COVER PAGE

MOBILE TOLERANT HYBRID NETWORK ROUTING PROTOCOL FOR WIRELESS SENSOR NETWORKS

By

Jacques Nicolaas Pretorius

Submitted in partial fulfilment of the requirements for the degree

Master of Engineering (Computer Engineering)

In the

Faculty of Engineering, the Built Environment and Information Technology

UNIVERSITY OF PRETORIA

March 2010

ABSTRACT

Mobile Tolerant Hybrid Network Routing Protocol for Wireless Sensor Networks

by

Jacques Nicolaas Pretorius

Supervisor: Prof GP Hancke

Department: Electrical, Electronic and Computer Engineering

Degree: Master of Engineering (Computer Engineering)

Wireless Sensor Networks (WSN) may consist of hundreds or even thousands of nodes and could be used for a multitude of applications such as warfare intelligence or to monitor the environment. A typical WSN node has a limited and usually irreplaceable power source and the efficient use of the available power is of utmost importance to ensure maximum lifetime of each WSN application. Each of the nodes needs to transmit and communicate sensed data to an aggregation point for use by higher layer systems. Data and message transmission among nodes collectively consume the largest amount of the energy available in a WSN. The network routing protocols ensure that every message reaches the destination and has a direct impact on the amount of transmissions to deliver a messages successfully. To this end the transmission protocol within the WSN should be scalable, adaptable and optimized to consume the least possible amount of energy to suite different network architectures and application domains.

This dissertation proposes a Mobile Tolerant Hybrid Energy Efficient Routing Protocol (MT-HEER), where hybrid refers to the inclusion of both flat and hierarchical routing architectures as proposed by Page in the Hybrid Energy Efficient Routing Protocol (HEER). HEER was previously developed at the University of Pretoria and forms the starting point of this research.

The inclusion of mobile nodes in the WSN deployment proves to be detrimental to protocol performance in terms of energy efficiency and message delivery. This negative impact is attributable to assuming that all nodes in the network are statically located. In an attempt to adapt to topological changes caused by mobile nodes, too much energy could be consumed by following traditional network failure algorithms. MT-HEER introduces a mechanism to pro-actively track and utilise mobile nodes as part of the routing strategy.

The protocol is designed with the following in mind: computational simplicity, reliability of message delivery, energy efficiency and most importantly mobility awareness. Messages are propagated through the network along a single path while performing data aggregation along the same route. MT-HEER relies on at least 40% of the nodes in the network being static to perform dynamic route maintenance in an effort to mitigate the risks of topological changes due to mobile nodes.

Simulation results have shown that MT-HEER performs as expected by preserving energy within acceptable limits, while considering the additional energy overhead introduced by dynamic route maintenance. Mobile node tolerance is evident in the protocol's ability to provide a constant successful message delivery ratio at the sink node with the introduction and increase in the number of mobile nodes.

MT-HEER succeeds in providing tolerance to mobile nodes within a WSN while operating within acceptable energy conservation limits.

Keywords:

Wireless sensor networks, flat routing, hierarchical routing, node mobility, energy consumption, energy efficiency, message delivery ratio

UITREKSEL

Bewegings Tolerante Hybriede Netwerk Roeterings Protokol vir Koordlose Sensor Netwerke

deur

Jacques Nicolaas Pretorius

Studieleier: Prof GP Hancke

Departement: Elektriese, Elektroniese and Rekenaar Ingenieurswese

Graad: Meester van Ingenieurswese (Rekenaar Ingenieurswese)

Koordlose Sensor Netwerke mag bestaan uit honderde of selfs duisende nodes en kan gebruik word vir ‘n legio van toepassings soos oorlogs intellegensie of om die omgewing te monitor. ‘n Tipiese node in so ‘n netwerk het ‘n beperkte en soms onvervangbare energie bron. Die effektiewe gebruik van die beskikbare energie is dus van uiterste belang om te verseker dat die maksimum leeftyd vir ‘n koordlose sensor network behaal kan word. Elkeen van die nodes in the network moet die waargeneemde data aanstuur oor die netwerk na ‘n versamelings punt vir latere gebruik deur applikasie vlak stelsels. Informasie en boodskap transmissie tussen die nodes is wel een van die aktiwiteite wat die meeste energie verbruik in the netwerk. Die roeterings protokol verseker dat die boodskappe die eindbestemming behaal en het ‘n direkte impak op die hoeveelheid transmissies wat kan plaas vind om dit te bewerkstellig. Die roeterings protokol moet dus skaleerbaar, aanpasbaar en verfyn word om die minste moontlike energie te verbruik in verskillende toepassings velde.

Hierdie verhandeling stel ‘n Bewegings Tolerante Hybriede Netwerk Roeterings Protokol vir Koordlose Sensor Netwerke (“MT-HEER”) voor. In hierdie konteks verwys hybried na die samesmelting van beide plat en hierargiese roeterings beginsels soos voor gestel deur Page in Hybriede Netwerk Roeterings Protokol (“HEER”). HEER was ontwikkel by die Universiteit van Pretoria en vorm die begin punt van hierdie navorsing.

Die insluiting van bewegende nodes in ‘n Koordlose Sensor Netwerk toon ‘n negatiewe tendens in terme van energie effektiwiteit en suksesvolle boodskap aflewerings by die eindbestemming. Die grootste rede vir hierdie negatiewe tendens is die aanname deur gepubliseerde werke dat alle nodes in die netwerk staties is. Te veel energie sal vermors word indien tradisionele fout korregerende meganismes gevolg word om aan te pas by die bewegende nodes. MT-HEER stel ‘n meganisme voor om die bewegende nodes te gebruik as deel van die roetering strategie en gevolglik ook hierdie nodes te volg soos hulle beweeg deur die netwerk.

Die protokol is ontwikkel met die volgende doelstellings: rekenkundig eenvoudigheid, betroubare boodskap aflewering, energie effektiwiteit en bewustheid van bewegende nodes. Boodskappe word langs ‘n enkele pad gestuur deur die netwerk terwyl boodskap samevoeging bewerkstellig word om die eind bestemming te bereik. MT-HEER vereis wel dat ten minste 40% van die netwerk nodes staties bly om die dienamiese roeterings instandhouding te bewerkstellig.

Simulasie toetse en resultate het bewys dat MT-HEER optree soos verwag in gevalle waar daar bewegende nodes deel uit maak van die netwerk. Energie bewaring is binne verwagte parameters terwyl die addisionele energie verbruik binne rekening gebring word om te sorg vir bewegende nodes. Die protokol se toleransie teen bewegende nodes word ten toon gestel deur die vermoë van die protokol om konstant ‘n hoë suksesvolle bookskap aflewerings verhouding te handhaaf.

MT-HEER behaal die uitgesette doel om ‘n toleransie teen bewegende nodes beskikbaar te stel, terwyl die protokol steeds funksioneer binne verwagte energie besparings limiete.

Sleutel Terme:
Koordlose sensor netwerke, plat roetering, hierargiese roetering, bewegende nodes, energie verbruik, energie effektiwiteit, boodskap aflewerings verhouding

LIST OF ABBREVIATIONS

ACK	Acknowledge Message
ADV	Advertisement Message
CPEQ	Cluster based Periodic, Event-driven and Query-based Routing Protocol
CSMA	Carrier Sense Multiple Access
DRB	Dynamic Route Beacon
DRM	Dynamic Route Maintenance
EDC	Event-driven Cluster Routing
GBR	Gradient Based Routing
GPS	Global Positioning System
GUI	Graphical User Interface
HEAR-SN	Hierarchical Energy-Aware Routing for Sensor Networks
HEER	Hybrid Energy Efficient Routing Protocol
IEEE	Institute of Electrical and Electronics Engineers, Inc.
IP	Internet Protocol
IT	Information Technology
LEACH	Low Energy Adaptive Clustering Hierarchy
MAC	Medium Access Control
MANET	Mobile Ad Hoc Networking
MDC	Mobile Data Collector
MECH	Maximum Energy Cluster-Head
MEMS	Micro Electro-Mechanical Systems
MF	Mobility Framework
MFCA	Minimum Cost Forwarding Algorithm
MT-HEER	Mobile Tolerant Hybrid Energy Efficient Routing Protocol
NIC	Network Interface Card
OMNeT++	Discrete Event Simulation Environment for Communication Networks
OSI	Open Systems Interconnection

PDA	Personal Digital Assistant
PEGASIS	Power-Efficient GAthering in Sensor Information System
PEQ	Periodic, Event-driven and Query-based Routing Protocol
QoS	Quality of Service
REQ	Request Message
RF	Radio Frequency
SEER	Simple Energy Efficient Routing Protocol
SNR	Signal to Noise Ratio
SPIN	Sensor Protocols for Information via Negotiation
TDMA	Time Division Multiple Access
TEEN	Threshold sensitive Energy Efficient sensor Network protocol
TTL	Time-to-live
WSN	Wireless Sensor Network

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Introduction

Technological advances during the last few years in the field of micro electro-mechanical systems (MEMS) and wireless communications have realised the idea of a network of small wireless nodes capable of “sensing” any measurable phenomena in its immediate vicinity [1]. The sensing electronics can measure ambient conditions of the environment surrounding the sensor and transform them to an electronic signal and ultimately a message. The concept is to have a huge, collaborative network of entities being able to sense and wirelessly convey messages containing the sensed data to higher-level systems for analysis and reactive measures. These entities (or sensor nodes) form a wireless connected network to meet the intended application domain’s requirements. Industry accepts this network of nodes as Wireless Sensor Networks (WSN). The network of nodes may consist of hundreds or even thousands of nodes, which implies that it should be very cheap to produce in large quantities and be adaptable to different application areas. The implementation and usage of WSN may vary from environmental monitoring to biological applications to military applications concerned with surveillance, reconnaissance and targeting [1].

Every node is bound to constraints of which a limited power supply due to the size and construction of a WSN node is the most important. The functional lifetime of each node directly impacts the performance of the WSN as a whole. Collectively all the nodes within the WSN should consume the least possible quantity of energy to prolong the lifetime of the total network.

The term “sink node” is accepted as a base station that will collect and possibly store all the sensed information from each of the sensing nodes. The WSN paradigm dictates that there always be at least one sink node, but need not be limited to one. Sink nodes typically do not have the same energy constraints as the majority of sensing nodes and should be the only interfaces to higher-level systems that will consume the sensed data. The sensing nodes are generally accepted as “source nodes” that will gather information from its sensors. The sensed information in the format of a message will be transmitted wirelessly across multiple source nodes in an attempt to deliver the original sensed information to the sink node. The decision to create messages from sensed information is based on the WSN application and may be requested from or initiated by a source node. A routing protocol is the mechanism implemented and utilised to determine the path a message should take to traverse the source nodes in an attempt to eventually reach the sink node. The design of a routing protocol determines the amount of message transmissions and directly impacts the energy efficiency of a WSN.

1.2 Scope of Work

This dissertation is focused on the investigation, design and implementation of a mobile tolerant energy efficient network routing protocol for WSN to aggregate and route sensed data to a predetermined data extraction point. This research aims to contribute to the current WSN body of knowledge by analysing the ever-increasing requirements from a WSN application perspective, yet at the same time considering theoretical advances in the field of WSN routing protocols. The major goal of this dissertation is to further previous research undertaken at the University of Pretoria to enhance the efficiency and reliability of WSN network routing protocols.

Possible WSN applications would certainly dictate the requirement of handling moving nodes within the WSN. Topological changes due to node failures and mobile nodes are inevitable and a multitude of literature proposes mechanisms to mitigate node failures. Continuing research on WSN should implicitly minimise the power consumption due to routing activities within the WSN as well as enhance the WSN tolerance to topological changes due to node mobility.

There are a few hidden risk factors in the design and analyses of WSN protocol specifications. The first is the amount and relevance of assumptions made to develop the protocol. All assumptions made a scientific process need to be clearly stated and its validity evaluated against industry accepted standards. The impact of each assumption should be critically analysed when researching and attempting to enhance existing protocol specifications. This is of high importance during the performance analysis of a network routing protocol. The second is the level of abstraction of complexity by making too much assumptions and setting too many variables as static. Implementation and analysis of the proposed protocol could lead to invalid results due to a high level of abstraction. The third is the simulation arena for theoretical and analytical protocol evaluation. This cannot be controlled from a governance point of view and can easily ensure that one protocol outperforms another. The research in this dissertation aims to clearly state, analyse and mitigate the aforementioned risk factors.

1.3 Research objectives

As with every science there is always a need to expand the body of knowledge by creating and enhancing the models and theories for that science. Research completed by Leuschner [2] and Page [3] at the University of Pretoria forms the starting point for this research, while embracing and consuming the ever expanding body of published literature on this topic. This research develops a methodology for defining a mobile tolerant and energy aware network routing protocol for WSN by identifying shortcomings and weaknesses in currently published literature. The proposed routing protocol adds a new approach to the existing collection of texts on WSN network routing protocols by extracting and combining existing methodologies to produce a new routing technique.

The research investigates, defines and implements a WSN network routing protocol that is energy aware to sustain effective network functionality for as long as possible, is scalable and has the ability to handle failures and mobile nodes within the network. A critical evaluation of routing protocols is conducted to determine the protocol's relevance and applicability to networks containing both static and mobile node.

The network architecture of choice would be a hybrid of both flat and hierarchical network routing structures by combining attributes and functionalities of both structures into an event driven operation. Event driven operation is chosen to eliminate the complexity and unnecessary network traffic introduced by other types of operations such requesting data from specific nodes. In principle, the protocol will analyse the energy availability of the all the nodes next in line for transmission to the destination as well as the energy availability for the total path to the destination. The two most prominent protocols for flat networks (directed flooding and SPIN) will be used as an extension to [2] and [3] to form a basic foundation with the additional benefits of hierarchical network protocols.

The hierarchical properties are manifested by dynamically selecting one or more aggregation nodes per grouping of nodes if the transmission path to the destination would exceed a calculated amount of hops. This calculation depends on the cost of reaching the destination in terms of energy usage to the next node, as well as to the final destination (sink node). The calculation of link transmission costs and data aggregation is attempted in sources such as [4], [5], [6] and [7] that implement this concept for hierarchical network protocols with fairly good results. These sources will form a foundation to extrapolate cost calculations. If this calculated value is sufficiently low enough, there is no need to add the complexity of a hierarchical routing protocol. The aggregating nodes will be responsible only for relaying messages but its role may change over time depending on its energy availability. The selection of the aggregation node will be based on the highest energy levels, amount of nodes possibly selecting that node as the next hop and fewest hops to the destination.

One problem identified in the literature is the ability of most protocols to handle topological changes in the network without consuming too much available energy. Some protocol specifications state that the protocol acts pro-actively against successful transmission failures, but this is not possible if the failure is caused by damage to the nodes, failure of components within the nodes or mobile nodes moving beyond transmission range. These failures might even cause network separation and will form part of the proposed protocol specification. This is specifically the reason why the chosen path to the destination is of utmost importance. Should the next hop selection only be made based on neighbour energy levels and the assumption that the neighbour node is still within transmission range, there is no assurance that the destination will be reached. Sources in literature (for example [8] and [9]) propose fault handling processes that will be enhanced to suite the proposed research. Extensive research in literature covering the techniques and applicability to networks containing mobile and static nodes is scarce. Literature that explicitly covers topological changes (such as [10]) is used to enhance the designed protocol.

Although node mobility is currently not one of the main requirements for WSN, it is the author's opinion that this is a short coming in some unique cases. Node mobility may either be accidental or engineered as part of the planned application arena. Under accidental node mobility, protocol performance may be negatively hampered under severe circumstances. This research however focuses on the planned application scenario. Possible applications could include:

1. Mining applications may require personnel to perform rescue operations in remote locations and the subsequent environmental data of those locations. Although static nodes may be deployed in those the areas, various nodes may have become damaged in accidents or may not monitor specific environmental parameters posing a threat. The mobile nodes can be attached to the rescue personnel and could monitor and forward additional information.
2. Security companies that are hired to protect and patrol security estates can be issued with WSN nodes. The static nodes can be deployed to sense break-ins, fire alarms or any security related data from residential properties. Mobile nodes can be issued to patrolling personnel that can gather information like the movement and speed of the patrolling guards, areas visited, etc.
3. Oceanographic research of conditions on reefs with scuba divers allocated sensor nodes to collect additional information during dives. The information collected by the divers could be in locations where keeping nodes static is not possible, such as mid-water analysis.
4. Oceanographic research on the sea conditions that specific fish species experience while migrate among different reefs. The reefs can house the static nodes and the mobile nodes can collect and forward parameters such as temperature, depth, surge, movement, travel speed, etc.

It should be noted that sub aqua research will definitely change the radio model to that suggested in this dissertation, but the protocol operation should still remain similar.

This research contributes to the body of knowledge by proposing the Mobile Tolerant Hybrid Energy Efficient network Routing (MT-HEER) protocol. MT-HEER incorporates a mechanism to mitigate the risks and promote message delivery when introducing mobile nodes to the WSN. Additional measures and mechanisms are proposed to mitigate the complexities that node mobility adds to the protocol performance. These mechanisms aims to balance the network routing protocol's tolerance to node mobility, the always present need for power conservation and an acceptable network packet delivery ratio at the sink node.

1.4 Research Design and Methodology

The research designs and proposes an implementation for a WSN routing protocol using a qualitative methodology. Within this paradigm, the researcher attempted to thoroughly understand and link methodologies, principles, processes and assumptions of existing routing protocol designs to the proposed protocol. The focus is on elaborating, enhancing and combining existing network routing protocol specifications by using deductive strategies to form a model for an enhanced network routing protocol to better handle topological changes within a WSN network.

The validity of the proposed protocol principles are highlighted by the execution of simulations and the comparison of the results between the proposed protocol design and other published protocols such as directed flooding [15], Simple Energy Efficient Routing Protocol (SEER) [2], Low Energy Adaptive Clustering Hierarchy (LEACH) [24] and Hybrid Energy Efficient Routing Protocol (HEER) [3]. The aforementioned protocols were selected due to the fact that the simulation model implementations were readily available and accepted as being correctly implemented. The comparison and analysis is manifested by performing theoretical and programmatic analysis by means of an accepted and academically used WSN simulator under the same set of assumptions and conditions for all protocol implementations. The assumptions and conditions are evaluated and derived from previous simulation models. The simulator employed to produce simulation results for the purpose of research analysis, is OMNet++ [37] with the addition of the Mobility Framework (MF) [39]. The combination of the core simulator and the MF extension provides a powerful simulation environment for WSN and mobile models.

The protocol comparison is evaluated on the following minimum set of performance criteria:

1. Scalability of the protocol (varying network size)
2. Computational complexity (number of messages sent during initialization of the network, stable conditions and failure conditions)
3. Ratio of successful message delivery to the amount of created messages at the destination.
4. Changing network architecture due to topological changes (tolerance to failures and node mobility)

1.5 Outline of Dissertation

The remainder of this dissertation is organised as follows:

Chapter 2: Literature Overview presents a brief overview and summary of the body of knowledge covering WSN network principles, the communication within a WSN, network routing protocols, energy efficiency and node mobility.

Chapter 3: WSN Routing Protocol Evaluation summarises the seminal contributions to WSN routing protocols. The research and design of a mobile tolerant routing protocol utilises selected mechanisms from the summarised work.

Chapter 4: WSN Routing Protocol Simulations provides some background on the simulation environment and simulation assumptions that are kept constant for all simulations.

Chapter 5: WSN Routing Protocol Design proposes the Mobile Tolerant Hybrid Energy Efficient network Routing (MT-HEER) protocol design and operation. The chapter includes all the design and simulation assumptions made during the course of this research.

Chapter 6: Results and Discussion presents the experimental principles and procedures used to produce a set of simulation results. The presented results are objectively analysed and discussed. The chapter continues with a discussion of factors that could influence WSN simulations and may unknowingly undermine or superimpose resulting protocol performance.

Chapter 7: Conclusion concludes this research.

1.6 Chapter summary

This research investigates, defines and implements a WSN network routing protocol that is energy aware, scalable, computationally simple and tolerant to topological changes due to node failures and mobile nodes by following the described research methodology.

Chapter 2

LITERATURE OVERVIEW

2.1 Introduction

The energy resources on WSN nodes differ from the normal mobile networks present in the IT and mobile telecommunication domain. Each node is subject to a set of constraints inherent to the nature of the technology. These constraints include limited processing speed, storage capacity, communication bandwidth and energy resources ([1], [11] to [14]). Of the aforementioned constraints, limited energy resources is the most important factor influencing the functional lifetime of the nodes due to the generally irreplaceable power sources of the nodes [1]. The limited power source [11] can practically not be replenished by conventional ways such as recharging the batteries, due to the nature of the WSN application domain.

Wired networks on the other hand are a more mature technology and have the advantage of established network principles and protocols; for example the IP routing standard. The type of constraints placed on WSN and wired networks differ significantly. Due to these differences there is little or no opportunity to inherit principles from the wired networks for application to WSN principles.

Literature states that transmission among nodes consumes the majority of the available energy on each individual node [14]. Protocol designs for WSN should be aware of this and aim to reduce the amount of transmissions required to successfully deliver a message at the sink. Proposals for new and innovative protocols on all of the levels of the protocol stack (such as [15] and [16]) consider these constraints to prolong functional network lifetime and enhance power efficiency.

Network topologies are dynamically changing structures due to mobility of nodes and unavailability of nodes over time. Node mobility is included for the purpose of the intended research and is discussed in more detail in 3.5. A compound network consisting of both static and mobile nodes is considered for the purpose of research and analysis in this dissertation. WSN nodes are prone to failure [1] and may fail due to depleted power sources or physical damage caused by the immediate environment surrounding the nodes. Environmental interference may cause a node being intermittently unavailable. None of the aforementioned failures or topological changes due to mobility should affect the overall effectively and functionality of the WSN. Literature presents a few examples [17] of providing multiple levels of path redundancy to ensure that every message reaches its destination successfully.

When performing research on such a rapidly growing body of knowledge as WSN, there are a number of approaches to the technology. A literature survey on WSN yields many scientific approaches to the design and implementation of these networks. A wide variety of approaches is available from an application and middleware viewpoint (application programmer's perspective where the underlying technology is of no relevance) down to the physical hardware design of the nodes (engineer's perspective to improve the physical characteristics and functionality of each node). Each approach attempts to fulfil its own need and in doing so, abstract certain difficulties away from the reality of WSN implementations. Every information source should be analyzed in the correct context to determine the relevance, integrity and validity of the assumptions made to deliver results (if presented in the information source) as well as to determine the applicability to the scope of research being done.

2.2 WSN Overview

2.2.1 Background

The WSN arena is filled with various authors' attempts to specify a new network routing protocol that might have the same impact to WSN as the Internet Protocol had for the internet. A thorough understanding of WSN and the technology enabling the WSN applications is of utmost importance to investigate and classify the relevance of the completed works. Technical databases such as the IEEE and University research databases are flooded with sources such as [11] to [14] to provide the uninformed and interested reader with a solid foundation in the fundamentals of WSN. These sources are solely written at a high level and intended to provide the reader with a general overview of the technology, research completed on the subject and the vast array of possible applications for WSN. Some documents (such as [12]) name specific research issues that need to be addressed to grow the current body of knowledge and this is a good starting point in identifying gaps that can be filled by research undertaken in these directions. Although this level of information is directly related to the chosen field of study, it is necessary to understand principles.

2.2.2 Properties and Constraints

The unique properties and constraints of WSN that will influence the design of any routing protocol [16] can be summarized as:

1. Constrained energy supplies limiting the nodes' functional lifetime
2. Limited computational power and memory
3. Unreliable wireless transmission medium
4. Network self organisation and configuration
5. Scalability in the order of 1000s of nodes which influences node addressing
6. Node mobility
7. Node failures are expected due to the possible environments where WSN are deployed
8. Communication is data centric and in small but frequent packets
9. Location awareness maybe required

2.2.3 WSN Protocol Stack

The commonly accepted model for the WSN protocol stack is adapted from the established Open Systems Interconnection (OSI) reference protocol stack that is used in IT and telecommunications. The WSN protocol stack enables integration between application and routing layer, power awareness, routing capability and efficient wireless transmission within the nodes. The layers within the protocol stack [15] are represented by Figure 2.1.

Application Layer
Transport Layer
Network Layer
Data Link Layer
Physical Layer

Figure 2.1 WSN Protocol Stack

The responsibilities of each layer can briefly be explained as:

1. The physical layer provides the wireless communication radio interface and ensures successful transmission of data by implementing robust signal modulation and transmission techniques (sending/receiving messages).
2. The protocol implemented on the data link layer is typically known as the MAC protocol. MAC protocols are typically expected to be power aware, perform noise filtering and minimize message collision with neighbouring nodes' transmission.
3. The network layer implements the routing protocol and is also the research area of this dissertation.
4. The transport and application layers are commonly fused to form a single layer. This layer ensures the flow of data and is responsible for hosting the application software depending on the sensor requirements.
5. Continuing academic research has shown an interest in breaking down the boundaries between the layers in the protocol stack. This research would produce a "cross-layered" approach to bridge the boundaries among any of the currently accepted layers or even merge some of these layers. It is however not the focus of this dissertation and will not be elaborated on further. The interested reader may consult sources such as [9] and to a lesser extent [3] for more information on "cross-layered" protocol designs.

There is mention in [15] of additional functional planes to support responsibilities on each of the layers of the protocol stack to improve energy efficiency in a WSN. These planes are responsible for power, mobility and task management. The power management plane dictates how power is managed within a node, such as sending out notifications when the available power reaches a predefined limit. The mobility plane is responsible for knowing neighbour nodes' and the node's mobility status to base decisions on neighbour status and routing on. The task management plane takes care of sharing tasks between nodes to ensure that some nodes do not get over utilised and prematurely become unavailable.

2.2.4 Routing Defined

A network routing protocol may be defined as a mechanism to propagate a message or data from a source node to a sink node. Message propagation may be with one single transmission or multiple transmissions traversing multiple nodes to reach the sink node. Literature accepts each transmission between two nodes as a hop towards an end point. An end point may be a message aggregation node, or it may be a sink node.

2.2.5 Routing Protocol Classification

There are different paradigms to network layer routing protocols depending on the network structure and the operation that the protocol would like to achieve [17]. The network structure of a WSN could be flat, hierarchical or location based, whereas the operation could be negotiation based, multi-path based, event based, query based or QoS (quality of service) based. Investigating and proposing a routing protocol would depend on the selection of network structure and operation. A survey and comparison is done of well-known WSN routing protocols in [15], [16] and [17] to show the positives and negatives of the mentioned protocols. These sources are highly relevant to determine the assumptions and base from which this research project starts and it provides guidance to established methodologies and practices.

2.3 Network routing protocols

2.3.1 Protocol Categories

As mentioned in the previous section WSN network routing protocols are traditionally categorized into direct, flat or hierarchical routing techniques depending on the network structure [17]. Alternatives to the traditional network routing paradigms is hybrid and mobile aware routing protocols.

1. *Direct routing* transmits messages directly to the destination or sink node. This paradigm however does not constitute a network routing scheme in the true definition of the term and is not considered of any consequence for the remainder of this dissertation. For informational purposes however; direct routing assumes that all source nodes are within the maximum transmission distance of the sink node and only one hop away from the sink. Practically this is not possible in a WSN as it would require very high transmission power in the source nodes placed far away from the sink. Each transmission would consume equally high power to transmit messages to the sink node. This operation directly opposes WSN implementation in general, due to the fact that nodes are constrained by limited power sources.
2. A *flat routing protocol* assumes that each node is equal and transmission to the destination node will propagate through the network on rules determined by the protocol. This family of protocols is discussed in more detail in section 3.2.
3. A *hierarchical or clustering routing protocol* divides the WSN into groups of nodes. Routing to a destination node will traverse the master node of the group, which in turn will then relay the message transmission towards the destination node. This family of protocols will be discussed in more detail in section 3.3.
4. A *hybrid routing protocol* is a relatively new approach that combines attributes from both flat and hierarchical routing techniques. Many scholars might classify these protocols as hierarchical protocols but for the purpose of this research, hybrid routing is handled separately as described in section 3.4.

5. In addition to the network structure a routing protocol could be classified by the protocol's *awareness of node mobility* within the network. Traditionally flat and hierarchical architectures primarily focus on networks with only static nodes. Node mobility awareness is consequently described separately in section 3.5.

2.3.2 Protocol Operations

The application utilising the WSN will determine the protocol operation of the routing technique. Routing protocols may use one or more of the operations mentioned below. These protocol operations can be based on [17]:

1. *Negotiation* routing minimises data transmissions by discarding redundant or duplicate data transmissions. The protocols using this type of operation will suppress duplicate messages and prevent redundant data from being transmitted. This operation is manifested by exchanging negotiation messages before actual data transmission commences.
2. *Multi-path* routing protocols enhance network reliability and fault tolerance by maintaining multiple paths towards the sink node. The alternative path information is kept up to date by periodic "broadcast"-like messages. If the primary path fails, the alternative will immediately be used. Increased reliability is gained at the expense of higher energy consumption with the increased overhead of maintaining an alternative path for message delivery. Message forwarding may also traverse both paths, hence increasing network traffic at even higher energy consumption. Direct Diffusion [23] utilises this method of operation for routing.
3. In *event driven* protocols source nodes react if the variable being sensed reaches a predefined threshold. This generates an event that will be handled by the design rules of the routing protocol to determine when and where the message will be forwarded.
4. *Query based* routing functions on the basis of a sink node requesting certain information from the WSN. This request may be sent to a predetermined node or may be propagated on the network ensuring all nodes with the requested data will reply. Direct Diffusion and Rumour Routing [23] utilises this method of operation for routing.

5. *QoS (Quality of Service)* based routing weighs energy consumption up against data quality. Routing decision logic is based on QoS metrics such as message delay, energy consumption to reach the sink node, message priority and delivery success ratio. These metrics are introduced in networks and applications with a low tolerance to a specific metric, for example a low tolerance to message delays in chemical production plant. Although energy efficiency is not the main focus in these protocols, it is still one of the main considerations in protocol evaluation.
6. Nodes in a WSN needs to cooperatively process different types of data as specified by the routing protocol's design, either within source nodes, data aggregator nodes or sink nodes. Research produced two types of data processing: *coherent and non-coherent processing*. Non-coherent data processing will process raw data in each node before forwarding a message to the next node towards the sink node. Coherent processing will perform minimal processing like header stripping, time stamping and redundant message suppression. Coherent data processing or data aggregation is selected in most cases in light of energy efficiency.

2.4 Considerations for WSN Routing Protocol Design

The factors influencing the design ([12] and [14]) of WSN routing protocols are important because they should provide protocol development guidelines. These factors could be used for the purposes of analysis and comparison among routing protocols. The aforementioned factors can be summarised as:

1. *Scalability:* Network size may exceed 1000's or even 10000's and a routing protocol should be able to handle this amount of nodes.
2. *Node density:* Nodes may be highly concentrated in specific sensing areas to maximise the sensing coverage area. Routing protocols should be able to select energy optimal routing paths within the dense node deployment area. The node density can typically be calculated as:

$$\mu(R) = \frac{N.\pi R^2}{A} \tag{1}$$

   Where N is the number of sensor nodes in region A and R is the maximum transmission range.

3. *Deployment:* Sensor node deployment may be extremely dense requiring a precise and energy efficient mechanism for handling of network topology maintenance. Node deployment may be a random distribution or at predetermined positions. Topology can typically be divided into three phases:
    a. Deployment of nodes can be randomly distributed or placed at predetermined positions.
    b. Post-deployment phase where topology changes may occur due to failures, mobility, interference, etc.
    c. Redeployment of nodes to augment areas where topology changes are causing detrimental network performance.
4. *Network topology:* A multi hop network of sensor nodes may form arbitrary paths from any source node to a sink node. An important property on the WSN is maximum amount of hops to reach the sink node as determined by the routing protocol. This affects network characteristics such as latency and path failure tolerance. Networks should be self configurable and also have the ability to reconfigure should topology change occur.
5. *Fault tolerance (Reliability):* Node failures are inevitable in WSN due to power depletion, damage or interference. A routing protocol's tolerance to fault conditions is measured against the effect of failing nodes on the remainder of the network's functionality. The reliability of a sensor node [14] is modelled by a Poisson distribution of the probability that a sensor node will not experience a failure within an interval *t:*

$$R_k(t) = e^{-\lambda_k . t} \quad (2)$$

    Where λ_k is the failure rate of sensor nodes in time period *t.*
6. *Mobility:* The movement of sensor nodes may be required by the application or may be an incidental effect and may apply to any percentage of the nodes in the network. Movement may be at a constant speed or at intervals with periods of remaining static. The movement pattern of nodes may be random or determined by a preselected model, for example soldiers patrolling the perimeter of a base. Random movement will impact network topology changes worse than a predefined movement model, due to the unpredictability of the movement. Each of the mentioned mobility parameters will influence the

network topology differently and the model assumed will impact the design and function of routing protocols.

7. *Transmission media:* Sensor nodes are linked with a wireless medium which eliminates the criteria for line-of-sight between nodes. Most of the sensor nodes available in the industry utilise radio frequencies (RF) for transmission among nodes.
8. *Connectivity:* Two nodes are connected if the nodes' positions are such that they are within a maximum transmission distance R from each other. If each source node is permanently connected to a sink node with an arbitrary amount of hops, the network is termed as connected. Connectivity is sporadic if network connectivity is broken due to node mobility, node failures or interference.
9. *Hardware Constraints:* The typical sensor node architecture comprises of a power module, sensing module, transceiver module, processing module and possibly additional application specific modules, as shown in Figure 2.2. The most important component is the power module that determines the total time-to-live (TTL). The additional modules can fulfil tasks such as positioning, but should be limited to the minimum, because all the components need to fit into a very small containment unit and should consume the minimum amount of energy.

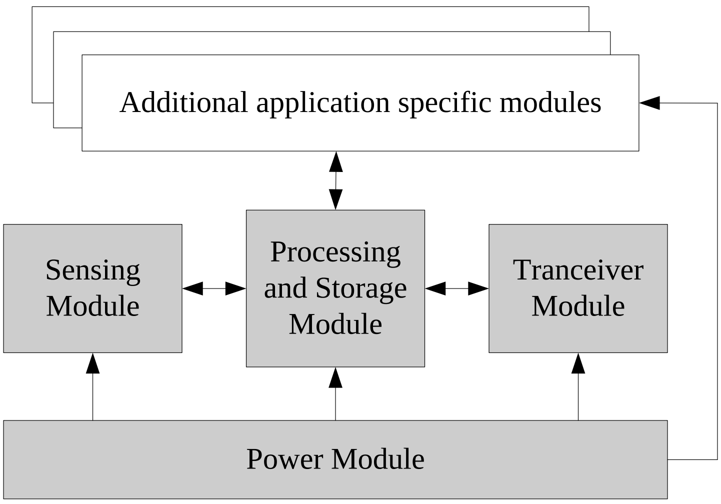

Figure 2.2 Typical sensor node architecture

10. *Environment:* Sensor nodes can be literally placed very close or even directly within the sensed phenomena.
11. *Power Consumption:* The limited power supply of a sensor node directly affects the functional lifetime of that node, and indirectly the functional lifetime of the network. Each node in the WSN may assume the role of data originator, data aggregator and/or data forwarder. The failure of a node may severely impact the network topology and negatively performance. Additional energy will be consumed in an attempt to rectify routing information. The main operations consuming power on a sensor node can be attributed to:
    a. Communication or transmission of messages which is the highest consumer of node energy.
    b. Sensing activities which are typically the second highest consumer of energy.
    c. Processing that involves calculations, routing protocol operations, application layer procedures, etc and consumes the least amount of energy.
12. *Lifetime:* The intended application determines the required functional lifetime of a WSN. The functional lifetime has a direct impact on the energy efficiency of the routing protocol and node robustness.
13. *Self configuration:* Given the possible number of nodes deployed in a WSN, the network should be able to self configure without any manual intervention. Self configuration of the network should consume the least amount of energy possible to enable a high degree of connectivity. The network should be able to automatically reconfigure to maintain the level of connectivity should network topology changes occur. Reconfiguration may be initiated by:
    d. Periodic intervals.
    e. Mechanisms to continuously update routing information. This type of network reconfiguration may potentially consume huge amounts of energy and its usage should be weighed against QoS properties.

2.5 Chapter Summary

Wireless sensor networks are a fairly new technology with multitudes of possibilities for new and interesting applications. These networks can be deployed randomly without physical restriction as with normal wired sensor systems.

Communication within these networks is of utmost importance and also the most energy consuming activity. Due to the physical restrictions of the network and its components, energy conservation should be the very first consideration when developing WSN and researching to further the current body of knowledge.

This chapter presented a brief overview on the communication with a WSN and in more detail concepts and principles that apply to the network routing protocol layer. To conserve energy and maximise the lifetime of the network, a routing protocol needs to be as simple and scalable as possible. The design of a routing protocol should consider the intended network topology and operations required from the application layer. The considerations for protocol design, simulation and evaluation are summarised at a high level.

Chapter 3

WSN ROUTING PROTOCOLS

3.1 Introduction

Energy awareness in WSN is of critical importance to the lifetime of the network. Many attempts have been made to create "intelligent" routing protocols that base routing decisions on energy levels within the nodes and in the network as a whole. All the routing protocols described in this chapter attempts to focus on energy awareness with trade-offs to certain design decisions.

Chapter 2 describes at a high level that WSN routing protocols can be categorised as implementing flat, hierarchical, hybrid and/or mobility routing techniques. Each of these protocol paradigms are described in this chapter. Hybrid routing is a fairly new concept and many scholars might classify these protocols as hierarchical protocols. For the purpose of this dissertation, hybrid routing is handled on its own due to the unique combination of both flat and hierarchical properties in these types of protocols (hence the term hybrid). Throughout the available literature flat and hierarchical architectures primarily focus on networks with only static nodes and to this end node mobility is also handled separately in this chapter.

3.2 Flat Routing Protocols

Flat routing protocols view all nodes as equal. In other words, all nodes will sense data as well as perform routing of messages and data. Literature accepts the following established flat routing protocols [18]: flooding, gossiping, SPIN (Sensor Protocols for Information via Negotiation), directed diffusion, rumour routing, MCFA (Minimum Cost Forwarding Algorithm), GBR (gradient based routing) and certain energy aware routing protocols as set out in [15], [16] and [17]. The seminal contributions for flat network architectures are summarised in the following sub-sections.

3.2.1 Flooding and Gossiping

Flooding can actually not be classified as a routing algorithm in the true meaning of the term. Flooding is rather a mechanism that functions on the same layer as routing protocols that will ensure data messages will reach a sink node. Flooding is a source initiated protocol and each received message is stored in the node's memory. There is a small amount of decision logic involved as every node will broadcast the received messages only if the message has not yet been forwarded. The broadcasting of a specific message will continue until the message reaches the sink, or the message has reached its maximum time to live (TTL). Maximum TTL is typically specified either by the allowable amount of hops or as a time limit. The operation of Flooding can be summarised by Figure 3.1 (adopted from [3]).

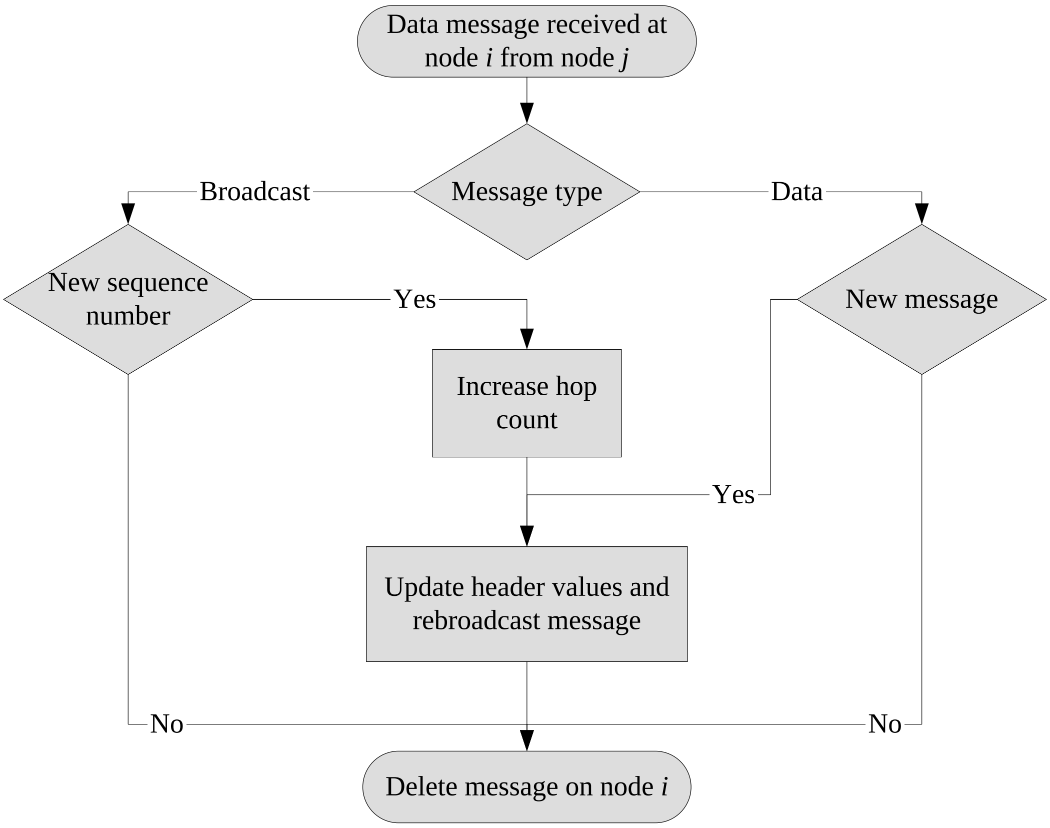

Figure 3.1 Message receiving process flow for Flooding

The major advantages of flooding are:

1. Computational *simplicity* in terms of routing and topology maintenance
2. High degree of fault *tolerance* and topology changes

Some disadvantages of flooding are:

1. *Implosion*: Data messages are always forwarded from a specific node if that node has not yet sent it. The same messages could potentially be sent to the same node but from different sources.
2. *Overlap*: Two source nodes may share the same sensing area and the same data is broadcasted and received by the two nodes with no data aggregation or dissemination. Overlap is a much harder problem to solve.
3. *High energy consumption*: Nodes are not energy aware and do not modify routing logic based on available energy levels.

Gossiping is a derived version of flooding and attempts to avoid implosion by randomly selecting the next node from its neighbours rather than broadcasting a message. The receiving node will randomly select a new neighbour and forward the message. The cost of avoiding implosion is the long propagation times it may take for a message to reach a sink.

Directed flooding [19] is adapted from flooding, but claims to consume much less energy and exhibit an even higher tolerance to network faults than flooding. This proposal however requires geographic positional and directional knowledge of the network such as GPS abilities.

Flooding and gossiping should theoretically be able to adapt easily to network topology changes due to the fact that no routing information is discovered and stored.

3.2.2 Sensor Protocols for Information via Negotiation (SPIN)

Sensor Protocols for Information via Negotiation [20] is a source initiated protocol and proposes mechanisms to overcome implosion and overlapping. Message transmission is dependent on inter node negotiation to ensure that only useful data is transmitted. Inter node negotiation is only executed with neighbour nodes which requires local knowledge of only those nodes.

Negotiation among nodes use message data descriptors called meta-data to describe the data that needs to be forwarded. Transmission of messages, and in particular data messages, is the most energy expensive operation in WSN. Meta-data enables the routing protocol to exchange information about the data rather than the data itself and to efficiently disseminate data messages. Data dissemination is achieved by determining what data has already been received by a node and what information still needs to be sent to the node. This mechanism reduces the amount of transmissions of duplicate data messages. The SPIN meta-data model relies on a common naming standard between the routing protocol and application layer, producing a layer integrated approach.

Every node also has a resource manager that keeps track of node resources such as energy levels. The resource manager will be consulted during routing activities in an attempt to reduce energy consuming activities when energy levels drop too low. An example of this is the unconditional forwarding of messages without negotiation. Interfacing to the resource manager is also available to the application level.

The operation of SPIN is based on the usage of three message types:

1. ADV: advertisement of new available data that can be forwarded to neighbour nodes, by sending a message containing meta-data of the data message
2. REQ: meta data reply message to an ADV message to acknowledge that the neighbour is interested in the advertised data (only for new data)
3. DATA: actual message containing the advertised data if a specific node received a REQ message

SPIN can be summarized by Figure 3.2 which was adopted and simplified from [20]. Node A offers data with an ADV message to node B. Node B responds with a REQ message, upon which node A forwards a DATA message. Node B now offers data with an ADV message to all its neighbours. The sink node, node C and D is interested and responds with a REQ message. Node B forwards the DATA message to all the interested nodes. The process continues until all data messages have reached the sink node.

The mechanisms implemented by SPIN overcome the disadvantages of flooding mentioned in the previous section as follows:

1. Negotiation eliminates implosion by not transmitting redundant data messages
2. Meta-data overcomes overlap by not sending duplicate data to nodes that is already aware of predetermined data.
3. Energy awareness prolongs the functional network lifetime

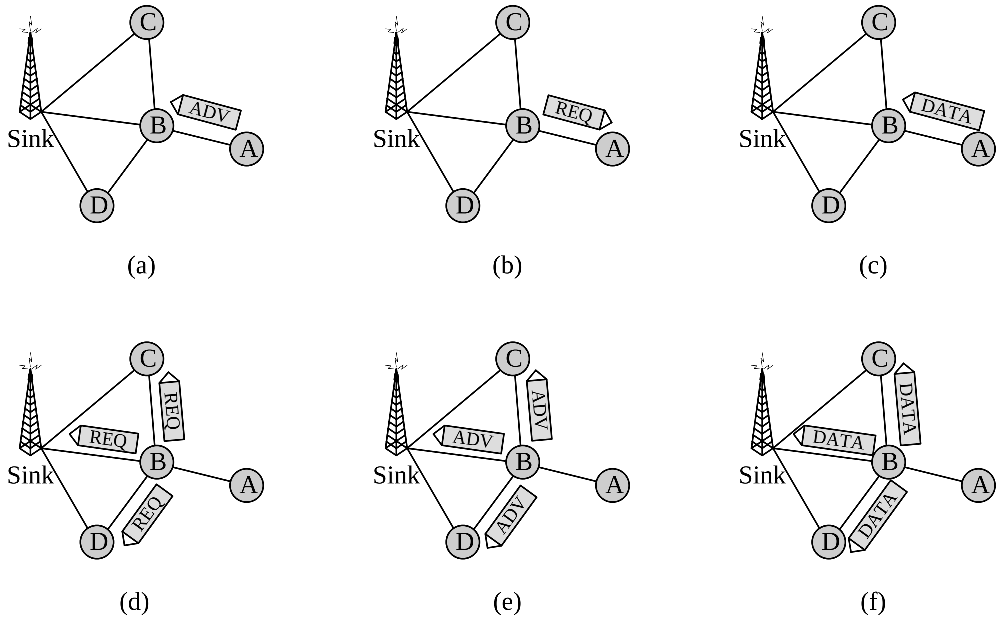

Figure 3.2 SPIN message sequence flow from (a) to (f)

3.2.3 Minimum Cost Forwarding Algorithm (MCFA)

The Minimum Cost Forwarding Algorithm [21] aims to deliver a data message from a source node to the sink node along a minimum cost path. MFCA is a source initiated protocol and assumes that data messages will always flow in the direction of the sink node. The implementation of MCFA does not require routing table storage at source nodes. The only information stored at each source node is the cost of forwarding a data message to the sink node. The minimum cost for a specific node is the optimal path from that node to the sink node. The cost may be one or a combination of: hop count towards the sink node, transmission energy consumed to reach the sink node, a function of the received message signal strength, the number of retransmissions or some other predefined metric.

MCFA operates within two modes in the WSN, the initialisation and operational mode.

1. *Initialisation mode:* Each source node initially sets its cost infinity. The sink node initialises the network with the generation of an advertisement (ADV) message with a cost of zero. A node receiving the ADV message will update its cost if the received cost plus the link cost is less than the locally stored cost. In the case of lower cost the ADV message will be broadcasted again from the receiving node while updating the cost of the ADV message with the locally stored cost. In the case of a higher cost, the ADV message will be discarded. The authors of [21] suggest a timeout period on each node to wait for ADV messages in order to minimise the overall amount of ADV messages being broadcast. This process will establish a minimum cost at each node.
2. *Operational mode:* Once the costs have been established at each of the nodes, transmitted data message will flow along the minimum cost path. If a data message originates at a source node, it will be broadcast to all nodes. All data messages will include the minimum cost from that source to the sink node as well as the consumed cost. A receiving node will only forward the message (by means of a broadcast) if the sum of the consumed cost and the cost at that node is the same as the source's cost. The process will be repeated until the message reaches the sink node. This mechanism routes a message using the minimum cost path without maintaining path information.

MCFA does however have its set of weaknesses and [22] proposes a few mitigating mechanisms. Unless minimum cost is updated periodically, the power sources of nodes along the *minimum cost path will be depleted prematurely*. Should the *network topology change* due to node failures or node mobility, data messages may be lost if a path with equal cost is not present. Although paths with equal cost will provide tolerance to network topology changes, *equal cost paths* will be much more energy consuming by sending the same messages along two different paths.

3.2.4 Directed Diffusion

Directed diffusion [23] is one of the most referenced routing protocols and has been used as base for many protocol designs. The routing protocol is designed upon a query based operation which directs queries at certain parts of the network for requested data. Data queries are initiated at the destination (sink nodes in WSN) and the network reacts to the request. The routing protocol is data centric using attribute value pairs and is tightly integrated into the application layer which causes the protocol to be highly application aware. The protocols promote energy savings by empirically selecting paths, aggregating data during routing activities in the network to reduce transmissions and by shutting down sensors when there is no interest for data.

An interest message specifies the “query” for the required sensed information the sink node is interested in and is supported by the network. The interest message also contains the reporting interval of the interest data and the interest expiry time. The reporting interval specifies the interval at which the sensed data should be forwarded to towards the sink node. The interest expiry time specifies a time when the sensed data should no longer be transmitted to the sink. The usage of attribute value pairs to explicitly specify the interested data ensures that no undesired data is forwarded to the sink and reducing irrelevant data transmissions.

Directed diffusion consists of several processes:

1. *Interest propagation:* The sink node periodically broadcasts the interest messages to its neighbour nodes. The first broadcast can be seen as an exploratory request to determine if there are nodes in the network that can provide the interest data. Every node stores an interest cache containing the information about the interest and the neighbour node from which the interest was received. Interest aggregation can only be achieved if interest types for the application are distinct. The interest message is rebroadcasted at every receiving node to propagate to the specified network segment.
2. *Gradient establishment:* As the interests are propagated through the network, a gradient is computed and stored for a unique interest on each of the nodes. The gradient specifies a value and a direction in which to forward data messages with the lenience of adopting different semantics to the value. The usage of the

gradient enables the protocol to route interest data towards the sink node. Data transmission from the source nodes as reply to the exploratory interest message will now commence. Interest data will be forwarded towards the sink using the gradient for the interest on each node. Upon receiving a data message, a node will determine if the message is new from the locally stored data cache. If the data has not been received, it will follow the protocol's data aggregation policies to eventually forward the message to the next node in the path towards the sink node. When the sink receives the exploratory interest data, it will resend an interest message along the same path of the data message to reinforce the path towards the sink. The reinforcement interest message now contains a higher data rate and extended interest expiry time.

3. *Data propagation:* Once the gradient of a source node has been reinforced, it will send the interest data at the lower interval specified in the reinforcement interest message.

Directed diffusion can be summarized by Figure 3.3 which was redrawn from [23].

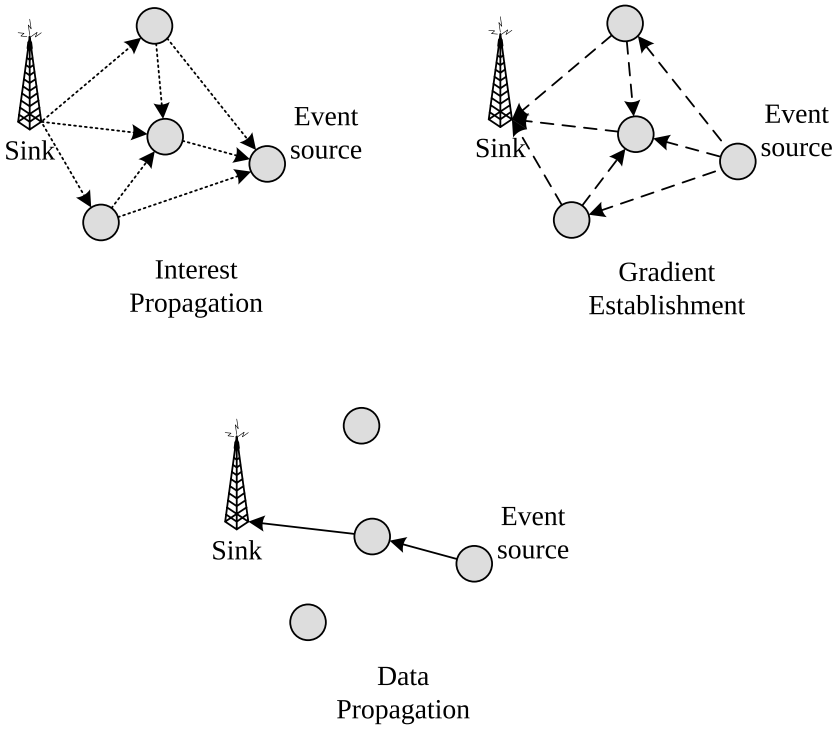

Figure 3.3 Directed diffusion processes

Directed diffusion is not optimised for energy efficiency due to the flooding required to propagate the interests and establish gradients on each node. This protocol may also require high amounts of memory to store the interest gradients and received messages.

3.2.5 Simple Energy Efficient Routing (SEER)

Simple Energy Efficient Routing (SEER) [2] proposes an event driven source initiated routing protocol to minimise energy consumption of each node to prolong the functional lifetime of the network during normal routing activities. SEER attempts to provide a scalable solution on a deployment of homogeneous sensor nodes due to the computational simplicity in the routing logic. No node is aware of the global network structure and only requires local network knowledge of its neighbour nodes. Every node will store information of the neighbour nodes such as node address, available energy of the node and hop count to reach the sink node.

The operation of SEER is based on the following phases:

1. After the nodes have been deployed the sink nodes initiates the *network discovery* phase. The sink broadcasts the initial broadcast message containing a unique sequence number. Upon receiving the initial broadcast message, the node will store the sequence number and update the neighbour node's information. Every successive broadcast message sent from the sink node will have an incremented sequence number to uniquely identify each broadcast phase. If the received broadcast sequence is a new sequence, the node will increment the hop count, current node energy levels and new source address in the message header and broadcast the updated message. If a node receives a broadcast message with a hop count lower than its own hop count, the receiving node will update its own value. The hop count of a node is indicative of the shortest route to reach the sink node. Once the broadcast has reached all the nodes the discovery phase has completed.
2. SEER relies on source nodes initiating data messages from sensed data. The data messages are forwarded towards a sink node and the network enters the *data transmission* phase. Routing logic is based on hop count and available energy levels of the neighbour nodes. The inclusion of available energy levels in decisions ensures that a single path to a sink node does not prematurely

deplete nodes in that path. Before data messages are sent to the next neighbour node, the transmitting node will decrease the neighbour node's energy levels in an attempt to modify routing paths to distribute network load. SEER allows for a distinction between critical and non-critical messages. Non-critical messages are routed on a single path towards a sink node. Critical messages are transmitted over two different paths towards a sink node. The neighbour selection process implemented by SEER (adopted from [2]) can be summarised by Figure 3.4. The node receiving a data message will update the neighbour information from which the message was transmitted and then forward the message using the same routing logic as for the data originator.

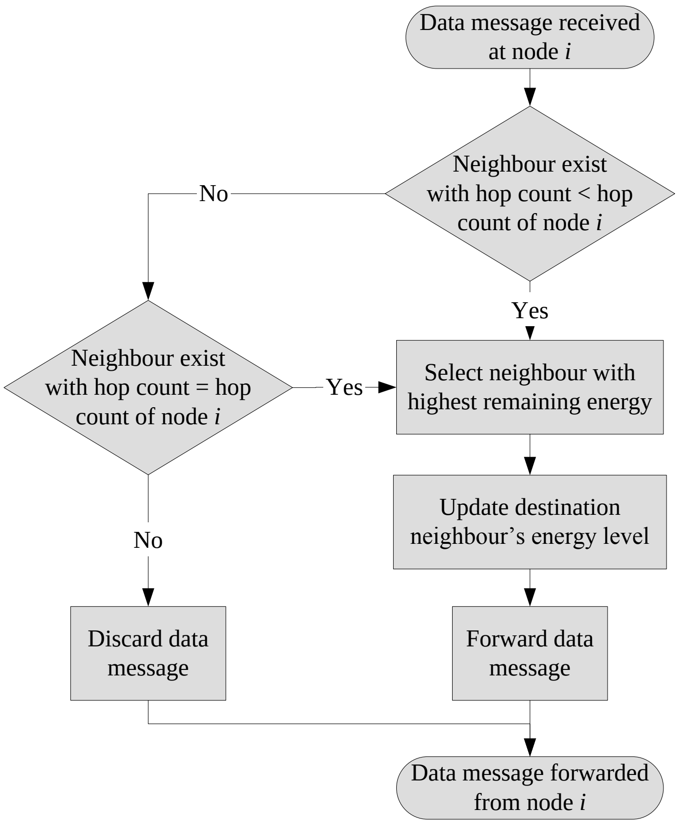

Figure 3.4 SEER neighbour selection process

3. Although SEER implements a mechanism to be neighbour node energy aware, that same neighbour nodes may receive data from other source nodes. *Energy updates* are broadcast to all neighbours once the energy levels fall below a predetermined threshold value.

4. In order to compensate for unknown network topology changes such as node failures and mobile nodes, SEER performs *network maintenance* by periodically initiating a broadcast message from the sink node with an incremented sequence number.

Some of the identified disadvantages of SEER are:

1. SEER proposes a single path to the sink node for non-critical messages. This disadvantage implies that should a single node in that path fail, all messages traversing that node will be lost. The changed topology will only be rectified after a broadcast has been re-initiated from the sink node. The same applies to mobile nodes created as part of the path to the sink and then moving beyond the transmission distance of the specified path.
2. The dual path implementation of critical messages could theoretically be very energy consuming should the amount of critical messages significantly increase. This is rather a trade-off between reliability and energy preservation.

The operation of SEER can be summarized by Figure 3.5 (adopted from [3]).

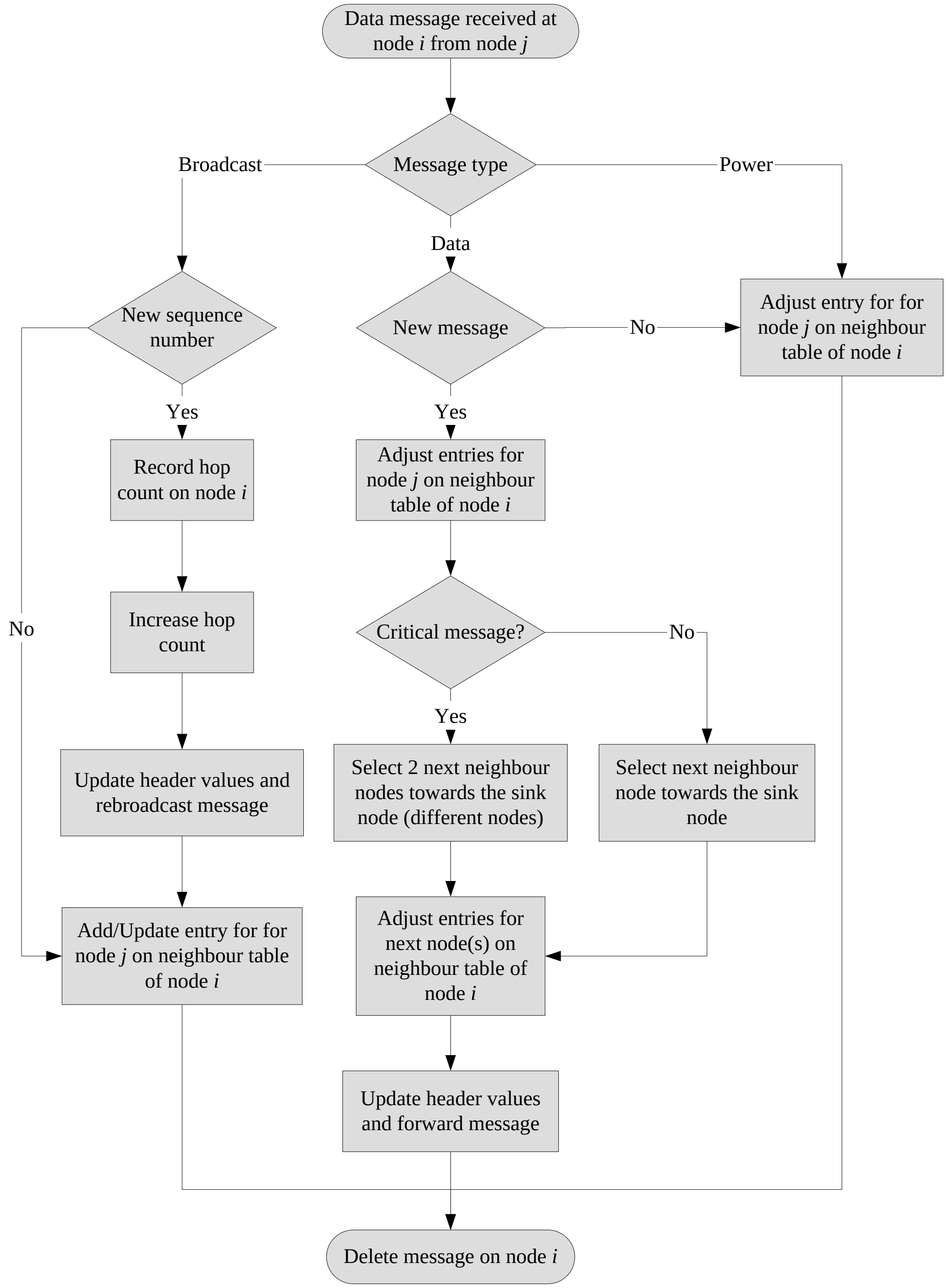

Figure 3.5 Message receiving process flow for SEER

3.2.6 Advantages of flat routing protocols

Topological changes are limited to neighbouring nodes. This local network knowledge at each node and the fact that each node functions equally promotes *protocol scalability* and *computational simplicity*. There is no requirement for complex calculations to determine or manage a hierarchical network topology.

Flat network topologies treat each node equally to simplify deployment strategies to the use of *homogeneous nodes*. There is no requirement of including and deploying specific nodes with higher specified functionality and resource capabilities.

3.2.7 Disadvantages of flat routing protocols

The *sink node may overload and cause latency in message delivery* in networks with a high density of source nodes surrounding a sink node. The high volumes of data messages traversing the sink neighbour nodes may cause *hot spots* and power sources to fail prematurely.

Low density networks may cause *separation and even isolation* of parts of the network should a single critical node fail in a deployment of homogeneous nodes. In this case a suggestion would be to add additional sink nodes.

3.3 Hierarchical Routing Protocols

Hierarchical routing protocols on the other hand treat nodes differently in terms of functionality and purpose in the WSN. Sensed data messages are aggregated at certain nodes and forwarded to the next aggregation node in the path towards the sink node. The aggregator nodes are predetermined nodes or dynamically assigned by the routing protocol. The purpose of the aggregator nodes may be for routing purposes only or both routing and data sensing. The majority of the nodes is still seen as sensing nodes and will propagate the sensed data originating in each node towards the aggregation node.

Literature accepts the following established hierarchical routing protocols: LEACH (Low Energy Adaptive Clustering Hierarchy [24]) and variants of it [25] and [26], PEGASIS (Power-Efficient GAthering in Sensor Information System [27]), EDC

(Event-driven Cluster Routing [28]), HEAR-SN (Hierarchical Energy-Aware Routing for Sensor Networks 0), BATR (Balanced Aggregation Tree Routing [30]) and TEEN (Threshold sensitive Energy Efficient sensor Network protocol [31]). The seminal contributions for hierarchical network architectures are summarised in the following sub-sections.

3.3.1 Low Energy Adaptive Clustering Hierarchy (LEACH)

Low Energy Adaptive Clustering Hierarchy (LEACH) [24] was ground breaking research work in terms of hierarchical network topologies. As with flat network topologies, LEACH is based on the deployment of homogeneous nodes and assumes that the sink node is stationary. The self configuration of clusters groups a selection of nodes that are geographically close to a cluster head into a logically separate network. This mechanism assumes that source nodes will always be a single hop away from the cluster head. Data message aggregation is performed within each cluster at the cluster head. The operation of LEACH assumes that cluster heads will always be one transmission away from the sink node and will forward the aggregated messages directly to the sink node. This is a high-power transmission and will drain node power factors more than normal source-to-cluster head transmissions. LEACH proposes a mechanism to randomly select and rotate cluster heads to distribute aggregation and sink transmission load across the network.

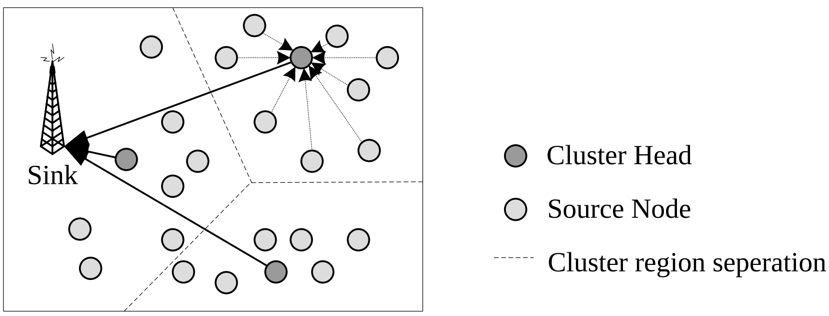

Figure 3.6 LEACH Cluster configuration and transmission

The operation of LEACH is broken up into time intervals where each time interval will consist of two phases: setup and steady state. The mentioned time interval is determined prior to node deployment. At the beginning of every time interval, the network will start with a setup phase and progress on to a data transmission phase as follows:

1. *Setup phase*: Each node decides independently whether or not it should become a cluster head, based on the predetermined percentage of cluster heads and the amount of times that that node has fulfilled the function of cluster head. Assume node n will determine its cluster head status by selecting a random number between 0 and 1 and compare it to a threshold value *T(n)* for the round. If the chosen random number is less than *T(n)*, then node *n* become a cluster head for the specific round. The threshold value [24] is calculated as:

$$T(n) = \frac{P}{1 - p(r \, mod \, \frac{1}{P})} \quad \text{if } n \in \text{G} \qquad (3)$$

   Where *P* is the desired percentage of the total amount of nodes to function as cluster heads at any given moment, *r* is the current round and G is the set of nodes involved in the cluster head election that has not been cluster head in the last $\frac{1}{P}$ rounds. Empirical research shows that the optimal percentage of cluster heads average at 5% of the total amount of nodes in the network.

   Every node that elected itself as a cluster head will now broadcast an advertisement message to all its neighbours. Upon receiving the broadcast message, source nodes will decide on a cluster head to join based on the highest received signal strength of the broadcast message. The receiving node will notify the cluster head of its joining. The wireless transmission during the setup phase utilises a Carrier Sense Multiple Access (CSMA) MAC protocol. Each cluster head creates a Time Division Multiple Access (TDMA) schedule to assign a time slot for every joined source node and will broadcast the schedule to all the cluster nodes. The timeslot assigned in the TDMA schedule will allow every node to transmit data messages without interference from other node transmissions.

2. *Steady state phase:* After all nodes have been informed of the cluster TDMA schedule, it will forward data messages to the cluster head if they have data available. Each source node may disable its radio outside the period of that

node's transmission timeslot in the TDMA schedule, as an additional method of conserving energy. After the time schedule of all nodes has elapsed, the cluster node aggregates the data into a single message to be transmitted to the sink node. The duration of the steady state phase is much longer than the setup phase to conserve energy.

The implementation of LEACH does have its disadvantages:

1. The assumption that any node in the network can assume the role of a cluster head implies that all nodes should have the capacity to transmit messages directly to the sink. Extensions to LEACH to utilise multi hop clusters propose to solve this disadvantage.
2. LEACH requires support for multiple MAC protocols to fulfil the functionality of the protocol.
3. Maximum Energy Cluster-Head (MECH) [25] is based on LEACH and identifies deficiencies in the cluster formation mechanism of LEACH. MECH proposes alternative solutions to overcome the following drawbacks of LEACH:
   a. The geographic distribution of nodes is not considered and cluster heads may be in one small location. This means that some source nodes may have to transmit across vast distances to reach that node's cluster head or may not even be within transmission distance of any cluster head.
   b. Source nodes are not distributed evenly among cluster heads. The cluster heads of clusters containing more source nodes will drain the supply of that head much faster.

The operation of LEACH can be summarized by Figure 3.7 (adopted from [3]).

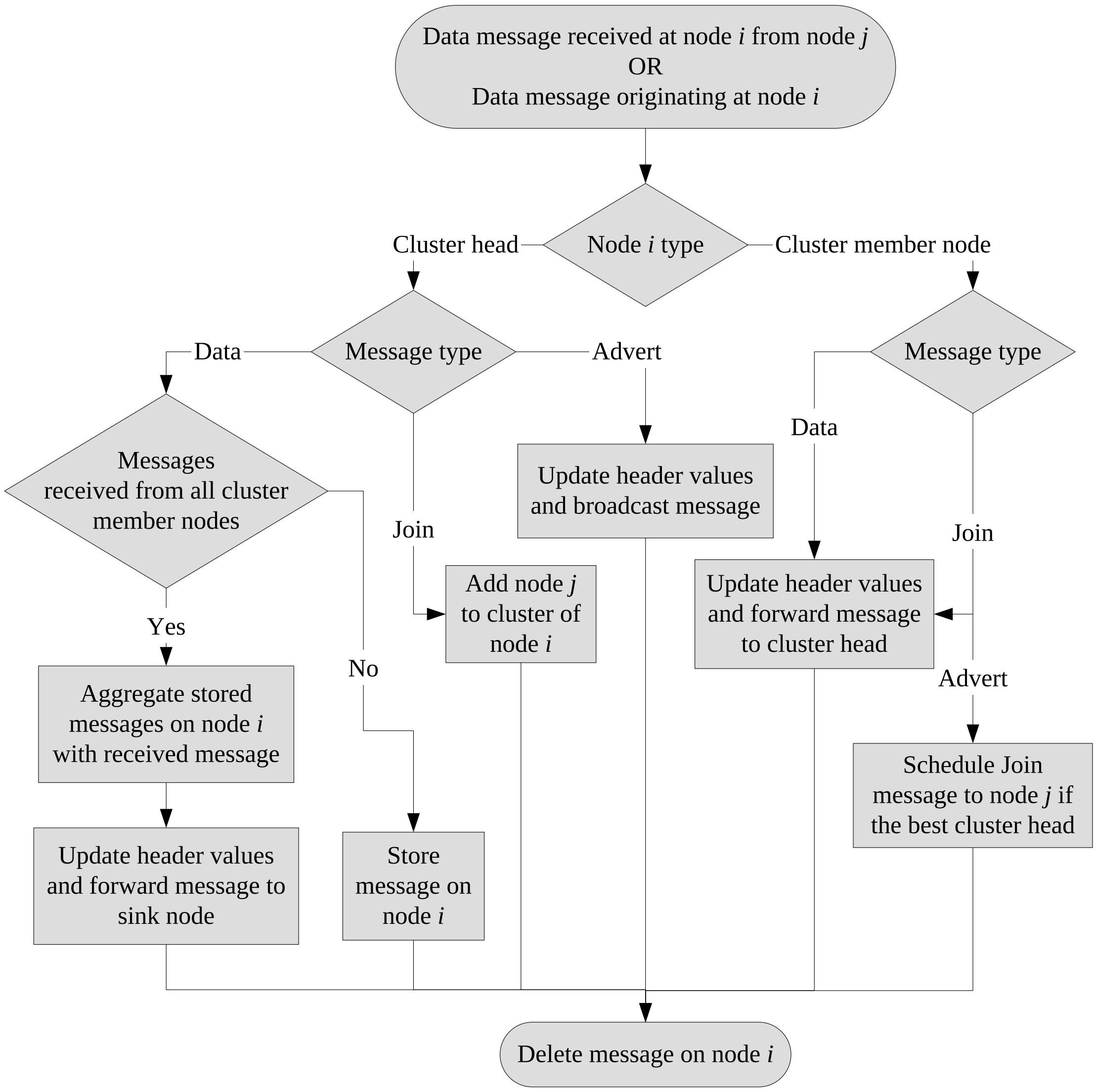

Figure 3.7 Message process flow of LEACH

3.3.2 Power-Efficient GAthering in Sensor Information System (PEGASIS)

Power-Efficient GAthering in Sensor Information System (PEGASIS) [27] proposes a chain based protocol as an enhancement of LEACH. As with LEACH, PEGASIS is based on the deployment of homogeneous nodes, assumes that the sink node is stationary and the energy consumed during message transmission is dependant on the distance of transmission. PEGASIS assumes that every node in the network has global network knowledge and can create a chain of nodes along which messages will be forwarded in an attempt to reach the sink. The protocol assumes that all source nodes have the same energy level and will transmit messages with the same amount of power. The closest neighbour is determined by using the signal strength to determine the distance to each neighbour. Chaining of a specific node will only be to the closest neighbour for that node.

The operation of the protocol is based on the principle that only one node in the chain (named the leader for the round) will transmit aggregated data messages to the sink node in a specific round of data transmission. Each node will only transmit data messages to that node's closest neighbour in the chain. The receiving neighbour will aggregate the received message with its own data and forward the aggregated message to the next node in the chain towards the leader. Every node in the chain will have a change in successive rounds to fulfil the role of transmission leader thus distribute the high energy consumption transmissions to the sink node among the whole network.

The improvements and changes in PEGASIS attempts to eliminate the overhead incurred by LEACH to setup the dynamic cluster formations in every round and to minimise the transmission distances that non-leader nodes have to transmit and limit the amount of transmissions using data aggregation.

Disadvantages of PEGASIS are:

1. The protocol operation may require high amounts of local memory to store the global network knowledge.
2. Global network knowledge does not contribute positively to protocol scalability.
3. Excessive delays are introduced to the most distant source nodes from the transmission leader.
4. The single leader may become the bottleneck in high density networks and may fail during a single transmission round. The data messages for this round will never be forwarded to the sink for the network in total.

3.3.3 Threshold sensitive Energy Efficient sensor Network protocol (TEEN)

Threshold sensitive Energy Efficient sensor Network protocol (TEEN) [31] proposes a reactive routing protocol specifically applicable to time critical applications. TEEN adopts the principle of hierarchical clustering. Sensor nodes are grouped together that is geographically close to each other with one common cluster head. For the purposes of this section these cluster heads are seen as first layer cluster heads. The first layer of cluster heads collect data messages from the sensor nodes within its cluster, aggregate the data messages and forward the aggregated message to a higher layer. The higher layer could be a sink node or a next layer of cluster heads that cluster the first layer of cluster heads together. This process is repeated to form a hierarchical cluster topology with the highest layer cluster heads transmitting data messages directly to the sink node. Figure 3.8 (adopted from [31]) illustrates a subset of a typical hierarchical network topology.

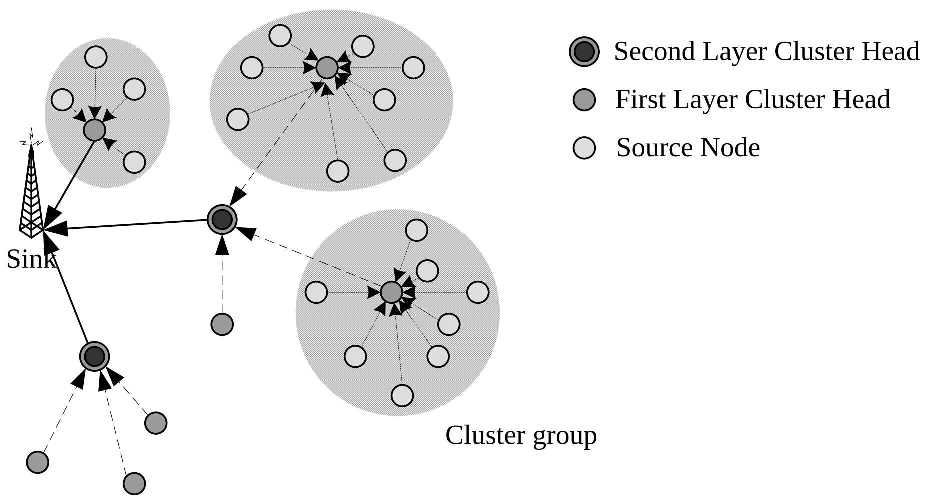

Figure 3.8 Typical hierarchical clustering network topology

After every cluster topology change the cluster heads broadcasts a message to its members containing the interested attributes with its hard and soft thresholds:

1. An *attribute* is the set of physical parameters the user of the system is interested in and reflects the data to be sensed.
2. A *hard threshold* is the absolute value of the attribute being sensed. If this value is exceeded the sensor node should switch its transmitter on and forward the data message to the cluster head. The hard threshold minimises the transmissions by only forwarding data within the range of interest.
3. A *soft threshold* is a small value of the attribute being sensed to trigger the sensor node should switch its transmitter on and forward the data message to the cluster head. The soft threshold minimises the transmissions by only forwarding significant attribute value changes.

The sensor nodes continuously sense the environment and will only forward data messages to the cluster node once the hard threshold has been exceeded. After this stage, the sensor node will only forward data if the change in the sensed attribute exceeds the soft threshold and still exceeds the hard threshold.

The main disadvantage of TEEN is that if the hard threshold is never reached, no data messages will be forwarded towards the sink node. TEEN is not suitable for applications requiring periodic data updates. This is however not a disadvantage but merely a projection of specific application implementations.

Adaptive Threshold sensitive Energy Efficient sensor Network (APTEEN) [32] proposes an extension to TEEN to mitigate the hard threshold problem described above. APTEEN proposes a mechanism to periodically transmit data messages in addition to the responsive threshold triggered data messages. In addition to the information broadcasted to the cluster members in TEEN, APTEEN will include a count time to reflect the maximum time between two successive initiations of data messages. If a source node is not triggered by the hard and soft threshold values within the count time period, the node will sense the specified attribute and forward the data message to the cluster head.

The disadvantage of both TEEN and APTEEN is the overhead and complexity of cluster formation, maintaining the threshold based functions and attribute based queries.

The energy consumption of both TEEN and APTEEN is well planned to minimise the energy hungry actions such as message transmission. Simulation results in literature show that both TEEN and APTEEN outperform LEACH in terms of energy conservation. The main benefit of both TEEN and APTEEN is that the user of the system can modify the sensed data requirements with every cluster topology change.

3.3.4 Periodic, Event-driven and Query-based Routing Protocol (PEQ)

Periodic, Event-driven and Query-based Routing Protocol (PEQ) [33] proposes a query based and event driven routing protocol to minimise energy consumption, minimise message delivery latency, improve reliability and provide fast path recovery due to topological changes. The design of PEQ is based on a deployment of homogeneous sensor nodes. No node is aware of the global network structure and only requires local network knowledge of its neighbour nodes such as hop count to reach the sink node. The hierarchical properties of PEQ are manifested by the concept of a tree structure, termed a hop tree for the purposes of this section. The hop tree is used as a message propagation mechanism towards the sink node and the aggregation structure is similar to the branching concept used by Hybrid Energy Aware Routing (HEER). PEQ introduces a mechanism to switch to a fast recovery by minimising data transmissions among neighbouring nodes mode in the case of network failures.

PEQ employs a mechanism based on publishing and subscribing to events [33]. Source nodes may forward event notifications to the sink node. The sink will express interest in certain events by either subscribing to received event notifications or by flooding the network with subscription request. When the source node detects the subscribed events, the node will forward those data messages to the sink. This mechanism reduce the amount of overhead in transmitted messages by only forwarding the required data and discarding data of no interest to the sink node.

The operation of PEQ can be summarised by the following phases:

1. The *hop tree configuration* is initiated by the sink node by performing a restricted broadcast of a configuration message. The configuration message contains the hop count value (starting with a hop count of zero), time stamp and the source address. The receiving node will compare the received hop count to the locally stored hop count. If the message hop count is less than the receiving node’s hop count it will locally store the message hop count. The node will increment the node hop count and update the new source address and hop count value in the message header and broadcast the updated message to its neighbours. If the message hop count is more than the receiving node’s

hop count, it will discard the message. This process is repeated until all the nodes in the network has received the configuration message and will complete the hop tree configuration phase. Each node will locally store the source node's information to be used for routing purposes.

2. After the initial network configuration, the sink may either flood the network with each interest or wait for event notifications and start the *subscription propagation* phase. Each node contains a locally stored subscription table containing an entry for every different subscription. Every subscription will contain the interest properties, the source node from which the interest was received with the lowest hop count and the sink node that subscribed to the interest. This information is used to forward the sensed data message back to the correct source.
3. When a source node senses a phenomena it will verify if a sink node has subscribed to that event. The *event message delivery* to the sink node propagates though the network towards the subscribed sink node, using the hop tree and subscription data configured on each node. Should no sink be subscribed to an event, the source node will forward an event notification to the sink nodes that follow the same route as event messages. Upon receiving an event notification, the sink node may choose to subscribe to the event by forwarding the subscription back to the source node, to repeat the event delivery process.
4. PEQ implements a message acknowledgement mechanism for every message forwarded as a means *repair broken paths*. Should a sending node not receive an ACK message after a predetermined time-out period, that node will assume that the destination node is no longer available. The sending node now attempts a path recovery by broadcasting a SEARCH message. Upon receiving a SEARCH message, the receiving node will reply with a configuration like message with its hop count and address. This will update the recovering node's routing tables.

A major advantage to the implementation of PEQ is the ability to forward specific data to a specific sink node. This ability could increase network protocol scalability by utilising multiple sink nodes and avoiding hot spots around a single sink node.

PEQ suffers from the single path to a specific sink problem. If a node is unable to recover from a node failure towards the sink, that node and possibly all nodes further away from the sink on the path will become isolated from the rest of the network. Isolation may have been avoided if nodes further away from the sink could use a different path by avoiding the node that cannot recover its path to the sink. An additional disadvantage to PEQ is the additional message transmissions introduced by the usage of an ACK message for every event message forwarded.

Clustering PEQ (CPEQ) [33] is summarised by [34] as a cross between LEACH and PEQ. Cluster formation and path configuration towards the sink node is similar to a hierarchical version of LEACH where cluster heads form a layer of multi hop nodes towards the sink node. Subscription propagation and event message delivery as proposed by PEQ will follow the same routes as source initiated data messages for a hierarchical version of LEACH would.

3.3.5 Advantages of hierarchical routing protocols

In general, hierarchical protocols promote the use of *data aggregation* and data fusion. This could include some additional functionality to eliminate duplicate messages. Data aggregation reduces network overhead and the amount of messages being transmitted, thus *reducing energy consumption*.

Some hierarchical network topologies simplify deployment strategies to the use of *homogeneous nodes*. There is no requirement of including and deploying specific nodes with higher specified functionality and resource capabilities.

The power consumed in clustering approaches is much lower to setup clusters rather than the network as a whole. *Power consumption is localised* to the cluster. Data messages also traverse much less nodes to reach the cluster heads. In general hierarchical routing protocols consume much less energy in the network as a whole.

3.3.6 Disadvantages of hierarchical routing protocols

Some hierarchical network topologies require the inclusion and deployment of specific nodes with much higher specified functionality and resource capabilities. This will influence the cost of the network in terms of deployment as well as manufacturing. Deploying nodes with *differing functionality* will complicate any attempt to rotate cluster heads.

To overcome this disadvantage some hierarchical protocols propose the use of *homogeneous nodes*. Routing strategies that require the cluster heads to communicate directly to the sink node, or over extra long distances will significantly consume more energy. To overcome these cluster head *hot spots*, protocol implementations utilise dynamic cluster head selection and rotation.

The *additional overhead* to randomise cluster heads in homogeneous network, add additional computations and *complexity*. Additional information needs to be gathered by the sensor nodes to setup and initialise the network, thus introducing additional overhead and transmitted messages for initialisation. Network initialisation could *consume much more energy* than flat network topologies.

The selection of cluster heads is determined by the protocol operation. *Scalability* could be influenced negatively with increasing network sizes and the requirement for an increased amount of cluster heads.

3.4 Hybrid Routing Protocols

Hybrid routing protocols utilises a combination of hierarchical and flat routing principles. Sensed data messages are aggregated at certain nodes and forwarded to the next node in the path towards the sink node. The aggregator nodes are predetermined nodes or dynamically assigned by the routing protocol. The purpose of the aggregator nodes are for routing purposes and data sensing. The majority of the nodes is still seen as sensing nodes and will propagate the sensed data originating in each node towards the aggregation node.

3.4.1 Hybrid Energy Efficient Routing (HEER)

Hybrid Energy Efficient Routing (HEER) [3] proposes an event driven source initiated routing protocol to minimise energy consumption of each node to prolong the functional lifetime of the network during normal routing activities. The design of HEER aims at computational simplicity, limiting transmission, energy awareness and limited tolerance to network topology changes. HEER was designed by expanding the basic principles of SEER and transforming the protocol into a semi hierarchical network protocol. No node is aware of the global network structure and only requires local network knowledge of its neighbour nodes. Every node will store information of the neighbour nodes such as node address, available energy of the node and hop count to reach the sink node. Additional information is computed, stored locally and used for routing purposes but this will be elaborated in the remainder of this section.

The operation of HEER is based on the following phases:

1. As with SEER, the sink node initiates the *network discovery* phase after the nodes have been deployed. The sink broadcasts the initial broadcast message containing a unique sequence number. Upon receiving the initial broadcast message, the receiving node will store the sequence number and update the neighbour node's information. Additionally the protocol dictates the calculation of an administrative distance value for each neighbour. This administrative distance value is reflection of transmission distance rather than the geographical distance and should take care of environmental interferences. The "admin" value is used by the protocol to determine the best next hop towards the sink node and is calculated by:

$$Admin = \frac{(H_c)^d}{P_n} \tag{4}$$

Where H_c is the hop count of the neighbour towards the sink node, *d* is the distance of the neighbour node to this node and P_n is the current energy level of the neighbour node. Every successive broadcast message sent from the sink node will have an incremented sequence number to uniquely identify each broadcast phase. If the received broadcast sequence is a new sequence, the node will increment the hop count, current node energy levels and new source address in the message header and broadcast the updated message. After every broadcast message received, the receiving node will select the node in its neighbour table with the highest admin value as that node's branch head. The concept of a branch head can be visualised by a tree structure joining at branches to eventually have the sink node as the last branch head. Every broadcast message received will update the branch head to ensure that the best branch head is always selected. After 5 minutes from the initial broadcast, the receiving node will send a join message to the branch head. The branch head will in turn broadcast an information message to inform all the branches of its current status. Once the broadcast has reached all the nodes the discovery phase has completed

2. HEER relies on source nodes initiating data messages from sensed data. The data messages are forwarded towards a sink node and the network enters the *data transmission* phase. Source nodes forward data messages to branch head nodes in an effort to reach the sink node. HEER allows for a distinction between critical and non-critical messages. A branch head will wait for non-critical data messages from all of that node's branches, aggregate all of the messages and forward the aggregated message to its branch head. Critical messages are forwarded immediately by the branch heads and does not allow for message aggregation. A branch head receiving a critical message, will acknowledge the receipt by sending an ACK message back to the sender node in an effort to improve reliability. Received data messages are used to update the transmitting node's information in the receiving node's neighbour tables. Before data messages are sent to the next neighbour node, the transmitting node will decrease the neighbour node's energy levels in an attempt to modify routing paths to distribute network load.
3. Although HEER implements a mechanism to be neighbour node energy aware, that same neighbour nodes may receive data from other source nodes. *Energy updates* are broadcast to all neighbours once the energy levels fall below a predetermined threshold value. Branch heads will remove the failing branch nodes and branch nodes will elect new branch heads in the case of a failing brand head. In order to compensate for unknown network topology changes

such as node failures and mobile nodes, HEER performs *network maintenance* by periodically initiating a broadcast message from the sink node with an incremented sequence number.

The operation of HEER however can be summarized by Figure 3.9. A critical analysis of the implementation, design principles, advantages and disadvantages of HEER is discussed in 5.2.

3.4.2 Advantages of hybrid routing protocols

The advantages of hybrid routing protocols are in short a combination of the flat and hierarchical advantages. Hybrid protocols promote the use of *data aggregation* and data fusion while remaining *computationally simpler* than hierarchical protocols. Data aggregation and computational simplicity reduces network overhead and the amount of messages being transmitted, thus *reducing energy consumption*.

Hybrid network topologies simplify deployment strategies to the use of *homogeneous nodes*.

3.4.3 Disadvantages of hybrid routing protocols

Hybrid network protocols aimed at networks with only static nodes have *little or no tolerance to node mobility*. Although network re-initialisation phases are executed in an effort to compensate for network topology changes, node mobility may impact path selection towards the sink negatively.

Network initialisation could also *consume much more energy* than flat network topologies.

Scalability could be influenced negatively with increasing network sizes and the requirement for an increased amount of branch heads causing *hot spots* in return.

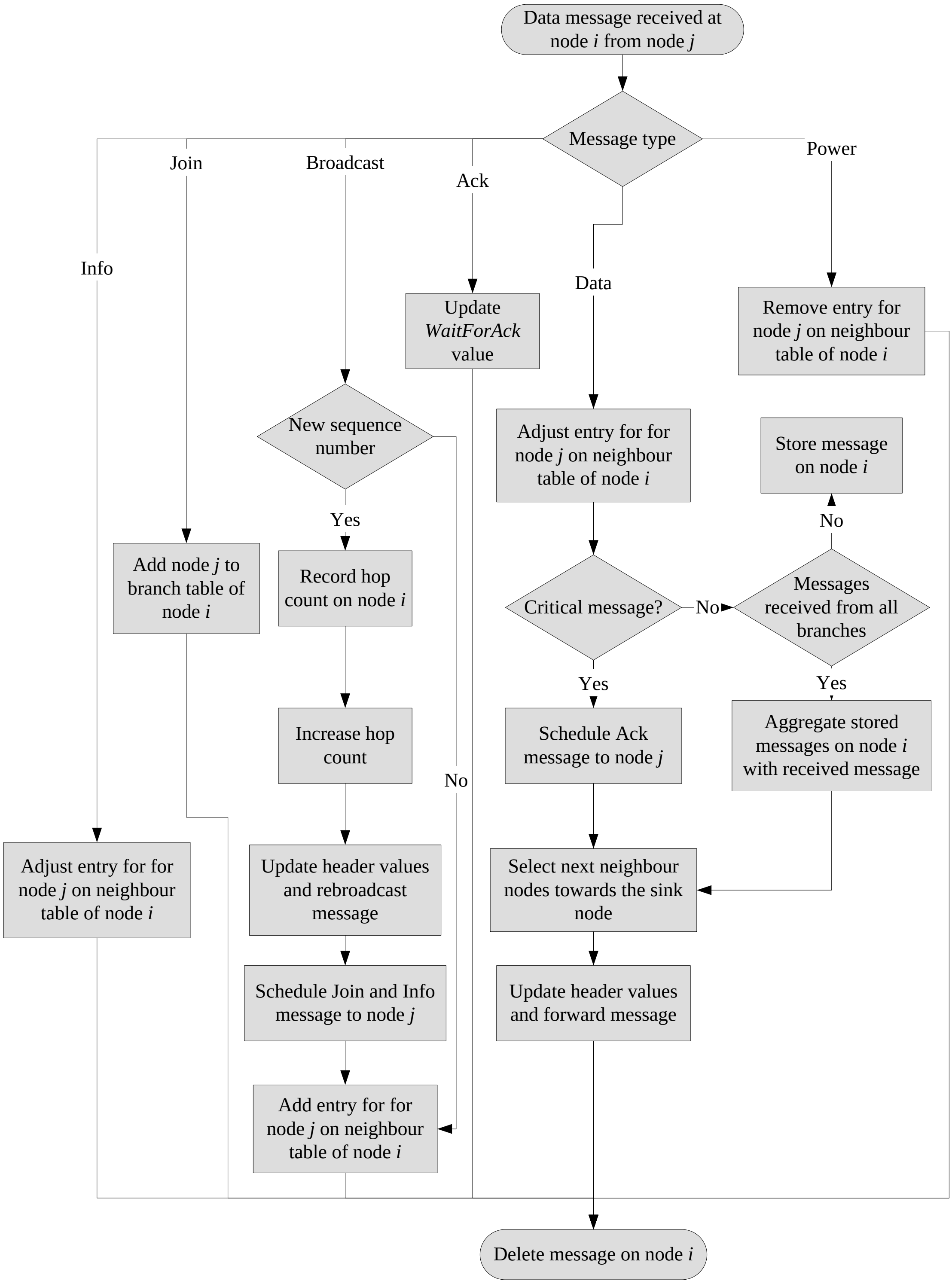

Figure 3.9 Message receiving process flow for HEER

3.5 Mobility aware routing

An interesting factor that influences the performance of routing protocols is node mobility. Most of the current protocols assume that the sensor nodes and sink nodes are stationary. However, there might be situations such as battle environments where the sink and possibly the sensors need to be mobile. Mobile nodes add additional complexity to network topology changes that routing protocols have to manage. Frequent updates of the position of the mobile nodes and the propagation of the updates through the network may excessively deplete the energy levels of nodes. Routing algorithms are needed in order to handle the overhead of mobility and topology changes in such an energy constrained environment, while still maintaining the full functionality of the network. Design considerations for networks containing mobile nodes include reliable packet delivery, energy efficiency and low message delivery latency.

3.5.1 Mobility principles

The wireless nature of WSN presents the ability to support mobile entities within a WSN. Mobility can be present in a WSN in three forms:

1. *Node mobility*: All or a subset of the source nodes are mobile in this context and depends on the intended application. Node mobility implies that the network topology will change with time if one or more source nodes move away from predetermined message paths to a sink node. There is a trade off between the energy consumed to maintain full network functionality and the tolerance of routing protocols to handle topology changes due mobility of source nodes.
2. *Sink mobility*: Sink nodes may be mobile in an attempt to reduce energy consumed to propagate a data message across the network to a sink node. The sink node may not even be part of the network, for example a user collecting data using a Personal Digital Assistant (PDA). The challenge is to design the routing protocol to be able to facilitate requests from a mobile sink node that will change its geographic position over time and to route data message to such a mobile sink node.

3. *Event mobility*: The phenomena required to be sensed by the source nodes may be mobile and moving between sensor nodes. This does not seem challenging at all if you consider that it is still just another phenomenon that needs to be sensed. However routing protocols such as TEEN (described in 3.3.3) may request specific attributes from source nodes in specific regions. These types of protocols need to be aware of event mobility and is tightly bound to the application. This type of mobility is not the focus of this research and will not be considered further.

3.5.2 Mobile Data Collectors/Periodic, Event-driven and Query-based Routing Protocol (MDC/PEQ)

Mobile Data Collectors/Periodic, Event-driven and Query-based Routing Protocol (MDC/PEQ) [10] proposes a low message delivery latency and mobile gathering mechanism for delay sensitive applications. MDC/PEQ proposes the deployment of mobile data collectors (MDC) that moves among the sensor nodes while collecting sensor data from the sensor nodes. MDC is utilised in an attempt to alleviate high traffic loads and message bottlenecks near static sink nodes and to reduce latency in data message propagation to the sink node. The MDC entities may be normal sensor nodes or some higher function device such as a personal digital assistant (PDA) or cellular phone.

The author provides a valid statement that mobile sink nodes will also enable the routing protocol to collect data messages from source nodes disconnected from the network. As mentioned previously, network disconnection can be caused by geographic location, failing nodes or mobile source nodes moving out of transmission range.

The design of MDC/PEQ is based on PEQ and CPEQ principles and assumes the deployment of homogeneous static sensor nodes with the introduction of mobile data collectors. None of the nodes require global network knowledge to promote network scalability. The intention of [10] is to provide a sensor network for delay sensitive applications and so depends on a hybrid of MDC and at least one static sink node. Every MDC fulfils the role of a cluster head whose sole responsibility is to collect and aggregate data messages while moving among the sensor nodes. The source nodes will forward data messages to the static sink, should a MDC not be in close proximity to the source node initiating a data message. Each node will store a hop count value for both an MDC and the shortest route to a static node.

The operation of MDC/PEQ can be summarised as:

1. *Cluster configuration* to the static sink node is manifested in exactly the same manner as for PEQ and CPEQ while storing hop counts on every node to the static sink node (h_s). Cluster configuration to the MDC is similar to PEQ with the MDC being the root of the hop tree and is initialised with the MDC broadcasting a beacon message. The beacon message contains an MDC identifier, message time-to-live (TTL) and current hop level. Upon receiving a beacon message, the sensor node will update the hop count of the receiving node to the MDC (h_m) using the same algorithm as CPEQ with one difference. When updating h_m the receiving node will take the signal strength of the message from the MDC into consideration.
2. Due to the fact that MDC will cause topology changes due to mobility, MDC/PEQ proposes a mechanism to dynamically perform *route maintenance* and handoff when a MDC moves out of transmission range. A predetermined reception threshold is specified based on the maximum transmission distance. Each node will only register to a single MDC cluster at any given instance in time, The receiving node will update the routing information for that cluster each time a node receives a beacon message from the same MDC, if the signal strength in above the reception threshold. If a source node receives a message from the node's MDC but the signal strength in below the reception threshold, it is a general indication that the MDC is moving away and the MDC will be removed from the routing tables. This node will broadcast a "low signal strength" message to its tree members (or branches as termed in HEER) to indicate that the route to the MDC is no longer valid. This in turn propagate the MDC hop tree of nodes and update the h_m on all the branch nodes away from the MDC. The node does no longer belong to a MDC cluster and will

perform dynamic cluster configuration to the next beacon message received with signal strength above the reception threshold.

3. As mentioned the implementation of [10] relies on a hybrid of MDC and at least one static sink node. *Data message propagation* is manifested by either forwarding the message to the sink node or the MDC. When a data message is initiated by a sensor node or receives a message from another node, that sensor node will base the routing decision on the lowest hop count (h_m or h_s). If the route to the MDC and static sink node has the same number of hops, the source node will forward the message to the static sink node.

The author of [10] show interesting experimental results that on average 80% of sent data messages are forwarded successfully to MDC. The results also indicate that the successful packet delivery ratio is much higher than compared to the static PEQ, and argues that the improvement is due to the minimised hops of message propagation to a MDC. The reduction in the number of hops to reach a destination should positively reduce energy consumption in the network as a whole. The successful packet delivery ratio is however influenced negatively if the movement speed of the MDC increases.

3.5.3 Advantages of mobility aware routing protocols

The main advantage of mobility aware routing protocols is its ability to *support static as well as mobile nodes* in the network.

The mechanisms introduced to offer this support attempts to *avoid network separation* due to mobile nodes while promoting high message *delivery success ratios* at the sink node.

3.5.4 Disadvantages of mobility aware routing protocols

The mechanisms utilised to provide mobility support may be detrimental to the protocol's *network scalability* with increasing network sizes and/or a higher fraction of mobile nodes in the network.

Mobility support adds additional *computational complexity* in order to determine and predict node movement and data collection strategies. The added computational complexity adds *additional power requirements* with an increase in the amount of mobile nodes to perform normal routing activities.

Mobile aware routing protocols may assume the deployment of homogeneous nodes or not. Some of the mobile nodes acting as data collectors may require *unlimited power resources*.

3.6 Chapter Summary

The design and improvement of any research project to positively contribute the current body of knowledge, requires a fundamental understanding of the leading and most referenced work available in literature. This chapter summarises the main contributors' research and implementations applicable to WSN routing protocols.

Routing protocols can be classified into flat, hierarchical and hybrid network topologies. The protocols influencing this research are covered at a high level with reference to more detailed documents. The protocols that are used as benchmark and for simulation analysis later in this dissertation are described in more detail. These are Flooding, SEER, LEACH and HEER.

Mobility is covered briefly with the presentation of a protocol providing tolerance to mobility in WSN. The concept of MDC provides the ground work that is adopted to complement this research.

Each of the protocols presented in this chapter does have its own set of challenges and unique approaches, but the intent is to use the advantages and powerful mechanisms to enhance the design of this research. The researcher need to consider and critically evaluate the types and relevance of assumptions made during protocol design and result presentation.

Chapter 4

WSN ROUTING PROTOCOL SIMULATIONS

4.1 Introduction

Lacking availability and additional cost of WSN hardware imposes difficulties on the research, practical implementation and verification of WSN routing protocols. To overcome these difficulties, routing protocols can be modelled in simulation tools to emulate the real world scenarios. Simulation of routing protocols can provide valuable information in terms of protocol validation and optimisation. Simulation environments assist in improving the accuracy and credibility of protocol performance by ensuring that results are repeatable and verifiable by other researchers. Available discrete event simulators can provide a base for researchers to analyse and debug the operation of a protocol design.

4.2 Wireless Discrete Event Simulation

4.2.1 OMNeT++

There are numerous commercial and non-commercial network simulation tools on the market today. OMNeT++ is a community developed, open source simulator available free for academic and non-profit purposes. OMNeT++ is widely used as a simulation platform in the scientific community.

OMNeT++ offers the following advantages:

1. Existing building blocks are modular and expandable to have the ability to create new modules independently and not being limited to existing modules.

2. Several open source simulation models such as IP, IPv6 and mobility are available on the OMNeT++ website [37] . The Mobility Framework [39] is extensively defined and is handled separately in the next section (4.2.2).
3. A network model consists of numerous modules connected by channels and connections. OMNeT++ offers the flexibility to define a hierarchical node topology consisting of multiple modules (with the possibility of sub module nesting) and communication among the modules. This network and node definition is supported by a GUI interface, but can also be defined by a flexible textual description (the NED language). The same format is used by the GUI.
4. Model and simulation behaviour can be defined in any C++ editor and compiled with the OMNeT++ GUI simulation libraries. OMNeT++ also link its GUI libraries to the debugging capability provided in the executables and assemblies of certain C++ compilers. This offers an excellent code debugging and tracing environment during the development and testing phases of network and routing models. It also provides a powerful GUI debugging platform providing inspector windows into user created objects and control over the execution of simulations.

4.2.2 Mobility framework

The Mobility Framework (MF) [39] is an extensively defined network simulation framework for wireless and mobile simulation models in OMNeT++. The core framework implements support for dynamic connection management, node mobility and a basic model for a sensor node.

The MF implements dynamic connection management by adding a global *channel control module* to the simulation environment. The channel control module keeps track of the position of every node within the simulation area to determine connectivity between nodes. The location tracking enables the channel control module to establish connectivity among sensor nodes that are within the specified maximum transmission distance of each other. All sensor nodes that are connected to a transmitting node will receive all data packets from that node, whether it is processed or not. The channel control module will dynamically establish and tear down connection as sensor nodes move within the simulation area, enabling management of connectivity among static and mobile nodes.

The mobility framework presents five basic modules for a wireless sensor node as shown in Figure 4.1 (adopted and redrawn from [39]). These layers and the respective functions can be described by the following:

1. The *network interface card (NIC) layer* provides physical layer functionality such as transmitting, receiving, signal modulation and medium access mechanisms. The MF divides the NIC module into three modules:
    a. The snrEval to compute signal to noise ratios (SNR)
    b. A decider to process the SNR information to determine if a message is correctly received, lost or has bit errors
    c. Medium access control (MAC) to provide network addressing and channel access control.
2. The design and implementation of routing protocols for WSN inherits from the *basic network layer* provided by the MF. The basic network layer module provides the basic communication with the lower layer modules, while the inherited implementation can concentrate on user definitions.
3. The *application layer* is of no consequence for the purposes of this research.
4. The *blackboard module* is used for communication among the modules and subsequently the layers mentioned above. Any information available within the sensor node can be published to the blackboard and retrieved by any other module subscribed to the information.
5. The *mobility module* is responsible for node positioning within the simulation environment and node mobility. This module updates the channel control module with the node position at any given moment.

The MF is soon to be part of a more complete merger project, currently named MiXiM (Mixed simulator) that will combine several OMNeT++ frameworks into a single ready-to-use framework for support of mobile and wireless simulations. MiXiM has however not yet been formally released and is planned to include the following frameworks for OMNeT++ simulations:

1. *ChSim*: wireless channel simulation framework
2. *Mac Simulator*: MAC layer simulation framework
3. *Mobility Framework*
4. *Positif Framework*: localisation algorithm testing framework

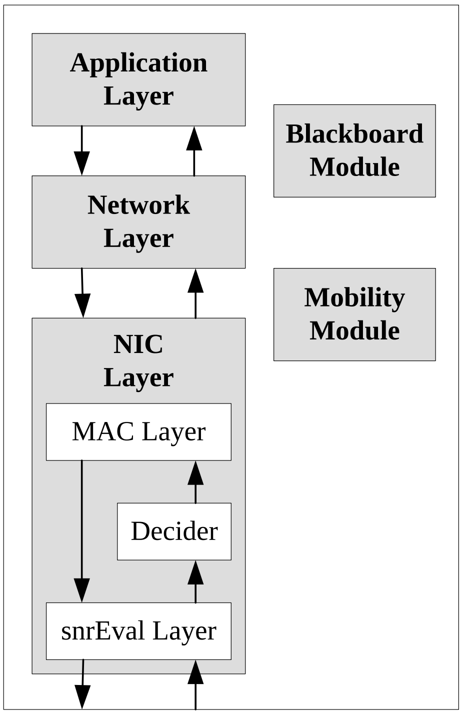

Figure 4.1 Mobility framework structure of a sensor node

4.2.3 Simulation Environment

OMNeT++ has been selected for WSN network protocol implementation and research, due to it's applicability to communication network simulations for both wired and wireless environments. Although implementations for all the protocols mentioned in Chapter 3 is not readily available for OMNeT++, the starting point for this research is HEER and the implementation is available for immediate use and analysis. The versions of the frameworks and development environments used in the protocol implementation and simulations for this research are:

1. OMNeT++ version 3.3 Win32 for Visual C++ 2005 (Express)
2. Mobility Framework version 2.0p3
3. Microsoft Visual C++ 2005 Express Edition

4.3 Simulation assumptions

This dissertation focuses on the network layer routing protocol and certain assumptions have to be made about the rest of the sensor node architecture artefacts and environment variables. These artefacts and variables influence the implementation and simulation results of this research and are discussed the following subsections.

4.3.1 Radio model

The radio model specifies the energy consumed during the transmission and reception of messages over a wireless medium. This research assumes the energy model as a first order radio model as described by Heinzelman et al. in [24] and adopted by both SEER and HEER.

This model assumes a dissipation of $E_{elec} = 50\, nJ/bit$ for the radio transmitter and receiver circuitry and a dissipation of $\varepsilon_{amp} = 100\, pJ/bit/m^2$ for the transmitter amplifier. Assuming that sensor *x* receives a message of *k* bits, the energy consumed by that sensor will be:

$$E_{Rx}(k) = E_{elec} * k \qquad (5)$$

Assuming that sensor *x* transmits a message of *k* bits over a distance *d*, the energy consumed by that sensor will be:

$$E_{Tx}(k,d) = E_{elec} * k + \varepsilon_{amp} * k * d^2 \qquad (6)$$

The parameters used in both the equations above are illustrated by Figure 4.2 (redrawn from [24]).

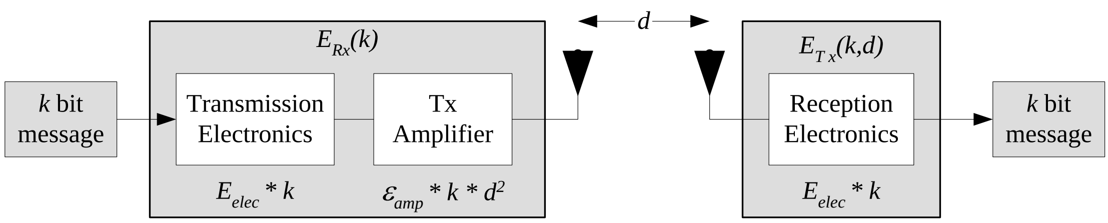

Figure 4.2 First Radio Model

This model specifies a high energy consuming radio model, which emphasises the need to minimise message transmissions. The same maximum transmission distance is used by all the simulation models in this research.

This research assumes that all radios can transmit at the same maximum transmission distance R. Kotz et al. [35] questions this assumption with a real-world data collection exercise and suggests using real-world data as input to simulators to compensate for factors like elevation of nodes, radio interference, etc. This recommendation is however ignored in order to simplify the analysis and comparison to other routing protocols and should be recognised as a weakness of this research.

4.3.2 Channel model

The channel model for the purposes of this research assumes symmetrical radio frequency (RF) links. Symmetry of the radio channel implies that the same amount of energy will be consumed to transmit a message from node *x* to node *y*, as compared to a transmission from node *y* to node *x*. This is not a perfect model but is assumed sufficient for the purpose of routing protocol modelling, as long as the same model is used for all the simulated routing protocols. Kotz et al. [35] empirically attempts to prove that a symmetrical and error free RF link is not valid by collecting transmission and reception data between multiple nodes. Although the results show that link symmetry is not always valid, a fair amount of inter node communication still conforms to this model. The conclusion is that protocol simulations should not assume symmetry and should include a fair amount of link asymmetry as part of the research. This factor should be recognised as a weakness of this research's dependency on symmetrical RF links.

The channel model is further not assumed to be totally error free and does consider channel irregularities and interference to a minor extent. The adopted free-space-loss model in [3] is used in the simulations where the received power (P_r) can be represented as a function of the transmitted power (P_t) by:

$$P_r = \frac{P_t}{({}^{4\pi}/_{\lambda})^2 * d^{\alpha}} \qquad (7)$$

Where λ is the signal wavelength, d is the distance between transmitting and receiving nodes and α is referred to as the path loss exponent. Environmental noise conditions present a huge problem to the errors generated in wireless transmissions. The path loss exponent attempts to cater for environmental interferences.

4.3.3 Medium Access Control Layer

It is assumed that the Medium Access Control (MAC) Layer will handle contention for the radio channel by implementing a Carrier Sense Multiple Access (CSMA) MAC protocol for all wireless communication in the proposed routing protocol. Simulation of other routing protocol for comparison purposes requires the implementation of Time Division Multiple Access (TDMA) MAC protocol as well as CSMA for broadcasting purposes.

4.3.4 Application layer

In an attempt to design a routing protocol adaptable to a wide variety of applications that function on a periodic data reporting mechanism, the application layer is specified very simplistic. The application layer expects all of the sensor nodes to forward the non-critical sensed data at a specific interval (every 15 minutes in the simulations). Although critical data may be forwarded at any instance in time, the simulation setup sends every tenth message as a critical message. This layer has been adopted from the research done in [3].

4.3.5 Node hardware and Mobility

This research assumes the deployment of homogeneous nodes across the simulation area with a single static node. The following properties are assumed for every sensor node:

1. The availability of local node memory to store routing information
2. Limited power source initialised at 5mJ
3. An always on radio transmitter with no requirement to shutdown and reactivate the transmitter
4. Variable power radio transmitter (adopted from the design of HEER)
5. Local knowledge of mobility status that could be determined before node deployment or by means of a movement sensor

The movement of mobile nodes is not orchestrated by the nodes but by means of external influence. The mobility could be manifested by attaching a sensor node to a moving vehicle or to a person for that fact. The movement pattern used for simulation is random and the movement speed may be modified in the middle of a simulation run. Mobility is one of the main research topics in this dissertation and all protocols are evaluated with a selection of different percentages of mobile nodes and speeds.

4.3.6 Cross layer communication

The cross layer design of HEER has been adopted for this research and protocol operation depends on communication among the layers. Support for node mobility requires intimate knowledge of received message signal strength. The signal strength is communicated directly to the routing layer in an effort to implement seamless connection management of mobile nodes moving through the simulation area.

4.4 Chapter Summary

Simulation of wireless sensor network routing protocols is very important to further the current body of knowledge. This chapter introduced the simulation environment used for this research as first choice due to the vast amount of available documentation and support for OMNeT++. The mobility framework provides the user with an easy mechanism to design and implement protocols at any layer of the WSN communication stack with support for node mobility and dynamic connection management.

A survey of literature has shown that all sources used in this dissertation directly or indirectly assumes certain properties and artefacts of the network and node design. The major problems identified are the amount, type and relevance of assumptions made to specify the routing protocol and for presenting results. These assumptions need to be considered, critically evaluated and the impact of each thoroughly understood when researching routing protocol specifications. The assumptions made during the design of the proposed protocol are described in as much detail as required to inform the reader of the possible weaknesses and future extensions of this research.

Chapter 5

WSN ROUTING PROTOCOL DESIGN

5.1 Introduction

Each of the seminal contributions briefly described in Chapter 3 was designed while considering the consequences of every design decision made. Literature dictates that the main consideration for the design of routing protocols should be energy efficiency. Other considerations should also be taken into account such as scalability, reliability and mobility. These considerations affect the final proposal, design and performance of a new routing protocol.

The scope of this dissertation includes the design and implementation of the proposed routing protocol named the *Mobile Tolerant Hybrid Energy Efficient Routing (MT-HEER)*. MT-HEER is based on HEER [3] to allow for enhanced packet delivery in networks where topological changes due to node mobility are unavoidable.

The operation of MT-HEER is discussed in an effort to inform the interested reader of every detail of the design. The level of detail is of utmost importance to promote repeatability and ratification of the completed work.

5.2 Critical Evaluation of HEER

Since HEER (as described in [3]) is the base starting point of this dissertation, this section covers a brief evaluation of the design rules and implementation of HEER. The OMNeT++ implementation of HEER was received from the author and was accepted as the author's final implementation. The OMNeT++ implementation of HEER was modified slightly to correct for miscalculations in packets delivered to the sink and miscalculations of message packet sizes as described in the remainder of this section.

5.2.1 Advantages of HEER

Critical analysis of HEER showed the following major advantages as compared to other published work. These advantages provide well developed building blocks used during this research.

1. The *energy efficiency* of HEER in general is acceptable with a combination of data aggregation for non-critical messages and an acknowledgement mechanism for critical messages.
2. HEER introduces an ingenious method of combining the advantages of both *flat and hierarchical* topologies.
3. The *cross layered approach* enables efficient communication between the layers in the communication stack.
4. *Tolerance to network topology changes* due to failing nodes is acceptable and within expected parameters. The protocol provides an easy mechanism to increase tolerance to node failures at the cost of higher energy consumption for network reconfiguration. This energy consumption cost is mitigated by *localised network maintenance* with power message transmissions.

5.2.2 Disadvantages and Reasoning errors of HEER

Critical analysis of the implementation of HEER produced the following disadvantages and reasoning errors that is addressed and improved on in this research.

1. The selection of a branch head is based purely on the hop count, “pseudo” distance, available energy and the amount of branches of the possible node. In unique cases two nodes may elect each other as branch heads creating the possibility of looping message transmissions.
2. Every branch head will wait for messages from all of the branch nodes. The design of HEER assumes that there will always be messages available and forwarded from all the branch nodes. The *aggregate data messages will never be forwarded* towards the sink node, should one or more of the branch nodes fail unexpectedly. *Message delivery latency* will be introduced if one or more of the branch nodes do not have any sensed data for a particular period.

3. The aggregated message forwarding mechanism implemented for HEER relies on the *aggregation node (or branch head) initiating data messages* as well. Aggregated messages will only be forwarded if the branch head senses data and creates a message to be forwarded. This situation can occur if the branch head's sensors malfunction or if there is no phenomena to report on.
4. The concept of including a *creator identifier* in forwarded messages is sound, but it is not implemented properly for aggregated messages. The creator identifier is not reflected in the message header for aggregated messages, unless the author's intention is to include the creator identifier as part of the message payload.
5. The protocol design presented in [3] reflects the network header size as 109 bits for a data message and this value was used in the OMNeT++ implementation as well. The correct header size is in fact 117 bits which has a performance impact on power calculations during simulations. This *mathematical error* questions the *validity of the results* presented in the HEER research. The message data header size has been fixed for the intention of presenting reliable results in this research.
6. The inclusion of a *destination address* in the broadcast message network header adds unnecessary overhead and consumes more energy on transmission than required. A destination address does not add any value to broadcast message as the message will be received and handled by all source nodes within the maximum transmission distance of the broadcasting node.
7. The results shown by the author for the number of data messages received by the sink node is suspect. The simulation implementation of received from the author of HEER *double counts the messages received* at the sink node. This has been fixed for the intention of presenting reliable results in this research.
8. The network routing configuration phase will select and update a branch head only with broadcast and power messages. This is a good strategy for energy efficiency and tolerance to network topology changes due to failing nodes. This is however not the best mechanism for successful message propagation through a network that includes a number of mobile nodes. There is mention of dynamic route calculation with every message being transmitted and

support for mobility in [3]. Support for mobility in HEER is only valid statement if the amount of mobile nodes, movement speed and pattern of the mobile nodes is such that the topology change is minimal between the periodic configuration broadcasts.

5.3 MT-HEER Protocol Design Considerations

This section discusses energy efficiency as well as the additional considerations taken into account during the design and implementation of MT-HEER.

5.3.1 Energy efficiency

Sources in literature state that transmission among nodes in a WSN is the highest energy consuming activity in the network [14]. The functional lifetime of the network as a whole can be prolonged by reducing the energy consumption in every node. The design of MT-HEER is based on the following mechanisms in an attempt to maximise energy conservation:

1. *Source initiated*: The source nodes will forward sensed data messages towards the sink node. There is no requirement for the sink node to forward interests or queries to one or many of the source nodes. In fact, any request from the sink node for any data messages should be ignored by the sensor nodes. This design choice reduces the global network knowledge of nodes and the number of messages transmitted to receive the data messages.
2. *Event driven*: Sensed data transmissions is dictated by the source nodes (source initiated) and will only forward data messages if relevant data is available. The transmission period can be based on a fixed time period or when predefined thresholds are reached. Source initiated and event driven mechanisms constitute the same operation in most cases.
3. *Single message path*: Data messages are forwarded to the single best node towards the sink node. A single message path towards the sink node minimises transmissions for a single message and contributes to the reduction in message duplication.
4. *Data aggregation*: The aggregation of data messages in static nodes reduces the amount of transmission overhead and the total amount of messages being

transmitted in the network as a whole. Data aggregation has been proven in literature to produce higher energy conservation.

5.3.2 Scalability

All flat network routing protocols and some of the recently developed hierarchical routing protocols assumes the deployment of homogeneous. The scalability of MT-HEER is manifested in the following properties:

1. *Hardware independence* requires that all nodes in the network have the same functionality, available resources and will equally participate in the protocol operation.
2. *Computational simplicity* and *dynamic route calculation* enables a higher scalability factor for the expansion of network on demand. Routing logic is localised to each node and its immediate neighbours. No global network knowledge is required. Routing data messages will only contain information to forward the message one hop closer to the sink node and the best node towards the sink node is computed with every message transmission. MT-HEER is computationally slightly more complex than HEER. The added complexity can be explained the tracking and routing mechanism for mobile nodes in an effort to provide a higher tolerance to node mobility.

5.3.3 Reliability of message delivery

The intention of sensor applications is to gather information of sensed phenomena in order to react to specific data if so required. The successful delivery of data messages at the sink node is a very important QoS property and requirement to fulfil the need of the application. Data messages are thus required to successfully propagate through the network to the sink node with the lowest possible latency. The reliability of the protocol needs to be weighed against the energy efficiency. Excellent energy efficiency does not add any value if the successful message delivery ratio is too low.

As mentioned the design of MT-HEER dynamically calculates the route from a source node towards the sink node. Dynamic route calculation should contribute to the routing protocol’s tolerance to topological changes due to node failures, node mobility and dynamic network expansion. To ensure critical messages are successfully delivered at the sink node, the transmission critical messages require the acknowledgement of message delivery. It is assumed that the application layer will handle unsuccessful message deliveries and retransmissions.

5.3.4 Node mobility

Traditional routing protocols rely on a static deployment of sensor and sink nodes in the network. Some parts of the network may become isolated from its path to the sink should any node in that same path fail or move beyond the maximum transmission distance. Routing protocols that allow no or little tolerance to node mobility will only reconnect the isolated path if the same mobile node moves with the transmission distance of its previous neighbours again. Network maintenance in the form of a network reconfiguration phase alleviates the problem posed by mobile nodes, but many data messages may be lost between configuration phases. The design of MT-HEER considers the following factors related to node mobility:

1. The *node movement model* can be a random movement pattern or predetermined. Random movement models should represent worse case scenario in terms of topological changes in the network as no predictions can be made about possible mobile node connectivity.
2. As mentioned *network connectivity* among nodes may be sporadic if mobile nodes are introduced to the network. If the node density in the network is high enough, connectivity and message routes towards the sink node should be able to be sufficient to deliver data messages at the sink node. This statement however assumes that mobile nodes are not selected as part of the path towards the sink node and that mobile nodes only use static nodes as a path towards the sink.
3. Mobility may introduce a *reduction in load* on the nodes closest to the sink by reducing packets traversing those nodes and rather using mobile nodes to reach the sink depending on available energy and traditional routing rules.
4. *Mobile node location* should not be a requirement for the protocol to function correctly.
5. Local *knowledge of node mobility status* is required to successfully alter the routing protocol behaviour for those nodes.
6. A number of the nodes deployed in a network should *still remain static* to anchor the static network configuration for energy efficiency.

5.4 Simulation model

Consider the deployment of *n* homogeneous wireless sensor nodes in an area *A*. To emulate real-life situations all the nodes and sink are distributed randomly across the area *A*. Assume that percentage *x* (where $x < 50\%$) of N nodes are marked as mobile nodes and the remainder percentage *y* as static nodes. The movement pattern of the mobile nodes across area *A* will be random for purpose of this model and nodes will move at a relatively slow constant speed. Assume that there is a single static sink node responsible for collecting all data from all nodes and interacting as the gateway to the wireless sensor network.

Each node uses it wireless radio link to communicate with its neighbours at a maximum transmission range R. No single node has global knowledge of the network and relays data to the sink in multiple hops across other nodes. The multi-hop simulation model implies that not all nodes will be in direct contact with the sink.

Each node requires local memory to store routing information and to store aggregated messages until being sent. Network addresses require memory allocation of 16 bits per address, which allows for 2^{16} (65536) unique addresses. Every mobile node would require an internal clock (or timing mechanism) to schedule dynamic routing updates for the node itself and its neighbours.

For the purpose of this research, assume that each node except the sink node will initiate and transmit a data message every 15 minutes and that every 150 minutes a critical data message will be created.

5.5 Protocol Design and Function

This section describes the detailed phases, design and implementation of MT-HEER.

5.5.1 Phase 1: Network Initialisation and Configuration

Once the nodes have been deployed and the network topology has stabilised, the network initialisation phase starts. The MT-HEER routing protocol uses a sink-initiated restricted broadcast flooding mechanism. The initialisation phase serves the purpose of configuring the initial routes from each of the nodes to the sink (similar to the Network Discovery step in SEER [2] and HEER [3]). It is assumed that no mobile node will start moving and that no data transmission will commence until the network initialisation phase has been completed.

5.5.2 Restricted Broadcast

The restricted broadcast initiates network discovery and configuration of the routing information. Within the MT-HEER scope, restricted broadcast implies that:

1. The sink node chooses a unique sequence number to distinguish this broadcast from other possible broadcast messages.
2. Each broadcast sequence (sequence number indicating uniqueness) is forwarded only once by each of the nodes in the network.
3. Each handled broadcast message is sent to all nodes on within the maximum transmission distance R.
4. Every broadcast message received is used to update neighbour information.

In contrast to SEER [2] and HEER [3] no destination address is used in the broadcast message, thus reducing the size of the broadcast message. All nodes within the maximum transmission distance R from the transmitting node will receive the broadcast message. The reduction in message size will drain less energy when transmitting and receiving a broadcast message.

The sink will initialise the hop count to 0 and each node resending the broadcast message will increment the hop count of the received broadcast message. Every node transmitting a broadcast message will add that node’s energy level and mobility status. The header for a broadcast message will contain the information specified in Table 5.1.

Table 5.1 Broadcast message Header

Field	Size (bits)	Description
Source Address	16	Address of the source node
Energy Level	16	Current energy level of the source node
Hop Count	8	Current hop count to the sink of the source node
Sequence Number	8	Unique sequence number for each restricted broadcast
Kind	3	Type of message as described in 5.5.3.
Is Mobile	1	Bit field showing if the source node is a mobile node
Total	52	Total bits used in header

5.5.3 Message Kind

The concept of a message kind enumeration included in every message (as adopted from [3]) serves the function to uniquely identify the type of message being sent or received. The size of the field requires 3 bits to allow for 23 (8) allowable values. The different values used are shown in Table 5.2.

Table 5.2 Possible Message Kind Values

Kind Type	Kind Value
Broadcast	0
Data	1
Power	2
Join	3
Acknowledge	4
Info	5
Beacon	6

5.5.4 Routing Information

Routing information is localised in each node without the knowledge of the global network. To this end, every node only contains the minimum amount of data to forward a message one hop closer to the sink. This requires memory to locally store and manage routing information in the form of a neighbour table. The information shown in Table 5.3 is stored as an array of neighbours in the neighbour table to reflect all of the nodes within the maximum transmission distance R from the specific node.

The information in the neighbour table is used to dynamically select the best path towards the sink based on hop counts and conserving energy network wide. The best path is determined by adopting the HEER admin value [3] as a Routing Coefficient:

$$Routing\ Coeffisient = \frac{(H_c)^d}{P_n} \tag{8}$$

Where: H_c is the hop count of the neighbour towards the sink node, *d* the distance of the neighbour node to this node and P_n the current energy level of the neighbour node. The distance *d* is calculated from the power of every message received and updated in the neighbour table which could be influenced if a free space loss model is not used for the radio transmission.

Table 5.3 Neighbour Table entries

Neighbour Parameter	Parameter Type
ID	Integer
Hop Count	Integer
Power	Integer
Branch	Integer
Distance	Decimal
Energy Coefficient	Decimal
Signal Power	Decimal
Is Mobile Node	Bit

A specific node will be aware of each neighbouring node's mobility status. Neighbouring node mobility is only an indication of whether or not the neighbouring node will be responsible for topological changes in the network structure. Although this parameter is not directly used in the routing calculations, it is indirectly used during the dynamic route maintenance as described in 5.5.7. During the calculation of routes towards the sink, mobile and static nodes will be handled in exactly the same manner based on the data stored in the neighbour table.

The node (assume node K) will now wait for 5 minutes to receive all available broadcast messages still being transmitted to update the neighbour table with all the node's neighbours as well as updating it's own hop count to the lowest possible value. Node K will now send a Join message to the "best" node in the neighbour table towards the sink. Assume that node K+1 is the "best" node towards the sink node. Node K will send a Join message to node K+1 to add node K as a branch in K+1's neighbour table. The concept of a branch of node K+1 only serves the purpose of indicating which nodes might be forwarding messages to node K+1. The "best" node is chosen by evaluating the neighbours with hop counts lower than node K's hop count, has the lowest Routing Coefficient value and with the least amount of branches. The selection of the "best" node will also ensure that the branches of node K is not selected to ensure that no routing loops may occur. The header for a join message will contain the information specified in Table 5.4.

Table 5.4 Join message Header

Field	Size (bits)	Description
Source Address	16	Address of the source node
Destination Address	16	Address of the destination node
Kind	3	Type of message as described in 5.5.3.
Total	35	Total bits used in header

Node K+1 will now wait another 5 minutes to receive all Join messages from its branches. This node will now send Info message to all its branches information them of the amount of branches currently to node K+1. Every Info message received will update the Routing Coefficient value in the receiving node's neighbour table for the transmitting node. The header for a join message will contain the information specified in Table 5.5.

Table 5.5 Info message Header

Field	Size (bits)	Description
Source Address	16	Address of the source node
Destination Address	16	Address of the destination node
Branch Size	4	Amount of nodes that may send messages to this node
Kind	3	Type of message as described in 5.5.3.
Total	39	Total bits used in header

As with HEER, every update of the neighbour table will force a recalculation of the routing information and selected node to use as the preferred path towards the sink. To this end, route recalculation will occur with the following message types received at a specific node: Broadcast, Join, Info, Data, Power and Beacon (Data, Power and Beacon message types and its impact will be discussed in detail in 5.5.5 and 5.5.7).

This completes the network initialisation phase.

5.5.5 Phase 2: Data Transmission

Data transmission can now commence without any negative impact on the successful delivery of packets to the sink. As mentioned in the previous section, every node has the predetermined routing knowledge of how to get a message one step closer to the sink. As with SEER [2] and HEER [3] the protocol allows for critical and non-critical messages. Every message is clearly marked as being critical or not, as shown in the header for a data message in Table 5.6.

A critical message will always be sent immediately to the branch head, whether it was created in the same node or received from a node further away from the sink. The same rule applies to both static and mobile nodes. Due to the nature of critical messages, it is of utmost importance that these messages reach the sink. A critical message is only sent to a single upstream node but expects an acknowledgement back that the message has been successfully received. If an acknowledgement is not received within a predefined time-limit, the message will be re-transmitted.

Table 5.6 Data message header

Field	Size (bits)	Description
Data	32	Message Data
Creator ID	16	Address of the node where the message originated to uniquely identify the message originating node.
Source Address	16	Address of the source node
Destination Address	16	Address of the destination node
Energy Level	16	Current energy level of the source node
Hop Count	8	Current hop count to the sink of the source node
Message Size	8	Total message size indicating aggregation of messages if the value is greater than 1
Kind	3	Type of message as described in 5.5.3
Critical	1	Bit field indicating criticality of message
Is Mobile	1	Bit field showing if the source node is a mobile node
Total	117	Total bits used in header

When a node receives a critical message, that node will use the source address of the received message to send back an acknowledgement (ACK) message to confirm that the critical message has been successfully received. The header for an ACK message will contain the information specified in Table 5.7. The critical message will immediately be forwarded to the next branch head until the message eventually reaches the sink. It is assumed that the network and application layer will handle the resending of critical messages if an ACK message was nor received by the sending node.

Table 5.7 ACK message header

Field	Size (bits)	Description
Source Address	16	Address of the source node
Destination Address	16	Address of the destination node
Kind	3	Type of message as described in 5.5.3.
Total	35	Total bits used in header

Non-critical messages will also be sent immediately if a static node has no branches or if the node is marked as mobile, using the routing information as described in 5.5.4. Non-critical messages however do not require acknowledgements of delivery in an attempt to minimise network transmissions. If a static node has one or more branches, messages will be aggregated as described in the next section.

5.5.6 Message forwarding and aggregation

As mentioned in the previous section, mobile nodes will always forward messages (critical and non-critical) without any message aggregation to increase the possibility of successful message delivery at the sink. If non-critical messages were to be aggregated at mobile nodes, there is a possibility that:

1. The mobile node may temporarily move outside the transmission distance of any node. Messages will not be lost, but timely message delivery will be decreased.
2. The path towards the sink may become much longer if the node physically moves away from the sink, thus increasing the mobile node's hop count. The movement would inherently decrease the possibility of successful or timely message delivery. This would not be a preferred scenario and data aggregation in mobile nodes would be detrimental to the performance of the protocol.
3. The path towards the sink may become shorter if the node physically moves towards the sink, thus reducing the mobile node's hop count. The movement would inherently increase the possibility of successful or timely message delivery. However due to the fact that the node movement pattern is random within this simulation model, there is no way to empirically predict the

possibility of this scenario or the advantage data aggregation in mobile nodes would add to the performance of the protocol.

Upon receiving a non-critical data message, a node will strip all the message headers except for the original message creator ID and the message data. These two parameters will be stored in a local cache on the node, until all the cached data are aggregated and forwarded towards the sink. The decision of a static node to forward a single message concatenated with all of the cached messages (specified as the aggregated messages) is based on two factors:

1. Messages have been received from all the branches of that node.
2. A predetermined time limit has been reached, but messages have not been received from all branches. This factor has been included to ensure messages will be forwarded even if message transmission from any neighbouring node fails due to unforeseen circumstances.

The second factor mentioned above is in contrast to the implementation of HEER that will always wait for all branch messages.

All of the data messages and respective creator IDs stored within the node's cache, will be aggregated within the data field of the forwarded message. The message size will be increased to the count of the aggregated messages and the original message being forwarded. The structure of the aggregated message payload with the data of the message is shown in Table 5.8. The main motivation for aggregating messages in static nodes is to reduce the amount of messages and overheads being transmitted.

Table 5.8 Aggregated message payload

Field	Size (bits)	Description
Data	32	Message Data
Creator ID	16	Address of the node where the message originated
Total	48	Total bits used in header

5.5.7 Power Maintenance

The failure of nodes due to depleted power sources is unavoidable due to the very nature of wireless sensor networks. Every failing node will influence the static routes determined during the network initialisation phase, if neighbouring nodes are not aware of the pending failure. To prevent routes becoming unusable when a node in a path to the sink fails, the concept of power updates is adopted from both SEER and HEER. If the power level of a node reaches a predetermined level, that node will be deemed as failing and will broadcast a power message to all the nodes within the maximum transmission distance R. Nodes receiving the power message will remove the failing node from the neighbour tables and recalculate the routing information. The header for a power message will contain the information specified in Table 5.9.

Table 5.9 Power message header

Field	Size (bits)	Description
Source Address	16	Address of the source node
Kind	3	Type of message as described in 5.5.3.
Total	19	Total bits used in header

5.5.8 Static Route Maintenance

As with SEER and HEER, the sink node will initiate a restricted broadcast at a predetermined interval. The same procedure will be followed as for the network initialisation phase as defined in 5.5.1 and will ensure that neighbour tables are still up to date and calculated routes are still accurate. Every restricted broadcast initiated form the sink will use a new sequence number to uniquely identify new routing initialisation phases.

Periodic broadcasts have an added advantage to update the routing tables if topological changes due to mobile nodes have occurred in the network. The validity of the aforementioned statement is dependant on two factors:

1. The movement speed of mobile nodes should be slow enough to remain within the neighbour’s transmission range for at least one data message transmission.

2. The ratio of mobile nodes to static nodes should be sufficiently low to maintain routes towards the sink using only the static nodes.

For the purpose of this design, assume that none of the two statements are true and that static route maintenance alone will not be adequate to ensure a high ratio of successful message deliveries.

5.5.9 Dynamic Route Maintenance

The main reasons that might cause routing failures due to node mobility if no extra route compensation is added can be summarised as:

1. Branch heads where data messages are aggregated might move outside the range of the branch node's maximum transmission distance R. Branch nodes still assume that the moving branch head is available to receive messages and keep on forwarding messages as such with no receiving node to forward messages towards the sink.
2. The moving node may aggregate messages within its data cache while moving away from its branch head. At the stage where the moving node tries to forward the aggregated messages towards the sink, the branch head will be unavailable to receive messages.
3. A moving node might isolate one or more nodes from the rest of the network with no means of routing the isolated nodes' messages towards the sink. If that same node does not pass within close proximity of its original position, the isolated nodes' messages will never reach the sink until static route maintenance attempts to recalculate routes towards the sink.

The introduction of Dynamic Route Maintenance (DRM) aims to mitigate the risk of possible message routing failures due to the movement of nodes within the simulation area as stated above. A mobile node will periodically broadcast a Dynamic Route Beacon (DRB) message (adopted from [10]) to all nodes within the maximum transmission distance R from that node. The message header the beacon message is described in Table 5.10.

Deduction through simulation analysis has shown that the interval of the DRB potentially has an impact on the energy efficiency and message delivery success rate to dynamically maintain routes to the sink node. A lower interval promotes the message delivery success rate by always having updated routing information. This however negatively impacts energy efficiency by transmitting more messages. A higher interval promotes energy efficiency by reducing message transmissions. This however negatively impacts the message delivery success rate by allowing route information to become outdated. The interval of the DRB was empirically tested to yield the best results at a 5th of the periodic restricted broadcast interval to maximise the packet delivery ratio and minimise the energy consumption impact.

Only static nodes will react on the DRB message and mobile nodes will discard the message. Upon receiving a DRB message in a static node, it will use information provided in the message header to recalculate the receiving node's routing table.

The following sequence logic will be used to update the neighbour tables of each of the receiving static neighbour nodes as well as the mobile node that initiated the DRB:

1. As mentioned in 5.5.4 the received power of the message is used to calculate the distance between a receiving static node and the transmitting mobile node. If the mobile node is already a neighbour to the static node, the distance between the nodes is increasing and the distance is beyond 98.5% of the maximum transmission distance R, the mobile node will be removed from the static node's neighbour table.
2. If the mobile node is already a neighbour to the static node, the distance between the nodes is decreasing (or remaining constant) and the distance is less than 98.5% of the maximum transmission distance R, the mobile node's information in the neighbour table will be updated. This will force a recalculation of the mobile node's Routing Coefficient.
3. If the mobile node is not yet a neighbour to the receiving static node the mobile node's information and calculated Routing Coefficient will be added to the static node's neighbour table.

4. The same procedure and rules as described in 5.5.4 will be followed to acknowledge the DRB with a Join and Info message. In this case the Join and Info messages will follow the DRB message and not on a Broadcast message. This step will force the recalculation of the best routes towards the sink, as the mobile node may provide a better path with less hops and energy consumption.

Table 5.10 Beacon message header

Field	Size (bits)	Description
Source Address	16	Address of the source node
Energy Level	16	Current energy level of the source node
Hop Count	8	Current hop count to the sink of the source node
Kind	3	Type of message as described in 5.5.3.
Total	43	Total bits used in header

In summary, the operation of MT-HEER can be illustrated by Figure 5.1.

5.6 Protocol evaluation and comparison

The protocols that are used as benchmark and for simulation analysis in this chapter are Flooding, SEER, LEACH and HEER. The simulation implementation of these protocols on OMNeT++ was received from the author of [3], analysed and modified to suite the analysis requirements of this dissertation. Each of the mentioned protocols is simulated in the mentioned environment to measure and analyse the simulated performance against MT-HEER.

The limited bandwidth and energy consumption available in WSN requires that the bandwidth of the transmission is limited by reducing the overhead of every message. Table 5.11 (adopted from [3] and updated) shows a comparison with the simulated routing protocols' overhead directly impacting the energy consumption during message transmission. This comparison clear shows a reduction in network overhead during message transmission as proposed by HEER and theoretically should show an improvement in energy efficiency.

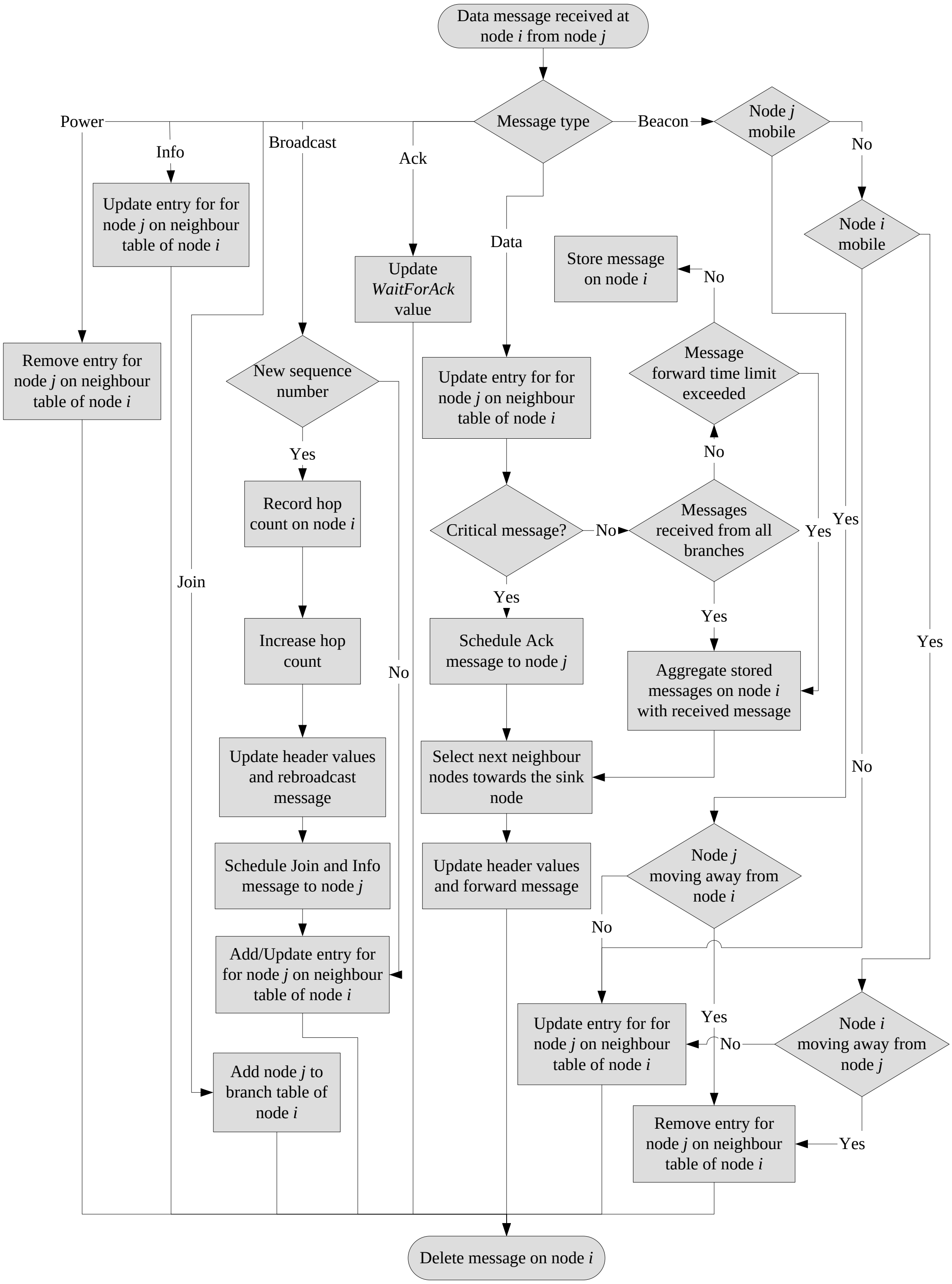

Figure 5.1 Message receiving process flow for MT-HEER

Table 5.11 Comparison of simulated protocols' overhead

Protocol	Broadcast	Data	Power	Advert	ACK	Join	Info	DRB
Flooding	52	108	-	-	-	-	-	-
SEER	68	125	68	-	-		-	-
LEACH	-	100	-	68	-	60	-	-
HEER	68	117	36	-	36	36	40	-
MT-HEER	52	117	19	-	35	35	39	43

5.7 Chapter Summary

Routing protocol design and operation is based on considerations dictated by literature and published works. The design considerations for the development of the proposed MT-HEER protocol are: energy efficiency, scalability, reliability of message delivery and support for node mobility.

The functional model of MT-HEER provides an explicit statement of the simulation model used to produce the results presented in the next chapter. The model is abstracted up to a level with sufficient information to ensure that the results can be independently repeated and verified.

The operation of MT-HEER is initialised by a restricted broadcast from the sink node to configure the routing information on each node. Each node will select a branch head as the next hop towards the sink node and will use that node to propagate data messages to the sink node. Data aggregation at the branch heads minimises the amount of message transmissions. Self configuration and tolerance to topological changes are manifested in the form of power, static route and dynamic route maintenance.

Chapter 6

RESULTS AND DISCUSSION

6.1 Introduction

The proposed protocol design for MT-HEER is implemented and simulated in OMNeT++ utilising the Mobility Framework (MF). The performance of MT-HEER as a routing protocol is compared to Flooding, SEER, LEACH and HEER and the implementation of each of these protocols only differs at the routing layer with minor differences in interlayer communication paradigms. Flooding was chosen due to the non-existence of any formal routing logic producing a high energy consuming protocol. The absence of any formal routing logic however increases the protocol's ability to handle topological changes. SEER and HEER were chosen as the introductory protocols for the starting point of this research. LEACH is one of the seminal contributions of hierarchical network architectures and is referenced multiple times in literature.

The artefacts and assumptions that do not differentiate the operation of any of the simulated routing protocols are used as base for the OMNeT++ implementation of all the evaluated protocols. These artefacts and assumptions are described in more detail in 4.3 and should ensure that the simulated protocol performance results are evaluated equally and under the same simulation conditions.

A critical evaluation of selected metrics produced by the results of protocol performance simulations provides an indication of the industrial applicability of the designed protocol.

6.2 Simulation metrics

When evaluating the performance of routing protocols, the research concentrates on predetermined performance metrics. These metrics are used to measure and critically analyse the simulated performance of MT-HEER against other leading routing protocols.

6.2.1 Time until first node dies

Message transmissions among nodes within the network are the highest energy consumer in WSN. The power source of a node is depleted with each transmission and reception of a message. The first complete depletion of a specific node’s energy will determine the point in time at which the full functionality of the network will start to decrease. This time will initiate the start of topological changes in the network. Although this metric does not indicate the routing protocol’s ability to handle topological changes due to failures or node mobility, it provides an indication of the energy consumption under no fault conditions and at maximum capacity.

6.2.2 Time until sink becomes unavailable

Routing protocol functionality depends on the ability to successfully deliver messages originating from source nodes to the sink node. The nodes within the maximum transmission distance from the sink node will relay messages towards the sink. The routing protocol’s message delivery ability is directly affected by the availability of the sink neighbouring nodes and the failure of these nodes over time, except for LEACH. As soon as all the power sources of the sink neighbour nodes are depleted, there is no way that a message can be delivered to the sink node. This metric benchmarks the point in time that a message will no longer be able to be reach the sink node and the functionality of the network will cease.

This metric is not evaluated for networks containing mobile nodes. A mobile node may pass within the maximum transmission distance from the sink node and act as a path towards the sink node even if all of the static neighbouring nodes have died.

6.2.3 Packet delivery ratio

Although energy efficiency should be the main design consideration for WSN routing protocols, the successful delivery of messages from source nodes to the sink node is a very important performance metric. If messages are not successfully delivered at the sink node, the network is wasting energy on unnecessary message transmissions and thus consuming more energy.

The packet delivery ratio can be defined as the ratio of the number of packets received at the sink node to the number of packets sent by all the source nodes. This ratio ranges between 0% to 100% where 100% indicates that all the data messages originating from the source nodes has successfully been delivered at the sink node. Packet delivery ratio is an important metric as it describes the message throughput of the routing protocol as well as the protocol's ability to handle topological changes and failures in the networks.

The implementation of the simulator environment is aimed at periodic data messages being initiated from each of the source nodes. The time period between message creations is 15 minutes. The packet delivery is measured over two fixed time periods. The first period will ensure that only a single message is created at each source node and allows 10 minutes for the messages to be delivered at the sink node. The second period is 610 minutes and allows for 40 messages to be created at each source node and 10 minutes for the last messages to be delivered at the sink node.

The packet delivery ratio depends mainly on path availability in terms of connectivity among nodes. Connectivity among nodes may be influenced by sporadic connectivity to mobile nodes, node failures, transmission collisions, etc.

6.3 Simulation Results: Static nodes

The simulation of the routing protocols for the purposes of this section consists only of homogeneous static nodes. The nodes are randomly deployed across the simulation area and the sink node is randomly chosen. This means that the sink node may be placed at any position within the simulation area. The variables of the simulator setup are kept constant for the simulation of each of the protocols. This includes the same source and sink node deployment for every protocol simulation. The results gathered for a static node deployment focuses on energy efficiency and packet delivery metrics.

Each of the simulation results are gathered for node deployments of 25, 50, 100, 200 and 500 nodes. The size of the simulation area is varied depending on the number of nodes for the simulation in an attempt to maintain an average distance between nodes. For example a node deployment of 50 nodes is performed on a simulation area of 400x400 area units; and a node deployment of 500 nodes is performed on a simulation area of 1000x1000 area units. Figure 6.1 shows a typical node deployment of 25 nodes in the OMNeT++ simulation environment with node 12 as the sink node.

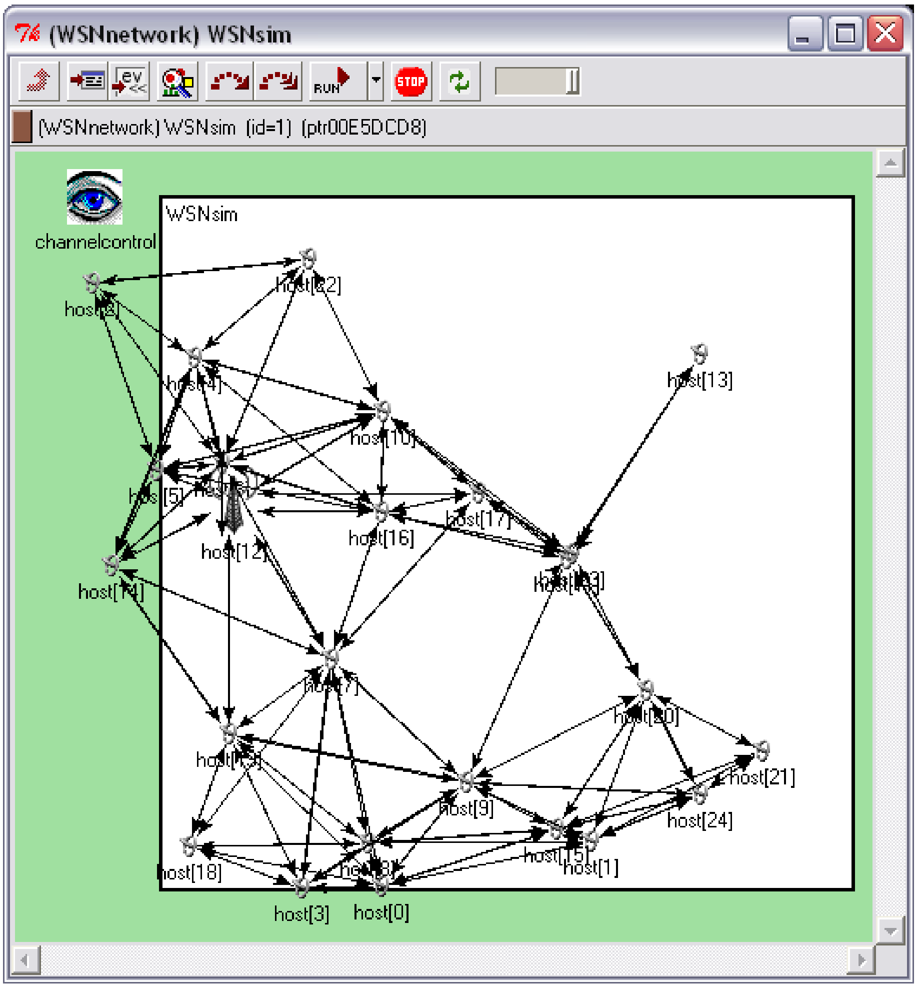

Figure 6.1 Typical node deployment in OMNeT++

6.3.1 Time until first node dies

The simulation results reflecting the first node failures for differing node numbers are shown in Figure 6.2. The flat network routing protocols are less scalable than the hierarchical network protocols. This can be seen from the performance of flooding and SEER which rapidly decline with an increase in network nodes. The time that the first node dies for LEACH on the other side, declines exponentially due to the possible large distances that the cluster heads have to transmit to the sink node. As expected one of the nodes that was elected as a cluster head is the first to deplete its energy source due to the deployment of homogeneous nodes. The performance of HEER and MT-HEER are similar and decline linearly with an increase in the number of nodes. The data aggregation strategy implemented in both HEER and MT-HEER proves to be less dependent on network size, implying better scalability and energy efficiency. HEER still out performs MT-HEER. An interesting phenomena observed during the simulation of MT-HEER is that that sink neighbouring nodes is not the first node to deplete its energy source. This means that MT-HEER reduces the hot spot problem close to the sink node.

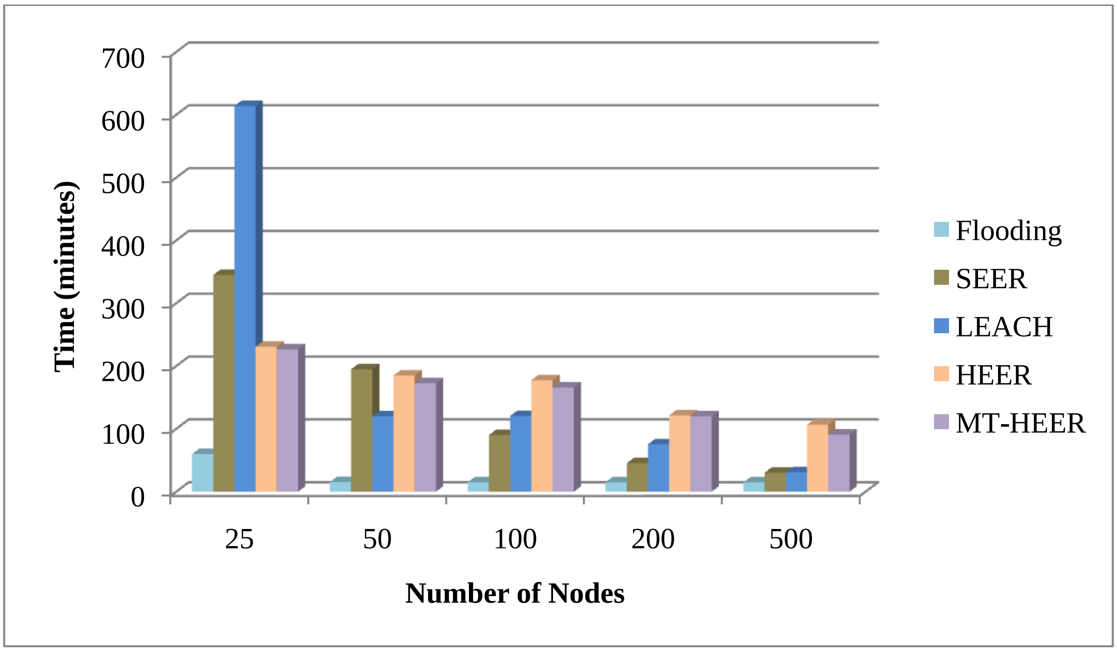

Figure 6.2 Time until the first node's power source depletes

6.3.2 Time until sink becomes unreachable

The routing protocol's ability to deliver messages to the sink node is dependant on the availability of the sink neighbouring nodes. The period in which the paths towards the sink node are accessible directly reflects the functional lifetime of the network and is shown in Figure 6.3. Flooding causes sink neighbouring nodes to deplete their energy sources much faster with the increase in node numbers due to the flooding of all messages. SEER, being a flat routing protocol, does not scale that well with an increase in network size. The main reason for this is the absence of a data aggregation strategy. HEER and MT-HEER perform much better than flat routing protocols with MT-HEER outlasting HEER due to less overhead in the transmitted messages.

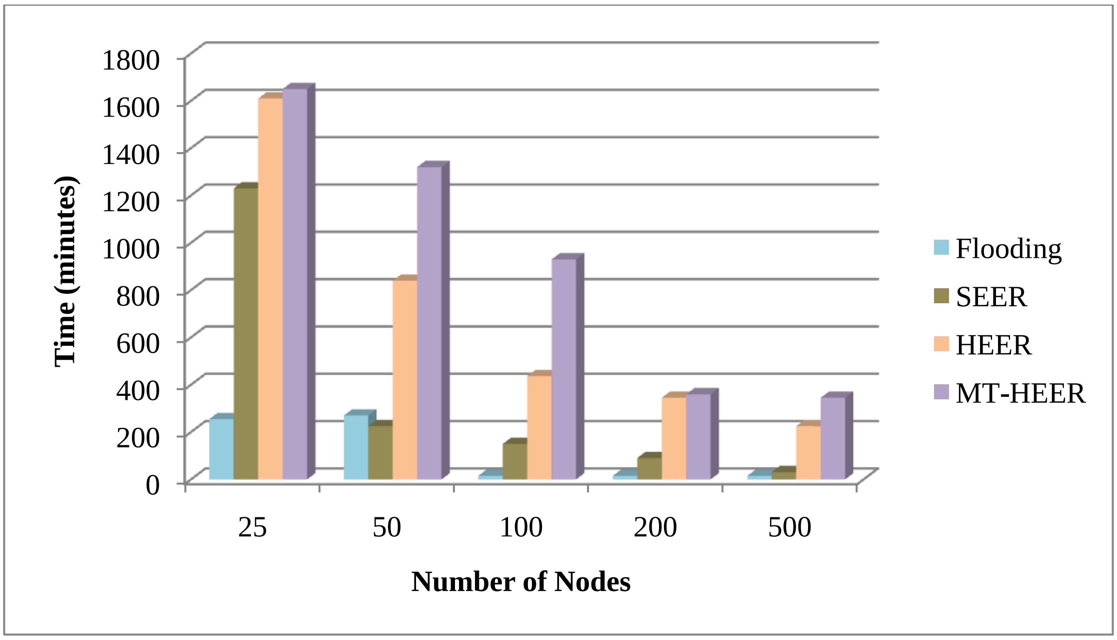

Figure 6.3 Time until sink node becomes unreachable

6.3.3 Message delivery ratio over fixed time period

Figure 6.4 shows the message delivery ratio results for a fixed period of 25 minutes and allowing a single message to be created in each of the source nodes. None of the simulated protocols caused a node failure during this period eliminating the protocol's ability to handle network failures. The immediate delivery of data messages by Flooding ensures that messages are delivered at the sink node due to the duplication of messages. With the increase in the number of nodes, flooding causes nodes to fail and reduces the delivery ratio. The energy efficiency of SEER complements the delivery ratio by producing low message delivery latency and no aggregation strategy. Messages are delivered to the sink node fast and efficiently. The direct communication from the LEACH cluster heads ensures that messages are consistently delivered to the sink node. The aggregation mechanisms implemented in both HEER and MT-HEER causes message delivery latency. At first glance the results show that the message delivery ratio decreases significantly, but this can be attributed to increasing latency with an increase in network size.

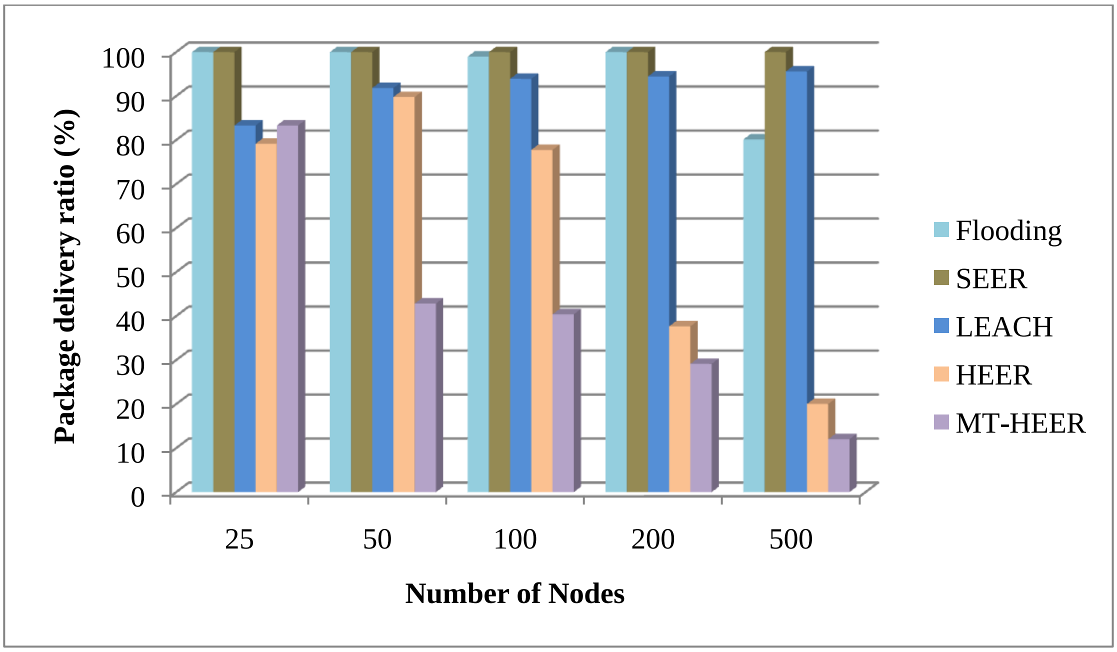

Figure 6.4 Message delivery ratio after 25 minutes

Figure 6.5 shows the message delivery ratio results for a fixed period of 610 minutes and allowing for the creation of 40 messages, should the individual source nodes still have sufficient energy levels available. A period of 610 minutes will ensure that at least one node failure occurs during the simulation and that the protocol's ability to handle network failures comes into play. The results produced by Flooding do not reflect the protocols true performance. The network dies out very fast and does not create new messages beyond 15 minutes for 500 nodes. Although a large ratio of the messages is delivered by, the amount of messages delivered is factors less than the other protocols. The linear reduction of message delivery ratios of LEACH, HEER and MT-HEER is expected with the increase in network size and message delivery latency due to data aggregation. These results do not indicate that the messages never reach the sink node, but merely that it may take longer. HEER outperforms the rest of the simulated protocols in terms of message delivery in static networks. The lower ratios of MT-HEER can be attributed to additional message aggregation mechanisms using a higher amount of power in individual nodes. These aggregations mechanisms are used to ensure higher message delivery for the handling of mobile nodes and for network fault tolerance. This does unfortunately add extra message delivery latency to the protocol when the network size increases and showing a lower message delivery ratio at the specific time period.

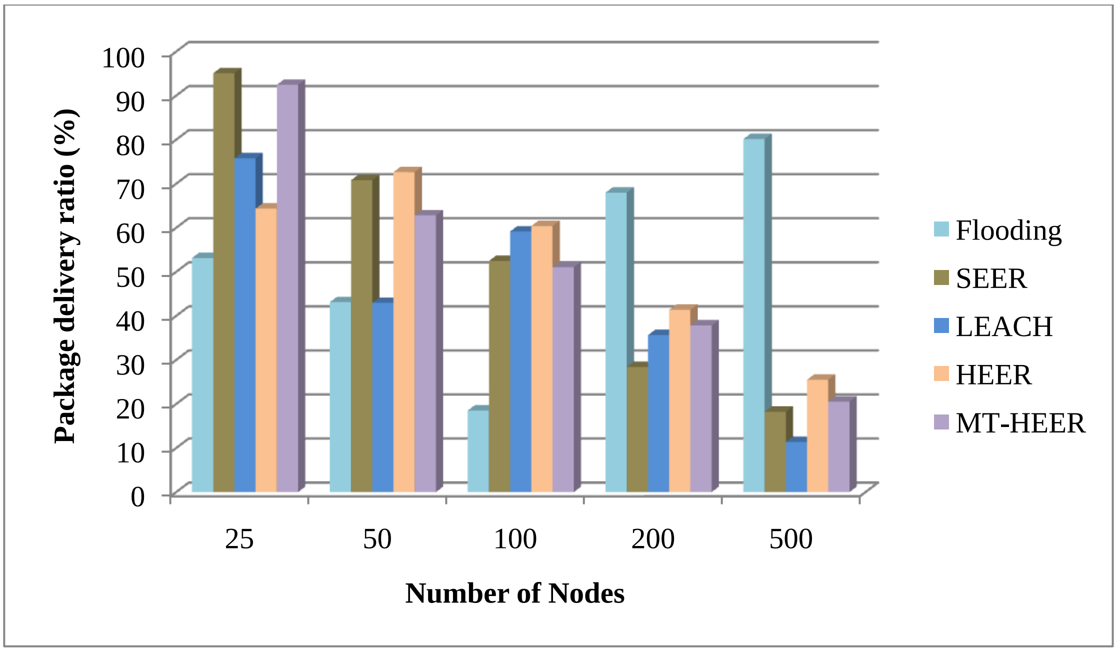

Figure 6.5 Message delivery ratio after 610 minutes

6.4 Simulation Results: Static and mobile nodes

The simulation of the routing protocols for the purposes of this section consists of a combination of static and mobile homogeneous nodes. All the nodes are randomly deployed and the sink node is randomly chosen as with the static scenario. A selection of nodes is marked as mobile nodes that will randomly move at a constant speed across the simulation area. The variables of the simulator setup are kept constant for the simulation of each of the protocols. This includes the same source and sink node deployment as well as the same movement patterns for every protocol simulation. The same protocol implementations are used for the mobile scenario as for the purely static deployment in the previous section. The results gathered for a static node deployment focuses on energy efficiency and packet delivery metrics.

The simulation results are gathered for node deployments of a 100 nodes deployed on a simulation area of 600x600 area units. Each of the results gathered are grouped by the percentage of mobile nodes in the deployed network and is 10%, 15%, 20%, 25%, 30% and 50% of the total amount of nodes deployed.

6.4.1 Time until first node dies

The simulation results reflecting the first node failures for differing node numbers are shown in Figure 6.6. The network size does not differ for the different simulations and as expected the performance of Flooding, SEER, LEACH and HEER do not depend on the amount of mobile nodes. None of the aforementioned protocols implements any mechanism to mitigate topological changes due to node mobility. The energy efficiency of MT-HEER does however decline rapidly with an increase in mobile nodes. The sharp decrease can be attributed to the additional mobility handling mechanisms for continuously update the routing protocols of the static nodes.

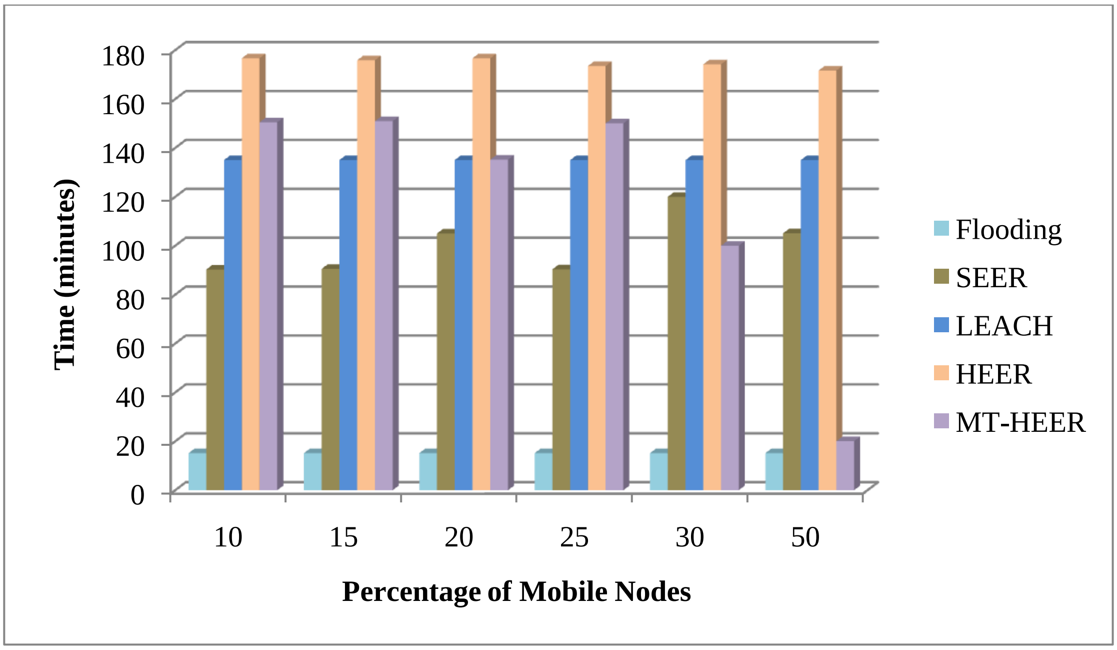

Figure 6.6 Time until the first node's power source depletes

6.4.2 Message delivery ratio over fixed time period

Figure 6.7 shows the message delivery ratio results for a fixed period of 25 minutes and allowing a single message to be created in each of the source nodes. None of the simulated protocols caused a node failure during this period eliminating the protocol's ability to handle network failures. Flooding is not affected by node mobility and can be attributed to the fact that there is no routing information stored. The flooded message will always reach the sink as long as node failures do not prohibit message propagation. The implementation of LEACH has not been altered and the variation in the delivery ratio does not follow any pattern. The only explanation for the erratic behaviour could be that the elected cluster heads may have been mobile nodes. Small amounts of mobile nodes do not affect HEER significantly, but beyond 20% mobile nodes the performance of MT-HEER exceeds the delivery ratio of HEER. HEER and MT-HEER do still have the message delivery latency as in the static scenario which could explain the lower delivery ratios. Due to the chosen low movement speed, node positions will not change much in the first 15 minutes and message delivery performance is similar to the static scenario in 6.3.1.

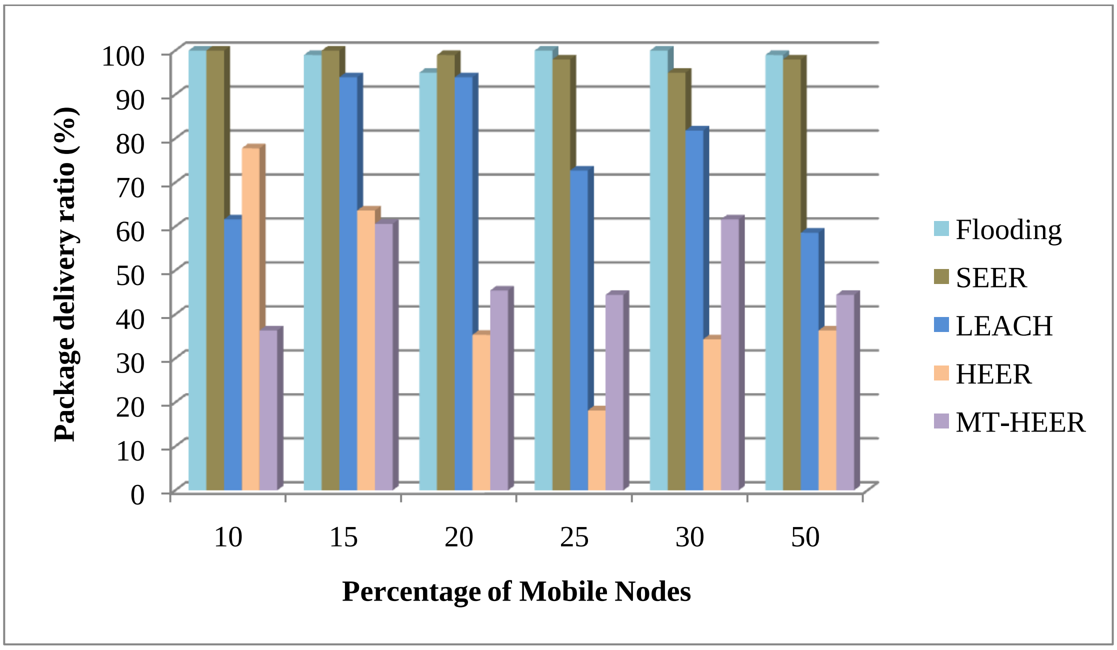

Figure 6.7 Message delivery ratio after 25 minutes

Figure 6.8 shows the message delivery ratio results for a fixed period of 610 minutes and allowing for the creation of 40 messages should the individual source nodes still have sufficient energy levels available. A period of 610 minutes will ensure that at least one node failure occurs during the simulation and that mobile nodes' position will change dramatically. This scenario will illustrate a protocol's true ability to handle topological changes due to network failures and node mobility. The erratic results gathered for the protocol simulations can be attributed mainly to node mobility. Of importance is the fact that 10% or less mobile nodes do not have a major impact on any of the protocols' performance. MT-HEER does however show superior performance and support for node mobility above 10 % mobile nodes. It can be extrapolated that MT-HEER will outperform other protocols in networks with more than 50% mobile nodes.

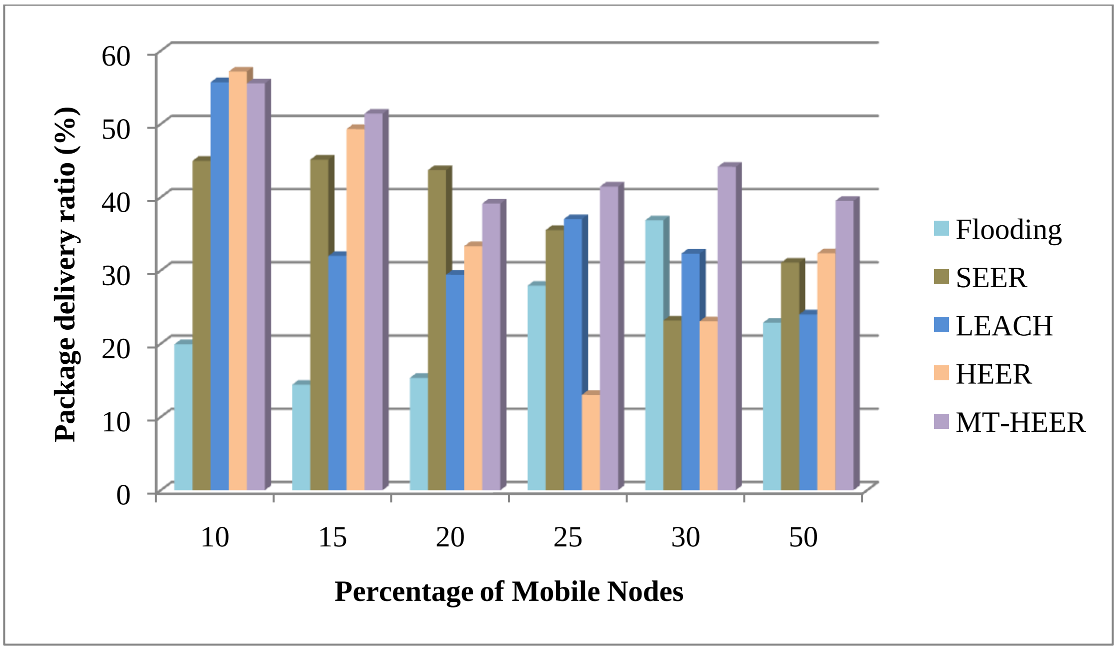

Figure 6.8 Message delivery ratio after 610 minutes

6.5 Impact of node mobility

The main areas of interest for this research, namely energy efficiency and successful message delivery, are influenced by node mobility as can be seen from the results shown in 6.3 and 6.4. The degree of influence can be seen by a comparison of the mentioned simulation metrics for static and mobile scenarios.

The first comparison is made on the energy consumption by comparing the results of the time period until the first node's energy source depletes for the same number of nodes. Figure 6.9 shows the difference between the static and mobile network simulation results for elapsed time until the first source node dies. Although this is a comparison between time variables, the time of transmission is a direct reflection of the energy consumed during normal network operation. A positive number on Figure 6.9 illustrates that the protocol caused higher energy consumption when operating in a network with mobile nodes than in a static network deployment. Inversely a negative number illustrates lower energy consumption in a network with mobile nodes than in a static network deployment. The higher energy consumption of MT-HEER is expected with the additional mechanisms to mitigate topological changes due to mobility. The graph confirms that an increase in the percentage of mobile nodes will transmit more messages in the network, thus increasing the energy expenditure. An unexpected result is the lower energy consumption of SEER with the introduction of mobile nodes.

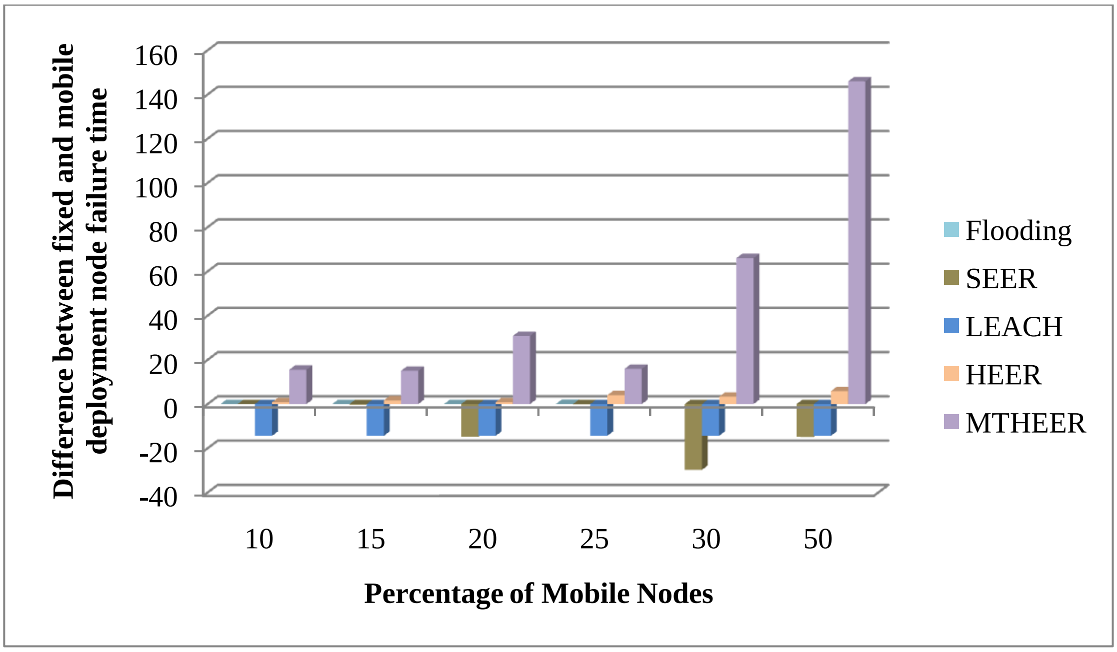

Figure 6.9 Difference in energy consumption

The energy consumption however cannot be evaluated in isolation. As mentioned before, low energy consumption does not add any value if the message delivery success rate is negligible. Figure 6.10 shows the difference between the static and mobile network simulation results for successful message delivery ratios. The chart illustrates higher positive values as more successful message deliveries in static network deployments and higher negative values as more successful message deliveries in network deployments containing mobile nodes. Lower positive values and preferably negative values indicate a higher tolerance to node mobility. The graph clearly confirms that MT-HEER has a much higher tolerance to node mobility. An interesting phenomenon is that Flooding actually delivers more successful messages to the sink in the mobile network scenario. This can be contributed to mobile nodes moving close to the sink node and relieving some of the pressure of the sink neighbouring nodes.

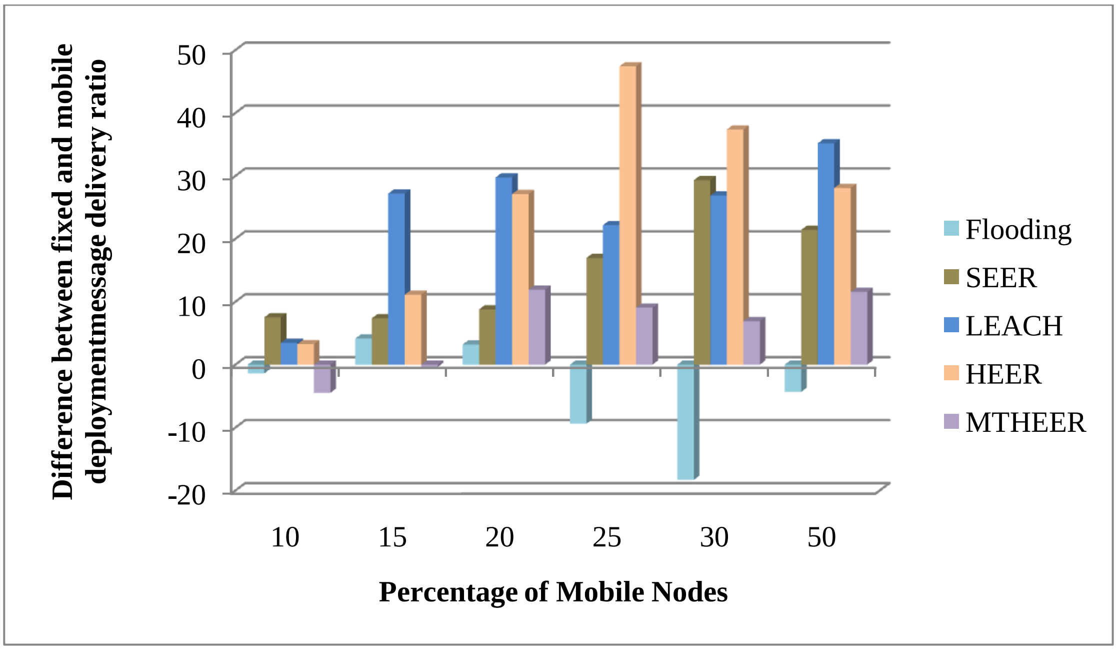

Figure 6.10 Difference in successful message delivery ratios

For the purpose of this research the author assumes that the higher than normal energy consumption of MT-HEER is acceptable if traded off against tolerance to node mobility. This trade-off between energy consumption and mobile node tolerance would however be best benchmarked in an actual industry application to determine if the power consumption limits is acceptable to the industry requirement.

6.6 Factors influencing simulation results

Simulation analysis and deduction during simulations of the routing protocols yielded some interesting characteristics of factors that influence the results presented in this dissertation. It is important to note that any variation on these factors will yield differing results by different researchers. Variation of one or a combination of the factors may even provide results indicating superior performance above others. These factors must carefully be considered and the impact of it analysed during network simulations.

The size of the *simulation area* and the number of nodes placed in this area affects the *initial node placement*, if random node placement is chosen. Random placement functions on a normal random number generator within the maximum x and y axis perimeters. A popular assumption during the design of many routing protocols in literature state that all nodes will be connected; where connectivity is defined as being within the maximum transmission distance of another node. Random placement of nodes by a simulation may place nodes far away from other node(s) causing *isolation between nodes* or even parts of a network. Without visual verification this may not be realised and could cause detrimental results. Protocol simulations could be written to ensure that all assumption is met before simulation runs are commenced.

The results presented in this dissertation assumed a *random placement of the sink node*. The sink node may be placed such that the sink will only have a single neighbouring node. User input on the selection may place the sink node at a more energy efficient position to ensure prolonged network functionality.

Node mobility introduces a few unknown variables that can be described as:

1. Although this dissertation assumes random movement of mobile nodes, literature provides a vast number of mathematical models describing the movement of mobile nodes. The *mobility model and the movement speed* however will be directly affected by the intended application, for example the movement of sensor nodes within a water line is explicitly known and routing protocols can be designed to anticipate that movement.

2. As shown in the results, the *number of mobile nodes* in a sensor network has a direct effect on the performance of the protocol. MT-HEER has been design with at least 50% of the nodes as stationary to form an anchor point.
3. Node *mobility during the initialisation* phase may also cause havoc in protocol operation, if not taken into consideration.
4. Moving source nodes may cause network and node isolation especially on the edges of the simulation area. The isolated nodes may only be able to forward messages to the sink node during sporadic network connectivity of mobile nodes passing within the maximum transmission distance.
5. The *mobility of the source nodes, sink nodes or both* has a direct impact on the strategy taken during message routing. Protocol design should explicitly state the assumptions of mobile nodes. All the protocols simulated in this chapter assume a static sink node.
6. A survey on literature mentioning node mobility assumes *knowledge of the location* of a mobile node in an effort to perform intelligent routing. Without proper scientific backing on the process of location determination, the researcher should be weary of these types of assumptions.

High density of nodes will provide more alternative paths towards a sink node. The *impact of node mobility* on the routing protocol in these deployments will be minimal as there will always be a next better path utilising the static infrastructure. The researcher’s opinion is that it is better to design for a network of sparsely deployed node to ensure that densely deployed nodes will yield the same result.

6.7 Chapter Summary

Simulation results and the verification thereof are crucial in academic research without the possibility of industrial application and testing. This chapter introduces the simulation metrics used for this research to evaluate and analyse the performance of routing protocols. The protocols included in the simulations are: Flooding, SEER, HEER, LEACH and MT-HEER.

The results presented in this chapter focus on energy efficiency and the reliability of message delivery for two different scenarios. The first scenario is a deployment of homogeneous static nodes and the second a deployment of static and mobile homogeneous nodes. The results are evaluated against the discussed metrics and possible reasons are stated for abnormalities in the simulation results. The performance metrics for the two scenarios are compared and clearly shows the impact of node mobility on the normal operation of a routing protocol.

During the simulation runs the author realised a few factors that may influence the simulation results of routing protocols presented by literature. These factors are discussed briefly.

Chapter 7

CONCLUSION

7.1 Introduction

Wireless sensor networks are a fairly new technology with multitudes of possibilities for new and interesting applications. These networks can be deployed randomly without physical restriction as with normal wired sensor systems.

Transmission of data and communication within these networks is of utmost importance and also the most energy consuming activities. Transmission of data relies on the network routing protocol to successfully deliver messages and data from the source to the sink. Due to the physical restrictions of the network components, energy conservation should be the very first consideration when developing and researching network routing protocols within a WSN. WSN network routing protocols is the focus of this dissertation to further the current body of knowledge.

7.2 Summary of Protocol Design

To conserve energy and maximise the lifetime of the network, a routing protocol needs to be as simple and scalable as possible. It should however not abstract away too much complexity, because by doing so it might fail to handle failures and topological changes within the network. These topological changes might even consume more energy network wide in an attempt to compensate for changes.

This dissertation proposes a mobile tolerant hybrid energy efficient network routing (MT-HEER) protocol for WSN. The routing protocol is required to handle failures as well as topological changes (node mobility) within the network, without consuming too much energy in the network as a whole to adapt to the changed network architecture. The development and evaluation of MT-HEER forms the foundation of this dissertation.

The development of MT-HEER required a thorough understanding of WSN routing principles and the seminal contributions to the field. Literature proves that a routing protocol contributes to the energy efficiency of a WSN by limiting the amount of message transmissions. MT-HEER is built upon the following principles and design decisions:

1. Combining the exceptional qualities of both flat and hierarchical network topologies.
2. Source initiated and event driven action.
3. Single message propagation path towards the sink node.
4. Data aggregation along the message propagation path.
5. Routing based on knowledge of neighbour hop count, available energy levels, distance (computed from transmission strength) and mobility status.
6. Computational simplicity.
7. Reliability of message delivery at the sink node.
8. Dynamic adaption to topological changes due to node failures and fault conditions, but more importantly due to node mobility.

7.3 Summary of Results

The implementation and simulation of MT-HEER was completed utilising OMNeT++ and the basic building blocks of the Mobility Framework (MF). The simulated MT-HEER performance was compared to Flooding, SEER, HEER and LEACH using metrics such as energy efficiency and the reliability of message delivery.

The simulation results presented in this dissertation represents two network architectures: static only nodes and a combination of static and mobile nodes. The performance metrics for the two scenarios are compared and clearly shows the impact of node mobility on the normal operation of a routing protocol. Static node deployment performance of MT-HEER exceeds flat network performance but does not quite match the performance of LEACH. Although MT-HEER consumes more energy with the inclusion of mobile nodes in the network, the reliability of message delivery more than compensates for this drawback.

The design of MT-HEER does present one major drawback when discussing true node mobility. Routing logic assumes that a number of static nodes be available for routing anchors. Simulation analysis has shown that once more than 60% of the network consists of mobile nodes, the performance of MT-HEER deteriorates rapidly. This however does not only apply to MT-HEER, but published protocols based on static networks perform even worse.

Acceptable evaluation, comparison and verification of published routing protocols require the availability of source code for the same simulator environment. The simulation code for all the protocols was accepted as correctly implemented. The source code for MT-HEER is readily available for analysis and performance verification.

7.4 Contribution

The proposed and developed Mobile Tolerant Hybrid Energy Efficient network Routing (MT-HEER) protocol succeeds in providing tolerance to mobile nodes in a WSN network. This tolerance is evident in the protocol's ability to ensure higher successful message delivery to the sink node with the introduction and increase of the number of mobile nodes, as compared to other protocols. The mechanism providing the tolerance to node mobility does however consume more energy in the fixed and mobile scenario than in the fixed only node scenario. This elevated energy consumption however still maintains the need for energy conservation within acceptable limits and is justified maintaining a high message delivery ratio with an increase of the fraction of mobile nodes in the network.

7.5 Future Work

Future research continuing on the design and principles used for the development of MT-HEER will verify the results and design presented in this dissertation. Possible research areas and topics could include:

1. Message aggregation and fusion of specific types of data and could also lead to discarding duplicate data. This would however require a cross layered design between the application and routing layer. The routing protocol would also require some knowledge of the application layer.

2. The investigation of specific mathematical mobility models for node movement based on specific applications in an attempt to verify the design of MT-HEER and possibly other routing protocols. The mobility models could include variable speed models.
3. The current implementation of the MT-HEER DRB transmission period is based on a fixed interval. The interval chosen for the simulations produced the best results for the chosen node movement speed. Variations in the node speed will affect the results positively or negatively due to the fixed period. Research on a variable DRB interval dependant on the current speed of the mobile node may prove to be invaluable.
4. The operation of MT-HEER assumes that the successful delivery of critical messages and the receipt of ACK messages are handled by the application layer. Further work could include intelligence in the routing layer to associate critical messages with corresponding ACK messages by means of a "conversation identifier". In this case the routing protocol could take over the responsibility to retransmit failed transmissions and dynamically removed failed transmission links to future use.
5. MT-HEER in its current relies strongly on the fact that nodes do not move during the initialisation phase. Future research could analyse the impact of constantly moving nodes, that is even during the initialisation phase and propose additional enhancements to MT-HEER to mitigate the added complexity.
6. This dissertation assumes the random placement of sensor nodes as well as the random placement of the sink node. The position of the sink node could drastically improve the energy efficiency of the network. A mathematical model determining the most energy efficient position for the sink node will provide a valuable contribution to the current body of knowledge.

7.6 Chapter Summary

The use of dynamic route maintenance might consume more energy but ensures an acceptable reliability of message delivery when topology changes occur due to node failures and mobility. MT-HEER succeeds in providing tolerance to mobile nodes within a WSN while operating within acceptable energy conservation limits.

REFERENCES

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: a survey", *Computer Networks*, Volume 38 (2002) pp 393–422, December 2001.

[2] C.J. Leuschner, "The design of a simple energy efficient routing protocol to improve wireless sensor network lifetime", Faculty of Engineering, the Built Environment and Information Technology, Dissertation, University of Pretoria, April 2005. [Online]. Available: http://upetd.up.ac.za/thesis/available/etd-01242006-091709/unrestricted/00dissertation.pdf

[3] J.G. Page, "Energy Efficient Hybrid Routing Protocol for Wireless Sensor Networks", Faculty of Engineering, the Built Environment and Information Technology, Dissertation, University of Pretoria, December 2006.

[4] C. Schurgers and M.B. Srivastava, "Energy efficient routing in Wireless Sensor Networks", Proceedings of MILCOM 2001 (Communications for Network-Centric Operations: Creating the Information Force). IEEE, Volume 1, pp 357-361, October 2001.

[5] R.C. Shah and J.M. Rabaey, "Energy Aware Routing for Low Energy Ad Hoc Sensor Networks", in Proceedings of WCNC2002 (Wireless Communications and Networking Conference, 2002), Volume 1, pp 350-355, March 2002.

[6] J. Zhang and H. Shi, "Energy-Efficient Routing for 2D Grid Wireless Sensor Networks", in Proceedings of International Conference on Information Technology: Research and Education, 2003, pp 311-315, August 2003.

[7] J.H. Chang and L. Tassiulas, "Maximum Lifetime Routing in Wireless Sensor Networks", IEEE/ACM Transactions on Networking, Volume 12, Issue 4, pp 609 – 619, August 2004.

[8] A. Datta, "A Fault-Tolerant Protocol for Energy-Efficient Permutation Routing in Wireless Networks", IEEE Transactions on Computers, Volume 54, Issue 11, pp 1409 – 1421, November 2005.

[9] Q. Liang, "Fault-Tolerant and Energy Efficient Wireless Sensor Networks: A Cross-Layer Approach", in Proceedings of IEEE MILCOM 2005, Volume 3, pp 1862-1868, 17-20 October 2005.

[10] R.W.N. Pazzi and A. Boukerche, “Mobile data collector strategy for delay-sensitive applications over wireless sensor networks”, Computer Communications, Volume 31, Issue 5, pp 1028-1039, March 2008.

[11] D. Culler, D. Estrin and M. Srivastava, “Overview of Sensor Networks”, *IEEE Computer Society*, Volume 37, Issue 8, pp 41-49, August 2004.

[12] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, “A Survey on Sensor Networks”, *IEEE Communications Magazine*, Volume 40, Issue 8, pp102-114, August 2002.

[13] V. Handziski, A. Kopke, H. Karl and A. Wolisz, “A common wireless sensor network architecture?”, Telecommunications Networks Group, Technische Universitat, Berlin, Technical Report TKN-03-012 pp 10-17, July 2003.

[14] E. Yoneki and J. Bacon, “A survey of Wireless Sensor Network technologies: research trends and middleware’s role”, Computer Laboratory, University of Cambridge, Technical report UCAM-CL-TR-646, September 2005. [Online]. Available: http://www.cl.cam.ac.uk/TechReports/

[15] S. Dai, X. Jing and L. Li, “Research and Analysis on Routing Protocols for Wireless Sensor Networks” in Proceedings International Conference on Communications, Circuits and Systems, 2005, Volume 1, pp 407-411, May 2005.

[16] M.A. Perillo and W.B. Heinzelman, “Wireless Sensor Network Protocols”, in Algorithms and Protocols for Wireless and Mobile Networks, Eds. A. Boukerche et al., CRC Hall Publishers, 2004

[17] J.N. Al-Karaki and A.E. Kamal, “Routing Techniques in Wireless Sensor Networks: A Survey”, IEEE Wireless Communications, Volume 11, Issue 6, pp 6-27, December 2004.

[18] K. Akkaya and M. Younis, “A Survey on Routing Protocols for Wireless Sensor Networks”, Ad Hoc Networks, Volume 3, Issue 3, pp 325-349, May 2005.

[19] R. Farivar, M. Fazeli and S.G. Miremadi, “Directed Flooding: A Fault-Tolerant Routing Protocol for Wireless Sensor Networks”, in Proceedings of Systems Communications, 2005, pp 395-399, 14-17 August 2005.

[20] J. Kulik, WR Heinzelman and H. Balakrishnan, “Negotiation-Based Protocols for Disseminating Information in Wireless Sensor Networks”, Wireless Networks, Volume 8, Issue 2-3, pp169-185, March 2002.

[21] F. Ye, A. Chen, S. Lu, and L. Zhang, “A scalable solution to minimum cost forwarding in large sensor networks,” in Proceedings of IEEE 10th Int. Conf. Computer Communications and Networks, Scottsdale, Arizona, pp 304-309, October 2001.

[22] W.D. Henderson and S. Tron, “Verification of the Minimum Cost Forwarding Protocol for Wireless Sensor Networks”, IEEE Conference on Emerging Technologies and Factory Automation, 2006 (ETFA 2006), pp 194 – 201, 20-22 September 2006.

[23] C. Intanagonwiwat, R. Govindan, D. Estrin and J. Heidemann, "Directed Diffusion for Wireless Sensor Networking", IEEE/ACM Transactions on Networking, Volume 11, No 1, February 2003.

[24] W.R. Heinzelman, A. Chandrakasan and H. Balakrishnan, “Energy-Efficient Communication Protocol for Wireless Microsensor Networks”, in Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, 2000, Volume 2, pp 10, 4-7 January 2000.

[25] Ruay-Shiung Chang and Chia-Jou Kuo, “An Energy Efficient Routing Mechanism for Wireless Sensor Networks”, 20th International Conference on Advanced Information Networking and Applications, 2006 (AINA 2006), Volume 2, pp 5, 18-20 April 2006.

[26] V. Mhatre and C. Rosenberg, “Homogeneous vs Heterogeneous Clustered Sensor Networks: A Comparative Study”, IEEE International Conference on Communications, 2004, Volume 6, pp 3646-3651, 20-24 June 2004.

[27] S. Lindsey and C.S. Raghavendra, “PEGASIS: Power-Efficient GAthering in Sensor Information Systems”, in Proceedings of IEEE Aerospace Conference, 2002, Volume 3, pp 3-1125 – 3-1130, 2002.

[28] Zheng Zeng-wei, Wu Zhao-hui and Lin Huai-zhong, “An Event-Driven Clustering Routing Algorithm for Wireless Sensor Networks”, in Proceedings of International Conference on Intelligent Robots and Systems, 2004 (IROS 2004), Volume 2, pp 1802 – 1806, 28 September - 2 October 2004.

[29] M. Hempel. H. Sharif and P. Raviraj, “HEAR-SN: A New Hierarchical Energy-Aware Routing Protocol for Sensor Networks”, in Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS), 2005, pp. 324a - 324a, 3-6 January 2005.

[30] H.S. Kim and K.J. Han, “A Power Efficient Routing Protocol Based on Balanced Tree in Wireless Sensor Network”, in Proceedings of the 1st International Conference on Distributed Frameworks for Multimedia Applications, Besançon, France, 2005, pp. 138-143, 6-9 February 2005.

[31] A. Manjeshwar and D.P. Agrawal, “TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks”, in Proceedings of 15th International Symposium on Parallel and Distributed Processing, pp 2009-2015, 23-27 April 2001.

[32] A. Manjeshwar and D.P. Agrawal, “APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks”, in Proceedings of International Symposium on Parallel and Distributed Processing, pp 195-202, 15-19 April 2002.

[33] R.B. Araujo, A. Boukerche and R.W.N. Pazzi, “Fault-tolerant wireless sensor network routing protocols for the supervision of context-aware physical environments”, International Symposium on Parallel Architectures, Algorithms, and Networks, 2008 (I-SPAN 2008), pp 247-253, 7-9 May 2008.

[34] A.Martirosyan, A. Boukerche and R.W.N. Pazzi, “A Taxonomy of Cluster-based Routing Protocols for Wireless Sensor Networks”, Journal of Parallel and Distributed Computing, Volume 66, Issue 4, pp 586-599, April 2006.

[35] D. Kotz, C. Newport and C. Elliot, “The mistaken axioms of wireless-network research”, Dartmouth College Computer Science Technical Report TR2003-467, July 2003.

[36] A. Varga, “OMNeT++ Discrete Event Simulation System, Version 3.2, User Manual”, 29 March 2005. [Online]. Available: http://www.omnetpp.org

[37] “OMNeT++ Discreet Event Simulation System,” 2008. [Online]. Available: http://www.omnetpp.org

[38] M. Lobbers and D. Willkomm, "A Mobility Framework for OMNeT++, User Manual, Version 1.0a4", 12 January 2007. [Online]. Available: http://www.omnetpp.org

[39] "Mobility Framework for OMNeT++," 2008. [Online]. Available: http://mobility-fw.sourceforge.net