A FRAMEWORK FOR ENERGY EFFICIENCY IN WIRELESS MULTI-HOP AD HOC AND SENSOR NETWORKS

BY

AHMED M. SAFWAT

A thesis submitted to the

School of Computing

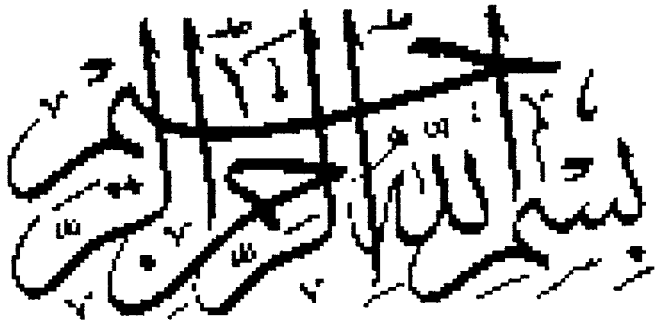
in conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

April 2003



IN THE NAME OF ALLAH,
THE MOST GRACIOUS, THE MOST MERCIFUL

إهـداء . .

إلى والدي ووالدتي

To my FATHER and MOTHER

دَعِ الأَيَّامَ تَفْعَلُ مَا تَشَاءُ ... وَطِبْ نَفْساً إِذَا حَكَمَ القَضَاءُ

وَلا تَجْزَعْ لِحَادِثَةِ اللَّيَالِي ... فَمَا لِحَوَادِثِ الدُّنْيَا بَقَاءُ

وَكُنْ رَجُلاً عَلَى الأَهْوَالِ جَلْداً ... وَشِيمَتُكَ السَّمَاحَةُ وَالوَفَاءُ

وَإِن كَثُرَتْ عُيُوبُكَ فِي البَرَايَا ... وَسَرَّكَ أَن يَكُونَ لَهَا غِطَاءُ

تَسَتَّرْ بِالسَّخَاءِ فَكُلُّ عَيْبٍ ... يُغَطِّيهِ كَمَا قِيلَ السَّخَاءُ

وَلا تُرِ لِلأَعَادِي قَطُّ ذُلاً ... فَإِنَّ شَمَاتَةَ الأَعْدَا بَلاءُ

وَلا تَرْجُ السَّمَاحَةَ مِن بَخِيلٍ ... فَمَا فِي النَّارِ لِلظَّمْآنِ مَاءُ

وَرِزْقُكَ لَيْسَ يُنْقِصُهُ التَّأَنِّي ... وَلَيْسَ يَزِيدُ فِي الرِّزْقِ العَنَاءُ

وَلا حُزْنٌ يَدُومُ وَلا سُرُورٌ ... وَلا بُؤْسٌ عَلَيْكَ وَلا رَخَاءُ

إِذَا مَا كُنْتَ ذَا قَلْبٍ قَنُوعٍ ... فَأَنْتَ وَمَالِكُ الدُّنْيَا سَوَاءُ

وَمَن نَزَلَتْ بِسَاحَتِهِ المَنَايَا ... فَلا أَرْضٌ تَقِيهِ وَلا سَمَاءُ

وَأَرْضُ اللَّهِ وَاسِعَةٌ وَلَكِن ... إِذَا نَزَلَ القَضَا ضَاقَ الفَضَاءُ

دَعِ الأَيَّامَ تَغْدِرُ كُلَّ حِينٍ ... فَمَا يُغْنِي عَنِ المَوْتِ الدَّوَاءُ

الإمام الشافعي

ABSTRACT

Multi-hop wireless ad hoc and sensor networks eradicate the costs of infrastructure deployment, setup, and administration. Ad hoc wireless networks allow anywhere, anytime network connectivity with complete lack of control, ownership, and regulatory influence. The intricate problem of energy conservation in wireless ad hoc and sensor networks is of great significance due to the limited battery capacity of the participating mobile devices. Hence, ad hoc routing protocols ought to be energy conservative. However, the simulation studies carried out for table-driven and on-demand ad hoc routing protocols fall short of examining essential power-based performance metrics, such as average node and network lifetime, energy-based protocol fairness, average dissipated energy per protocol, and standard deviation of the energy dissipated by each individual node. In this thesis, we present a thorough energy-based performance study of power-aware routing protocols for wireless mobile ad hoc networks. Our energy consumption model is based on a detailed implementation of the IEEE 802.11 physical layer convergence protocol (PLCP) and medium access control (MAC) sublayers. To our best knowledge, this is the first such detailed performance study.

Moreover, we propose, based on the findings of our energy-based performance study, two novel optimal energy-efficient schemes for path selection and load assignment, namely, Energy-Constrained Path Selection (ECPS) and Energy-Efficient Load Assignment (E2LA). Both schemes employ probabilistic dynamic programming techniques and utilize cross-layer interactions between the network and MAC layers. Despite the fact that energy conservation is achieved, there is no provision of a deterministically guaranteed minimum system lifetime.

We next develop a novel deterministic framework that allows the admission of flows without jeopardizing the limited energy of the wireless stations. This framework alleviates congestion by using multiple routes and through contention mitigation. Similar to ECPS and E2LA, it may be used with any existing energy-efficient routing scheme. Our experiments reveal that load balancing is achieved amongst the routes and the nodes in the wireless ad hoc network without violating any of the energy constraints, and while adhering to a deterministically pre-computed minimum system lifetime.

To achieve further energy savings, we propose a power-aware and responsive wireless infrastructure. The wireless infrastructure, namely, Power-Aware Virtual Base Stations (PA-VBS), facilitates energy-efficient routing and load balancing and enforces semi-centralized medium access. Although the proposed wireless backbone forms a dominating vertex set in the network graph, its construction is based upon localized decisions.

Keywords: Wireless Mobile Ad hoc Networks, Sensor Networks, Energy Efficiency, Routing, Medium Access Control, Wireless Infrastructure, Quality of Service

ACKNOWLEDGEMENTS

My thankfulness and gratitude are to Allah, the Holy, the Creator, the Sovereign, the justly Proud, the Originator, and the Wise.

It is He Who brought you forth from the wombs of your mothers when ye knew nothing; and He gave you hearing and sight and intelligence and affections: that ye may give thanks to Allah.[1]

But say, "O my Lord! Increase me in knowledge."[2]

They said: "Glory to Thee, of knowledge we have none, save what Thou Hast taught us: in truth it is Thou who art perfect in knowledge and wisdom."[3]

"O my Lord! So order me that I may be grateful for Thy favors, which Thou hast bestowed on me and on my parents, and that I may work the righteousness that will please Thee: and admit me, by Thy Grace, to the ranks of Thy Righteous Servants."[4]

I'm grateful to Professors Hossam Hassanein and Hussein Mouftah. Their supervision and mentorship were a source of help and guidance throughout the course of my research. Their confidence in my work made me strive to achieve my goals. I would also like to thank Professor Patrick Martin for his help and advice.

[1] The Holy Qurán, surat: "Al-Nahl", aya: 78.
[2] The Holy Qurán, surat: "Taha", aya: 114.
[3] The Holy Qurán, surat: "Al-Bakara", aya: 32.
[4] The Holy Qurán, surat: "Al-Naml", aya: 19.

I'm totally indebted to my parents and my dearest sisters, Najla'a and Eman. Their love provided me with the motivation and endeavor to accomplish this piece of work. They have always been a great source of encouragement and support. They gave me so much that I can never pay them back. I will owe them gratitude and thankfulness forever.

Finally, I would like to thank everyone else who has supported me in one way or another throughout my doctoral studies.

STATEMENT OF ORGINALITY

I hereby certify that this Ph.D. thesis is original and that all ideas and inventions attributed to others have been properly referenced.

LIST OF ACRONYMS

3G	Third Generation Wireless Systems
4G	Fourth Generation Wireless Systems
4G+	Beyond Fourth Generation Wireless Systems
ACK	Acknowledgement
AFECA	Adaptive Fidelity Energy-Conserving Algorithm
AODV	Ad hoc On-demand Distance Vector
AODV-BR	Ad hoc On-demand Distance Vector Backup Routing
ARP	Address Resolution Protocol
BECA	Basic Energy-Conserving Algorithm
BMT	Border Mobile Terminal
BTMA	Busy Tone Multiple Access
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
CGSR	Clusterhead Gateway Switch Routing
CI	Contention Indicator

CMMBCR	Conditional Max-Min Battery Capacity Routing
CRB	Contention-Resolution-Based
CRB/CT	Contention-Resolution-Based with Contention Threshold
CSMA	Carrier Sense Multiple Access
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CT	Contention Threshold
CTS	Clear-To-Send
DAG	Directed Acyclic Graph
DBTMA	Dual Busy Tone Multiple Access
DCF	Distributed Coordination Function
DFS	Depth-First-Search
DIFS	DCF InterFrame Space
DLAR	Dynamic Load-Aware Routing
DSDV	Destination-Sequenced Distance Vector
DSR	Dynamic Source Routing
DSSS	Direct Sequence Spread Spectrum
E2LA	Energy-Efficient Load Assignment
ECPS	Energy-Constrained Path Selection
EDP	Energy-Dependent Participation
EMC	Expected MAC Contenders
FCC	Federal Communications Commission

GAF	Geographical Adaptive Fidelity
GPS	Global Positioning System
GSR	Global State Routing
IARP	IntrAzone Routing Protocol
IBSS	Independent Basic Service Set
IERP	IntErzone Routing Protocol
IETF	Internet Engineering Task Force
IP	Integer Program/Programming
ISM	Industrial, Scientific, and Medical
LAR	Location-Aided Routing
LCC	Least Cluster Change
LP	Linear Program/Programming
MAC	Medium Access Control
MACA	Multiple Access Collision Avoidance
MACA-BI	Multiple Access Collision Avoidance By-Invitation
MANET	Mobile Ad hoc NETworking working group
MBCR	Minimum Battery Cost Routing
MMBCR	Min-Max Battery Cost Routing
MP	Max Power
MSC	Mobile Switching Center
MT	Mobile Terminal

MTPR	Minimum Transmission Power Routing
NIC	Network Interface Card
NMP	Normalized Max Power
NPV	Normalized Power Value
PAMAS	Power-Aware Multi-Access with Signaling
PARO	Power-Aware Routing Optimization
PA-VBS	Power-Aware Virtual Base Stations
PDP	Probabilistic Dynamic Programming
PHY	Physical Layer
PLCP	Physical Layer Convergence Protocol
PSTN	Public Switched Telephone Network
Q-GSL	Quasi-Guaranteed System Lifetime
QoS	Quality of Service
RB	Random Backbone
RERR	Route ERRor
RREP	Route REPly
RREQ	Route REQuest
RTR	Ready-To-Receive
RTS	Request-To-Send
SIFS	Short InterFrame Space
SNR	Signal-to-Noise Ratio

SSA	Signal Stability Adaptive routing
TDMA	Time Division Multiple Access
TORA	Temporally Ordered Routing Algorithm
UP	UpPeriod
VB	Virtual Backbone
VBR	Variable Bit Rate
VBS	Virtual Base Station
ZRP	Zone Routing Protocol

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In conventional wireless communication networks, a fixed amount of frequency spectrum is allocated to a cellular operator by the national regulator. In Germany alone, the 3G spectrum was sold to carriers for more than $46 billion. On the contrary, in a homogeneous wireless ad hoc network, mobile devices such as cellular phones and laptop computers are equipped with wireless transceivers that operate in the 2.4 GHz (2400-2483.5 MHz) Industrial, Scientific, and Medical (ISM) band. The use of this band requires no licensing from the Federal Communications Commission (FCC), and hence no costs are incurred for its use. Nevertheless, wireless ad hoc networks may operate in licensed frequency bands and constitute parts of a heterogeneous wide area network.

The interest in wireless ad hoc networks stems from their suitability for different types of application scenarios ranging from home and road to office and school [1]. Since there is no fixed infrastructure, a wireless mobile ad hoc network can be deployed quickly. Thus, such networks can be used in situations where either there is no other wireless

communication infrastructure present or where such an infrastructure cannot be used because of military tactics, an emergency (as a result of a natural disaster or an enemy attack, for example), cost reasons, etc.

In ad hoc networks, the non-existence of a centralized authority complicates the problem of medium access regulation. The centralized medium access control (MAC) procedures, undertaken by base stations in cellular networks, have to be enforced in a distributed, and hence collaborative, fashion by mobile stations in the ad hoc network. Mobile stations may contend simultaneously for medium access. Consequently, transmissions of packets from distinct mobile terminals are more prone to overlap, resulting in packet losses. As a result, retransmissions take place and noticeable delays result. Likewise, the performance of the MAC scheme has a great effect on the performance of the routing method employed, and on the energy consumption of the wireless network interface card. Third Generation (3G) wireless networks use a variation of the time division multiple access (TDMA) and code division multiple access (CDMA) schemes. MAC protocols with slot reservation capabilities for handling multimedia ATM constant bit rate (CBR) and variable bit rate (VBR) traffic cannot be extended to ad hoc networks since they assume a centralized access technique that is dependent on fixed base stations. For ad hoc networks, the IEEE 802.11 carrier sense multiple access with collision avoidance (CSMA/CA) distributed protocol [47] is utilized. CSMA/CA is a random access protocol and cannot support bandwidth reservation, nor provide QoS routing.

Ensuring QoS-communications in a wireless mobile ad hoc network is, apparently, highly dependent upon routing and MAC. Obviously, conforming to QoS measures, such as delay bounds, depends on the chosen route from the source mobile station to the destination mobile station. In addition, complying with QoS guarantees imposes the use of a MAC method that guarantees the successful transmission of packets under high mobility and/or heavy load circumstances.

In addition to cooperatively resolving the problem of simultaneous medium access, wireless ad hoc stations must resolve the problem of hidden terminals. A neighbor of the destination that is out of the wireless range of the source may interfere with the transmissions of the source. It is noteworthy that neither the source nor the interferer will be aware of the collision. Therefore, immediate recovery cannot be undertaken by either of them. The source will only attempt retransmission if it does not receive an acknowledgment from the destination within a specific time frame. Accordingly, the source will be aware of the occurrence of a collision, or data corruption. On the other hand, a node that is in range of the sender but not the receiver is called an exposed terminal. An exposed terminal can transmit at the same time as the sender, without causing a collision to occur. In spite of that, using the traditional CSMA/CA scheme, the exposed terminal will defer from accessing the channel. As a result, the capacity of the wireless ad hoc network will be reduced. Hence, energy is wasted groundlessly.

Evidently, most of the research performed in the field of wireless ad hoc networks has focused on the problems of routing and MAC. Nevertheless, the limited battery capacity

of the mobile devices making up the ad hoc network draws our attention to the importance of power awareness in wireless ad hoc network design. Hence, ad hoc routing and MAC protocols ought to be energy conservative. More importantly, in independent basic service set (IBSS) mode, wireless ad hoc stations may not switch off their wireless network interface cards (NICs) to conserve power because they must always be ready and willing to receive traffic from their neighbors.

On the contrary, a wide range of energy conservation schemes may be implemented in conventional wireless networks. In a wide area cellular environment, for example, a wireless node may be scheduled to sleep during idle periods [33]. Incoming packets are buffered by base stations and wireless access points and delivered to sleeping nodes when their transceivers are switched back on.

In a wireless ad hoc network, the energy consumption of the NIC may be significant compared with conventional wireless networks. As a result, a thorough energy-based comparative and performance study is essential to any bandwidth-based study. Energy-efficient schemes must still attain high throughput and packet delivery without a considerable tradeoff.

In previous research, the so-called power-aware routing schemes used battery capacity as the sole means for achieving energy-based fairness and system longevity. It is our conjecture that energy conservation may only be attained if valuable MAC (and PHY)

input is used by the network layer to take effective energy-aware routing decisions and resolve contention. Such premises and presumptive conclusions are supported by an exhaustive energy-based performance study. Consequently, and unlike previous proposals, we develop schemes which are not only energy-efficient but also performance-preserving and cultivate the performance of existing routing algorithms with respect to traditional metrics.

Hence, in this thesis, we propose novel schemes whose objective is to enhance the operation of existing power-aware multi-path routing protocols and to mitigate contention via cross-layer designs and optimal load assignments. In addition, we develop an infrastructure for the infrastructure-less wireless mobile ad hoc networks. This is of utmost importance as it reduces the problem of wireless mobile ad hoc communications, from a multi-hop problem, to a single-hop one, as in conventional cellular communications. It also fosters energy conservation in infrastructure-less networks. The previous apprehensions originate from the intuitive realization of the problems of power consumption, medium access, and routing in wireless mobile ad hoc networks and their effects on QoS-guaranteed communications.

This thesis is organized as follows. The next chapter presents a review of the literature and state of the art. This is followed in Chapter 3 by our detailed energy-based performance study. In Chapter 4, novel energy-efficient cross-layer designs and schemes are proposed and studied. A framework for deterministically imposing a quasi-guaranteed minimum system lifetime and for contention mitigation is introduced, explained, and

studied in Chapter 5. We describe our power-aware infrastructure scheme and study its performance via event-driven simulations in Chapter 6. Finally, Chapter 7 presents the conclusions drawn from the thesis in addition to future work.

CHAPTER 2

RELATED WORK

To study energy conservation in wireless multi-hop networks researchers must have a solid and thorough understanding of the different functions of the ad hoc protocol stack whose performance has a direct impact on energy consumption. Many innovative routing schemes take advantage of the previously proposed mobility tracking protocols [3-11] for cellular networks. Ad hoc schemes that utilize conventional cellular architectures come up with a hierarchical network architecture of the wireless mobile ad hoc network [12-14]. Most of the routing protocols proposed for ad hoc networks [15-21], [24], [28-29] do not construct hierarchical network architectures as in cellular networks. The normal flat-routed network architecture is used instead. As a result, packet routing is done based on peer-to-peer connections. Distributed medium access regulation, for wireless ad hoc networks, has been extensively studied [50], [37], [47], [49], [54-57], and several techniques have been proposed for collision avoidance. Recently, schemes that utilize geographical information using the Global Positioning System (GPS) were developed

[63-64]. Numerous energy-conserving schemes have been proposed for wireless ad hoc networks [72]. In particular, we have been witnessing a proliferation in unconventional energy conservation techniques [71], [78], [82]. Minimum-energy ad hoc networks composed of wireless nodes with an unlimited supply of power are studied in [91]. Minimum-energy routing has been also addressed in [92-96], where the approach is to minimize the total energy expenditure to reach the destinations. However, some nodes are overused and their batteries are thus depleted. In [97], bounds on the system lifetime are derived for static sensor networks. Nevertheless, these derivations can only be used for data-gathering wireless sensor networks. To date, energy efficiency and routing reliability remain the main impediments towards the commercial realization of wireless multi-hop networks.

This chapter is organized as follows. Section 2.1 surveys the proposed schemes for mobility tracking in conventional cellular networks. Section 2.2 presents several protocols and architectures developed to construct a hierarchical ad hoc network. Some of the existing schemes used to route packets in flat-routed ad hoc networks, are described in Section 2.3. Utilizing geographical information in wireless ad hoc networks is investigated in Section 2.4. In Section 2.5, distributed wireless medium access regulation and its challenges are presented. Application-level power-awareness techniques for multi-hop networks are studied in Section 2.6. Finally, Section 2.7 presents the chapter summary and conclusions.

2.1 MOBILITY TRACKING IN CELLULAR NETWORKS

The field of mobility tracking in cellular communications is considered the driving force for some novel research in the area of routing in wireless mobile ad hoc networks. Therefore, it is important to understand the basic concepts of cellular communications and how the problem of mobility tracking is addressed in conventional cellular networks. As a matter of fact, many of the innovative ideas that came out recently to serve in the field of routing in wireless mobile ad hoc networks are enhancements of well-known cellular schemes that proved to be successful in the field of cellular communications.

Two basic operations are involved in mobility tracking: location update and terminal paging. Location update is carried out by the mobile terminals to inform the network of their latest cell location. Three dynamic location update strategies, namely time-based, movement-based and distance-based, are studied in [4]. Under these three schemes, location updates are performed based on the time that has elapsed, the number of movements performed, and the distance traveled, respectively, since the last location update. A mobility tracking method similar to that used in mobile-IP networks is introduced in [5], [6].

Other schemes [8] focus on dividing the service area of a mobile switching center (MSC) into partitions, each of which is composed of a number of cells. This, consequently, limits the location update events to when partition boundaries are crossed. A movement-based location update scheme is combined with a selective paging scheme in [9]. The scheme in [9] does not require each mobile terminal to store information about the arrangement and

the distance relationship of all the cells. In the proposed scheme, a mobile station performs a location update when the number of movements since the last location registration is equal to a predefined value called the location update movement threshold.

A location update policy that minimizes the cost of mobile terminal location tracking is developed in [10]. A mobile terminal dynamically determines when to update after moving to a new cell based on the distance traveled since the last location registration and the incoming call arrival probability.

The concepts of broadcasting and forwarding pointers are introduced in [11]. In broadcasting, when the exact location of a mobile terminal is not known, the location server broadcasts the search message to a group of base stations, which includes the base station currently serving the mobile terminal. In the forwarding pointers scheme, when a mobile terminal moves into another cell, it leaves a pointer to its new location at the previous cell. Searching is done by following a chain of pointers, from the last known position to the current position. However, in [8], it is shown by simulation that the performance of the forwarding pointers scheme degrades at the low "calls per move" end. The pointer chain is longer when no boundary crossing occurs, and there is no call delivery. These two basic schemes are combined in [8]. At the low "calls per move" end, the algorithm switches to the limited broadcast scheme from the forwarding pointers scheme to improve performance.

2.2 HIERARCHICAL WIRELESS AD HOC ARCHITECTURES

Several hierarchical architectures for ad hoc networks have been proposed [12], [13], [14]. A mobile infrastructure is developed in [12] to replace the wired infrastructure and base stations in conventional cellular networks. Moreover, an adaptive hierarchical routing protocol is devised. The routing protocol utilizes the mobile infrastructure in routing packets from a source mobile terminal (MT) to a destination MT. The algorithm divides the wireless mobile ad hoc network into a set of clusters, each of which contains a number of MTs, which are at most two hops away. However, clusters are only a logical arrangement of the nodes of the mobile network. Thus, unlike clusterheads, cluster centers do not coordinate transmissions. Nodes that relay packets between clusters are called repeaters. Such nodes play a major role in performing the routing function. Since route maintenance was not studied by means of simulation, it is not possible to evaluate its efficiency.

In [13], the Random Backbone (RB) algorithm requires that the clusterheads form a dominating and independent set in the network graph. RB is used to derive a solution for the basic problem of connecting the clusterheads into a virtual backbone (VB) using fewer connections compared to the deterministic case. The scheme provides a probabilistic connectivity guarantee. RB constructs a virtual backbone that is connected with a probability that approaches 1 for large and dense networks. It is apparent that for this type of approach to address the communications problem in wireless mobile ad hoc networks, it would add great overhead to the MTs in the network since mobile stations might have to switch roles from time to time. Every time this happens, clusterheads will

have to be determined from the beginning and the virtual backbone must be re-built from scratch. Also, route construction and maintenance are not addressed and, as a result, the effect of the proposed architecture on packet delivery is not studied.

A heuristic, called max-min, is proposed in [14] and forms d-hop clusters. For the sake of determining the set of clusterheads, each run of max-min uses 2d rounds of flooding. It is when the mobile stations with the largest node IDs are d hops apart that max-min's solution is optimal, which cannot be guaranteed in a dynamic environment such as that of an ad hoc network. In addition, since max-min forms d-hop clusters, it does not suit the dynamic nature of ad hoc networks, where topology changes due to node mobility are unpredictable. Therefore, determining (or predicting) the time as to when the heuristic should be best run is significant.

In the *Virtual Base Station (VBS)* scheme [53], [60], some of the MTs, based on an agreed-upon policy, become in charge of all the MTs in their neighborhood, or a subset of them. This can be achieved by electing one to be a VBS. If a VBS moves or stops acknowledging its presence via its so-called *hello* packets, for a period of time, a new VBS is elected. A number of issues need to be resolved. The first issue is the way in which the VBSs are chosen. Electing a single VBS from a set of nominees should be done in an efficient way. Another issue to be addressed is the handing over of responsibilities from one VBS to another. Every MT (*zone_MT*/VBS) has a *sequence number* that reflects the changes that occur to that MT, and a *my_VBS* variable, which is used to store the VBS in charge of that MT. If an MT has a VBS, its my_VBS variable is

set to the ID number of that VBS, else if the MT is itself a VBS, then the my_VBS variable is set to 0, otherwise it is set to –1. Hello messages sent by VBSs contain their current knowledge of the ad hoc network, i.e., the whole ad hoc network. A VBS accumulates information about all other VBSs and their lists of MTs and broadcasts this information in its periodic hello messages. On the other hand, unlike VBSs, zone_MTs accumulate information about the network from their neighbors between hello messages, and their network information is cleared afterwards.

After being chosen as a VBS, the node stores information about all other nodes in the network. An MT is chosen by one or more MTs, to act as their VBS based on an agreed-upon rule, viz., the MT with the smallest ID number. A noteworthy remark is that a node with a smaller ID number than another may be thought of as one that is more capable (in terms of processing speed, battery capacity, or any other criterion) than the latter. MTs announce their ID numbers with their periodic hello messages. An MT sends a *merge-request* message to another MT if the latter has a smaller ID number. The receiver of the merge-request responds with an *accept-merge* message, increments its sequence number by 1 to indicate that some change took place, and sets its my_VBS variable to 0. When the MT receives the accept-merge, it increments its sequence number by 1 and sets its my_VBS variable to the ID number of its VBS. If an MT hears from another MT whose ID number is smaller than that of its VBS, it sends a merge-request message to the former. When it receives an accept-merge message, it increments its sequence number by one and updates its my_VBS field. The MT then sends a *dis-join* message to its previous VBS (the dis-join message is only sent if an accept-merge was received, otherwise it

would not be sent). When the old VBS receives the dis-join, it removes the sender from its list of MTs, which it is in charge of, and it increments its sequence number by one.

Unlike other one-hop infrastructure creation protocols, such as the least cluster change (LCC) protocol proposed in [59], VBS puts more emphasis on a node becoming a VBS rather than being supervised by a VBS. Hence, if a node receives a merge request, it responds by sending an accept-merge message, even if it is being supervised by a VBS. Moreover, it is noteworthy here that this neither degrades intra-cluster nor inter-cluster communications by any means. This is because if the node that became a VBS was originally acting as a gateway, or, equivalently, a border mobile terminal (BMT), for its previous VBS, it will still be a gateway, besides being a VBS. This becomes of great significance if the criterion upon which VBSs are elected is one that relies on the assets possessed by the MTs of the ad hoc network. Processing speed, main and secondary storage, MAC contention experienced in the neighborhood of the MT, and more, can be amongst such assets. Consequently, if an MT chooses not to become a clusterhead, as in [59], only because it is under the supervision of another MT, even though it possesses the required assets to become one, the node requesting to merge might experience demoted communications to other nodes in the ad hoc network because it does not have the proper resources and is not able to be associated with a VBS. A comprehensive comparative study of VBS and LCC has been conducted in [42].

2.3 ROUTING IN AD HOC AND SENSOR NETWORKS

Routing protocols in wireless mobile ad hoc networks can be classified as *table-driven*,

on-demand, and *hybrid.* In table-driven routing protocols, all the mobile stations are required to have complete knowledge of the network. Table-driven routing is also known as proactive routing. On-demand routing, however, implies that routes are built between nodes only as desired by source nodes. Hence, the terms on-demand and reactive can be used interchangeably. Hybrid routing schemes combine the functionality of both table-driven and on-demand routing. Location information is utilized in geography-informed routing protocols to route data packets to their destinations. In energy-conserving ad hoc routing protocols, the wireless nodes can trade off energy dissipation and data delivery quality according to application requirements.

Examples of table-driven routing protocols include global state routing (GSR) [15], destination-sequenced distance vector (DSDV) [16], and VBS-based routing [34]. Dynamic Source Routing (DSR) [18], ad hoc on-demand distance vector (AODV) [19], ad hoc on-demand distance vector backup routing (AODV-BR) [79], and depth-first-search (DFS) routing [80], are examples of on-demand routing protocols. The zone routing protocol (ZRP) [20] is both proactive and reactive, and is thereby a hybrid wireless mobile ad hoc routing scheme.

2.3.1 TABLE-DRIVEN AD HOC ROUTING

In Global State Routing (GSR) [15], similar to the link state routing scheme, nodes exchange vectors of link states among their neighbors during routing information exchange. Consequently, nodes maintain a global knowledge of the network topology and optimize their routing decisions locally. A key difference between GSR and

traditional link state routing is that link state packets are flooded into the network whenever a node detects topology changes. The dissemination method used by the distributed Bellman-Ford algorithm is the one used here because it has the advantage of no flooding. Nodes running GSR periodically exchange their link state tables with their local neighbors only. This ensures no flooding as in the conventional link state algorithm. Each node maintains the following: a neighbor list, a topology table, a next hop table, and a distance table. A node's immediate neighbors are stored in its neighbor list. For each destination in the ad hoc network, the topology table contains the link state information as reported by the destination. In addition, for each destination, the next hop table contains the node to which the packets for this destination must be forwarded. The shortest distances to all destinations are stored in the distance table. The link state with a greater sequence number renders the one with a smaller sequence number obsolete. The main disadvantage of GSR is that it maintains knowledge of the full network topology, which becomes unnecessary when topological changes are less frequent, and hence, the demand on such knowledge is less. Moreover, even when the network topology changes rather quickly, but the active ongoing communications are not at all affected by such changes, it is again unnecessary to trigger any updates to reflect the topological changes. This is also not desirable in the case of large populations of mobile nodes.

The Destination-Sequenced Distance Vector routing protocol (DSDV) [16] has been proposed as a variant of the distance vector routing method by which wireless mobile nodes cooperate to form an ad hoc network. In DSDV, a routing table is maintained by every mobile node in the network. The table contains all the possible destinations within

the network and the number of hops to each destination, as well as the sequence number assigned by the destination node. The sequence numbers provide a means to enable mobile nodes to distinguish old routes from new ones. Thus, the formation of loops is avoided. Mobile stations running DSDV transmit both time-driven and event-driven topological information. They periodically transmit their routing updates to their one-hop neighbours. Moreover, a node transmits its routing table if it has been significantly changed since the last update. For the sake of efficiency, the routing table updates that can be sent include a full table or an incremental update. However, in a dynamic mobile ad hoc network, incremental updates may not be of any useful value for nodes that have been offline. New route broadcasts include the sequence number of the destination, along with its address and the number of hops to reach it. The route with the highest sequence number is always used. If two routes have the same sequence number, ties are broken using the shortest distance. However, DSDV is inefficient because of the requirement of periodic update transmissions, regardless of the number of changes in the network topology. In addition, in highly dynamic networks, where topological changes are very frequent, incremental updates will grow in size and will be as bad as full dumps in terms of traffic generation, and will even become unnecessary when the routes are not needed for any ongoing communications sessions.

The VBS routing protocol [34], [40], [101] utilizes the VBS infrastructure-creation protocol described in Section 2.2 to route packets from a source node to a destination node. The VBS protocol is used in successfully developing a fast and reliable routing scheme that provides a solution to the intricate problem of routing in wireless mobile ad

hoc networks. All the MTs run the VBS protocol, and hence, some are elected to act as VBSs. The wireless mobile ad hoc network relies on the wireless mobile infrastructure created by VBS to provide communications between the different MTs. Each and every VBS is in charge of a set of MTs including itself. In VBS routing, only certain nodes, i.e., the nodes that form the dynamic infrastructure, are eligible to acquire the knowledge of full network topology. Route requests are not flooded to the rest of the ad hoc network due to the existence of the wireless mobile infrastructure. As explained earlier, a BMT is an MT that lies within the wireless transmission range of one or more nodes that are being supervised by different VBSs, or are VBSs themselves. When an MT wishes to send a packet, it sends it to its VBS, which forwards the packet to the VBS in charge of the destination or to the correct BMT. The sent packet contains the address of the destination. When the VBS receives the message, it looks up the destination address in its table. If the destination is found, the VBS of the source MT will forward the packet to the VBS in charge of the destination. This is done by consulting the BMT field of that VBS. The message is then forwarded to the MT (which might also be a VBS) whose ID number is stored in the BMT field. The BMT, after receiving the message, forwards it to its own VBS (if it was not itself one). This process is repeated until the message reaches the destination. Thus, path maintenance is an on-the-fly, built-in process.

MTs are neither responsible for the burden of discovering new routes, nor maintaining existing ones. Hence, this novel routing scheme eliminates the initial search latency, introduced by the discovery phase in on-demand routing protocols, which degrades the performance of interactive and/or multimedia applications due to the lack of a wireless

mobile infrastructure that can efficiently handle routing. In addition, unlike table-driven protocols, MTs are not required to store the whole network topology. Proactive-reactive routing is also possible using the VBS architecture.

Clusterhead gateway switch routing (CGSR) [59], as opposed to VBS-based routing, requires all the nodes in the ad hoc network to maintain routing information about each and every other node in the network. Other pitfalls of this scheme include the problem of pseudo links.

In table-driven routing protocols, since all the mobile stations are required to have complete knowledge of the network through their periodic and incremental/triggered updates, unnecessary periodic updates may have a negative effect on energy conservation. A number of routing proposals for ad hoc networks take energy conservation into consideration so as to extend the lifetime of the wireless nodes by wisely using their battery capacity [73], [74], and [75].

2.3.2 ON-DEMAND AD HOC ROUTING

On-demand routing has two major components: *route discovery* and *route maintenance.* The route discovery function requires a source to use some form of flooding. The transit nodes, upon receiving a query, "learn" the path to the source and enter the route in their forwarding tables. The destination node responds using the path traversed by the query. A route can then be established between source and destination. Route maintenance is responsible for reacting to topological changes in the network, and its implementation

differs from one algorithm to the next.

A simple method of route discovery is proposed in [24]. It is very similar to the Address Resolution Protocol (ARP). Flooding is used to propagate a request packet (similar to that of ARP) in order to reach other mobile stations that are not immediate neighbours. As the request propagates, each MT adds its own address to a route being recorded in the packet, before broadcasting the request on to its neighbours. If an MT finds its own address already recorded in the route in the request message, it discards the copy of the request. The main disadvantage of this scheme is the growing size of the address chain in large ad hoc networks, which means that the scheme is not scalable to large networks.

Source routing has been used in a number of contexts for routing in wired networks, using their pre-defined or dynamically constructed source routes. Dynamic Source Routing (DSR) [18] is an on-demand routing protocol that allows nodes to dynamically discover a route in a multi-hop network to any destination. In DSR, mobile nodes are required to maintain route caches that contain the source routes of which the mobile is aware. When new routes are learned, entries in the route cache are immediately updated. When a mobile station needs to establish a connection to another mobile station, it dynamically determines one based on cached information and on the results of a route discovery protocol. An MT initiating route discovery broadcasts a Route Request (RREQ) packet to its one-hop neighbours. The RREQ contains the address of the destination. Every RREQ packet is uniquely identified by the pair <source address, broadcast id>. To limit the number of RREQs propagated on the outgoing links of a

mobile node, it does not forward the RREQ if it had been previously seen it. Therefore, when a mobile node receives a RREQ packet it checks the <source address, broadcast id> pair. It appends its own address to the list of addresses in the route record of the RREQ and forwards the packet to its neighbors, only if it has not previously received a RREQ with an identical <source address, broadcast id> pair, and it is not the destination. A route reply (RREP) packet is sent to the to the source by the destination itself or a node that has can supply a route to the destination, after including a copy of the reversed route record, of the received RREQ, in the RREP packet. Because of the source routing requirement, DSR is not scalable to large networks. Furthermore, the need to place the entire route in both route replies and data packets causes a significant overhead.

Another on-demand routing protocol designed for wireless mobile ad hoc networks is Ad hoc On-demand Distance Vector (AODV) [19]. AODV acquires routes dynamically. When a source node desires a route to a destination for which it does not already have a route, it broadcasts a RREQ packet across the network. Similar to DSR, the RREQ packet contains the source address, a broadcast id, and the destination address. It also contains the destination sequence number, and a hop count. A RREQ is uniquely identified by the pair <source address, broadcast id>. As the RREQ packet travels from a source to various destinations, the reverse path from all the nodes back to the source is set up. To set up a reverse path, a node records the address of the neighbor from which it received the first copy of the RREQ. These reverse route entries are maintained for a sufficient time period to allow the RREQ to traverse the network. Nodes receiving this packet update their information for the source node and set up backwards pointers to the source node in the

route tables. A node receiving the RREQ packet may send a RREP if it is either the destination or if it has a route to the destination with a corresponding sequence number greater than or equal to that contained in the RREQ. If this is the case, it sends a RREP back to the source. Otherwise, it rebroadcasts the RREQ. A RREP packet is unicasted to the node from which the RREQ was received, and the node sets a forward pointer to the issuer of the RREP. Every mobile node forwards the RREP packet to the neighbor to which it previously set a reverse path. By the time the RREP reaches the source, a route is set up from the source to the destination. The source node may now begin to forward data packets to the destination.

The signal stability based adaptive (SSA) routing protocol [58] is a distributed adaptive routing protocol for finding and maintaining stable routes, based on their longevity, in an ad hoc network. In [58], an architecture is presented for the implementation of SSA. The proposed protocol is novel in its use of signal strength and stability of individual hosts as route selection criteria. The objective of this scheme is to reduce route maintenance by using longer-lived routes. In this protocol, a host initiates route discovery on-demand as in DSR and AODV. Only when a route is needed to send data, will the host discover one. The source broadcasts search packets which will propagate to the destination, allowing the destination to choose a route and return a route-reply. Figure 2.1 shows an example ad hoc network that runs SSA. If a link on the path that was chosen by the destination (upper part of Figure 2.1) becomes weak, a new route can be formed (lower part of Figure 2.1). A solid edge indicates a strong link, and a dashed one indicates a weak link. The performance results show that the use of signal strength consistently decreases the

route maintenance required by providing more stable routes. To choose long-lived routes, SSA utilizes the information available at the link layer.

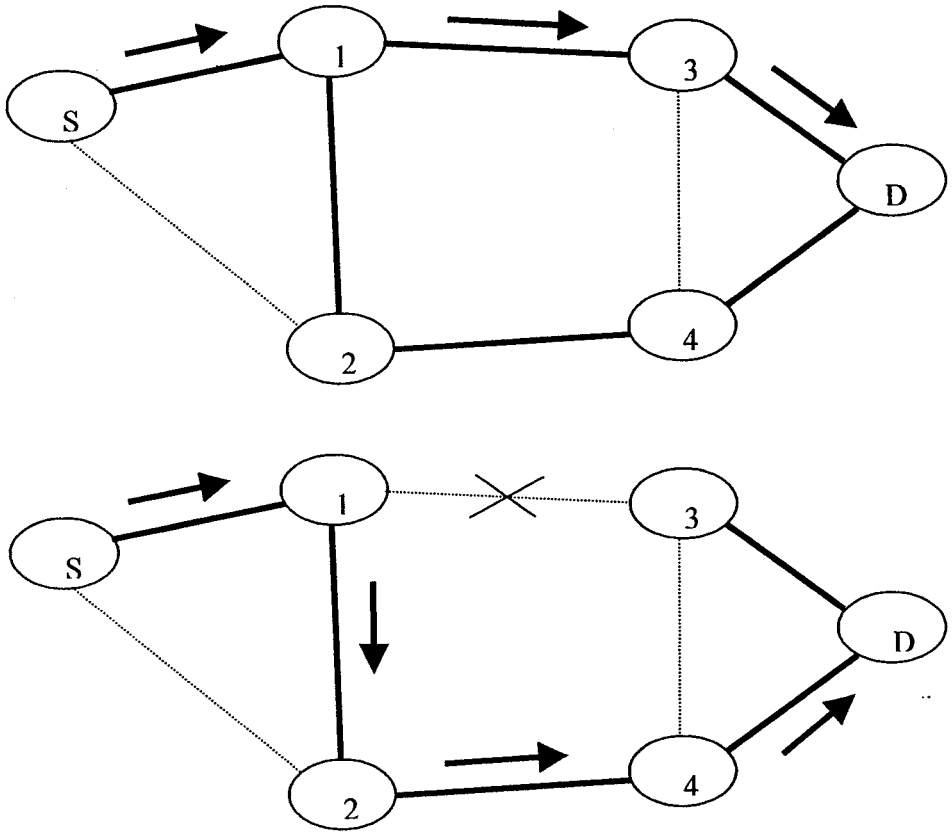


Figure 2.1: Example ad hoc networks running SSA [58]

The signal strength criteria allow SSA to differentiate between strong and weak channels. Each channel is characterized as strong or weak by the average signal strength at which packets are exchanged between the two hosts.

As can be seen in Figure 2.2, it is possible that SSA fails to find a strongly connected route from the source S to the destination D. In this case, a search for any route must be initiated by the source. In such scenarios, SSA functions in exactly the same way as DSR. This introduces a drawback in SSA since more delay is added by looking for a stable route. Only when a source fails to find one, does it look for any route.

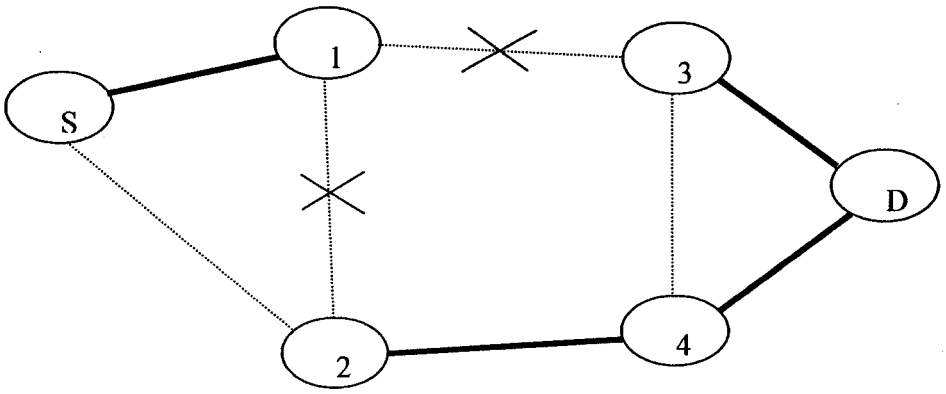


Figure 2.2: The non-existence of a strong route to the destination [58]

The temporally ordered routing algorithm (TORA) [65] is a link reversal algorithm [66] for routing in ad hoc networks. For each possible destination in the ad hoc network, TORA maintains a destination-oriented directed acyclic graph (DAG). In a DAG, starting from any node, the path leads to the sink (destination). TORA uses the notion of heights to determine the direction of each link. Despite dynamic link failures, TORA attempts to maintain the destination-oriented DAG such that each node can reach the destination, as illustrated below. Figure 2.3 illustrates how link reversal is performed in TORA. An arrow connecting a pair of nodes in this figure implies that the two nodes can communicate with one another. That is, the physical link between the two nodes is

bidirectional. However, the TORA algorithm imposes a logical direction on the links, as illustrated in Figure 2.3(a) – this figure shows the destination-oriented DAG with node 6 being the destination. Observe that, starting from any node in the graph, the destination 6 can be reached by simply following the directed links. Now, assume that the link between nodes 4 and 5 breaks (perhaps because node 5 moves away from node 4). Then, in the destination-oriented DAG, node 4 does not have any outgoing logical link. In response, TORA reverses the logical direction of the (4, 2) link, as shown in Figure 2.3(b). Now, node 2 does not have any outgoing logical link. In response, the logical direction of link (1, 2) is reversed, resulting in the graph in Figure 2.3(c). Link reversal continues until a node finds an outgoing link other than the one that has just been reversed by its neighbour. Now since node 1 has an outgoing logical link, (1, 3), each node (other than the destination 6) has an outgoing logical link, and is able to reach the destination node 6 by following the directed links.

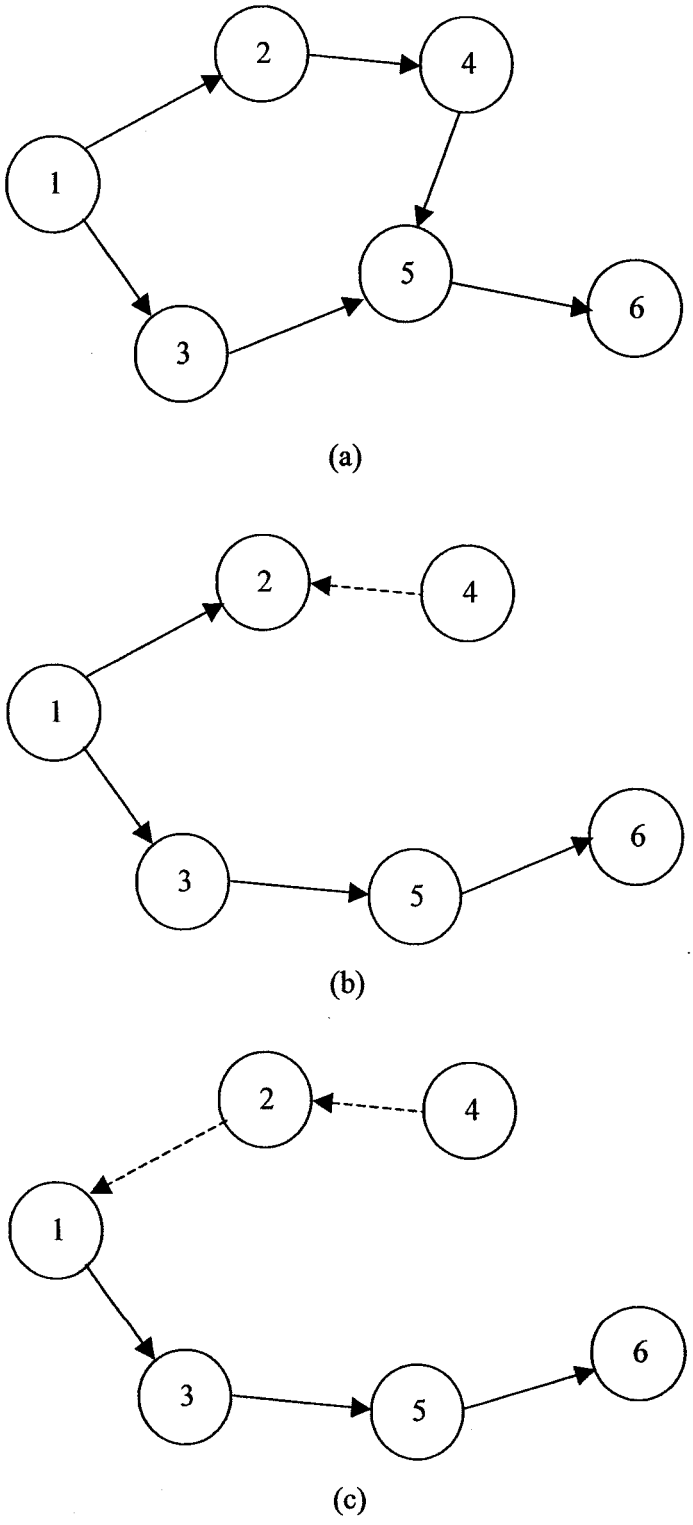


Figure 2.3: Route maintenance in TORA. Reversed links are shown as dashed lines.

Dynamic load-aware routing (DLAR) [76, 77] also builds routes on demand. Unlike other on-demand schemes, the destination learns more than one route to the source prior sending the RREP. The route with the minimal load is then chosen by the destination. During route setup, nodes attach their load information to the RREQ prior to forwarding it. Three different algorithms are proposed for route selection. The number of buffered packets by a node is considered its nodal load. The three schemes differ in the way they calculate the routing cost using the nodal loads. In the first scheme, the routing load of each node is simply added. However, this can cause contention, and reduce the lifetime of the network. Nodes may become part of multiple routes and may thus become congested and not able to deliver packets along the route. Ties are broken via choosing the route with the smallest number of hops. The second algorithm uses the average number of packets buffered at each intermediate node along the path. Ties are broken using the shortest delay. The route with the least number of intermediate nodes that have their load exceeding a certain threshold is selected using the third scheme. Simulation analysis shows that the three DLAR schemes perform very well regardless of the mobility degree. However, the second algorithm achieves the smallest packet delivery fraction, and creates the largest hop count among the DLAR [77] schemes.

As shown above, in route discovery a source uses flooding to acquire a route to its destination. This degrades system-wide energy conservation. The transit nodes, upon receiving a query, learn the path to the source and enter the route in their forwarding tables. Route discovery and maintenance may become inefficient under heavy network load since intermediate nodes will have a higher probability of moving due to the delay in

packet transmissions attributed to MAC contention. Hence, routes will also have a higher probability of breaking as a result of mobility. This wastes battery power, and thus the lifetime of the wireless nodes decreases. Moreover, flooding of route requests and route reply packets in on-demand routing protocols may result in considerable energy drains under a realistic energy-consumption model that takes idle time and promiscuous mode power into account. Every station that hears the route request broadcasts consumes an amount of energy proportional to the size of the broadcast packet. In addition, stations that hear a corrupted version of a broadcast packet still consume some energy. Besides, the simulation studies carried out for table-driven and on-demand routing protocols fall short of providing necessary power-based performance metrics.

Two energy-aware DSR-based protocols are proposed in [83]. The first protocol is called request-delay, while the second is called max-min. The idea of request-delay is borrowed from [84], [85]. Nodes with lower energy levels delay re-broadcasting route request packets, thus decreasing the probability of becoming intermediate nodes. Max-min uses the same concept proposed in [74].

2.3.3 HYBRID AD HOC ROUTING

The Zone Routing Protocol (ZRP) [20] is the only hybrid ad hoc routing scheme in the literature. Routing using ZRP relies on the notion of a routing zone. It is required that every node knows the exact topology within its routing zone. IntrAzone Routing Protocol (IARP) operates inside the routing zones. IARP can use any table-driven protocol such as distance vector or link state routing.

The IntErzone Routing Protocol (IERP) is used for finding routes between the routing zones themselves. It is used to reach the destinations that lie outside the same routing zone as the source. A source node that wishes to send a packet to a destination node sends a query to all the nodes on the border of its routing zone. The border nodes, upon receipt of the query packet, re-broadcast the query packet to their border nodes if the destination is not in their routing zones. This process continues until the packet reaches the destination, or a node that lies in the routing zone of the destination. The receiving node then responds to the query packet and indicates the forwarding path in the reply packet. A zone radius of one hop turns ZRP into a pure reactive protocol. On the other hand, for large values of the zone radius, ZRP becomes a pure proactive protocol. The adjustment of the routing zone parameter is not done dynamically. This is a drawback of this protocol, since there is no way to adapt to the ever-changing demands of the ad hoc network.

2.4 GEOGRAPHY-INFORMED WIRELESS AD HOC NETWORKS

Novel schemes that utilize location information to reduce the routing overhead in ad hoc networks are proposed in [17], [21]. Two Location-Aided Routing (LAR) protocols are presented. These protocols limit the search for a route to the so-called request zone, which is determined based on the expected location of the destination node at the time of route discovery. The LAR algorithms are essentially identical to flooding. However, a node that is not in the request zone does not forward a route request to its neighbours. Thus, implementing LAR algorithms requires that a node be able to determine if it is in the request zone for a particular route request—the two proposed LAR algorithms differ

in the manner in which this decision is made. LAR nodes obtain location information using GPS [22], [23], [63]. The proposed schemes are based on a flooding algorithm that is used to deliver packets to nodes within a location-based multicast group. In the basic scheme [17] a mobile station, upon receiving the multicast packet, compares the multicast region's coordinates with its own location. It accepts the packet if it lies within the multicast region. To prevent routing loops, the node forwards the packet only if it had not previously received it. This method, of sending packets to nodes that lie within specific boundaries by means of flooding, is not efficient.

As an enhancement to the basic flooding method, the notion of a forwarding zone is introduced (see Figure 2.4). The forwarding zone is the smallest rectangle that includes the sender and the multicast region with the sides of the rectangle parallel to the horizontal and vertical axes. To direct other nodes in their forwarding decisions, the source explicitly specifies the forwarding zone. A node forwards the packets intended to the multicast group only if it lies within the forwarding zone specified by the source, unlike the multicast flooding algorithm. In Figure 2.4, the multicast region is represented by the triangle OPB. Therefore, the smallest rectangle that includes the source and the multicast region is SABC, and it represents the forwarding zone. If the source lies within the multicast region, as in Figure 2.5, then, by definition, the forwarding zone is the multicast region itself. Hence, the forwarding zone and the multicast region, in Figure 2.5, are both represented by the rectangle ABCD.

Another location-based multicast scheme was also developed in [17], [21]. The

forwarding zone approach is not used explicitly this time; it is used rather implicitly to determine whether to forward or discard the multicast packets. In order for the mobile nodes to be able to determine whether to forward or drop the packets sent by a source node, the source node must include the following three pieces of information in the multicast packet: the specification for the multicast region, its geometrical center (specifically, its coordinates), and the source's coordinates.

The two proposed enhancements in LAR result in lower message delivery overhead, as compared to flooding. Using GPS is considered a drawback of LAR algorithms as it adds more overhead to the route discovery process since all the mobile terminals will have to be equipped with GPS transceivers.

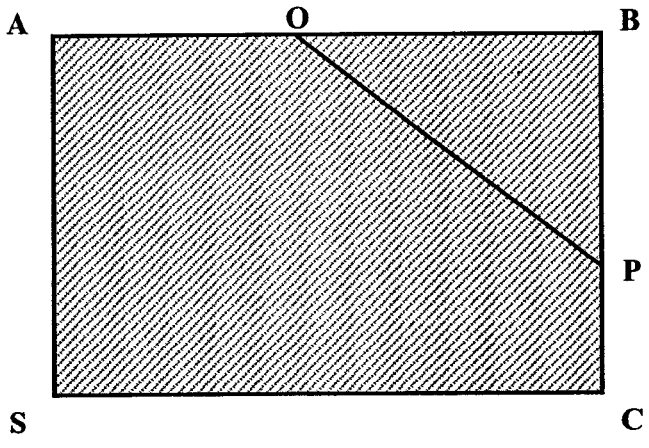


Figure 2.4: LAR, source node outside the multicast region [17, 21]

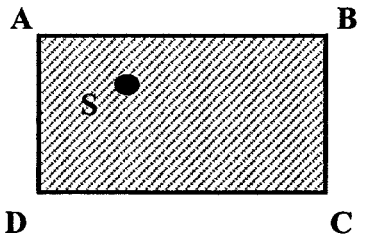


Figure 2.5: LAR, source node inside the multicast region [17, 21]

Geocasting is the problem of sending packets to the set of all nodes in a specific geographical region [61]. Unlike multicasting, a host does not join a geocast group. Instead, it automatically becomes a member if its location belongs to the geographical region specified for the group. A geocasting protocol based on flooding is proposed in [62].

GeoTORA [64], [39] is based on the TORA routing protocol. TORA is modified to be able to perform anycast [67, 68] and the GeoTORA protocol is then obtained using a small variation of the anycasting protocol. In anycasting, the message is delivered to any member of the anycast group. The anycasting algorithm would maintain a single DAG for a given anycast group. A DAG in which all anycast nodes are sinks is maintained. Figure 2.6 illustrates the anycast scheme. In this case, assume that nodes 1 and 2 belong to the anycast group. The present DAG structure is shown in Figure 2.6(a). If node 3 moves, breaking link (1, 3), the logical direction of link (3, 4) is reversed, resulting in the DAG shown in Figure 2.6(b). Now all nodes that are outside the anycast group have an outgoing link (and a path to at least one node in the anycast group).

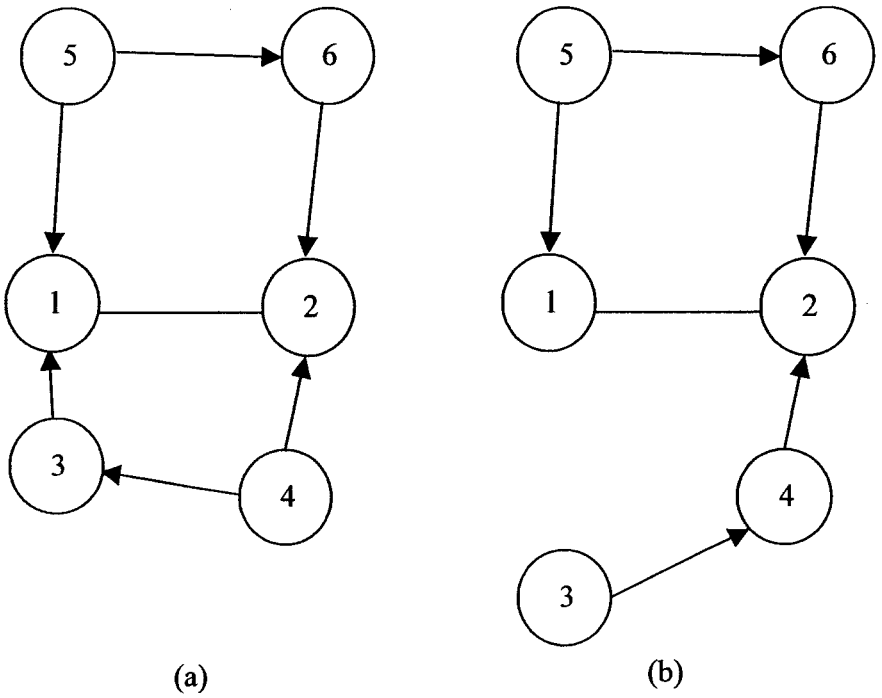


Figure 2.6: Anycasting using modified TORA

To perform geocasting using GeoTORA, first, a sender node essentially performs an anycast to the geocast group members – similar to the above anycast protocol. In GeoTORA, logical directed links are set up such that a node wishing to perform a geocast can reach any one node in the geocast group by simply forwarding the packet on any outgoing link. When any node in the geocast group receives the packet, it floods the packet such that the flooding is limited to the geocast region. For instance, if node 9 in Figure 2.7 wants to perform a geocast, it forwards the geocast packet to node 8, along the outgoing link (9, 8). Node 8, in turn, forwards the packet to node 2. Since node 2 is in the geocast region, it initiates a flooding of the packet limited to the geocast region. Nodes 1 and 4, on receiving the packet from node 2, forward the packet to their neighbors. When node 2 receives the packet from node 1 or 4, it does not forward the packet, since node 2

has already once forwarded the packet to its neighbours. In this manner, the packet will reach nodes 1, 2, 3, and 4 that belong to the geocast region. As expected, Geocast flooding performs better than GeoTORA in terms of packet delivery. However, Geocast flooding incurs more packet delivery overhead. This is because GeoTORA limits the scope of flooding to the geocast region. Besides, since the overhead of GeoTORA consists of data packets as well as control packets used to create and maintain routes, the overhead increases with increasing node mobility or, equivalently, decreasing pause times.

The studies based on an energy model that considers energy dissipation in sent/received packets, and idle time, suggest that energy optimizations must turn off the radio, and not simply reduce packet transmission and reception. In [51], measurements show standby:receive:transmit ratios of 1:1.05:1.4. Ratios of 1:2:2.5 and 1:1.2:1.7 are reported elsewhere ([52] and [54]). Consider, for example, a single-hop ad hoc network (like the one shown in Figure 2.8) consisting of k nodes, and a MAC scheme that uses an RTS-CTS-DATA-ACK message exchange sequence. Any transmission by one node will be heard by all the k-1 other nodes. Consequently, a single successful packet transmission and reception would consume E_RTS_{Tx} + (k-1).E_RTS_{Rx} + E_CTS_{Tx} + (k-1).E_CTS_{Rx} + $E\_P\text{-}2\text{-}P\_Data_{Tx}$ + (k-1).$E\_P\text{-}2\text{-}P\_Data_{Rx}$ + E_ACK_{Tx} + (k-1)E_ACK_{Rx}. This is in addition to the energy dissipated by all k nodes during the defer/backoff period.

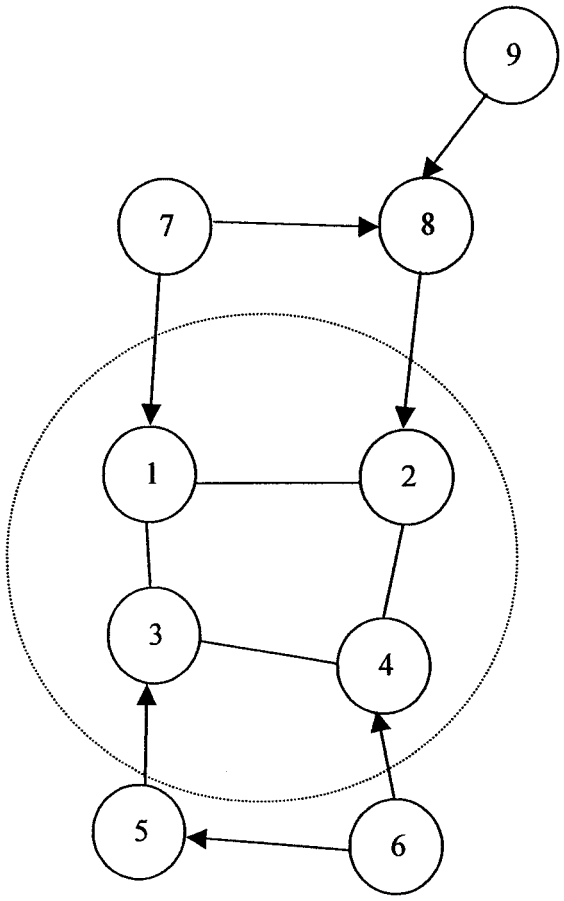


Figure 2.7: Geocasting using GeoTORA

Therefore, it is important to turn off the radio whenever possible. Powering off the radio conserves energy both in overhearing due to data transfer, and in idle state energy dissipation when no traffic exists [98], [99]. Therefore, it is recommended that nodes try powering off their radio most of the time.

Power-Aware Routing Optimization (PARO) [81] adjusts the transmission power and chooses the route that results in the minimum total transmission power. Using PARO,

nodes are required to stay awake and idle time power consumption is not accounted for. Thus, this is somewhat limited and of minor practical value.

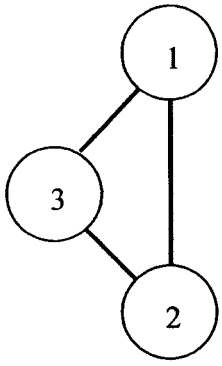


Figure 2.8: Wireless NIC Energy consumption in a single-hop scenario

Achieving energy conservation via turning off the radio is further explored in [69]. This approach is similar to the use of Time Division Multiple Access (TDMA) for power conservation [70], and the Power-Aware Multi-Access protocol with Signaling PAMAS [54]. However, unlike these approaches, the proposed scheme employs information from above the MAC layer to control radio power. Obviously, when there is significant node redundancy in an ad hoc network, multiple paths exist between nodes. Thus, some intermediate nodes can be powered off while still maintaining connectivity. For example, in Figure 2.9, if node 1 is awake, nodes 2 and 3 are extraneous for communication between nodes S and D.

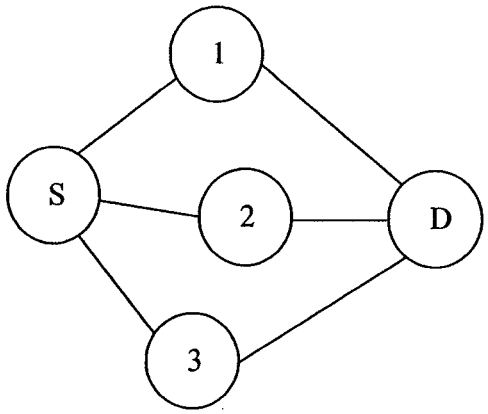


Figure 2.9: Example of node redundancy in ad hoc routing

However, it is not trivial to find equivalent nodes routing-wise in an ad hoc network. Nodes that are equivalent between a pair of nodes may not be equivalent between others. In Figure 2.10, although nodes 1 and 2 are equivalent for communication between S and D, only node 1 is acceptable between S and 3 since S may reach 3 only via node 1.

The proposed scheme, namely geographical adaptive fidelity (GAF), makes use of application and system information to turn off node radios for extended periods of time. Node duty cycles are influenced by application endpoints and node movement patterns to preserve communication fidelity. In addition, it also makes use of node deployment density to adaptively adjust routing fidelity. Denser node deployment is used to extend the lifetime of the wireless ad hoc network. GAF uses location information and virtual grids to determine node equivalence. Location information used in GAF is provided by GPS. GAF addresses the node equivalence problem by dividing the whole area where nodes are distributed into small virtual grids. The grid is defined such that all the nodes in adjacent grids can directly communicate together. Thus, nodes that lie in one grid are

equivalent from a routing point of view. Thus, in Figure 2.11, only 3 nodes need to be awake at any one time to maintain routing fidelity. This means that the number of nodes that need to be awake is equivalent to the number of virtual girds.

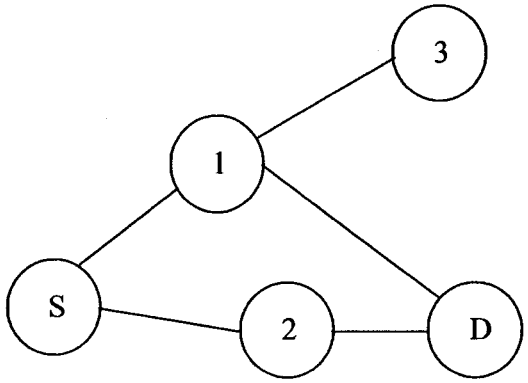


Figure 2.10: The problem of node equivalence

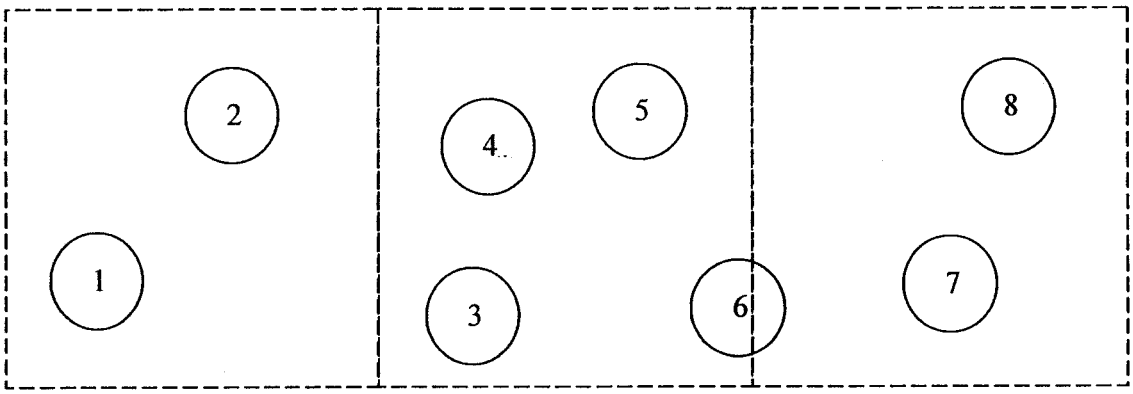


Figure 2.11: Example of virtual grids in GAF

2.5 MEDIUM ACCESS AND ENERGY CONSERVATION

Cellular networks have a wealth of centralized MAC schemes such as TDMA/TDD [102], MASCARA [103], DS/CDMA [104], and others [43]. However, the wireless

stations in an ad hoc network must resolve the problem of simultaneous medium access in a distributed manner. Likewise, wireless ad hoc stations must resolve the problem of hidden terminals. A neighbour of the destination that is out of the wireless range of the source may interfere with the transmissions of the source. In Figure 2.12, H is the hidden terminal [55]; if H and S transmit to D at the same time, then a collision takes place. It is noteworthy that neither the source nor the interferer are aware of the collision. It is only the lack of a positive acknowledgement (ACK) from the destination that notifies them of a possible collision and triggers the source to retransmit.

A node that is in range of the sender but not the receiver is called an exposed terminal. An exposed terminal can transmit at the same time as the sender, without causing a collision to occur. In spite of that, using the traditional carrier sense multiple access (CSMA) scheme, the exposed terminal defers from accessing the channel. As a result, the capacity of the wireless ad hoc network is reduced. Hence, energy is wasted needlessly.

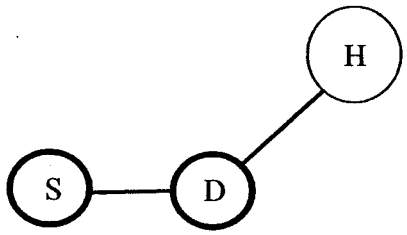


Figure 2.12: The hidden terminal problem in wireless networks

In [55], a scheme that solves the hidden terminal problem using a busy tone is proposed. The protocol, named busy tone multiple access (BTMA), was developed for cellular networks. A busy tone is sent only while the base station is receiving. Thus, transmitters

are prevented from accessing the channel. All nodes, including hidden terminals, within the cell site receive the busy tone and back-off. Wireless nodes are allowed to transmit only in the absence of busy tones. Once a node detects a busy tone on the secondary channel, it cannot use the data channel, even if no signal is detected. This implies that a node is allowed to use the data channel even while another node's transmissions are being heard. Apparently, the BTMA protocol may not be used in a multi-hop wireless ad hoc network since there are no base stations. In addition, BTMA uses two widely separated channels instead of one and increases the hardware costs and complexity of the Network Interface Card (NIC).

In contrast with BTMA, Dual busy tone multiple access (DBTMA) [36], [37], [57] uses two busy tones instead of one. In particular, the hidden terminal problem is resolved using the receive busy tone signal, whereas the exposed terminal problem is overcome using the transmit busy tone signal. DBTMA wastes battery capacity by forcing the wireless node to continuously sense the medium for the transmit busy tone and the receive busy tone signals. A more energy-efficient MAC protocol would consider turning off the transceiver during standby time to save power.

In multiple access collision avoidance (MACA) [48], a node wishing to transmit a data packet to a neighbour first sends a request-to-send (RTS) frame to the neighbour. All nodes that receive the RTS are not allowed to transmit. Upon receipt of the RTS, the neighbour that the RTS was sent to replies with a clear-to-send (CTS) frame. Also, any node that hears the CTS transmission is prevented from using the channel. Hence, the

RTS-CTS message exchange clearly alleviates the hidden terminal problem present in wireless networks. Due to this scheme, data frames are, at least in theory, delivered collision-free. As a result, collisions can only affect control packets. In this case, the IEEE exponential backoff is used to resolve MAC contentions for control packets. In practice, however, collisions may still affect data frames. Nodes that do not properly receive a CTS frame are eligible to use the medium and their transmissions might overlap with those of the source. Hence, this MAC scheme will only decrease the probability of data collisions, and data remains vulnerable to corruption. From the standpoint of energy consumption, corrupted CTS frames result in idle time energy losses for those neighbours that successfully receive the CTS frame. The neighbourhood-wide energy loss is proportional to the number of neighbours of the sending station. MACA also does not use link-layer positive or even negative ACKs, but rather end-to-end ACKs. It is worth noting that the transmission of an RTS frame considerably reduces the energy costs of data collisions. RTS frames result in a favourable reduction in the consumed energy in case of a collision, as compared with the consumed energy, in addition to the larger delay due to collision time, if the actual data frame is sent. These energy and delay savings are achievable in most cases. The use of the RTS frame is nullified whenever the size of data is comparable to that of the RTS frame. A threshold is used to specify the size of the data frames for which an RTS frame ought to be sent.

MACAW [49], on the other hand, uses link layer ACKs to increase data throughput. The IEEE 802.11 MAC and PHY standard [33], [47], as in MACAW, utilizes link layer positive ACKs for all unicast traffic. The ACK frame allows the conveyance of fast

positive acknowledgements, and fast recovery in the event of its absence. ACK frames may only be used with unicast traffic. ACK frames sent in response to a broadcast message have a large collision probability and waste energy and network bandwidth. In a wireless ad hoc network, almost all the proposed routing protocols rely to a large extent on broadcasts. Periodic updates in table-driven schemes, and route discovery, route setup, and route maintenance messages in on-demand protocols are all examples of network-level broadcasts. An extensive amount of battery capacity would be wasted on ACK transmissions and ACK collisions throughout the wireless network. Apparently, because of the use of explicit ACKs, the exposed terminal problem was reintroduced, and thus the energy supply of the wireless nodes is gradually depleted by idle time power consumption. If node 1, or node 4, in Figure 2.13, wants to send a data packet to node 5, it will needlessly wait until the end of the transmission. This contributes to increasing the total idle time energy [100].

In MACA by-invitation (MACA-BI) [56], data must be foreseen beforehand by the receivers. Hence, MACA-BI may only be used by periodic, and not unpredictable, traffic. In the presence of bursty traffic, more efficient protocols have to be used. This is true since it is impossible for the receivers to predict the time instances at which the transmitters send their data frames. Incorrect predictions cause all the neighbouring nodes to waste an amount of energy proportional to the size of the ready-to-receive (RTR) frame, which is sent by the receiver in place of the CTS frame. The sender then responds with the actual data, and the use of the RTS frame is rendered obsolete. In a wireless ad hoc environment, in which nodes are allowed to move freely at all times, anticipation of a

source's transmissions would be extremely difficult due to the unpredictable patterns of contention for medium access by nodes neighbouring to the source. Therefore, it would still not help the receiver to know the transmission schedule of the sender. Complex application-level and contention prediction schemes would be required. The authors claim that the RTR-DATA dialogue achieves improved performance over MACAW. As previously mentioned, the efficiency of this scheme relies on the ability to predict when the sources have data to send.

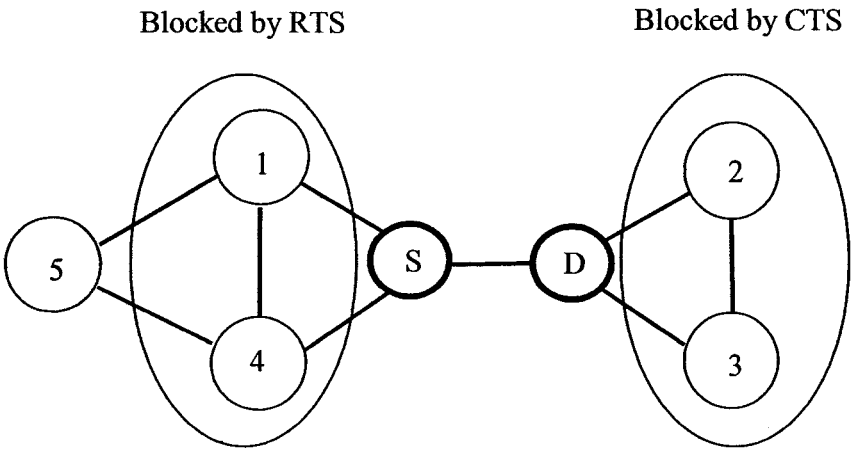


Figure 2.13: MACAW suffers from the exposed terminal problem

In all the above MAC schemes, nodes consume most of their battery capacity while in their idle states, that is, while doing nothing. No special MAC-based energy conservation measures are adopted by any of them. Unlike other approaches that employ information from above the MAC layer to control radio power, the Power-Aware Multi-Access protocol with Signaling (PAMAS) [54] is an energy conservative MAC protocol proposed for wireless ad hoc networks. In addition to the primary data channel, PAMAS uses a secondary channel, called the signaling channel [50], for control traffic. The

neighbours of the traffic pair are prohibited from using the channel during an ongoing transmission via an RTS-CTS dialogue. A node sets its wireless NIC to sleep mode in case it overhears a neighbour's transmission. It is not clear if the energy-conserving behaviour of PAMAS negatively affects the end-to-end delay. In this scheme, a current receiver can force a neighbour to defer using the medium by sending a busy tone over the signalling channel. This can be used to force the sender of an RTS to defer. All nodes that successfully receive the busy tone are prevented from using the medium for the period of the data transmission over the primary channel. Even if the RTS is detected by the receiver, its CTS transmission is corrupted by that of the busy tone since they are both sent simultaneously over the control channel. Since most of the consumed energy in a wireless ad hoc network is attributed to the idle time, PAMAS obtains high energy-conservation gains only in networks with high traffic loads. In networks with low to modest traffic, the energy-based performance of PAMAS is close to that of other CSMA protocols. Therefore, nodes running PAMAS will still consume most of their battery capacity while in their idle states. Based on the aforementioned, PAMAS takes no measures so as to reduce idle time energy consumption unless there is an ongoing transmission.

2.6 UNCONVENTIONAL ENERGY CONSERVATION

Energy-conserving algorithms that operate above on-demand ad hoc routing protocols, such as DSR and AODV, introduce a trade-off between routing fidelity and extended network lifetime. Such algorithms promise minor modifications to the underlying routing protocols. The major contribution of such algorithms lies in turning off the wireless

transceiver while being idle. Two application-level energy conservation algorithms designed to run on top of existing on-demand schemes are proposed in [71]. Mobile nodes running the basic energy-conserving algorithm (BECA) power-off as long as possible. BECA does not take node density into consideration. A node is interrupted from its sleeping state only if it has data to send. The on-demand routing protocol must retry route requests a number of times in order to increase the probability of getting at least one route request through. This is the main disadvantage of this algorithm as it obviously becomes a source of energy loss and MAC contention. Evidently, route requests may collide, and none may ever successfully get through. Consider, for example, if the neighbour had been sleeping, and the route requests sent while it was listening were corrupted by another neighbour's transmissions.

The second application-level scheme is the adaptive fidelity energy-conserving algorithm (AFECA). In contrast with BECA, AFECA utilizes node density to switch off radios for longer time durations. Therefore, an estimation of the number of neighbors must be carried out by each node. Neighborhood size is not measured accurately by AFECA. As a result, the neighborhood size may be underestimated by a node. Underestimation of neighborhood size causes a node to sleep for shorter periods, and hence consume power needlessly. On the other hand, its overestimation will increase the node's sleep time and will in many cases result in routing latency, or even packet losses.

Span [78] is a distributed algorithm where nodes make local decisions on whether to sleep, or to join a forwarding backbone as a coordinator. Each node bases its estimate on

how many of its neighbour will benefit from it being awake. However, the performance evaluation experiments show that Span has higher loss rates than IEEE 802.11. Plus, using fewer nodes to forward packets may decrease potential channel utilization since each time a node exponentially backs off to avoid collision, there is a greater chance that the channel becomes unoccupied for a longer period of time. Span also elects more coordinators than is necessary. This is due largely to non-uniform node density which often causes more nodes to become coordinators.

In [82], a new power-based cost metric is used. The metric is a function of the remaining battery capacity and the number of one-hop neighbors. Nodes that are running low on battery power are allowed to increase the cost of using them as routers proportional to the number of their neighbors. Nodes are also allowed to sleep as long as they do not lie on any active routes and have no frames in their MAC queue. However, deciding how long a node should turn off its radio imposes a tradeoff.

The objective of [87] is to develop a metric that is useful for energy conservation and conventional performance metrics as well. This was motivated by the work in [88], [89]. Finding an optimal, or sub-optimal, value for the weighting parameter in the energy-performance metric is non-trivial. A value of one is shown, empirically, to be the most reasonable. However, there exists no solid guarantee that it is indeed the best value.

A scheme, namely Energy-Dependent Participation (EDP), in which a node's participation in the network is based on the ratio of its residual energy to its initial battery capacity is proposed in [90]. On the downside, there is no consideration for the network density. EDP may therefore result in temporary network partitioning in dynamic sparse networks after intermediate nodes transit to their sleep states. Thus, optimizing the sleep period is of critical value here.

2.7 SUMMARY AND CONCLUSIONS

In a wireless multi-hop ad hoc network, the lack of a fixed infrastructure dictates that the wireless stations must always be awake. On the contrary, in a wide area or local area cellular environment, wireless stations may be scheduled to sleep. Base stations and wireless access points, being central controllers, will be in charge of buffering all incoming packets to sleeping nodes. Thus, in a wireless ad hoc network, wireless stations may not sleep. Consequently, all of the wireless nodes may consume power unnecessarily due to overhearing the transmissions of their neighboring nodes.

The hidden terminal problem in multi-hop wireless ad hoc networks is alleviated using the RTS-CTS dialogue, which only represents a partial solution. As a result, savings in energy and delay are achieved, but collisions may still take place. Nevertheless, the use of the RTS frame is nullified whenever the size of data is comparable to that of the RTS frame. Although an exposed terminal may transmit at the same time as the sender,

without causing a collision, the use of link-layer ACKs in the current IEEE 802.11 standards inhibits exposed terminals as well from using the medium.

Amongst all the proposed wireless mobile ad hoc routing protocols, Dynamic Source Routing (DSR) and Ad hoc On-demand Distance Vector (AODV) are the most eminent. Compared to the rest of the ad hoc routing protocols, their reported performances are marked by distinction [25]. Therefore, both protocols, i.e., DSR and AODV, are candidates for standardization by the Internet Engineering Task Force (IETF) Mobile Ad hoc NETworking (MANET) working group [26]. Nevertheless, hierarchical and hybrid routing techniques may prove useful for wireless ad hoc networks [45, 46].

Unnecessary periodic updates in table-driven routing protocols may have a negative effect on energy conservation. Nevertheless, the flooding of route request and route reply packets in on-demand routing protocols also results in considerable energy drains under a realistic energy-consumption model that takes idle time and promiscuous mode power into account. Every station that hears the route request broadcasts will consume an amount of energy proportional to the size of the broadcast packet. In addition, stations that hear a corrupted version of a broadcast packet will still consume some energy.

The need for energy conservation reaches beyond routing and medium access. In recent studies, assessment of the neighborhood density and incoming traffic has been realized as factors that influence energy-conservation and wireless nodes turn off their transceivers

accordingly. As a result, such approaches have higher packet loss rates than schemes that require nodes to be constantly awake.

CHAPTER 3

PLCP AND MAC-BASED PERFORMANCE EVALUATION OF ENERGY-AWARE ROUTING SCHEMES

Although wireless mobile ad hoc stations have limited battery capacity and, consequently, ad hoc routing schemes ought to be energy conservative, previous simulation studies carried out for table-driven and on-demand ad hoc routing protocols fall short of examining essential power-based performance metrics. Besides traditional performance metrics, we must examine important energy-based performance metrics, including average node and network lifetime, energy-based protocol fairness, average dissipated energy per protocol, and standard deviation of the energy dissipated by each individual node. In this chapter, we present a thorough energy-based performance study of power-aware routing schemes for wireless mobile ad hoc networks. Our energy

consumption model is based on a detailed implementation of the IEEE 802.11 physical layer convergence protocol (PLCP) and medium access control (MAC) sublayers. To our best knowledge, this is the first such detailed performance study. Moreover, we discuss the implications that our conclusions have on energy conservation in wireless mobile ad hoc networks.

3.1 INTRODUCTION

Most of the research performed in the field of wireless ad hoc networks has focused on the problem of routing. Bandwidth efficiency and end-to-end delays are the main concern of the developed protocols and the conducted simulation studies. Due to the limited energy capacity of the wireless devices and the importance of power awareness in ad hoc routing protocol design, a thorough energy-based performance study is essential to any bandwidth-based study.

Our performance study is based on a complete simulation of the IEEE 802.11 MAC sublayer. In addition to the energy spent on cooperatively resolving the problem of simultaneous medium access, wireless ad hoc stations must resolve the problem of hidden terminals. As explained earlier, the capacity of the wireless ad hoc network is reduced as a result of hidden and exposed terminals and, thus, energy is wasted groundlessly.

This chapter is organized as follows. The following section surveys the proposed routing schemes for wireless ad hoc networks. Section 3.3 describes the energy model adopted in the chapter. Our simulation experiments follow in Section 3.4. Finally, Section 3.5 is the chapter summary and conclusions.

3.2 ROUTING AND ENERGY CONSERVATION

In this chapter, we use the broad classification of routing protocols in wireless mobile ad hoc networks into table-driven and on-demand. As explained in Chapter 2, in table-driven routing protocols, all the mobile stations are required to have complete knowledge of the network through their periodic and incremental/triggered updates. Unnecessary proactive periodic updates may have a negative effect on energy conservation.

The use of flooding to acquire routes in on-demand routing degrades system-wide energy conservation. In on-demand protocols, route discovery and maintenance may become inefficient under heavy network load since intermediate nodes have a higher probability of moving due to the delay in packet transmissions attributed to MAC contention. Hence, routes also have a higher probability of breaking as a result of mobility, and battery power is wasted. It is worth mentioning that the flooding of route request and route reply packets in on-demand routing protocols may result in considerable energy drains under a realistic energy-consumption model that takes idle time and promiscuous mode power into account. Every station that hears the route request broadcasts consumes an amount of

energy proportional to the size of the broadcast packet. In addition, stations that hear a corrupted version of a broadcast packet will still consume some amount of energy.

A number of routing proposals for ad hoc networks take energy conservation into consideration so as to extend the lifetime of the wireless nodes by wisely using their battery capacity [73], [74], and [75]. In [73], minimum total power routing (MTPR) is proposed. If the total transmission power for route R is P_R, then the route can be obtained from $P_{MTPR} = \min_{R \in S} P_R$, where S is the set containing all possible paths. On the downside, this approach tends to select routes with more hops than other approaches. This is realizable since transmission power is inversely proportional to distance [73]. Thus, more energy may be wasted network-wide because a larger number of nodes are now involved in routing as all nodes that are neighbors to the intermediate nodes will also be affected, unless they were in sleep mode.

Minimum battery cost routing (MBCR) [74] utilizes the sum of the inverse of the battery capacity for all intermediate nodes as the metric upon which the route is picked. However, since the summation must be minimal, some hosts may be overused because a route containing nodes with little remaining battery capacity may still be selected. Min-max battery cost routing (MMBCR) [74] treats nodes more fairly from the standpoint of their remaining battery capacity. Smaller remaining battery capacity nodes are avoided and ones with larger battery capacity are favored when choosing a route. The route is found using the equation

$$P_{MMBCR} = \min_{R \in S}[\max_{n \in R}(1/Battery\_Capacity_n)]$$

However, more overall energy is consumed throughout the network since minimum total transmission power routes are no longer favored. In [75], MTPR is used when all the nodes forming a path (note that one path is sufficient) have remaining battery capacities that are above a so-called battery protection threshold, and MMBCR is used if no such path exists. The joint protocol is called conditional max-min battery capacity routing (CMMBCR).

Wireless ad hoc networks composed of wireless nodes with an unlimited supply of power are studied in [91]. [92], [93], [94], [95], and [96], addressed minimum-energy routing. The approach is to minimize the total energy expenditure to reach the destinations. Consequently, some nodes are overused and their batteries are thus depleted. In [97], bounds on the system lifetime were derived for static sensor networks. Nevertheless, these derivations may not be used for wireless mobile ad hoc networks.

3.3 ENERGY CONSUMPTION MODEL

In a multi-hop ad hoc network, or in a wireless LAN operating in independent basic service set (IBSS) mode, wireless stations must always be ready and willing to receive traffic from their neighbors. Base stations and wireless access points will be in charge of buffering all incoming packets to sleeping nodes. Thus, in a wireless ad hoc network, due to the nonexistence of a base station or central controller, the network interface card cannot sleep in ad hoc mode. Likewise, all of the wireless nodes will consume power

unnecessarily due to overhearing the transmissions of their neighbors. Although this obviously wastes an extensive amount of the total consumed energy throughout the lifetime of the wireless station, on-demand routing protocols require that nodes remain powered on at all times so as to participate in on-the-fly route setup using route request broadcasts and route reply packets. Besides, table-driven protocols also require constant operation in the active state in order to exchange periodic updates and participate in packet routing.

In [51], measurements show idle:receive:transmit ratios of 1:1.05:1.4. Ratios of 1:2:2.5 and 1:1.2:1.7 are reported elsewhere ([52] and [54]). Nonetheless, we adopt the most recent results reported in [107], which are shown in Table 3.1 below for a 2 Mbps wireless network interface card. However, unlike the linear model in [107], our study takes into account transmission failures. This is critical to the proper quantification of energy losses in a wireless ad hoc network. Intuitively, idle-time energy dissipation will completely dominate system-wide energy consumption.

Table 3.1
2 Mbps power consumption figures

Mode	Measurement
Transmit	1.327 W
Receive	966.96 mW
Idle	843 mW

In our simulations, we provide a simple implementation for the PLCP sublayer. This is useful for our energy-based study so as to quantify most of the sources of energy loss since PLCP preambles and headers are transmitted at lower rates than those used for data

frames. The PLCP preamble and a PLCP header are always attached to the start of each packet for synchronization purposes. Also, the length of the data packet and the rate are both obtained from the PLCP header by the receiver. To enable a lower-rate network interface card to communicate with a higher-rate one, the PLCP preambles and headers are transmitted at the smallest allowable rate, 1 Mbps. Taking the 1 Mbps PLCP rate into account in the performance experiments is of utmost importance since the lower the rate the more the consumed energy. Various IEEE 802.11 parameters were considered in the study. Table 3.2 lists the IEEE 802.11 delay parameters, for direct sequence spread spectrum (DSSS), used in our study:

Table 3.2
Interframe spaces for DSSS

InterFrame Space	Value Used in Simulations
Short InterFrame Space (SIFS)	10 μs
Random Backoff Slot	20 μs
DCF InterFrame Space (DIFS)	50 μs

3.4 SIMULATION EXPERIMENTS

A packet-level discrete-event simulator was developed in order to monitor, observe and measure the performance of five routing schemes: minimum-hop routing, MTPR, MBCR, MMBCR, and CMMBCR. All the simulation experiments were conducted for ad hoc networks with 30 mobile nodes. Initially, each mobile station was assigned a unique node ID and a random position in the x-y plane. All the data packets had a fixed size of 1024 octets long so as to avoid fragmentation of longer packets. SIFS delays were used

for control frames (RTS, CTS, and ACK), whereas DIFS delays were used between successive data frames.

3.4.1 PERFORMANCE METRICS

The simulator measures the following noteworthy statistical performance metrics:

1. *Average Dissipated Energy* – This is measured in W.s (Joules). The smaller this value, the more energy-efficient the routing protocol is, and vice versa.
2. *The Standard Deviation of the Dissipated Energy* - This is also in Joules. The computed value reflects whether the routing scheme overused any number of nodes or not. This is a very important performance measure since it is a measure of protocol fairness. Therefore, the sought value for this metric is one that is as close as possible to zero. The smaller this value, the fairer the routing criterion is.
3. *Energy Consumption for Control Frame Transmissions and Receptions* – This shows the percentage of the total energy consumed throughout the simulation by all stations for RTS transmissions, RTS receptions, CTS transmissions, CTS receptions, ACK frame transmissions, and ACK frame receptions. These experiments show which frame type dominates energy consumption at the control frame level.
4. *Energy Consumption for Defer, Data Transmissions and Receptions* - This shows the percentage of the total energy consumed throughout the simulation by all stations for defer periods (including SIFS, random backoff, and DIFS), point-to-point data transmissions, point-to-point data receptions, broadcast data transmissions, and broadcast data receptions.

The simulation experiments are set up for wireless mobile ad hoc networks that cover a 100 x 100 unit grid. The wireless transmission range of the mobile nodes was 20 units. The velocity of the mobile nodes was uniformly distributed between 0 and 10 units/second, and they were allowed to move randomly in any direction with two mobility models. The first model uses pause periods to simulate a low-mobility scenario, whereas the second uses no pause periods and simulates a high-mobility scenario. The results of the corresponding experiments were collected for all the five routing schemes and were carefully examined.

3.4.2 SIMULATION RESULTS

Unexpectedly, the simulation results shown in Figure 3.1 show that minimum-hop routing achieves good performance under both mobility scenarios; 3.79 and 3.89 Joules of energy were dissipated on average, respectively. This implies that in most cases the shortest routes were also efficient from the standpoint of energy consumption. This is because minimum-hop routing almost always involves the smallest number of wireless stations so the total energy consumed system-wide will be close to minimal. MBCR consumes more energy than the other four schemes. With MBCR the nodes that lie on the minimum battery cost paths can be easily overused and this results in contention at the MAC level, which means that the overused nodes consume more energy while deferring for a busy medium. On the other hand, MMBCR produces the least average consumed energy under low mobility, whereas the drained energy increases slightly under high mobility. This is because as nodes move, a chosen route based on the remaining battery capacity might no longer be the best route from the point of view of the MAC layer since

more nodes might be simultaneously contending for medium access in the vicinity of the intermediate nodes. Other, possibly, shorter routes might be produced due to the movement of the wireless nodes, but, the chosen route remains the same. In such scenarios, minimum-hop routing chooses the shorter route. In other words, the chosen route using MMBCR might in some cases be a much longer route that happens to have the smallest maximum battery cost among all possible routes. Minimum-hop routing consumes around 58% more energy than MMBCR under low mobility, whereas MMBCR consumes approximately 37.5% more under high mobility. Therefore, the trade-off here is between system-wide energy efficiency and fair battery usage.

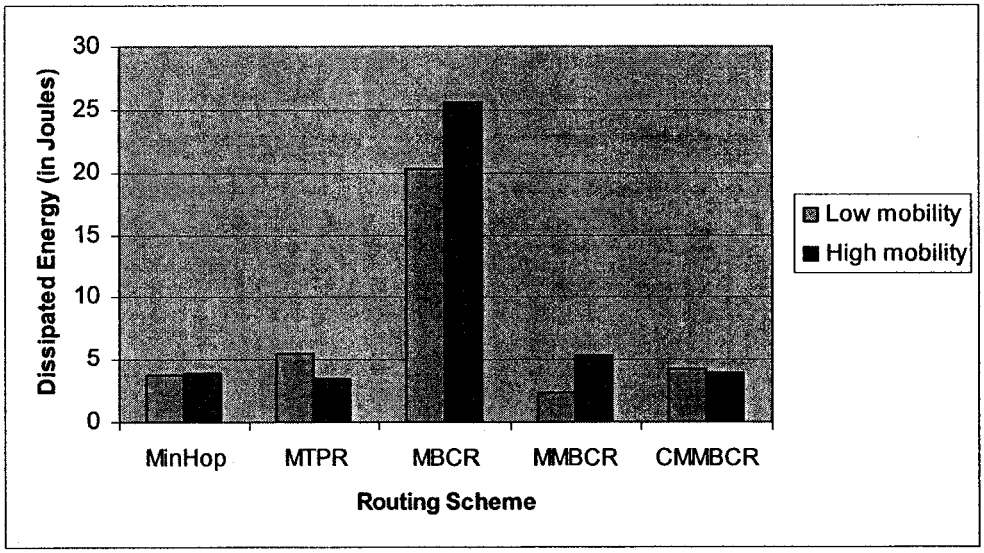


Figure 3.1: Average dissipated energy (in Joules)

Figure 3.2 shows the standard deviation of the dissipated energy in Joules. Surprisingly, minimum-hop routing exhibits the best performance among the five protocols. Upon careful examination, all the other schemes (except MTPR) put great emphasis on the

remaining battery capacity rather than the amount of energy drained thus far due to the communication-related activities in which the wireless stations are involved. A node running MBCR, MMBCR, or CMMBCR with large remaining battery capacity (and hence small battery cost) is prone to be overused as compared to the nodes with the least remaining battery capacity. The standard deviation is 0.93 Joules in the case of minimum-hop routing, whereas it is 1.88 Joules in the case of MMBCR which scored the second smallest standard deviation. As expected, based on the results of Figure 3.1, MBCR scores the worst value for the standard deviation and is hence the most unfair scheme. The standard deviation for CMMBCR lies between that of MTPR and that of MMBCR. In our simulations, the battery capacity protection threshold was set to half the maximum battery capacity. This explains the value of the standard deviation for CMMBCR.

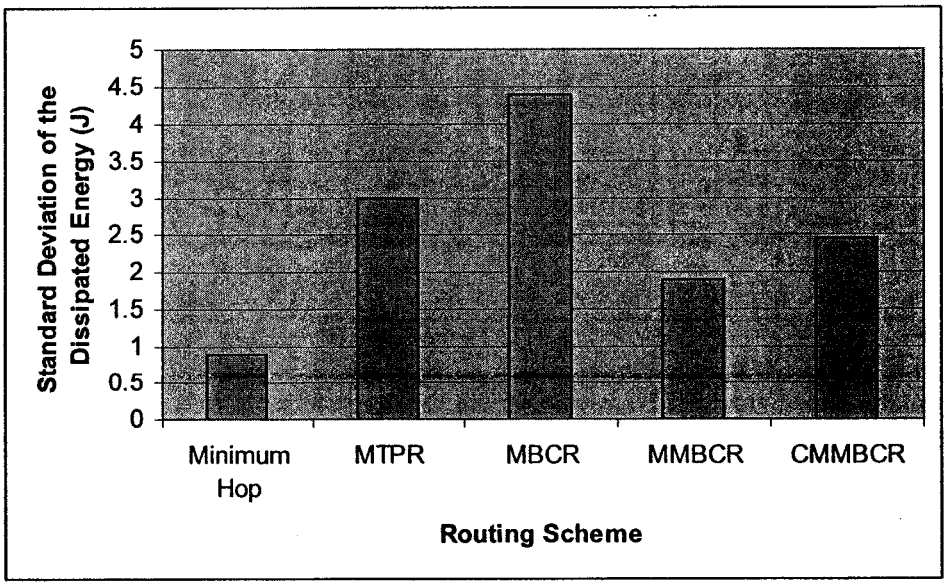


Figure 3.2: Protocol fairness based on energy consumption

Figure 3.3 shows some very interesting findings. Regardless of the routing scheme employed, when the dominant idle time consumed energy is neglected, most of the drained energy is attributed to the defer periods. This includes all three types of defer: SIFS, random backoff, and DIFS. Moreover, if the idle time energy consumption were not to be neglected, a wireless ad hoc network under heavy traffic load would still encounter the same energy consumption behavior. This is due to the fact that the mobile nodes will spend most of their time deferring for successful medium access, or sending, or receiving frames. Even though defer-related energy consumption is only a special case of idle time energy consumption, some care can be exercised in the design of power-aware MAC protocols by making some nodes based on node density transit to their sleep states while deferring rather than wasting this significant amount of energy. This also implies that the random exponential backoff algorithm needs revisiting.

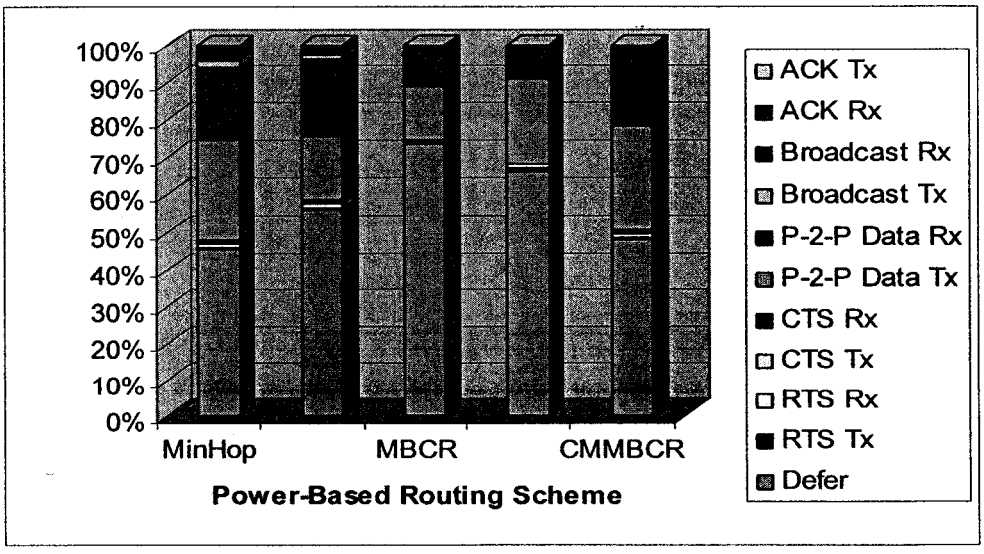


Figure 3.3: Energy consumption shares

RTS receipt energy and CTS receipt energy are considerably large for all schemes. For MBCR and MMBCR, RTS transmissions (actual transmissions and retransmissions) account for a large share of the consumed energy for control frames (see Figure 3.4). As explained earlier, MBCR and MMBCR can both result in MAC contentions in the vicinity of a node with large remaining battery capacity. This leads to several unsuccessful RTS retransmissions and more wasted energy due to these retransmissions. As shown in Figure 3.5, the defer-related energy for minimum-hop routing is minimal. The second best is that of CMMBCR which is approximately 1.05 that of minimum-hop routing.

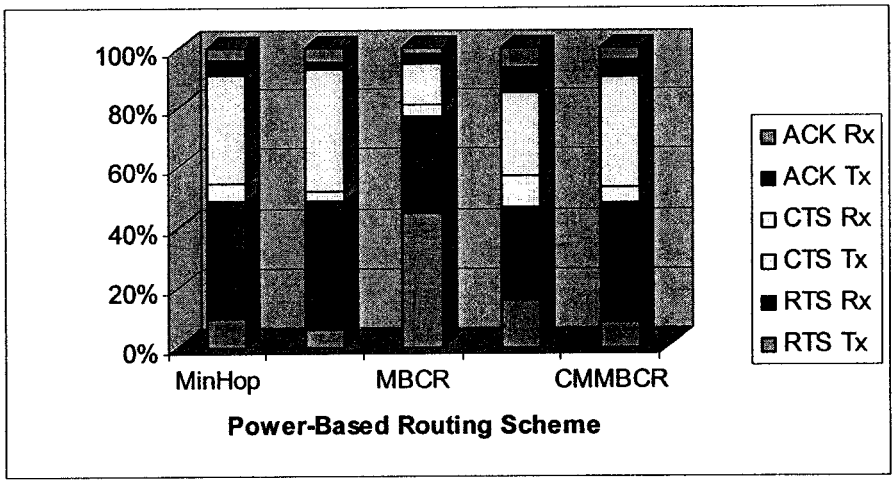


Figure 3.4: Energy consumption shares for MAC control frames

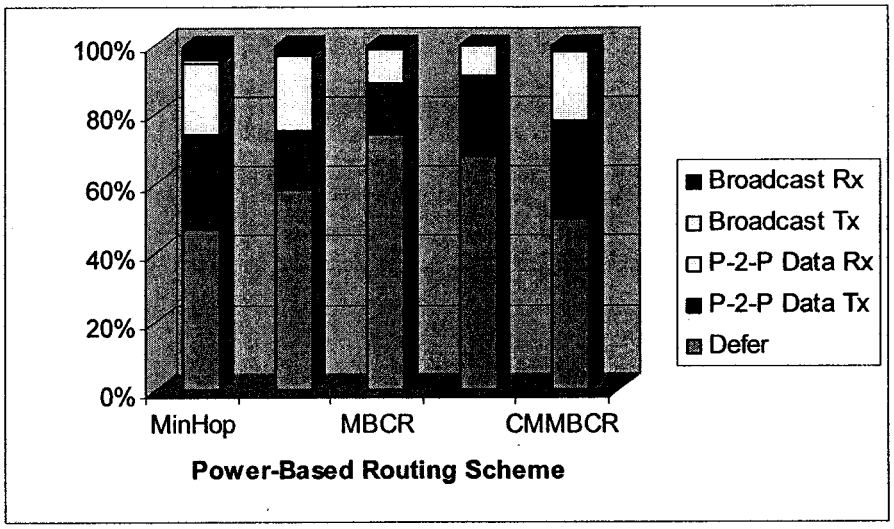


Figure 3.5: Energy consumption shares for defer, data transmissions and receptions

3.5 SUMMARY

In summary, we presented a detailed performance study of power-aware routing schemes in wireless ad hoc networks. MBCR attempts to fairly use the battery capacity of the wireless stations, but consumes more total energy than the other four schemes. MMBCR produces the least average consumed energy under low mobility. However, its performance degrades slightly in the presence of high mobility. Minimum-hop routing exhibits the best performance among the five schemes in terms of energy-based fairness. This is because MBCR, MMBCR, and CMMBCR all put great emphasis on the remaining battery capacity. Hence, a node may be overused only because its remaining battery capacity is larger than others. Interestingly, most of the drained energy is attributed to the defer periods and not actual point-to-point data transmissions.

The fact that minimum-hop routing, a non-power-aware scheme, performs better than the proposed power-aware schemes, in many cases, suggests that there is still more to energy-efficient routing protocol design than what has been unveiled so far. Our MAC-based investigations reveal that battery capacity alone cannot be used to satisfy our requirements for a power-efficient routing scheme. Power-based fairness and energy conservation can both be attained if valuable MAC and PHY input is passed to the network layer above so as to avoid high-contention zones and penetrate low-contention ones, and if nodes are allowed to disable their transceivers as long as routing fidelity is maintained. The latter is achieved via a dynamic wireless mobile infrastructure.

CHAPTER 4

ENERGY-EFFICIENT CROSS-LAYER DESIGNS

Contrary to present conjectures, in the previous chapter our MAC-based performance study revealed that battery capacity alone cannot be used as the sole means for achieving energy-based fairness and system longevity. Moreover, energy conservation may be attained only if valuable MAC (and PHY) input is passed to the network layer. Hence, in this chapter, we propose two schemes whose objective is to enhance the operation of existing power-based multi-path routing protocols via cross-layer designs and optimal load assignments. Our proposed *Energy-Constrained Path Selection (ECPS)* scheme is described in detail in the next section. In Section 4.2, another novel scheme, namely *Energy-Efficient Load Assignment (E2LA)*, is proposed, and its optimal algorithm follows in Section 4.3. Section 4.4 contains a variety of experiments conducted on E2LA and its variants. Finally, in Section 4.5 we summarize our findings and present the conclusions drawn from this chapter.

4.1 ENERGY-CONSTRAINED PATH SELECTION (ECPS)

Our goal herein is to maximize the probability of sending a packet to its destination in at most n transmissions. Assuming that a single transmission results in a fixed amount of dissipated energy, then for the total expended energy to be no more than γ it is required that $n \le \left\lfloor \frac{\gamma}{e} \right\rfloor$, where e is equal to the dissipated energy per hop.

Energy-Constrained Path Selection (ECPS) achieves the above objective by employing probabilistic dynamic programming (PDP) techniques and assigning a unit reward if the favorable event (in this case, reaching the destination in n, or less transmissions) occurs, and assigns no reward otherwise. Hence, it can be shown that maximizing the expected reward is equivalent to maximizing the probability that the packet reaches the destination in at most n transmissions.

Let $f_t(i)$ be the probability of success immediately before the t^{th} transmission in n or less transmissions given that the transmitting node, i, acts optimally. Thus the following three axioms apply:

$$f_t(i) = 0 : t > n, i \in N - \{D\} \quad (1)$$

$$f_t(D) = \begin{cases} 1 : t \le n+1 \\ 0 : t > n+1 \end{cases}, \text{ D is no more than t-1 hops away from S} \quad (2)$$

$$f_n(i) = 0 : i \notin n_D \quad (3)$$

In the above, N is the set of all wireless nodes in the network, and n_i is the set of all single-hop neighbors of node i, including i itself. The first axiom simply states that it is

infeasible to reach the destination in at most n transmissions if an intermediate node, i, receives the packet in more than n transmissions. From (2), the probability of delivering the packet to the destination in at most n transmissions from D itself prior to the t^{th} transmission is 1 provided that D is at most $t-1$ hops away from the source. The final axiom establishes the impracticality of satisfying our objective prior to the n^{th} transmission by node i unless i is a one-hop neighbor of D. Let $p_{j_R}^{\alpha}$ be the probability that the $(j+1)^{th}$ node on route R successfully receives the packet transmitted by the j^{th} node on R after α attempts when $1 \le \alpha \le MMR+1$, and the probability of failure in $MMR+1$ transmissions when $\alpha = MMR+2$. We now use the following PDP recursion, namely ECPS, to find the probability of success at node i right before the t^{th} transmission $f_t(i)$:

$$f_t(i) = \begin{cases} 1 : i = D \\ \max\limits_{j} \sum\limits_{k} p_i^k f_{t+k}(j) \end{cases} \tag{4}$$

In (4), j is the next hop towards the destination, D, and k is the number of transmissions until the packet is successfully received by j. Let $H(R,i)$ for route R and node $i \in R$ be equal to the number of hops between the source S and node i. Thus,

$$f_t(i) = 0 : H(R,i) \ge t \tag{5}$$

Apparently, a route R may be used as an input to the PDP if and only if $H(R,D)$ is smaller than n. Obviously, our objective is to calculate $f_1(S)$, and ties are broken by choosing the next hop which results in less total energy expenditure (i.e., fewer retransmissions). Figure 4.1 shows a small wireless ad hoc network to which the above

PDP may be applied. In this example, we are interested in maximizing the probability of reaching D in at most n=3 transmissions. There are 3 S-D routes. Table 4.1 lists the set of probabilities for reaching the next hop from every node i on any of the 3 S-D routes, excluding D. For example, S successfully sends a packet to any of its neighbors in 1 transmission with probability 0.25, in 2 transmissions with probability 0.35, and in 3 transmissions with probability 0.4. We use (4) above and work backward until we are able to determine $f_1(S)$.Once $f_1(S)$ is determined we are capable of finding the route that contributed to its value.

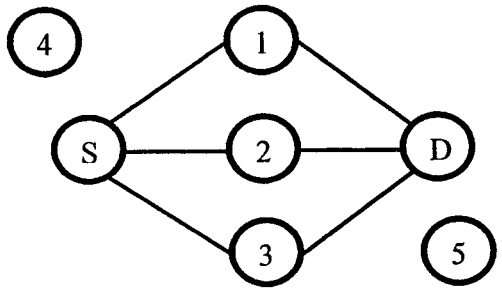


Figure 4.1: A sample wireless ad hoc network with 3 S-D routes

Table 4.1
Transmission probabilities for the nodes in Figure 4.1

Node	Number of attempts	Probability
S	1	0.25
	2	0.35
	3	0.4
1	1	0.5
	2	0.2
	3	0.3
2	1	0.3
	2	0.2
	3	0.5
3	1	0.7
	2	0.2
	3	0.1

The following sets of computations are performed:

Stage 3 (n) Computations:

$$f_3(D) = 1 \tag{6}$$

$$f_3(1) = p_1^1 f_4(D) + p_1^2 f_5(D) + p_1^3 f_6(D) = p_1^1 f_4(D) = 0.5 \tag{7}$$

$$f_3(2) = p_2^1 f_4(D) + p_2^2 f_5(D) + p_2^3 f_6(D) = p_2^1 f_4(D) = 0.3 \tag{8}$$

$$f_3(3) = p_3^1 f_4(D) + p_3^2 f_5(D) + p_3^3 f_6(D) = p_3^1 f_4(D) = 0.7 \tag{9}$$

Stage 2 (n-1) Computations:

$$f_2(1) = p_1^1 f_3(D) + p_1^2 f_4(D) + p_1^3 f_5(D) = p_1^1 f_3(D) + p_1^2 f_4(D) = 0.7 \tag{10}$$

$$f_2(2) = p_2^1 f_3(D) + p_2^2 f_4(D) + p_2^3 f_5(D) = p_2^1 f_3(D) + p_2^2 f_4(D) = 0.5 \tag{11}$$

$$f_2(3) = p_3^1 f_3(D) + p_3^2 f_4(D) + p_3^3 f_5(D) = p_3^1 f_3(D) + p_3^2 f_4(D) = 0.9 \tag{12}$$

Stage 1 (n-2) Computations:

$$f_1(S) = \max \begin{cases} p_S^1 f_2(1) + p_S^2 f_3(1) + p_S^3 f_4(1) = 0.35 \\ p_S^1 f_2(2) + p_S^2 f_3(2) + p_S^3 f_4(2) = 0.23 \\ p_S^1 f_2(3) + p_S^2 f_3(3) + p_S^3 f_4(3) = 0.47 \end{cases} \tag{13}$$

$$\therefore f_1(S) = 0.47$$

Equation (6) follows from (2), while (7)-(12) follow from axioms (1), (2). Finally, (13) follows from (3), (4). From (13), the optimal route (the one that maximizes the probability of success in at most n transmissions) is *S-3-D*. Note that we did not need to calculate $f_2(S)$ in stage 2. Although ECPS chooses a single route, from a set of tentative energy-efficient routes, that maximizes the probability of success and thus save the energy that would otherwise be wasted in case of failure, it does not take advantage of the multiple routes which are readily available to the ECPS scheme through any on-demand

or table-driven routing algorithm. In addition, utilizing multiple routes frequently results in overlapping transmissions which in turn reduce the system-wide consumed energy.

4.2 ENERGY- EFFICIENT LOAD ASSIGNMENT (E2LA)

Based on our findings from Chapter 3 and Section 4.1, energy conservation is attained through all the following:

1. Battery-based energy-aware routing
2. Utilization of multiple simultaneous routes
3. Load distribution for energy efficiency and contention resolution
4. Turning off the radios while achieving routing fidelity by means of a power-aware wireless mobile infrastructure.

In this section, we address the first three objectives, whereas the fourth is the topic of Chapter 6. Thus, now we turn to the problem of efficiently distributing the routing load among |Z| energy-aware routes. In our proposed scheme, packets are allotted to routes based on their willingness to save energy. At the MAC layer, each node may compute the probability of successfully transmitting packets in α attempts using time averages, where $\alpha \leq MMR+1$ and MMR is the maximum MAC retransmissions. Thus, if the source has a flow of P packets, and $r_t(P_t)$ is the expected reward of P_t packets assigned to route t, and $f_t(x)$ is the maximum expected reward when x packets are assigned to routes t, t+1,..., |Z|, then we need to solve the following recursion for t=1, 2, ..., |Z|-1:

$$f_t(x) = \max_{P_t}\{r_t(P_t) + f_{t+1}(x - P_t)\} \tag{14}$$

Clearly, P_t is the number of packets assigned to route t, and $f_{|Z|}(x)$ is by definition the maximum expected reward assigned to route $|Z|$, and $f_{|Z|}(x) = r_{|Z|}(x)$. This novel scheme is called *Energy-Efficient Load Assignment (E2LA)*.

We now unveil the reward schemes used herein, and introduce the concepts of *boolean*, *signed-unity*, and *contention-resolution-based* reward schemes, *contention indicator (CI)*, and *contention threshold (CT)*.

Definition 1 (Boolean Reward)

*A reward scheme is termed **boolean** if a unit reward is assigned if a packet is carried to the destination without encountering any errors along the way, and the reward is equal to zero otherwise.*

Definition 2 (Signed-Unity Reward)

*A reward scheme is termed **signed-unity** if a unit reward is assigned if a packet encounters no errors along its way to the destination, and the reward is equal to minus unity otherwise.*

Definition 3 (Contention Indicator)

*The **contention indicator** for a route R, CI_R, is a number reflecting the probability of transmission failure mainly due to contention for medium access, mobility, etc.*

Definition 4 (Basic Contention-Resolution-Based Reward)

A reward scheme is termed ***contention-resolution-based (CRB)*** *if it is signed-unity and a negative reward is assigned to route R as a function of the number of packets to deliver and* CI_R.

Definition 5 (CRB Reward with Contention Threshold)

A reward scheme is termed ***contention-resolution-based with contention threshold (CRB/CT)*** *if it is signed-unity and a negative reward is assigned to route R as a function of the number of packets that exceed a threshold,* γ_R, *and* CI_R.

In Figure 4.2, the destination D may be reached from the source S using 2 routes: *S-1-D* and *S-2-D*. Table 4.3 lists the set of probabilities for carrying Y consecutive packets to the destination without retransmissions. These probabilities may be computed using the set of probabilities for reaching the next hop in at most one attempt for every node $i \in$ the 2 S-D routes, which are shown in Table 4.2. Thus, for a route R consisting of $H(R, D_R)$ hops, the probability to route Y consecutive packets to the destination D_R all in exactly Y transmissions is equal to

$$P_R^Y = \left[\prod_{j=1}^{H(R,D_R)} p_{j_R}^1 \right]^Y \tag{15}$$

We may now compute the optimal allocation of packets to the routes in Figure 4.2. We assume that the source has P=3 packets it wishes to send to the destination. Depending on the reward scheme, we might need to perform the following computations:

Let *P(M, N, R)* be the probability of a total of *M* errors (also equal to the number of retransmissions) by the *N* nodes of route *R*. This probability may be computed as follows:

$$\sum_{k_1=0}^{\min(MMR+1,M)} \sum_{k_2=0}^{\min(MMR+1,M-k_1)} \cdots \sum_{k_{H(R,D_R)}=0}^{\min(MMR+1,M-\sum_{l=1}^{H(R,D_R)-1} k_l)} p_{1_R}^{1+k_1} p_{2_R}^{1+k_2} \cdots p_{H(R,D_R)_R}^{1+k_{H(R,D_R)}},$$

$$\text{where } 0 \le M \le (MMR+1).H(R,D_R) \text{ and } \sum_{l=1}^{H(R,D_R)} k_l = M \tag{16}$$

Although the above sum may sometimes be useful, depending on the employed reward scheme, it is the individual terms that mainly matter in the calculation of the expected reward and not necessarily the actual value of the summation. Hence, it is important to find the number of terms in the summation. We do so using the following recursion:

If M is the sum of the retransmissions carried out by all the N nodes of route R, MMR is the maximum allowable MAC retransmission, and *S(M, N, MMR)* is the number of terms in the above summation, then:

$$\mathrm{S}(0, \mathrm{n}, \mathrm{MMR}) = 1 \tag{17}$$

$$\mathrm{S}(1, \mathrm{n}, \mathrm{MMR}) = \mathrm{n} \tag{18}$$

$$\mathrm{S}(\mathrm{m}, 1, \mathrm{MMR}) = \begin{cases} 1 : m \le MMR+1 \\ 0 : m > MMR+1 \end{cases} \tag{19}$$

$$\text{and } S(M,N,MMR) = \sum_{l=0}^{M} S(l, N-1, MMR).S(M-l, 1, MMR) \tag{20}$$

Fortunately, however, boolean reward is assumed in the example of Figure 4.2. Thus, in calculating the expected reward $r_t(P_t)$ (using Table 4.2) only the terms that contain the probability of carrying a packet/packets to the destination without errors (P_R^1, P_R^2, P_R^3-we only need these 3 because P=3) are considered and all the rest are equal to 0 because the reward in this case will be equal to 0. We may also consider negative reward schemes

based on the amount of consumed energy and the MAC contention experienced in the route's vicinity. We first compute $f_2(0), f_2(1),..., f_2(3)$, and since $|Z|=2$ then $f_2(x) = r_2(x)$.

Table 4.2
Transmission probabilities for the nodes in Figure 4.2

Node	α	Probability
S	1	0.5
	2	0.2
	3	0.3
1	1	0.5
	2	0.3
	3	0.2
2	1	0.5
	2	0.4
	3	0.1

It can be shown that the expected reward of assigning g packets to route R, $r_R(g)$, is equal to:

$$\binom{g}{g} P_R^g . Q_R^{\,g-g} . g + \binom{g}{g-1} P_R^{g-1} . Q_R^{\,g-(g-1)} . (g-1) + ... + \binom{g}{g-(g-1)} P_R^{g-(g-1)} . Q_R^{\,g-[g-(g-1)]} . 1 + \binom{g}{g-g} P_R^0 . Q_R^{\,g-[g-g]} . 0$$

$$= \binom{g}{g} P_R^g . Q_R^{\,0} . g + \binom{g}{g-1} P_R^{g-1} . Q_R^{\,1} . (g-1) + ... + \binom{g}{1} P_R^1 . Q_R^{\,g-1} . 1 + \binom{g}{0} P_R^0 . Q_R^{\,g} . 0 \quad (21)$$

$$= \sum_{i=0}^{g} \binom{g}{i} P_R^i . Q_R^{\,g-i} . i \quad (22)$$

where $Q_R = 1 - P_R = \sum_{l_1=1}^{MMR+2} \sum_{l_2=1}^{MMR+2} ... \sum_{l_{H(R,D_R)}=1}^{MMR+2} P_{1_R}^{l_1} P_{2_R}^{l_2} ... P_{H(R,D_R)}^{l_{H(R,D_R)}}$, such that for route R:

$$l_i = 1 \Rightarrow \exists l_j \neq 1, j \neq i$$
$$\textit{and} \quad (23)$$
$$l_i = MMR + 2 \Rightarrow P_{j_R}^{l_j} = 1, \forall j > i$$

Table 4.3
Success probabilities for consecutive packets calculated using Table 4.2

Route	Number of Consecutive Packets Routed Error/Collision-Free	Probability
R_1	1	$P_{R_1}^1 = 0.25$
	2	$P_{R_1}^2 = 0.0625$
	3	$P_{R_1}^3 = 0.015625$
R_2	1	$P_{R_2}^1 = 0.25$
	2	$P_{R_2}^2 = 0.0625$
	3	$P_{R_2}^3 = 0.015625$

In equation (22), i is the number of packets, out of g, that are delivered error-free.

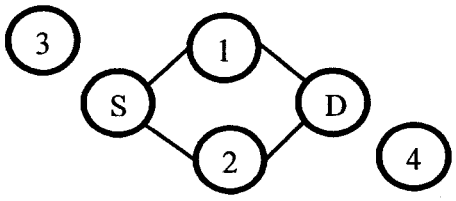


Figure 4.2: A sample wireless ad hoc network with 2 S-D routes

4.3 E2LA: THE ALGORITHM

For G packets and $|Z|$ routes, there exist $|Z|^G$ possible ways of assigning packets to routes. This is a "brute force" way for maximizing the expected reward and is infeasible for large values of G. For simplicity, let $z = |Z|$. Thus, determining the maximum expected reward by enumerating all possible ways and calculating the expected reward for each one requires $\Omega(z^G)$ time. Instead, we propose the following recursive algorithm:

```
Recursive E2LA(t,i)
{
  /**
    * Maximum expected reward
    * calculated by initially calling
    * Recursive E2LA(1,G)
    */

  if (t == z)
      f_z(i) = r_z(i)
  else
      f_t(i) = max_{0≤j≤i} {r_t(j) + Recursive E2LA(t+1, i-j)}
}
```

The total number of basic operations (calculations involving $f_a(b)$, where $1 \le a \le z$ and $1 \le b \le G$) is (in the following analysis $z = 4$):

$$
\begin{aligned}
T_{\text{Rec}}(1,G) &= \sum_{i=0}^{G} T_{\text{Rec}}(2,i) \\
&= \sum_{i=0}^{G}\sum_{j=0}^{i} T_{\text{Rec}}(3,j) \\
&= \sum_{i=0}^{G}\sum_{j=0}^{i}\sum_{k=0}^{j} T_{\text{Rec}}(4,k) \\
&= \sum_{i=0}^{G}\sum_{j=0}^{i}\sum_{k=0}^{j} 1 \\
&= \sum_{i=0}^{G}\sum_{j=0}^{i} (j+1) \\
&= \sum_{i=0}^{G} \frac{i(i+1)}{2} + \sum_{i=0}^{G} (i+1) \\
&= \sum_{i=0}^{G} \frac{i(i+1)}{2} + \Theta(G^2) \\
&= \frac{1}{2}\left(\sum_{i=0}^{G} i^2 + \sum_{i=0}^{G} i\right) + \Theta(G^2) \\
&= \frac{1}{2}\left(\frac{G(G+1)(2G+1)}{6} + \frac{G(G+1)}{2}\right) + \Theta(G^2) \\
&= \Theta(G^3)
\end{aligned}
\tag{24}
$$

For $z = 5$,

$$T_{Rec}(1,G) = \sum_{h=0}^{G} \Theta(h^3)$$
$$= \Theta\left(\sum_{h=0}^{G} h^3\right)$$
$$= \Theta(G^4)$$ [1] (25)

The time complexity of the recursive solution depends on the value of z. The larger the value of z the slower the scheme. Thus, the recursive algorithm is computationally expensive. Hence, we propose the following enhanced (non-recursive) algorithm the performance of which is asymptotically independent of z.

```
E2LA(G, z)
{
   for(i = 0; i ≤ G; i++)
        f_z(i) = r_z(i)

   for(h = z − 1; h ≥ 2; h−−)
        for(i = 0; i ≤ G; i++)
             f_h(i) = max_{0≤j≤i} {r_h(j) + f_{h+1}(i − j)}

   f_1(G) = max_{0≤j≤G} {r_1(j) + f_2(G − j)}
}
```

The total number of basic operations (calculations involving $f_a(b)$, where $1 \le a \le z$ and $1 \le b \le G$) is:

$$T_{E2LA}(G,z) = (G+1) + (z-2).\left(\sum_{i=0}^{G}(i+1)\right) + (G+1)$$
$$= 2.(G+1) + (z-2).\left(\sum_{i=0}^{G} i + \sum_{i=0}^{G} 1\right)$$

---

[1] $\sum_{h=0}^{G} h^3 = \frac{G^2(G+1)^2}{4} = \Theta(G^4)$

$$= 2.(G+1) + (z-2).\left(\frac{G(G+1)}{2} + (G+1)\right)$$
$$= 2.(G+1) + (z-2).\left(\frac{(G+2)(G+1)}{2}\right)$$
$$= \Theta(zG^2) = \Theta(G^2) \quad (26)$$

From (26), this new quadratic performance is an enhancement over (25), and is asymptotically independent of z. Clearly, for larger values of G our enhanced E2LA algorithm computes the optimal allocation of packets much faster than the recursive E2LA.

4.4 EXPERIMENTS AND ILLUSTRATIONS

We study the characteristics of E2LA via a number of experiments. We also demonstrate the effect of several parameters on the optimum reward and the carried load by each of the $|Z|$ routes. Moreover, we prove that energy savings are achieved due to using E2LA. We use the sample network in Figure 4.2 as a basis for our experiments.

In the first couple of experiments, the results of which are shown in Figures 4.3 and 4.4, signed-unity reward is used, and we vary $p_{i_{R_x}}^{y}$ for $x \in \{1,2\}, i \in R_x - \{D\}, y \in \{1,2,3\}$, and compute the carried routing load for route R_2, L_{R_2}, and the optimum expected reward achieved. The routing load carried by R_1, L_{R_1}, is equal to the difference between the total number of packets and L_{R_2}. Table 4.4 shows the transmission probabilities used in the first experiment for all the nodes on routes R_1 and R_2. The transmission probabilities may vary for different next hop nodes. This is in recognition of the fact that medium access

conditions vary across the network and are dependent on the neighbors of both the sender and the receiver. As shown in Table 4.4, $p^1_{S_{R_2}}$ is equal to 0. Thus, R_1 is always favored regardless of the number of packets. On the other hand, in Experiment II, a different set of success probabilities is used for R_2, whereas they are unchanged for R_1. Both $p^1_{S_{R_2}}$ and $p^1_{2_{R_2}}$ are set to 1.0 (i.e., any number of packets will be delivered to the destination error-free). This results in the total number of packets being carried solely by R_2.

Table 4.4
Transmission probabilities used in Experiment I

Node	Next Hop	α	Probability
S	1	1	0.45
		2	0.5
		3	0.05
S	2	1	$\approx$0.0
		2	0.0
		3	1.0
1	D	1	0.65
		2	0.15
		3	0.2
2	D	1	1.0
		2	0.0
		3	0.0

Figure 4.4 shows that the maximum expected reward decreases linearly when the number of packets increases in the first experiment. It can be shown that this is due to the fact that the reward per packet is a negative constant. The maximum expected reward for a single packet is equal to -0.415, and -8.3 in the case of 20 packets. In the second experiment, increasing the number of packets results in increasing the maximum expected reward. Moreover, the slope of the curve at any point is equal to the maximum expected reward per packet. We observe that in all cases so far the aggregate flow is solely carried by a

single route. That is to say, 100% of the traffic is carried by one route and all other routes remain idle. This remains true even if more than two routes are being used. In the following experiments we show how to eliminate this deficiency by using a more realistic reward assignment scheme. We first examine the two most basic reward schemes. When boolean reward is employed, the maximum expected reward per packet is simply:

$$\begin{aligned}&\max_x\left(\left(\prod_{j=1}^{H(x,D_x)} p_{j_x}^1\right).1+Q_x.0\right)\\&=\max_x P_x^1\end{aligned} \tag{27}$$

In the case of signed-unity reward, the maximum expected reward per packet becomes:

$$\begin{aligned}&\max_x\left(\left(\prod_{j=1}^{H(x,D_x)} p_{j_x}^1\right).1+Q_x.-1\right)\\&=\max_x\left(P_x^1-Q_x\right)\end{aligned} \tag{28}$$

Both results may be verified using (22). For the first scheme,

$$\begin{aligned}r_R(g)&=\sum_{i=0}^{g}\binom{g}{i}P_R^i.Q_R{}^{g-i}.i\\&=\sum_{i=1}^{g}\frac{ig!}{(g-i)!i!}P_R^i.Q_R{}^{g-i}\\&=gP_R\sum_{i=1}^{g}\frac{(g-1)!}{(g-i)!(i-1)!}P_R^{i-1}.Q_R{}^{g-i}\\&=gP_R\sum_{k=0}^{g-1}\frac{(g-1)!}{(g-1-k)!k!}P_R^k.Q_R{}^{g-1-k}\\&=gP_R\sum_{k=0}^{g-1}\binom{g-1}{k}P_R^k.Q_R{}^{g-1-k}\\&=gP_R(P_R+Q_R)^{g-1}\\&=gP_R\\\therefore\max_x r_x(1)&=\max_x P_x^1\end{aligned} \tag{29}$$

which is the same as the result in (27).

Now, for the second scheme,

$$r_R(1) = \sum_{i=0}^{1} \binom{1}{i} P_R^i . Q_R^{1-i} . (i - (1-i))$$
$$= \sum_{i=0}^{1} \binom{1}{i} P_R^i . Q_R^{1-i} . (2i - 1)$$
$$\therefore \max_x r_x(1) = \max_x \left(2.P_x^1 - 1\right)$$
$$= \max_x \left(P_x^1 - Q_x\right) \qquad (30)$$

where (30) follows from the binomial theorem and (29), and is the same as the result in (28). Table 4.5 shows all four reward schemes and the corresponding expected reward for assigning g packets to route R.

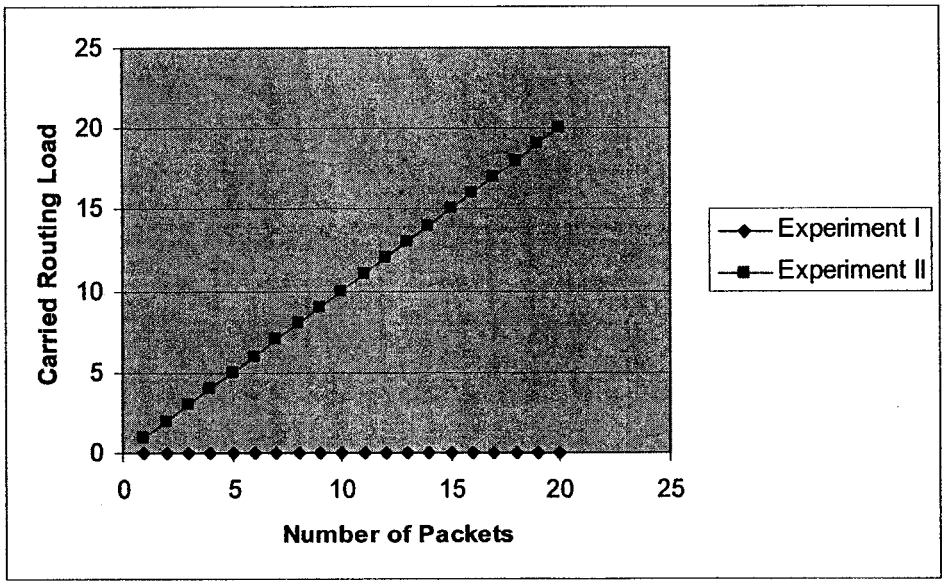


Figure 4.3: E2LA with signed-unity reward and the optimum load carried by R_2

Table 4.5
E2LA reward schemes

E2LA Reward Scheme	Expected Reward for g packets
Boolean	$g \cdot P_R^1$
SU	$g \cdot \left(P_R^1 - Q_R\right)$
BCRB	$\sum_{i=0}^{g} \binom{g}{i} P_R^i . Q_R^{\,g-i} . (2i - g - CI_R^i)$
CRB/CT	$\sum_{i=0}^{g} \binom{g}{i} P_R^i . Q_R^{\,g-i} . (2i - g - CI_R^{(i-\gamma_R)})$

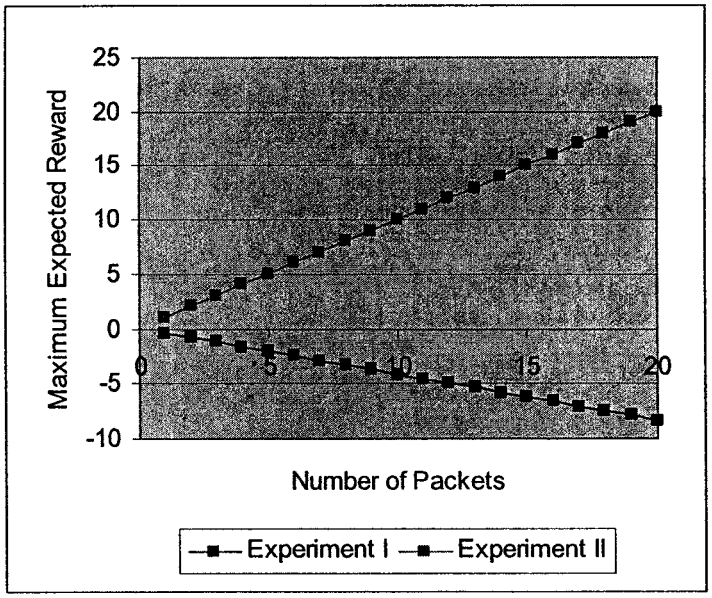


Figure 4.4: E2LA with signed-unity reward and the optimum expected reward

Figures 4.5 and 4.6 show the results of two sets of experiments where CRB reward is used and CI_{R_1} and CI_{R_2} are set to 1.235 and 1.2, respectively. The transmission probabilities for all the nodes on routes R_1 and R_2 are shown in Table 4.6, and are the same for both sets of experiments. In Experiment III, *CI* is taken into account only if the number of packets delivered error-free exceeds 1, whereas it is used for any number of

error-free packets in Experiment IV. As shown in Figure 4.5, R_2 is used exclusively in Experiment IV until the number of packets exceeds 11. The load is then shifted completely to R_1. This should not have been at all possible. However, the expected reward of assigning the whole flow to a route that has been previously inactive may be greater than that of distributing the load amongst a set of routes of which one or more have been active in the case of a smaller aggregate flow. In Experiment III, the number of packets assigned to R_2 increases from 14% of the aggregate flow to approximately 37% when the aggregate flow becomes 19 packets. Figure 4.6 shows that the optimum expected reward remains almost the same for both experiments.

Table 4.6
Transmission probabilities used in Figures 4.5 and 4.6

Node	Next Hop	α	Probability
S	1	1	0.1
		2	0.0
		3	0.9
S	2	1	0.12
		2	0.0
		3	0.88
1	D	1	0.1
		2	0.0
		3	0.9
2	D	1	0.08
		2	0.2
		3	0.72

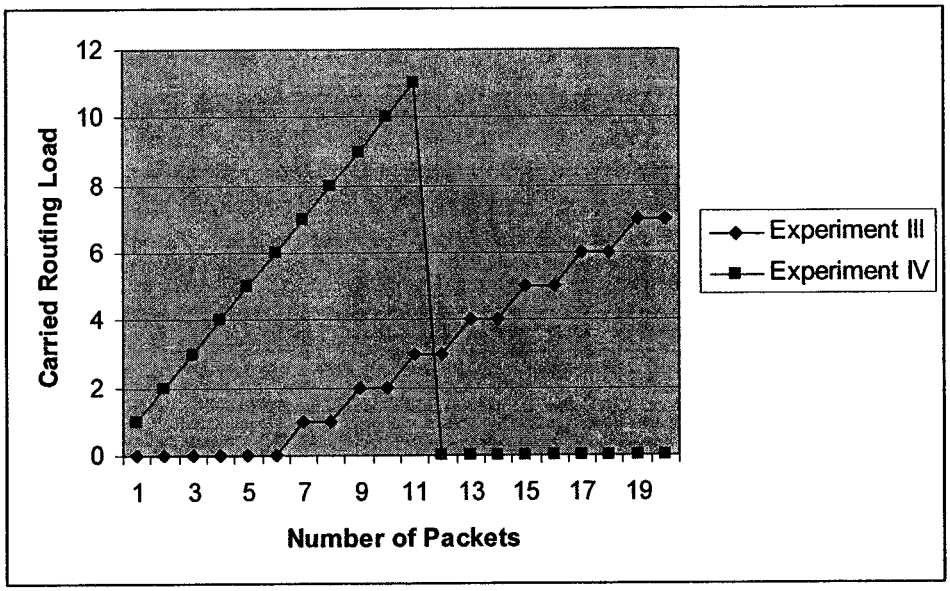


Figure 4.5: E2LA with CRB reward and the optimum carried load by R_2

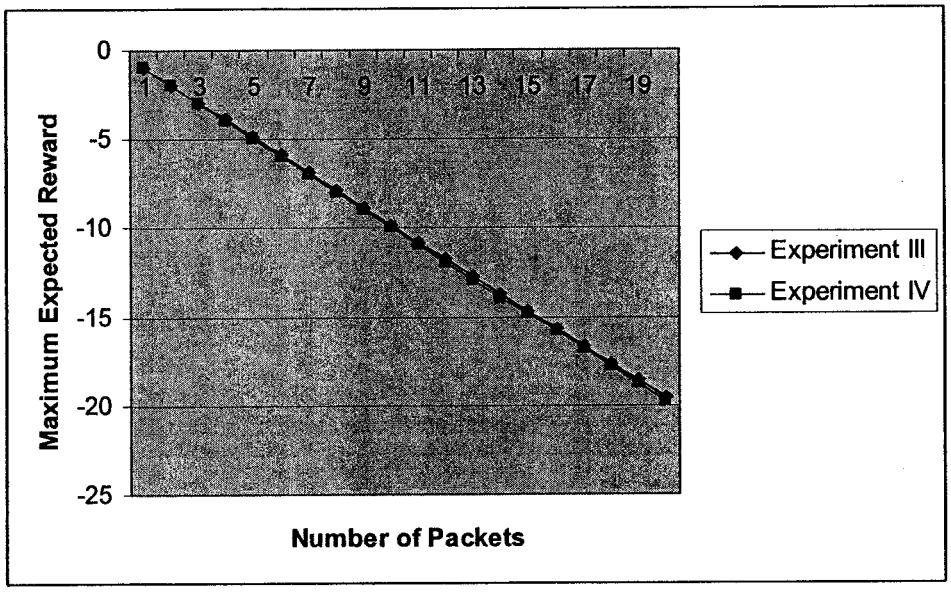


Figure 4.6: E2LA with CRB reward and the optimum expected reward

In Figure 4.7, we study the effect of CRB/CT reward on the distribution of the routing load. As shown in Figure 4.7, the optimum load is always a non-decreasing function of the aggregate flow. This is not always true with the previous reward schemes. Unlike

Experiment III, the load may never be shifted completely between routes. The same set (which is shown in Table 4.7) of success probabilities is used for both experiments. In Experiment V, CI_{R_1} and CI_{R_2} are set to 1.2 and 1.4, respectively, and to 1.4 and 1.2, respectively, in Experiment VI. In both experiments, CT is equal to 2. In Experiment V, L_{R_2} remains constant for $G \geq 13$ because the slight edge that R_2 has over R_1 in terms of the success probabilities may no longer offset the effect of its larger CI on the expected reward when the load exceeds 6 packets. On the other hand, for $CI_{R_2} \leq CI_{R_1}$ and the same set of success probabilities, or a better one, L_{R_2} is lower bounded by L_{R_1}.

Table 4.7
Transmission probabilities used in Figure 4.7

Node	Next Hop	α	Probability
S	1	1	0.5
		2	0.3
		3	0.2
S	2	1	0.5001
		2	0.3
		3	0.1999
1	D	1	0.5
		2	0.3
		3	0.2
2	D	1	0.5001
		2	0.3
		3	0.1999

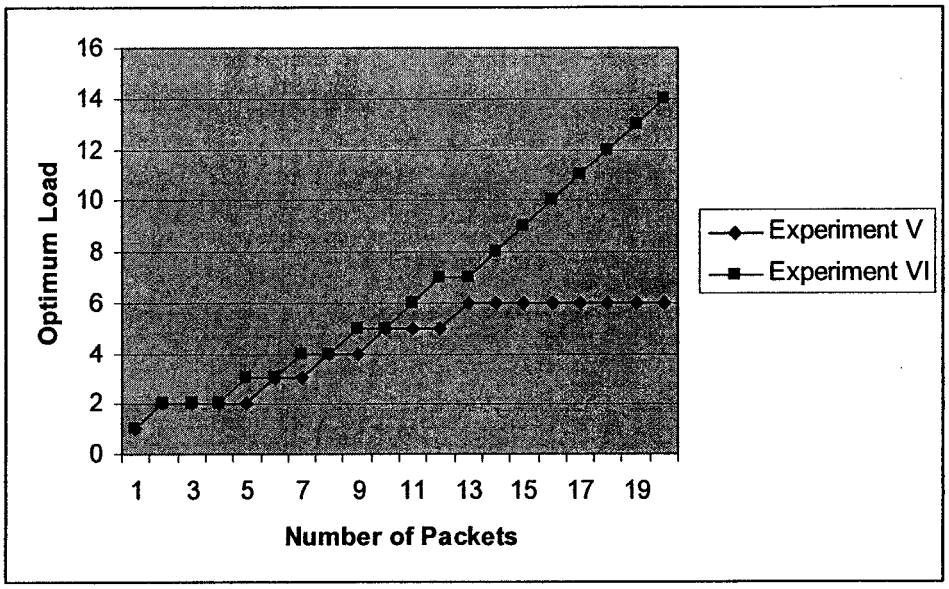


Figure 4.7: E2LA with CRB/CT reward and the optimum carried load by R_2

Figure 4.8 shows the expected energy savings for both routes using the same set of success probabilities as in the experiments of Figure 4.7. In our calculations, a worst-case energy consumption scenario is assumed. In this scenario, a node is allowed to be in one of the "transceiver-enabled" states. Thus, both the source and its neighbors are prone to energy consumption. The dissipated power due to a single transmission is equal to $1.33W$ and receiving and processing a frame by a neighbor requires $1W$. All frames are $1K$ bytes long, and the data rate is $22Mbps$. It is assumed that every node has an average of 20 neighbors. CI_{R_1} and CI_{R_2} are set to 1.2 and 1.4, respectively. The expected energy savings for R_1 vary between approximately 10 and 60 mJ, when the aggregate flow is 1 and 20 packets, respectively. The expected energy savings for R_2 are larger than those of R_1, as the carried optimum load is always less. The expected energy savings for R_2 are 0 when the whole flow is carried solely by R_2, but grows until it reaches 140 mJ when the total flow is equal to 20 packets/sec. It is noteworthy that whenever the slope of the

corresponding piece-wise linear curve in Figure 4.7 is equal to 0, the slope of the energy savings curve in Figure 4.8 is equal to the expected energy saved due to a single packet. For example, between 14 and 19 packets, inclusive, the slope in Figure 4.7 is equal to 0, whereas that of Figure 4.8 is approximately equal to 10.

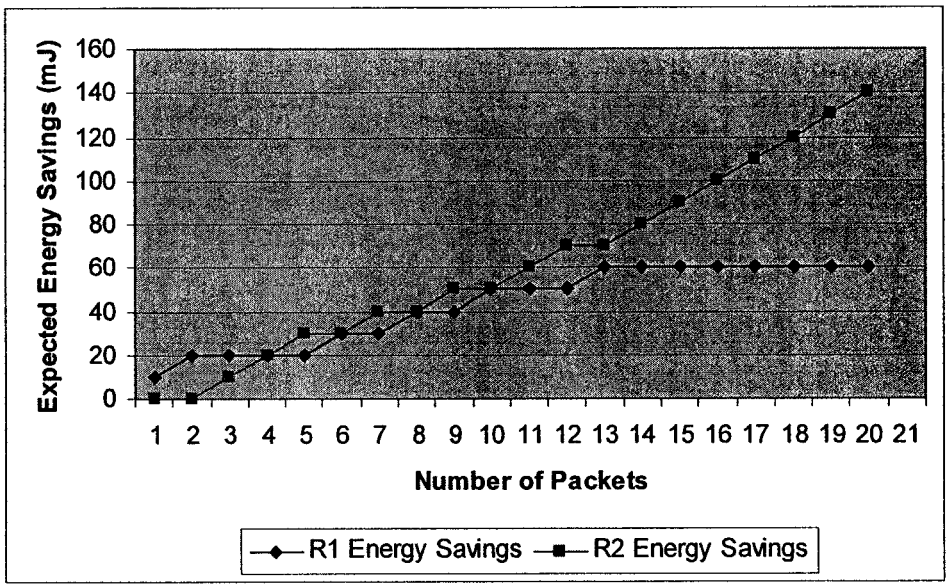


Figure 4.8: Expected Energy Savings for *R1*, *R2*, $CI_{R_1} = 1.2$ and $CI_{R_2} = 1.4$

In Figure 4.9, CI_{R_1} and CI_{R_2} are set to 1.4 and 1.2, respectively, with all the other parameters unaltered. The R_1 and the R_2 expected energy savings curves in Figure 4.9 resemble the R_2 and the R_1 curves, respectively, in Figure 4.8, except for the first three routing loads. In case of R_2 in Figure 4.9, this is because CI_{R_2} is now smaller than CI_{R_1} and packets that were once inadmissible may now be carried by R_2. More importantly, R_2's expected energy savings curve always forms a lower bound for R_1's expected energy savings curve (as can be seen in Figure 4.9). This is attributed to R_2's

success probabilities being greater and its contention indicator being smaller than those of R_1.

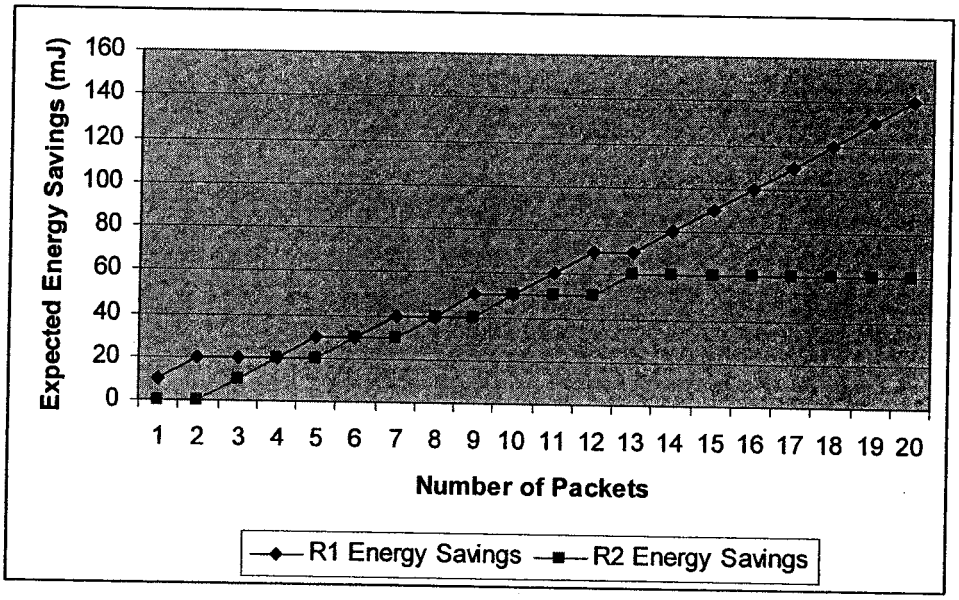


Figure 4.9: Expected Energy Savings for *R1*, *R2*, $CI_{R_1} = 1.4$ and $CI_{R_2} = 1.2$

In Figure 4.10, we investigate the expected energy savings for R_1 with variable CI_{R_1} to CI_{R_2} ratios. Ratios of 1:1, 3:1, and 9:1 are used. All the other parameters, including the frame size and transceiver speed, remain unaltered. It is worthy of notice that the greater the ratio, the more the energy savings. This adds to the advantages of E2LA since MAC contention is avoided through distributing the load in such a way that the chances of success are maximized. For $1 \le G \le 5$, the average energy savings are the same for the three distinct ratios because the number of packets allotted to R_1 is also the same. However, this changes for $G \ge 6$. For $G = 20$, for example, 70% more energy is saved, on average, when the CI_{R_1} to CI_{R_2} ratio is equal to 9:1, as compared to the 1:1 ratio.

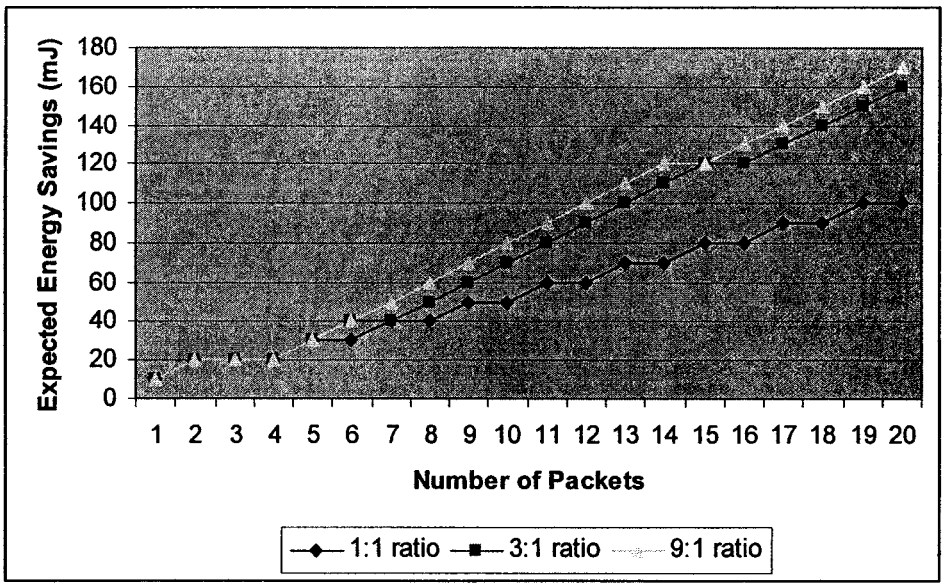


Figure 4.10: The impact of the *CI* ratio on the expected energy savings for *R1*

4.5 SUMMARY

In this chapter, we proposed ECPS, a novel energy-efficient scheme for wireless ad hoc and sensor networks. ECPS utilizes cross-layer interactions between the network layer and MAC sublayer to achieve energy conservation. Any set of energy-aware routes may be used as an input to ECPS. The MAC sublayer provides the network layer with information pertaining to successfully receiving a CTS or an ACK frame, or failure to receive one. ECPS, in turn, chooses the route that minimizes the probability of error or, equivalently, maximize the expected reward. On the downside, ECPS has no means of using multiple routes simultaneously.

On the other hand, we proposed another energy-efficient scheme, E2LA, which uses readily available multiple routes to its advantage, and alleviates the shortcomings of ECPS. We provided an in-depth explanation of the operation of our proposed E2LA algorithm. Similar to ECPS, E2LA also employs probabilistic dynamic programming techniques and utilizes cross-layer interactions between the network and MAC layers. Besides, we developed four distinct reward schemes for which E2LA assigns routing loads accordingly. In addition, illustrative examples of E2LA were presented, and its diverse properties were introduced and validated.

E2LA evidently relies on the accuracy of the MAC-originated statistics and the ability to maintain them dynamically and timely. Moreover, despite the fact that energy conservation is achieved through E2LA, there is no provision of a deterministically guaranteed minimum system lifetime. In summary, the significance of the following has been apprehended:

1. Load distribution among multiple simultaneous routes
2. Establishing a guaranteed minimum network lifetime
3. Power-aware infrastructure that is capable of running quasi-centralized MAC schemes and routing.

CHAPTER 5

QUASI-GUARANTEED SYSTEM LIFETIME AND CONTENTION MITIGATION

In this chapter, we propose a novel deterministic framework, namely *Quasi-Guaranteed System Lifetime (Q-GSL)*, that allows the admission of flows without jeopardizing the limited energy of the wireless stations. A noteworthy feature of our framework is that the upper bound on the packet rate is computed in the absence of the energy overheads associated with routing, contention resolution, channel sensing, etc. Hence, the upper bounds derived herein would always constitute valid upper bounds on the amount of data transmitted by a source node per unit time. We derive upper bounds for networks with disjoint routes and for non-mutually exclusive paths as well. Likewise, we study the main characteristics of the proposed techniques. Our experiments reveal that load balancing is achieved amongst the routes and the nodes in the wireless ad hoc network without violating any of the energy constraints, and while adhering to a pre-computed minimum

system lifetime. Moreover, Q-GSL alleviates congestion by using multiple routes and through contention mitigation.

5.1 INTRODUCTION

None of the techniques and protocols developed thus far provides solid guarantees on the minimum system lifetime in the presence of congestion. In this chapter, our main objective is not to propose a new energy-aware protocol that leads to greater system lifetime. Our goal is rather to propose a new framework that would allow the transfer of data in a wireless ad hoc network from the source to the sink while maintaining a minimum network lifetime. We derive bounds on the aggregate packet flow in the presence of system-wide and nodal energy constraints.

This chapter is organized as follows. The next section discusses prefatory material of relevance to our optimal flow framework. In Section 5.3, the proposed optimal Q-GSL framework is described in detail. This is followed in Section 5.4 by a variety of experiments conducted on the new technique. Finally, in Section 5.5 we summarize our findings and the conclusions drawn from the chapter.

5.2 PRELIMINARIES

Consider a directed graph $G(N,E)$ where N is the set of all wireless nodes, and E is the set of all directed links $<i,j>$ where $i,j \in N$. L_R is the set containing all the directed links that constitute route R. In the most basic energy consumption models [52] and [54],

the system-wide consumed energy due to the transfer of a single packet of fixed size from node i to node j is $\sum_{g \in n_i} f_{ij}(g)$ where n_i is the set containing all neighbors of node i including itself, and where

$$f_{ij}(g) = \begin{cases} f(Pw_{ij}^{tx}) : g = i, Pw_{ij}^{tx} > 0 \\ f(Pw_{ij}^{rx}) : g = j, Pw_{ij}^{rx} > 0 \\ 0 : g \neq i, g \neq j \end{cases} \quad (1)$$

In the above, $f(Pw_{ij}^{tx})$ is the energy per packet consumed by the sender, i, when the required transmission power to reach j is Pw_{ij}^{tx}, and $f(Pw_{ij}^{rx})$ is the energy per packet consumed by the receiver, j, when the received power is Pw_{ij}^{rx}. The energy of the rest of the network remains unaffected. Hence, the total dissipated energy attributed to $P_{i,j}$ packets from i to j is equal to $P_{i,j}.(f(Pw_{ij}^{tx}) + f(Pw_{ij}^{rx}))$. Every node has a limited residual battery capacity. We denote by BC_i the residual battery capacity for node i. Initially, node i has $BC_i(0)$ of residual battery capacity, and its instantaneous battery capacity after t time units is given by

$$BC_i(t) = BC_i(0) - \int_0^t P_i(\tau).d\tau \quad (2)$$

$P_i(\tau)$ is the power dissipated in i at time τ. Let T_{SD} be equal to the duration of the flow from S to D.

We now introduce the concepts of constrained link-based, zone-based, and node-based energy dissipation in wireless ad hoc networks. Moreover, we unveil our interpretation of the system lifetime, which we also adopt in this chapter.

Definition 1 (Viable Link)

A directed link l is termed ***viable*** *for a number of packets P if the total amount of energy expended in the wireless network due to transmitting the P packets per unit time on l does not exceed a predefined link threshold* (γ_{LINK}).

Definition 2 (Viable Node)

A node i is termed ***viable*** *for a number of packets P if the total amount of energy expended in the wireless network due to transmitting the P packets per unit time by i does not exceed a predefined zone threshold* (γ_{ZONE}).

Definition 3 (Feasible Node)

A node i is termed ***feasible*** *for a nonempty set of directed links L and a specific assignment of packets per unit time for each link* $l \in L$ *if the total amount of dissipated energy by i during* T_{SD} *does not exceed* BC_i.

Definition 4 (Expected MAC Contenders-EMC)

The EMC for a wireless node i is the sum of the expected number of application layer packets, λ_i, *and PHY frames,* μ_i *per unit time.*

Definition 5 (System Lifetime)

This is the time until the first wireless node runs out of battery power.

5.3 THE Q-GSL FRAMEWORK

In this chapter, we want to address and optimally solve the following problem:

Given a source (S), a destination (D), a set M (where $|M| \neq 0$) of directed power-aware paths connecting S to D, the initial energy/battery capacity (BC) in each node $i \in R_m, \forall m \in \{1,...,|M|\}$, a lower bound on the system lifetime, the modulation scheme, the required bit error rate (BER), and the receiver sensitivity $(Pw_{\min}^{rx})$, what is the upper bound on the total number of packets per unit time that S can pump into the network?

Thus, we are in fact proposing a new framework that would allow the transfer of data in a wireless ad hoc network from the source to the sink without jeopardizing the limited energy of the wireless terminals. This framework would also help us better understand the intricate problem of energy conservation in wireless networks, and thus solve it for disjoint routes and for non-mutually exclusive paths as well. A noteworthy feature of our framework is that we compute the upper bound on the packet rate in the absence of the energy overheads associated with periodic broadcasts, route discovery, route maintenance, medium access control, physical channel sensing, etc. Hence, the maximum packet rate found in this chapter would always constitute a valid upper bound on the amount of data transmitted by a source node per unit time.

A. Mutually Exclusive (Disjoint) Routing

One can observe that the problem of maximizing the packet flow from S to D given a set M of disjoint paths (i.e., $\forall R_1, R_2, ..., R_{|M|} \in M$, $R_i \cap R_j = \{S, D\}$ for $i \neq j$) and the constrained energy dissipation per unit time is equivalent to a maximum flow integer programming (IP) problem given by Table 5.1.

Table 5.1
Integer program for determining optimal packet flow

Maximize $z = \sum_{n \in n_s - \{s\}} P_{s,n}$ (3)

s.t.

Conservation of flow constraint:

$$P_{h,i} = P_{i,j}, \forall < h,i >, < i,j > \in L_R, \forall R \in M \quad (4)$$

Non-negativity constraint:

$$P_{i,j} \geq 0, \forall < i,j > \in L_R, \forall R \in M \quad (5)$$

Energy constraints:

$$T_{SD} \sum_{s' \in n_s - \{s\}} (P_{s,s'} \cdot f(Pw_{ss'}^{tx})) \leq BC_s \quad (6)$$

$$T_{SD}(P_{i,j}\ f(Pw_{ij}^{tx}) + P_{h,i} f(Pw_{hi}^{rx})) \leq BC_i, \forall < h,i >, < i,j > \in L_R, \forall R \in M \quad (7)$$

$$P_{i,j}(f(Pw_{ij}^{tx}) + f(Pw_{ij}^{rx})) \leq \gamma_{LINK}, \forall < i,j > \in L_R, \forall R \in M \quad (8)$$

In Table 5.1, $\sum_{n \in n_s - \{s\}} P_{s,n}$ is the total number of packets transmitted by the source per unit time. The first set of constraints (4) constitutes the conservation of flow constraints. Since

only disjoint paths are allowed, the number of packets on any link on route R must equal the number of packets carried by the following link. The second set of constraints is obvious. It makes no physical sense to sustain a negative flow. The third and fourth constraints state that the total dissipated energy by any node cannot exceed its residual battery capacity. The last constraint (8) guarantees that the upper bound for the consumed energy per unit time due to the data transfer from i to j is γ_{LINK}. In other words, it ensures that $<i, j>$ is a viable link for $P_{i,j}$. Although we put a limit on the link, the system-wide energy is still not constrained. This might be sufficient for sparse networks. However, for large networks the system-wide consumed energy must be constrained. Figure 5.1 shows a wireless ad hoc network to which the above integer program can be applied. The set of $|M| = 2$ disjoint routes is shown in boldface. The rectangular objects represent sources of interference and fading.

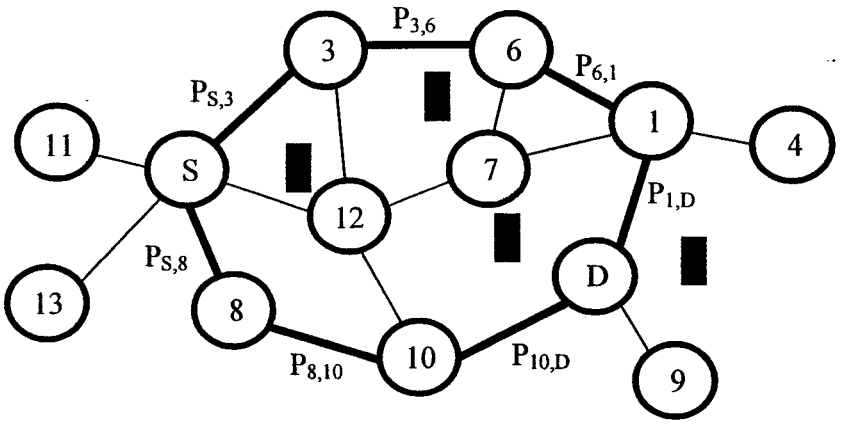


Figure 5.1: An example network with two S-D routes carrying traffic

Shannon's theorem imposes an upper bound on the channel capacity, and this upper bound should be reflected in the optimization problem. According to Shannon's theorem,

the channel capacity (C) is constrained by both the channel bandwidth (W) and signal-to-noise ratio (SNR):

$$C \le W \log_2(1 + S/N) \tag{9}$$

Hence, the total number of bits per second transmitted by the source must satisfy (9):

$$PKT\_SIZE. \sum_{s' \in n_s - \{s\}} P_{s,s'} \le W \log_2(1 + S/N) \tag{10}$$

Equivalently, it is sufficient to satisfy the data rate $(R \le C)$ of the wireless network interface card (NIC):

$$PKT\_SIZE. \sum_{s' \in n_s - \{s\}} P_{s,s'} \le R \tag{11}$$

We now consider a more realistic and sophisticated power consumption model in which the system-wide consumed energy due to the transfer of a single packet of fixed size from node i to node j is $\sum_{g \in n_i} f_{ij}(g)$ where

$$f_{ij}(g) = \begin{cases} f(Pw_{ij}^{tx}) : g = i, Pw_{ij}^{tx} > 0 \\ f(Pw_{ij}^{rx}) : g = j, Pw_{ij}^{rx} > 0 \\ f(Pw_{ig}^{rx}) : g \ne i, g \ne j, Pw_{ig}^{rx} > 0 \end{cases} \tag{12}$$

In the above, $f(Pw_{ig}^{rx})$ is the energy dissipated by a non-receiver node, g, that is a neighbor to i $(g \in n_i - \{i, j\})$ due to processing the MAC header of a single frame. Hence, the total dissipated energy attributed to $P_{i,j}$ packets from i to j is equal to $P_{i,j} \sum_{g \in n_i} f_{ij}(g)$, and the total amount of energy consumed by the source in a single time unit now becomes

$$\sum_{s' \in n_s - \{s\}} P_{s,s'} f(Pw_{ss'}^{tx}) + \sum_{s' \in n_s - \{s\}} \left(\sum_{s'' \in (n_{s'} - \{s'\})} P_{s',s''} f(Pw_{s's''}^{rx}) \right) \tag{13}$$

while the consumed energy by node $i \neq s$ is altered to

$$\sum_{j \in n_i - \{i\}} P_{i,j} f(Pw_{ij}^{tx}) + \sum_{h \in n_i - \{i\}} \left(\sum_{h' \in (n_h - \{h\})} P_{h,h'} f(Pw_{hi}^{rx}) \right) \tag{14}$$

We can now replace the program in Table 5.1 with the one in Table 5.2. In the new integer program, Q-GSL constraints (18) and (19) replace (6) and (7), respectively, and ensure that the source s and each and every viable node i are also feasible nodes with respect to the assignment of packets to each $l \in L = \bigcup_{R \in M} L_R$. Despite the fact that the set of constraints in Table 5.2 ensure that the energy consumed due to the packets traversing a single link in one unit time does not exceed a certain threshold, the system-wide consumed energy is unconstrained and may in some cases be large.

Table 5.2
Integer program for determining optimal packet flow for disjoint routes

Maximize $z = \sum_{n \in n_s - \{s\}} P_{s,n}$ (15)

s.t.

Conservation of flow constraint:

$$P_{h,i} = P_{i,j}, \forall < h,i >, < i,j > \in R, \forall R \in M \tag{16}$$

Non-negativity constraint:

$$P_{i,j} \geq 0, \forall < i,j > \in R, \forall R \in M \tag{17}$$

Energy constraints:

$$T_{SD}\left(\sum_{s^{'}\in n_s-\{s\}} P_{s,s^{'}} f(Pw_{ss^{'}}^{tx}) + \sum_{s^{'}\in n_s-\{s\}}\left(\sum_{s^{"}\in(n_{s^{'}}-\{s^{'}\})} P_{s^{'},s^{"}} f(Pw_{s^{'}s^{"}}^{rx})\right)\right) \le BC_s \tag{18}$$

$$T_{SD}\left(\sum_{j\in n_i-\{i\}} P_{i,j} f(Pw_{ij}^{tx}) + \sum_{h\in n_i-\{i\}}\left(\sum_{h^{'}\in(n_h-\{h\})} P_{h,h^{'}} f(Pw_{hi}^{rx})\right)\right) \le BC_i,$$

$$\forall i \in R_m - \{s\}, \forall m \in \{1,...,|M|\} \tag{19}$$

$$P_{i,j}\sum_{g\in n_i} f_{ij}(g) \le \gamma_{LINK}, \forall <i,j> \in L_R, \forall R \in M \tag{20}$$

Channel capacity constraint:

$$PKT\_SIZE.\sum_{s^{'}\in n_s-\{s\}} P_{s,s^{'}} \le R \tag{21}$$

B. Non-Mutually Exclusive Paths

In practical networks, M might contain two or more intersecting routes. For instance, in the example wireless mobile ad hoc network shown in Figure 5.2, |M| = 3, and $R_1 = \{S,3,6,1,D\}, R_2 = \{S,3,12,7,10,D\}, R_3 = \{S,8,10,D\}$. Therefore, we clearly have $R_1 \cap R_2 \neq \{S,D\}$ and $R_2 \cap R_3 \neq \{S,D\}$. Thus, paths R_1 and R_2, and paths R_2 and R_3 are not mutually exclusive. This allows for the division of the outgoing flow at some nodes, and the aggregation of the incoming flows into others. For example, the flow into node 3 is split amongst nodes 6 and 12, and that from nodes 8 and 12 is augmented at node 10.

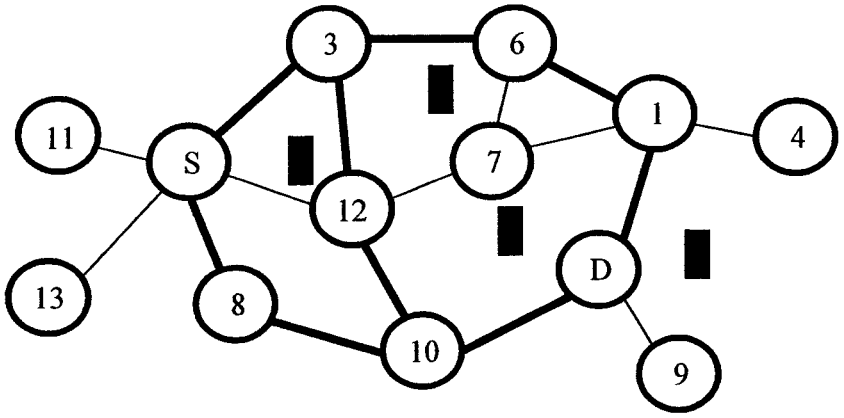


Figure 5.2: An example network with non-mutually exclusive S-D routes

One modification to the linear program in Table 5.2 enables us to capture this "packet splitting" routing problem by substituting (22) below for (16):

$$\sum_{m=1}^{|M|} \sum_{h \in R_m, h \neq i} P_{h,i} = \sum_{m=1}^{|M|} \sum_{j \in R_m, j \neq i} P_{i,j}, \forall i \in \bigcup_{m=1}^{|M|} R_m \tag{22}$$

C. System-Wide Constraint of Energy Dissipation

While the IP so far limits energy dissipation on a per-link basis it would be more meaningful to limit it on a per-node basis. That is to say, all nodes participating in routing must be viable nodes with respect to the total number of packets traversing them per unit time. This would have a favorable effect on MAC contention in the vicinity of the transmitter in addition to reducing energy consumption. This is achieved by adding the following constraint

$$\sum_{m=1}^{|M|} \sum_{j \in R_m, j \neq i} P_{i,j} \sum_{g \in n_i} f(g) \leq \gamma_{ZONE}, \forall i \in R_k, \forall k \in \{1,\ldots,|M|\} \tag{23}$$

Nevertheless, the network must only accommodate flows that will not cause congestion. Thus, we replace (23) with (23').

$$\sum_{m=1}^{|M|} \sum_{j \in R_m, j \neq i} P_{i,j} \leq \frac{\gamma_{ZONE}}{E[EPP]} - \sum_{g \in n_i} EMC_g, \forall i \in R_k, \forall k \in \{1,...,|M|\} \quad (23')$$

$E[EPP]$ is the expectation of the consumed energy per packet and $\sum_{g \in n_i} EMC_g$ is the sum of the EMC for node i and its neighbors. Obviously, $E[EPP]$ is dependent on the power consumption model employed. Figure 5.3 shows an example in which there exists a single ongoing flow and each node advertises the sum of its own and its neighbors' EMC values. In this example, the links shown in boldface are ones pertaining to the route used between the source and the destination (nodes 1 and 4, respectively). Table 5.3 shows λ_i and μ_i for all nodes i in Figure 5.3. The flow from the source is 5 *pkts/sec*. From Table 5.3, $\mu_i = 5, \forall i \in \{2,3,4\}$; this is so because nodes 2, 3, 4, in addition to 1, constitute the route between the source and the destination.

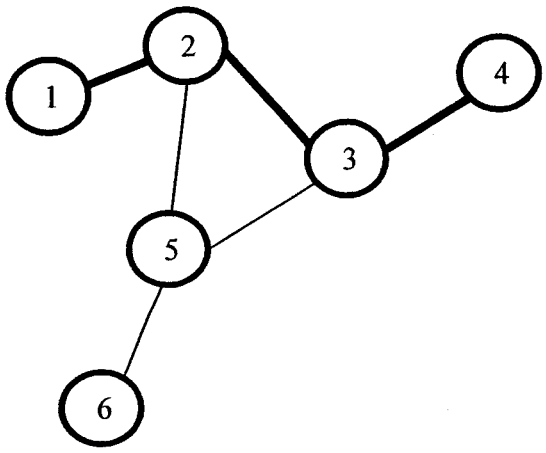


Figure 5.3: A wireless ad hoc network with a single active flow

It can be shown that $\sum_{g \in n_i} EMC_g$ is equal to $\lambda_i + \mu_i + \sum_{g \in n_i - \{i\}} (\lambda_g + \mu_g - P_{i,g} - P_{g,i})$. This is the value i advertises in its periodic beacons and is simply calculated using node i's knowledge of active flows. Therefore, $EMC_i = \begin{cases} 5 : i \in \{1,2,3,4\} \\ 0 : i \in \{5,6\} \end{cases}$. Apparently, the destination advertises a nonzero EMC value despite the fact that it acts merely as a receiver. However, it still responds to RTS frames with CTS messages which have the same effect as a data frame assuming all data frames are of equal length. Thus, this should be reflected in its EMC.

Table 5.3
EMC for the nodes in Figure 5.3

Wireless Node (i)	λ_i	μ_i
1	5	0
2	0	5
3	0	5
4	0	5
5	0	0
6	0	0

Furthermore, we would like to enhance Q-GSL further by introducing the constraint of system-wide energy dissipation. This can significantly enhance the lifetime of the wireless ad hoc network by reducing the volume of traffic traversing it. Thus, we introduce a new threshold, namely, the system threshold (γ_{SYSTEM}). Also, we realize that all the energy constraints introduced thus far are necessary for preserving battery capacity, but are not sufficient to guarantee a specific minimal system lifetime (as defined in Section 5.2). A single modification allows our program to capture this problem:

$$\left(\sum_{j \in n_i - \{i\}} P_{i,j} f(Pw_{ij}^{tx}) + \sum_{h \in n_i - \{i\}} \left(\sum_{h' \in (n_h - \{h\})} P_{h,h'} f(Pw_{hi}^{rx}) \right) \right) \leq \gamma_{NODE}, \forall i \in R_k, \forall k \in \{1,...,|M|\} \tag{24}$$

Or, equivalently,

$$\left(\sum_{m=1}^{|M|} \sum_{j \in R_m, j \neq i} P_{i,j} f(Pw_{ij}^{tx}) \right) + \left(\sum_{m=1}^{|M|} \sum_{h \in (R_m - \{i\}) \bigcap n_i} \sum_{h' \in (n_h - \{h\})} P_{h,h'} f(Pw_{hi}^{rx}) \right) \leq \gamma_{NODE}, \forall i \in R_k, \forall k \in \{1,...,|M|\} \tag{25}$$

We simply add (24) or (25) to our new integer program shown in Table 5.4. Either one ensures that the sum of the energy per unit time spent by any one node never exceeds a specific threshold. For example, for node i and $BC_i = 2$ kJ, and $\gamma_{NODE} = 0.56$ W, node i is guaranteed a minimum lifetime of approximately one hour. In the case where each node may have its own initial battery capacity, we can guarantee a minimum system lifetime of X hours by letting

$$\gamma_{NODE} = \frac{\min_i BC_i}{X} \tag{26}$$

The new program is shown in Table 5.4. This means that Q-GSL is NP-hard. The solution obtained by the linear programming (LP) relaxation of the Q-GSL IP and then rounding it off (or downward) does not necessarily yield the optimal solution. In fact, the candidate solution may even turn out to be an infeasible one for Q-GSL in the first place. Nevertheless, our Q-GSL framework may be formulated as a minimum cost network flow problem, which can be solved in polynomial time. For the sake of readability, however, in Table 5.4 we show the original Q-GSL IP before the conversion. In spite of the fact that the resultant optimal solution due to the LP relaxation is not the optimal

solution of the IP (i.e., it is actually less than or equal to the optimal), it is a feasible one. Most importantly, our approach herein is conservative and is desirable because it somewhat indirectly compensates for the absence of other energy overheads.

Table 5.4
Q-GSL for joint and disjoint routes

$$\text{Maximize } z = \sum_{n \in n_s - \{s\}} P_{s,n} \tag{27}$$

s.t.

Conservation of flow constraint:

$$\sum_{m=1}^{|M|} \sum_{h \in R_m, h \neq i} P_{h,i} = \sum_{m=1}^{|M|} \sum_{j \in R_m, j \neq i} P_{i,j}, \forall i \in \bigcup_{m=1}^{|M|} R_m \tag{28}$$

Non-negativity constraint:

$$P_{i,j} \geq 0, \forall < i, j > \in R, \forall R \in M \tag{29}$$

Energy constraints:

$$T_{SD}\left(\sum_{s' \in n_s - \{s\}} P_{s,s'} f(Pw_{ss'}^{tx}) + \sum_{s' \in n_s - \{s\}} \left(\sum_{s'' \in (n_{s'} - \{s'\})} P_{s',s''} f(Pw_{s's''}^{rx}) \right) \right) \leq BC_s \tag{30}$$

$$T_{SD}\left(\sum_{j \in n_i - \{i\}} P_{i,j} f(Pw_{ij}^{tx}) + \sum_{h \in n_i - \{i\}} \left(\sum_{h' \in (n_h - \{h\})} P_{h,h'} f(Pw_{hi}^{rx}) \right) \right) \leq BC_i, \tag{31}$$

$$\forall i \in R_m - \{s\}, \forall m \in \{1,...,|M|\}$$

$$\sum_{m=1}^{|M|} \sum_{j \in R_m, j \neq i} P_{i,j} \leq \frac{\gamma_{ZONE}}{E[EPP]} - \sum_{g \in n_i} EMC_g, \forall i \in R_k, \forall k \in \{1,...,|M|\} \tag{32}$$

$$\sum_{m=1}^{|M|} \sum_{<i,j> \in L_m} P_{i,j} \sum_{g \in n_i} f_{ij}(g) \leq \gamma_{SYSTEM} \tag{33}$$

$$\left(\sum_{m=1}^{|M|}\sum_{j\in R_m, j\neq i} P_{i,j} f(Pw_{ij}^{tx})\right)+\left(\sum_{m=1}^{|M|}\sum_{h\in(R_m-\{i\})\bigcap n_i}\sum_{h'\in(n_h-\{h\})} P_{h,h'} f(Pw_{hi}^{rx})\right)\leq \gamma_{NODE}, \forall i\in R_k, \forall k\in\{1,...,|M|\} \tag{34}$$

Channel capacity constraint:

$$PKT\_SIZE.\sum_{s'\in n_s-\{s\}} P_{s,s'} \leq R \tag{35}$$

5.4 EXPERIMENTS AND ILLUSTRATIONS

We present a few experiments that demonstrate the effect of the IP parameters on the optimum value for packet flow and the carried load by each of the $|M|$ routes. We use the sample network in Figure 5.2 as a basis for our experiments. However, there are no sources of interference other than overlapping transmissions. All packets are $1K$ bytes long, and the data rate (R) is equal to $2Mbps$. $Pw_{ij}^{tx}=1.33W$, $Pw_{ij}^{rx}=0.967W$. $f(Pw_{ij}^{tx})$, $f(Pw_{ij}^{rx})$ are 0.00545 and $0.004J/pkt$, respectively, and $f(Pw_{ig}^{rx})$ is equal to 0 otherwise. $T_{SD}=1s$, and $BC_i=100J:\forall i\in N$. In Figures 5.4 to 5.9, $EMC_i=0,\forall i$.

In the first couple of experiments (whose results are shown in Figures 5.4 and 5.5), we vary γ_{SYSTEM} between 0.05 and $0.5J/s$ and both γ_{ZONE}, γ_{NODE} are set to γ_{SYSTEM}. This is equivalent to a minimum system lifetime between 0.56 and 0.11 hour, respectively. Observing the simulation results in Figure 5.4 shows that the IP optimum increases with the increase in γ_{SYSTEM}. Evidently, for smaller values of γ_{SYSTEM}, any small increase produces a large increase in the total flow out of the source. For example, a 40% increase in γ_{SYSTEM} (from 0.05 to 0.07 J/s) results in a 100% increase in packet flow (from 1 to

2pkts/s). On the other hand, larger values of γ_{SYSTEM} have a less intense effect on the aggregate flow. For instance, an 11% increase (from 0.45 to 0.5 J/s) only results in increasing packet flow by 13% (from 15 to 17pkts/s).

Figure 5.5 shows that the routing load is distributed amongst routes A and C for optimal packet flow, whereas route B remains idle 100% of the time. This experiment raises the issue of the fair distribution of the aggregate flow. We observe that in the best case the aggregate flow is evenly distributed between routes A and C (50% each). In the worst case, however, 100% of the traffic is carried by route C. This is due to the fact that as γ_{SYSTEM} increases more and more packets can be transported to the destination using the shorter route (route C) that will result in the least amount of system-wide expended energy. In the following experiments we show how to eliminate this deficiency by using $\gamma_{NODE} < \gamma_{SYSTEM}$ rather than $\gamma_{NODE} = \gamma_{SYSTEM}$.

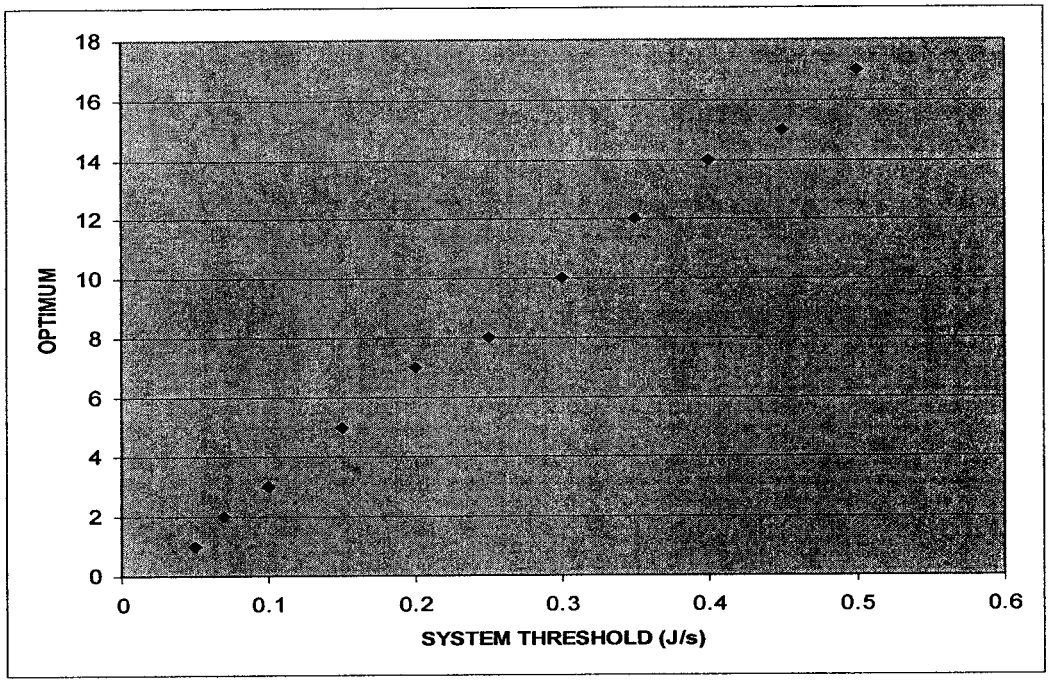


Figure 5.4: The effect of γ_{SYSTEM} on the optimum aggregate flow

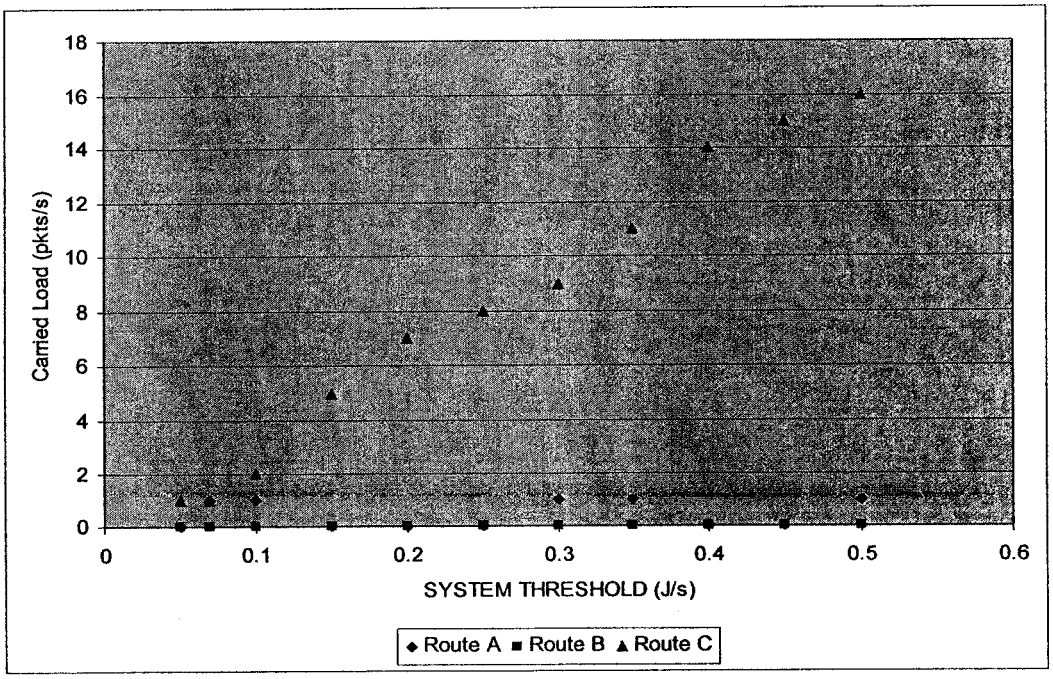


Figure 5.5: The effect of γ_{SYSTEM} on the load distribution

Figure 5.6 shows the results of two sets of experiments where γ_{NODE} is set to $\gamma_{SYSTEM}/10, \gamma_{SYSTEM}/5, \gamma_{SYSTEM}/3, \gamma_{SYSTEM}/2$, respectively. In the first set of experiments, γ_{SYSTEM} is equal to 0.5 J/s and is set to 5 J/s in the second set of experiments. When γ_{NODE} is 10% of γ_{SYSTEM} only 52% of the packets that may possibly be carried by the network without violating the system-wide energy constraint are actually admitted. Thus, γ_{NODE} restrains the flow of packets despite their adherence to the rest of the constraints. Nevertheless, when γ_{NODE} is only 33% of γ_{SYSTEM} all the packets which were to be admitted in the absence of γ_{NODE} (17 pkts/s when γ_{SYSTEM}=0.5 J/s and 176 pkts/s when γ_{SYSTEM}=5 J/s) are carried by the network.

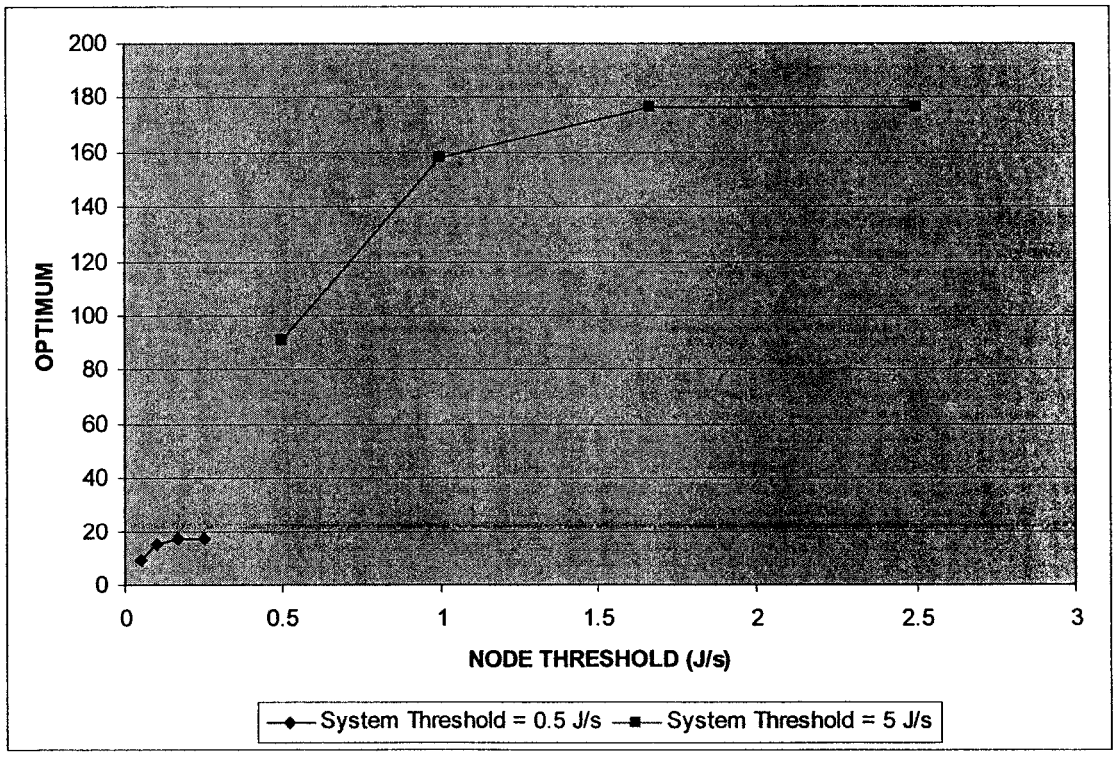


Figure 5.6: The effect of γ_{NODE} on the optimum aggregate flow

We study the effect of γ_{NODE} on the fair distribution of the routing load in Figures 5.7 and 5.8. From Figure 5.7, we notice that in the best case about 56% (5 out of 9 packets) of the traffic is carried by route A and the rest is carried by route C (the shorter of the two routes). This suggests that we achieve load balancing amongst the routes and the nodes in the wireless ad hoc network without violating any of the energy constraints, and while still guaranteeing a minimum system lifetime. However, the drawback here is that there is a reduction in the aggregate flow of packets from the source to the destination. However, this can be overcome by increasing γ_{NODE} slightly. The same conclusions may be drawn from Figure 5.8. However, there is a steep decline in the packets carried on route A as γ_{NODE} approaches γ_{SYSTEM}.

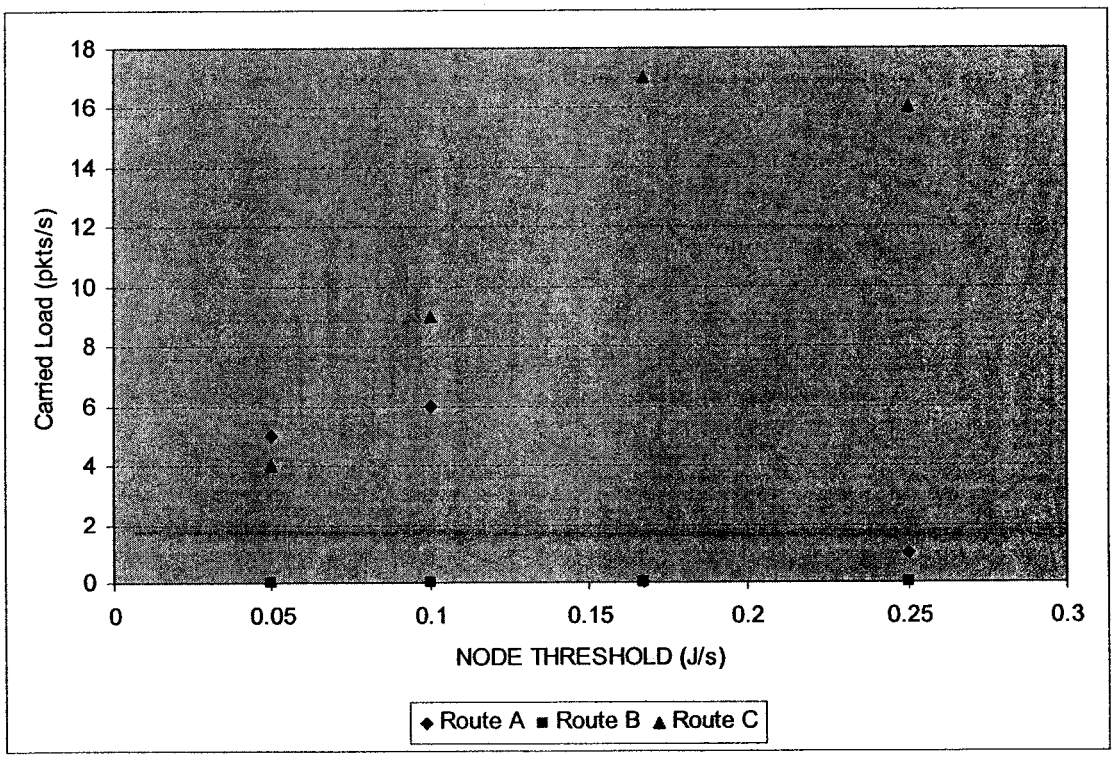


Figure 5.7: The effect of γ_{NODE} on the load distribution, $\gamma_{SYSTEM} = 0.5J/s$

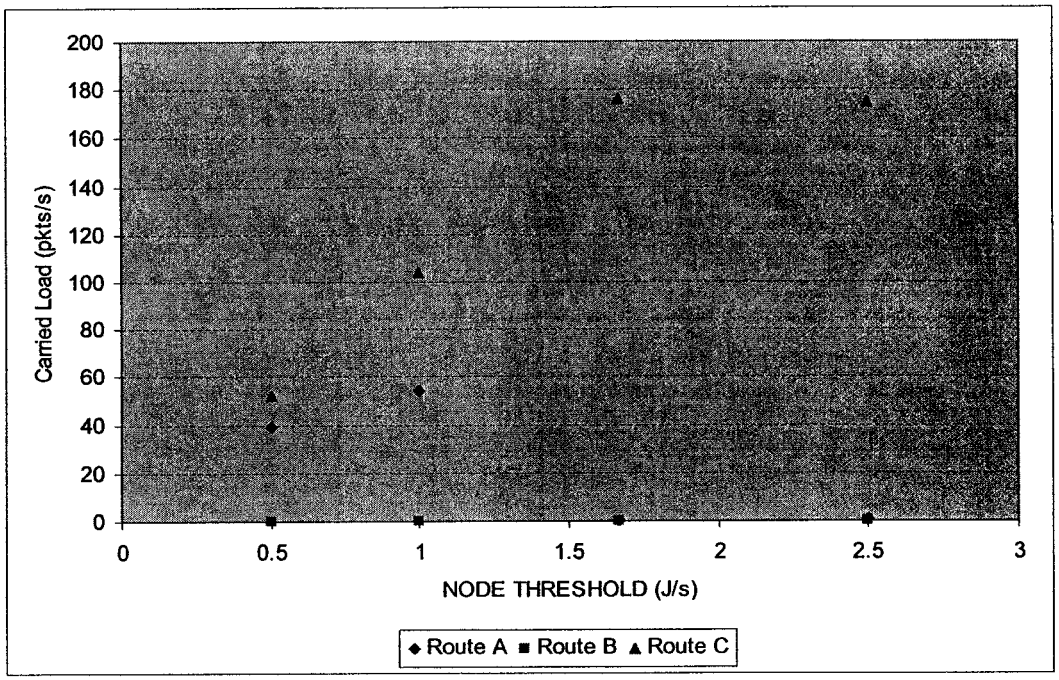


Figure 5.8: The effect of γ_{NODE} on the load distribution, $\gamma_{SYSTEM} = 5J/s$

The effect of γ_{ZONE} on the load distribution is shown in Figure 5.9. Table 5.5 shows the EMC values for all the nodes in Figure 5.2. Despite the fact that route C is the shortest route, node 8 has an EMC equal to 50 and at most 55 more packets are admissible when $\gamma_{ZONE} = 1J/s$. Route C's load drops by more than 68% (when $\gamma_{ZONE} = 1J/s$), as compared to the one in Figure 5.8 when $\gamma_{NODE} = 2.5J/s$, and constitutes 55% of the aggregate flow. When γ_{ZONE} is increased to 1.67*J/s* C's load makes up 77% of the total flow. In this case, a maximum of 126*pkts/s* may be routed via route C. Apparently, increasing γ_{ZONE} yields assigning a greater share of the flow to the shortest route. Thus,

the flow is carried solely by route C as γ_{ZONE} is increased to $2.5J/s$. In other words, the increment in γ_{ZONE} compensates for the large EMC value at node 8.

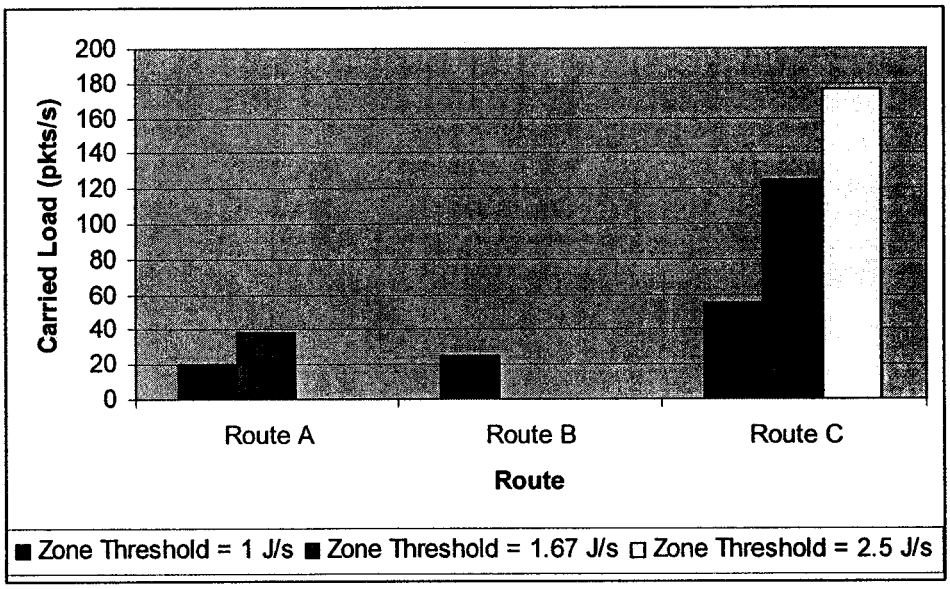


Figure 5.9: The effect of γ_{ZONE} on the load distribution, $\gamma_{SYSTEM} = 5J/s$

Table 5.5
EMC for nodes in Figure 5.2

Wireless Node	EMC
S	5
3	6
12	2
6	3
1	8
10	25
8	50
D	0

We study the effect of the EMC on the per-link load for routes A, B, and C, in Figures 5.10 to 5.12, respectively. The EMC at node 8 is set to 50, 75, 100, and 125, respectively, whereas the EMC values for the rest of the nodes are equivalent to the ones shown in

Table 5.5. γ_{ZONE} is set to $1J/s$. For route C and links <S,8> and <8,10>, which carry the same load, the number of packets per second decreases as the EMC increases. Hence, our framework routes around congestion and allocates more packets to less congested routes. This may be verified from Figures 5.10 and 5.11. When the EMC increases the portion of the load that was previously carried by route C is either shifted to A or B, or is efficiently distributed between both routes. Links <3,6>, <6,1>, and <6,D> solely carry the "extradited" load of route C when the EMC at node 8 is equal to 100, while links <3,12>, <12,10>, and <10,D> compensate for node 8's congestion.

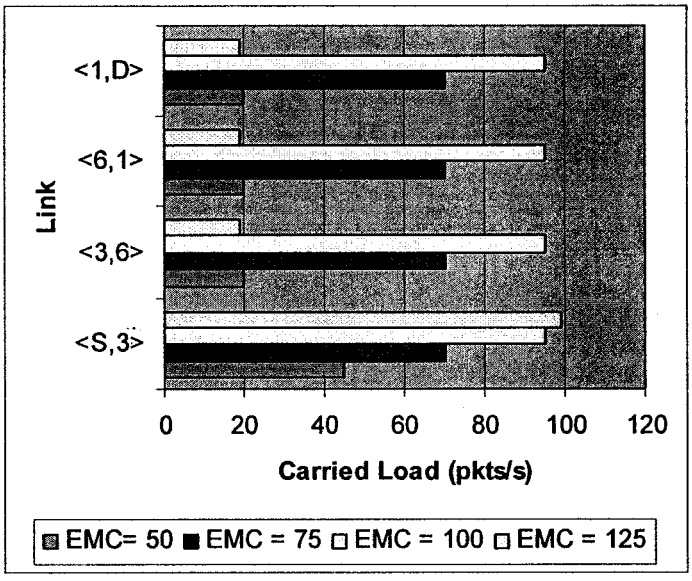


Figure 5.10: The effect of the EMC on the per-link load for route A

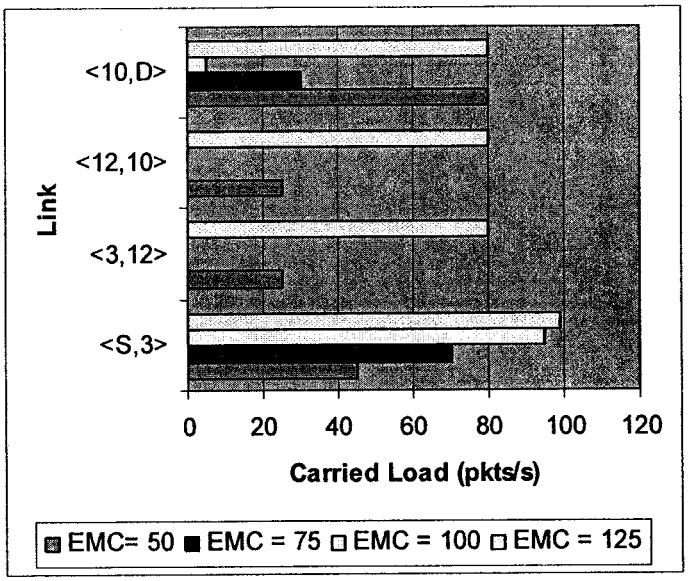


Figure 5.11: The effect of the EMC on the per-link load for route B

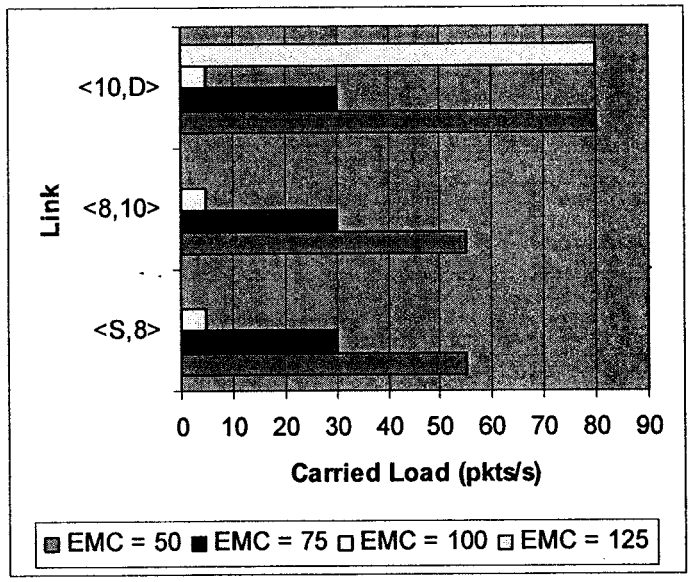


Figure 5.12: The effect of the EMC on the per-link load for route C

Figure 5.13 shows the impact of congestion on the optimum flow. In the two sets of experiments, γ_{ZONE} is equal to $1J/s$ and the EMC at node 8 is varied. In the first set of experiments, the EMC for all nodes, excluding node 8, are kept small enough so as not to congest routes A and B. It is evident from Figure 5.13 that despite the fact that route C is

congested at node 8, the optimum flow remains constant regardless of the value of 8's EMC. Nevertheless, there is a slight decrease (from 100 to 99*pkts/s*) in the optimum flow when the EMC=125. On the other hand, when both routes A and B become congested, the optimum flow drops sharply by 35 and 40% with EMC=100 and EMC=125, respectively. Table 5.6 shows the EMC used in the second set of experiments (except for node 8).

Table 5.6
EMC for nodes in Figure 5.2

Wireless Node	EMC
S	5
3	6
12	75
6	75
1	8
10	25
D	0

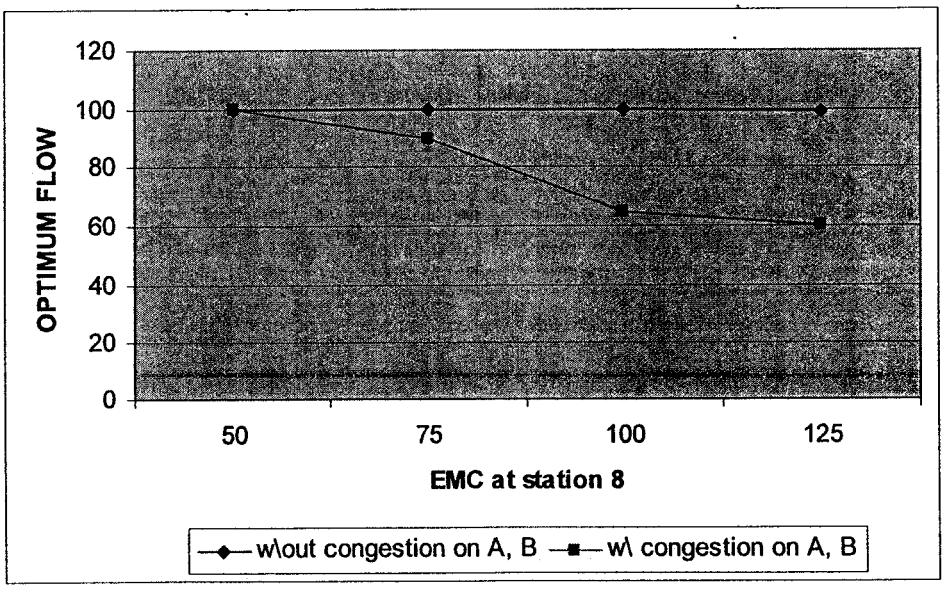


Figure 5.13: The impact of congestion on the optimum flow

5.5 SUMMARY

In previous chapters, we showed that minimum-hop routing performed better than some proposed power-aware schemes that use battery capacity solely as a means for energy conservation. This revealed that battery capacity alone cannot be used to satisfy our requirements for a power-efficient ad hoc wireless network. In this chapter, we devised a novel framework that would allow the transfer of data in a wireless ad hoc network using a set of $|M|$ energy-aware routes from the source to the destination while maintaining a pre-computed minimum system lifetime. Upper bounds on the aggregate packet flow in the presence of system-wide and nodal energy constraints were derived in addition to the individual loads carried by each of the $|M|$ routes. It is important to emphasize that γ_{SYSTEM} significantly enhances the lifetime of the wireless ad hoc network by reducing the volume of traffic traversing it. Although γ_{SYSTEM} affects the flow out of the source, the flow must still satisfy the battery capacity constraints at all the nodes en route to the destination. Increasing γ_{SYSTEM} adversely affects the fair distribution of the load since shorter routes that would result in the least amount of system-wide dissipated energy are always favored. However, this is mitigated through the incorporation of γ_{NODE}, the nodal threshold, into our optimizations. Unlike γ_{SYSTEM}, γ_{NODE} suffices to guarantee a specific minimum system lifetime and fairly distributes the routing load between different routes. Nevertheless, our experiments unveil a tradeoff between the fair distribution of the load and the aggregate flow admitted by the network. Moreover, we showed how Q-GSL routes around congestion and achieves load balancing despite its presence. Furthermore, Q-GSL may be used to provide QoS guarantees and maintain multiple active flows

without degrading its energy-conserving behavior and QoS-preserving performance provided that information about active flows is conveyed to source nodes requesting new connections/calls.

While Q-GSL accommodates fairly sophisticated power consumption models, other practical concerns, chief amongst them the energy expended in route discovery, route maintenance, and medium access, remain to be incorporated. It is our hope that the techniques reported here will provide a starting point for constructing the ultimate bounds on the aggregate flow for energy-limited wireless ad hoc networks. Based on our findings from Chapter 3, it is also of utter importance to develop an efficient and responsive wireless infrastructure to effectively maintain the framework proposed in this chapter. Such an infrastructure, which is thoroughly described in the following chapter, achieves further energy savings.

CHAPTER 6

POWER-AWARE VIRTUAL BASE STATIONS (PA-VBS)

In this chapter, we propose a novel infrastructure formation scheme for wireless mobile ad hoc networks. The proposed architecture, namely, *Power-Aware Virtual Base Stations (PA-VBS)*, mimics and maintains the operation of the conventional fixed infrastructure in cellular networks. In the PA-VBS protocol, a mobile node is elected from a set of nominees to act as a temporary base station within its zone based on its residual battery capacity. Likewise, we study the characteristics and performance of PA-VBS by means of simulation. It is shown that PA-VBS scales well to large networks of mobile stations, and that it outperforms other infrastructure-formation protocols in terms of load balancing. The PA-VBS architecture facilitates the development of a comprehensive and promising framework for quality of service (QoS) management in wireless mobile ad hoc networks once the proper integration of the MAC protocol with the routing and call admission control mechanisms is established. Moreover, it lays the groundwork for

assigning bandwidth, and/or implementing priorities, and hence for QoS-based routing by conveying the quality of a path prior to call setup. To the authors' best knowledge, this is the first time that energy is used as a basis for developing a wireless mobile infrastructure, and achieving load balancing.

This chapter is organized as follows. The next section discusses the inevitable need for a wireless infrastructure. Existing infrastructure-creation protocols in wireless mobile ad hoc networks are discussed in Section 6.2. In Section 6.3, the PA-VBS protocol is described in detail. This is followed in Section 6.4 by a thorough performance evaluation study of PA-VBS. Finally, Section 6.5 presents the conclusions drawn from the chapter.

6.1 A WIRELESS INFRASTRUCTURE: THE INEVITABLE NEED

Developing an infrastructure for the infrastructure-less wireless mobile ad hoc networks is of utmost importance. Such an infrastructure reduces the problem of wireless mobile ad hoc communications, from a multi-hop problem, to a single-hop one, as in conventional cellular communications. The previous apprehension originates from the intuitive realization of the problems of medium access and routing in wireless mobile ad hoc networks and their effects on QoS-guaranteed communications. Infrastructure-based communications enable the use of simple variations of the widely used cellular protocols in ad hoc networks. Such a dynamic infrastructure will form the basis for developing and running MAC protocols, and routing algorithms, such as ECPS, E2LA, and our deterministic scheme in Chapter 5, which can utilize it to conform to the different QoS requirements. Infrastructure-based wireless mobile ad hoc communications will help

circumvent routing and medium access control issues. Therefore, in this chapter, an energy-conserving wireless mobile infrastructure is developed for ad hoc networks. We propose a novel infrastructure-creation scheme for wireless mobile ad hoc networks, namely, *Power-Aware Virtual Base Stations (PA-VBS)*. Nevertheless, the developed infrastructure is, essentially, a mobile one. The wireless mobile infrastructure acts as the executive regulatory authority that carries out mobility tracking, and, hence, routing in wireless mobile ad hoc networks. The proposed infrastructure-formation scheme demonstrates quick response to topological changes in the ad hoc network. Additionally, the protocol is scalable to networks with large populations of mobile stations. It outperforms current infrastructure-creation schemes in stability and load balancing among the mobile stations forming the infrastructure.

6.2 RELATED WORK

In [12], the node with the highest connectivity is chosen to be the centre of the cluster. This, in fact, introduces a major drawback to the stability of the various clusters since under high mobility cluster re-formations frequently take place. According to the scheme proposed in [59], a wireless node may only become a clusterhead if it is not associated with one. In our virtual base stations (VBS) protocol [53], [60], MTs are elected as clusterheads based on their ID numbers. However, the VBS protocol puts more emphasis on a node becoming a clusterhead, or a VBS, rather than being supervised by one. Hence, if a node receives a merge request, it responds by sending an accept-merge message, even if it is being supervised by a VBS. Moreover, it is noteworthy that this neither degrades intra-cluster nor inter-cluster communications by any means. If the node that became a

VBS was originally acting as a gateway for its previous VBS, it remains a gateway, besides being a VBS. This becomes of great significance if the criterion upon which VBSs are elected is one that relies on the assets possessed by the MTs of the ad hoc network. Processing speed, main and secondary storage, and MAC contention experienced in the neighborhood of the MT can be amongst such assets. Consequently, if an MT chooses not to become a clusterhead, only because it is under the supervision of another MT, even though it possesses the required assets to become one, the node requesting to merge might experience demoted communications to other nodes in the ad hoc network because it does not have the proper resources, nor is it able to be associated with one that does.

Figure 6.1 shows two illustrations of the VBS scheme. In Figure 6.1 (A), all the MTs are within radio range of one another, except 1 and 4. Due to the asynchronous transmission of the hello messages, MT 2 may broadcast its hello message before MT 1. Therefore, MTs 3, and 4, receive MT 2's hello message first. They send *merge-request* messages to MT 2 [a] and it sends *accept-merge* messages back to each one of them [b]. The scenario in Figure 6.1 (B) starts with MT 1 sending its hello message. MT 3, realizing that it heard from an MT whose ID number is smaller than that of their VBS, sends a merge-request message to MT 1 [a]. MT 1 sends back a merge-accept message [b]. After receiving the merge-accept messages, 3 sends a *dis-join* message to 2 [c], which removes them from its list of supervised MTs. 2 is still the VBS of 4, which did not send a dis-join message to MT 2 since it is not in the transmission range of MT 1.

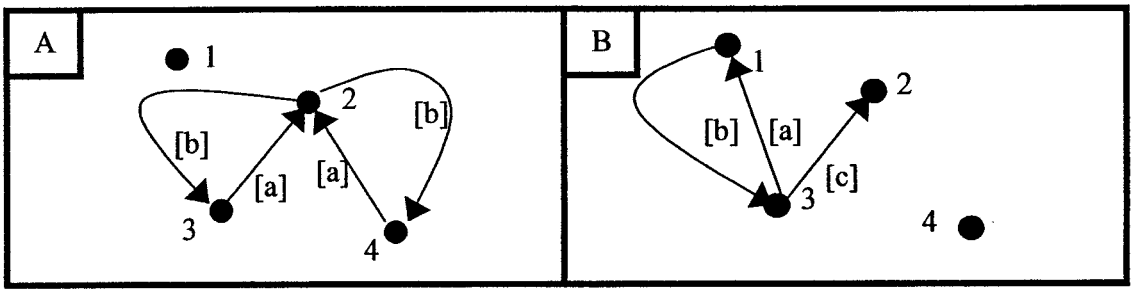


Figure 6.1: Finding the VBSs

6.3 THE POWER-AWARE VIRTUAL BASE STATIONS (PA-VBS) SCHEME

We herein adopt a new approach to achieve energy conservation by developing a novel infrastructure formation scheme for wireless mobile ad hoc networks based on the residual battery capacity of the wireless nodes. An overview of the proposed PA-VBS scheme is presented in Section 6.3.1. Moreover, a detailed description of the protocol, including the pseudo code of the algorithms that are of special importance to the operation of PA-VBS, is provided in Section 6.3.2. This is also in addition to the finite state machine describing the protocol, and its accompanying state transition table.

6.3.1 OVERVIEW OF THE SCHEME

In our scheme, some of the MTs, based on their current residual battery capacity, become in charge of all the MTs in their neighborhood, or a subset of them. This can be achieved by electing one to be a *Power-Aware Virtual Base Station (PA-VBS)*. If a VBS moves or stops acknowledging its presence via its periodic beacons, also called *hello* messages, for a period of time, a new one is elected. Electing a single VBS from a set of nominees is

done in an efficient way. Every MT has a *sequence number* that reflects the changes that occur to that MT. Sequence numbers are not only used for the sake of taking proper routing decisions, as in the case of the VBS protocol, but are also used to save battery energy whenever possible. In addition to the sequence number, a *myVBS* variable is used to store the ID number of the VBS in charge of that MT. If an MT has a VBS, its myVBS variable is set to the ID number of that VBS, else if the MT is itself a VBS, then the myVBS variable is set to 0, otherwise it is set to –1.

MaxPower$_i$ *(MP*$_i$*)* is defined as the battery capacity whose value is in one-to-one correspondence with the amount of time in seconds that MT$_i$ when used by a class-1 user, would last, starting from the time it had a fully charged battery, without having to be re-charged, provided that (1), (2), and (3), below are true.

(1) It remains a class-1 user during the whole MP$_i$ period

(2) It does not become a VBS during the whole MP$_i$ period

(3) It has inactive neighbors during the whole MP$_i$ period.

Likewise, *MAX_POWER* is a constant defined as the minimum required battery capacity for a class-1 user's MT to last, starting from the time it had a fully charged battery, for exactly one day without having to be re-charged, provided that the following are true:

(1) It remains a class-1 user during the whole one-day period

(2) It does not become a VBS during the whole one-day period

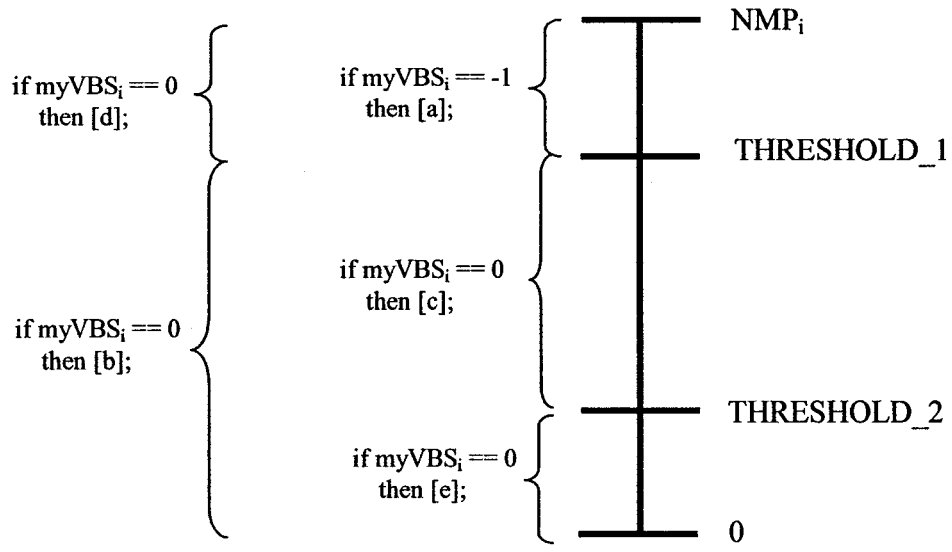
(3) It has inactive neighbors during the whole one-day period

Besides, *NormalizedMaxPower$_i$ (NMP$_i$)* is equal to MP_i / MAX_POWER. Hence, NMP_i can be equal to 1 only when MP_i is equal to MAX_POWER. The current *NormalizedPowerValue* for MT_i, *NPV$_i$*, is equal to the instantaneous battery capacity of MT_i, divided by MAX_POWER.

As will be explained in Section 6.4, PA-VBS facilitates load balancing between the wireless nodes. This is largely due to the fact that PA-VBS takes the nodal activity of the MTs a VBS is in charge of into account when calculating the amount of consumed power.

An MT is chosen by one or more MTs, to act as their VBS based on some power thresholds. MTs announce their NPVs in their periodic hello messages. An MT sends a *merge-request* message to another MT if the latter has an NPV greater than or equal to the former's, and a predetermined energy threshold. The receiver of the merge-request responds with an *accept-merge* message only if its NPV is above the first threshold, namely *THRESHOLD_1*, at the time it receives the merge request message, in which case it increments its sequence number by 1 to reflect the change, and sets its myVBS variable to 0. When the MT receives the accept-merge, it increments its sequence number by 1 and sets its myVBS variable to the ID number of its new VBS. If an MT hears from another MT whose NPV is larger than that of its VBS, it does not send a merge-request message to it as long as its VBS's NPV is above THRESHOLD_1 (also called the *association threshold*). A *dis-join* message is sent by an MT to its VBS only if the transmissions of the VBS have not been heard by the MT for some timeout period. If the VBS receives the dis-join message, it removes the sender from its list of MTs, which it is

in charge of, and it increments its sequence number by one. Figure 6.2 summarizes PA-VBS decision-making based on the energy thresholds.



[a]: send a merge request message to an MT with NPV greater than mine.
[b]: merge request messages can be sent by the MTs supported by this VBS, provided that the receiver's NPV is greater than THRESHOLD_1.
[c]: discard any merge request.
[d]: no merge request messages will be sent by the MTs supported by this VBS.
[e]: set the iAmNoLongerYourVBS flag to true.

Figure 6.2: PA-VBS decision making based on the energy thresholds

6.3.2 Detailed Description Of The Protocol

This section contains the pseudo code for the algorithms that are of special importance to the operation of PA-VBS. MTs broadcast their current NPVs as part of their hello messages. Hello messages contain other useful pieces of information, such as, the sequence number and the *iAmNoLongerYourVBS* flag. The iAmNoLongerYourVBS flag is used by a node acting as a VBS (see Figure 6.3) to convey to the MTs it is currently in charge of whether it can support them for another hello period or not. When this flag is set to false, the MTs receiving the hello message know that they will still be served by

their VBS for another hello period, and therefore do not need to look for a new one for at least one more hello period. However, if that flag is set by the VBS to true, then the MTs return to their initial state (see Figure 6.4). This actually takes place when the NPV of the VBS drops below the second energy threshold, namely *THRESHOLD_2*.

```
1.    if (myVBS == 0)
2.        if (iAmBelowThreshold_2())
3.            N = 0;
4.            cancelAllMyMTsTimers();
5.            iAmNoLongerYourVBS = true;
6.            mySequenceNumber++;
7.            myVBS = -1;
8.        else
9.            iAmNoLongerYourVBSFlag = false;
```

Figure 6.3: Using the iAmNoLongerYourVBS flag by a VBS for service denial

```
1.    if ( (myVBS > 0 ) && (theSenderOfTheHelloIsMyVBS()) )
2.        if (iAmNoLongerYourVBS == true)
3.            cancelMyVBSTimer();
4.            mySequenceNumber++;
5.            myVBS = -1;
6.        else
7.            restartMyVBSTimer();
8.            lastPowerValueReportedByMyVBS =
              myVBS.normalizedPowerValue;
```

Figure 6.4: Using the iAmNoLongerYourVBS flag by an MT to detect service denial

Upon receiving a hello message, the MT sends a merge request message to the sender, if and only if one of the following two cases is satisfied:

(1) The MT is neither a VBS nor being supported by one, and the following hold:

a. Its NPV is less than the NPV of the sender of the hello message.

b. The NPV of the sender is above THRESHOLD_1.

(2) The MT is currently supported by a VBS, and the following hold:

a. The last reported NPV by the current VBS is below THRESHOLD_1.

b. The NPV of the MT is less than the NPV of the sender of the hello message.

c. The NPV of the sender is above THRESHOLD_1.

The pseudo code of the algorithm executed by the MT to determine whether to send a merge request to its neighbor, upon receiving a hello message from it, is shown in Figure 6.5. The pseudo code shows that the MT does not send a dis-join to its current VBS. This can be explained as follows. Even though the MT sends the merge request based on the conditions listed above, this does not necessarily mean that the node receiving the merge request will accept the merge at the time the request is actually received. If the merge request is sent at time t_0 when the VBS is capable of supporting the MT, it might be received at time t_1 when its NPV is below THRESHOLD_1.

```
1.    if
      ((received NPV>=THRESHOLD_1) && (received NPV> my NPV))
      &&
      ((myVBS == -1)
      ||
      ((myVBS > 0) && (lastEnergyValueReportedByMyVBS < THRESHOLD_1)))
2.                sendMergeRequestMessageTo(senderOfHello);
```

Figure 6.5: The merge decision process

The pseudo code of the routine executed by an MT when it receives a merge request from one of its neighbors is shown in Figure 6.6. As stated in line 1, if and only if the receiving MT's NPV is currently above THRESHOLD_1 will it then proceed to accept the merge request. If the condition in line 1 is satisfied, the receiver increments the number of MTs it is in charge of by 1, sets its myVBS variable to 0 to reflect that it is currently a VBS, and increments its own sequence number by 1 (see lines 2-4). In line 5, the VBS starts a timer to trigger the initiation of a timeout period. The timer is reset every time the VBS receives a hello message from the MT. If the timer expires, the VBS will no longer be in charge of the MT. This can happen if the VBS, or, equivalently, the MT, moves out of the wireless transmission range of the MT. If it happens that they become within the wireless transmission range of each other after the expiration of the timer, then the MT must send a new merge request to the VBS. The VBS may then accept or reject the merge request based on its NPV at the time of the reception of the merge request. If the receiver of the merge request is being supported by another node acting as its VBS, then it cancels the timer corresponding to its VBS, and sends a dis-join message to it (lines 7-9). In the case where a node is incapable of serving a neighbor, PA-VBS supports two means of notifying an MT of a rejected merge request. The first is an *implicit merge reject* method where the MT assumes that its merge request is rejected if it does not receive anything back from the VBS. The other method relies on sending a *merge reject* message back to the MT informing it that its request cannot be supported (line 11). The implicit merge reject method can be utilized in highly congested zones, while merge reject messages can be used otherwise.

```
1.    if (MTisAboveThreshold_1())
2.          N++;
3.          myVBS = 0;
4.          mySequenceNumber++;
5.          startTimerForTheMT();
6.          sendAcceptMergeMessageTo(senderOfMergeRequest);
7.          if (myVBS > 0)
8.                cancelTimerOfMyOldVBS();
9.                sendDisjoinMessageToMyOldVBS();
10.   else
11.         do nothing; or sendMergeRejectMessageTo(senderOfMergeRequest);
```

Figure 6.6: The mergeRequestReceipt() algorithm

When an MT receives an accept-merge message it performs the algorithm in Figure 6.7. The receiver sends a disjoin message back to the issuer of the accept-merge message, as in line 2, if it is either a VBS or an MT supported by one. Otherwise, the receiver of the accept-merge message executes lines 3-6. It first sets its myVBS variable to the ID number of the node from which it received the accept-merge message (line 4). It also stores the NPV sent by its VBS, for future merge-related decisions, and starts a timer corresponding to its VBS. If the timer expires, the MT will no longer be associated with its VBS. The MT can then be associated with another VBS, regardless of the last NPV reported by its previous VBS. Hence, even if the last reported power value was above THRESHOLD_1, the MT can still issue new merge request messages.

```
1.    if (myVBS >= 0)
2.          sendDisjoinMessageTo(senderOfAcceptMerge);
3.    else if (myVBS == -1)
4.          myVBS = ID of the sender of the accept-merge;
5.          storeLastEnergyValueReportedByMyVBS();
6.          startTimerForMyVBS();
```

Figure 6.7: The acceptMergeReceipt() algorithm

The pseudo code shown in Figure 6.8 is used by an MT whenever a dis-join message is received. If the MT is a VBS, it increments its sequence number by one, decrements the number of nodes it is in charge of by one, and cancels the timer corresponding to the sender of the dis-join message, since it is no longer responsible for it (lines 2-4). Otherwise, the received dis-join message is discarded because the receiver is not a VBS (see lines 7-8).

```
1.    if (myVBS == 0)
2.          mySequenceNumber++;
3.          N--;
4.          cancelMTTimer(senderOFDisjoinMessage);
5.          if (N == 0)
6.                myVBS = -1;
7.    else
8.          discard the received dis-join message;
```

Figure 6.8: The disjoinReceipt() algorithm

Figure 6.9 shows the finite state machine at node i, FSM_i, running the PA-VBS scheme. The state transition table is shown in Table 6.1. FSM_i shows that any node i in the wireless ad hoc network, can be in one of four states. A node is in state 1 if and only if it is a VBS, and its NPV is greater than THRESHOLD_1. State 2 is reached when a node is a VBS, but its NPV lies between the two energy thresholds. If a node is neither a VBS, nor being supported by one, then it reaches state 3. Finally, a node in state 4 is one being supported by a VBS. An MT can be in state 1 and reach state 2 in one step if it consumes an amount of normalized energy equal to the difference between its NPV and THRESHOLD_2. Transitions from state 2 to state 1 become feasible as a result of battery recharging. If a VBS in state 1 is supporting a single MT, and receives a disjoin message

from it, or, equivalently, if the timer for that MT expires due to its, or the VBS's, motion, then the VBS will no longer be in charge of any other MT, and hence transits to state 3. Transitions in the opposite direction are also possible since an MT, that is not currently serving any other node, can receive a merge request message and move to state 1 on the condition that its NPV is above THRESHOLD_1. FSM_i clearly shows that transitions are not possible between states 2 and 4. Moreover, transitions from states 1 to 4, and states 3 to 2, are infeasible. An MT in state 2 must first reach state 3 before reaching state 4. This is because a direct transition from state 2 to state 4 is infeasible. In addition, an MT in state 4 can only reach state 2 via state 1. This is due to the fact that a node that is not acting as a VBS can become one provided that its NPV is above THRESHOLD_1. Once it is a VBS, its normalized battery capacity can drop below THRESHOLD_1, but still

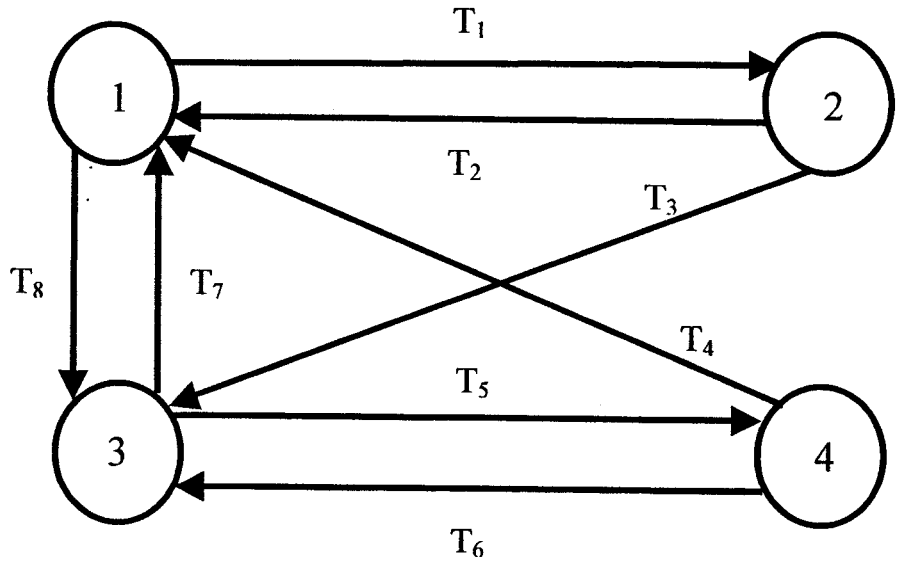


Figure 6.9: The finite state machine at node i, FSM_i, running the PA-VBS architecture

remain above THRESHOLD_2. In this case, it remains a VBS (and transits to state 2) provided that its NPV does not drop below THRESHOLD_2. Likewise, an MT in state 1 can only reach state 4 via state 3. In PA-VBS, a node cannot ask for another MT's

support if it is acting as a VBS. Therefore, a VBS cannot be supported by another MT in one step. Besides, a node in state 3 ought to first become a VBS with an NPV greater than THRESHOLD_1 (which implies transiting to state 1) before reaching state 2.

Table 6.1
The state transition table for FSM_i (TH1: THRESHOLD_1; TH2: THRESHOLD_2)

Transition	Event	Condition	Action
T_1	Power Consumption	$TH2 <= NPV_i <= TH1$	No more accept-merge messages
T_2	Charging	-	Send accept-merge messages to received merge request messages, N++, myVBS = 0, and mySequenceNumber++
T_3	The VBS's timer for its only MT expires, or a dis-join message is sent by the MT	The MT becomes associated with another VBS, or the MT is out of range	myVBS = -1, mySequenceNumber++, and N = 0
T_4	Merge request received from a neighbor	$NPV_i >= TH1$	Accept merge sent to neighbor, dis-join sent to my VBS, myVBS=0, mySequenceNumber++, N = 1, and start timer
T_5	Accept merge received from a VBS	$NPV_{VBS} > NPV_i$, and $NPV_{VBS} > TH1$	$myVBS = ID_{VBS}$, mySequenceNumber++, start timer, and store the reported NPV
T_6	The iAmNoLongerYourVBS flag is set to true, or the MT's timer for its VBS expires	-	myVBS = -1, and mySequenceNumber++
T_7	Merge request received from neighbor	$NPV_i > NPV_{neighbor}$, and $NPV_i > TH1$	Accept merge sent to neighbor, myVBS=0, mySequenceNumber++, and start timer
T_8	The timer, for the only MT, expires, or a dis-join message is sent by the MT	The MT becomes associated with another VBS, or is out of the wireless range	MyVBS = -1, mySequenceNumber++, and N = 0

6.4 PA-VBS PERFORMANCE EVALUATION

In this section, the performance of the PA-VBS infrastructure-creation scheme is studied. Section 6.4.1 describes our power consumption model. This is followed in Section 6.4.2 by a description of the simulation model. Besides, a description of the performance metrics that were taken into consideration in evaluating the performance of PA-VBS is given in the same section. The results of the conducted simulation experiments are presented in Section 6.4.3. The simulation results show that PA-VBS surpasses VBS in its overall performance.

6.4.1 POWER CONSUMPTION MODEL

For the sake of simplifying our analysis, both MP_i, for all i, and MAX_POWER, are assigned the corresponding unique duration in seconds. In addition, NPV_i is made equal to the unique value in seconds, corresponding to the current battery capacity value of MT_i, divided by MAX_POWER. Depending on the user's activity, every user can be in 1 of 10 classes at any time. A user who frequently toggles between the ON and OFF modes of operation can be possibly classified as a class-1 user. On the contrary, a user who uses his/her MT for palm-computing, playing games, or listening to music, is of a class other than 1, depending on the level of processing involved. The larger the number assigned to a user's class, the more processing done to accommodate the needs of the user, and, hence, the more power consumed.

The consumed energy during a period of time Δt is directly proportional to Δt. Therefore, and for the sake of finding out the amount of consumed energy, the operation period,

called *UpPeriod$_i$ (UP$_i$)*, from the last time the consumed energy by MT_i was calculated until the present, is considered in the energy consumption calculations. The average nodal activity, α, is a measure that reflects the percentage of time a cluster's medium is used to carry packets originating from the node, or packets delivered to the node by its VBS. α_1 / *CONSTANT$_1$* is defined as the amount of normalized energy drained from the battery of any MT in one day as a result of being a neighbor to exactly one node whose average nodal activity during the one-day period is α_1. In addition, α_2 / *CONSTANT$_2$* is, by definition, the amount of normalized energy drained from the battery of any VBS in one day due to serving only one MT, provided an *average nodal activity* or, equivalently, *carried routing load factor* equal to α_2. This is regardless of the user class of the VBS and its MP. α_2 can be any value between 1 and 10 depending on how active the mobile node is. An α_2 value of 1 means that the MT does not lie on any routing path, and that it only broadcasts its periodic hello messages to its neighbors. In other words, a value of 1 accounts for the minimum possible consumed/dissipated energy by a VBS; the consumed energy due to supporting a single mobile unit. Even though an MT can be inactive routing-wise, there is still some minimum processing required by the VBS to support the wireless node. The clusterhead not only processes the periodic hellos of its nodes to be able to provide routing support for them, but also regulates medium access in its cluster, and carries out packet scheduling.

The energy of an MT is drained due to two factors. The first factor is user-driven. On the contrary, the other factor is neighbor-driven. User-driven power consumption was explained earlier. However, neighbor-driven power consumption exists if and only if the

MT has one or more neighbors. If not, then the contribution of the neighbor-driven power consumption function to the total consumed power by the MT is 0. The user-driven consumed energy for MT_i is equal to: $UserClass_i$ * (UP_i / MAX_POWER). Besides, the neighbor-driven share of the consumed energy is equal to: [α_1 * (n_i / $CONSTANT_1$) + α_2 * (N_i / $CONSTANT_2$)] * (UP_i / MAX_POWER), where N_i is the number of MTs the VBS is currently supporting. In the case of a non-VBS node, MT_i, n_i translates to the number of wireless units which are neighbors to MT_i, whereas it is interpreted as the number of neighboring wireless nodes of MT_i which are not being supported by MT_i, otherwise. In addition, an MT can limit the number of mobile units it can support by setting its N_MAX_i variable to a value greater than or equal to N_MAX. Consequently, and to guarantee fair clustering, and achieve load balancing, an MT must be willing to support at least N_MAX mobile units. However, in our simulations, all the MTs had an N_MAX_i equal to N_MAX, and N_MAX is set to ∞. N_MAX_i can be used whenever the user experiences some unacceptable performance, in relation to the stand-alone applications running on the wireless unit, while it is a VBS.

It is noteworthy that the exact value of MAX_POWER for class-1 applications can be found using field-testing. Whenever a value is obtained, it can be substituted for the value in the formulas above. However, this should not affect the performance of PA-VBS. In addition, different MT units can have different MAX_POWER values. This also does not affect the correctness of our formulas. The only difference is that MAX_POWER_i is now used instead of MAX_POWER. The same can be also done for $CONSTANT_1$ and $CONSTANT_2$. Experimental values can be obtained and substituted for $CONSTANT_1$

and $CONSTANT_2$ into our power consumption formulas. Again, different MT units can have different values.

6.4.2 SIMULATION MODEL AND PERFORMANCE METRICS

A packet-level discrete-event simulator was developed in order to monitor, observe and measure the performance of the PA-VBS protocol. Initially, each mobile station is assigned a unique node ID, a random position in the x-y plane, and a random battery capacity between the nodes' maximum power value and THRESHOLD_2. The conducted simulation experiments are set up for wireless mobile ad hoc networks covering a 200 x 200 unit grid. The wireless transmission range of the MTs is set to 20 units. Hello messages are broadcasted every 1 second. The velocity of the mobile nodes is uniformly distributed between 0 and 10 units/second, and they are allowed to move randomly in any direction. Each simulation is run for 8 simulated hours (except for the simulations performed for the fourth experiment, which are run for 24 simulated hours), and the ad hoc network was sampled every 1 second. α_1 is neglected and set to 0. Likewise, α_2 is randomly chosen between 1 and 10, and $CONSTANT_2$ is set to 10. The second and sixth experiments are conducted for variable values of THRESHOLD_1. THRESHOLD_1 is otherwise set to 0.75, whereas THRESHOLD_2 is set to 0.25. The first three simulation experiments are conducted for ad hoc networks with 25, 50, 75, and 100 mobile nodes. However, in the rest of the simulation experiments, the wireless ad hoc network consists of 25 MTs, and the total energy consumed by each and every MT throughout the simulation experiment is obtained.

An explanation of the four noteworthy statistical performance metrics, measured by the simulators, follows:

1. *Average Number of VBSs* - The smaller this number, the more the number of mobile nodes that have to be served by each VBS, and vice versa.
2. *Average VBS Duration* - The average time duration (in seconds) for which a mobile node remains a VBS. This is a very important performance measure since it is a measure of system stability. This is due to the fact that the larger the duration, the more stable the scheme. Therefore, the sought value for this measure is actually infinity, as in conventional cellular networks where a base station serves as a base station during its lifetime, or the whole lifetime of the cellular network.
3. *Total Number of Mobile Nodes Elected as VBS* - The total number of mobile nodes elected as VBSs during the whole simulation run-time. A small value of this statistic reflects the system's tendency to elect the same set of VBSs. This implies that a small fraction of the mobile nodes is elected as VBSs in the case of small values.
4. *Total Energy Consumed by Every MT* – The total energy consumed by each and every MT throughout the simulation experiment. The closer these values are, the more evenly distributed the load is amongst the nodes of the wireless mobile ad hoc network.

The results of the corresponding experiments were compared against the VBS infrastructure-formation scheme that was shown in [53], [60] to overcome the drawbacks of the previously proposed infrastructure-creation protocols.

6.4.3 SIMULATION RESULTS

Observing the simulation results of Figure 6.10, shows that PA-VBS produces a larger number of clusterheads than VBS. This is actually because PA-VBS distributes the work load amongst a number of nodes that satisfy the power requirements, unlike VBS which tends to elect the same set of nodes with the smallest IDs again and again. The growth in the number of clusterheads is linear with the number of MTs in the case of PA-VBS. On the other hand, the number of clusterheads in the case of VBS remains almost constant. In networks with much larger populations, especially those that have CBR-traffic sources, this causes considerable MAC delays due to the large number of MTs that are simultaneously contending for medium access at a constant rate.

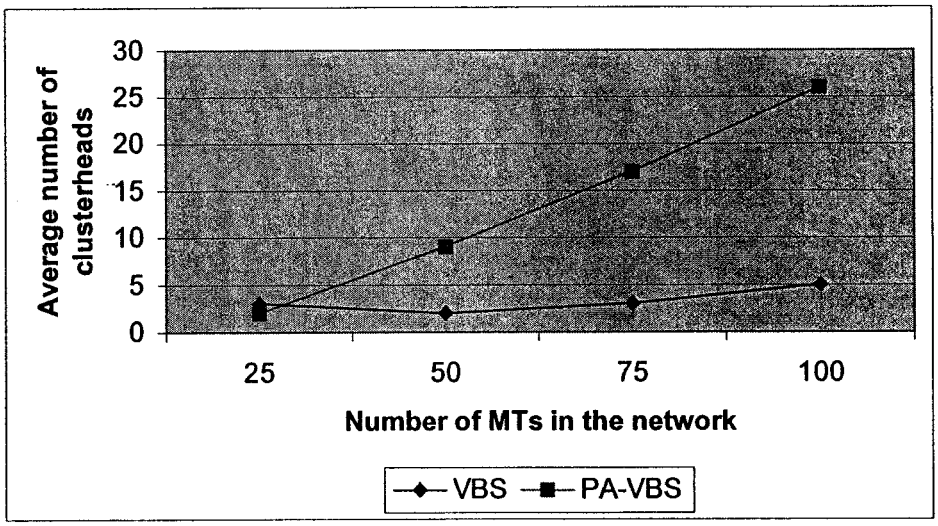


Figure 6.10: Impact of network density on number of VBSs

The average VBS duration, as explained before, is an important measure of the stability of any clusterhead-formation protocol. Figure 6.11 clearly shows, regardless of the value of THRESHOLD_1, that since VBS elects nodes based on their ID numbers, they remain

as clusterheads for longer periods than in the case of PA-VBS. However, in practice, this cannot be achievable since clusterheads consume more power than other MTs, and their battery power drains quicker. Hence, VBS is more prone to undergo disorder. The results show that the clusterhead duration, on average, is between 2.5 and 4 times more in the case of VBS. This implies that using VBS drains all the battery power of the clusterheads until they can no longer operate.

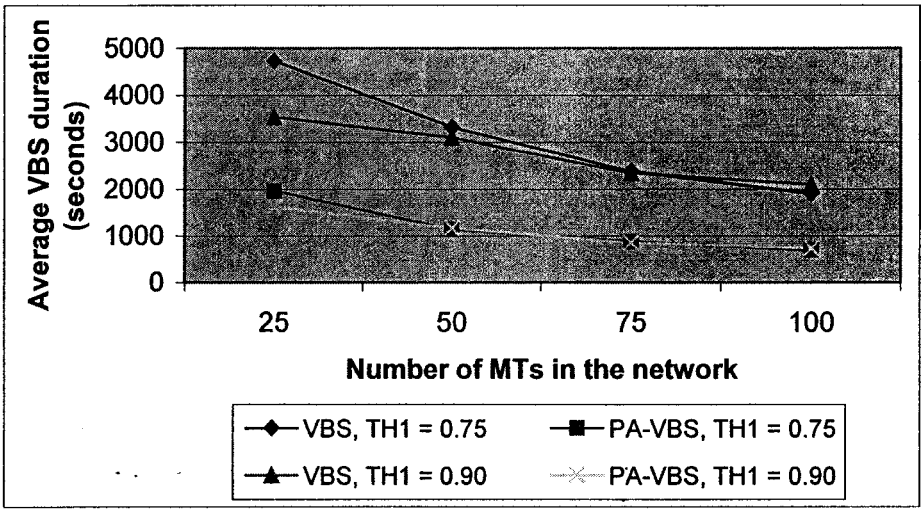


Figure 6.11: Impact of network density on clusterhead duration

Figure 6.12 shows that PA-VBS achieves load balancing amongst the nodes in the wireless ad hoc network. Every node is elected as a clusterhead at least once during the simulation run-time, regardless of the wireless transmission range. On the contrary, VBS elects a smaller fraction of the total number of mobile nodes as clusterheads during the entire simulation run-time. In addition, the total number of nodes elected as clusterheads decreases as the wireless range increases. This implies that VBS does not guarantee fairness amongst the wireless nodes as with PA-VBS. As a result of increasing the

wireless range, there is a 40% drop in the total number of VBSs in the case of 50 nodes, and around 29% with 75 and 100 MTs. This result proves that PA-VBS attains fair clustering.

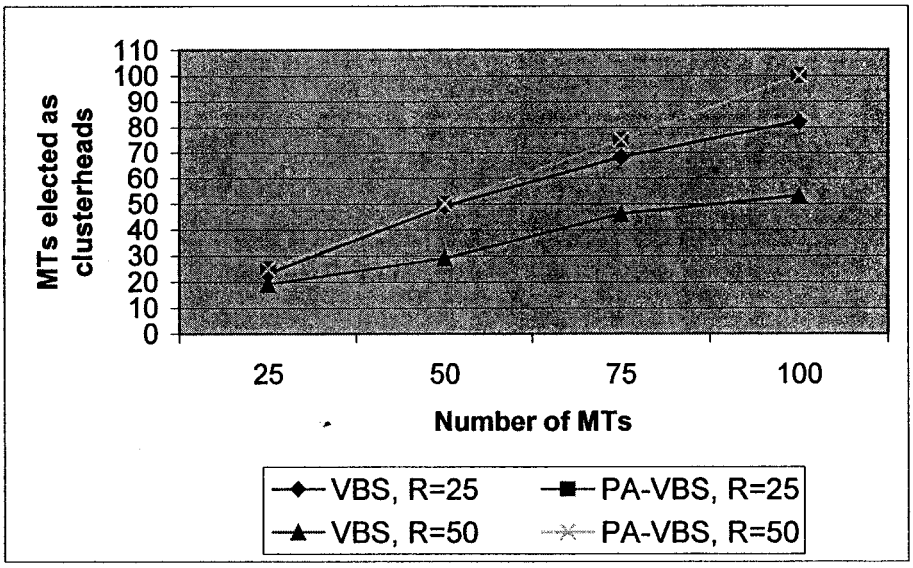


Figure 6.12: Impact of network density on total number of elected clusterheads

In Figure 6.13, the energy consumed by the lowest ID nodes running VBS is considerable as opposed to the energy drained by the rest of the nodes. For example, the MT with an ID equal to 1 consumes more than 3 times its MP when the wireless transmission range is equal to 10 units. However, most of the other MTs consumed less than their MP. The amount of consumed energy in the case of VBS increases when the wireless range is increased. On the other hand, the MTs running PA-VBS consume almost the same amount of energy throughout the simulation regardless of the wireless transmission range. This simulation experiment further proves that PA-VBS achieves load balancing. PA-VBS guarantees that a node has a fair share of becoming a clusterhead, and that its

power is used fairly, as compared to all the other MTs.

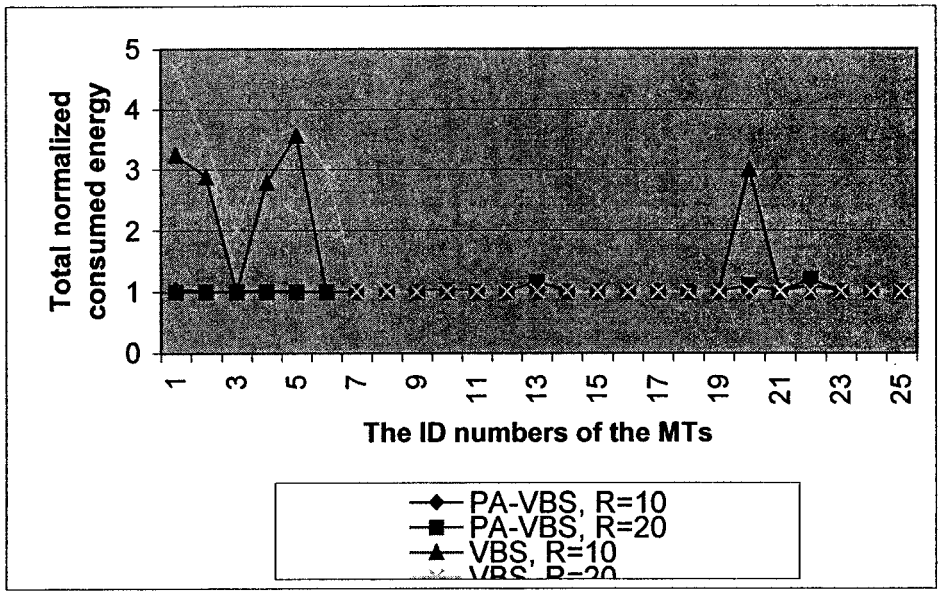


Figure 6.13: Total consumed energy throughout the simulation by each MT

Observing the simulation results of Figure 6.14 shows that PA-VBS achieves load balancing between the wireless nodes, regardless of the routing load carried by the clusterheads. However, the total energy drained from the batteries of the MTs does increase with the increment of the routing load. For example, a 67% increase in the carried routing load results in more than double the amount of consumed battery capacity. The difference between the total energy consumed by any two MTs is never more than 14%. Besides, in most cases, the wireless nodes consumed an equal amount of battery capacity.

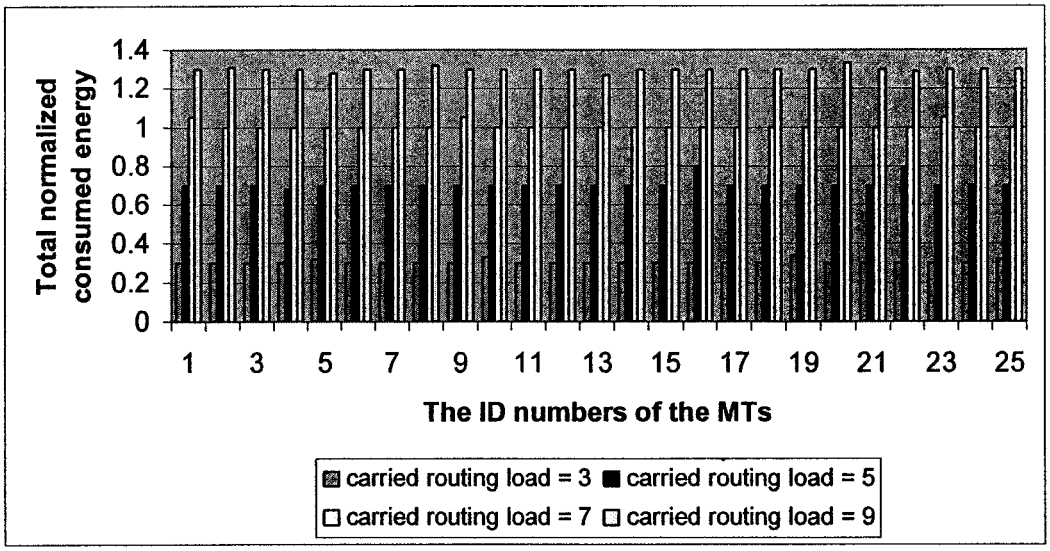


Figure 6.14: The impact of the carried routing load on the total consumed energy

As shown in Figure 6.15, the amount of consumed energy by the wireless nodes is affected by the value assigned to the association threshold. This is attributed to the fact that the nodes initially elected as clusterheads, and whose instantaneous energy values are well over the association threshold, remain as clusterheads for very long periods of time. Consider, for example, a node whose initial normalized energy is equal to 90%. This node may, in the worst case, operate as a clusterhead until it consumes 65% of its battery capacity. When the association threshold is raised to 0.5, a considerable decrease in the consumed energy is noted. Fewer discrepancies in the values of the consumed energy can be seen.

The simulation results of Figure 6.16 show that PA-VBS achieves load balancing between the wireless nodes, regardless of the neighbor activity experienced by the wireless nodes. However, the total energy drained from the batteries of the MTs does increase with the increment of the neighbor activity factor. For example, a 100% increase

in the neighbor activity results in a 75% increase in the amount of consumed battery capacity. The total energy consumed by any two MTs is never remarkably different. In most cases, the wireless nodes consumed an equal amount of battery capacity.

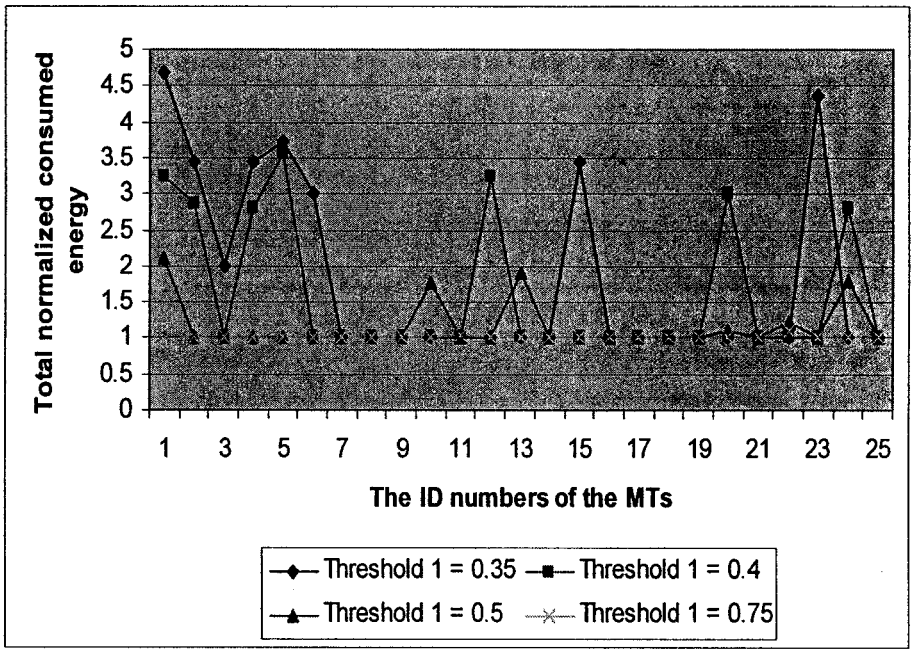


Figure 6.15: The impact of the association threshold on the total consumed energy

In Figure 6.17, the energy consumed by the wireless nodes running PA-VBS with varying charging periods is examined. The amount of drained energy is found to be sometimes considerably different under a long charging period (16 hours). Once nodes are elected as PA-VBSs, some are forced to remain PA-VBSs for longer periods in case their neighbors are being charged.

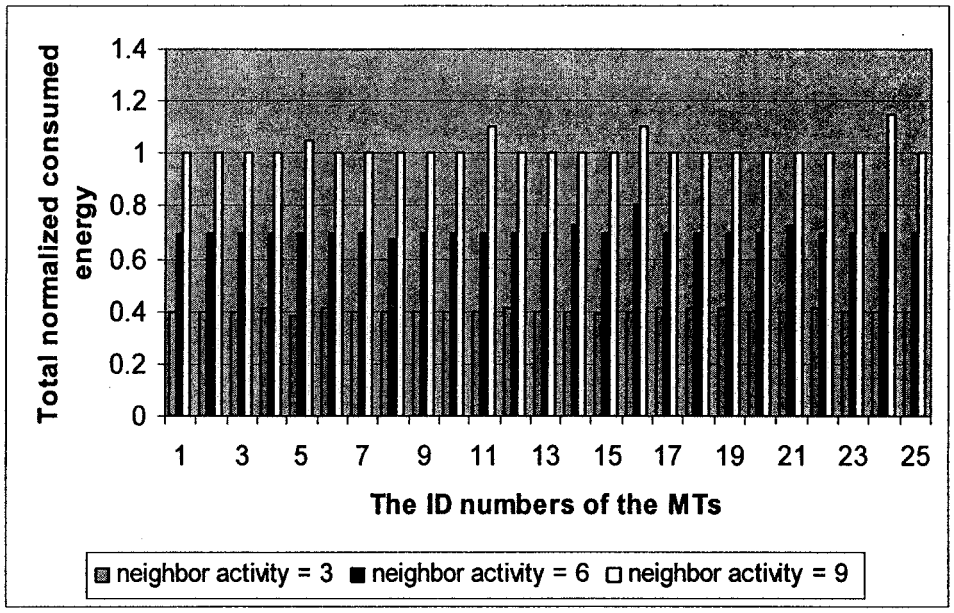


Figure 6.16: The impact of the neighbor activity factor on the consumed energy

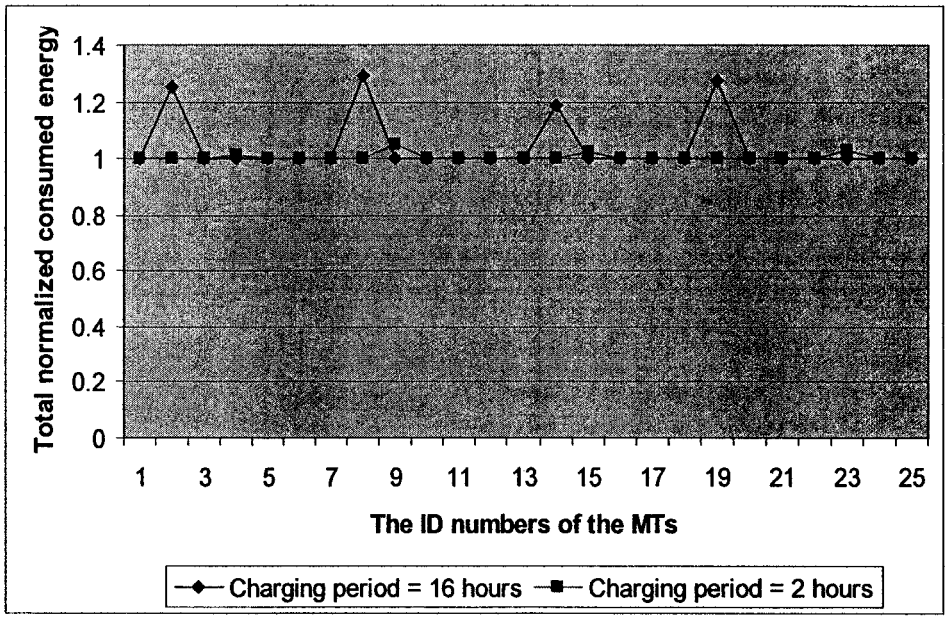


Figure 6.17: The impact of the charging period on the consumed energy

6.5 SUMMARY

In this chapter, we introduced a novel infrastructure-creation scheme for wireless mobile ad hoc networks, namely PA-VBS. The chapter also provided an in-depth explanation of the operation of our proposed PA-VBS infrastructure-creation protocol. Moreover, packet-level simulation experiments of wireless mobile ad hoc networks, with variable node densities, were conducted, and the results were examined. PA-VBS surpasses VBS in balancing the load amongst the nodes of the wireless mobile ad hoc network. Unlike other clustering protocols, PA-VBS allows the mobile nodes to use their valuable battery power fairly. On the contrary to VBS, PA-VBS does not drain all the battery capacity of the clusterheads since clusterheads do not accept any more merge requests when below THRESHOLD_1. Plus, PA-VBS introduces the concept of service denial. Our experiments showed that PA-VBS always elects 100% of the wireless nodes as clusterheads. Hence, it attains fair clustering. All the nodes serve as a clusterhead, at least once, during their operation. Most importantly, PA-VBS keeps the total consumed energy by the MTs almost constant. This guarantees fair energy consumption amongst the wireless mobile nodes.

PA-VBS balances the load amongst the wireless nodes regardless of the carried routing load. Consequently, PA-VBS can be utilized as a basis for routing in ad hoc networks. Moreover, PA-VBS can form the cornerstone for the wise distribution of the network load amongst all the viable paths between a source and destination pair.

CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH

7.1 CONCLUSIONS

In a multi-hop wireless ad hoc or sensor network, wireless mobile stations form a network without the intervention of a dedicated infrastructure. Wireless ad hoc networks eradicate infrastructure deployment, setup, and administration costs. Although energy conservation in wireless ad hoc and sensor networks is of great significance due to the limited battery capacity of their constituent mobile devices, it remains the main impediment towards their deployment.

In this thesis, we have shown the indispensable need for energy conservation in wireless multi-hop ad hoc networks. It is noteworthy that the behavior of any protocol entity and the operations carried out by such entities impact the energy costs. Moreover, separation between the different layers of the OSI model has been a common practice. Nevertheless, it is the unpredictability and unreliability of the underlying wireless medium, the complexity of contention-based medium access, and the dynamic nature of wireless

multi-hop networks that account for our visionary anticipation of resorting to cross-layer interactions between the network and lower layers to achieve energy efficiency in multi-hop networks.

Conventionally, energy efficiency was addressed in isolation of performance-only metrics. This led to an imminent tradeoff between energy conservation and traditional networking performance. Consequently, the network becomes vulnerable to packet losses and throughput is compromised in the case of an energy-only metric. On the other hand, unfair battery utilization and network partitioning result in the case of a performance-only metric. Thus, initially, we conjectured the requisite development of new joint energy-based and performance-preserving metrics and schemes for wireless multi-hop networks.

Schemes that incorporate the wireless nodes instantaneous battery capacities with load, contention, and PHY information, along with others, are more prone to achieve our ultimate goal of developing reliable, fast, and long-lived wireless multi-hop ad hoc and sensor networks. On the face of it, this adds to the analytical complexity of such networks which are already far more intricate than their cellular counterparts as we have shown. The challenge is, nonetheless, to invent computationally affordable algorithms and metrics. This thesis has been clearly developed in previous chapters. In spite of that, our schemes must only use information that is readily available to the network layer through route setup and route maintenance. It is intuitive that our energy-efficient algorithms should not introduce new overheads to which a portion of the consumed power becomes attributed, but rather rely on existing ones. Hence, we guarantee that our combined

energy and performance schemes, as well as future ones that utilize them, do not render the power consumption modeling of wireless multi-hop networks infeasible. This not only applies to network layer signaling but also signaling between adjacent OSI layers.

Our initial premises were then strongly supported by the energy-based performance study presented in Chapter 3. Previous studies have modeled energy consumption theoretically rather than empirically. Likewise, they relied, on the whole, on basic CSMA/CA-like medium access schemes. These factors, and others, have hindered achieving reliable results, conclusions, and implications for power-aware design of wireless multi-hop networks. On the contrary, our energy-based performance study is based on an empirical power consumption model and a detailed implementation of the IEEE 802.11 physical layer convergence protocol (PLCP) and MAC sublayers.

Our simulations showed that a non-power-aware routing scheme sometimes performs as well, and in other times even better, than the so-called power-aware schemes. Thus, our MAC-based investigations reveal that battery capacity cannot be used solely to satisfy our requirements for a power-efficient routing scheme. In fact, a node may be overused only because its remaining battery capacity is larger than others. This will adversely affect traditional performance metrics as well. We have also observed that most of the drained energy is attributed to MAC contention, queuing delays and idle periods and not actual point-to-point data (or broadcast) transmissions and receptions. Instead, power-based fairness and energy conservation can both be attained if valuable MAC and PHY input is passed to the network layer above so as to avoid high-contention zones and

penetrate low-contention ones. These conclusions highly favor hierarchical wireless ad hoc and sensor network architectures as a means of acquiring further energy savings. This becomes increasingly realizable in populous networks.

We proposed four new schemes that utilized the previous observations and conclusions. The first two, namely, Energy-Constrained Path Selection (ECPS) and Energy-Efficient Load Assignment (E2LA), achieve energy conservation via probabilistic dynamic programming and cross-layer designs. Unlike ECPS, E2LA was designed to extend existing power-aware single-path routing protocols to simultaneously use multiple routes through the optimal distribution of the network load. We developed four distinct reward schemes to be used in conjunction with E2LA. In addition, we illustrated the characteristics of E2LA in a number of experiments and studied its energy-conserving potential. It was shown that E2LA shuns MAC contention, distributes the load in such a way that the chances of delivery success are maximized and thus saves energy. Nevertheless, there is no provision of a deterministically guaranteed, or at least quasi-guaranteed, minimum system lifetime.

Moreover, we devised a novel scheme that would allow transferring data from the source to the destination in a wireless ad hoc network using a set of energy-aware routes while maintaining a pre-computed minimum system lifetime. The proposed framework is computationally inexpensive and is based on the remaining battery capacities, a set of energy thresholds and the expected MAC contenders at the wireless nodes. Our framework was shown to route around congestion and allocates more packets to less

congested routes. While the framework here accommodates fairly sophisticated power consumption models, other practical concerns, chief amongst them the energy expended in route maintenance and medium access, remain to be incorporated and are the target of our future research.

For the sake of complementing our energy-efficient framework, we devised a novel infrastructure-creation scheme for wireless mobile ad hoc networks, namely Power-Aware Virtual Base Stations (PA-VBS). It surpasses prominent schemes in balancing the load amongst the nodes of the wireless mobile ad hoc network. Unlike other clustering protocols however, PA-VBS allows the mobile nodes to use their valuable battery power fairly. PA-VBS will be utilized in our future research as the basis for developing cellular-like somewhat centralized power management algorithms.

In conclusion, and in view of the foregoing, we had an acumen and discernment to advance the state of the art of wireless multi-hop ad hoc and sensor networks. The outcomes of our research will enable the deployment of sustainable wireless ad hoc networks. Likewise, the optimization of the energy consumption of the network layer and MAC sublayer creates a battery power reservoir for applications. Hence, existing Internet applications may be ported, with the merest modifications, to a power-limited multi-hop wireless environment. This not only rids us of the expenditures required for new software but also of the development and testing time and labor. Such expenditures and labor may encumber the opportune development of a reliable and energy-efficient wireless multi-hop network, and hence hamper its commercial realization.

7.2 FUTURE RESEARCH

In our future work, we will incorporate the overheads associated with routing and medium access into our system lifetime framework. Efficient mechanisms to estimate the resulting energy overheads are yet to be invented. Enhancing our framework through the addition of conventional performance-based constraints, and hence providing QoS guarantees, is among our current research activities. The challenge therein is to maintain the computational efficiency of the scheme.

We will also devise tractable (polynomial-time) approximation algorithms for solving hard (NP-complete) optimization problems pertaining to QoS and energy-aware dynamic wireless ad hoc and sensor network infrastructures. The approximation algorithms should solve the corresponding infrastructure problem with a reasonable relative error. In this respect, we will investigate the possibility of utilizing classical algorithm design methods, such as greedy algorithms, dynamic programming, etc.

The first generation of wireless multi-hop ad hoc infrastructures will be used solely for autonomous networking and will thus be deployed in home and local area networks. Prior to providing support for wide area networking via the second generation of wireless ad hoc infrastructures, optimization problems involving Virtual Base Stations (VBSs) of radiuses larger than one, need to be studied. In the meantime, it is not at all clear how control information relevant to the formation of the infrastructure would be exchanged between non-neighboring stations without considerable overhead.

The development of industrial-grade test-beds for multi-hop ad hoc and sensor networks is another area of ongoing and future research. The designed test-beds ought to be not only energy-efficient, but also QoS-preserving. We will then deploy the test-beds in a controlled environment which would enable us to carefully monitor and analyze its performance.

Finally, we strongly infer that despite the complexities and challenges ad hoc networking and communications are undoubtedly the future of wireless. Consequently, over the next few years we will witness a proliferation in the integration of heterogeneous wireless systems, including multi-hop networks. Discovering novel means for resource reservation over multiple hops, along with efficient multi-channel support, will facilitate their 4G+ deployment.

BIBLIOGRAPHY

[1] E. Royer and C. Toh, "A review of current routing protocols for ad-hoc mobile wireless networks," *IEEE Personal Communications Magazine*, April 1999, pp. 46-55.

[2] I. Chlamtac, A. Farago, A.D. Myers, V.R. Syrotiuk, and G. Zaruba, "ADAPT: A Dynamically Self-Adjusting Media Access Control Protocol for Ad Hoc Networks," *Proceedings of GLOBECOM 1999*, pp. 11-15.

[3] A. Bar-Noy and I. Kessler, "Tracking mobile users in wireless communication networks," *IEEE Transactions on Information Theory*, vol. 39, no. 6, pp. 1877-1886.

[4] A. Bar-Noy, I. Kessler, and M. Sidi, "Mobile users: to update or not to update? ," *ACM-Baltzer J. Wireless networks*, vol. 1, no. 2, pp. 175-186, July 1994.

[5] C. Huitema, "Routing in the Internet," Englewood Cliffs, NJ Prentice-Hall, 1995.

[6] S. Rajagopalan and B. Badrinath, "An adaptive location management strategy for mobile IP," *ACM Mobicom 1995*, pp. 170-180, Nov. 1995.

[7] Network Working Group, RFC 2002: "IP mobility support," Category: Standards Track, Editor: C. Perkins, Oct 1996.

[8] A. Hac and B. Liu, "Database and location management schemes for mobile communications," *IEEE/ACM Transactions on Networking*, vol. 6, no. 6, pp.851-865.

[9] I. Akyildiz, S. Ho and Y. Lin, "Movement-based location update and selective paging for PCS networks," *IEEE/ACM Transactions on Networking*, vol. 4, pp. 629-638.

[10] I. Akyildiz and J. Ho, "Dynamic mobile user location update for wireless PCS networks," *Wireless Networks*, (1995), pp. 187-196.

[11] T. Imielinski and B. Badrinath, "Querying locations in wireless environments," *Wireless Communications: Future Directions*, J. M. Holtzman and D. J.

Goodman, Eds. Boston, MA: Kluwer, 1993, pp. 85-108.

[12] C. Lin and M. Gerla, "A Distributed Architecture for Multimedia in Dynamic Wireless Networks," *Proceedings of IEEE GLOBECOM 1995*, pp. 1468-1472.

[13] I. Chalmtac and A. Farago, "A new approach to the design and analysis of peer-to-peer mobile networks," *ACM-Baltzer J. Wireless Networks*, pp. 149-156, 1999.

[14] A. Amis, R. Prakash, T. Vuong and D. Huynh, "Max-min d-cluster formation in wireless ad-hoc networks," *IEEE Infocom 2000*, pp. 32-41.

[15] T. Chenand M. Gerla, "Global State Routing: a new routing scheme for ad hoc wireless networks," *Proceedings of IEEE ICC'98*, pp. 171-175.

[16] C. Perkins and P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers," *Computer Communications Review*, pp. 234-244, October 1994.

[17] Y. Ko and N. Vaidya, "Location-Aided Routing (LAR) in mobile ad hoc networks," *ACM/IEEE Int. Conf. On Mobile Computing and Networking (MobiCom'98)*, pp. 66-75, October 1998.

[18] D. Johnson and D. Maltz, "Dynamic source routing in ad hoc wireless networks," *Mobile Computing*, Kluwer Academic Publishers, 1996.

[19] C. Perkins and E. Royer, "Ad-hoc on demand distance vector routing," *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, Feb 1999, pp. 90-100.

[20] Z. Haas, "A New Routing Protocol for the Reconfigurable Wireless Networks," *Proceedings of IEEE ICUPC'97*, San Diego, CA, Oct. 1997, pp. 562-66.

[21] Y. Ko and N. Vaidya, "Geocasting in mobile ad hoc networks: location-based multicast algorithms," *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, Feb 1999, pp. 101-110.

[22] Iowa State University GPS page, *http://www.cnde.iastate.edu/gps.html.*

[23] Trimble GPS Tutorial, *http://www.trimble.cm/gps/fsections/aa_f1.htm.*

[24] D. Johnson, "Routing in ad hoc networks of mobile hosts," *Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications*, pp. 158-163, Dec 1994.

[25] S. Das, C. Perkins and E. Royer, "Performance comparison of two on-demand routing protocols for ad-hoc networks," *IEEE Infocom 2000*, vol. 1, pp. 3-12.

[26] IETF MANET Working Group Charter, *http://www.ietf.org/html.charters/manet-charter.html.*

[27] C. Hedrick, "Routing information protocol," *RFC 1058*, June 1988.

[28] S. Corson and J. Macker, "Mobile Ad hoc NETworking (MANET): "routing protocol performance issues and evaluation considerations," *RFC 2501*, January 1999.

[29] S. Lee, M. Gerla and C. Toh, "A simulation study of table-driven and on-demand routing protocols for mobile ad hoc networks," *IEEE Network Magazine*, July/August 1999, Vol. 13, No. 4, pp. 48-54.

[30] ATM Forum, "Traffic management specification, Version 4.0," April 1996.

[31] ETSI TC-RES, Radio Equipment and Systems (RES); HIgh PErformance Radio Local Area Network (HIPERLAN); Functional Specification. ETSI, 06921 Sophia Antipolis Cedex -- France, July 1995. Draft prETS 300 652.

[32] European Telecommunications Standards Institute, ETSI HIPERLAN/1 standard, *http://www.etsi.org/technicalactiv/hiperlan1.htm*, October 2000.

[33] B. Bing, "High-speed wireless ATM and LANs," Artech House Publishers, 2000.

[34] A. Safwat and H. Hassanein, "Infrastructure-based Routing in Wireless Mobile Ad hoc Networks," *The Journal of Computer Communications*, Vol. 25, No. 3, 2002, pp. 210-224.

[35] J. Cano and P. Manzoni, "Evaluating the energy-consumption reduction in a MANET by dynamically switching off network interfaces," *IEEE ISCC*, pp. 186-191, July 2001.

[36] Z. Haas and J. Deng, "Dual Busy Tone Multiple Access (DBTMA) - Performance Evaluation," *IEEE VTC'99*, vol. 1, pp. 314-319, May 1999.

[37] Z. Haas and J. Deng, "Dual Busy Tone Multiple Access (DBTMA): a New Medium Access Control for Packet Radio Networks," *IEEE International Conference on Universal Personal Communications*, vol. 2, pp. 973-977, October 1998.

[38] Y. Ko and N. Vaidya, "Geocasting in mobile ad hoc networks: Location-based multicast algorithms," *Technical Report TR-98-018*, Texas A&M University, September 1998.

[39] Y. Ko and N. Vaidya, "Anycasting and geocasting in mobile ad hoc networks," *Technical Report 00-015*, Dept. of Computer Science, Texas A&M University, June 2000.

[40] A. Safwat and H. Hassanein, "Structured routing in wireless mobile ad hoc networks," *Proceedings of the 6th IEEE Symposium on Computers and Communications (ISCC) 2001*, pp. 332-337, July 2001.

[41] T. Sheltami and H. Mouftah, "Efficient cluster-based routing protocol for wireless ad hoc networks," *IEEE International Symposium on Advances in Wireless Communications (ISWC) 2002.*

[42] T. Sheltami, and H. Mouftah, "A comparative study of two clustering protocols," *Proceedings of IASTED International Conference on Wireless and Optical Communications*, pp. 251-256, June 2001.

[43] I. Akyildiz, J. McNair, L. Carrasco, R. Puigjaner and Y. Yesha, "Medium access control protocols for multimedia traffic in wireless networks," *IEEE Network*, Vol. 13, No. 4, July/August. 1999, pp. 39-47.

[44] D. Raychaudhri, et al, "WATMnet: a prototype wireless ATM system for multimedia personal communication," *IEEE JSAC*, Vol. 15, Nov. 1, Jan. 1997, pp. 83-95.

[45] J. Hubaux, "Terminodes : Toward self-organized mobile networks," *Technical Report No. SSC/1999/022*, Swiss Federal Institute of Technology, June 1999.

[46] J. Hubaux, J.Le Boudec, M. Vojnovi'c, S. Giordano, M. Hamdi, L. Blazevi'c, and L. Butty'an, "Toward mobile ad hoc WANs: Terminodes," *Technical Report No. DSC/2000/006*, Swiss Federal Institute of Technology, January 2000.

[47] IEEE LAN/MAN standards committee, "Wireless LAN medium access protocol (MAC) and physical layer (PHY) specifications," *IEEE Standard 802.11*, 1999.

[48] P. Karn, "MACA--a new channel access method for packet radio," *Proceedings of the 9th ARRL/CRRL Amateur Radio Computer Networking Conference*, pp. 134-140, 1990.

[49] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A Media Access Protocol for Wireless LAN's," *ACM SIGCOMM Computer Communication Review*, vol. 24, no. 4, pp. 212-225, 1994.

[50] C. Wu and V. Li, "Receiver-Initiated Busy-Tone Multiple Access in Packet Radio Networks," *Proceedings of ACM SIGCOMM Workshop*, vol. 17(5), pp. 336-342, 1987.

[51] M. Stemm and R. Katz, "Measuring and reducing energy consumption of network interfaces in handheld devices," *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Science*, Special Issue on Mobile Computing, 80(8):1125-1131, August 1997.

[52] O. Kasten, "Energy consumption," Swiss Federal Institute of Technology, *http://www.inf.ethz.ch/~kasten/research/bathtub/energy_consumption.html*, 2001.

[53] H. Hassanein and A. Safwat, "Virtual Base Station Protocol for Wireless Mobile Ad hoc Networks," *Proceedings of the International Conference on Computer Theory and Applications*, Alexandria, Egypt, Aug 2002.

[54] S. Singh and C. Raghavendra, "Power-efficient MAC protocol for multihop radio networks," *Proceedings of IEEE PIRMC*, vol. 1, pp. 153-157, September 1998.

[55] F. Tobagi and L. Kleinrock, "Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in CSMA and Busy-Tone Solution," *IEEE Trans. on Communications*, COM-23, pp. 1417-1433, December 1975.

[56] F. Talucci, M. Gerla and L. Fratta, "MACABI (MACA By-Invitation): A Receiver Oriented Access Protocol for Wireless Multiple Networks," *Proceedings of PIMRC*, pp. 1-4, September 1997.

[57] Z. Haas and J. Deng, "A collision-free medium access control scheme for ad hoc networks," *Proceedings of WCNC'99*, September 1999.

[58] R. Dube, C. Rais, K. Wang, and S. Tripathi, "Signal stability based adaptive routing (SSA) for ad hoc mobile networks," *IEEE Personal Communications*, pp. 36-45, Feb. 1997.

[59] C. Chiang, H. Wu, W. Liu, and M. Gerla, "Routing in clustered multihop, mobile wireless networks with fading channel," *Proceedings of the IEEE Singapore International Conference on Networks*, 1997, pp. 197--211.

[60] H. Hassanein and A. Safwat, "Virtual base stations for wireless mobile ad hoc communications: an infrastructure for the infrastructure-less," *The International Journal of Communication Systems*, Vol. 14, 2002, pp. 763-782.

[61] J. Navas and T. Imielinski, "Geocast - geographic addressing and routing," *Proceedings of ACM/IEEE Int. Conf. on Mobile Computing and Networking (MobiCom'97)*, September 1997.

[62] Y. Ko and N. Vaidya, "Geocasting in mobile ad hoc networks: Location-based multicast algorithms," *Proceedings of IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*, pp. 101-110, February 1999.

[63] "NAVSTAR GPS operations," *http://tychi.usno.navy.mil/gpsinfo.html.*

[64] Y. Ko and N. Vaidya, "GeoTORA: A protocol for geocasting in mobile ad hoc networks," *Technical Report 00-010*, Dept. of Computer Science, Texas A&M University, March 2000.

[65] V. Park and M. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," *IEEE Infocom*, vol. 3, pp. 1405-1413, 1997.

[66] E. Gafni, D. Bertsekas, "Distributed Algorithms for Generating Loop-free Routes with Frequently Changing Topology," *IEEE Trans. on Communications*, Vol. COM-29, No. 1, pp. 11-18, January, 1981.

[67] S. Deering, and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," *RFC 2460*, December 1998.

[68] C. Partridge, T. Mendez, and W. Milliken, "Host Anycasting Service," *RFC 1546*, November 1993.

[69] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad hoc routing," *Proceedings of the Seventh Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 70-84, July 2001.

[70] G. Pottie and W. Kaiser, "Embedding the internet: wireless integrated network sensors," *Communications of the ACM*, 43(5):51-58, May 2000.

[71] Y. Xu, J. Heidemann, and D. Estrin, "Adaptive energy-conserving routing for multihop ad hoc networks," *Technical Report TR-2000-527*, USC/Information Sciences Institute, October 2000, ftp://ftp.isi.edu/isi~pubs/tr-527.pdf

[72] F. Bennet et al., "Piconet: embedded mobile networking," *IEEE Personal Communications*, 4(5): 8-15, October 1997.

[73] K. Scott and N. Bambos, "Routing and channel assignment for low power transmission in PCS," *Proc. ICUPC '96*, vol. 2, pp. 498-502, October 1996.

[74] S. Singh, M. Woo, and C. Raghavendra, "Power-aware routing in mobile ad hoc networks," *Proc. Mobicom '98*, pp. 181-190, October 1998.

[75] C. Toh, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks," *IEEE Communications Magazine*, pp. 138-147, June 2001.

[76] S. Lee and M. Gerla, "Dynamic load-aware routing in ad hoc networks," *ICC '01*, vol. 10, pp. 3206-3210.

[77] S. Lee and M. Gerla, "Dynamic load-aware routing in ad hoc networks," *Technical Report*, Computer Science Department, University of California, Los Angeles, August 2000.

[78] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, "Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks," *Proc. Mobicom '01*, pp. 85-96.

[79] S. Lee and M. Gerla, "AODV-BR: Backup Routing in Ad hoc Networks," *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, September 2000.

[80] I. Stojmenovic, "DFS based routing with guaranteed delivery in wireless networks," *The IEEE 29th International Conference on Parallel Processing*, pp. 461-466.

[81] J. Gomez, A. Campbell, M. Naghshineh and C. Bisdikian, "PARO: Conserving Transmission Power in Wireless ad hoc Networks," *IEEE 9th International Conference on Network Protocols (ICNP'01)*, Riverside, California. November 2001.

[82] N. Gupta and S. Das, "Energy-Aware On-Demand Routing for Mobile Ad Hoc Networks," *Proceedings of IWDC 2002*, pp. 164-173.

[83] W. Yu and J. Lee, "DSR-Based Energy-Aware Routing Protocols in Ad Hoc Networks," *Proceedings of ICWN 2002.*

[84] W. Zhao and K. Ramamritham, "Virtual Time CSMA Protocols for Hard Real-Time Communications," *IEEE Transactions on Software Engineering*, Vol. SE-13, No. 8, pp. 938-952, August 1987.

[85] W. Zhao, J. Stankovic, and K. Ramamritham, "A Window Protocol for Transmission of Time Constrained Messages," *IEEE Transactions on Computers,* Vol.39, No. 9, pp. 1186-1203, Sept 1990.

[86] NS-2 Manual, *http://www.isi.edu/nsnam/ns*.

[87] A. Bhatnagar and J. Teifel, "An Energy-Performance Metric for Mobile Ad hoc Networks," *Internet Article.*

[88] A. Martin, "Towards an Energy Complexity of Computation," *Information Processing Letters*, V77, pp 181-187, 2001.

[89] A. Martin, A. Lines, R. Manohar, M. Nystr¨om, P. Penzes, R. Southworth, U. Cummings, and T. Lee, "The Design of an Asynchronous MIPS R3000

Microprocessor," *Proceedings of the 17th Conference on Advanced Research in VLSI*, September 1997.

[90] M. Pearlman, J. Deng, B. Liang, and Z. Haas, "Elective Participation in Ad hoc Networks Based on Energy Consumption," *Proceedings of IEEE Globecom 2002.*

[91] V. Rodoplu and T. Meng, "Minimum energy mobile wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 1, pp. 22-31, 2000.

[92] D. Baker and A. Ephremides, "The architectural organization of a mobile radio network via a distributed algorithm," *IEEE Transactions on Communications*, vol. COM-29, no. 11, pp. 56-73, Jan. 1981.

[93] A. Ephremides, J. Wiselthier, and D. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling," *Proceedings of the IEEE*, vol. 75, no. 1, pp. 56-73, Jan. 1987.

[94] M. Ettus, "System capacity, latency, and power consumption in multihop-routed SS-CDMA wireless networks," *Proceedings of IEEE Radio and Wireless Conference (RAWCON) 98*, pp. 55-58.

[95] T. Meng and V. Rodoplu, "Distributed network protocols for wireless communication," *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems (ISCAS'98)*, vol. 4, pp. 600-603.

[96] T. Shepard, "Decentralized channel management in scalable multihop spread spectrum packet radio networks," *Tech. Rep. MIT/LCS/TR-670*, July 1995.

[97] M. Bhardwaj and A. Chandrakasan, "Bounding the lifetime of sensor networks via optimal role assignments," *Proceedings of IEEE Infocom 2002*, vol. 3, pp. 1587-1596.

[98] A. Safwat, H. Hassanein, and H. Mouftah, "Energy-Aware Routing in MANETs: Analysis and Enhancements," *Proceedings of The Fifth ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2002),* Atlanta, Georgia, USA, pp. 46-53, Sep 2002.

[99] A. Safwat, H. Hassanein, and H. Mouftah, "A MAC-based Performance Study of Energy-Aware Routing Schemes in Wireless Ad hoc Networks," *Proceedings of IEEE Globecom 2002,* Taipei, Taiwan, Nov 2002.

[100] A. Safwat, H. Hassanein, and H. Mouftah, "Power-Aware Wireless Mobile Ad hoc Networks," *Handbook of Ad hoc Wireless Networks,* Chapter 22, CRC Press, 2002.

[101] A. Safwat, H. Hassanein, and H. Mouftah, "Structured Proactive and Reactive Routing for Wireless Mobile Ad hoc Networks," *Handbook of Ad hoc Wireless Networks,* Chapter 13, CRC Press, 2002.

[102] D. Raychaudhri, et al, "WATMnet: a prototype wireless ATM system for multimedia personal communication," *IEEE JSAC*, Vol. 15, Nov. 1, Jan. 1997, pp. 83-95.

[103] N. Passas, et al, "Quality of service-oriented medium access control for wireless ATM networks," *IEEE Communications Magazine*, Vol. 35, No. 11, Nov. 1997, pp. 42-50.

[104] R. Pichna and Q. Wang, "A medium access control protocol for cellular packet CDMA carrying multirate traffic," *IEEE JSAC*, Vol. 14, Dec. 1996, pp. 1728-1736.

[105] B. Bing, "Wireless local area networks: the new wireless revolution," *John Wiley and Sons*, June 2002.

[106] M. Gast, "802.11 wireless networks: the definitive guide," *O'Reilly and Associates*, April 2002.

[107] L. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment," *Proceedings of IEEE Infocom 2001*, vol. 3, pp. 1548-1557.