (number of GAF grids), and therefore its efficiency decreases when the scale of network size is increased. Additionally, grid-based flooding (network-wide propagation) results in limited channel re-use introducing system capacity constraints. Moreover, VGP may not be effective when offered traffic load increases and/or the number of mobile sinks decreases.

A series of analyses conducted to demonstrate the efficiency of VGP in terms of overhead under different network scenarios are presented in Section 5.2. The robustness of proposed VGP is also studied comprehensively through simulations; these results are provided in Section 5.4.

4.4 Virtual Zone-based Registration and Paging Routing (VZRP)

The core of VZRP utilises the concept of Register and Page accomplished by the use of an emulated Mobile Switch Centres (MSCs) akin to cellular networks (refer to Section 4.2.2); but importantly in WSNs, no bespoke infrastructure exists to provide the MSC-like functionality. To overcome this, VZRP first groups all virtual grids into a tetragon zone and selects the most central grid as the zone header which emulates the functionalities of MSCs. VZRP target goal is to further reduce the network overhead and improve the scalability in terms of physical network size compared with VGP. Further, an Implicit Distance-based Routing (IDR) algorithm is proposed to facilitate the implementation of Register and Page with minimum overhead taking into consideration network efficiency and reliability.

Motivated by cellular networking that efficiently delivers services to an individual or a group of mobile users by tracking their location, VZRP tracks mobile sink(s) via two basic operations: Register and Page which are the key to improving the scalability in terms of physical network size. In more detail, Register: mobile sinks register their location (current grid id) to the zone header periodically; Page: when data needs to be forwarded to mobile sinks, the source node only sends data to the zone header, and then data is paged to the last grid location of each mobile sink recorded in the zone header. Similar to VGP, the connection between source node and mobile sinks is implicit, since there is no need to establish and maintain the routing path. The details of zone formation, required data format, IDR algorithm, and the Register and Page process are described in the following sections.

4.4.1 Zone Formation

Recall GAF uses location information to divide the network into virtual grids. Every mobile sensor node and mobile sink has a predefined mapping of a location following its grid coordinates. The size of the grid is *r* * *r* as illustrated in Figure 4.3. Assuming the physical network size is *X* * *Y*, VZRP groups all virtual grids into a single zone i.e. the zone contains grids numbering between G(0, 0) and G(*ZX*, *ZY*). *ZX* and *ZY* represent the column and row of *x,y*-coordinate, respectively, and can be calculated as:

$$ZX = \left\lfloor \frac{X}{r} \right\rfloor, \qquad ZY = \left\lfloor \frac{Y}{r} \right\rfloor \tag{4.4}$$

where $\lfloor k \rfloor$ is the largest integer less than *k*. After the grid and zone are formatted, a grid located in the centre of the zone is then selected as the zone header. Each mobile sensor node and mobile sink can calculate the grid coordinates of the zone header by:

$$G(\left\lfloor \frac{ZX}{2} \right\rfloor, \left\lfloor \frac{ZY}{2} \right\rfloor) \tag{4.5}$$

For example, in Figure 4.5, assuming that the grid size is *r* * *r* and the physical network size is 5*r* * 5*r*, according to Equation (4.4) and Equation (4.5), the grid coordinate of zone header is G(2, 2). GAF ensures there is only one active node in each grid (including zone header), except in the rare case of a void grid. Thus, each virtual grid (acting like a base station) is able to communicate with the zone header (acting as a MSC) via direct or multi-hop connections.

4.4.2 Packet Formats

Two packet formats - REG and PAG - are used for Register and Page operations (Section 4.4.3). The REG packet is generated periodically by mobile sinks and forwarded to the zone header, registering each mobile sinks' current location in relation to the zone header. Therefore, each zone header can track the location of registered mobile sinks in real time. The transmission of a REG packet is governed by the Implicit Distance-based Routing (IDR) algorithm, further explained in Section 4.4.4. The REG packet format consists of the sink id, sequence number (seq num), sink grid id (the current gird coordinate of sink), dest grid id (the destination grid coordinate viz. the grid coordinate of the zone header) and Distance Vector (DV). Similar to Flooding, the sink id field and seq num fields are used to avoid endless packet circulations. The sink grid id and dest grid id fields are used for mobile sinks to register their location in respect to zone header and indicates its last grid coordinates. Note that the dest grid id field always contains the grid coordinates of the zone header, since the zone header is solely responsible for the tracking of mobile sinks. So far, the values of aforementioned fields are fixed and given by the initialised mobile sink. It is worth noting, the value of the DV field is dynamic and overwritten by intermediated mobile nodes (active nodes) to direct data towards the zone header. DV is the relative distance in grid co-ordinates between generator or forwarder of a REG packet to the zone header; such information is critical for routing decisions to be made in the proposed IDR algorithm (refer to Section 4.4.4).

When an event is detected, a PAG packet is generated by the source node and paged to the zone header. Based on the last grid coordinates of mobile sinks recorded by the zone header, the PAG packet is further paged to each mobile sink individually. Decisions on the transmission of PAG packets are also governed by the proposed IDR algorithm. The PAG packet format consists of source id, seq num (sequence number in the PAG packet), dest grid id, DV, and payload. Similar to REG packets, the source id field and seq num field are used to avoid endless packet circulations. The dest grid id field contains either the grid coordinates of the zone header when the PAG packet is initialised by the source node or one of the mobile sinks to which the PAG packet is further paged. The DV field is also dynamic and represents the relative distance between generator or forwarder of PAG packets and intended destination (recorded in dest grid id). The payload field contains the data generated by the source node. The packet formats used in the VZRP are shown in Figure 4.7.

REG	sink id 2bytes	sep num 2bytes	sink grid id 2bytes	dest grid id 2bytes	DV 2bytes

PAG	source id 2bytes	sep num 2bytes	dest grid id 2bytes	DV 2bytes	payload

Figure 4.7 Packet formats in VZRP.

4.4.3 Register and Page

In order to overcome these drawbacks of VGP, the proposed VZRP utilises Register and Page to cope with the degree mobility of mobile sinks. For registration, each mobile sink periodically records its current location (grid coordinate) in the REG packet and sends it to the zone header (resembling registration) using Implicit Distance-based Routing, described in the next Section. The upper bound on reporting interval *RI* is calculated based on the physical size of the grid, *r;* and maximum velocity of mobile sink *V*:

$$RI = \frac{r}{V} \tag{4.6}$$

Such an upper bound on reporting interval ensures that the zone header is afforded an acceptable accuracy (the ability to trace the mobility of sink movements from one grid id to another) to determine a mobile sink's current location. The zone header is responsible for the creation of a table for tracking the location of participating mobile sinks. The inhabitant of the zone header (the active node at the zone centre) may change dynamically according to movement and the remaining energy resources of mobile nodes (Section 4.2.3). A 'new' zone header carries no past information on the locations of mobile sinks until the mobile sinks report through REG packets after time period *RI*/2 (on average). Before the mobile sink reports, the zone header is not functional resulting in potential packet losses. In order to avoid such a detrimental condition, the previous zone header (active node) will always handover the tracking table to the new zone header so that communication between new zone header and all mobile sinks can continue without disruption.

For the page, once the source node has acquired data, it generates a PAG packet and transmits it using IDR. Since the zone header tracks the location of each mobile sink, it takes responsibility for forwarding the packet to each sink. On receipt of a PAG packet, the zone header copies it (if more than one mobile sink is present) and then pages the current location of each mobile sinks individually. After the last copy is paged and arrives at the last mobile sink, the transmission of PAG packets is terminated.

4.4.4 Implicit Distance-based Routing (IDR)

The objective of the proposed IDR is to facilitate the operation of Register and Page with minimum overhead. The IDR algorithm has three features. Firstly, it utilises implicit connection which makes it robust to dynamic changes in network topology. Secondly, IDR determines the next preferable grid on the air without any prior knowledge of network topology; therefore both bandwidth and energy resources is conserved. Finally, IDR utilises the concomitant advantage of overhearing to avoid unnecessary packet transmission; the efficiency of IDR is optimised without sacrificing packet delivery.

When a REG or PAG packet is intended for a specific gird coordinate (say) G(x, y), (shown in the dest grid id in Figure 4.7), the packet generator first calculates the distance vector $DV_{G(x,\ y)}$ defined as the relative distance between its current grid coordinates G(cur_x, cur_y) and intended destination grid coordinates G(x, y). Formally:

$$DV_{G(x,y)} = \sqrt{(x - \mathrm{cur\_}x)^2 + (y - \mathrm{cur\_}y)^2} \qquad (4.7)$$

After $DV_{G(x,\ y)}$ is calculated, the packet generator records the value of $DV_{G(x,\ y)}$ into the packet field of DV and broadcasts the packet using CSMA/CA (to avoid collision if the channel is busy temporally). When the packet is received by nearby active nodes, each calculates their distance vector $DV_{G(x,\ y)}$ using Equation (4.7) and compares the value of DV recorded in the packet. Active nodes with smaller $DV_{G(x, y)}$ then overwrite the DV value in the packet and rebroadcast the packet using CSMA/CA with a preferable back-off time. Otherwise, active nodes drop the packet immediately. The back-off time is calculated by;

$$back\text{-}off = DV_{G(x,y)} * CW_{slot} + rand(0, slot) \tag{4.8}$$

where CW_{slot} is the minimum contention window in slots. Note that the first part of Equation (4.8) is to ensure that the active node with smallest distance vector has the highest priority to rebroadcast the packet. The second part of Equation (4.8) is to prevent collisions of the packet when multiple active nodes are addressed with the same distance vector. Due to the nature of broadcast, active nodes that attempt to rebroadcast the packet may overhear the rebroadcasted packet from the active node addressed with smaller $DV_{G(x, y)}$. In this case, those active nodes cancel the task of rebroadcasting and drop the packet immediately, since the packet has been forwarded already, there is no need to waste additional bandwidth. Such packet rebroadcasts repeat until the packet arrives at its intended destination. According to the rules of rebroadcasting, the generated packet is routed through the best possible path (shortest path) using implicit connections which utilise on the air apriori knowledge of network topology. Therefore, IDR adapts well to changes of network topology maintaining existing connections even when nodes are on the move.

In order to provide a better understanding of IDR, consider the following example. In Figure 4.8, a large number of mobile nodes are scattered over a sensing field. GAF divides the network into virtual grids which ensures that one active node is present in each. These active nodes by default drive the remaining mobile nodes (not shown in

Figure 4.8) into sleeping mode to conserve energy. Now assuming that the active node located at G(2, 2) is the source node which generates data for destination node located at G(7, 2). The source node first calculates the corresponded distance vector $DV_{G(7, 2)}$ using Equation (4.6); in this case $DV_{G(7, 2)} = 5$. It embeds $DV_{G(7, 2)}$ into the packet and broadcasts the packet. The active nodes in the neighbouring grids receive the broadcasted packet and immediately calculate their distance vectors $DV_{G(7, 2)}$ and compare it with the DV value recorded in the packet. Only active nodes located on G(3, 1), G(3, 2) and G(3, 3) satisfy the rule for rebroadcast, since only this subset of active nodes are able to propagate the packet closer to its intended destination grid G(7, 2); the remainder simply drop the received packet. Those appropriate active nodes then overwrite the DV value in the received packet and compete to rebroadcast the packet using CSMA/CA with a back-off given by Equation (4.7). In this case, active node located at G(3, 2) will negotiate access to the channel first and will rebroadcast the packet since it has smallest $DV_{G(7, 2)} = 4$ (larger than G(3, 1) and G(3, 3) which are $DV_{G(7, 2)} = 4.12$) viz. it is closet to the intended destination in this rebroadcast cycle. The rebroadcasted packet is then received by active nodes located at G(4, 0), G(4, 1), G(4, 2), G(4, 3), G(5, 1), G(5, 2) and G(5, 3) closer (their $DV_{G(7, 2)}$ values are smaller than 4) to the intended destination grid. Meanwhile, the rebroadcast packet is overheard by active nodes located at G(3, 1) and G(3, 3) which then immediately cancel the task of packet rebroadcasting to save energy and bandwidth. For the next cycle, the active node at G(5, 2) rebroadcasts the packet further while rebroadcasts scheduled by the active nodes at G(4, 0), G(4, 1), G(4, 2), G(4, 3), G(5, 1) and G(5, 3) are cancelled due to overhearing. After the second cycle is completed, the packet successfully arrives at the intended destination node located on grid G(7, 2). The active node on the intended destination grid carries out the third and final rebroadcast to suppress rebroadcasts from active nodes located at G(6, 0), G(6, 1), G(6, 2) and G(6, 3), and terminates the transmissions for its packet. In this case, the packet has delivered from source node to destination with a minimum total overhead of 4 transmissions and in which active nodes at G(3, 2) and G(5, 2) served as packet forwarders. Note that the IDR algorithm consumes no overhead in the establishment and maintenance of the routing path; the connection is implicit with no routing overhead.

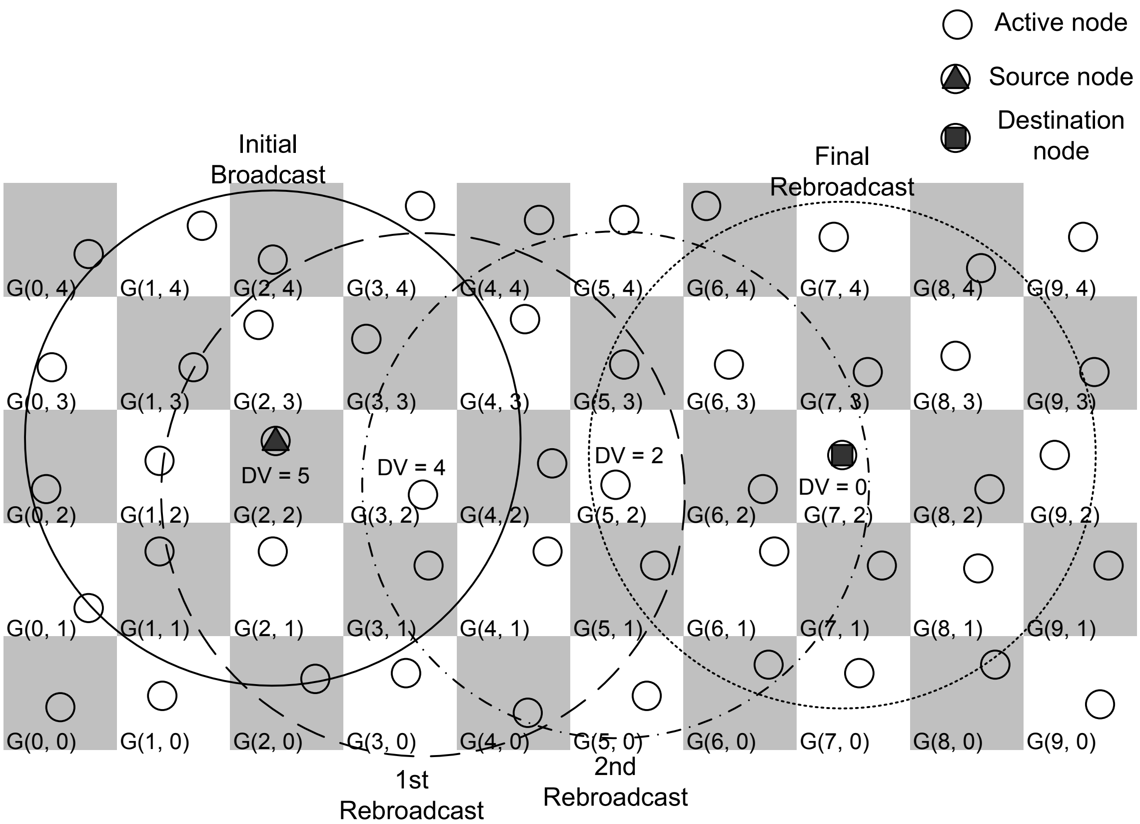

Figure 4.8 The implicit connection between source node and destination node in IDR algorithm.

The principle of IDR is that the packet is addressed towards its destination using the

DV value without taking a pre-established (explicit) routing path. From the perspective of routing, such an approach provides a certain degree of data delivery against node mobility. Note that the packet will fail to arrive at the destination node, if no neighbouring active nodes being addressed have smaller DV values after a complete rebroadcast cycle. For instance, the packet will be lost after the initial broadcast, if G(3, 1), G(3, 2) and G(3, 3) are all void grids. However, this scenario is unlikely to happen in WSNs, since, typically, a large number of mobile nodes are scattered within the network and the routing fidelity is maintained by GAF (Section 4.2.3).

4.4.5 Discussion

VZRP is designed to provide better network stability in terms of physical network size and improves the efficiency of mobile connections through a Register and Page strategy. The VZRP adopts an approach similar to the Always-Register mechanism, one of two extreme cases for tracking in cellular networks (Section 4.2.2). There are two major reasons for adopting this approach; firstly, the default size of a REG packet is 10bytes which is typically smaller than a PAG packet (including the payload). Therefore it is more cost-effective for the zone header to track the location of mobile sinks using REG packets. Secondly, the traffic flow for REG packets is many-to-one (all mobile sinks register to the same zone header) which is much easier to manage than PAG packet flow which is many-to-many.

The current design of VZRP fails to correct any errors caused by channel interference such as collisions due to the hidden terminal problem [Tobagi and Kleinrock 1975].

A retransmission mechanism is needed for IDR to provide end-to-end secure transmission. Such a requirement is critical for some WSN applications such as battlefield surveillance, intrusion detection, or detection and tracking applications and this would be the subject of future work (Section 6.2).

4.5 Conclusions

This chapter presents two novel implicit routing schemes – the Virtual Grid Paging (VGP) and Virtual Zone Registration and Paging (VZRP) protocol - inspired by the concept of radio paging systems and cellular networking, respectively. To leverage those concepts and make them applicable to infrastructure-less WSNs, the topology control algorithm - the Geographic Adaptive Fidelity (GAF) - is deployed to form a virtual grid network. Both VGP and VZRP with implicit connection provide reliable connections robust against node mobility. Moreover, since implicit connection does not have to establish and maintain a path, they are routing overhead free.

In VGP, data is propagated to multiple mobile sinks through grid-based flooding, in which the communication overhead remains constant regardless of the degree of node mobility and density. Therefore, VGP is able to support unlimited node mobility and an unlimited number of mobile node. However, the communication overhead is a direct function of physical network size. VZRP is proposed to overcome this constrain by implementing a Register and Page strategy. Mobile sinks periodically register their last location to a zone header. The location information is then available for the zone header to page the data to each participating mobile sink directly resulting in a significant reduction of the overheads. That communication overhead is further reduced by use of Implicit Distance-based Routing (IDR) protocol. Due to the inherent advantage of implicit connection and packet overhearing, the communication process of Register and Page under the management of IDR yields robust connectivity with minimum communication overhead.

A series of analyses are conducted to demonstrate the performance of the two proposed routing protocols in terms of network overhead under different network scenarios (Section 5.2). The performance of VGP is also studied comprehensively through simulations, the results of which are provided in Section 5.4.

Chapter 5

Mathematical Analysis and Simulation Results

5.1 Introduction

This chapter evaluates the performance of the two proposed implicit routing protocols VGP and VZRP and a comparison is made with a well established explicit routing solution, AODV. A mathematical analysis model is developed to determine an idealised level of routing overhead for both VGP and VZRP. Since an analytic framework can not fully treat the complexity of full communication scenarios, this chapter then describes results obtained through simulation that indicate the performance of the proposed protocols in terms of normalised routing load, packet delivery ratio, delay, how and when the proposed protocols outperform AODV and whether or not the VZRP is superior than VGP.

5.2 Overhead Analysis

This section analyses the efficiency and scalability of VGP and VZRP by evaluating the communication overhead for a number of sinks set to capture a certain amount of data from a number of sources, with all entities assumed to be moving at a constant speed. In order to re-enforce the potential of implicit connection, the results are then compared with an explicit connection approach (henceforth referred to as ECA), in which the source first floods a discovery message throughout the whole network, and then the sink responds with a reply message that travels back to the source via the (reverse) path taken by the discovery message. An explicit connection is established through a series of intermediated mobile nodes before initiating the delivery of data. AODV and DSR both adopt this approach, although each utilises different types of acknowledgement (active or passive) to ensure the success of packet transmissions over the path. Because both VGP and VZRP do not employ acknowledgements (they naturally prevent packet losses) like ECA does, acknowledgements are irrelevant when performing the overhead analysis.

The analysis focuses on the worst-case communication overhead for each protocol. The goal is to keep the analysis simple and easy to follow while capturing the fundamental differences between VGP, VZRP and ECA. For a fair comparison, note that the topology control algorithm GAF is not only considered in the analysis of VGP and VZRP, but also in the analysis of ECA. However, the overhead introduced by GAF does not materially impinge on the analyses since it performs independently and the overhead for each protocol is identical.

5.2.1 Model and Notations

This analysis considers a square sensor field $A \cdot R$ in which N sensor nodes are uniformly distributed, where R is transmission range of all entities; A is a parameter (integer) representing the scale of physical network. There are *k* mobile sinks and *s* mobile sources within the field, moving at a constant speed *V*, while receiving *d* data packets from a single mobile source. Each of *k* mobile sinks receive $s \cdot d$ data packets in total during a time period *T*. Each data packet (including headers) are of unit size *L* for all protocols and all signalling messages (such as REG packet in VZRP or RREP packet and RREQ packet in AODV) are of comparable size *l*. Due to GAF, the communication overhead consumed in flooding throughout the sensor field is upper-bounded, proportional to the number of virtual grids formed to cover the sensor field. The overhead cost of sending a message along a multiple (both implicit and explicit) connection via greedy geographical forwarding is also upper-bounded, proportional to the number of virtual grids in the connection.

GAF divides the sensor field into virtual grids; the length of grid is set to its maximum R/√5 so that each side of the sensor field has $\lceil (A \bullet R)/(R/\sqrt{5}) \rceil = \lceil \sqrt{5}A \rceil$ virtual gids; thus there are $\lceil \sqrt{5}A \rceil^2$ virtual grids in the entire field. Assuming that no void grids are created, the density of mobile nodes is sufficient to ensure more than one node in each grid, considering $N \gg \lceil \sqrt{5}A \rceil^2$. All entities - mobile sources, mobile sinks and mobile nodes - traverse *m* grids in time *T*, and *m* is upper bounded by $1 + \frac{VT}{(R/\sqrt{5})}$. For a stationary entity, $m = 1$.

In the case of unicast communication, the packet traverses a horizontal grid followed by a vertical grid to reach its destination instead of a straight-line path, and no jumping of grids occurs when a packet traverses between two distinct grids. For example, a packet traversing from G(0, 0) to G(3, 2), traverses from G(0, 0) through G(1, 0), G(2, 0), G(3, 0), G(3, 1) finally arriving at G(3, 2). In the case of flooding, the packet will traverse outward from the origin through every grid in the sensor field.

5.2.2 Overhead Analysis

The analysis starts with consideration of the communication overhead for VGP, the most straightforward but nevertheless robust solution. With VGP, the source utilises grid-based flooding to propagate each data packet throughout the entire sensor field regardless of the number or mobility of mobile sinks. The overhead for a unit size data packet to be flooded throughout the entire field is equal to the number of virtual grids i.e. $\left\lceil \sqrt{5}A \right\rceil^2$. Then the total communication overhead (CO) for all s mobile sources to communicate with all k mobile sinks during a time period of T, without considering the overhead attributed to GAF, is;

$$CO_{VGP} = s\left\lceil \sqrt{5}A \right\rceil^2 dL \tag{5.1}$$

From Equation (5.1), it is clear that the overhead burden created by VGP is independent of m (mobility of all entries) and k (number of mobile sinks). VGP adapts well to changes in network topology and satisfies the desired scalability in the number of mobile sinks. However, this overhead increases exponentially when the physical length of the network is increased.

With VZRP, each mobile sink has to register its current location with the zone header whenever it moves from one grid to another. In total, this registration will be performed *m* times in duration *T*. Each mobile sink sends a REG (signalling) packet to the zone header using IDR. The REG packet traverses from one grid to another towards the zone header located at the most central grid of the field. The overhead for a REG packet to reach the zone header is;

$$\frac{1}{2}c\left\lceil\sqrt{5A}\right\rceil l \tag{5.2}$$

where $c\left\lceil\sqrt{5A}\right\rceil$ is the average number of grids along the grid-line path from the mobile sink to a given grid in the sensor field and c ranges between 0 and 2 ($0 < c \leq 2$). When the mobile sink and the zone header are close, $c \approx 0$. When the mobile sink and the zone header are located at two, diagonally opposite corners (the farthest two points in the sensor field), $c = 2$. Because the location of the zone header is located at the most central grid, the overhead for a REG packet to travel from mobile sink to the zone header is reduced by a factor of $\frac{1}{2}$.

The overhead for a mobile source to deliver *d* data packets to the zone header, or indeed for the latter to deliver *d* data packets to a mobile sink, is equal to $\frac{1}{2}c\left\lceil\sqrt{5A}\right\rceil dL$. For *s* mobile sources and *k* mobile sinks that traverse *m* grids in duration *T*, the total overhead becomes;

$$CO_{VZRP} = km\left(\frac{1}{2}c\left\lceil\sqrt{5A}\right\rceil l\right) + s\left(\frac{1}{2}c\left\lceil\sqrt{5A}\right\rceil dL\right) + sk\left(\frac{1}{2}c\left\lceil\sqrt{5A}\right\rceil dL\right)$$

$$= \frac{1}{2}c\left\lceil\sqrt{5A}\right\rceil \cdot (kml + sdL + skdL) \tag{5.3}$$

From Equation (5.3), it is clear to see that the overhead increases as a linear function of $\frac{1}{2}c\left\lceil \sqrt{5A} \right\rceil$ as parameters *m*, *s* and *k* increase. This means that VZRP effectively manages increases in the mobility of sinks as well as increases in the number of mobile sources and/or mobile sinks.

With ECA, the mobile source initially floods a discovery message throughout the entire network, and the mobile sinks respond to mobile sources with a unicasting reply message. Once the mobile source receives the reply message, an explicit connection is established between the mobile source and mobile sink via a serious of intermediate active nodes. The overhead for performing such explicit connection establishment is;

$$\left\lceil \sqrt{5A} \right\rceil^2 l + c\left\lceil \sqrt{5A} \right\rceil l \tag{5.4}$$

where $\left\lceil \sqrt{5A} \right\rceil^2 l$ is the overhead associated with discovery message flooding, and $c\left\lceil \sqrt{5A} \right\rceil l$ is the overhead for the reply message to traverse from the mobile sink to mobile source ($0 < c \le 2$).

The explicit connection has limited flexibility in maintaining the connectivity for mobile entities. Because of GAF, explicit connections can only be guaranteed as long as all mobile sources, mobile sinks and intermediated active nodes do not move outside the grid they belong to. When any of them move outside of its grid, a new mobile node from the void grid wakes up and assumes the role of active node; the previous active node is then forced to sleep to conserve energy. Hence the explicit connection is broken at this point. The mobile source then must restart the discovery process in order to re-establish the connection to the mobile sink n.b. it is assumed that there is no alterative path available in the cache, The worst-cast overhead for a mobile source to maintain connection to a mobile sink, is governed by the need for each mobile source to execute the discovery process *M* times in duration *T*, and on average *M* is given by;

$$M = 1 + \frac{VT}{(R/\sqrt{5})} * c\left\lceil \sqrt{5}A \right\rceil \tag{5.5}$$

Since *M* increases linearly with the physical length of network *A*, the explicit connection is prone to be broken easily and needs to be established more frequently when the number of intermediated active nodes in the connection between mobile source and mobile sink increases. For a static network, $M = 1$.

For *s* mobile sources and *k* mobile sinks, the overhead for maintaining *sk* explicit connections (many-to-many) between each mobile source and mobile sink, without considering potential query aggregation, is;

$$skM \cdot \left(\left\lceil \sqrt{5}A \right\rceil^2 l + c\left\lceil \sqrt{5}A \right\rceil l \right) \tag{5.6}$$

Once the explicit connection is established, each mobile source then starts the delivery *d* data packets to each mobile sink using unicasting. The overhead to deliver *d* data packet from a mobile source to a mobile sink is thus $c\left\lceil \sqrt{5}A \right\rceil dL$. For *s* mobile sources and *k* mobile sinks, the overhead to deliver *d* data packets, without considering potential data aggregation, is;

$$sk \cdot \left(c\left\lceil \sqrt{5}A \right\rceil dL \right) \tag{5.7}$$

The total overhead (CO) of ECA is the sum of Equation (5.6) and Equation (5.7);

$$CO_{ECA} = skM \cdot \left(\left\lceil \sqrt{5}A \right\rceil^2 l + c\left\lceil \sqrt{5}A \right\rceil l \right) + sk \cdot \left(c\left\lceil \sqrt{5}A \right\rceil dL \right) \tag{5.8}$$

To compare VGP and ECA, the ratio of communication overheads is;

$$\frac{CO_{VGP}}{CO_{ECA}} \approx \frac{s\left\lceil \sqrt{5}A \right\rceil^2 dL}{skM \cdot \left\lceil \sqrt{5}A \right\rceil^2 l}, \qquad A \gg 1 \tag{5.9}$$

Equation (5.9) can be rewritten and simplified as;

$$\frac{CO_{VGP}}{CO_{ECA}} \approx \frac{d}{kM}, \qquad L \sim l \tag{5.10}$$

Thus, in a large physical size network with small data packets (comparable to the size of signalling packets), VGP asymptotically outperforms ECA in terms of worst-case overhead when the number of mobile sinks (*k*) increases and/or the network topology (characterised by *M*) changes dynamically due to increases in the moving speeds of entities and/or increasing physical network size. Conversely ECA asymptotically outperforms VGP when the traffic load increases (characterised by *d*).

To compare VZRP and ECA, the ratio of communication overheads is;

$$\frac{CO_{VZRP}}{CO_{ECA}} = \frac{\frac{1}{2}c \cdot (kml + sdL + skdL)}{skM \cdot \left(\left\lceil \sqrt{5}A \right\rceil l + cl \right) + sk \cdot (cdL)}$$

$$\approx \frac{\frac{1}{2}(kml + sdL + skdL)}{skm\left\lceil \sqrt{5}A \right\rceil \cdot \left(\left\lceil \sqrt{5}A \right\rceil l + cl \right) + skdL}, \qquad M \approx m * c\left\lceil \sqrt{5}A \right\rceil$$

$$\approx \frac{\frac{1}{2}kml}{skm\left\lceil \sqrt{5}A \right\rceil \cdot \left(\left\lceil \sqrt{5}A \right\rceil l + cl \right)}, \qquad m \gg 1$$

$$\approx \frac{\frac{1}{2}}{s\left\lceil \sqrt{5}A \right\rceil^2 + \left\lceil \sqrt{5}A \right\rceil c} \tag{5.11}$$

Thus, in high node mobility scenarios, VZRP has asymptotically lower worst-case communication overhead compared with ECA as the number of mobile sources (*s*) and/or the physical network size (characterised by *A*) increases.

Finally, to compare VGP and VZRP, the ratio of communication overheads is;

$$\frac{CO_{VGP}}{CO_{VZRP}} \approx \frac{s\left\lceil \sqrt{5}A \right\rceil^2 dL}{\frac{1}{2}c\left\lceil \sqrt{5}A \right\rceil \cdot (sdL + skdL)}, \qquad d \gg 1 \tag{5.12}$$

Equation (5.12) can be rewritten and simplified as;

$$\frac{CO_{VGP}}{CO_{VZRP}} \approx \frac{\left\lceil \sqrt{5}A \right\rceil}{\frac{1}{2}c + \frac{1}{2}ck} \tag{5.13}$$

Ignoring the constant $\frac{1}{2}c$, Equation (5.13) becomes;

$$\frac{CO_{VGP}}{CO_{VZRP}} \approx \frac{2\left\lceil \sqrt{5}A \right\rceil}{ck} \tag{5.14}$$

Thus, in a large physical size network with high traffic loads, VGP asymptotically outperform VZRP in terms of the worst-case communication overhead when the number of mobile sinks (*k*) increases. On the other hand, VZRP asymptotically outperforms VGP when physical network size increases.

5.2.3 Numerical Results

In order to provide further understanding on the worst-case communication overhead for each protocol, this section provides numerical results by coding the analysis models in MATLAB [Gilat 2004] allowing the overhead to be calculated and compared precisely under more realistic networking scenarios without simplifying the models.

As a concrete example, the analysis in this section considers the parameter values summarised in Table 5.1. The values of those physical parameters are based on the network scenario containing multiple source nodes and sinks, together with a large number of sensor nodes. All entities are mobile within a large physical area. The traffic flow is assumed to be many-to-many.

Table 5.1 Parameters utilised for numerical analysis.

s	k	N	R	A	V	l	T	L	d	c
5	10	5000	200m	20	10m/s	10	100s	50	100	1

Figure 5.1 shows the computed overhead for both ECA and VGP as a function of parameters A and d, Equation (5.8) and Equation (5.1), respectively. When the physical network size factor A is increased, the overhead for ECA increases exponentially. This significant increase is caused by the need to maintain the explicit routing path. The overhead for VGP also increases exponentially with increasing physical network size factor A. However, it is notably lower than that for ECA since implicit connections are overhead-free for path establishment and maintenance. When the number of generated packets per source is increased, the overhead for ECA increases marginally compared with VGP for which the increase is a direct function of parameter d. This result shows that;

- VGP has better scalability than ECA for increasing physical network size
- ECA has better scalability than VGP for increasing number of packets generated per source.

Both trends agree with Equation (5.10, the ratio of communication overhead between ECA and VGP.

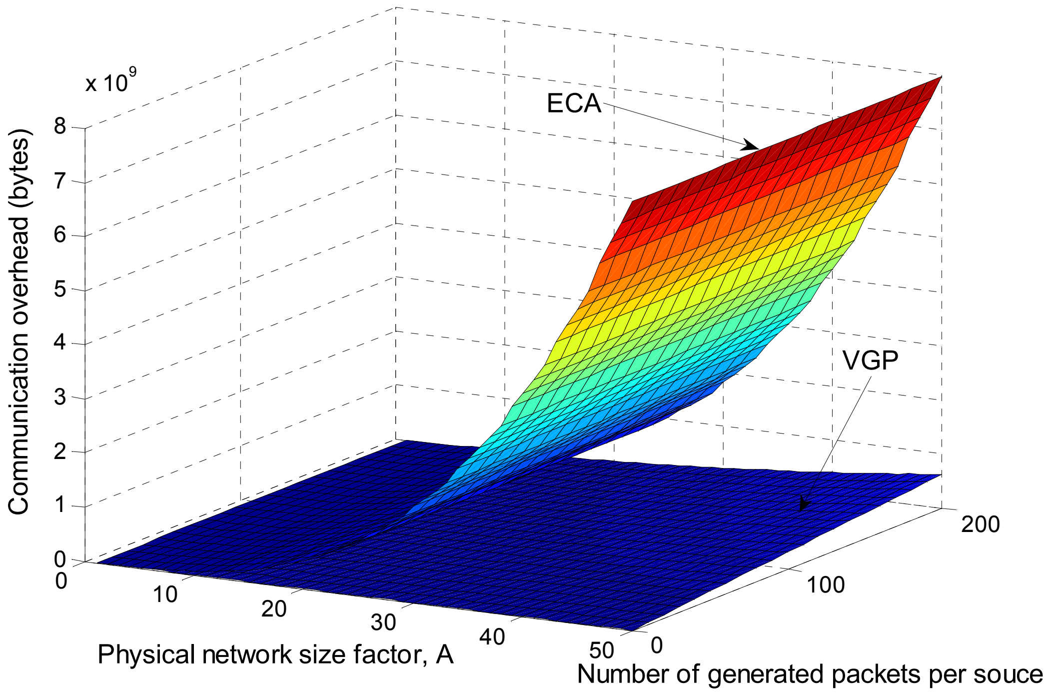

Figure 5.1 A comparison of the worst-case communication overhead for ECA and VGP.

Figure 5.2 shows the computed overhead for ECA and VZRP as a function of parameters *A* and *V*, Equation (5.8) and Equation (5.3), respectively. Note that the overhead is displayed on the z-axis in a logarithmic scale. When the physical network size factor A is increased, as expected, the overhead for ECA increases exponentially owing to the need to maintain the explicit routing path. The overhead for VZRP increases linearly as a direct function of the physical network size factor A, and is notably lower than that of the ECA. When the velocity of an entity increases, the overhead for VZRP increases slightly compared with VGP which increases linearly with velocity V. This result shows that VZRP has better scalability than ECA in terms of increasing physical network size and the velocity of mobile nodes. Those results are in agreement with Equation (5.11), the ratio of communication overhead between ECA and VZRP.

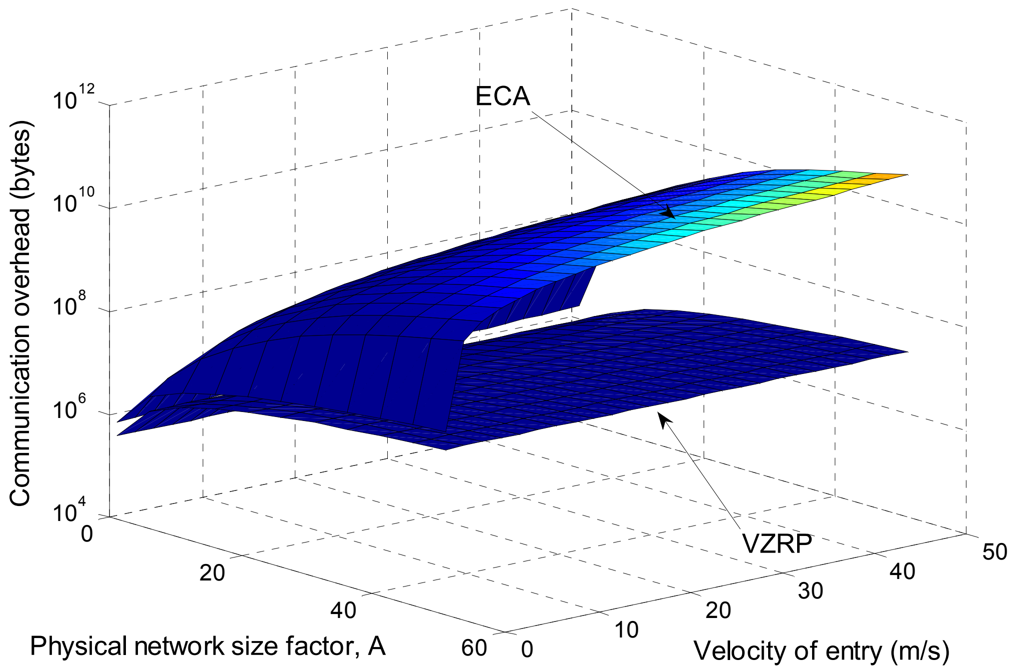

Figure 5.2 A comparison of the worst-case communication overhead for ECA and VZRP. n.b. the communication overhead is displayed on the z-axis in a logarithmic scale.

Figure 5.3 shows the computed communication overhead of both VGP and VZRP as a function of parameters A and k, Equation (5.1) and Equation (5.3), respectively. Again, the communication overhead for VZRP increases slightly compared with the VGP which increases exponentially with increases in the physical network size factor A. Moreover, the overhead for VGP is independent of the number of mobile sinks k while for VZRP, the overhead increases as a function of increases of k. This result shows that;

- VGP offers better scalability than VZRP with increasing number of mobile sinks.

- VZRP has better scalability than VGP with increasing physical network size.

Those results are in agreement with Equation (5.14), the ratio of communication overhead between VZRP and VGP.

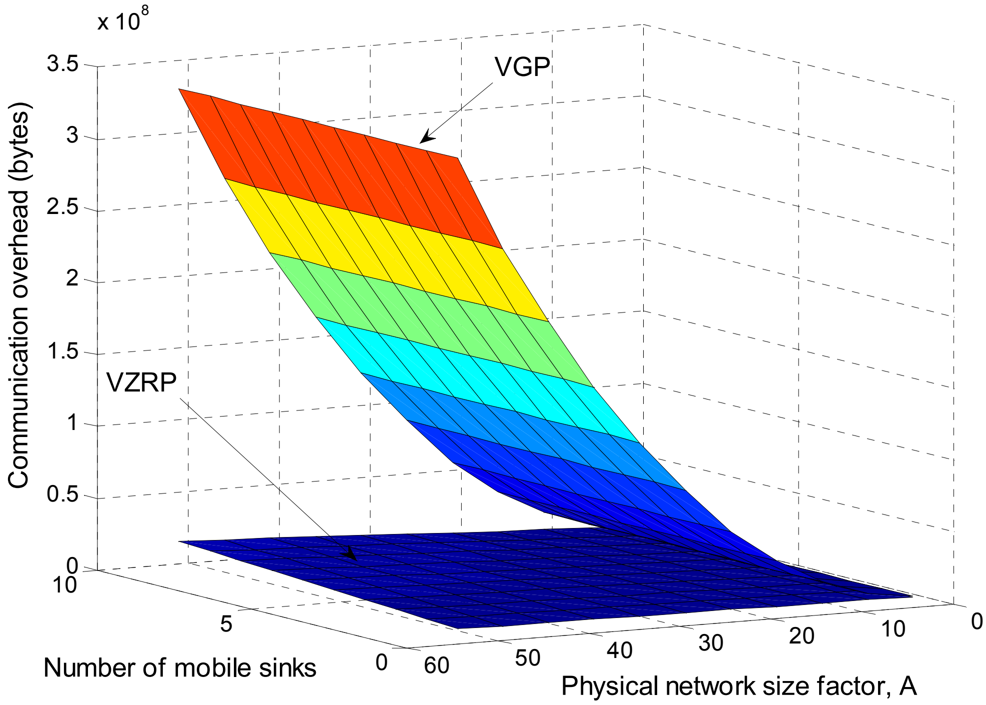

Figure 5.3 A comparison of the worst-case communication overhead for VGP and VZRP.

5.2.4 Summary

This section has evaluated the communication overheads for VGP and VZRP and compared them with an ECA protocol. The comparisons are made by defining the ratio of the communication overheads between any two protocols. Then further comparisons are made through numerical analysis without simplifying the mathematical models. In summary the results reveal that;

- VGP has asymptotically a lower worst-case communication overhead than ECA as the physical network size and node velocity increase.
- VZRP has asymptotically a lower worst-case communication overhead than VGP as the physical network size and traffic load increase, since VZRP efficiently delivers data traffic by tracking the location of mobile sinks.
- VZRP also has asymptotically a lower worst-case communication overhead than ECA under a number of network conditions due to the use of implicit routing and tracking strategy of VZRP.

In the following Section, simulation is used to explore the proposed VGP and VZRP performance under more realistic network conditions.

5.3 Simulation Framework

It is difficult to capture the details of VGP and VZRP performance purely within an analytical framework. Therefore, the Network Simulator 2 (NS-2) [NS] is used to implement all the functionalities of VGP and VZRP, evaluating variations in node movement, node density, traffic pattern, traffic load, packet size and physical network size.

Although there are several well-known simulators that can be used for the analysis of WSNs such as Network Simulator 2 (NS-2), GloMoSim [GloMoSim], QualNet [QualNet], OPNET [OPNET], OMNeT++ [OMNeT++], this thesis employs the NS-2 as the platform for several reasons:

- Widely utilised in analysing WSNs.
- Open source.
- License free.
- Many existing routing protocols and algorithms are available under NS-2, including AODV and GAF.
- New modification can be easily integrated into the existing platform.

Detailed comparisons of NS-2 and other simulators can be found in [Begg et al. 2006; Caro 2003; Halvardsson and Lindberg 2004]. A short overview of NS-2 is given in the next Section.

5.3.1 NS-2 Overview

Network Simulator 2 developed at UC Berkeley is an object-oriented, discrete event driven network simulator targeted at networking research including wired and wireless network functions and protocols. In general, NS-2 provides users with a way of specifying such network protocols and simulating their corresponding behaviours. NS-2 has been significantly improved by the open source community and its current release includes wireless network support provided by the CMU Monarch extensions [CMU] which are widely used by many research groups.

The simulator is written in C++ and the script language OTcl (MIT's object extension to Tool command language) [OTcl]. C++ is a powerful programming language which enables fast execution of applications. However, usually, some modifications are needed in order to perform specific simulations; the main simulation structure is retained but modification of parameters is required to execute on tasks such as performance comparisons of routing protocols. This implies that additional time and effort is taken to recompile C++ code every time a modification is needed. OTcl is a language which offers the key advantage that such modifications do not need to be recompiled but its execution capability is slower than the compiled language. NS-2 makes this unification through tclcl, i.e. OTcl linkage. The main objects of a simulation such as nodes and protocols are implemented using C++ and the configuration of parameters such as number or position of nodes, time of the simulation, etc. are implemented in OTcl. Hence the combination of the two languages offers a compromise between performance and ease of use. Implementation and simulation under NS-2 consists of four steps:

- Implement the protocol under investigation by adding a combination of C++ and OTcl code to NS-2's base.
- Describe the simulation through OTcl script.
- Start the NS-2 simulation engine.
- Analyse the generated trace files used to do subsequent data processing (calculating delay, throughput, received rate, etc.).
- Visualisation of the outputs of the simulation are realised with program Network Animator (NAM).

Implementing a new protocol requires adding C++ code to represent the it's functionalities, as well as updating key NS-2 OTcl configuration files so NS-2 can recognise the new protocol and its default parameters. C++ code also describes which parameters and methods are available for OTcl. An overview of how a simulation is executed is shown in Figure 5.4.

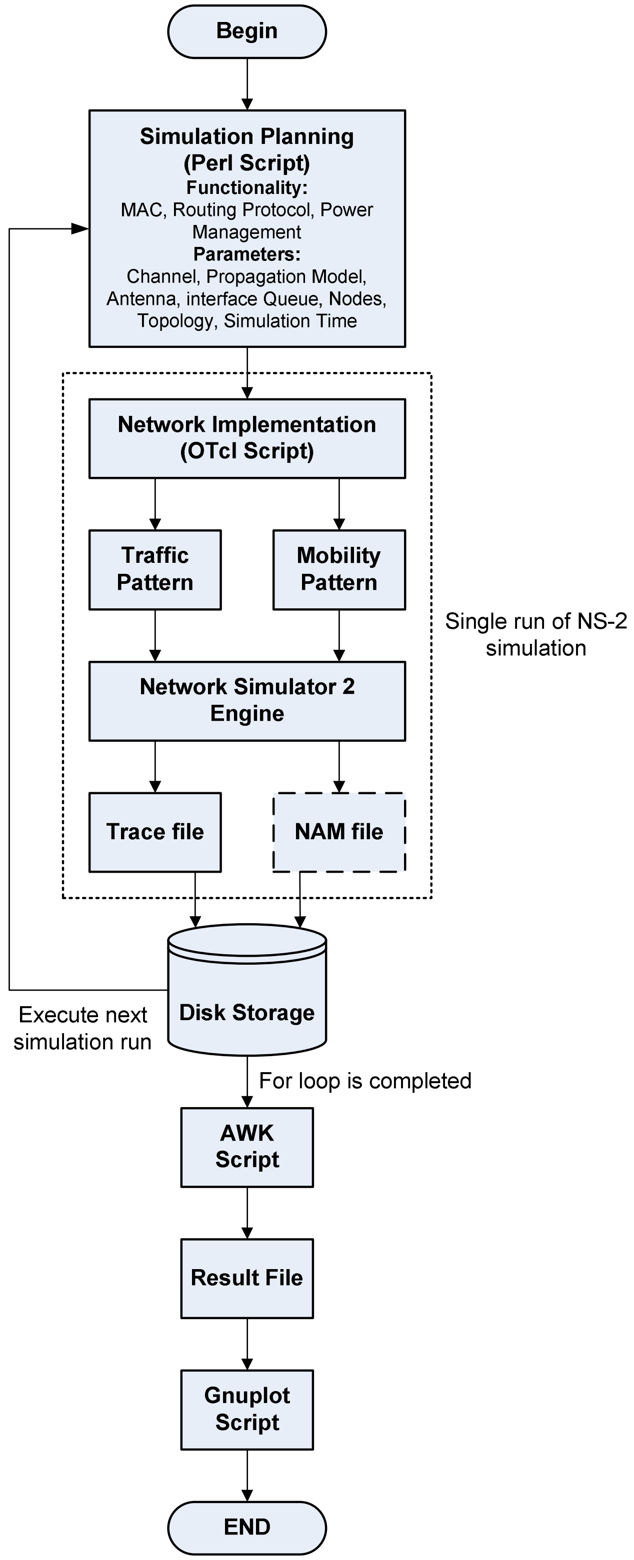

Figure 5.4 NS-2 simulation processes.

In order to eliminate human error and ease the simulation process, a perl script is used to initialise the simulation and process the generated trace files in a predefined order. The perl script is written as a 'for loop' which allows the simulations to be sequentially initialised with the specific network functionalities and parameters are arranged and given in the script. The network implementation is mainly described in OTcl script where construction of the networking scenario and preparation for the simulation according to the imported network functionalities and parameters from the previous perl script are carried out. The detail of the network implementation is discussed in the next Section.

A series of traffic patterns and mobility patterns are generated using embedded OTcl script and compiled C++ script, respectively, (both supplied by CMU) for the use in a variety of simulation runs. The traffic patterns are Continuous Bit Rate (CBR), traffic common in many existing routing protocol evaluations such as AODV and DSR; it is also used in this research and is the basis for the validation and comparison. The packet size, packet rate, source-destination pair and communication paradigm (one-to-one, one-to-many, many-to-one, many-to-many, refer to Section 2.2.2) are varied according to the setting of network to be studied. The simulated mobility patterns are based on Random Waypoint (Section 2.4.1.2), a common and most widely-used mobility model in the literature; it is used in this work for the validation and comparison.

The network field size, the degree of mobility, moving speed and the pause time are also varied according to the network scenario to be studied. Once the simulated network is implemented, both the appropriate traffic pattern and mobility pattern are uploaded. Output trace files are generated and stored to the disk when each simulation run is completed. After all runs are completed, AWK script [Aho et al. 1986] is used to analysis the generated traces and saved into a results file. Finally, Gnuplot script [Gnuplot] is used to produce the results graphs based on the post-analysed data from the results file. Note that NAM files are only generated and examined for validation and for debugging purposes.

5.3.2 Network Implementation

For a fair comparison the proposed implicit routing schemes VGP and VZRP and the AODV routing protocol are evaluated under the same networking scenario. The reasons for choosing AODV are as follows:

- AODV is widely accepted as a benchmark push-based scheme and standard. In a range of multiple mobile sinks scenarios, push-based routing is typically more efficient than pull-based approaches [Liu et al. 2007b].
- AODV is designed particularly to facilitate wireless communication in a range of mobile scenarios. Its performance has been proved to be reliable against mobility. The study in [Broch et al. 1998] shows that AODV delivers over 95% of the packets regardless of mobility rate.
- Validated implementations of AODV are available with NS-2 [Broch et al. 1998; Perkins et al. 2001].

- Xu et al. [2001] have studied the performance of GAF over unmodified AODV and have proved that the integration of GAF with AODV is more efficient than AODV alone.

Figure 5.5 shows an overview of network implementation which incorporates GAF supported mobile node model for the study of varied routing protocols. Below are the details of modules that make up the GAF supported mobile node model.

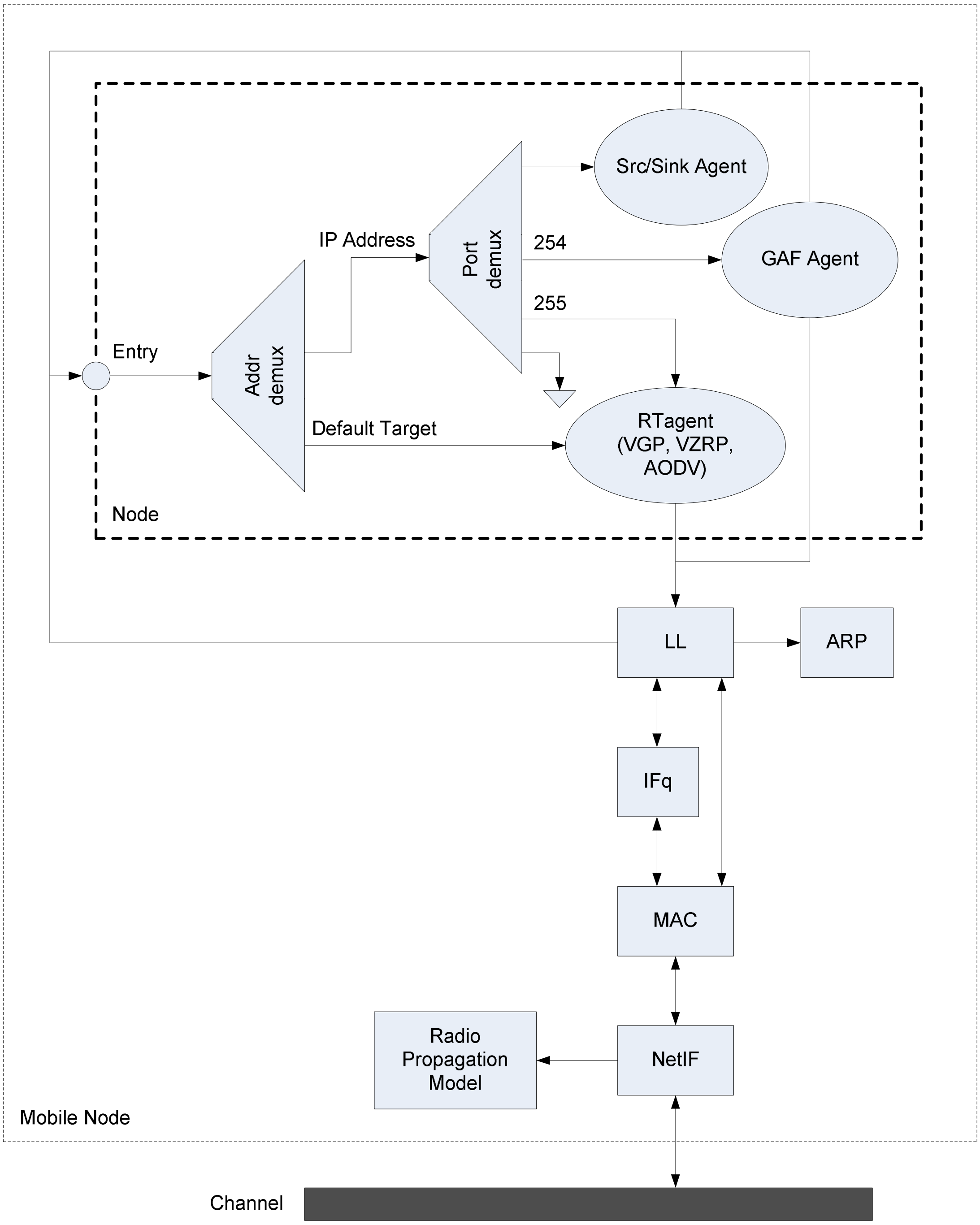

Figure 5.5 Mobile node model.

- *Node class:* the node itself is a standalone class in OTcl. In the case of an unicast node, the node class consists of two objects: an address classifier (addr demux) and a port classifier (port demux). The function of those classifiers is to distribute incoming packets to the correct agent or outgoing link (lower layer).

- *Src/Sink agent:* packets sent by the source (Src) are handled by the node's entry point which passes them to the routing agent (RTagent) through the classifier and sent to the immediate lower layers. The sink agent will only receive packets through the classifier, if the packet is addressed to it.
- *GAF agent:* by default, the GAF agent is added to port 254. This agent is used to control network topology, powering the radio off as much as possible while maintaining a constant level of routing fidelity (refer to Section 4.2.3).
- *RTagent:* this agent is used to perform the routing functionality. A port number of 255 is used to attach the routing agent to a mobile node. The mobile node also uses a default target to point to the routing agent. In the event the incoming packet is not for the mobile node itself, the packet is handed to the routing agent which assigns the routing information and sends it down to the link layer (LL).
- *LL class and ARP module:* the link layer class is responsible for simulating data link protocols. The LL class uses an Address Resolution Protocol (ARP) to determine the hardware address of the neighbouring nodes and map IP addresses to their correct interfaces.
- *IFq class:* an Interface queue (IFq) class 'CMUPriQueue' is used to queue all outgoing packets prior to transmission governed by the MAC. By default, the IFq has a maximum size of 50 packets, maintaining a queue with two priorities each served in First In First Out (FIFO) order. GAF packets are assigned a higher priority than routing packets and data packets.

- *MAC module:* this module is an implementation of the IEEE 802.11 standard [IEEE Standard 2007], the protocol most commonly used by the WSNs modelling community, and an appropriate candidate protocol to be used in this analysis.
- *Radio propagation model:* the two-ray ground reflection model [Glover and Grant 1998] and a unity gain omni-directional antenna is used in order to make the results obtained from this study comparable with existing work. Both the transmitter and receiver antennas are placed at a height of 1.5m above the ground. The antenna gain is set to 0dBm.
- *NetIF model:* the wireless network interface (NetIF) uses the characteristics of the 914MHz Lucent WaveLAN [Tuch 1993] or 916MHz CC1000 radio chipset [CC1000] for the purpose of model validation or network evaluation, respectively. WaveLAN is a shared media radio with a nominal bit rate of 2Mbps and nominal radio range of 250m. CC1000 is a commonly used radio hardware platform for WSNs with a nominal bit rate of 76.8Kbps and nominal radio range of 157m. The detail configuration parameters for the wireless interface can be found in [Tuch 1993] and [CC1000].

5.3.3 Model Validation

For the simulation results to be credible, it is essential and critical that the simulation framework undergoes validation, especially for AODV which serves as a reference with which to compare new routing protocols. Therefore, this section presents results of a preliminary simulation-based analysis of performance in order to ensure that the implementation of AODV – together with the GAF, integral to the functionalities of the new protocols - under NS-2 is faithful to the scheme's specifications. The implementation of each is validated by comparing the generated results with that of a known baseline [Perkins et al. 2001; Xu et al. 2001].

5.3.3.1 AODV Validation

The validation of AODV is carried out with the same simulation parameters as stated in [Perkins et al. 2001]. The model uses characteristics similar to the radio interface WaveLAN with nominal bit rate of 2 Mbps and radio range of 250m. The movement pattern of mobile nodes is governed by Random Waypoint with a speed between 0m/s-20m/s and a pause time which varied between 0, 30, 60, 120, 300 and 500s. During each simulation, 100 mobile nodes move in an area of 2200mx600m for 500s of simulated time. The applied traffic sources are continuous bit rate (CBR), commonly used for the protocol evaluation. Source-destination pairs are spread randomly over the network; the packet size is 512bytes; and the packet rate is set to 4 packets per second.

Figure 5.6 and Figure 5.7 show the packet delivery ratio (the overall percentage of data packets successfully delivered using AODV) and normalised routing load (the number of routing packets transmitted per data packet delivered at the destination) as the pause time is varied. Each graph represents the average of 10 simulation runs with different mobility patterns. For all cases, the results obtained from simulations (represented by the line with cross symbol) are close to the results reported in [Perkins et al. 2001] (represented by bar) as both simulations are conducted using the same simulator and network settings. Specifically, AODV delivers almost all data packets (over 95%) successfully regardless of pause time. The packet delivery fraction approaches 100% when the pause time is set to 500s (representative of a stationary network); similarly, the normalised routing overhead is very small when the pause time is set to 500s. As expected, the normalised routing overhead increases significantly when the pause time decreases as more routing packets are generated to maintain the routing path. The slightly difference in packet delivery fraction and normalised routing overhead is due to the dynamic nature of the network.

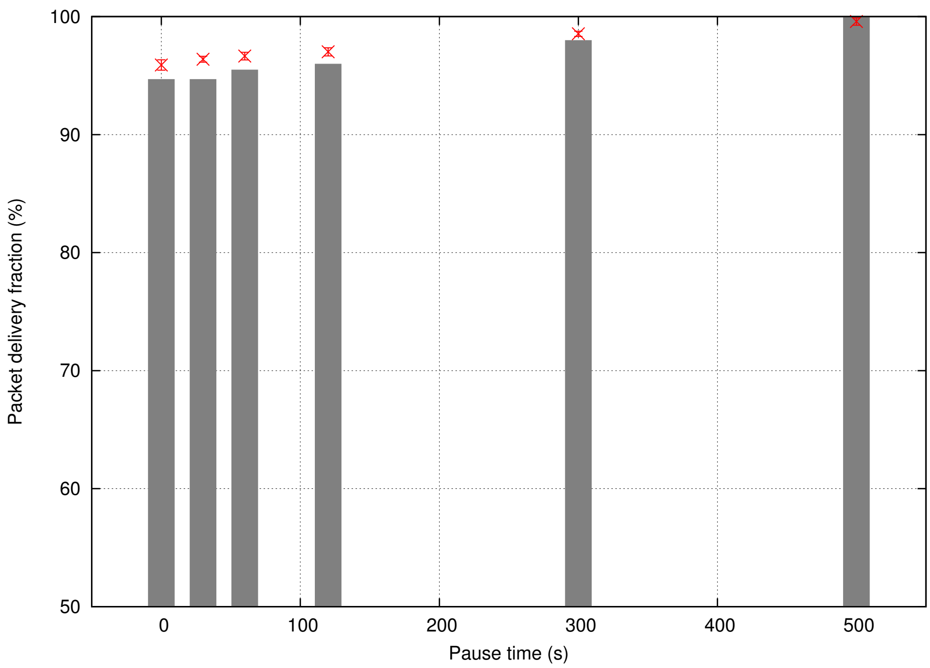

Figure 5.6 Packet delivery fraction as a function of pause time.

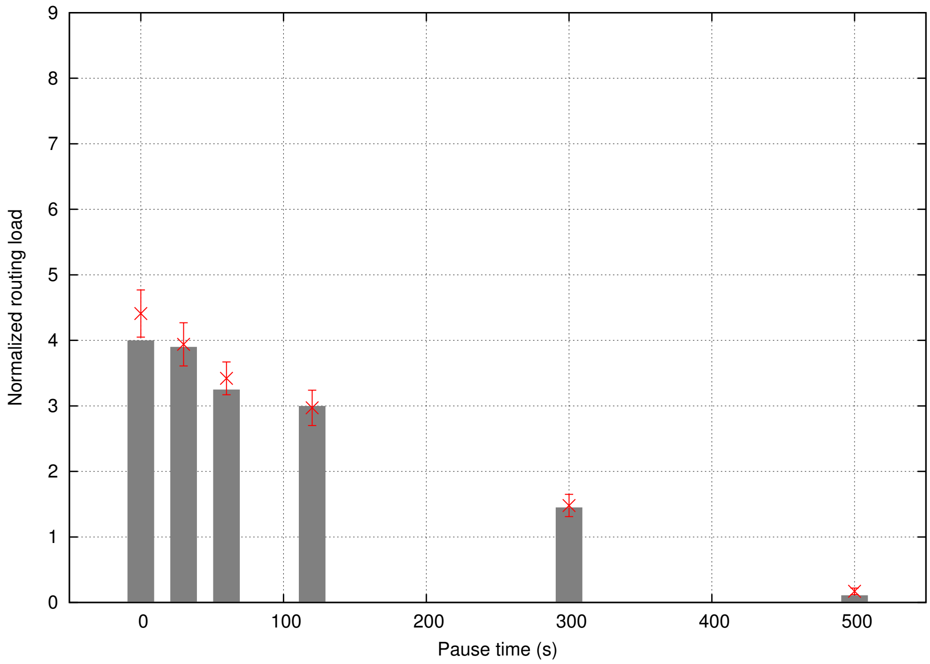

Figure 5.7 Normalised routing overhead as a function of pause time.

5.3.3.2 GAF Validation

GAF model validation is carried out with the same simulation parameters as stated in [Xu et al. 2001] in terms of energy dissipation and packet delivery quality when GAF integrates with AODV protocol. The simulation environment consists of 110 nodes that include 10 CBR traffic nodes (acting as sources and sinks) with unlimited energy and 100 transit nodes (acting as intermediated nodes and forwarding traffic using AODV) with limited energy; the transit node can only listen for ~450 simulated seconds. The movement pattern of the all 110 nodes is governed by Random Waypoint with a speed between 0m/s-20m/s and a fixed pause time of 0s (constant moving) within an area of 1500mx300m for 3600s of simulated time. The packet size is 512bytes and the packet rate is set to 1 packet per second. Without GAF, no energy balancing occurs between all 100 transit nodes and the network dies off after 450s due to constant 'listening'; consequently the packet ratio decreases dramatically. The purpose of this simulation is to study how GAF extends network lifetime.

Figure 5.8 and Figure 5.9 show the packet delivery ratio and fraction of survived nodes over the network lifetime. Each graph represents the average of 7 simulation runs at different mobility patterns. The results obtained from simulations (represented by the line with cross symbols) are close to the results reported in [Xu et al. 2001]. Specifically, the network lifetime is arbitrarily defined as equivalent to the time when there is a 20% decrease of the monitored packet delivery ratio. In Figure 5.8, it is clear that GAF is able to extend network lifetime to 1400 seconds, about 3 times compared to that provided by unmodified AODV (without GAF). Figure 5.9 shows the manner the transmit node survives in the extended network lifetime. At a time of 1400s, the number of nodes that have survived is less than 10% which explains why the packet delivery ratio is compromised since such a small fraction of nodes can no longer maintain a fair routing fidelity for packet delivery. Again, the slightly difference in results can be attributed to the dynamic nature of the network.

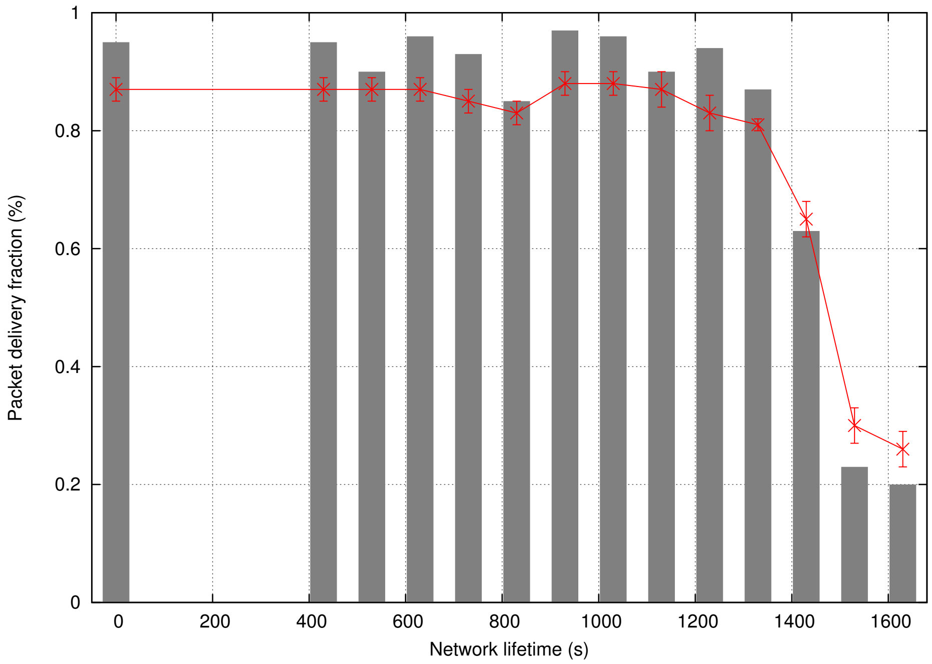

Figure 5.8 Packet delivery ratio over network lifetime.

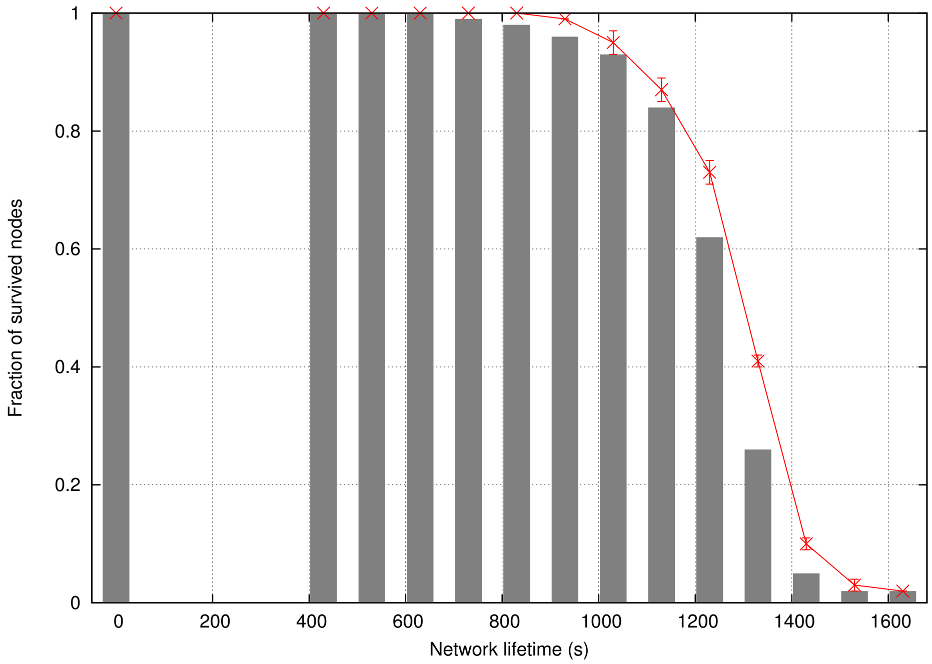

Figure 5.9 Fraction of survived nodes over network lifetime.

5.3.4 Network Settings and Configurations

In order to evaluate the performance of the proposed VGP and VZRP protocols in depth, a wide range of different network scenarios are studied. As VGP and VZRP need to support scalability under different communication paradigms (one-to-one, one-to-many, many-to-one and many-to-many, refer to Section 2.2.2 for detail) whilst managing high node mobility, the effect of the following network parameters are investigated: physical network size, node density, number of source nodes, number of mobile sinks and maximum moving speed. Packet size and packet rate are also taken into account to determine the efficiency of the proposed protocols under different traffic loads. In total, more than 1000 simulation runs were carried out; all combinations of 3 protocols (VGP, VZRP and AODV), 3 physical network sizes (700mx700m, 1400mx1400m and 2100mx2100m), 2 transit node densities (1.2 nodes per virtual grid and 2.4 nodes per virtual grid where a virtual grid is 70mx70m in size), 2 source node numbers (1 source node and 4 source nodes), 4 mobile sink numbers (1 mobile sink, 2 mobile sinks, 4 mobile sinks and 8 mobile sinks), 3 maximum speeds (10m/s, 20m/s, 30m/s, all with 0 pause time), 5 packet sizes (8bytes, 16bytes, 32bytes, 64bytes and 128bytes) and 4 packet rates (1pkt/s, 2pkt/s, 4pkt/s and 8pkt/s). The pre-defined (fixed) parameters capturing the hardware performance have been described in Section 5.3.2. Note that unlike the GAF validation that focused on a study of network lifetime, evaluations of the 3 protocols does not consider energy conservation since network lifetime falls outwith the scope of this thesis; it is undoubtedly a potential area for future study.

Sets of simulation are conducted into 6 distinct configurations. Each of them is carried out to compare the performances of the 3 protocols under specific network scenarios for the range of network settings detailed above. In the first configuration, the packet rate is varied to study the protocols' response to changes in traffic loads under one-to-one communication. In the second configuration, the number of mobile sinks is varied to study scalability as well revealing the routing performance for both one-to-one and one-to-many communication paradigms. The third configuration is similar to the second but it also varies the physical network size to further study network scalability for both one-to-one and one-to-many communication paradigms. In the fourth configuration, the number of mobile sinks and physical network size are varied to study the network scalability for both many-to-one and many-to-many communication paradigms. In the fifth configuration, the number of mobile sinks and packet size are varied to study the sensitivity to packet size for both many-to-one and many-to-many communication paradigms. Finally, in the sixth configuration, the pnode density and maximum speed are varied to study network scalability and the ability of the protocols to manage high node mobility when only one-to-one communication paradigm is considered. The details of network settings for the 6 configurations are summarised in Table 5.2.

Table 5.2 Network settings for 6 distinct configurations.

	Config 1	*Config 2*	*Config 3*	*Config 4*	*Config 5*	*Config 6*
Physical network size	$700m^2$	$700m^2$	Varied	Varied	$700m^2$	$1400m^2$
Transit node density	1.2 node/grid	1.2 node/grid	1.2 node/grid	1.2 node/grid	1.2 node/grid	Varied
Source node number	1	1	1	4	4	1

Mobile sink number	1	Varied	Varied	Varied	Varied	1
Maximum speed	10m/s	10m/s	10m/s	10m/s	10m/s	Varied
Packet size	64bytes	64bytes	64bytes	64bytes	Varied	64bytes
Packet rate	Varied	2pkt/s	2pkt/s	2pkt/s	2pkt/s	2pkt/s

5.3.5 Performance Metrics

Four performance metrics are used to evaluate and compare the protocols:

- *Routing Overhead:* the total number of routing packets transmitted during a simulation. This metric represents the cost of a routing scheme.
- *Normalised Routing Overhead (bytes):* the number of transmitted bytes per data byte at the destination. Note that the MAC layer overheads such as RTS, CTS and ACK are taken into account with the AODV protocol for a fair comparison. This metric indicates the efficiency of a routing scheme viz. the degree it maintains functionality in congested or low-bandwidth environments.
- *Packet Delivery Ratio:* the ratio of the number of data packets delivered to destinations to those generated by all source node(s). This metric characterises the robustness of a routing scheme.
- *Average Delay:* a measure of time between when a data packet is generated to when it reaches its destination. This metric represents a degree of Quality of Service (QoS).

5.4 Simulation Results

The purpose is to study when and how the proposed protocols VGP and VZRP outperform an existing, well-known protocol such as AODV. In Section 5.4.1, the protocols' performances as a function of packet rate when considering the most basic one-to-one communication paradigm is presented. Section 5.4.2 then presents the protocols performances as a function of mobile sink number. The protocols' performance under the one-to-many communication paradigm, one of the more typical scenarios in WSNs (considering the multiple mobile sinks present in the network) is then explored. Section 5.4.3 presents the protocols' performances as a function of physical network size considering both one-to-one and one-to-many communication paradigms. Their performances for both many-to-one and many-to-many communication paradigms are reported in Section 5.4.4. Section 5.4.5 shows the protocols' performance as a function of packet size considering both many-to-one and many-to-many communication paradigms, also typical scenarios in WSN. Section 5.4.6 shows the protocols performance as a function of maximum speed considering only one-to-one communications which allows the effect of mobility to been studied along without other networking effect such as network congestion and radio interference. Section 5.4.7 discusses these simulation results.

5.4.1 Varying Packet Rate with one-to-one communication paradigm

Figure 5.10 depicts the variation of routing overhead as a function of packet rate for the three protocols. As the packet rate is increased, the routing overhead increases linearly for both VGP and VZRP, but only increases marginally for AODV. AODV shows better scalability than VGP and VZRP at high packet rates due to the fact that its explicit routing path can be reused efficiently. VGP consumes significantly more overhead, especially when the packet rate is high with one-to-one communication. VZRP consumes less overhead than VGP and its performance is comparable with AODV due to the use of the 'Register and Page' strategy.

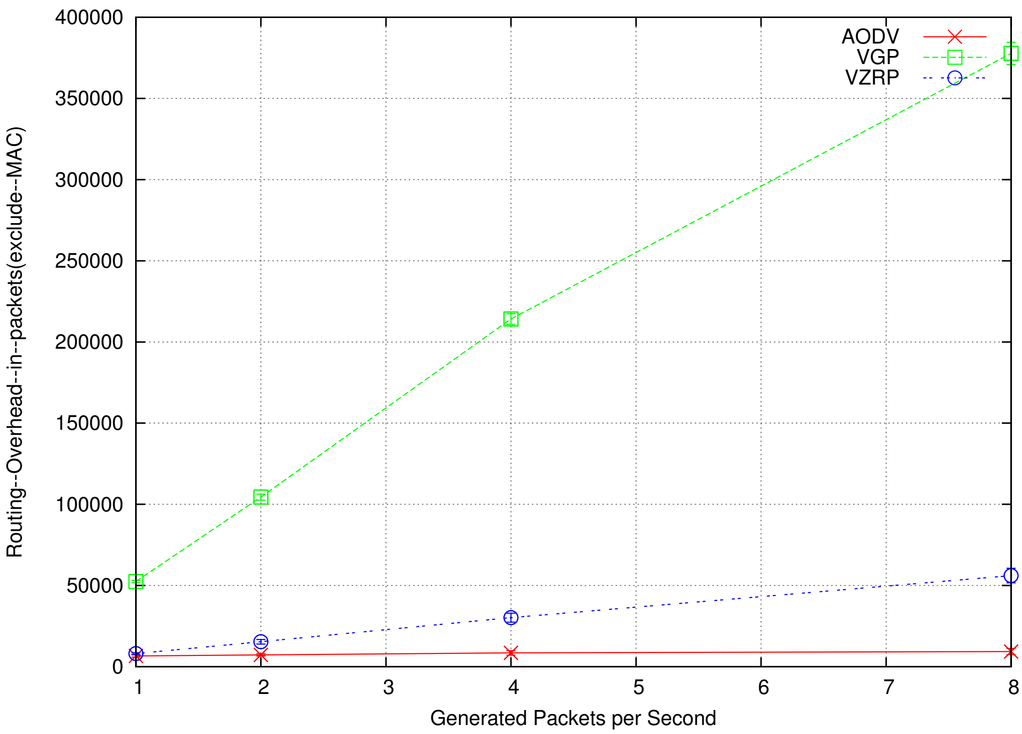

Figure 5.10 Routing overhead as a function of packet rate under one-to-one communication.

Although VGP consumes a significant amount of overhead for packet delivery, it provides the best performance in terms of packet delivery ratio (Figure 5.11). However, the packet delivery ratio of VGP drops dramatically when the packet rate is increased to 8pkt/s. Such degradation in packet delivery ratio is also observed for AODV, in this case attributed to the increased probability of packet collisions and network congestion. This occurs when high traffic loads implicitly propagate throughout the entire network with VGP as well as the high traffic load delivery via an explicit routing path (likely to be congested) with AODV. The result is clear to see;

VZRP is less sensitive to increasing data packet rate.

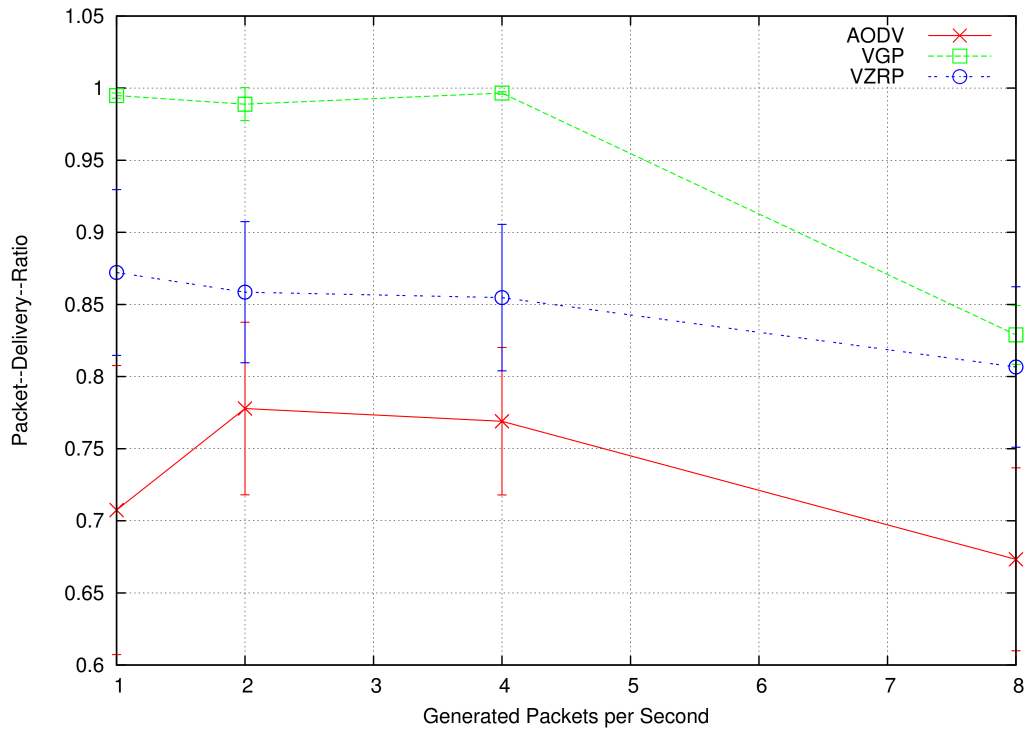

Figure 5.11 Packet delivery ratio as a function of data packet rate under one-to-one communication paradigm.

Figure 5.12 shows the delay of the protocols as a function of packet rate. When the packet rate is small (<=4pkt/s), both VGP and VZRP exhibit significantly shorter delays than AODV, since there is no need for VGP and VZRP to establish a routing path before the packet can be delivered. When the packet rate is as large as 8pkt/s, the delay of AODV continuously decreases as the established connections are re-used frequently. The delay of VGP increases under the traffic load of 8pkt/s which implying that the network has become congested. The delay of VZRP remains constant regardless of the packet rate, since the traffic load is still well below network capacity.

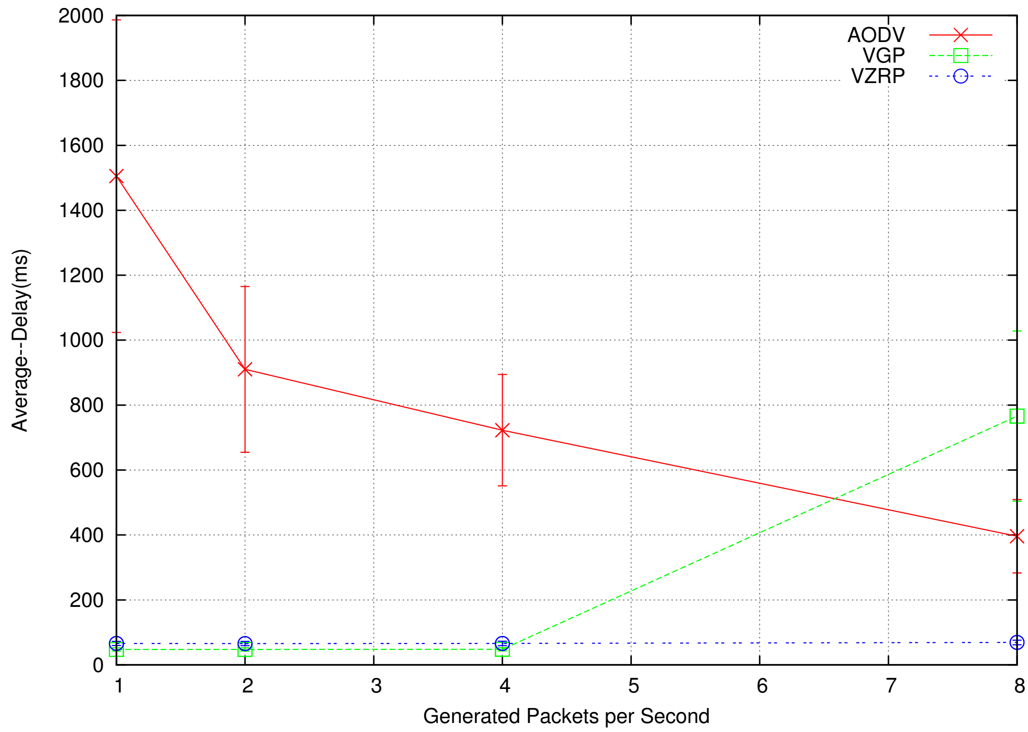

Figure 5.12 Delay as a function of data packet rate under one-to-one communication paradigm.

5.4.2 Varying Number of Mobile Sinks with one-to-one and one-to-many communication paradigms

Figure 5.13 shows the routing overhead of the protocols as a function of number of mobile sinks. For VGP, the constant value overhead implies that it is independent of the number of mobile sinks viz. VGP is capable of supporting the network that contains an unlimited number of mobile sink. However, the cost of packet delivery using VGP is enormous suggesting that it is only viable as a potential solution when a large number of mobile sinks is present in the network. For AODV, the overhead increases linearly with the number of mobile sinks. For VZRP, results show that it has better scalability than AODV as it introduces far less overhead than AODV as the number of mobile sinks increases.

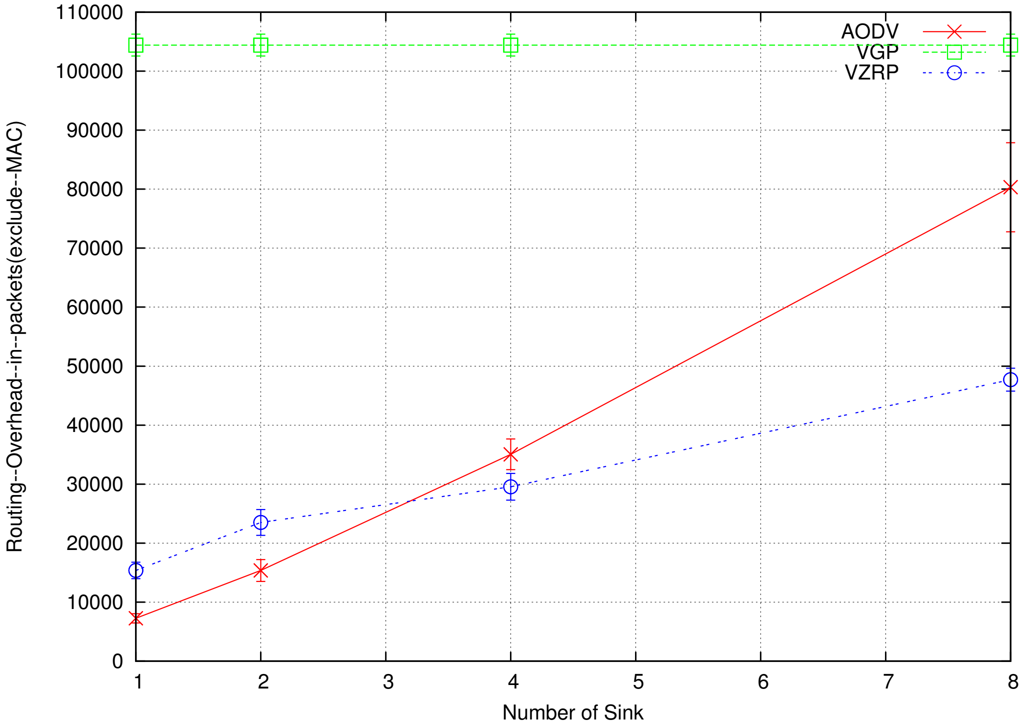

Figure 5.13 Routing overhead as a function of the number of sinks under one-to-many communication paradigm.

Figure 5.14 shows that the packet delivery ratio is independent of the mobile sink number. VZRP outperforms AODV and tends to be more robust than AODV when the number of mobile sinks is increased.

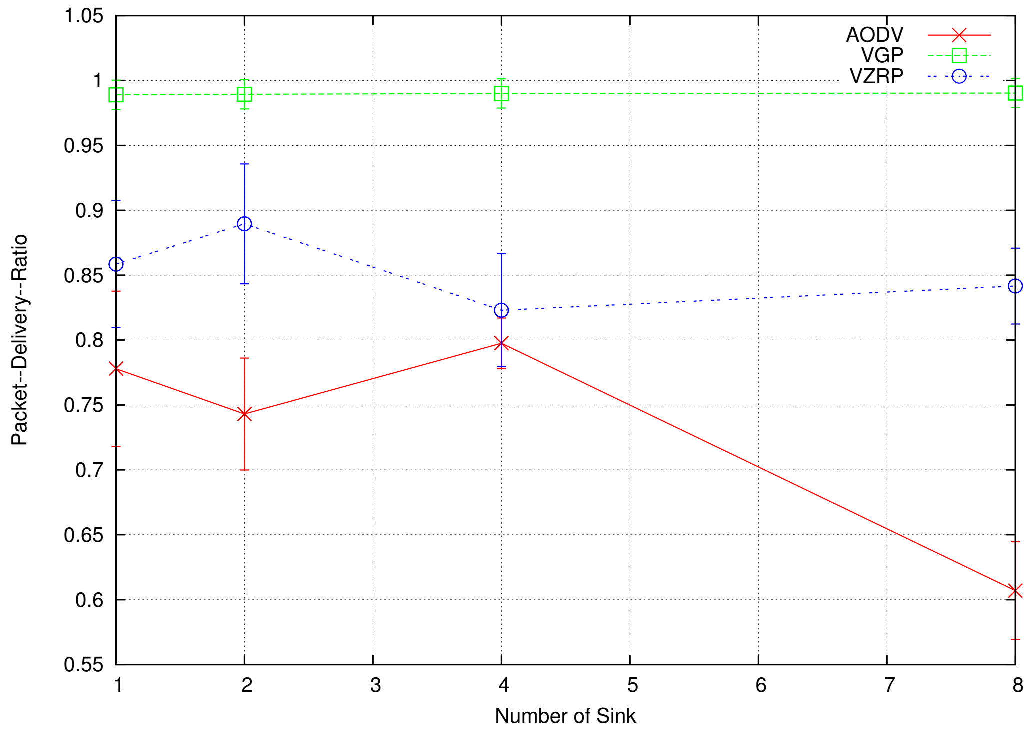

Figure 5.14 Packet delivery ratio as a function of the number of sink under one-to-many communication paradigm.

Figure 5.15 shows that both VGP and VZRP outperform AODV as they can deliver packets in a more straightforward manner in contrast to explicit routing paths which require a lengthy process to establish. Networks under VGP and VZRP show no sign of congestion when 8 mobile sinks are present.

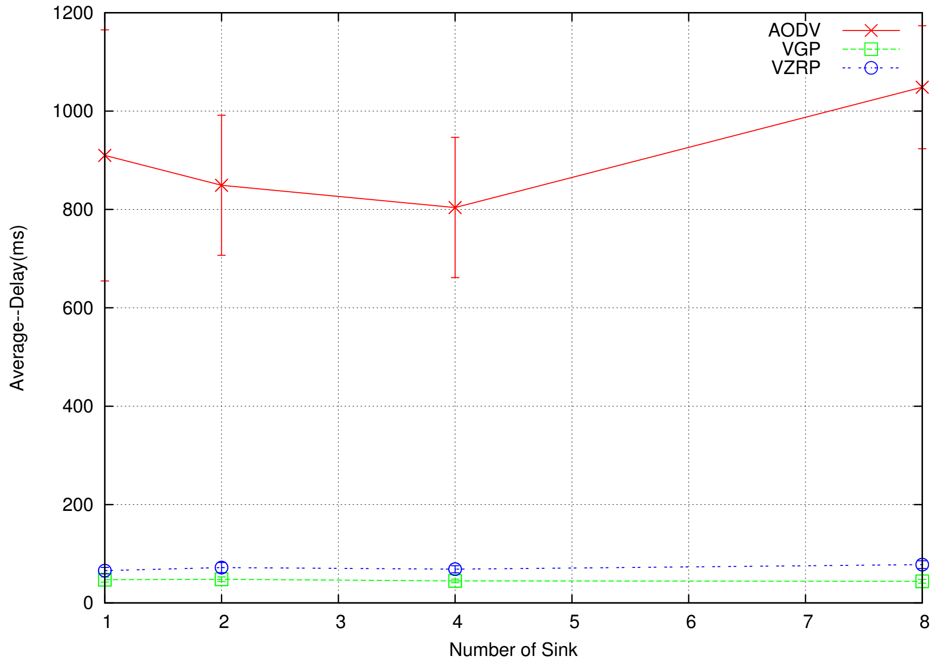

Figure 5.15 Delay as a function of the number of sinks under one-to-many communication paradigm.

5.4.3 Varying the Physical Network Size with one-to-one and one-to-many communication paradigms

Figure 5.16 shows the packet delivery ratio as a function of network size for both one-to-one and one-to-many communications. Firstly, considering one-to-one communications, VGP outperforms VZRP and AODV for all cases with a packet delivery ratio of ~1 even when the physical network size is increased to 2100m x 2100m, as opposed to both the packet delivery ratios of VZRP and AODV which decrease significantly. The latter is as a consequence of the source nodes and mobile sinks tending to be further apart, communicating through a path comprising more intermediated transit nodes and thus resulting in a higher probability of packet delivery failure. VZRP is more reliable than ADOV as the physical network size increases. Except VGP (independent of mobile sink number as discussed earlier), the performance of both VZRP and AODV decreases further for one-to-many communication paradigm. This result implies that both VGP and VZRP have better scalability than AODV and are more suitable for use in WSNs.

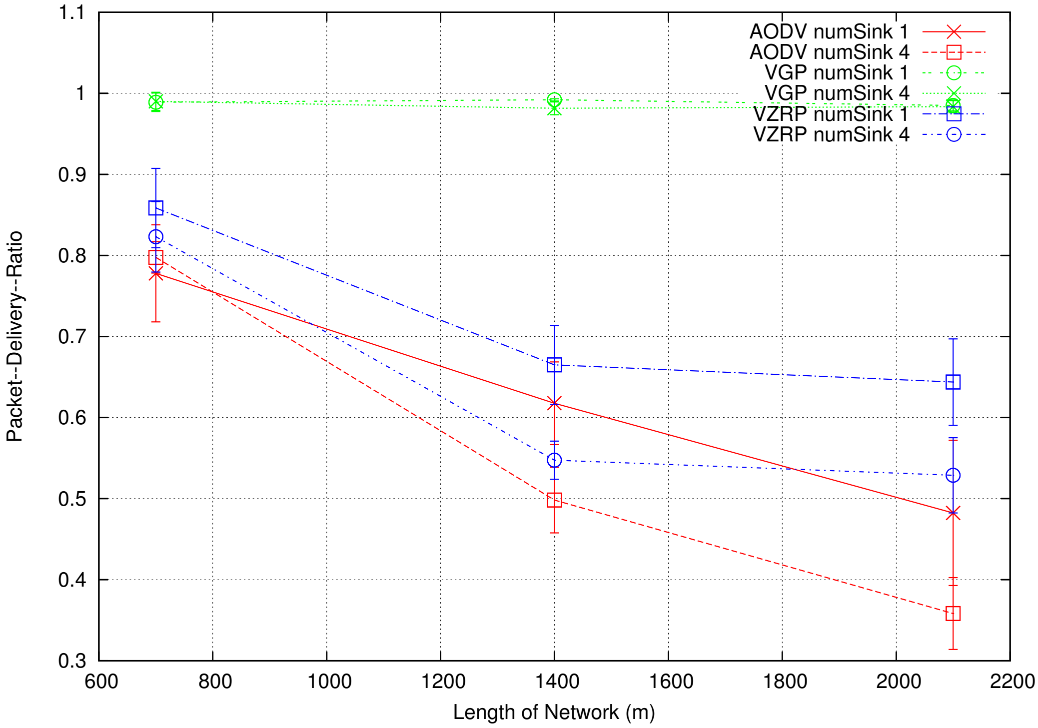

Figure 5.16 Packet delivery ratio as a function of the physical length of a network under one-to-one and one-to-many communication paradigms.

Although VGP has outstanding performance in terms of packet delivery ratio, the efficiency of VGP is shown to be very poor (Figure 5.17) for one-to-one communications. Results also show that VZRP exhibits the best efficiency and scales well with increasing physical network size. Forone-to-many communications, the normalised routing loads of both VGP and VZRP decrease. Noteworthy, is the fact that the normalised routing load of VGP decreases to a quarter as the routing overhead is shared by all four mobile sinks present in the network. Conversely, the normalised routing load of AODV increases for one-to-many communications, since AODV is specifically designed for this paradigm.

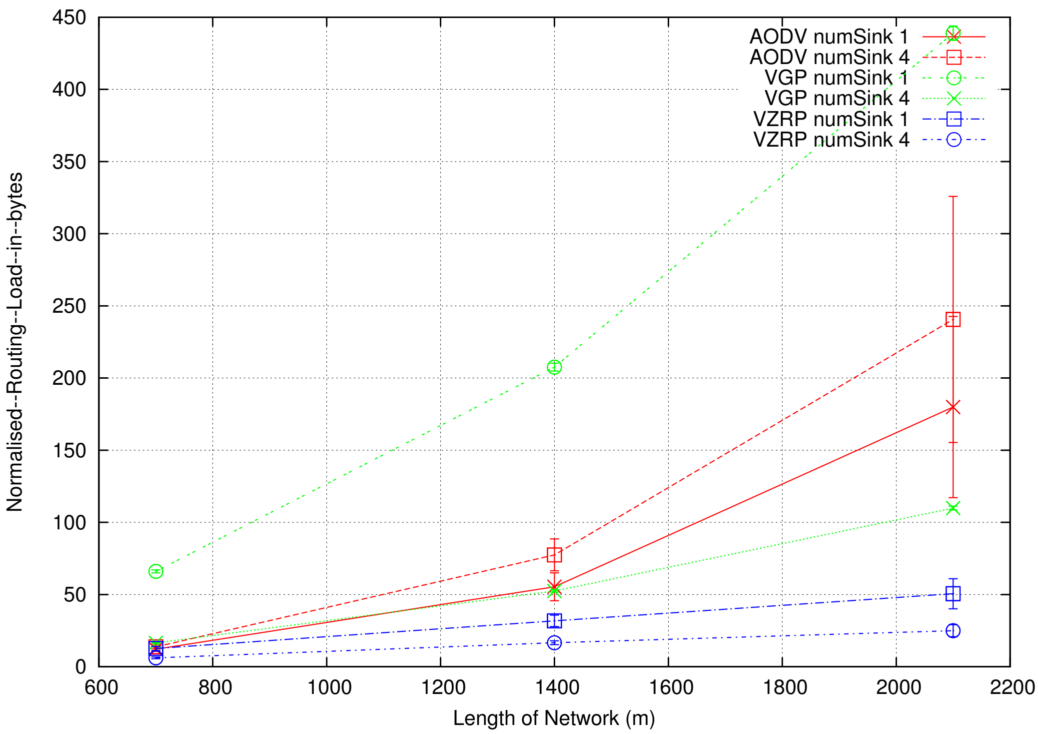

Figure 5.17 Normalised routing load as a function of the physical length of network under one-to-one and one-to-many communication paradigms.

Both VGP and VZRP outperform AODV in terms of average packet delay (Figure 5.18). As expected, the delay for AODV increases significantly as packet delivery may experience additional delays on retransmission as a consequence of packet collisions and broken paths. VGP and VZRP suffer only slight increases in delay with increases in the physical network size.

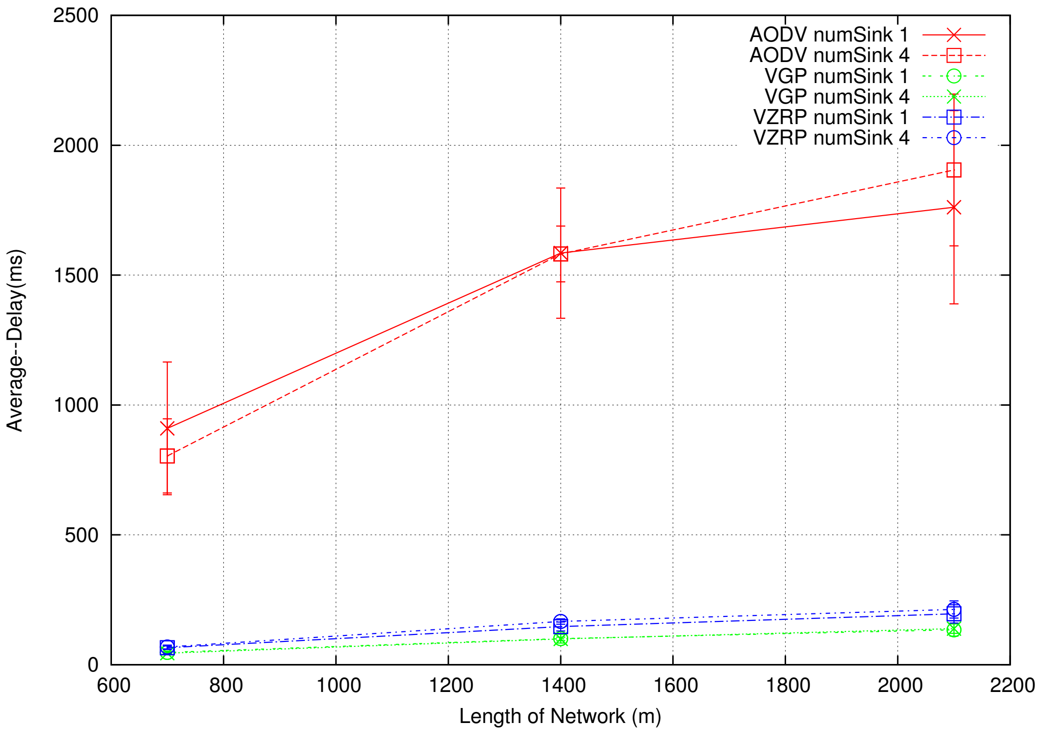

Figure 5.18 Delay as a function of physical length of network under one-to-one and one-to-many communication paradigms.

5.4.4 Varying Physical Network Size with many-to-one and many-to-many communication paradigms

Figure 5.19 shows packet delivery ratio for many-to-one and many-to-many communication paradigms. The trends are similar to the those reported in the last Section; VGP has the best performance in terms of packet delivery ratio, which remains constant regardless of the physical size of network and number of mobile sinks; VZRP exhibits better scalability than AODV for increasing physical network size and adapts well to manage many-to-many communications.

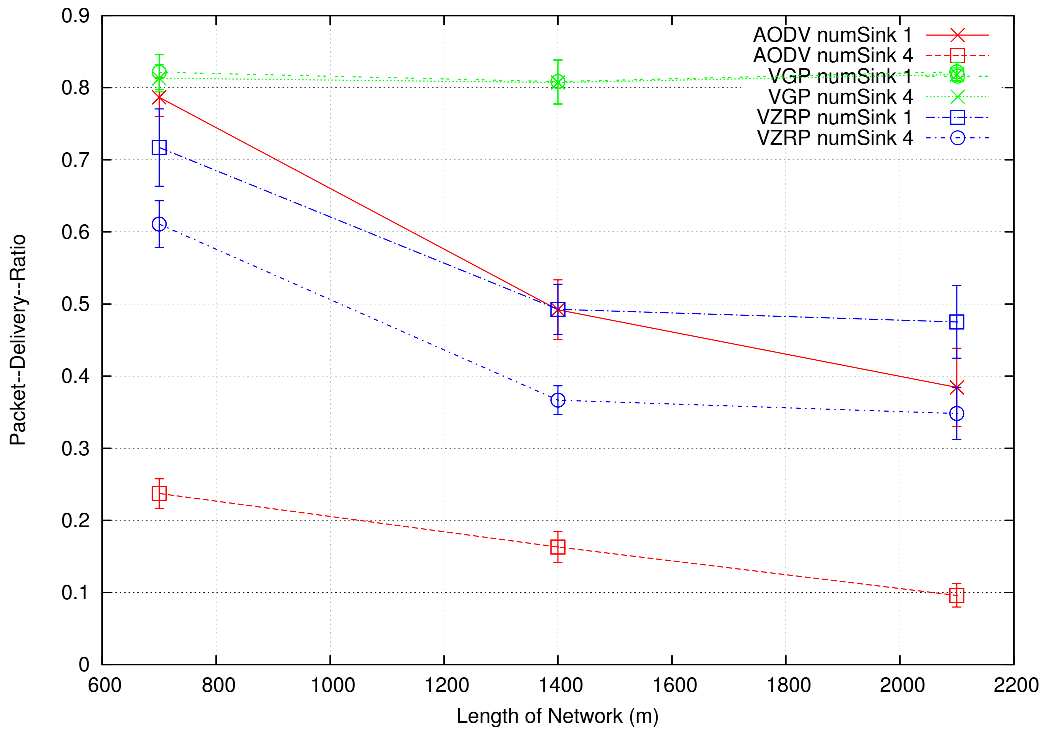

Figure 5.19 Packet delivery ratio as a function of the physical length of network under many-to-one and many-to-many communication paradigms.

The normalised routing load of VGP is significantly higher compared with AODV for one-to-one communications. The results depicted in Figure 5.20 show that the normalised routing loads of VGP and AODV are comparable viz. VGP can achieve a similar level of efficiency as AODV can for many-to-one communications. VZRP provides an extremely high efficiency compared with VGP and AODV. When considering many-to-many communications, both the efficiencies of VGP and VZRP show improvement, but that of AODV degrades. Such a difference is caused by the fact that both the routing overheads of VGP and VZRP can be shared inherently by multiple mobile sinks during packet delivery, but the operation of AODV is performed individually when packets are delivered to multiple mobile sinks resulting in a decreasing efficiency as more contention is expected.

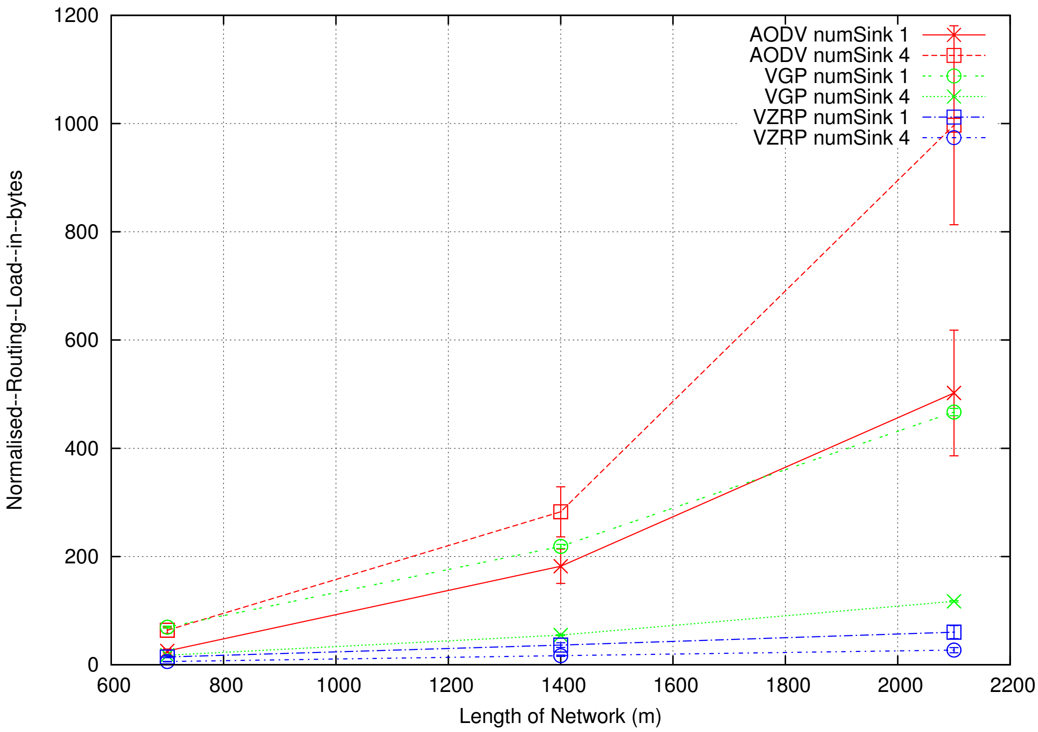

Figure 5.20 Normalised routing load as a function of physical length of network under many-to-one and many-to-many communication paradigms.

Again, both VGP and VZRP outperform AODV in terms of delay. The significant increases in delay of AODV implies more contention under many-to-many communications (Figure 5.20).

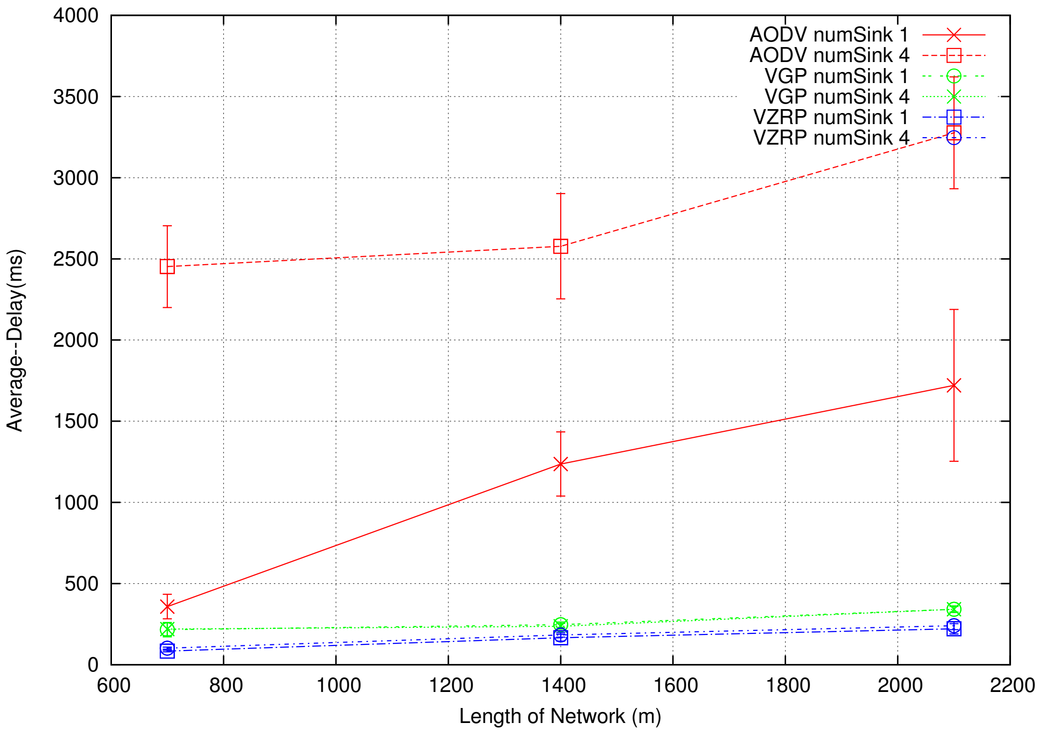

Figure 5.21 Delay as a function of physical length of network under many-to-one and many-to-many communication paradigms.

5.4.5 Varying Packet Size with many-to-one and many-to-many communication paradigms

Figure 5.22 shows the packet delivery ratio as a function of packet size. As the packet size is increased, it effectively emulates a network with increased traffic load. It is clear that the performances of all protocols decreases as the network is more likely to be congested, consequently more packet losses are expected. VGP is shown to be more sensitive to packet size than others as the network becomes highly congested which can be observed from the delay performance of VGP (Figure 5.24). The results also show that VZRP presents better scalability than AODV in terms of packet size.

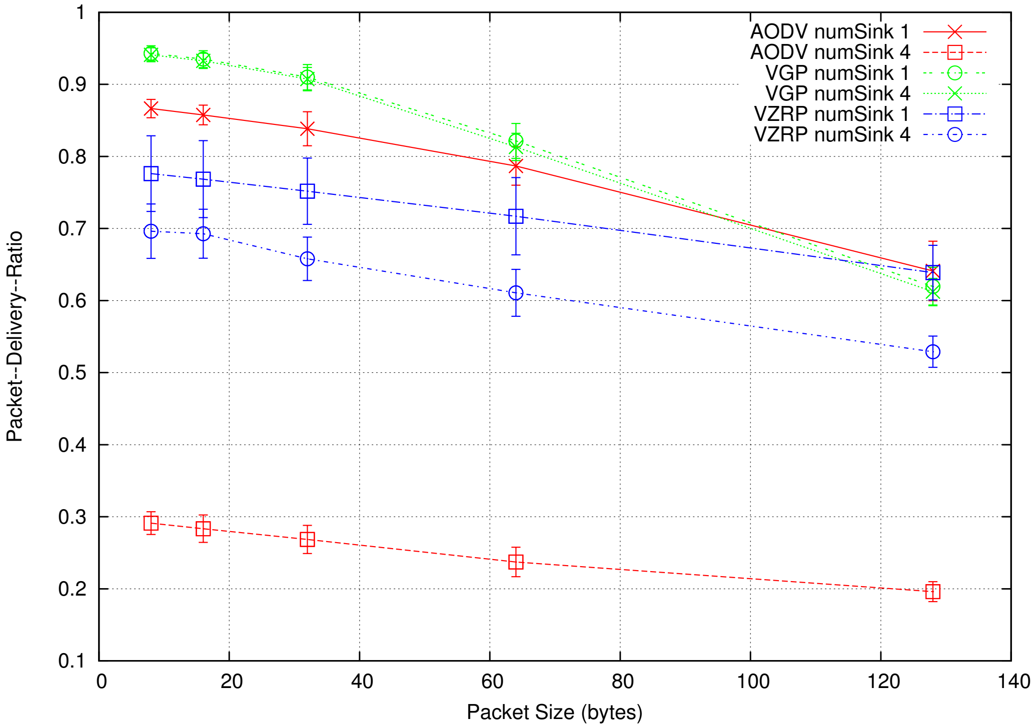

Figure 5.22 Packet delivery ratio as a function of packet size under many-to-one and many-to-many communication paradigms.

Figure 5.23 shows the normalised routing load as a function of packet size; the performance of AODV is highly inefficient when the packet size is small. This is because AODV is swamped with MAC layer overheads (such as RTS, CTS and ACK) which are necessary to facilitate communication over the established routing path. On the other hand, both VGP and VZRP are MAC layer overhead free. This result implies that both VGP and VZRP are more applicable in general WSN applications that typically support smaller packet sizes.

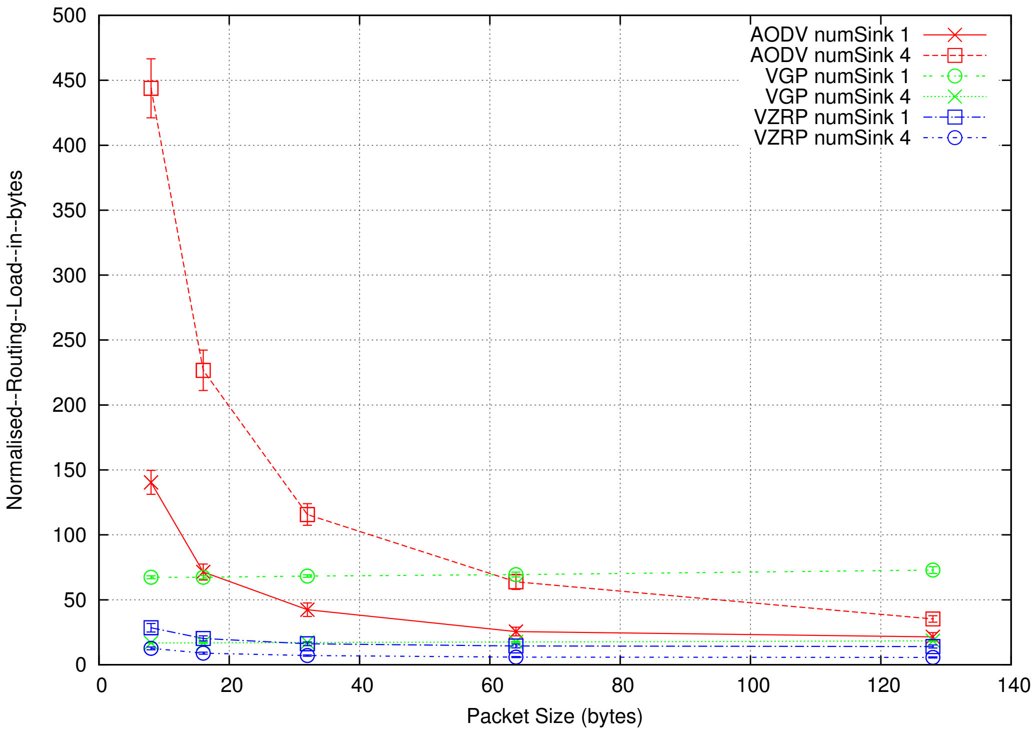

Figure 5.23 Normalised routing load as a function of packet size under many-to-one and many-to-many communication paradigms.

Figure 5.24 shows the delay as a function of packet size. VGP and VZRP have shorter delays than AODV for most cases, except the scenario when the packet size is set to 128 bytes for VGP. Such an exception is caused by the fact that the network becomes highly congested with increasing packet size.

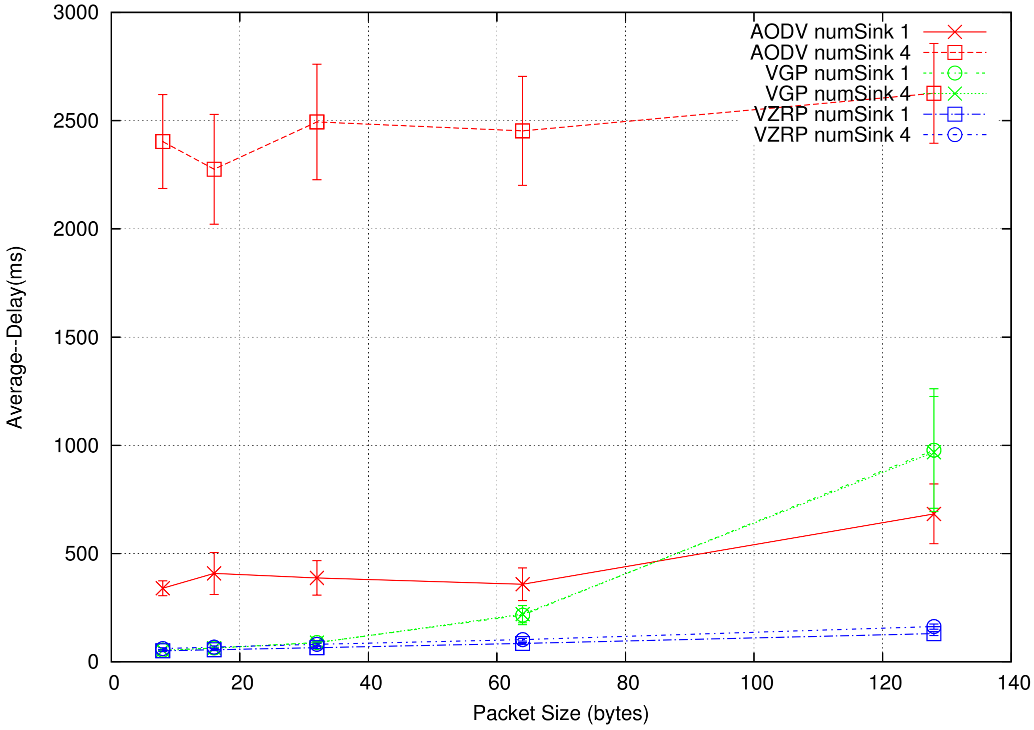

Figure 5.24 Delay as a function of packet size under many-to-one and many-to-many communication paradigms.

5.4.6 Varying Maximum Speed with one-to-one communication paradigm

Figure 5.25 shows the packet delivery ratio as a function of maximum speed for one-to-one communications only. As the speed of nodes increases, the probability of explicit routing path failure is increased resulting in compromised packet delivery ratio for AODV. The packet delivery ratios of VGP and VZRP remain constant even at a maximum speed up to 30 m/s (corresponding to high mobility). Such results corroborate that implicit routing protocols adapt well to treat mobility, outperforming AODV. When number of transit node is increased to 2.4 node/grid, the packet delivery ratios of VGP and VZRP increase as the success of implicit routing is dependent on whether each virtual grid contains an active node for packet forwarding, guaranteed at high transit node density. Conversely the packet delivery ratio of AODV decreases with increasing transit node density caused by the edge effect [Lim et al. 2002].

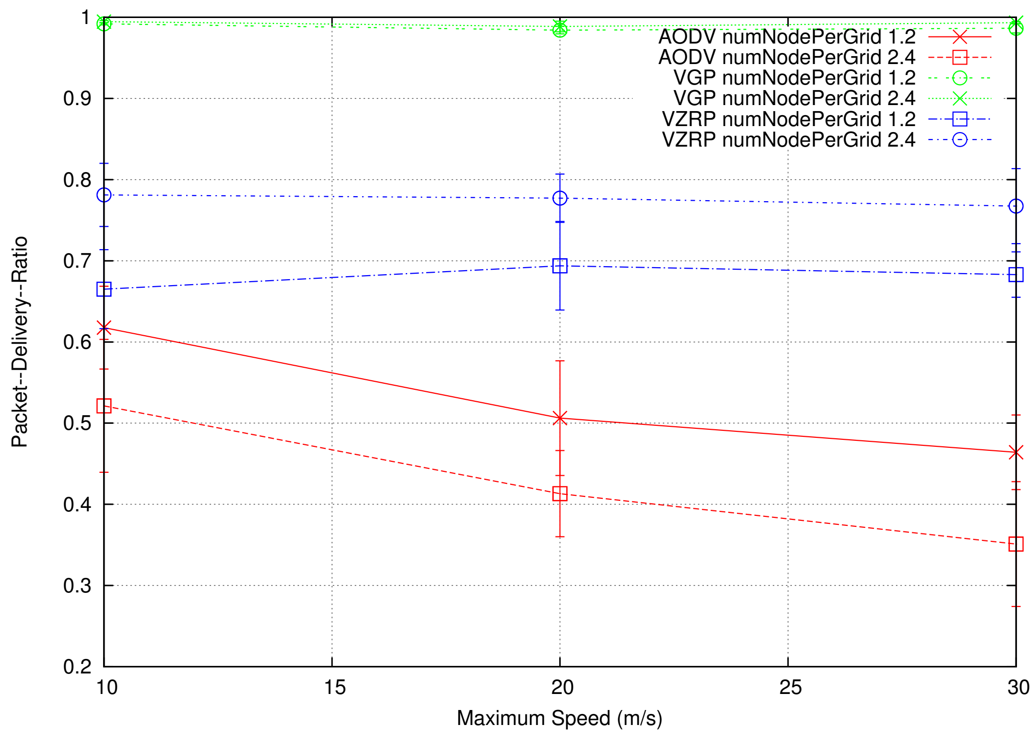

Figure 5.25 Packet delivery ratio as a function of maximum moving speed under one-to-one communication paradigm.

Figure 5.26 shows the normalised routing load as a function of maximum speed. VZRP provides the highest efficiency for all cases compared with AODV and VGP. The efficiency of VGP is affected highly by the over-deployment of transit node in the case of 2.4 node/grid; VGP is not able to provide significant improvement on packet delivery ratio like VZRP, but still consumes additional routing overhead compromising efficiency as the transit node density increases.

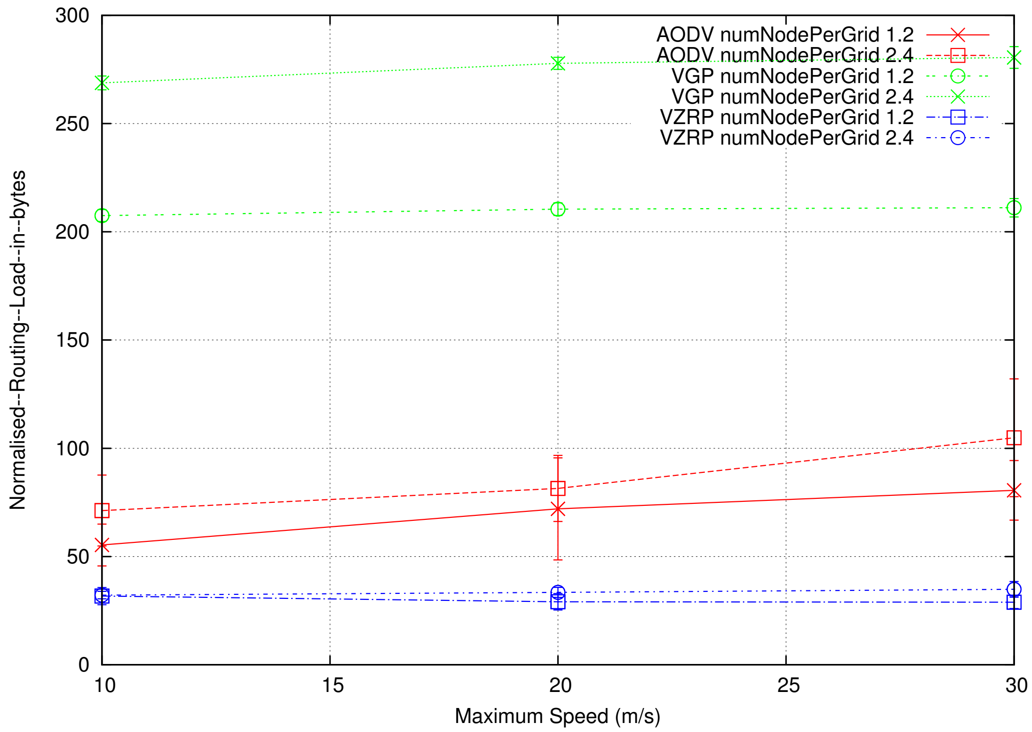

Figure 5.26 Normalised routing load as a function of maximum speed under one-to-one communication paradigm.

Figure 5.27 shows the delay as a function of maximum speed. Only the delay for AODV is affected by increasing maximum speed due to frequent packet retransmissions. Again, both VGP and VZRP outperform AODV in terms of delay with increasing maximum speed.

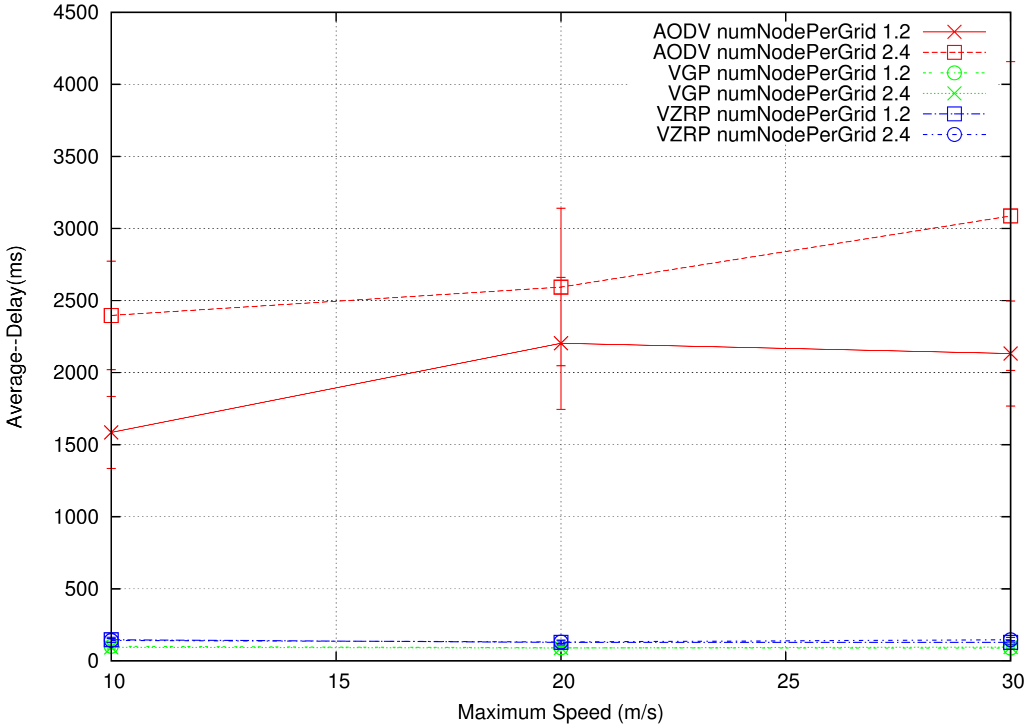

Figure 5.27 Delay as a function of maximum speed under one-to-one communication paradigm.

5.4.7 Discussion

The performances of VGP, VZRP and AODV have been studied systematically through a serious of simulations. In order to provide clear conclusions on the performance comparison, several observations can be made;

- AODV is only suitable for one-to-one communications, not a common scenario in WSNs.
- VGP provides a most robust connection service against mobility and scalability in terms of increasing mobile sink number and physical network

size. However, it also introduces a considerable additional routing overhead resulting in poor efficiency.

- VGP is most suitable when a certain number of mobile sinks are present in the network and/or the packet size is relatively small.
- In most cases, VZRP is the most efficient routing protocol due to the use of the 'Register and Page' and the IRP algorithm.
- VZRP also adapts well to changes in network conditions.
- Both VGP and VZRP have a remarkably low average delay due to the use of implicit routing.

5.5 Conclusions

This chapter has presented the results of a performance comparison of VGP, VZRP and AODV protocols obtained from both mathematical analysis and simulation. Numerical results confirm that VZRP is the most efficient solution for delivering packets from multiple sources to multiple mobile sinks. Rigorous simulation results show that the VGP scheme provides the best reliability (in most cases) on delivering packets. Even at high mobility, large number of entities and large physical network size, the VGP scheme scales well yielding high packet delivery ratios (above 0.9 for most of cases). Moreover, simulation results show that both VGP and VZRP outperform AODV and are suitable for applications which require high packet delivery ratio at minimum delay whilst managing high mobility and scalability.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis has presented the concept of implicit routing and demonstrated its potential on improving the routing performance within Wireless Sensor Networks where the both the source and constituent nodes are mobile. Two novel implicit routing protocols are proposed to gracefully manage node mobility by integrating with a topology control algorithm for large scale mobile WSN implementations.

The dissertation starts with the background of the research area, containing a survey of WSN and defining the goal to be achieved. Unlike many existing studies that focus on the routing solution for static WSN, this work addresses the development and analysis of routings solution that support highly dynamic WSN applications. In addition, network scalability is also been taken into account in the designs of the routing strategies.

In an attempt to provide the founding design principles to manage mobility while enhancing network scalability, the concept underpinning implicit routing strategies has been evaluated through analysis using a series of metrics. Table 6.1 summarises the outcome of an analysis of a range of metrics that are most appropriate to compare the behaviour of explicit and implicit connections as a function of increasing node mobility, node density and network scale. Results show that the implicit connection approach provides an enhanced degree of tolerance (higher reliability and stability) to increasing node mobility and physical network size compared with traditional explicit connection approaches. Results also show that the implicit connection approach adapts well to the increasing node density and overcomes the "edge effect" occurring in explicit connection approach. The former is able to take the full advantage of a fully connected network resulting in a significant improvement in connection reliability and stability.

Table 6.1 The outcome of metrics analysis that facilitates the comparison between explicit and implicit connection approaches.

	Node mobility	*Node density*	*Network scale*
Explicit connection approach	Sensitive	Sensitive	Sensitive
Implicit connection approach	Tolerant	Adaptive	Tolerant

Two novel implicit routing schemes have been presented – referred to as the Virtual Grid Paging (VGP) and Virtual Zone Registration and Paging (VZRP) routing protocols - inspired by concepts from radio paging systems and cellular networks, respectively. The objective of VGP is to efficiently transport data to individual or groups of interested mobile sinks via implicit connections while the communication overhead remains constant regardless of the degree of node mobility and node density. To further improve the efficiency and scalability of the routing protocol, VZRP has been developed based on the concept of the 'Register and Page' used in cellular network. VZRP has been proven to diminish the impact of node mobility and achieve a highly scalable and reliable wireless communication platform with minimum overhead.

A mathematical analysis model is employed to estimate the worst-case communication overhead under simplified network conditions to evaluate the performance of the proposed implicit routing protocols and to facilitate a comparison with a well established explicit routing protocol. Table 6.2 summaries the outcome of analysis depicting worst-case communication overhead trends for each protocol under increasing of mobile sinks number, node mobility and network scale. Results indicate that VGP is the most applicable approach for network scenarios where number of mobile sinks and/or node mobility increase. The estimated worst-case communication overhead of VGP remains constant while for both VZRP and ECA it increases linearly. Additionally, VZRP has lower (only increases linearly) worst-case communication overhead than VGP and ECA with increasing network scale (Section 5.2.3).

Table 6.2 The estimated worst-case communication overhead for each protocol.

	VGP	*VZRP*	*ECA*
Number of mobile sinks	Constant	Linearly increased	Linearly increased

Node mobility	Constant	Linearly increased	Linearly increased
Network scale	Exponentially increased	Linearly increased	Exponentially increased

The performance of the candidate protocols was also evaluated through extensive simulations. The use of simulation allows these protocols to been studied under more realistic network conditions. Studies include analysis of their performance under various network conditions, the scale of the network, the reliability of connections, as well as a comparison of implicit routing protocol with the traditional explicit routing protocol. Table 6.3 summaries the outcome of simulation results depicting the most applicable approach from the perspective of robustness, efficiency and delay for each simulated network configuration. For the purpose of clarity, the upper side of the Table states the varied network parameter and communication paradigm(s) for each network configuration. Results indicate that VGP is a promising approach to achieve enhanced robustness (packet received rate) for all network configurations (in most cases except in configuration 1 and configuration 4 where the results show VZRP has higher adaptability than others). VZRP exhibits the best efficiency compared with others for all network configurations except in configuration 1 for one-to-one communications (Section 5.4.1). VZRP and VGP perform equally well in most cases and always outperform AODV from the perspective of delay. In summary, both analytical and simulation results indicate that implicit routing can be used for a large scalable mobile networks, providing to a reliable yet efficient solution for future WSNs.

Table 6.3 Performance comparison summary.

	Config 1	*Config 2*	*Config 3*	*Config 4*	*Config 5*	*Config 6*
Varied parameter(s)	Packet rate	Mobile sink number	Network size/mobile sink number	Network size/mobile sink number	Mobile sink number/ packet size	Node density/ maximum speed
Communication paradigm(s)	One-to-one	One-to-many	One-to-one/ One-to-many	Many-to-one/many-to-many	Many-to-one/many-to-many	One-to-one
Robustness	VZRP	VGP	VGP	VGP	VZRP	VGP
Efficiency	AODV	VZRP	VZRP	VZRP	VZRP	VZRP
Delay	VZRP	VGP/VZRP	VGP/VZRP	VGP/VZRP	VZRP	VGP/VZRP

6.2 Future Work

Implicit connections offer a tremendous potential in advancing the development of large scale mobile WSNs. While some issues have been considered in this thesis, the following areas should be explored further to determine the feasibility of the proposed routing protocols in terms of implementation:

- *More independent implicit connection approach:* The proposed implicit connection approaches operate in tandem with a geographic location information-based technique such as GAF. Both solutions are too application-dependent for generic WSN implementations, since every node may not always have location information available due to hardware or energy limitations. In order to make the concept of implicit connection more practical in terms of routing, future research should be seeking an alterative method to facilitate implicit connection without the need for any geographic location information.

- *Enabling retransmission mechanism for IDR:* the designed IDR algorithm lacks a retransmission mechanism. Such requirement is critical for some WSN applications such as battlefield surveillance applications, intrusion detection applications, or detection and tracking applications which are much less tolerant to packet losses. Intuitionally, it may be difficult to enable such a feature because no end-to-end acknowledgment (ACK) is used to indicate whether the packet has reached its destination successfully. One possible solution is to employ a transport layer algorithm, usually more flexible and independent of other layers; whether a transport layer implementation is feasible remains as an open issue for study.
- *Network lifetime:* in addition to network performance metrics such as packet delivery ratio and packet delay, energy consumption is one of the primary considerations in WSN system design. It would be valuable to study the relationship between network lifetime and implicit connections.

References

Ababneh and Selvakennedy 2007 N. Ababneh, S. Selvakennedy, "OTC: An Optimized Topology Control Algorithm for Wireless Sensor Networks," PDCAT 2007. Adelaide, Australia. December 2007, pp. 426-434.

Abouei et al. 2011 J. Abouei, K.N. Plataniotis and S. Pasupathy, "Green modulations in energy-constrained wireless sensor networks," Communications, IET , vol.5, no.2, Jan. 21 2011, pp.240-251.

Afanasyev et al. 2009 A. Afanasyev and K. Mayoral and Z. Zhu and S. Y. Oh, "DTCAST: Delay and Disruption Tolerant Multicasting Protocol," Proc. 11th Youth Technological Conference High Technologies and Intellectual Systems, Moscow, Russia, April 2009.

Ahn et al. 2009 S.-W. Ahn; H.-S. Jung; Y.-W. Lee, C. Yoo, "Network Condition Adaptive Real-Time Streaming of an Intelligent Ubiquitous Middleware for U-City," Ubiquitous Information Technologies & Applications, 2009. ICUT '09. Proceedings of the 4th International Conference on, 20-22 Dec. 2009, pp.1-5.

Aho et al. 1986 A.V. Aho, B.W. Kemighan, and P.J. Weinberger, "The AWK Programming Language," Reading, Mass, Addison Wesley, 1986.

Al-Hader et al. 2009 M. Al-Hader, A. Rodzi, A.R. Sharif, N. Ahmad, "Smart City Components Architicture," Computational Intelligence, Modelling and Simulation, 2009. CSSim '09. International Conference on, 7-9 Sept. 2009, pp.93-97.

Ailamaki et al. 2003 A. Ailamaki, C. Faloutsos, P. Fischbeck, M. Small and J. VanBriesen, "An environmental sensor network to determine drinking water quality and security," in SIGMOD Record, vol. 32, no. 4, Dec. 2003.

Akan and Akyildiz 2005 O.B. Akan and I.F. Akyildiz, "Event-to-sink reliable transport in wireless sensor networks," Networking, IEEE/ACM Transactions on, vol.13, no.5, Oct. 2005, pp. 1003- 1016.

Akyildiz et al. 2002 L. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and Erdal Cayirci Georgia Institute of Technology, "A Survey on Sensor Networks," Proc. IEEE Communications Magazine, August 2002, pp. 102–114.

Alam et al. 2008 M.M. Alam, M. Mamun-Or-Rashid and C. S. Hong, "WSNMP: A Network Management Protocol for Wireless Sensor Networks," Proc. Advanced Communication Technology, 2008. 10th International Conference on, vol.1, 17-20 Feb. 2008, pp.742-747.

Albowicz et al. 2001 J. Albowicz, A. Chen, L. Zhang; "Recursive position estimation in sensor networks," Network Protocols, 2001. Ninth International Conference on, 11-14 November 2001, pp. 35- 41.

Altaan 2010 A.S. Altaan, "Effects of Sensor Properties on Power Consumption in Wireless Sensor Network," Computer Research and Development, 2010 Second International Conference on, 7-10 May 2010, pp.335-339.

Bahl and Padmanabhan 2000 P. Bahl and V. N. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System," *Proc. IEEE INFOCOM* 2000, Tel Aviv, Israel, March 2000, pp. 775–784.

Bahl et al. 2004 P. Bahl, R. Chandra, and J. Dunagan, "SSCH: Slotted Seeded Channel Hopping for Capacity Improvement in IEEE 802.11 Ad-Hoc Wireless Networks," Proc. Mobicom, Philadelphia, September 2004.

Bai et al 2004 F. Bai, N. Sadagopan, and A. Helmy, "The IMPORTANT framework for analyzing the Impact of Mobility on Performance Of RouTing protocols for

Adhoc NeTworks," in Wireless Ad Hoc and Sensor Networks, Kluwer Academic Publishers, 2004.

Bansal et al. 1999 D. Bansal, A. Chandra, R. Shorey, A. Kulshreshtha and M. Gupta, M, "Mobility models for cellular systems: cell topography and handoff probability," Vehicular Technology Conference, vol.3, July 1999, pp.1794-1798.

Bar-Noy and Kessler 1993 A. Bar-Noy and I. Kessler, "Tracking mobile users in wireless communications networks," Proc. Twelfth Annual Joint Conference of the IEEE Computer and Communications Societies. Networking: Foundation for the Future. IEEE, 1993, pp.1232-1239.

Begg et al. 2006 L. Begg, W. Lin, K. Pawlikowski, S. Perera, H. Siriisena, "Survey of Simulations of Next Generation Networks for Studying Service Availability and Resilience," Technical Report, University of Canterbury, New Zealand, February 2006.

Bettstetter and Wagner 2002 C. Bettstetter and C. Wagner, "The spatial node distribution of the random waypoint mobility model," *Proc. German Workshop on Mobile Ad Hoc Networks (WMAN)*, Ulm, Germany, March, 2002.

Bianchi 2000 G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," Selected Areas in Communications, IEEE Journal on, vol.18, no.3, Mar 2000, pp.535-547.

Biradar et al. 2009 R. V. Biradar, V. C. Patil, S. Sawant, and R. R. Mudholkar, "Classification and comparison of routing protocols in wireless sensor networks", UbiCC Journal, Vol. 4. 2009, pp. 704-711.

Boleng 2001 J. Boleng, "Normalizing mobility characteristics and enabling adaptive protocols for ad hoc networks," Proc. Local and Metropolitan Area Networks

Workshop (LANMAN), March 2001, pages 9–12.

Bonnet et al. 2000 P. Bonnet, J. Gehrke and P.Seshadri, "Querying the physical world," IEEE Pers. Commun., vol. 7, no. 5, 2000, pp. 10-15.

Braun 2007 J. E. Braun, "Intelligent building systems—Past, present, future," American Control Conference, 9-13 July 2007, pp.4374-4381.

Broch 1998 J. Broch, D. Maltz, D. Johnson, Y. Hu, and J. Jetcheva, "Multi-hop wireless ad hoc network routing protocols," Proc. ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM), Dallas, Texas, USA, October 1998, pages 85-97.

Broch et al. 1998 J. Broch, D.A. Maltz, D.B. Johnson, Y. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad-hoc network routing protocol", Proc. fourth annual ACM/IEEE international conference on Mobile computing and networking, Dallas, TX, October1998, pp. 85-97.

Camarda et al. 1996 P. Camarda, P, G. Schiraldi, and F. Talucci, "Mobility modeling in cellular communication networks," Local Computer Networks, 1996., Proceedings 21st IEEE Conference on, 13-16 Oct 1996, pp.518-525.

Campista et al. 2008 M.E.M. Campista, L.H.M.K. Costa, O.C.M.B. Duarte, "WPR: A Proactive Routing Protocol Tailored to Wireless Mesh Networks", Global Telecommunications Conference, IEEE GLOBECOM 2008, December 2008, pp. 1-5.

Caro 2003 G.A. Di Caro, "Analysis of simulation environment for mobile ad-hoc networks," Technical Report, Dalle Molle Institute for Artificial Intelligence, Manno, Switzerland, December 2003.

Cayirci and Coplu 2007 E. Cayirci and T. Coplu, "Sendrom: Sensor Networks for Disaster Relief Operations Management," Wireless Networks, vol. 13, no. 3, 2007, pp. 409-423.

CC1000 CC1000, Single chip Very Low Power RF Transceiver, http://focus.ti.com/lit/ds/symlink/cc1000.pdf

Chang and Tassiulas 2004 J.-H. Chang and L. Tassiulas, “Maximum lifetime routing in wireless sensor networks,” Networking, IEEE/ACM Transactions on, vol.12, no.4, Aug. 2004, pp. 609- 619.

Chang et al. 2007 S.-H. Chang, M. Merabti and H. M. Mokhtar, “Coordinate Magnetic Routing for Mobile Sinks Wireless Sensor Networks,” Advanced Information Networking and Applications Workshops, 2007, 21st International Conference on, vol.1, 21-23 May 2007, pp.846-851.

Chehri et al. 2008 A. Chehri, P. Fortier, and P.M. Tardif, “Geo-location with wireless sensor networks using non-linear optimization,” Proc. International Journal of Computer Science and Network Security (IJCSNS), January 2008, pp. 145-154.

Chen et al. 2001 B. Chen, K. Jamieson, H. Balkrishnan and R. Morris, “Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless network,” Proc. Mobicom, 2001, pp. 85–96.

Cheng et al. 2004 X. Cheng, A. Thaeler, G. Xue, D. Chen, “TPS: a time-based positioning scheme for outdoor wireless sensor networks,” INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies, vol.4, 7-11 March 2004, pp. 2685- 2696.

Chiang et al. 1997 C.C. Chiang, H.K. Wu, W. Liu, and M. Gerla, “Routing in Clustered Multihop Mobile Wireless Networks with Fading Channel”, Proc. IEEE Singapore International Conference on Networks SICON’97, April 1997, pp. 197-212.

Chintalapudi et al. 2006 K. Chintalapudi, T. Fu, J. Paek, N. Kothari, S. Rangwala, J. Caffrey, R. Govindan, E. Johnson and S. Masri, “Monitoring civil structures with a wireless sensor network,” Internet Computing, IEEE , vol.10, no.2, March-April 2006, pp. 26-34.

Choi et al. 2007 Y. Choi, S. Park, F. Yu and S.-H. Kim, “An Implicit Cluster-Based Overlay Multicast Protocol Exploiting Tree Division for Mobile Ad Hoc Networks,” Proc. ICOIN 07, Estoril, Portugal, January 2007, pp. 3507-3512.

CMU The CMU Monarch Project’s Wireless and Mobility Extensions to NS, August 1998, http://www.monarch.cs.rice.edu/cmu-ns.html

Cormen et al. 1998 T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein, “Introduction to Algorithms,” McGraw-Hill, New York, 1998.

Crossbow 1995 Crossbow Technologies Inc. 1995, http://www.xbow.com

Cugola and Migliavacca 2009 G. Cugola and M. Migliavacca, “A Context and Content-Based Routing Protocol for Mobile Sensor Networks,” Proc. 6th European Conference on wireless sensor networks, Cork, Ireland, February 11-13, 2009.

Das et al. 2000 S.R. Das, C.E. Perkins, E.M. Royer, “Performance comparison of two on-demand routing protocol for ad-hoc networks,” Proc. INFOCOM 2000, TelAviv, Israel, March 2000.

Deb et al. 2003a B. Deb, S. Bhatnagar and B. Nath, “Information Assurance in Sensor Networks,” Proc. WSNA'03, San Diego, California, USA, September, 2003.

Deb et al. 2003b B. Deb, S. Bhatnagar and B. Nath, “ReInForM: Reliable Information Forwarding Using Multiple Paths in Sensor Networks,” Proc. 28th Annual IEEE Int. Conf. on Local Computer Networks, Bonn/Königswinter, Germany, 2003, pp.406-415.

Doherty et al. 2001 L. Doherty, K. S. Pister, and L. E. Gaoui, "Convex position estimation in wireless sensor networks," Proc. IEEE INFOCOM, vol. 3, 2001, pp. 1655-1663.

Di et al. 2010 C. Di, T.-T. Wu, H. G. Goh, B. Stephen, K.-H. Kwong, C. Michie, I. Andonovic, "Exploitation of Wireless Telemetry for Livestock Condition Monitoring," The 17th Wolrd Congress of the International Commission of Agricultural Engineering (CIGR 2010), Québec City, Canada, 13-17 June, 2010.

Doumit and Agrawal 2002 S. S. Doumit and D. P. Agrawal, "Self-Organizing and Energy-Efficient Network of Sensors," MILCOM 2002. Proceedings, vol.2, 7-10 Oct. 2002, pp. 1245-1250.

Dulman et al. 2003 S. Dulman, T. Nieberg, J. Wu and P. Havinga, "Trade-off between traffic overhead and reliability in multipath routing for wireless sensor networks," Proc. Wireless Communications and Networking Conference, New Orleans, Louisiana, USA, March, 2003.

Feeney 1999 L. M. Feeney, "A Taxonomy for Routing Protocols in Mobile Ad Hoc Networks," SICS Technical. Report T99/07, October 1999.

Filipponi et al. 2010 L, Filipponi, A. Vitaletti, G. Landi, V. Memeo, G. Laura, and P. Pucci, "Smart City: An Event Driven Architecture for Monitoring Public Spaces with Heterogeneous Sensors," Sensor Technologies and Applications (SENSORCOMM), 2010 Fourth International Conference on, 18-25 July 2010, pp.281-286.

Förster and Murphy 2010 A. Förster and A. L. Murphy, "A Critical Survey and Guide to Evaluating WSN Routing Protocols," Proc. the CONET 2010 workshop in conjunction with IPSN 2010, Stockholm, Sweden, 2010.

Fujiwara et al. 2004 T. Fujiwara, H. Makie, and T. Watanabe, "A framework for data collection system with sensor networks in disaster circumstances," Int. Workshop on Ad-Hoc Networks, June 2004, pp. 94-98.

Garcia-Luna-Aceves and Madrga 1999 J.J. Garcia-Luna-Aceves and E.L. Madrga, "A multicast routing protocol for ad-hoc networks," Proc. Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), 1999, pp. 784-792.

Gilat 2004 A. Gilat, "MATLAB: An Introduction with Applications," John Wiley & Sons, 2nd Edition, 2004.

Gislaso 2008 D. Gislaso, "Zigbee wireless networking" Elsevier, 2008.

GloMoSim GloMoSim (Global Mobile Information System Simulation Library), http://pcl.cs.ucla.edu/projects/glomosim/

Glover and Grant 1998 I. Glover and P. Grant, "Digital Communications," Prentice Hall, Europe, 1998.

Gnuplot Gnuplot homepage, http://www.gnuplot.info/

Goense et al. 2005 D. Goense, J. Thelen and K. Langendoen, "Wireless sensor networks for precise Phytophthora decision support," Proc. 5th European Conference on Precision Agriculture, Uppsala, Sweden, June 2005.

Goldsmith 2006 A. Goldsmith "Wireless Communications," Cambridge University Press, 2006.

Gomez and Campbell 2007 J. Gomez and A.T. Campbell, "Variable-Range Transmission Power Control in Wireless Ad Hoc Networks," Mobile Computing, IEEE Transactions on, vol.6, no.1, Jan. 2007, pp.87-99.

Gupta and Kumar 2000 P. Gupta and P. R. Kumar, \The capacity of wireless networks", IEEE Trans. on Information Theory, Vol. 46, March 2000, pp. 388-404.

Haas et al. 2002 Z.J. Haas, M.R. Pearlman, and P. Samar, "The Zone Routing

Protocol (ZRP) for Ad Hoc Networks," IETF MANET, Internet Draft, July 2002.

Hajek et al. 2008 B. Hajek, K. Mitzel, and S. Yang, "Paging and Registration in Cellular Networks: Jointly Optimal Policies and an Iterative Algorithm," Information Theory, IEEE Transactions on, vol.54, no.2, 2008, pp.608-622.

Halvardsson and Lindberg 2004 M. Halvardsson and P. Lindberg, "Reliable Group Communication in Military Ad-hoc Network," Technical Report, Vaxjo University, 2004.

Hamma et al. 2006 S. Hamma, E. Cizeron, H. Issaka and J.P. Guédon, "Performance evaluation of reactive and proactive routing protocol in IEEE 802.11 ad hoc network," International Society for Optical Engineering, vol. 6387, 2006.

Hay and Giaccone 2009 D. Hay and P. Giaccone, "Optimal routing and scheduling for deterministic delay tolerant networks," Proc. IEEE WONS, Feb. 2009.

He et al. 2003a T. He, C. Huang, B.M. Blum, J.A. Stankovic, and T.F. Abdelzaher, "Range-free localization schemes for large scale sensor networks," Proc. MOBICOM, 2003, pp.81-95.

He et al. 2003b T. He, J.A. Stankovic, C. Lu and T. Abdelzaher, "SPEED: a stateless protocol for real-time communication in sensor networks," Proc. ICDCS, Rhode Island, USA, May 2003, pp. 46-55.

Heinzelman et al. 1999 W. R. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks," Proc. ACM MobiCom '99, Seattle, WA, 1999, pp. 174–85.

Heinzelman et al. 2004 W. B. Heinzelman, A. L. Murphy, H. S. Carvalho, M. A. Perillo, "Middleware to Support Sensor Network Applications," IEEE Network, vol.18, no.1, Jan/Feb 2004, pp. 6-14.

Huang and Tseng 2003 C.-F. Huang and Y.-C. Tseng, “The coverage problem in a wireless sensor network,” Proc. ACM Int’l Workshop on Wireless Sensor Networks and Applications (WSNA), San Diego, CA, 2003, pp. 115–121.

Huang et al. 2006 X. Huang, H. Zhai, Y. Fang, “Lightweight Robust Routing in Mobile Wireless Sensor Networks,” Military Communications Conference, MILCOM 2006. IEEE, 23-25 Oct. 2006, pp. 1-6.

IEEE Standard 1999 LAN/MAN Standards Committee of the IEEE Computer Society, “IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirement, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,” 1999.

IEEE Standard 2003 LAN/MAN Standards Committee of the IEEE Computer Society, “IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (LR-WPANs),” October 2003.

IEEE Standard 2007 LAN/MAN Standards Committee of the IEEE Computer Society, “IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,” IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999), vol., no., June 12 2007, pp.C1-1184.

Intanagonwiwat et al. 2000 C. Intanagonwiwat, R. Govindan, and D. Estrin, “Directed Diffusion: a Scalable and Robust Communication Paradigm for Sensor Networks", Proc. ACM MobiCom, Boston, MA, 2000, pp. 56-67.

Jaquet et al. 2001 P. Jaquet, P. Muhlethaler, A. Qayyum, “Optimized Link State Routing Protocol,” IETF Draft, 2001. http://www.ietf.org/internetdrafts/draft-ietf-manet-olsr-06.txt

Johansson et al. 1999 P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degemark, “Scenario-based performance analysis of routing protocols for mobile ad-hoc networks,” Proc. Fifth Annual ACM international Conference on Mobile Computing and Networking, Seattle, Washington, August 1999, pp. 195-206.

Johnson and Maltz 1996 D. Johnson and D. Maltz, “Dynamic source routing in ad hoc wireless networks,” In T. Imelinsky and H. Korth, editors, Mobile Computing, Kluwer Academic Publishers, 1996, pages 153-181.

Johnson et al. 2001 D. Johnson, D. Maltz, Y-C. Hu, and J. Jetcheva, “The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks,” Internet Draft, at,Mar. 2001. http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-05.txt

Kappler and Riegel 2004 C. Kappler and G. Riegel, “A Real-World, Simple Wireless Sensor Network for Monitoring Electrical Energy Consumption,” Proc. EWSN 2004, Berlin, Germany, Jan. 2004.

Karl and Willig 2005 H. Karl and A. Willig, “Protocols and Architectures for Wireless Sensor Networks,” John Wiley & Sons, 2005.

Karp and Kung 2000 B. Karp and H. T. Kung. GPSR: Greedy perimeter stateless routing for wireless sensor networks. Proc. the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking, Boston, MA, August 2000.

Khan et al. 2010 A. Khan, A. Abdullah and N. Hasan, “Maximizing Lifetime of

Homogeneous Wireless Sensor Network through Energy Efficient Clustering Method," International journal of Computer Science & Network Security, 3 (6), 2010, pp. 583-594.

Kulkarni et al. 2006 S. Kulkarni, A. Iyer, C. Rosenberg, "An Address-Light, Integrated MAC and Routing Protocol for Wireless Sensor Networks" Networking, IEEE/ACM Transactions on, August 2006, pp. 793-806.

Kurata et al. 2004 N. Kurata, B.F. Spencer Jr. and M. Ruiz-Sandoval, "Building Risk Monitoring Using Wireless Sensor Network," Proc. World Conference on Earthquake Engineering, Vancouver, B.C., Canada, 1-6 August 2004.

Kweon et al. 2009 K. Kweon, H. Ghim, J. Hong and H. Yoon, "Grid-Based Energy-Efficient Routing from Multiple Sources to Multiple Mobile Sinks in Wireless Sensor Networks," Wireless Pervasive Computing, 2009. 4th International Symposium on, 11-13 Feb. 2009, pp.1-5.

Kwong et al. 2011 K. H. Kwong, T. T. Wu, H. G. Goh, K. Sasloglou, B. Stephen, I. Glover, C. Shen, W. Du, C. Michie, I. Andonovic, "Implementation of Herd Management Systems with Wireless Sensor Networks," IET Wireless Sensor Systems, Vol. 1 , Issue 2, pp. 55 – 65, March 2011, DOI: 10.1049/iet-wss.2010.0057

Kyasanur and Vaidya 2006 P. Kyasanur and N.H. Vaidya, "Routing and link-layer protocols for multi-channel multi-interface ad hoc wireless networks," Proc. SIGMOBILE Mobile Computing and Communications Review 10 (1) (2006), pp. 31-43.

Lea-Cox and Ross 2001 J. D. Lea-Cox and D. S. Ross, "Water and Nutrient Management Planning - Monitoring of Irrigation and Fertilization Practices," Combined Proceedings International Plant Propagators' Society, 2001.

Li et al. 2003 S. Li, S. Son and J. Stankovic, "Event Detection Services Using Data Service Middleware in Distributed Sensor Networks," Proc, the 2nd International Workshop on Information Processing in Sensor Networks, 2003.

Li et al. 2006 Y. Li, Z. Wang, Y.Q. Song, "Wireless Sensor Network Design For Wildfire Monitoring," Proc. of The Sixth World Congress on Intelligent Control and Automation, Vol.1, Dallan, 2006, pp. 109-113.

Li et al. 2011 C. Li, H. Zhang, B. Hao and J. Li, "A Survey on Routing Protocols for Large-Scale Wireless Sensor Networks," Sensors 11, no. 4, 2011.

Liang and Haas 1999 B. Liang and Z. Haas, "Predictive distance-based mobility management for PCS networks," Proc. Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), March 1999.

Lim et al. 2002 G Lim, K Shin, S Lee, H Yoon and J S Ma "Link Stability and Route Lifetime in Ad Hoc Wireless Networks," Proc. IEEE International Conference on Parallel Processing Workshops 2002.

Liu and Towaley 2003 B. Liu and D. Towaley, "On the Coverage and Detectability of Large-scale Wireless Sensor Networks," Proc. of the Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks Conference, INRIA Sophia-Antipolis, France, March, 2003.

Liu and Towaley 2004 B. Liu and D. Towsley, "A study of the coverage of large-scale sensor networks," Proc. IEEE International Conference on Mobile Ad-hoc and Sensor Systems, Fort Lauderdale, Florida, USA, October, 2004.

Liu et al. 2004 H.-H. Liu; Wu, J.-L.C.; C.-J. Wang, "A multi-hop implicit routing protocol for sensor networks," Vehicular Technology Conference, 2004. VTC2004-Fall, vol.4, 26-29 Sept. 2004, pp. 2946-2950.

Liu et al. 2007a Y. Liu, C. Li, Y. He, J. Wu and Z. Xiong, "A Perimeter Intrusion Detection System Using Dual-Mode Wireless Sensor Networks," Communications and Networking in China, 2007. Second International Conference on, 22-24 Aug. 2007, pp. 861-865.

Liu et al. 2007b X. Liu, Q. Huang and Y. Zhang, "Balancing Push and Pull for Efficient Information Discovery in Large-Scale Sensor Networks," Mobile Computing, IEEE Transactions on, vol.6, no.3, March 2007, pp. 241-251.

Luo et al. 2003 Haiyun Luo, Fan Ye, Jerry Cheng, Songwu Lu, Lixia Zhang, "TTDD: Two-tier Data Dissemination in Large-scale Sensor Networks," ACM Mobile Networks and Applications (MONET) Journal, 2003.

Madhavapeddy et al. 1995 S. Madhavapeddy, K. Basu, and A. Roberts, "Adaptive paging algorithms for cellular systems," Proc. 45th IEEE Vehicular Technology Conference, vol. 2, 1995, pp. 976–980.

Mainwaring et al. 2002 A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk and J. Anderson, "Wireless sensor networks for habitat monitoring," Proc. of the 1st ACM international workshop on Wireless sensor networks and applications, Atlanta, Georgia, USA, September, 2002.

Malatras et al. 2008 A. Malatras, A. Asgari and T. Bauge, "Web Enabled Wireless Sensor Networks for Facilities Management," Systems Journal, IEEE, vol.2, no.4, Dec. 2008, pp. 500-512.

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

Maroti et al. 2005 M. Maroti, B. Kusy, G. Balogh, P. Volgyesi, A. Nadas, K. Molnar, S. Dora and A. Ledeczi. "Radio interferometric geolocation," Proc. the 3rd international Conference on Embedded Networked Sensor Systems, San Diego, California, USA, Nov. 02-04, 2005, pp. 1-12.

McDonald and Znati 1999 A.B. McDonald and T.F. Znati, "A mobility-based framework for adaptive clustering in wireless ad hoc networks", IEEE Journal on Selected Areas in Communications, vol 17, No. 8, Aug 1999, pp. 1466-1487.

Mittal et al. 2009 S. Mittal and P. Kaur, "PERFORMANCE COMPARISION OF AODV, DSR and ZRP ROUTING PROTOCOLS IN MANET'S," Proc. International Conference on Advances in Computing, Control, and Telecommunication Technologies, Trivandrum, Kerala, India, December 2009.

Molisch 2005 A. F. Molisch, "Wireless Communications," John Wiley & Sons Inc, New York, NY, USA, 2005.

Monks et al. 2001 J.P. Monks, J.-P. Ebert, A. Wolisz, W.W. Hwu, "A study of the energy saving and capacity improvement potential of power control in multi-hop wireless networks," Proc. Local Computer Networks, 2001, pp.550-559.

Mousavi et al. 2007 S. M. Mousavi, H. R. Rabiee, M. Moshref, A. Dabirmoghaddam, "MobiSim : A Framework for Simulation of Mobility Models in Mobile Ad-Hoc Networks," Proc. The 3rd IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (IEEE WiMob 2007), New York, USA, October, 2007.

Murthy and Manoj 2004 C. S. R. Murthy and B. S. Manoj, "Ad Hoc Wireless Networks. Architectures and Protocols," Prentice Hall, PTR, 2004.

NS NS-2 (The Network Simulator 2), http://www.isi.edu/nsnam/ns/

OMNeT++ 2003 OMNeT++ SimulCraft Inc 2003, http://www.omnetpp.org/

OPNET OPNET (Optimized Network Engineering Tool), http://www.opnet.com/

OTcl Object Tcl Extensions, http://otcl-tclcl.sourceforge.net/otcl/

Pakzad et al. 2008 S.N. Pakzad, G.L. Fenves, S. Kim, D.E. Culler, "Design and Implementation of Scalable Wireless Sensor Network for Structural Monitoring," ASCE Journal of Infrastructure Engineering, 14(1), 2008, pp. 89-101.

Papadimitriou et al. 2010 A. Papadimitriou, D. Katsaros and Y. Manolopoulos, "Topology control algorithms for wireless sensor networks: a critical survey," CompSysTech 2010, pp. 1-10.

Parkinson and Spilker 1996 B. W. Parkinson and J.J. Spilker Jr., "Global Positioning System: Theory and Applications, vols. 1 and 2," American Institute of Aeronautics, 370 L'Enfant Promenade, SW, Washington, DC, 1996.

Pei et al. 2000 G. Pei, M. Gerla and T-W. Chen, "Fisheye state routing in mobile ad hoc networks", Proc. ICDCS Workshops, Taipei, Taiwan, April 2000, pp. D71-D78.

Pelusi et al 2006 L. Pelusi, A. Passarella, and M. Conti "Opportunistic networking: data forwarding in disconnected mobile ad hoc networks," Communications Magazine, IEEE, vol.44, no.11, November 2006, pp. 134-141.

Perkins et al. 2001 C.E. Perkins, E.M. Royer, S.R. Das, M.K. Marina, "Performance comparison of two on-demand routing protocols for ad hoc networks," Personal Communications, IEEE, vol.8, no.1, February 2001, pp. 16-28.

Perkins and Bhagwat 1994 C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," Proc. Conference on Communications Architectures, Protocols and Applications, August-September 1994, pp. 234-244.

Perkins and Royer 1999 C. E. Perkins and E. M. Royer, "Ad-hoc On-demand Distance Vector (AODV) routing," in the 2nd IEEE Workshop on Mobile Computing System and Applications, Feb 1999, pp. 90-100.

Pogkas et al. 2005 N. Pogkas, G. Karastergios, C. Antonopoulos, S. Koubias and G. Papadopoulos, "An ad-hoc sensor network for disaster relief operations," 10th IEEE Conference on Emerging Technologies and Factory Automation, vol.2, 19-22 Sept. 2005, pp. 131-139.

QualNet QualNet, Scalable Network Technologies, http://www.scalable-networks.com/

Radenkovic and Wietrzyk 2006 M. Radenkovic and B. Wietrzyk, "Mobile ad hoc networking approach to detecting and querying events related to farm animals," *Proc. ICNS 2006*, Santa Clara, California, USA, 2006, pp. 109-115.

Raju et al. 2010 S.R. Raju, K.Runkana, and J.Mungara, "ZRP versus AODV and DSR: A comprehensive study on ZRP performance," International Journal of Computer Applications 1(12), February 2010, pp. 35-40.

Rappaport 2002 T. S. Rappaport, "Wireless Communications – Principles and Practice," Prentice Hall, Upper Saddle River, NJ, 2002.

Rashid et al. 2008 R.A. Rashid, S. Arifin, M. Rahim, M.A. Sarijari and N.H. Mahalin, "Home healthcare via wireless biomedical sensor network," RF and Microwave Conference, 2-4 Dec. 2008, pp. 511-514.

Roads 1997 U.S. Roads, "Study Compares Older and Younger Pedestrian Walking Speeds," Road Engineering Journal, 1997.

Romer 2004 K. Romer, "Tracking real-world phenomena with smart dust," Proc. 1st European Workshop on Wireless Sensor Networks, Berlin, Germany, January 2004, pp. 28-43.

Romer and Mattern 2004 K. Romer, and F. Mattern, "The Design Space of Wireless Sensor Networks," IEEE Wireless Communications, December 2004, pp. 54-61.

Royer et al. 2001 E. Royer, P.M. Melliar-Smith, and L. Moser, "An analysis of the optimum node density for ad hoc mobile networks," Proc. IEEE International Conference on Communications (ICC), 2001.

Rubin and Choi 1997 I. Rubin and C. Choi, "Impact of the location area structure on the performance of signaling channels in wireless cellular networks," IEEE Communications Magazine 1997, pp. 108-115.

Saha and Matsumoto 2007 S. Saha and M. Matsumoto, "A WSN Protocol for Disaster Management," Proc. of IEEE IDC, Feb 2007, pp. 209-213.

Shen et al. 2001 C. Shen, C. Srisathapornphat and C. Jaikaeo, "Sensor Information Networking Architecture and Applications," IEEE Personal Communications, August 2001.

Sikka et al. 2006 P. Sikka, P. Corke, P. Valencia, C. Crossman, D. Swain, and G. Bishop-Hurley, "Wireless ad hoc sensor and actuator networks on the farm," *Proc IPSN 2006*, Nashville, TN, USA, April, 2006, pp. 492-499.

Singh and Raghavendra 1998 S. Singh and C.S. Raghavendra, "PAMAS: Power aware multi-access protocol with signalling for ad hoc networks," ACM Computer Communication Review, vol. 28, no. 3, July 1998, pp. 5-26.

Singh et al. 2010 S. K. Singh, M.P. Singh, and D.K. Singh, "Routing Protocols in Wireless Sensor Networks – A Survey," International Journal of Computer Science and Engineering Survey (IJCSES), vol. 1, no. 2, Nov. 2010, pp. 63-83.

Sridharan and Krishnamachari 2007 A. Sridharan and B. Krishnamachari, "Maximizing Network Utilization with Max-Min Fairness in Wireless Sensor Networks," Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops, 2007. 5th International Symposium on, 16-20 April 2007, pp. 1-9.

Stemm and Katz 1997 M. Stemm and R. H. Katz, "Measuring and reducing energy consumption of network interfaces in hand-held devices," IEICE Transactions on Communications, E80-B(8), Aug. 1997, pp. 1125-1131.

Tanenbaum 1981 A.S. Tanenbaum, "Computer Networks," Prentice-Hall, 1981.

Tilak et al. 2002 S. Tilak, N.B. Abu-Ghazaleh, and W. Heinzelman, "A taxonomy of wireless micro-sensor network models," ACM Mobile Computing and Communications Review (MC2R), vol. 6, no. 2, April 2002, pp. 28-36.

Tobagi and Kleinrock 1975 F.A. Tobagi and L. Kleinrock, "Packet Switching in Radio Channels: Part Ⅱ The Hidden Terminal Problem in CSMA and Busy-Tone Solutions," IEEE Transactions on Communications, 24(8) 1417-1433, 1976.

Tolety 1999 V. Tolety, "Load reduction in ad hoc networks using mobile servers," Master's thesis, Colorado School of Mines, 1999.

Tolle et al. 2005 G. Tolle, J. Polastre, R. Szewczyk, D. Culler, N. Turner, K. Tu, S. Burgess, T. Dawson, P. Buonadonna, D. Gay and W. Hong, "A macroscope in the redwoods," Proc. 3rd ACM Conf. on Embedded Networked Sensor Systems, San Diego, CA, 2005, pp. 51-63.

Tseng et al. 2002 Y-C Tseng, S-Y Ni, Y-S Chen and J-P Sheu, "The broadcast storm problem in a mobile ad hoc network," Wireless Networks, v.8 n.2/3, March-May 2002, pp.153-167.

Tuch 1993 B. Tuch, "Development of WaveLAN, an ISM Band Wireless LAN", AT&T Technical Journal, 72, (4), 1993, pp. 27-33.

Virone et al. 2006 G. Virone, A. Wood, L. Selavo, Q. Cao, L. Fang, T. Doan, Z. He, R. Stoleru, S. Lin, and J. A. Stankovic, "An Advanced Wireless Sensor Network for Health Monitoring," Proc. Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare, Arlington, VA, April 2-4, 2006.

Wang et al. 2005 T.-Y. Wang, Y.S. Han, P.K. Varshney and P.-N. Chen, "Distributed fault-tolerant classification in wireless sensor networks," Selected Areas in Communications, IEEE Journal on , vol.23, no.4, April 2005, pp. 724-734.

Wang et al. 2009 X. Wang, X. Wang, X. Fu, G. Xing, and N. Jha, "Flow-based real-time communication in multi-channel wireless sensor networks," Proc. in 6th European Conference on Wireless Sensor Net. 2009, pp. 33-52.

Want et al. 1992 R. Want, A. Hopper, V. Falcão and J. Gibbons, "The Active Badge Location System," ACM Transactions on Information Systems, 1992. 10(1), pp. 91-102.

Wark et al. 2007 T. Wark, C. Crossman, W. Hu, Y. Guo, P. Valencia, P. Sikka, P. Corke, C. Lee, J. Henshall, K. Prayaga, J. O'Grady, M. Reed and A. Fisher. "The design and evaluation of a mobile sensor/actuator network for autonomous animal control," Proc. Int. Conf. on Information Processing in Sensor Networks (IPSN), Cambridge, Massachusetts, April, 2007, pp. 206-215.

Wu et al. 2010a T.-T. Wu, C. Di, B. Stephen, H. G. Goh, K.-H. Kwong, C. Shen, W. Du, C. Michie, I. Andonovic, "A Practical Data Reporting Solution for Free-ranging Cattle Monitoring Applications Using Wireless Sensor Networks," ASABE 2010 Annual International Meeting, Pittsburgh, Pennsylvania, United States, 20 - 23 June 2010.

Wu et al. 2010b T.-T. Wu, K.-H. Kwong, C. Shen, C. Michie, I. Andonovic, "The Impact of Physical Conditions on Network Connectivity in Wireless Sensor Network," The 7th IEEE Consumer Communications and Networking Conference (CCNC 2010), Las Vegas, Nevada, United States, 9 - 12 January 2010.

Xie and Goodman 1993 H. Xie and D. J. Goodman, "Mobility Models and Biased Sampling Problem," IEEE ICUPC, Ottawa, Canada, vol. 2, October 1993, pp. 803-807.

Xie et al. 1993 H. Xie, S. Tabbane, and D. J. Goodman, "Dynamic Location Area Management and Performance Analysis," Proc. 43rd IEEE VTC, Secaucus, NJ, May 1993, pp. 536-539.

Xie and Goodman 1993 H. Xie and D. J. Goodman, "Mobility models and biased sampling problem," Universal Personal Communications, 1993. Personal Communications: Gateway to the 21st Century. Conference Record., 2nd International Conference on, vol.2, 12-15 Oct 1993, pp. 803-807.

Xu et al. 2001 Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed Energy Conservation for Ad-hoc Routing," ACM MobiCom, Rome, Italy, July, 2001.

Xue and Kumar 2004 F. Xue and P.R. Kumar, "The Number of Neighbors Needed for Connectivity of Wireless Networks," Proc. Wireless Networks, 2004, pp. 169-181.

Yan and Gharavi 2006 B. Yan , H. Gharavi, "Multi-Path Multi-Channel Routing Protocol," Proc. the Fifth IEEE International Symposium on Network Computing and Applications, July, 2006, pp. 27-31.

Yang et al. 2002 X. Yang, K. Ong, W. Dreschel, K. Zeng, C. Mungle and C. Grimes, "Design of a wireless sensor network for long-term, in-situ monitoring of an aqueous environment," Sensors, vol. 2, pp. 455-472, 2002.

Ye et al. 2002 W. Ye, J. Heidemann, and D. Estrin, "An Energy Efficient MAC Protocol for Wireless Sensor Networks," Proc. the 21st Int'l Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM 2002), New York, June 2002, pp. 1567-1576.

Ye et al. 2004 W. Ye, J. Heidemann and D. Estrin, "Medium Access Control With Coordinated Adaptive Sleeping for Wireless Sensor Networks," IEEE/ACM Transactions on Networking, Vol. 12, No. 3, June 2004, pp. 493-506.

Yu and Leung 2001 F. Yu and V.C.M. Leung, "Mobility-based predictive call admission control and bandwidth reservation in wireless cellular networks," INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, vol.1, 2001, pp. 518-526.

Zhang 2006 Z. Zhang, "Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges," Communications Surveys & Tutorials, IEEE , vol.8, no.1, First Quarter 2006, pp. 24-37.

Zhang and Wolff 2008 M. Zhang and R. S. Wolff "Routing protocols for vehicular Ad Hoc networks in rural areas," Communications Magazine, IEEE , vol.46, no.11, November 2008, pp. 126-131.

Zhang et al. 2004 P. Zhang, C. M. Sadler, S. A. Lyon, and M. Martonosi, "Hardware Design Experiences in ZebraNet," Proc SenSys 2004. The Second ACM Conference on Embedded Networked Sensor Systems, Baltimore, Maryland, November 2004.

Zhang et al. 2009 X. Zhang, M. Chang, M. Y. Man, W. K. Mo and Y. C. Hang, "A PDA-Based Intelligent Building Service System," Testing and Diagnosis, 2009. ICTD 2009. IEEE Circuits and Systems International Conference on, 28-29 April 2009, pp. 1-3.

Zhou and Shi 2006 J. Zhou and J. Shi, "A scalable routing protocol for low energy Wireless Sensor Networks," Wireless, Mobile and Multimedia Networks, 2006 IET International Conference on, 6-9 Nov. 2006, pp. 1-4.

Zhou et al. 2006a S. Zhou, R. P. Liu and Y.J. Guo, "Energy Efficient Networking Protocols for Wireless Sensor Networks," Industrial Informatics, 2006 IEEE International Conference on, 16-18 Aug. 2006, pp. 1006-1011.

Zhou et al. 2006b G. Zhou, C. Huang, T. Yan, T. He, J. A. Stankovic and T. F. Abdelzaher, "MMSN: Multi-Frequency Media Access Control for Wireless Sensor Networks," INFOCOM 2006. 25th IEEE International Conference on Computer Communications, April 2006, pp. 1-13.