Towards Efficient Power Consumption in Ad Hoc Networks

Iordan Grigoriev Ignatov
BScEng
University of Natal, South Africa

Submitted in fulfillment of the academic requirements
for the degree of Master of Science in Engineering
in the School of Electrical, Electronic and Computer Engineering
at the University of KwaZulu-Natal, Durban, South Africa

January 31, 2006

To Cheryl

This document was created in LaTeX

As the candidate's supervisor I have approved this dissertation for submission.

Signed: ____________________

Name: Mr. Stephen A. McDonald

Date: January 31, 2006

Abstract

An ad hoc network is a collection of an arbitrary number of mobile nodes which communicate with each other in a multi-hop wireless fashion. The operation of this type of network does not depend on a pre-established infrastructure and thus is easily deployable. Ad hoc networks operate without the need for central control units which eliminates the *single point of vulnerability* situation making them suitable for military applications. As network management is a responsibility of each participating node, it is done in a distributed way which is the origin of some of the greatest challenges faced by ad hoc network developers.

In the early days of ad hoc network development additional constraints were posed by factors such as limited *communication bandwidth*, *processing power* and *battery capacity.* The first two factors have since evolved significantly. Improving battery capacity and endurance has also been at the forefront of research efforts, however, the achieved advances do not yet meet the battery endurance requirements for ad hoc networking. The lifetime of an ad hoc network is dependent on the battery depletion rate by each participating node. Battery replacement in the field is often impractical or even impossible. To that end, power consumption optimisation is one of the top requirements in the development of ad hoc networks.

This dissertation outlines some of the most popular power optimisation techniques used with ad hoc networks. It also presents an analytical model which reveals the suboptimal efficiency of these methods as they promote power consumption optimisation solely through control of transmission power. The model leads to the conclusion that optimal power efficiency is achieved only if factors such as efficient channel use arbitration, re-

duced control and communication overhead and energy consumption during idle periods are considered. Due to a lack of protocols that address these factors, a new method is presented in the form of a power efficient Medium Access Control protocol. The novel scheme is based on a TDMA approach for structured channel use. It incorporates the idea of contention for channel use in order to offset some of the drawbacks of pure TDMA. For reduced control overhead the scheme makes use of positional information. Energy waste during idle states of the RF transceivers is minimised with extensive use of low-power mode.

The performance of the novel scheme is firstly studied through analytical modelling. With the help of mathematical derivations, characteristics such as *energy efficiency*, *congestion*, *throughput* and *channel access delay* are established. These analytical findings are later verified through simulations of the protocol. In addition to that its performance is compared with that of the popular IEEE 802.11 MAC. It was found that the proposed scheme provides 40 to 60 percent improved energy efficiency in active mode over that of IEEE 802.11 MAC. That together with the reduction of idle state energy consumption though use of low power mode, resulted in a network lifetime extension of 55 percent.

Preface

The research work presented in this dissertation was performed by Iordan Grigoriev Ignatov, under the supervision of Mr Stephen A. McDonald. in the School of Electrical, Electronic and Computer Engineering, University of KwaZulu-Natal. This work was supported by ARMSCOR, the Armaments Corporation of South Africa.

Publications from this work include:

Iordan G. Ignatov and S. McDonald, Position Aided STDMA MAC Protocol for Mobile Ad Hoc Netwroks, *in Proc. Military Information and Communications Symposium of South Africa (MICSSA'05), 2005.*

Iordan G. Ignatov and S. McDonald, Power Use Optimisation in Ad Hoc Netwroks, *in Proc. Southern African Telecommunication Networks and Application Conference (SATNAC'05), September, 6-8, 2005.*

The entire dissertation, unless otherwise indicated, is the student's own original work and has not been submitted in part, or in whole, to any other university for degree purposes.

Contents

List of Figures

List of Tables

List of Abbreviations

ABAM	Associativity-Based Ad hoc Multicast
ACK	Acknowledge
AMRIS	Ad hoc Multicast Routing protocol utilising Increase id-numberS
AMRoute	Ad hoc Multicast Routing
AOA	Angle of Arrival
AODV	Ad hoc On-demand Distance Vector
ARP	Address Resolution Protocol
BER	Bit Error Rate
BIP	Broadcast Incremental Power
BSD	Berkeley Software Distribution
CAMP	Core Assisted Mesh Protocol
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
$CSMA\backslash CA$	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
CW	Contention Window
DIFS	Distributed Inter-Frame Space

DMIP3S	Distributed Multicast Incremental Power with Potential Power Savings
DSDV	Destination Sequence Distance Vector
DSR	Dynamic Source Routing
DSSS	Direct Sequence Spread Spectrum
DVMRP	Distance Vector Multicast Routing Protocol
EADSR	Energy Aware DSR
EER	End-to-End Retransmission
EWMA	Embedded Wireless Multicast Advantage
GPS	Global Positioning System
HHR	Hop-by-Hop Retransmission
IF	Intermediate Frequency
IP3S	Incremental Power with Potential Power Savings
LARP	Localised Power-Aware Routing Protocol
LBM	Location-Based Multicast
LEAR-AODV	Local Energy-Aware AODV
LPR-AODV	Life Prediction Routing AODV
MAC	Medium Access Control
MANET	Mobile Ad Hoc Network
MAODV	Multicast Ad hoc On-demand Distance Vector
MIP3S	Multicast Incremental Power with Potential Power Savings
MST	Minimum Spanning Tree
NAV	Network Allocation Vector

ODMRP	On-Demand Multicast Routing Protocol
OSI	Open System Interconnect
PAR-AODV	Power-Aware Routing AODV
PA-STDMA	Position Aided Spatial-TDMA
PDR	Packet Delivery Ratio
RF	Radio Frequency
RRP	Route Request Packet
RTS	Ready To Send
SIFS	Short Inter-Frame Space
SINR	Signal to Interference plus Noise Ratio
SIR	Signal to Interference Ratio
SNR	Signal to Noise Ratio
SPF	Shortest Path First
S-REMiT	Refined energy-based multicast tree
TDMA	Time Division Multiple Access
TORA	Temporary Ordered Routing Algorithm
WBA	Wireless Broadband Advantage
WM	Weight Arithmetic Mean
WNIA	Wireless Network Interface Adapter

Chapter 1

Introduction

The field of *Communications* has grown significantly over the past twenty years as a result of major technological advances. It has reached a point where one of its subfields - *mobile communications* has become an irreplaceable part of everyday life in the form of the *cellphone.* A topic under the same subfield is *Mobile Ad Hoc Networks*(MANETs). Originating from *Packet Radio Networks*, which were in development more than twenty years ago for military applications, the concept of mobile ad hoc networking has shown potential to revolutionise contemporary wireless communications.

A mobile ad hoc network is a collection of an arbitrary number of mobile wireless nodes[1] that forms spontaneously. The spontaneous formation of ad hoc networks is a product of their distinctive characteristic i.e. network operation that does not require a pre-established infrastructure (e.g. base station, access point, router) for the purpose of network coordination. To achieve this, ad hoc networks are managed solely by the nodes that are involved. Such management is done in a fully distributed fashion whereby each node contributes to the operation of the network as a whole.

The self-organised operation of MANETs brings about their *ease of deployment* characteristic which makes them suitable for a large number of applications, including surveillance

[1] a "node" is a platform that has computing capabilities, performs routing and makes use of a wireless transceiver

and search and rescue. The absence of infrastructure and the ability of the network to dynamically determine data routes renders ad hoc networks less vulnerable in military environments. The ad hoc concept forms part of mesh networking which is currently under development to provide broadband communication in urban areas.

Another distinguishing characteristic of MANETs is *multi-hop* communication. Often, in ad hoc networks, the communicating pair of nodes is out of each other's communication range. Thus a forwarding process is needed with the help of intermediate nodes which lasts until the data packet reaches the destination node. Subject to the distance between the source-destination pair of nodes and the node density of the network, the data packet could be forwarded by a number of nodes referred to as *hops* and hence multi-hop data transmission is said to take place.

Developmental challenges are posed by the mobile nature of the ad hoc network nodes. In general, the direction and velocity of the nodes is unrestricted. That, together with the error prone wireless communication, leads to constantly changing network topology and an unpredictable level of network connectivity. In addition, there is a requirement that routing and medium access control (MAC) is performed by distributed algorithms. Thus the focus of research and development in the field has been predominantly in the following areas:

- robust data routing
- efficient medium access control

Routing in ad hoc networks is done on the basis of data forwarding through pre-established data paths. For the purpose of *route discovery*, topology information is exchanged by immediate neighbours while *route maintenance* requires frequent updating of that information. The more rapid the topology changes, the more frequent the updating should be to maintain reliable communication. There are a number of developed and tested techniques that use different heuristics for efficient data routing. A comprehensive overview of some of the existing routing mechanisms is presented in Section 2.3.

Distributed medium access control, which provides efficient bandwidth utilisation, has also been at the forefront of ad hoc network development. Most of the research work to date has been focused on arbitration of a single channel in an ad hoc network environment. An inherent consequence of single channel communication in ad hoc networks is transmission collisions. These collisions are the reason for efficiencies as low as 18% in some of the earliest MAC protocols [9]. Collisions adversely affect important performance characteristics such as throughput, energy efficiency and end-to-end propagation delay. From the early days of ad hoc MAC protocol development there has been a trade-off between contention based and time slotted transmission schemes. While the latter is less prone to collisions and therefore provides better efficiency, the former has a superior end-to-end data propagation delay. A common cause of collisions in ad hoc networks is the *hidden terminal* problem. Due to transmission range limits, not all nodes in the network can hear other transmitting nodes. As a result a node could be simultaneously receiving the transmissions of two or more nodes that are out of each other's range. The most common technique used to mitigate the occurrence of the hidden terminal problem is the Ready-To-Send/Clear-To-Send (RTS-CTS) handshaking procedure (proposed as part of the MACA protocol from [10]) between communicating peers. This procedure helps the nodes reserve the channel for the duration of the transmission through exchange of control messages.

More recent research work has been focused on energy conservation. Efficient energy consumption is essential in ad hoc networks as in most of their applications the mobile nodes are powered by batteries with limited capacity. Battery depletion forces nodes to shut down and as a result compromise the connectivity of the network. Therefore the operation of an ad hoc network is dependant on the lifetime of its nodes which is directly related to the nodal power requirements for data processing, transmission and reception and hence, typically, on battery depletion rate. In the presently perceived applications battery replacement in the field is often highly impractical. This has been the motivation behind research efforts focused on the development of power consumption reduction mechanisms which promote power efficiency through optimisation of the operation of individual Opens System Interconnection (OSI) layers [11] with respect to energy consumption.

Energy use optimisation, in the context of ad hoc networks, is the focus of the research work reported in this thesis. The objective is to investigate and show whether a solution providing an increase in power efficiency, can be achieved at the medium access control level of the network hierarchy.

The thesis is organised as follows: Chapter 2 presents some popular power optimisation techniques applicable to ad hoc networks. Chapter 3 investigates power consumption in ad hoc networks through the development of an analytical model. The analysis focuses on comparison of energy consumption reduction in the cases of efficient medium access control protocol versus that of efficient transmission power control. Chapter 4 is dedicated to the main contribution in this thesis and that is the proposed Position Aided Spatial-TDMA MAC protocol. It gives a full description of the scheme and a comprehensive analytical discussion of its expected performance characteristics. Chapter 5 presents simulation results of the protocol and compares them with those predicted by the analytical model performance in terms of energy efficiency, throughput and channel access delay. Chapter 6 concludes and provides a summary of the contributions of the reported research work.

Chapter 2

Standard Power Optimisation Techniques

2.1 Introduction

This chapter presents a survey of network layer and physical layer power consumption optimisation techniques applicable to ad hoc networks. The techniques considered are divided into *power control algorithms* and *power-aware routing protocols*. A summary of the existing literature that describes and analyses the advantages and disadvantages of the techniques is also included. The chapter is organised as follows: Section 2.2 describes a number of popular techniques which use power control for improved energy efficiency and network connectivity; Section 2.3 summarises popular unicast and multicast power-aware routing techniques.

2.2 Power Control Algorithms

Power control in ad hoc networks is the intelligent selection of an adequate transmit-power level by a node when it communicates with its neighbours. Power control algorithms try to use the lowest possible transmit power level that provides reliable communication between peers. Such algorithms have the following objectives:

- maintain network connectivity and avoid network partitioning;

- provide power efficient operation;

Power control algorithms determine the most appropriate power level with the help of some knowledge about the current status of the communication channel. This knowledge is either provided by an existing channel model or by information from the attenuation of the received signals. Such information is usually represented by metrics such as. signal to noise ratio (SNR), signal to interference ratio (SIR) and bit error rate (BER).

Power control algorithms are very popular within the existing CDMA cellular networks where good power control significantly improves network characteristics such as interference level, capacity, throughput and quality of service [12]. In ad hoc networks, although these improved characteristics are also welcomed, power control is primarily used for the purpose of improved network connectivity and power resource conservation.

What follows is an overview of some of the most prominent power control algorithms applicable to ad hoc networks. The discussions include brief descriptions of the algorithms accompanied by highlights of their advantages and disadvantages.

2.2.1 COMPOW power control algorithm

The COMPOW algorithm from [13] has three objectives:

- maximise network capacity;
- extend battery life by providing low power routes;
- reduce contention at the MAC layer;

The algorithm is based on the consideration that bidirectional links between the nodes of an ad hoc network would improve its performance. According to [13], bidirectional links would result in better channel management by the MAC layer and would decrease the number of packet retries between nodes (due to guaranteed acknowledgments from the receiving nodes). They are easily established when the nodes transmit at the same power level. Furthermore, because common power level does not guarantee common SINR at all receivers in an ad hoc network, the transmit power needs to be kept low. This would keep the transmitting neighbourhoods small and with the help of the spatial dispersion of the nodes in the field, more or less equal SINR could be achieved.

Proofs presented in [13] show that the use of the lowest common transmit power by the nodes of an ad hoc network results in significantly improved network capacity and power efficiency. Therefore the task of the algorithm is to identify and assign for use the lowest common transmit power level.

To identify the lowest common power, the algorithm creates and investigates all of the routing tables resulting from the use of each available transmit power level. Once created, each routing table is compared with the routing table RP_{max} that corresponds to the highest transmit power level P_{max}. If there exists a routing table RP_i, where $P_i < P_{max}$, that presents the same connectivity as the routing table of the highest power level RP_{max}, then that table RP_i is chosen for the operation of the node together with the corresponding power level P_i. To illustrate the process, Figure-2.1 shows a node with its neighbours and its transmission ranges relative to the different power levels.

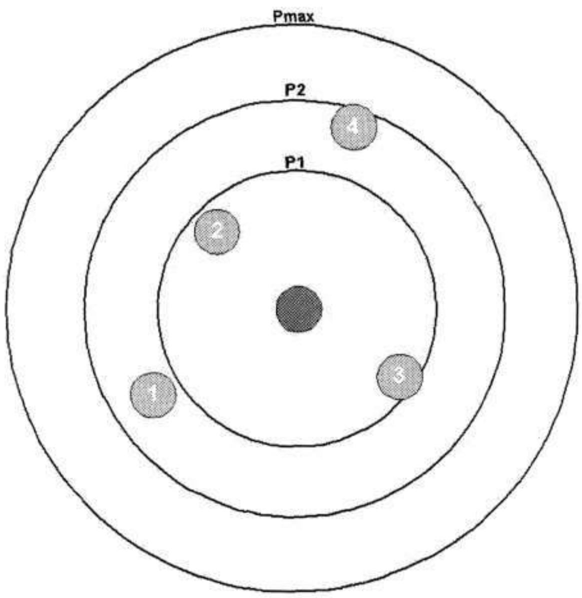


Figure 2.1: Node discovery at different transmit power levels

If the node used transmission power P_{max}, its routing tables would contain available routes to all four neighbouring nodes. The same would be observed if the node used transmission power P_2 ($P_2 < P_{max}$). However, if the node used transmission power P_1, it would have connections only to nodes 2 and 3. Therefore according to the rules of the algorithm, the chosen transmit power level would be P_2 with its corresponding routing table RP_2.

It is interesting to note that the COMPOW algorithm is one of the few power control algorithms implemented on a general platform ad hoc network (Compaq Presario laptops with Linux and CISCO Aironet 350). Unfortunately, no information has been presented in [13] that gives an indication of the achieved efficiency in terms of network capacity and power conservation.

The foreseeable problems with the algorithm are related to issues such as configuration overhead, latency and spatial distribution of the nodes.

An estimation presented in [13] suggests that in a static ad hoc network with an assumed six power levels and six neighbours per node, the resulting configuration overhead would consume approximately 60 kbit/sec of the bandwidth until the setup is complete. Although for the static case such bandwidth is acceptable, if the algorithms were to be implemented in a mobile network, the required bandwidth would be a multiple of the required route updates. Increased number of neighbours and available power levels would also contribute to a larger configuration communication overhead.

The discussion of related implementation issues in [13] indicates that the process of transmit power level shifting introduces significant latencies. This is due to hardware time constraints presented by most of the "off the shelf" wireless adapters.

In a realistic network environment (finite number of nodes distributed non-homogeneously) the convergence of the algorithm depends on the spatial distribution of the nodes. As suggested in [12], the more uneven it is, the higher the chances that the algorithm will converge to a suboptimal power level.

2.2.2 IEEE 802.11 MAC with power control

The IEEE 802.11 MAC [2] wireless communication standard uses two types of messages: broadcast messages and messages for specific hosts within the range of the transmitting radio (unicast). In the later case. before data is exchanged, signalling is used to establish the communication link between the peers and inform the neighbouring nodes of the forthcoming transmission. The signalling process is illustrated by Figure-2.2 while a good overview of the process is presented in [2].

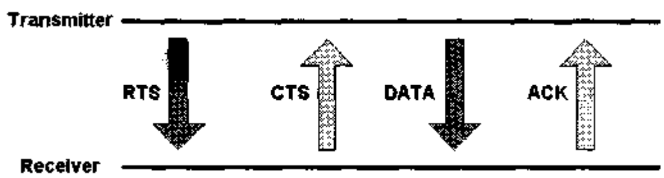


Figure 2.2: IEEE 802.11 MAC signalling process

An algorithm proposed by Agarwal *et al*, [14], uses the signalling process to implement power control. It places signal strength information in all exchanged signalling and data packets. This information serves as feedback according to which the most appropriate transmit power level is chosen. More specifically, the signal strength information represents the ratio of the strength of the received signal to the minimum acceptable signal strength.

When the receiver receives the RTS signal it compares its strength with the pre-set value for the minimum tolerable signal strength. It embeds the result in the CTS signal, which is sent back to the transmitter. As a result the transmitter will correct its power level (if required) and will carry out the same strength estimation for the CTS signal. The result is transmitted to the receiver with the next data packet. Further level adjustments are possible with the help of the following data packets and the corresponding ACK packets. As a result of the process both nodes will inform each other about the incoming signal strengths and both will have a chance to adjust their transmit power levels to an acceptably low value.

As described in [14], the implementation of the algorithm involves the maintenance of small tables, which contain information about the power control settings related to other nodes with which the node has recently been in communication. These tables contain the history of the received signal strength for both successfully transmitted and dropped packets. There is also a mechanism (a count-down timer), which prevents rapid fluctuations in transmit power levels. The simulation of the algorithm investigates two operational cases:

- the communication between two peers starts from the highest power level that is available and settles down to a level just above which packets are dropped - this is referred to as "full blast";
- the communication between two peers starts from the lowest power level and reaches a level at which packets are received successfully- this is referred to as "low blast";

The simulation results provided by [14] show that the low blast mode has a slightly lower energy consumption pattern and better throughput. However, it has a greater latency due to the time taken by the nodes to reach a power level at which packets are not dropped.

The overall performance of the two modes show similar results of 10% - 15% increase in throughput and 10% -20% improved power consumption in comparison to the unchanged MAC layer.

The overall efficiency of the algorithms is considered to be poor because power control takes place only during unicast communication. According to [14] this algorithm does not provide power control for broadcast/multicast messages. Considering the fact that very often routing in ad hoc networks takes place with the help of broadcasting, it is expected that the proposed algorithms would not significantly improve the power efficiency of the network.

The algorithm relies exclusively on information about the strength of the incoming signal. This sets a requirement on the wireless transceiver to be able to facilitate accurate signal strength measurements. Unfortunately most of the "off the shelf" wireless network adapters do not provide such measurements.

Since [14] does not provide evidence of a working prototype of the protocol, it is not clear how the issue of power level switching latency would be addressed.

2.2.3 Topology control for power efficiency

In addition to power efficiency, power control algorithms are useful for optimal topology control. Proper use of such control results in improved network connectivity and throughput. This concept is investigated by Wattenhofer *et al* [15], where it is shown that careful management of the degree of a node[1] leads to improved throughput and prolonged network lifetime.

The goal of the algorithm proposed by [15] is to provide a location-based, distributed topology control for optimal energy efficiency whereby:

- nodes use their local information to determine their operational power and connectivity degree;
- local decisions are made to guarantee global network connectivity;
- power efficiency is achieved by utilisation of minimum power paths;
- small node degrees are maintained to minimise interference and maximise throughput;

The algorithm uses a two phase approach: i) the initial phase where the nodes identify their neighbours in all directions and ii) an optimisation phase where suboptimal connections to neighbouring nodes are removed.

As described in [15], in the initial phase, a node starts beaconing with an increasing transmit power. When the neighbouring nodes detect the beaconing they reply to it. With the help of the reply, the beaconing node determines the relative location of the replying nodes (in case GPS is not used). This phase can be terminated by one of two conditions being reached: a) A sufficient number of neighbouring nodes have been found such that for any cone with an angle α where the union of all cones covers the whole 2π angle, there is at least one neighbouring node; b) the maximum transmission power has been reached. At the end of the initial phase, a node has established connections with all of its neighbours and is ready to perform a selection of the optimal connections.

[1] node degree is the average number of neighbouring nodes that a node has within its transmission range

In phase two, the number of neighbouring nodes is revised. As depicted in Figure-2.3, if in a particular direction there are two or more neighbours within the given cone, the node that requires the lowest transmit power level to be reached is kept in the list of neighbours. Following the same rule, the beaconing node in Figure-2.3 would disregard node 2 and leave only node 1 in its neighbours list. Similarly, if two neighbouring nodes within the same cone are reached with the same transmit power and at least one of them could be reached indirectly (over two hops), then that node is removed from the neighbours list of the original node to keep the node degree small.

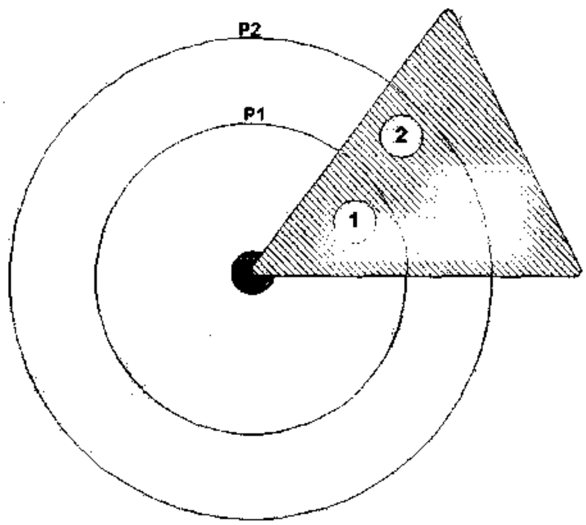


Figure 2.3: Directional node detection

A simulation of the algorithm is provided in [15] for the case of a static ad hoc network. It shows that network life time is prolonged by 30% - 40%. Also the throughput of the network is four times higher due to the controlled number of connections to a given node.

A drawback of the algorithm is related to its dependence on a positioning system such as Global Positioning System (GPS). In cases where it is not available, a solution to the Angle-of-Arrival (AOA) problem is required with the help of more than one directional antenna per node [16].

In a static ad hoc network, frequent updates of neighbour lists are not required. However, if the algorithm were to be applied in a mobile ad hoc network, the updates would have to be more frequent depending on node mobility. This would create excessive configuration overhead and interference. In addition to that, the authors have not specified an ordered procedure for the beaconing phase (phase one), which would result in increased contention for the communication channel and therefore extra energy expenditure.

2.2.4 Power control in clustered ad hoc networks

The COMPOW power control algorithm discussed in Section 2.2.1 was designed to achieve convergence to a common transmit power level for all nodes in the network. According to the investigation from [17], in a case where the nodes are grouped in clusters and the clusters are within different distances of each other (i.e. a non-homogeneous nodal dispersion), the COMPOW algorithm would be highly inefficient. It would converge to the highest communication power level that is needed to cover the distance between the two most remote clusters in the network. Hence the nodes in a cluster would use higher then adequate transmit power levels and that would result in unnecessary power expenditure.

Improved versions of the COMPOW algorithm for clustered ad hoc networks are presented in [17]. The algorithms choose the transmit power levels in a way that provides low transmit power within a cluster. High transmit power levels are only used for inter-cluster communication. The algorithms proposed in [17] identify clusters on the basis of the required transmit power. This technique is in contrast to the commonly used method of cluster identification based on location. An example of transmit power level clustering is given by Figure-2.4.

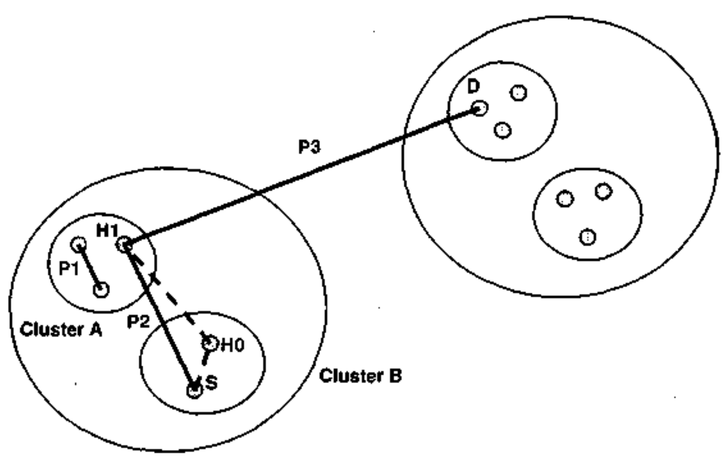


Figure 2.4: Ad hoc network clustering

Figure-2.4 shows two types of clusters A and B. Type A is identified as a cluster within which transmit power level P_1 is required for successful communication. Cluster B is formed by nodes that require transmit power P_2 to reach each other. Communication between clusters of type B is done using transmit power P_3.

The authors of [17] propose two cluster based algorithms for power control in a non-homogeneous ad hoc network. The first one is named CLUSTERPOW. It has been designed for routing through power control. Similar to COMPOW, it works with a number of routing tables that correspond to the available transmit power levels. The next hop is determined by the table that indicates an existing path to the desired destination and requires the lowest transmit power level. Following this rule, node S on Figure-2.4 would check all of its routing tables for nodes that provide routes to node D and compare the required transmit power levels to reach them. As in the case depicted by the figure, the next hop that provides a route to destination and requires the lowest possible transmit power P_2 is H_1.

It is has been proved by a number of investigations on the topic [13] [17] [18] , that numerous low power hops are preferable to fewer high power hops. This is the reason for the optimisation of the CLUSTERPOW, which results in the proposal of the second cluster based algorithm named Recursive CLUSTERPOW. The scheme performs a recursive lookup amongst the existing routing tables to verify if it is possible to use a few low transmit power hops on the route to the initially chosen next hop. In the case of Figure-2.4, the recursive algorithm would try to reach H_1 with the help of a few lower power hops. The result of the search would show that it is possible to break down that route into two hops if node H_0 is used as an immediate next hop. The new route to H_1 is shown by the dashed line in Figure-2.4.

As in the case of COMPOW, the two cluster based algorithms deal with multiple routing tables that correspond to different transmit power levels. Thus the efficiency of the algorithm depends on the number of the available power levels and node degrees. The larger the number is, the more time and power consuming the table maintenance procedures would be. The work presented in [17] does not provide information on the percentage power reduction achieved. More importantly, no estimations are provided for the resulting gain in power efficiency as a ratio between power expenditure for frequent table updates and power saved by the algorithms.

It is indicated in [17] that the Recursive CLUSTERPOW achieves better results. however, its performance is significantly more dependent on the number of existing routing tables. This is based on the fact that it performs a number of additional recursive searches for the optimal routes from source to destination. The algorithm is associated with a finite probability of packets getting stuck in infinite loops. To rectify this, a method of packet tunnelling has been suggested in [17]. To avoid the infinite loops, the packets are recursively encapsulated with the address of the hop for which the lookup is being done.

2.2.5 Distributed power control in ad hoc networks

Power control algorithms for cellular networks have been a focus of investigation for a number of years. Fully developed and efficient algorithms are used by the presently deployed cellular networks. In a typical cellular network the power control is dispensed by the base station associated with a given cell. In the case of an ad hoc network, however, such centralised power control methods are not directly applicable. To be used under the distributed nature of an ad hoc network, these methods require a number of modifications.

The power control algorithm suggested by [12] is based on a popular power control scheme used by presently operating cellular networks. To adapt the scheme for use in an ad hoc network, it has been transformed from a centralised to a distributed version. Its applicability to an ad hoc scenario is achieved by enabling all of the participating nodes to control the transmit power levels of their neighbours. As a result, the nodes carry out power control functions themselves and the need for a centralised infrastructure such as a base station is eliminated.

The fact that each node is able to control the power used by its neighbours necessitates the following routines:

- control set selection method - to reduce the signalling overhead and inter-node interference, each node is restricted to control the transmit power level of only a limited number of neighbouring nodes. The neighbours to be controlled by a node are selected via this method and grouped in a control set;

- there is a high probability that a node would receive power control commands from more than one neighbouring node at the same time. Thus a node needs a method to resolve contradicting power control commands. Such methods are referred to by [12] as power adaptation methods;

The author of [12] spends a considerable amount of time in discussion of the proposed control set selection and power adaptation methods. For the selection of a control set the following methods are suggested:

Power Threshold - a node controls the transmit power level of nodes that transmit with power greater than some predetermined threshold value. It is shown in [12] that in many cases, as the algorithm progresses, nodes are instructed to decrease their transmit power which after a few iterations results in their exclusion from the control set. This could lead to the occurrence of unwanted network partitioning. Hence a careful identification of the threshold value is important.

Distance Based Connectivity - a node controls the transmit power of all neighbouring nodes within distance d_{max} . The distance to a node is determined either with the use of a Global Positioning System (GPS) or by an estimation of the strength of a pilot signal of the transmitting neighbouring node.

Greatest K Nodes - a node controls the nodes from which it receives the strongest signals. The node ranks the received signals (evaluating the corresponding pilot signals) and only sends power control commands to the K nodes which transmit with the strongest signals.

The Greatest SIR Received - a node will send power control commands to a set of K nodes from which it receives the greatest signal-to-interference ratio.

After the exchange of power control commands, each node would have received a number of transmit power increase/decrease commands. [12] provides a number of ways for selection of commands to be obeyed by a node. These are summarised as follows:

Satisfy all - a node would decrease its transmit power only if all of the controlling nodes require that, that is if all of the received power control commands require transmit power reduction. The idea here is to minimise consecutive power decrease, which would lead to network partitioning. The main drawback of this method is related to the positive feedback effect. When power is increased, interference is also increased. Hence to sustain the targeted SIR, the transmit power needs to be increased again.

Satisfy mean of links - a node adjusts its transmit power in the direction which is required by the majority of the controlling nodes. That is if there are more nodes instructing a given node to increase its transmit power than there are nodes instructing it to decease its transmit power, then the node will increase its transmit power. If the numbers are equal, then the power level will be decreased to reduce interference. Such an approach is more balanced and less susceptible to positive feedback.

Satisfy the best link only - a node decreases its transmit power if any other node instructs it to do so. Power is only increased when the SIR targets of all its neighbours are not met. This method further decreases the susceptibility to the positive feedback effect. However, there is an increased chance of network partitioning.

Satisfy weighted arithmetic mean (WM) - a node adjusts it transmit power in the direction which is calculated to be more desired by the user. With the help of scaling factors (weights) a particular trend could be introduced. For example if the weight of the "increase" command is higher than the weight of the "decrease" command, then the trend is that "increase" commands are more important.

The efficiency of the algorithm is verified against the achieved optimisation of the following metrics: network connectivity, efficiency in terms of transmit power, system outage - the ability of the system to sustain satisfactory performance with changing operational conditions.

Simulation results presented by [12] indicate that maximum network connectivity is achieved when WM 1:2(weights of increase : decrease) is used in conjunction with SIR information from the greatest K with K set to five. If transmission power conservation is more important than connectivity a good choice would be the use of WM 1:1 with SIR data from the greatest K where $K = 2$. The lowest average transmission power is achieved by satisfying the best link only. However, this method produces the worst connectivity results.

Based on simulation results, it is concluded in [12] that the convergence of the algorithm depends on the initial transmit power level. The greater the initial transmit power the longer it takes for the algorithm to converge.

As mentioned earlier, the proposed algorithm stems from existing power control algorithms for cellular networks. Therefore it inherits characteristics like computational expensiveness and reliance on signal strength measuring facilities provided by the wireless transceivers. Although these appear to be typical for cellular networks, increased computational intensity would have a negative impact on the performance of an ad hoc network where nodes are characterised by low computational power and power conservation is critical. In addition, most of the available "off the shelf" wireless adapters do not facilitate signal strength measurement.

The power control algorithm has been tested in a custom designed simulation environment, which does not allow for a performance based comparison with other existing power control algorithms. In addition, some of the reported results appear unrealistic and are only useful in verification of the efficiency of the different variations of the algorithm as proposed by the author. As a result the real world performance characteristics of the algorithms are inconclusive.

2.2.6 Discussion on power control algorithms

Power control algorithms attempt to identify the most appropriate transmit power levels with the help of some knowledge about the current state of the communication channel and the topology of the network. As in most of the discussed algorithms, such information is obtained with the help of time consuming, highly iterative procedures, which often involve computation of signal and channel characteristics. Hence, power control algorithms are characterised with long, computationally heavy routines, which require collaboration with sophisticated wireless transceiver hardware.

In modern cellular and local area networks, fixed infrastructure units are responsible for the network management functions. These units are designed with the capability to execute computationally expensive routines while facilities are provided for gathering the required information.

The nature of ad hoc networks is different from that of conventional cellular networks. Within an ad hoc network, in addition to its regular operations, a node is required to carry out network management including power control. Excessive computational load related to these functions is not desirable and should be minimised as much as possible for the following reasons:

- intensive computation tasks (related to network configuration and maintenance) would hinder the execution of ordinary node functions by consuming excessive processing time;
- generally ad hoc network nodes are not designed to deal with excessive computational loads and in most of the cases do not have the required sophisticated hardware (in contrast to a typical base station used in a cellular network);
- frequent computationally intensive tasks results in additional power expenditure;

Evidence from the investigated power control algorithm simulations suggests that these tend to struggle in the presence of ad hoc network characteristics such as mobility, non-homogenous nodal distribution, difference in initial power levels between nodes and latency associated with shifts between power levels.

High mobility requires frequent network reconfiguration. As a result, the computational intensity is greater and the associated communication overhead is further increased giving rise to inter-node interference. To remedy the situation the nodes usually boost their transmit power, which results in poor power efficiency.

Non-homogeneous nodal distribution results in convergence of the algorithms to suboptimal (some times unacceptable) power levels. Delays posed by power switching procedures result in unwanted end-to-end packet delivery latencies that could be crucial in some of the applications with high QoS requirements.

It is the opinion of the author that power control algorithms should be used in network environments that are designed for the purpose. That is. the design of the nodes allows for intensive computation tasks and provides the required hardware facilities. On the other hand. implementation of power control algorithms on "off the shelf" general platforms could prove to be a difficult task resulting in suboptimal gain of power efficiency. The focus of the ongoing research work in this project is related to general platform ad hoc networks therefore other ways for power use optimisation in an ad hoc network are sought.

2.3 Power Aware Routing

Routing is the process of assigning communication links between the nodes of a network. Once established, the links are used for the transfer of data packets from source to destination. Depending on the number of destination nodes, routing could be subdivided into two fields: *unicast* routing and *multicast* routing. The former type of routing is associated with data forwarding from a source to a single destination node whereas the latter is used for simultaneous data delivery to many destination nodes.

For a number of years research work has produced an array of routing protocols [19] [20] [21] [22] that provide robust routing with no power consumption considerations. Some of the techniques result in an unacceptable power expenditure which significantly reduces the life-span of the network [1].

Woo *et al* [23] formulate the need for modification of the existing routing protocols for the purpose of power efficiency. It has been clearly stated in [23] that significant power conservation could be achieved by the routing protocols if the selection of routes that they are responsible for is based on power conservation criteria. This formulation has since stood as the main driver behind power-aware routing.

The following section describes and compares some of the current routing protocols and their power-aware versions.

2.3.1 Unicast routing

Unicast routing subdivides into three different types [24]: *source-initiated* (reactive), *table-driven* (proactive) and a *hybrid* of the two.

In the table-driven (proactive) approach, each node maintains routing information in the form of routing tables. To keep that information consistent with the network topology there is a frequent table update process triggered either by a timer or by an event such as a link error. According to an investigation in [25], maintenance of the routing tables results in poor scalability of this group of protocols. At the same time the frequent table updates result in increased routing overhead. However, when packets are routed with the help of pre-established routing tables, the end-to-end packet delivery delays are significantly improved [24].

Source-initiated (reactive) routing eliminates the disadvantages of table maintenance and frequent route updates. Routing protocols of this type discover routes only when they are needed. Hence, they are referred as on-demand routing protocols. Source-initiated routing is associated with lower communication overhead and higher packet delivery delay [25] [26] [27]. It has been established in [27], however, that at high node mobility, the control message overhead of some reactive routing protocols could exceed that of the table-driven protocols.

At present, there are a number of fully established unicast routing protocols such as DSR [21], AODV [22], DSDV [19] and TORA [20]. Most of these protocols have been studied and results from [25] [26] [27] indicate their adequate performance characteristics. Broch *et al* [27] evaluate a number of these protocols subject to different mobility and traffic conditions. Johanneson *et al* [25] carry out a similar study whereby the focus is on performance characteristics such as: communication overhead, throughput packet latencies and packet loss.

Before discussing the power consumption performance of some of the state of the art routing algorithms, a closer look should be taken at their different operation specifics and how they relate to the observed power consumption trends. Of particular interest is the communication overhead produced by each of the algorithms as it is found in [1] that it is directly responsible for the resulting power consumption.

AODV - Ad Hoc On-demand Distance-Vector

AODV is a source initiated routing protocol that uses *route requests* and *route replies*, which constitute a route discovery procedure, for the establishment of source-destination paths.

AODV as described in [22] stipulates that when a source node requires the transmission of data to an arbitrary destination node it first establishes the transmission route. This procedure is initiated by the source node transmitting route request packets (RRPs) to its neighbours. The RRPs contain the address of the source and destination nodes and a request ID. The intermediate nodes make note of the node that has transmitted the RRP which they then retransmit to their neighbours. The process continues until the destination node is reached by the RRP. Since on reception of an RRP, every intermediate node records from whom it has received it, the route between source and destination can be established. Once the destination node receives the RRP, it uses this route to send a route reply packet. Route reply packets indicate to the intermediate nodes that the route is valid and that they will be used as routers for the forthcoming data transmission. If an intermediate node does not receive a route reply over a period of time, it is free to overwrite the recorded route request. Once the source node receives a route reply that matches its route request, it can start data transmission.

If previously established data routes become unavailable due to node mobility, the node that detects that must inform the source node. As a result, the source node would initiate a route maintenance procedure.

DSR - Dynamic Source Routing

The Dynamic Source Routing (DSR) protocol from [21] is considered to be an improved version of the AODV protocol. The improvement comes with the introduction of route cache tables. In the case of AODV, the routing nodes keep information only about the last route request. In contrast to that, DSR allows the maintenance of small tables, which keep the information for a number of available source-destination paths. To keep the number of entries small, information about unused paths are regularly erased.

As a result of the available route caches, the route request procedure is slightly different from that carried out by AODV. When a source node requires data transmission, it first consults its route cache. If the required route exists as an entry, the node does not have to initiate a route request procedure. If a route is not found in the cache, the source node sends RRPs to its neighbours as in AODV.

Similarly, on the reception of an RRP, every intermediate node first checks its route cache. If a route to the required destination is found, the intermediate node replies to the source node with a route reply packet, which indicates the desired route. If the intermediate node does not have a route to the destination, it propagates the route request to its neighbours.

As stipulated by DSR, the route that a data packet has to traverse is embedded in the header of the packet by the source node. Thus each intermediate node only needs to read the header of the data packet to determine the next-hop. This constitutes the second major difference between DSR and AODV.

TORA - Temporary Ordered Routing Algorithm

TORA as described in [20], is a source initiated routing protocol. As it is similar to DSR and AODV, it makes use of the route discovery and route maintenance procedures.

What makes TORA different from the other the on-demand routing algorithms is the fact that it maintains and makes use of all available routes from source to destination. In contrast to that, DSR and AODV use a single route that is identified to require the least number of hops. This renders the TORA protocols highly adaptive in a mobile environment where, as a result of the mobility, the existence of a particular link is short in time. Its disadvantage, however, comes from the effect that multiple route maintenance has on communication overhead.

DSDV - Destination Sequenced Distance Vector

DSDV [19] is a modification of the Distributed Bellman-Ford routing algorithm [28] [29] which is popular with wired networks. The idea behind DSDV is based on the fact that the nodes of an ad hoc network are frequently used as routers. The protocol stipulates that they have to therefore maintain an updated routing table at all times and hence it is a classic example of table-driven routing. The required routing tables contain the next-hop information for every destination of which the node is aware. Each entry has an update sequence number and a value for the cost of the path to the destination in terms number of hops.

The process of routing information update requires each node to exchange such information with its neighbours. The exchange is either time or event driven. Over a given period of time all nodes perform a *full dump*. This is when the entire routing tables are exchanged between neighbouring nodes so any information about route changes will propagate through the network. However, if a node becomes aware of a route alteration it immediately informs its neighbours by transmitting only the updated table entry that corresponds to the changed path. This represents an event driven update. The receivers of the update make note of it and inform the rest of the nodes with the next full dump.

To differentiate between new and aging routing information, nodes pay attention to the update sequence numbers associated with each entry in the routing tables. Upon table updates, entries in the routing tables are only updated if the sequence numbers of the new routing information are higher than the ones present. Route update information is disregarded if the update sequence numbers match and there is no change in the routing information for a given route.

The Distributed Bell-Ford algorithm is known for the high probability of data packets getting stuck in infinite loops. These looping packets are infinitely exchanged between routing nodes which prevents them from reaching their final destinations. DSDV overcomes the infinite loop problem with the introduced *destination-sequence* numbers. They are updated in a manner which ensures that the next hop has either equal or a higher sequence number. As a result data routes form nondecreasing number sequences which are followed by the data packets in order to avoid infinite loops.

2.3.1.1 Performance investigation

Broch *et al* [27] investigate the packet delivery ratio and communication overhead characteristics of each of the four routing protocols DSR, AODV, DSDV and TORA. For simulation purposes, the protocols are subjected to different scenarios, relating to mobility conditions and number of source nodes.

Simulation results presented in [27] show that although AODV and DSR are similar, AODV produces five times more communication overhead than DSR in the worst case scenario of 0 pause time (random waypoint mobility model) and 30 source nodes. It has been found that AODV transmits 110 000 route request packets as opposed to 300 propagated packets by DSR. According to the information from [27], the observed difference is a result of the route cache tables introduced by the DSR protocol. Caching routes minimises the number of route request packets transmitted by each intermediate node and therefore lowers the communication overhead.

Further investigation in [27] shows that DSDV generates constant overhead, regardless of the loading conditions and number of source nodes. This was expected to be a result of the the periodic routing table updates which were set to occur every 15 seconds. However, the observed communication overhead appears to be generated by the triggered rather than the scheduled route updates. Simulations show that triggered updates are carried out at the maximum allowed rate of one per node per second. The resulting communication overhead is 45 000 packets send for a 900-seconds, 50-nodes, 0-seconds pause time, 10-sources simulation in comparison to 13 500 and 61 000 for DSR and AODV under the same conditions.

The worst communication overhead characteristic is presented by TORA. In the worst case scenario of pause time of 0 seconds and 30 source nodes the protocol undergoes congestive collapse. In the cases when congestion does not hinder the operation of the protocol, the produced overhead is 639 000 packets. In the best case of 900 seconds pause time and 10 source nodes, the protocol is responsible for 47 000 routing packets.

The general trend (with regards to produced communication overhead) is that on-demand routing protocols have the potential to perform equal to or better than the table-driven protocols even in cases of high mobility. This is achieved with the added benefit of no or limited maintenance of routing tables. The best performance in terms of generated communication overhead, under the the worst case simulation scenarios as reported in [27], is presented by the DSR protocol.

Assuming that the excess power consumed by a routing protocol is due to the transmission and reception of routing packets (neglecting power consumption due to computation and idling), the communication overhead produced should be a clear indication of its contribution to total power consumption.

2.3.1.2 Power consumption comparison

Investigation of the energy consumption behaviour of DSR, AODV, DSDV and TORA is the topic of [1]. For investigation purposes Cano *et al* [1] subject the protocols to a variety of scenarios. These scenarios are achieved by variation of the following five parameters typical for mobile ad hoc networks: number of mobile nodes, network area, mobility, number of traffic sources and data traffic patterns.

The simulation results presented in [1] are obtained using ns-2 [8]. The simulator uses a network interface model based on the WaveLan wireless adapter by Lucent. It has the following specifications:

- current consumption of 230mA in receiving mode and 330mA in transmitting mode;
- required voltage either 3.3V or 5V;

The consumed energy for reception and transmission is given in [1] as:

$$E_{RX} = \frac{230 \times 5 \times PaxketSize}{2 \times 10^6} \tag{2.1}$$

$$E_{TX} = \frac{300 \times 5 \times PaxketSize}{2 \times 10^6} \tag{2.2}$$

The investigation in [1] starts by subjecting the network to varying mobility conditions. The graphs in Figures-2.5,2.6 show the amount of energy consumed by each protocol for different node speed and pause time.

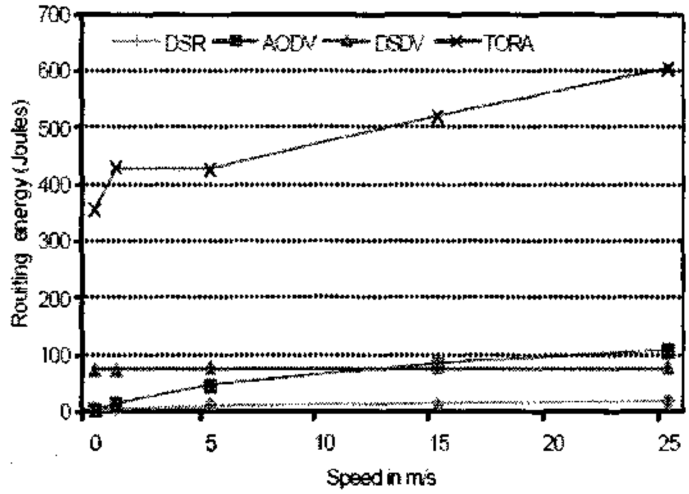


Figure 2.5: Energy consumption comparison as a function of node speed [1]

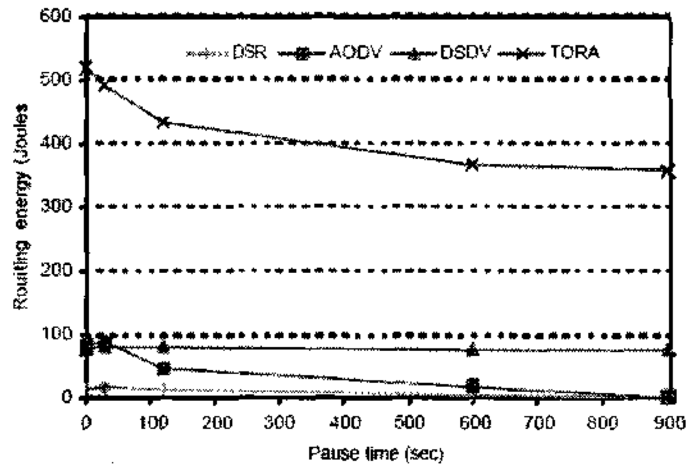


Figure 2.6: Energy consumption comparison as a function of node speed [1]

According to the simulation results DSR clearly outperforms the other protocols. Its performance is only slightly affected by node mobility. On-demand protocols present a tendency to use less energy when the rate of motion decreases. The table-driven routing protocol DSDV shows no change in energy consumption with mobility while the worst performance is presented by TORA. As expected, the energy consumption results are very close to the communication overhead performance results from [27](discussed in the previous Section 2.3.1.1).

Figures-2.7,2.8 present the energy consumption of the protocols as the number of nodes and traffic sources increases. Such increase is expected to result in larger numbers of issued routing packets.

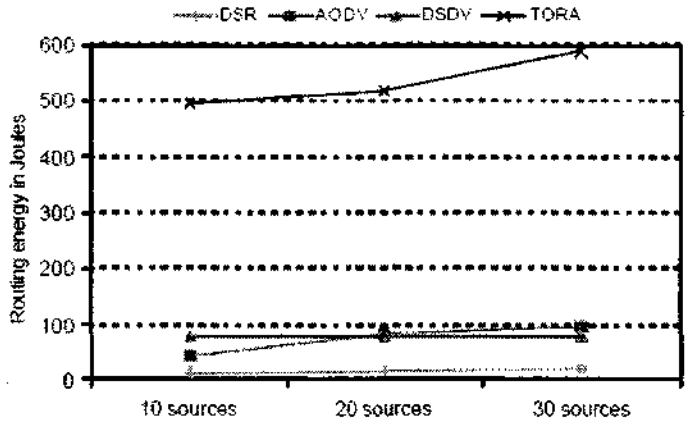


Figure 2.7: Energy consumption comparison as a function of traffic sources [1]

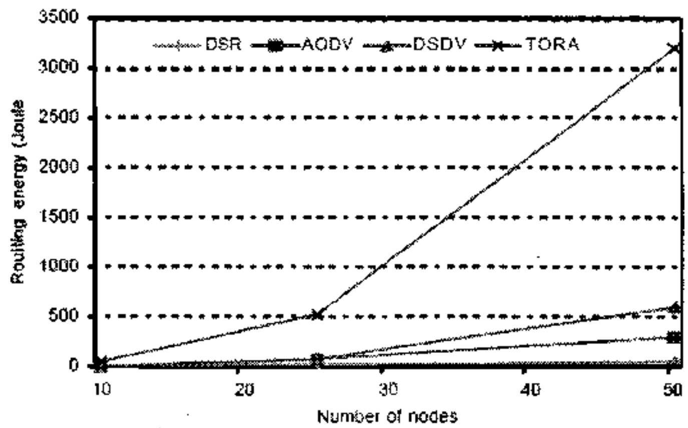
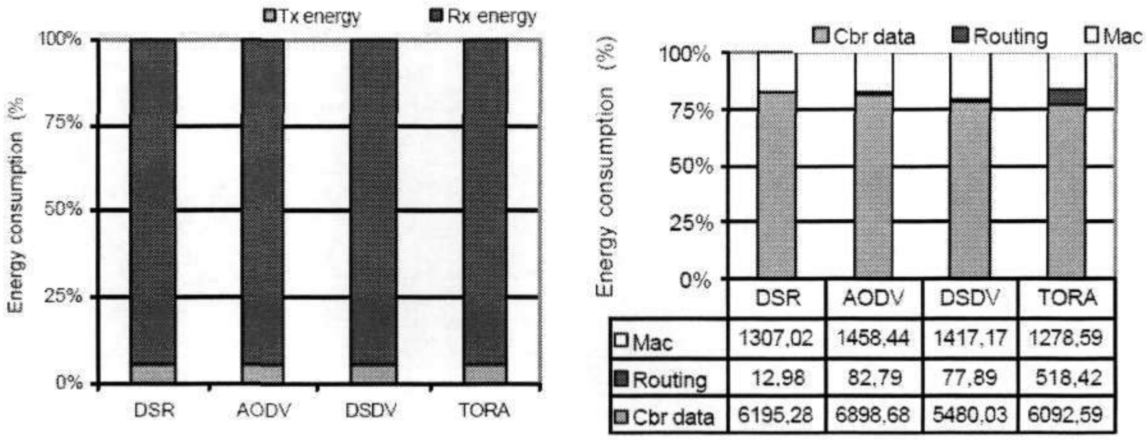


Figure 2.8: Energy consumption comparison as a function of node number [1]

However, the rise in energy consumption appears to be small except in the case of TORA. As shown by Figure-2.8 the number of participating nodes has minimal effect on the energy consumption of DSR and AODV. According to [1] this is related to the slight increase of route maintenance procedures. However, the same factor has a greater effect on DSDV and significant effect on TORA. The energy consumption pattern shown by DSDV is a direct result of the increased number and size of the exchanged routing tables. TORA shows an energy consumption increase represented by a steep line (for more than 25 nodes), which according to [1] is an indication of the poor scalability of the protocol.



(a) Percentage energy consumption per transmission and reception operations [1]

(b) Percentage energy consumed per packet type [1]

Figure 2.9: Energy consumption breakdown

Figure-2.9 generalises the performance characteristics of the investigated algorithms in terms of energy consumption. Figure-2.9(b) shows that DSR presents the best performance by spending as little as 12.98 joules of energy for routing purposes. Again this result is consistent with the previously discussed simulation results obtained by Broch *et al* [27] which indicate that DSR produces the least communication overhead.

An important result is shown in Figure-2.9(a). The energy consumption split between transmission and reception operations suggests that energy spent for data transmission is responsible for less than 10% out of the total consumed energy. According to [1], the bulk of the used energy is spent for reception, most of which is due to overhearing of remote data transmissions. It is the conclusion of the author that this result renders all techniques, which attempt to achieve energy efficiency solely through optimisation of transmit power use, significantly inefficient. Note that the estimations for the total power consumption does not included power expenditure due to node idling. Such expenditure could be of a great significance in ad hoc applications with low transmission duty cycles as noted in [1].

2.3.1.3 Modified DSR

The Dynamic Source Routing protocol shows excellent performance characteristics under a number of different network conditions. In addition to that, this routing mechanism produces minimal communication overhead, which is directly responsible for its superior energy consumption performance. However, the protocol does not actively employ any energy conservation mechanisms. This leaves room for modification, which could further improve its energy consumption efficiency. Modifications aimed at making DSR a power-aware routing protocol have been suggested by [30] [31] [32].

As it was described in Section 2.3.1, in the presence of more than one route between a given source-destination pair, DSR will choose the one that requires the least number of hops. In other words, DSR tries to optimise the hop-count metric.

Bhandare *et al* [31] suggest that the power consumption of the protocol could be minimised by substituting the hop-count metric with a metric related to power efficiency. That concept is the origin of the Energy Aware Dynamic Source Routing (EADSR) protocol proposed in [31]. The routing mechanism of EADSR is based on estimations of transmit power per source-destination route. In other words, a route between a source-destination pair is no longer chosen because it consists of the smallest number of hops but because it requires the lowest overall transmit power for a successful communication.

For the purpose of transmit power estimation, the header of the route request packet in EADSR is modified to contain power information. Each node that transmits such a packet adds to it the value of the transmit power P_{TX} that it will use. Each intermediate node that receives a route request estimates the required power P_{hop} for the corresponding hop. The estimation is done according to:

$$P_{hop} = P_{TX} + P_{thresh} - P_{RX} + M_d \tag{2.3}$$

where P_{thresh} is the sensitivity threshold of the receiving node, P_{RX} is the signal strength of the received signal and M_d is a margin added to ensure successful communication under channel variability.

After P_{hop} is obtained. it is logged in the route cache with the rest of the routing information for the corresponding hop. If the intermediate node does not have the route to the destination in its cache. it adds its ID and the obtained P_{hop} to the route request packet and forwards it.

When the destination node receives the route request packet. it extracts the route to the source and the information about the transmit power required by each hop. That information is then inserted in the route reply packet. which is sent back to the source node.

At the end of the route discovery procedure. the source node has the route to the destination with the exact power values required by each hop. This enables it to calculate the total required power to reach the destination node along the obtained route. Prior to data transmission. the source node chooses from its cache the route that requires the lowest total transmit power, extracts the corresponding routing and power-per-hop information, inserts it into the header of the data packet and begins transmission.

Additional modifications of the DSR proposed by [30] [31] include :

- route replies for intermediate nodes are disabled. This is based on the study by Maltz *et al* [33]. which proves that more than 60% of the routes provided from intermediate nodes are not valid;
- nodes keep a timeout associated with each route entry in their cache. Thus routes expire if they are not used for a given period of time. This reduces the chance of a source node selecting a route. which does not exist;
- nodes are enabled to collect route replies which are not directed to them. They compare the information contained by the snooped packets with the information in their route caches. If the comparison suggests that a node sits on a route that requires lower overall transmit power then the one advertised by the route reply, the intended source node is informed;
- on transmission of data packets, the transmit power requirements for each hop are checked and compared with the initially advertised values (during the route discovery

procedure). Thus changes in the total transmission power requirements are traced;

The EADSR protocol has been implemented and tested. Results documented in [31] are used for comparison between the performance of DSR and its modified version. The hardware test-bed (described in [34]) used for the tests consists of laptops equipped with Cisco Aironet 350 series wireless adapters. Each node is running Linux Red Hat 7.2. The experimentation procedures show that EADSR uses up to 30 times less power than the original DSR as result of the power efficient routes that it chooses.

Simulations of the EADSR protocol presented in [30], illustrate how mobility affects its performance. At first, the protocol is simulated in a static scenario. The choice of energy-optimal routes and the absence of mobility provide energy savings of 95% over similar DSR scenarios.

In the case of increasing mobility, the number of energy-optimal routes decreases as a result of the nodes being more wide spread in the network area. Therefore the achieved energy savings are lower than the ones documented for the static case. With increasing node velocity, EADSR tends to use routes of high stability such as minimum hop routes. As a result its power consumption performance converges to that of DSR.

2.3.1.4 Modified AODV

Senouci *et al* [35], propose three energy efficient routing protocols based on the AODV protocol. They are designed to increase network survivability with the help of energy-efficient route selection and operation based on residual battery power estimations.

- Local energy-aware routing based on AODV (LEAR-AODV)

The purpose of LEAR-AODV is to balance the energy consumption rates network-wide. This is done by allowing the nodes to choose whether they will be part of a route or not. The choice is based on the remaining battery power that a node has. In other words, a node can chose to reduce its participation in data forwarding and therefore conserve power. The protocol incorporates a mechanism that is used to avoid shortage of forwarding nodes due to selfish behaviour. To make all of the above possible, the Route Discovery and Route Maintenance procedures of AODV are modified.

During Route Discovery when an intermediate node receives a route request packet, it first examines its remaining battery power. If it is less than some predetermined threshold, the route request packet is dropped and the node announces that by broadcasting ADJUST_thr packet. This automatically means that the node will not forward data packet on behalf of the source node that sent the route request. Otherwise, if the intermediate node has sufficient battery power it retransmits the packet. To that end, it is guaranteed that the destination node will receive a route request along a route of nodes with sufficient battery power.

The Route Maintenance procedure in AODV is triggered by the unavailability of a hop along a source-destination route. An intermediate node that identifies a missing hop reports back to the source node and as a result a new route discovery is initiated. In the case of LEAR-AODV, a Route Maintenance procedure could be initiated by a node with a decreasing battery power. Nodes of the network are continuously checking their remaining battery power. If it becomes lower than the threshold value as a result of an ongoing data transfer, the node issues a route maintenance packet to the source node indicating that it will be no longer a part of the corresponding route.

LEAR-AODV provides a mechanism for a real-time adjustment of the threshold battery power values. This is to avoid the situation in which route request packets do not reach the destination node due to low battery power of the intermediate nodes. In such a case after an unsuccessful route request, the source node issues its following route request with an indication that the intermediate nodes must decrease the battery power threshold value.

- Power-aware routing based on AODV (PAR-AODV)

PAR-AODV assigns costs to each hop that lies on a source-destination route. They are based on the residual battery power of each node. Using these costs, all available routes are evaluated. The protocol uses the route that minimises the following function:

$$C(\pi, t) = \sum_{i \varepsilon \pi} C_i(t) \tag{2.4}$$

where

$$C_i(t) = \rho_i (\frac{F_i}{E_i(t)})^{\alpha} \tag{2.5}$$

and ρ_i is the transmit power of node i, F_i is the full charge battery capacity of the node i, E_t is the remaining battery capacity of node i in time t and α is a positive weighting factor.

During route discovery, prior to the transmission of a route request packet, each intermediate node calculates its link cost using Equation (2.5) and adds it to the header of the packet. Thus, when the destination node receives the route request packet it sends a route reply back to the source that contains the overall cost of the route. The source node selects the route that offers the lowest cost.

Additional *compute_cost* packets could be sent by the intermediate nodes in case they receive route request packets with a lower link cost than that currently in use. The compute_cost packets are sent to the destination node, which then informs the source node of the new, more cost-effective route. using a route reply.

- Life prediction routing routing based on AODV (LPR-AODV)

The last of the power-aware routing protocols proposed by [35] is LPR-AODV. It routes traffic through paths with a predicted long life-time. As in the case PAR-AODV the protocol assigns a cost to each link. The cost used by LPR-AODV is related to the battery life time of a node. The chosen route is the one that maximises the function:

$$\max_{\pi}(T_{\pi}(t)) = \max_{\pi}(\min_{i \varepsilon \pi}(T_i(t)) \tag{2.6}$$

where $T_{\pi}(t)$ is the life time of path π and $T_i(t)$ is the predicted lifetime of node i in path π.

The battery life time prediction of a node is based on its past activities. A good indication of the amount of traffic crossing the node is achieved by keeping a log of recent data routing operations. Every time the node sends a data packet it records its residual battery energy $E_i(t)$ at the given time instance t. The node also logs its residual energy $E_i(\dot{t})$ at time instance $\dot{t}$ when exactly N packets are sent/forwarded.

Similar to LEAR-AODV, each intermediate node calculates their costs in terms of predicted life time T_i using the following formulas:

$$T_i(t) = \frac{E_i(t)}{discharge\_rate_i(t)} \tag{2.7}$$

where

$$discharge\_rate_i(t) = \frac{E_i(\dot{t}) - E_i(t)}{t - \dot{t}} \tag{2.8}$$

and $E_i(t)$ is the remaining energy of node i at time t.

The estimated node cost is inserted by the intermediate nodes in the header of the propagated route request packet. On reception of a route request, the destination node issues a route reply which contains the overall route cost. If an intermediate node receives a route request packet with lower cost the destination node is informed by a *compute_liftime* packet. Thereafter the destination node informs the source node about the new route with a route reply packet.

The three algorithms are simulated and compared to the unmodified AODV protocol under two different scenarios: fixed and mobile. The improved network life time is studied in terms of:

- time taken for K nodes to die
- the time taken for the first node to die
- the time taken for all nodes to die

In the static case, the best performance is observed for LPR-AODV where the first node to switch off due to exhausted power resources under AODV routing appears 3244 seconds before a node malfunctions under LPR-AODV. This protocol outperforms the others by taking into account the battery discharge rates in addition to residual battery capacity.

In the mobile case, the LPR-AODV protocol once again offers the best performance in terms of network life-time extension. All three algorithms outperform the unmodified AODV algorithms under all mobility instances with an average network life extension of 1033 seconds at node speed of 4 m/s. As in the case of EADSR, with increasing mobility, the energy consumption performance of the modified versions of AODV converge to that of the original protocol.

2.3.1.5 Localised power-aware routing protocol (LARP)

Zhang *et al* [36] propose a protocol that operates at two levels: a lower level at which a traditional (DSR, DSDV, ADOV) routing protocol operates to provide global connectivity either proactively or reactively and a higher level at which alternative route selection is carried out based on power-efficiency.

The routing protocol that operates at the lower level guarantees that data packets will be successfully delivered to the intended destinations. As discussed earlier, the operation of the traditional routing protocols present at this level attempt to minimise the hop count, which results in the use of power inefficient routes. To provide power efficiency, a routine present at the higher level of the protocol enables each node to keep track of the links to their one-hop neighbours and the corresponding transmission power levels. This information is stored in a power table where minimum power routes between neighbouring nodes are stored.

Once the required data transmission routes are established by the routing protocol, intermediate routing nodes try to substitute the power inefficient links with a number of power efficient hops. The knowledge for these hops is extracted from the established one-hop power tables. In essence, if an intermediate node receives a packet that has to be transmitted to a next hop-π , it first consults its power table. If a route exists to π that is more power efficient than a direct transmission, the packet is routed along that route. Otherwise the packet is directly transmitted to π as indicated by the routing protocol.

The power tables are regularly updated with the transmission of *hello* messages which contain the identity of the sender, the identity of all its known one-hop neighbours and the minimum transmission power required to reach them.

This protocol is an alternative approach for improvement of the power performance of a traditional routing protocol without actually modifying it. Simulations of the protocol in [36] provide evidence for power efficiency improvement of up to 90% in comparison to cases where a traditional routing protocol is used. However, the simulation results are obtained with the assumption of an ideal MAC protocol, which suggests that there are no packet collisions. In the presence of increased routing overhead with the addition of frequent hello messages, packet collision would increase. Therefore such an assumption is expected to lead to unrealistic results because it does not account for the energy loss due to the potentially increased number of packet retransmission retries. Energy expenditure due to MAC signalling (RTS, CTS and ACK) is also not considered.

2.3.1.6 MAC signalling

Path-cost assignment on the basis of some energy efficiency criteria is a typical way for converting a regular routing protocol to power-aware routing protocol. However, according to Zhu *et al* [37] most of the proposed power-aware routing protocols assign costs based on models that do not capture the entire energy consumption by a node for communication purposes. In most cases, energy expenditure due to MAC signalling (RTS, CTS and ACK packets which are transmitted at full power as defined by IEEE 802.11) and data packet retransmissions is not accounted for.

For an improved power efficiency, power consumption models are proposed in [37] that account not only for energy consumed by data packet transmission but also for transmission of MAC layer control packets. Each of these models corresponds to different channel access methods. One of the models is based on the IEEE 802.11 MAC. It is identified in [37] as a *hop-by-hop retransmission*(HHR) mechanism where all intermediate nodes provide link-layer retransmissions (as opposed to an *end-to-end retransmission*(EER) mechanism).

There are two ways of transmitting data frames over the channel when using IEEE 802.11, either with the help of the *two frame exchange scheme* (also known as the *basic* scheme) or *four frame exchange scheme* . Because these two methods are associated with different numbers of transmitted control packets they have different energy consumption models.

In the two frame exchange scheme, a node transmits a data packet if the channel is idle for a period that exceeds the Distributed Inter Frame Space (DIFS). If the channel is busy it will defer transmission until it is idle for a period of (DIFS) and then it will start a backoff timer with a random backoff time. If during the backoff time the channel is busy the timer will be paused. When the timer reaches zero, the node transmits a data packet immediately. The receiver replies with an ACK if it receives the packet successfully. If the sender does not receive an ACK within a given period of time the whole process will be repeated. A time diagram of the scheme is shown in Figure-2.10 and a full description of the process can be found in [2].

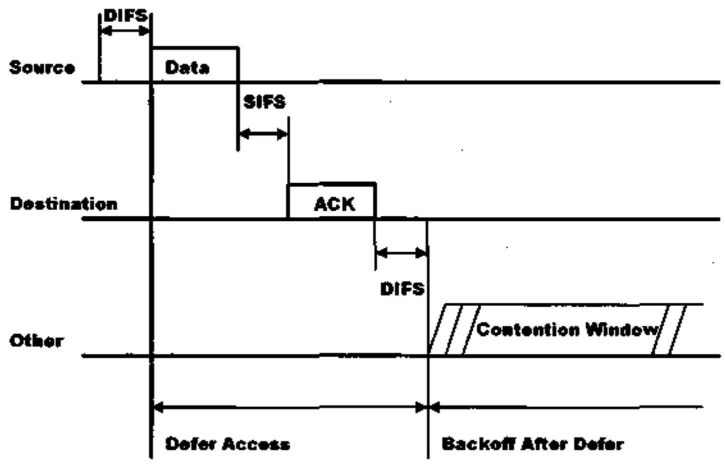


Figure 2.10: IEEE 802.11 two frame exchange [2]

A data transmission process between nodes i and j under IEEE 802.11 is associated with a packet error rate denoted as $p_{i,j}$, and an ACK packet error rate $p_{a,j,i}$. Under the given channel mechanism, according to [37] the total transmission power to transmit a packet from a node i to a node j is given by

$$\overline{P_T(i,j)} = \frac{P_{i,j} + P_{j,i}\frac{N_a}{N_8}p^*_{i,j}}{p^*_{i,j}p^*_{a,i,j}} \tag{2.9}$$

where $P_{i,j}$ is the transmission power from node i to node j and $p^*_{i,j}$ represents $1 - p_{i,j}$. The length of IEEE 802.11 symbols are represented by $N8$ and Na where:

$$N_8 = N + N_{802} + N_{phy} and N_a = N_{ack} + N_{phy} \tag{2.10}$$

N - data packet size

N_{802} - IEEE 802.11 header size

N_a - ACK packet size

N_{phy} - physical layer overhead packet size

Similarly, the total power consumed in receiving a packet from node i to node j is given by:

$$\overline{P_R(i,j)} = P_r(\frac{1}{p^*_{a,i,j}} + \frac{N_a}{N_8}) \quad (2.11)$$

where P_r is the power consumed for packet reception.

Using Equations (2.9) and (2.12), the average total transmission power for a packet sent from node i to node j successfully according to [37] is:

$$\overline{P(i,j)} = \overline{P_T(i,j)} + \overline{P_R(i,j)} \quad (2.12)$$

and the average total power consumed along a path from node 0 to node M is given by:

$$\overline{P_{total}} = \sum_{i=0}^{M-1} (\overline{P_T(i,j) + 1} + \overline{P_R(i,j) + 1} \quad (2.13)$$

In the four frame exchange scheme RTS/CTS packets are used for channel reservation and a virtual carrier-sense mechanism. Prior to data transmission the sender will transmit an RTS packet if the channel is available for a period longer than DIFS. The receiver answers with a CTS packet indicating that it is ready for transmission. If CTS is not received within a given time interval the RTS packet is retransmitted. Data transmission may commence only if a CTS packet is received and the channel is idle for a period longer than the DIFS. As in the previous case, data packets are acknowledged by ACK packets. RTS/CTS packets are used to inform the rest of the participating nodes that the channel will be unavailable for the duration of the intended data transmission. A timing diagram representing the scheme is depicted by Figure-2.11.

The total average receive and transmit power is increased by an amount required for the successful exchange of the RTS/CTS packets between communicating nodes. Therefore

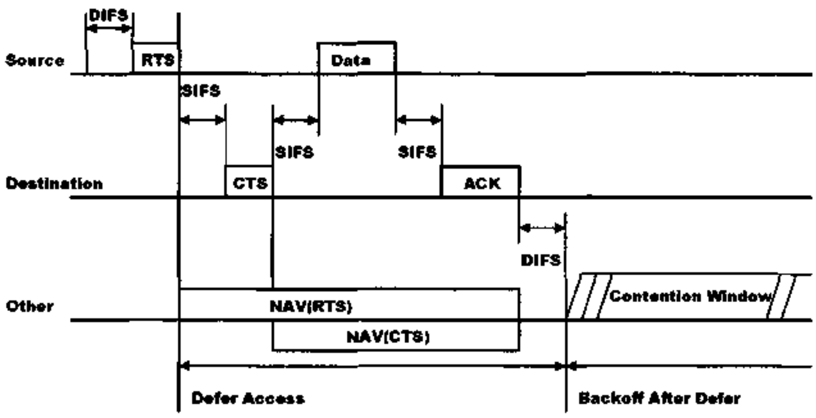


Figure 2.11: IEEE 802.11 four frame exchange [2]

Equations (2.9) and (2.12) are changed to:

$$\overline{P_T(i,j)} = \frac{P_m(\frac{N_r}{N_8} + \frac{N_c}{N_8}p^*_{a.i,j})}{p^*_{r,i,j}p^*_{c,i,j}p^*_{a,i,j}} + \frac{P_{i,j} + P_{j,i}\frac{N_a}{N_8}p^*_{i,j}}{p^*_{i,j}p^*_{a,i,j}} \tag{2.14}$$

$$\overline{P_R(i,j)} = P_r\frac{\frac{N_r}{N_8} + (\frac{N_c}{N_8} + p^*_{i,j} + \frac{N_a}{N_8}p^*_{i,j}p^*_{a.i,j})p^*_{a.i,j}}{p^*_{c,j,i}p^*_{i,j}p^*_{a,j,i}} \tag{2.15}$$

where P_m is the maximum transmit power available, $p^*_{r,i,j}$ and $p^*_{c,j,i}$ are the rates of successful transmission of RTS and CTS packets and Nr and Nc are packet sizes given by:

$$N_r = N_{rts} + N_{phy} and N_c = N_{cts} + N_{phy} \tag{2.16}$$

N_{rts} - RTS packet size

N_{cts} - CTS packet size

The average total consumed power along the path for source node 0 to source node M is calculated with Equations (2.14) and (2.15) which are substituted to Equation (2.13).

To provide energy efficiency, any routing algorithm that is designed to select communication routes while trying to minimise a cost function could be modified to use the energy consumption model described above. For simulation purposes the authors of [37] have modified the AODV routing protocol to select routes that offer the lowest cost obtained by Equation (2.13). The simulation results are compared with the performance of the *Retransmission Energy-Aware* routing protocol which uses simplified cost functions representing energy consumption for data transmission from [38]. The results have also been compared with the performance of an AODV routing protocol. that operates under a power control scheme that adjusts the transmission power according to the distance between the sender and receiver.

Simulation results show that as far as energy efficiency is concerned, the AODV routing protocol operating with the proposed energy model outperforms the schemes used for comparison in both cases - two and four frame exchange schemes. Under two frame exchange operation the energy consumption efficiency is improved by up to 32% in comparison to the performance of the simplified cost function and up to 52% in comparison to the performance of the transmit power control scheme. Under four frame exchange operation the energy consumption improvement is up to 23% and 26% accordingly.

2.3.2 Multicast routing

Multicast communication techniques allow a single source node to simultaneously deliver data to a group of nodes in the network (in broadcasting data is delivered to all nodes in the network). Multicast routing techniques for ad hoc networks are inspired by the broadcast nature of the wireless channel where a transmitted data packet will be received by all nodes that lie within the communication range of the packet source (equipped with an omnidirectional antenna) [39]. This fundamental network characteristic is referred as the *wireless broadcast advantage* (WBA).

According to [24], multicast routing algorithms for ad hoc networks can be subdivided into five groups: *group-based*, *source-based*, *core-based*, *mesh-based* and *flooding-based.*

Group-based routing protocols rely on a *forwarding* group of nodes for the correct delivery of multicast data packets. This technique does not require the maintenance of multicast trees (as is the case in most of the wired multicast routing algorithms) which according to [24] significantly reduces the communication overhead generated by a tree configuration. Some of the group-based routing protocols are: ODMRP (On-Demand Multicast Routing Protocol)[40] and LBM (Location-Based Multicast)[41].

In contrast to the group-based approach, source-based routing is done with the help of multicast routing trees rooted at the source nodes. Source-based routing protocols construct multicast trees on the basis of some efficiency criteria such as: battery life, latency, number of hops etc. Research has shown that finding a minimum-energy broadcast spanning tree is an NP-complete problem [34]. However, acceptable solutions could be obtained with the use of a number of different heuristics which results in a variety of source-based routing protocols. Examples for this type of multicast routing are the following protocols: DVMRP (Distance Vector Multicast Routing Protocol)[42] and ABAM (Associativity-Based Ad hoc Multicast)[43].

The core-based idea for multicast routing comes as an improvement to the source-based technique. To reduce the incurred communication overhead for tree maintenance purposes, core-based routing protocols use a single shared tree for each multicasting group. However, in the presence of increased node mobility, sharing a common multicast tree could prove to be unreliable. In such cases, according to [44], the decreased mean time between route discovery cycles results in prolonged communication interruptions. Examples of core-based multicast routing include: MAODV (Multicast Ad hoc On-demand Distance Vector)[45], AMRoute (Ad hoc Multicast Routing)[46]and AMRIS (Ad hoc Multicast Routing protocol utilising Increasing id-numberS) [47].

In an attempt to address the frequent link breakages in a highly mobile ad hoc network and the associated poor performance of the previously discussed multicast routing mechanisms, the mesh-based routing method was proposed. Protocols of this type use a number of different multicasting paths to deliver data. In cases of a link breakage, there is thus always a set of alternative data routes to multicast group members, resulting in a high reliability for this type of protocol. The Core Assisted Mesh Protocol (CAMP)[48] is a mesh-based multicast routing protocol.

An alternative to the mesh-based protocols (for highly mobile environments) are the flooding-based protocols. Instead of relying on pre-established links, these protocols broadcast data packets to all nodes in the network. According to [24] there are five different types of flooding: *blind flooding*, *probability-based* floods, *area-based* floods and *neighbour-*

knowledge floods.

2.3.2.1 Power-aware multicast routing algorithms

Most of the research work dedicated to development and optimisation of power efficient broadcast/multicast routing protocols is founded on source-based multicast routing. As previously discussed, this particular multicast routing approach operates with the help of multicast trees rooted at a source node. The difference between ordinary source-based multicast routing protocols and their power-aware counterparts is that the latter construct and use multicast trees that require minimum overall transmission power.

During analysis ad hoc networks are often represented as a complete directed graph $G = (V, E)$ where V is the set of nodes participating in the network and E is the set of direct links between the nodes. A model, which reflects the real-world situation is the geometric case where the set G is considered in the Euclidean space and the cost of each edge that belongs to E is defined as the Euclidean distance. Finding a multicast tree in an ad hoc wireless network is the same as finding a sub-graph $H = (U, A)$ in G with a special vertex $r \in V$ (the source node) where $U \subseteq V - \{r\}$ represents the multicast group members and $A \subseteq E$ represents a span of all links that belong to the multicast tree. In the case where $D = V - \{r\}$ the formed tree is a broadcast tree. To construct energy-efficient multicast trees, power-aware multicast routing algorithms attempt to minimise the sum $\sum_{i \in U} p[i]$ where p is the transmission power assigned to each node i that belongs to U.

Power-efficient multicast routing protocols could be subdivided into two categories as it is done in [49]: *local search* algorithms and *augmentation* algorithms. The difference between the two types is related to their initial conditions. The latter type starts with the construction of an energy efficient multicasting tree from an empty entry set, whereas the former type starts with an existing tree and transform it to a power efficient multicast tree.

2.3.2.2 Augmentation algorithms

Augmentation algorithms operate on $G = (V, E)$ where V is a finite set of nodes and E is initially an empty set of links between the nodes that belong to V. Such algorithms typically construct broadcast trees gradually, through steps, in which structures of links are added to E. The resulting trees could be represented as a directed graph $B = (D, U)$ with a distinct vertex r (the source node) where $D \in V - \{r\}$ and $U \in E$.

- Minimum spanning trees (MST)

MST is originally a routing method from wired networks where routing protocols construct minimum spanning trees, over which multicast/broadcast traffic is distributed [50]. In wired networks, the formation of such trees is usually based on the optimisation of a cost function. In that case, a cost is assigned to each available link. In wired networks, an MST problem could be solved by Prim's or Kruskal's algorithms [51] in polynomial time. However, when subjected to a wireless environment, MST does not make use of the wireless multicast advantage, which reflects its higher applicability to link-based networks such as all wired networks. Although there are suboptimal generalised link-based MST solutions for ad hoc wireless networks, its corresponding node-based problem (typical for ad hoc wireless networks) of finding minimum-energy broadcast routing is proved to be NP-complete [4]. Because of the high complexity related to the construction of energy-efficient broadcasting trees, ad hoc wireless multicast routing algorithms make use of a number of heuristics to obtain suboptimal energy-efficient solutions. Often suboptimal MST solutions are taken as a starting point by a number of ad hoc wireless local search algorithms.

- Shortest path first (SPF)

SPF [52] is another link-based method for the construction of broadcast trees in a network. Similar to MST, SPF takes as an input the complete network topology (available routes and their costs) and as an output it provides the most cost effective broadcasting tree rooted at a given source node. Used by the widely known Open Shortest Path First protocol [53] (for wired networks), SPF is a tree construction approach that is purely link-based. As such, the solutions that it provides for ad hoc wireless networks (where for power optimisation purposes the cost of each link is the required transmission power) are suboptimal. As in the case of MST, the solutions of SPF are often used by local search algorithms for further power optimisations. The difference between MST and SPF is that the latter is usually solved by Dijkstra algorithm [52] whereas MST solutions are obtained either by the use of Prim's or Kruskal's algorithms.

- Broadcast incremental power (BIP)

The authors of [5] have recognised the inadequate solution provided by a link-based MST/SPF approach for the construction of a broadcast tree in a wireless ad hoc network. They have also taken into consideration the fact that a similar node-based approach is highly complex. This has led to the proposal of the heuristic based algorithm from [5] for the construction of a minimum-power broadcast tree, rooted at the source node that spans all nodes in an ad hoc network. The heuristic used is - *minimum incremental cost.* It reflects the required increase of transmission power by an already transmitting node so a new node could be added to its transmission list and hence the name of the algorithm is *Broadcast Incremental Power* (BIP). Figure-2.12 shows the step-by-step creation of a broadcast tree with the help of the BIP algorithm. In step one shown in Figure-2.12(a),

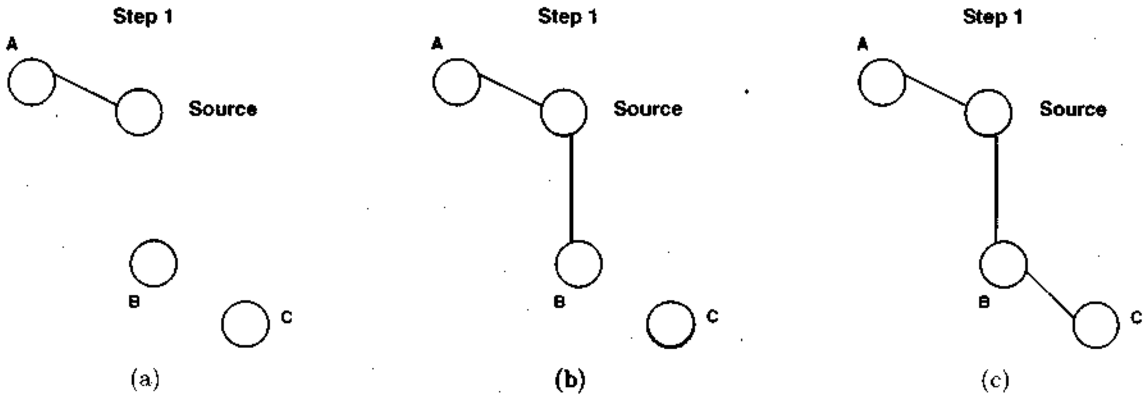


Figure 2.12: Broadcast tree construction through the BIP method

the source node must determine which of its neighbouring nodes requires the least transmit power for a successful communication. The required power to reach node A is less than the one required to reach node B so the first node added in the tree structure is A.

At the beginning of step two Figure-2.12(b), there are two options: 1) either the source node increases its power to add a new member in the tree structure or 2) node A looks for a neighbour to link with at a reasonable cost. In the given case, the incremental power required by the source node to reach a new node (node B) is less than the one required by node A to reach node B or C. Therefore the next node added to the tree structure is B.

At the beginning of the last step from Figure-2.12(c) there are three options: 1) source node further increases its power to reach node C resulting in a new incremental cost; 2) node A reaches node C at a given cost; 3) node B reaches node C at given cost. After a comparison between the three options, it appears that it is most cost effective if node B links with node C.

Although BIP is suited to the *node-based* nature of ad hoc wireless networks, since it makes full use of the wireless broadcast advantage, it does not always provide the most power efficient solution as will be discussed in Section 2.3.2.4.

- Incremental power with potential power savings (IP3S/MIP3S)

An alternative to the BIP algorithm is the Incremental Power with Potential Power Savings (IP3S) algorithm from [3]. It is a node-based broadcast tree construction method that makes use of the *potential power savings* idea proposed by its authors. The difference between IP3S and BIP is in the criteria for the selection of nodes for expansion.

In addition to the IP3S algorithm, Mehra *et al* have proposed in [3] a second algorithm - MIP3S for the construction of multicast trees in small scale ad hoc networks. The need for this algorithm is justified by the argument that in a small scale ad hoc network, achieving a multicast tree from a broadcast tree with the help of the pruning method (a common local search method discussed in the Section 2.3.2.3) is highly inefficient. MIP3S is also based on the potential power savings concept.

The potential power savings idea as defined in [3] is related to the overall tree power increase at each expansion step. Every time a node increases its power in order to reach new nodes, it checks if it has covered nodes that have already been covered. If it has, then the current power assignment at some other node is redundant. This indicates potential power savings that could be realised if it were possible to eliminate the redundant power assignment. Figure-2.13 shows an example of the concept. If node v on Figur-2.13 has to

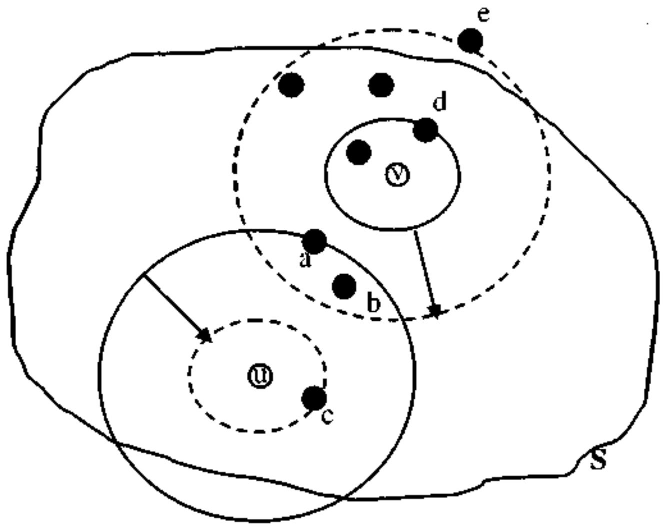


Figure 2.13: Potential power savings [3]

increase its power in order to reach node e, it will also cover nodes a and b. However, a and b are covered by node u so the power assignment at node u is redundant. Applying the potential power saving idea would lead to the reduction of the power level of node u. The achieved overall power saving is equal to the increased power at v minus the amount of reduced power at node u. The IP3S algorithm constructs broadcasting trees by expanding nodes $v \in S$ such that the increase in power at v minus the sum of power savings at all the u's is as small as possible. However, power reductions are only allowed if they do not compromise the connectivity of the broadcast tree.

The MIP3S algorithm (designed for the construction of multicast trees in small scale ad hoc networks) is based on the SPT method with the added benefit of the potential power savings idea. It maintains a set of nodes S reachable from the source node, which is grown until all multicast group members are included. In order to add a new (multicast group member) node to S the algorithm undergoes three steps. Step one finds the shortest path from the current set of nodes to the uncovered multicast group member. This could involve nodes which are not multicast group members. At step two, the required power levels are assigned to the relaying nodes that lie on the route to the uncovered multicast group member. Step three is an optimisation step where the potential power savings idea is applied. By design the MIP3S algorithm is centralised. However, a distributed implementation of the algorithm called DMIP3S is described in [54].

2.3.2.3 Local search algorithms

Local search algorithms transform already established broadcast trees $B = (D, U)$ into power efficient multicast trees $M = (R, T)$. This is done by reducing D to a subset $R \in D$ representing the multicast group members and optimising the set of edges U so that the overall cost of the resulting new set T is less than the overall cost of U.

- The Sweep method

The Sweep method from [5] is used to minimise the overall power consumption of an already created broadcast/multicast tree. The minimisation is achieved by removing redundant links between nodes that belong to the tree structure. To illustrate the concept of the Sweep algorithm Figure-2.14 depicts a simple broadcast tree scenario. Figure-2.14(a)

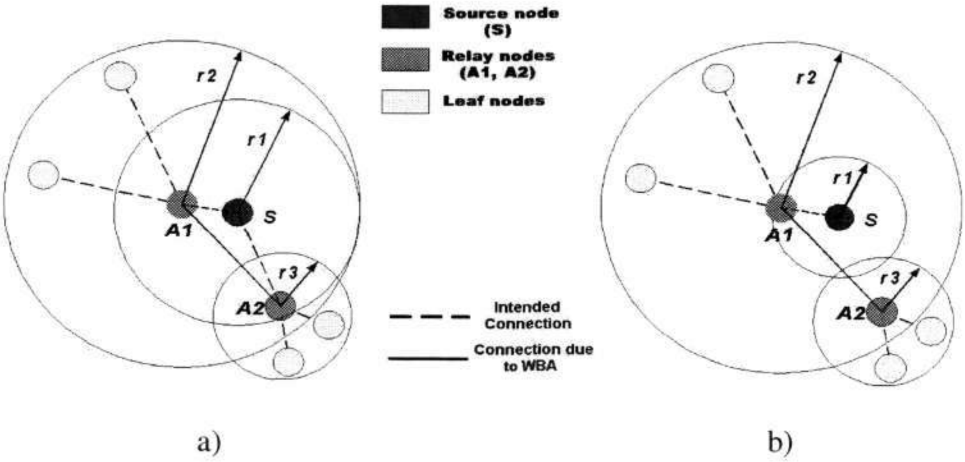


Figure 2.14: Potential power savings [3]

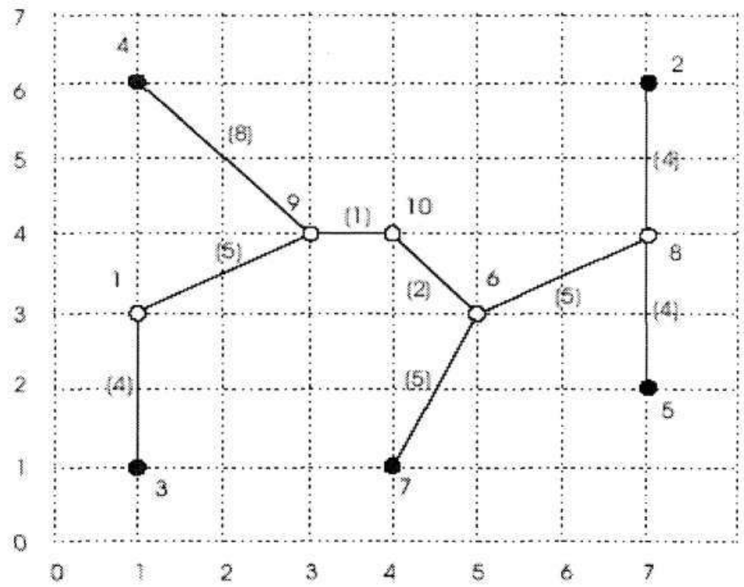
shows a suboptimal broadcast tree construction which results from the application of BIP. At execution, the Sweep algorithm would investigate the two relaying nodes $A1$ and $A2$ and the source node S. It will discover that the range $r2$ of node $A1$ is enough to reach the relaying node - $A2$. However, $A2$ is also a downstream node to S as shown by Figure-

2.14(a). Hence the connection between S and $A2$ is deemed redundant and S can lower its transmission power (without compromising the connectivity of the network) to a level sufficient just to cover $A1$ as shown in Figure-2.14(b).

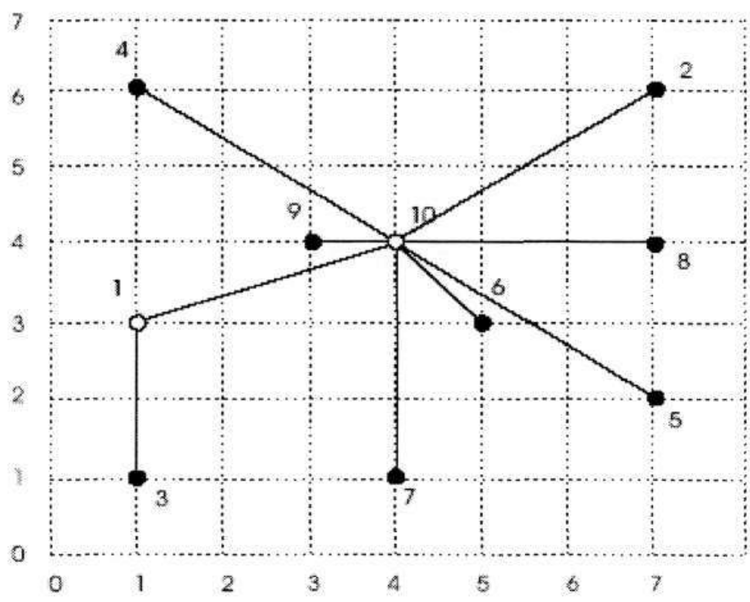
- Embedded wireless multicast advantage (EWMA)

In the case of the Sweep algorithm, the power consumption of suboptimal broadcast/multicast trees is optimised by removal of the present redundant links. The EWMA algorithm from [4] takes this idea one step further whereby the number of redundant links is maximised before their removal. This is achieved by increasing the transmission range of a number of appropriately chosen relay nodes.

EWMA operates in two phases. Phase one finds a suboptimal MST solution. Phase two transforms the MST to a power optimised broadcast/multicast tree. This is done in a number of iterations by defining and manipulating three sets of nodes C, F and E. C is the set of covered nodes, F is the set of transmitting nodes and E is the set of excluded relay nodes associated to the redundant links. The optimisation phase of the algorithm starts with $C = \{r\}$ and $F = E = \{0\}$, where r is the source node. Each iteration considers for power expansion the nodes that belong to the set $C - F - E$. The algorithm is terminated when all leaf nodes are included in C. At the end of the optimisation phase, F contains only transmitting nodes with high associated *energy gain*. Reference [4] defines the energy gain of a power expanded node as the decrease in the total energy of the tree, obtained by excluding some of the transmitting nodes from the original MST in exchange for an increase in the transmission power of the expanded node. An example provided in [4] and illustrated by Figure-2.15 shows the execution of the algorithm. Figure-2.15(a) shows the broadcast tree obtained by the MST method in the initial phase of EWMA (the required energies for the available links are given in the brackets). The total energy of the tree is eMST = 23. The second phase of the algorithm starts with $C = \{r\}$ and $F = E = \{0\}$ where the source node is 10. Therefore the algorithm considers for expansion $C-F-E = \{10\}$. Each iteration determines the energy gains associated with the exclusion of transmitting nodes as a result of the power expansion of the nodes that belong to the $C - F - E$ set. For example, in the present case, in order to exclude node 8, node 10 has



(a)



(b)

Figure 2.15: 2.15(a) MST broadcast 2.15(a) EWMA broadcast tree [4]

to increase its energy to cover 2 and 5 (the leaf nodes of 8) by:

$$\triangle e_{10}^{8} = \max_{i \in \{2,5\}} \{e_{10,i}\} - e_{10} = 13 - 2 = 11 \tag{2.17}$$

If the energy of 10 is increased by the calculated amount. it will also be able to cover the leaf nodes of relay nodes 6 and 9. So the resulting gain in excluding 8 by expanding 10 is calculated to be:

$$g_{10}^{8} = e_6 + e_8 + e_9 - \triangle e_{10}^{8} = 5 + 4 + 8 - 11 = 6 \tag{2.18}$$

Similar energy gain computations for node 10 are carried out with respect to the rest of the transmitting nodes in the initial MST solution. The highest possible energy gain is achieved when node 10 increases its energy to eliminate the need for node 8. Therefore after the first iteration the sets C,F and E look as follows: covered nodes so far - $C = \{1, 2, 4, 5, 6, 7, 8, 9, 10\}$; transmitting nodes so far - $F = \{10\}$; excluded nodes so far - $E = \{6, 8, 9\}$. However, the set of covered nodes C shows that node 3 is still uncovered. Therefore more iterations are required. The second iteration will examine and compare the gains of the following nodes $C - F - T = \{1, 2, 4, 5, 7\}$. At the end of iteration two, node 1 will be identified as the node with the highest possible energy gain so the final solution is $C = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $F = \{1, 10\}$ and $E = \{6, 8, 9\}$. The energy of the new broadcast tree shown by Figure-2.15(b) is eEWMA = **17**.

- Multicast incremental power (MIP) - The Prune method

The Multicast Incremental Power [5] is an extension of BIP for the construction of multicast trees. As such, it initially applies the BIP algorithm for the construction of a power efficient broadcast tree. This is followed by a tree transformation which results in a multicast tree. The transformation is done with the help of the *prune* process. Prune removes from the tree structure all leaf nodes that do not belong to the multicasting group together with all the nodes that are not needed to reach the group. Thus the set of nodes D is reduced to a set R which contains leaf nodes that belong to the multicast groups and nodes responsible for data relaying to the groups. Figure-2.16 shows the BIP tree and the associated data traffic before and after the pruning process.

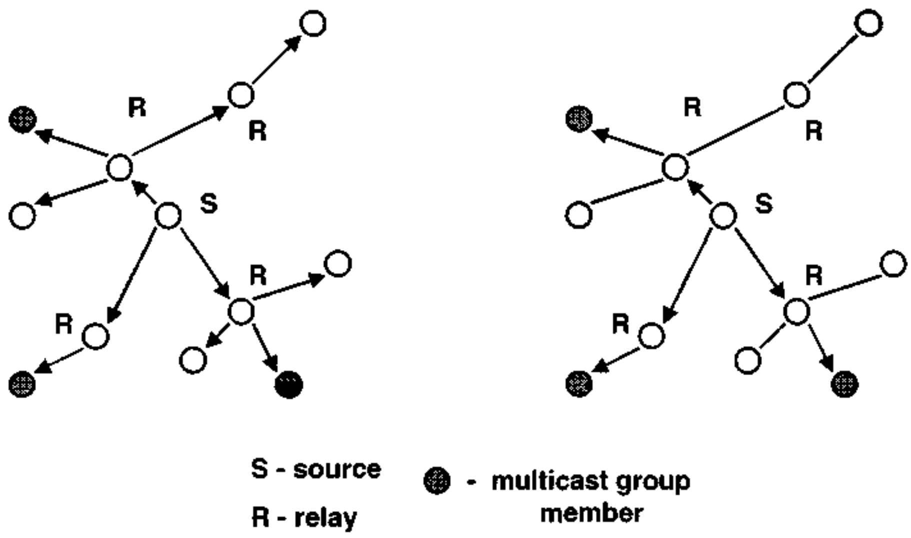


Figure 2.16: BIP and data traffic (a) before and (b) after the prune process [5]

- Refining energy-efficient source-based multicast tree (S-REMiT)

The S-REMiT [6] algorithm minimises the total power consumption of an existing broadcast tree through reassignment of downstream nodes to new parent nodes that require less energy. When long range radios are used, S-REMiT starts by creating a *minimum spanning tree*. Conversely in the case of shore range transmissions, the initial tree is a shortest path tree.

The refinement of the initial tree is done with the use of the $Change_i^{x,j}$ function, which changes the parent node of node i from node x to node j. The only condition for execution of the function is sustained connectivity. It stipulates that node j cannot be a descendant node of node i prior to the parent change. After the execution of the function, the nodes with affected energy levels are only i,x and j.

The selection of the new parent node j for a given node i is based on the energy efficiency gain that would result from the exchange. It is defined as the difference in energy consumption before and after the exchange. An example provided in [6] illustrates the steps of the parent change process. It considers the multicast tree shown in Figure-2.17. The transmission power levels of the two relaying nodes 9 and 6 depend on their most distant leaf nodes. In this case they are 2 and 8 respectively. In an attempt to decrease

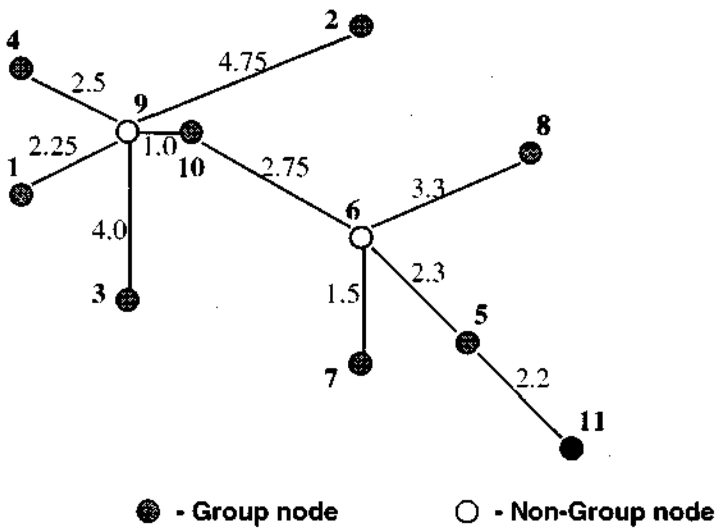


Figure 2.17: Multicast tree [6]

these power levels, the refinement phase of the algorithm would determine if there exists a positive energy gain in linking the leaf nodes to different parent nodes.

Considering a parent change from node 9 to node 6 for node 2, the algorithm first computes the current energy consumption of each relay node as $E_i(T, s) = K(r_{i,j})^\alpha$ (see footnote [2]). The computed values are: $E_6(T, 10) = 10.89$ and $E_9(T, 10) = 22.56$. The same calculations are made for the refined version of the tree T' which results form $Change_2^{9,6}$. If node 2 changes its parent node to be node 6, the new energy values would be: $E_6(T', 10) = 12.96$ and $E_9(T', 10) = 16$. The energy gain form the parent change would be:

$$g_2^{9,6} = (E_6(T, 10) + E_9(T, 10)) - (E_6(T', 10) + E_9(T', 10)) = 33.45 - 28.96 = 4.49 \quad (2.19)$$

It is also possible that node 2 switches to nodes 10 and 8. However, the resulting gains are unsatisfactory: -1.88 and -7.88. Hence for energy optimisation purposes node 2 is reassigned from its old parent node 9 to its new parent node 6.

[2] T - the original tree, T' - the resulting refined tree, s - source node, K - a constant related to the antenna properties. r - the Euclidean distance between the nodes; α - environment dependant constant

2.3.2.4 Performance investigation of multicast power-aware routing algorithms

Athanassopoulos *et al* [49] investigate by how much power efficiency is improved by local search algorithms when they operate on trees created by the basic augmentation algorithms. Simulation results presented in [49] suggest that the performance of the optimisation algorithms is strongly dependent on the propagation loss constant α. For lower values of the constant ($\alpha = 2$), local search algorithms such as EWMA and *Sweep* introduce significant power consumption optimisations. However, for higher values of $\alpha(\alpha = 4)$ the improvements are only marginal. The same trend is confirmed by [4] which suggest that node-based algorithms, which make use of the wireless broadcast advantage (eg. EWMA, BIP and BIP), perform significantly better for lower values of α and converge to the efficiency of a link-based MST solution for higher values of α.

Kang *et al* [55] have evaluated the performance of EWMA, BIP and MST. Through simulation, the following performance measures have been investigated: total transmit power, static network lifetime, total receive and interference power, hop count and ratio of transmitting to receiving nodes.

The best performance in terms of minimum total transmit power for all network sizes (between 10 and 300 nodes) is achieved by EWMA. It is followed by BIP while the least efficient algorithm being MST. These results are consistent with the results presented in [4], using the same metric.

In terms of the network lifetime span of a broadcast tree in a static network, the best performance is shown by BIP and the worst by EWMA. This is attributed to the fact that EWMA eliminates links by increasing the range (and therefore the power consumption) of a small number of nodes. Hence the batteries of these nodes are depleted faster which results in shorter network lifetime.

The investigation of the ratio of transmitting to receiving nodes shows that EWMA indeed makes use of a small number of transmitting (relaying) nodes that operate at higher transmit power levels. Although this strategy has a negative effect on the network lifetime, it provides satisfactory end-to-end packet delays because it is associated with a low hop count.

Gupta *et al* [6] have provided simulation results that display the total energy reduction achieved by the S-REMiT algorithm in comparison to EWMA and MIP. Simulations are carried out for the construction of both multicasting and broadcasting trees in two different scenarios: short and long range data transmissions. In the short range data transmission case which resembles the ad hoc wireless environment, S-REMiT outperforms EWMA by up to 71% and MIP by up to 63% in terms of the energy efficiency of the constructed multicast trees. However, for the construction of broadcast trees S-REMiT and EWMA display similar performances.

2.4 Summary and conclusion

Energy is a limited resource in ad hoc network environments and a number of energy consumption optimisation techniques exist. They can be classified in two categories: *power control algorithms* and *power aware routing protocols*. Both of these methods promote energy efficiency through optimised transmission power use.

A number of power control algorithms stem from similar techniques applied to centralised networks such as cellular networks. Under such algorithms, in the absence of a controlling unit, the nodes of an ad hoc network have the authority to regulate the transmission power of their neighbours through use of control messages. Other power control algorithms seek network-wide convergence to an optimal transmission power level that guarantees connectivity, lower interference and improved energy consumption. The intelligent choice of such a level is usually computationally expensive and requires information about the characteristics of the communication channel. Both of these conditions require sophisticated communications equipment which is often not available. It has been shown that node mobility is responsible for an increase in computational intensity and control and configuration communication overhead, thereby leading to degraded network performance. The investigated protocols have been found to converge to suboptimal power levels as a result of non-homogeneous nodal distribution and mobility.

Power-aware routing protocols just as routing protocols in general, can be divided into two fields: *unicast* and *multicast* routing. The former is responsible for data delivery between any two peers in an ad hoc network while the latter establish routes for data delivery from one node to many. Both operations (point to point and point to many points data delivery) are equally important and widely used with ad hoc networks.

There are a number of unicast routing protocols with well established performance characteristics. The three most commonly observed metrics are generated communication overhead, throughput and end-to-end transmission delay. Investigation has shown that the first metric is directly related to the energy consumption performance of any routing protocol. Section 2.3 presented a number of extensions to general unicast routing protocols which provide energy-aware operation. Such techniques substitute the criteria for a route choice to be - associated energy consumption. Simulation results have shown that these techniques achieve energy consumption efficiency under low node mobility. Further investigations provide evidence showing that the performance of the extended protocols converge to the energy consumption performance of their original versions as node mobility increases. The results suggest that the addition of such routing protocol extensions for networks operating under moderate to high mobility is ineffective.

Multicast algorithms base their operation on structures such as multicast groups or trees. Their construction is inherently expensive in terms of bandwidth and energy. Some of the existing techniques are practically inapplicable for ad hoc networks with rapidly changing topology. A number of power-aware multicast routing techniques were presented in Section 2.3.2. Simulation results have shown that they have the potential to construct power efficient multicast groups and trees. However, their investigation has been limited to static ad hoc network scenarios which suggests that they are inapplicable to mobile ad hoc networks.

Evidence has been found and presented in Section 2.3.1.2 suggesting that there is potential limitation with all techniques that promote power efficiency through optimisation of the transmission power. The limitation of these techniques is as a result of attempting to optimise transmit energy only which is, as low as 15% of the total consumed energy [1]. In addition, it is standard procedure that all control and configuration traffic is transmitted at a mandatory full power for reliability (e.g. in the case of IEEE 802.11 [2] protocol). Hence it can be concluded that the optimisations presented in this chapter are relevant to only a small fraction of the actual consumed energy and therefore achieve suboptimal results. It is thus one of the goals of this thesis to identify activities other than data transmission where energy conservation can be achieved. For the purpose of further investigation on this issue, Chapter 3 presents a study of the factors responsible for the energy consumption profile of ad hoc networks.

Chapter 3

Power Consumption Analysis of Ad Hoc Networks

3.1 Introduction

Optimal energy consumption efficiency in ad hoc networks can only be established through analysis of the processes that shape the profile of their energy consumption. Knowledge is required spanning: hardware design of mobile nodes; network dynamics typical for ad hoc networks and analytical tools that help to describe and quantify the factors responsible for the observed energy consumption performance.

The power consumption characteristic of mobile ad hoc networks is generally determined by the following:

- the design of the mobile nodes;
- the design of the network in terms of topology and protocols used;
- network dynamics;

The problem of *energy consumption optimisation* can be defined as: seeking optimal energy consumption of a node which has known hardware design and is subjected to expected operation conditions. The literature survey presented in Chapter 2 did not provide evidence suggesting that the discussed energy optimisation techniques have been established after an analysis of the energy consumption characteristics of ad hoc networks. Instead an observation was made that most of the existing schemes simply rely on the fact that level change of the transmission power is often facilitated, while the fact that transmission power optimisation could lead to suboptimal energy efficiency improvement is overlooked.

The objectives of this chapter are to: provide information regarding typical hardware design of an ad hoc network node; present an analysis which illustrates how network dynamics affect the operation of the network nodes and their energy consumption and report on the identified by the analysis weaknesses of the existing network protocols leading to inefficient energy consumption.

The chapter is structured as follows: Section 3.2 describes a typical hardware design of an ad hoc network node; Section 3.3 describes the typical operation states of a communicating node. Section 3.4 combines the states of operation with network topology factors to represent the network dynamics in terms of operation scenarios. The section also describes an analytical model which quantifies the energy consumption in each scenario. Section 3.5 presents a necessary extension to the discussed energy consumption model and the analysis of the factors shaping the energy consumption performance. Section 3.6 illustrates the effects of the identified factors through graphs representing numerical solutions of the analytical model.

3.2 Mobile wireless nodes

To understand how energy is consumed by the network nodes, a certain degree of familiarity with their hardware design is required.

Each mobile wireless node is a combination of a *host device* and a *wireless network interface adapter*(WNIA). The host device is an autonomous, battery-powered unit with data processing capabilities. The WNIA on the other hand requires power from the host device. It contains electronic components that carry out basic communication functions typical of the low level layers of the OSI hierarchy. The interaction between the host and the WNIA takes place over a standard communication interface with the help of a *device* *driver* on the host side. The responsibility of the driver is to map the activities of the user to functions that are specific for the device. The mapping process is described in detail by [56]. The combination between a host and a WNIA is depicted in Figure-3.1.

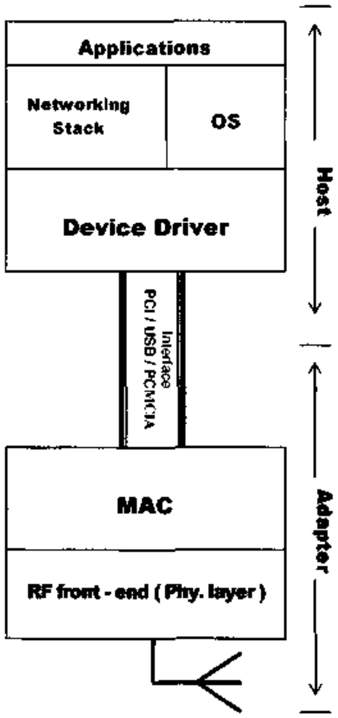


Figure 3.1: Mobile Wireless Node

On the host side, the device driver interacts with the part of the operating system that is responsible for networking, namely the networking subsystem. This subsystem takes care of the asynchronous arrival and generation of data packets. It consists of memory buffers and software modules that represent the networking protocols and layers from the OSI model. It forms a bridge between the device driver and the applications that use the device.

The hardware design of a typical WNIA includes two functional modules: a *radio* and a *MAC controller*. As described in [57], the radio is responsible for transformation of digital data and RF-signals for the purpose of wireless communication. To carry out these functions, the radio consists of the following components: *baseband processor*, *IF modem*, *RF-IF converter*, *low noise amplifiers and filters*. A component diagram of a generic hardware design for WNIAs is shown by Figure-3.2.

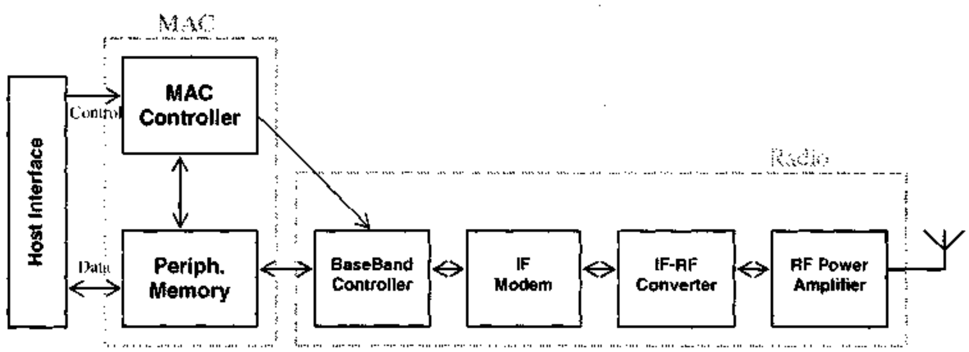


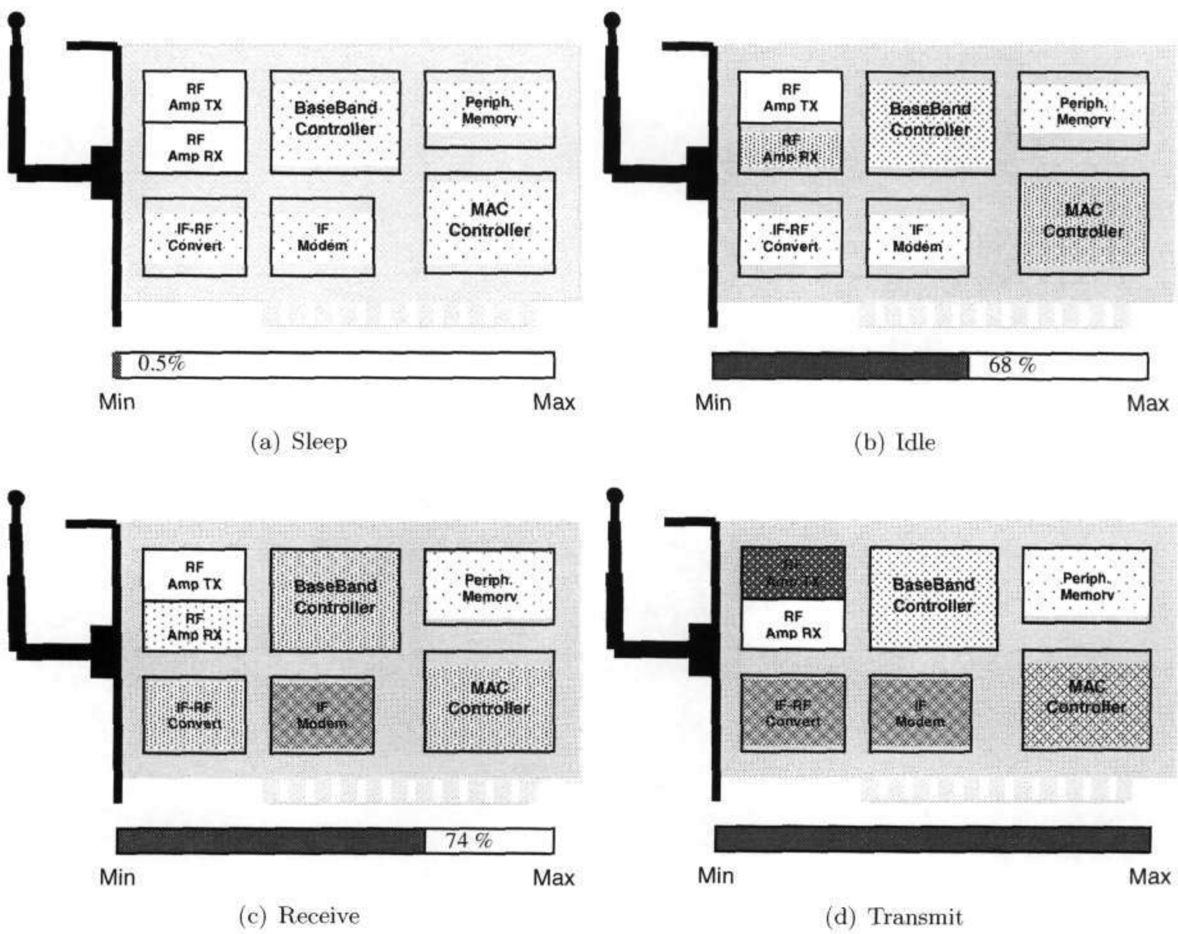
Figure 3.2: Components of a WNIA

The functions of the baseband processor are channel coding, modulation and symbol shaping of the outgoing digital data and synchronisation, channel equalization, demodulation and error correction of incoming signals. In the outgoing direction, the symbols generated by the baseband processor are converted to an analogue form and then passed to the *radio front-end*. Its responsibility is to transform the signal into an appropriate frequency which is followed by amplification, filtering and transmission through the antenna. The reception process is the inverse of that above. The baseband processor receives the signal in its base frequency from the radio front-end and converts it to a digital format. At the end of the signal transformation the processor outputs the regenerated bit stream.

The MAC controller contains a firmware implementation of a medium access control protocol that is responsible for regulation and arbitration of channel use. It has a peripheral memory module used for buffering of data packets. The implementation of the device driver is dependant on that of the MAC protocol as they interact with each other across the boundary between the host and the WNIA. It is becoming a trend to implement the MAC protocol as part of the device driver in order to reduce on-board code memory and thus reduce the price of the network adapter. Such an implementation is described in [58].

3.3 States of Operation

The energy consumption of a node is the sum of the consumption of the host for data processing plus the energy required by the electronics of the WNIA (assuming that there are no other peripheral devices). Optimisation of the energy consumed by the host is a separate research topic and it is not considered here. The energy consumption of the network adapter depends on two factors: the typical energy consumption of each of its hardware components and its current state of operation. Typically, a WNIA could be found in the following four states: *transmit*, *receive*, *idle* and *sleep*. The states of operation determine the instantaneous power consumption of the device as each requires a particular number of active hardware components. Furthermore, each of the components involved have different energy consumption requirements that depend on the functions executed in a given state.



(a) Sleep

(b) Idle

(c) Receive

(d) Transmit

Figure 3.3: Energy consumption breakdown

Figure-3.3 illustrates which components are functional during the four states and their contribution towards the corresponding energy consumption. Darker component correspond to higher energy consumption contribution while components that are not used are left unshaded. The figure also shows a comparison of the overall energy consumption of the device in each state. Experimental work presented in [59. 60. 61] has unanimously shown that the largest energy consumption occurs during data transmission. This is attributed to the mandatory amplification of the RF signal before it is transmitted though the antenna. Measurements during data reception have shown that it is on average 25 to 45 percent less energy expensive than data transmission. The active components during reception, as shown by Figure-3.3(c), are the IF modem, the IF-RF converter and the MAC controller. Energy consumption measurements have also suggested that the difference between the consumed energy during idle periods and that during reception is marginal. The reason being that most of the electronics required for reception is active during idle states for the purpose of channel scanning. As explained in [60], the small difference in energy consumption is introduced by the MAC controller which is less active during idle periods. It is important to note the low energy consumption during the sleep state. Emphasis should be placed on the opportunity that it provides for energy conservation. Sleep state, however, is often responsible for unreliable communication coordination. This is a result of the suspended operation of the WNIA during which channel information is unavailable. To that end, its use in environments of asynchronous data transmissions requires some form of transmission planning.

3.4 Scenarios of operation and the model of Nilsson

To capture the energy consumption profile of an ad hoc network, one has to establish its dynamics. Network dynamics representation could be done through description of all possible *scenarios* in which a node could be found. The scenarios are therefore a product of the network topology changes and the procedures executed by the governing protocols. In other words, the scenarios are a permutation of all possible locations a node could have with respect to the rest of the nodes and all possible states of operation in which it could be found.

As described by Nilsson *et al* [62], in any instance of data transmission in a mobile ad hoc network, a node could be found in any of the following four situations: in the vicinity of the source node, in the vicinity of the destination node, in the vicinity of both, or out of range of both. To capture the full set of possible states of operation, those described in Section 3.3 have to be further scrutinised and modified as follows:(a) the wireless medium allows for broadcast and unicast transmissions hence the transmit state is divided into *broadcast transmit* and *point-to-point transmit* and (b) the receive state is divided into *receive* - corresponding to reception of dedicated data, *promiscuous receive* - corresponding to overhearing and processing of all data transmissions and *discard* - corresponding to traffic overhearing but processing only dedicated data packets. The combination of all states and positions result in the following operation scenarios:

- broadcast transmit
- point-to-point transmit
- broadcast receive
- point-to-point receive
- promiscuous receive in the vicinity of the source
- discard in the vicinity of the source
- promiscuous receive in the vicinity of the destination
- discard in the vicinity of the destination
- idle

Every state (part of a scenario) requires execution of specific functions by the WNIA. As a result each scenario is associated with finite amounts of consumed energy. Knowledge of the WNIA components involved in the different states and their energy consumptions provides a way for quantification of the energy consumption of the scenarios. Estimations of the order and frequency of scenario occurrence in a particular ad hoc network allows for its energy consumption analysis. For example, in a mobile ad hoc network that operates with on-demand routing, the energy consumption profile will be shaped by the frequent broadcast route request transmissions. These broadcasts will be prominent in a network of highly mobile nodes where link breakages are frequent. In addition to that, it is a standard technique to use packet overhearing as a method of improving routing efficiency therefore allowing the state of *promiscuous receive* to affect the power consumption profile as well. The choice of a medium access control protocol also plays a significant role in state ordering. This factor is influenced by the mandatory procedures imposed by the MAC protocol for the purpose of channel acquisition and collision avoidance and by its channel arbitration efficiency.

Quantification of consumed energy in the duration of the scenarios can be obtained through a model proposed in [62]. The model estimates the consumed energy per processed (transmitted or received) data packet in the context of each scenario with the help of the following linear equation :

$$P = m \times size + b \qquad (3.1)$$

The equation consists of two components: a fixed component b and an incremental component $m \times size$. The latter reflects the energy cost for processing of a data packet. The value of m represents the energy consumed per byte in a particular state. It is directly related to the hardware specifications of the WNIA. Minimisation of the values that it takes can be achieved through one or more of the following actions: use of energy efficient hardware; use of sleep mode or/and control of the signal amplification before transmission. The value of $size$ represents the length of the processed packet in bytes. The component b is the fixed energy costs of MAC communication overhead generated for successful packet transmission. It is in the form of once-off channel acquisition and collision avoidance control message exchange and mandatory frame headers. It also represents the efficiency of the MAC in terms of generated communication overhead. The model, however, does not capture the per-packet energy penalties attributed to incurred overhead due to inefficient channel use. It will be shown that such communication overhead has considerable impact on energy efficiency and other performance characteristics as it includes transmission of redundant data.

The analysis presented in [62] quantifies the power consumption of a generic IEEE 802.11b WNIA in the context of the identified scenarios. The values of m and b (tabulated in [62]) are obtained through measurements. They are then combined for comparison of the per-packet energy consumption of the corresponding scenarios.

The model suggests that *point-to-point transmit* is the most energy expensive operation. The reason being that the transmit state has the highest energy consumption with corresponding high values of m. In addition to that point-to-point transmission takes place with the full MAC communication overhead so it has a high fixed cost as well.

The second most energy expensive scenario is *broadcast transmit*. It differs from point-to-point transmit only by the lower fixed cost as broadcast data packets are not transmitted with channel reservation and collision avoidance MAC communication overhead.

Instantaneous energy consumption of data reception is less than that of data transmission. That is evident from the smaller values of m in receive state. However, it could easily become an expensive process when *promiscuous data reception* is allowed whereby the frequency of the reception scenarios would significantly increase. Alleviation of that cost could be achieved by using *discard reception*. Although irrelevant data packets are ignored in discard mode, reception and processing of control packets is still mandatory. In actual fact, discard mode only shortens the time a WNIA spends in receiving. If, however, in the duration of the saved time the device is placed in an idle state, the energy consumption efficiency would not change significantly. The reason being the marginal difference in energy consumption between reception and idle states.

Energy consumption investigations summarised by the literature survey in Chapter 2 indicated that energy consumption optimisation techniques relying on transmit power control are inefficient. Further investigations were required to establish the validity of that statement. The model of Nilsson discussed in this section and the analytical results that it provides suggest that transmit power regulation is relevant only to a third of the possible operation scenarios. Therefore it confirms that such techniques are incapable of delivering optimal power efficiency on their own. Similar observations regarding the fixed component of MAC communication overhead indicates that it affects all of the functional scenarios. Further investigation is required to analyse and quantify the effect of such overhead on energy consumption. Congestion related communication overhead has also to be accounted for. The limitations of the model of Nilsson does not allow such analysis, it therefore requires modification.

3.5 Congestion overhead and extended model of Nilsson

The energy model of Nilsson has shown that communication overhead affects the energy consumption in all scenarios. It, however, does not account for overhead such as retransmission of data, control and routing packets. In mobile ad hoc networks of moderate mobility this overhead is a result of congestion and therefore it is considered here as *congestion related communication overhead.* Unable to account for it, the model from [62] is true only for ideal networks where such communication overhead do not exist.

A network is said to be congested when the offered load has placed it in a state of *saturation throughput* and further increase of the load results in worsening of the network operation. It has been concluded in [63] [64] that in the case of contention based medium access control mechanisms the degraded network operation is attributed to high numbers of packet collisions. Hence it is said that congestion results in high collusion probability.

Every successfully transmitted unicast packet of payload data is responsible for the transmission of additional data which includes: routing, control data for channel acquisition and collision avoidance, frame header information and its own unsuccessful transmission attempts. Most of it is essential for reliable operation of the network. However, as it will be shown, inefficient channel access arbitration often leads to transmission of excess communication overhead.

Point-to-point data transmission is initiated with route discovery if a route to the desired destination does not already exist. This is followed by channel reservation and collision avoidance procedures (such as RTS-CTS handshaking), payload data transmission and finally confirmation of data reception by the destination node. Packet collisions in a congested network result in transmission failures. Every unsuccessful transmission attempt fails either because the channel reservation does not succeed or because the transmitted payload data is not acknowledged by the receiver. It is standard procedure to retransmit failed control messages and payload data until the transmission is successful or a certain threshold number of attempts is reached. These retransmissions constitute redundant communication overhead that has an energy consumption cost proportional to the size of the control and payload data and number of transmission attempts. They also contribute to worsening the congestion condition of the network. The consumed energy as a function of procedural steps, packet size and transmission attempts is depicted by Figure-3.4. In the diagram, unsuccessful packet transmission is shown by a dashed box. With an effective

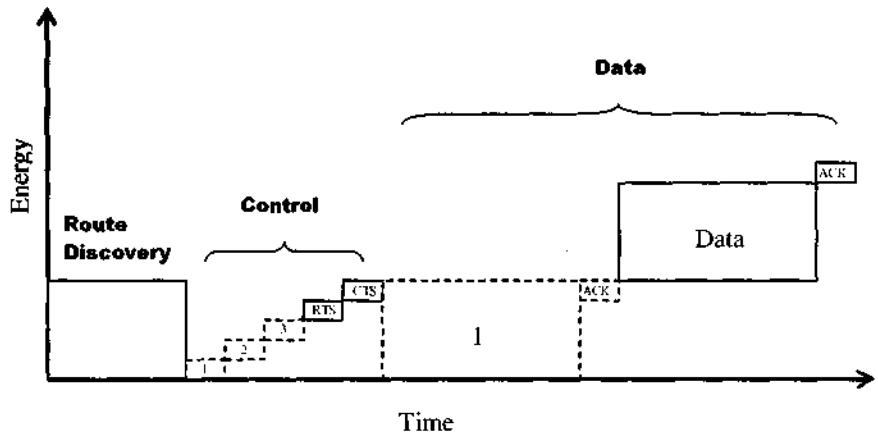


Figure 3.4: Point-to-point data transmission

channel reservation and collision avoidance procedure in place, the retransmissions are more likely to occur during the channel reservation phase. This is favourable from an energy consumption prospective as control packets are smaller than the payload data to ensure bandwidth efficiency. Nevertheless there is always a finite probability that a data packet will collide or acknowledgment will not be received even after a successful channel reservation. Furthermore, channel reservation is carried out only with unicast traffic.

Every transmission of data packets is started with route discovery unless a route has already been established between the source-destination pair. Every time a transmission retry threshold is reached, a route maintenance procedure is initiated. The impact of routing communication overhead on energy efficiency is analysed in [65]. The proposed analytical model shows that the larger the network, the smaller the probability that an established route will be reused under dynamic topology changes. The analysis also concludes that in the cases of both, reactive and proactive routing, the dominant energy consumption in the network is associated with transmission of route discovery control packets.

The discussion so far provides adequate reasoning as to why congestion related communication overhead has to be considered. To account for such overhead the model proposed by Nilsson has to be extended to include the energy consumption of redundant data. To achieve that, the author of this thesis has proposed that the per-packet energy should be represented as:

$$E_{pkt} = \sum_{0}^{k_1} E_{data} + \sum_{0}^{k_2} E_{control} + E_{routing} \tag{3.2}$$

The per-packet energy is the sum of the energy consumed for transmission of the packet itself and all of the associated redundant data. This includes the sum of the energy consumed for retransmission of the payload data packet, transmission and retransmission of control data and the energy required for route discovery. The number of retries is represented by k_1 and k_2 for payload data and control packets respectively. The energy associated with each processed unit of data is found by use of Equation (3.1).

Under the assumption that perfect channel reservation is in place providing collision free payload data transmission, Equation (3.2) can be simplified to the following:

$$E_{pkt} = E_{data} + \sum_{0}^{k} E_{control} + E_{routing} \tag{3.3}$$

The extended energy model accounts for energy consumption related to congestion generated overhead through k which represents the number of retransmissions of MAC control messages. As in the case of [63] and [64], quantification of k is established by linking it to congestion through collision probability. The problem can be defined as follows: how does the number of data retransmissions change with varying collision probability.

The probability of a collision for a given packet, as seen by the transmitting node, is the probability that in the duration of the same transmission slot at least one of the $n-1$ remaining nodes will initiate a transmission. If a node has a probability of transmission τ then according to [64] the collision probability as seen by the transmitter is:

$$P_c = 1 - (1-\tau)^{n-1} \tag{3.4}$$

The derivation of the transmission probability τ is subject to the chosen medium access control mechanism. A detailed derivation of τ in the case of IEEE 802.11 MAC can be found in [64].

Establishing the relation between the number of data retransmissions and collision probability is done with the help of successful transmission probability P_t. According to [63], it is possible to approximate the number of packet retries k as a geometric distribution with a probability function the probability of successful data transmission which is defined in [64]:

$$P_t = \frac{\tau n(1-\tau)^n}{1-(1-\tau)^{n-1}} \tag{3.5}$$

P_t is the ratio between the probability of packet transmission and the probability that only one node is transmitting in the slot. The geometric distribution illustrates how the probability of a successful transmission decreases as the number of retries k increases. The distribution is shown by Figure-3.5(a).

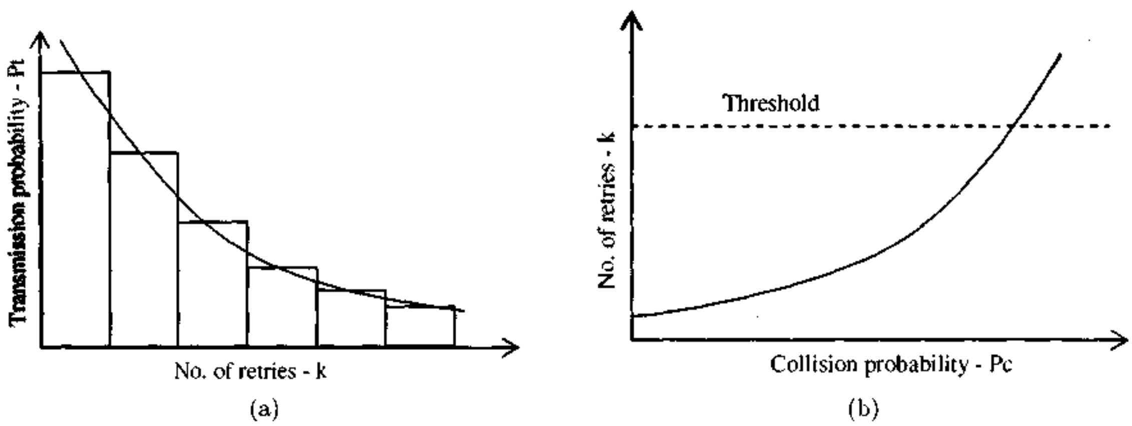


Figure 3.5: (a) Probability of transmission as a geometric distribution; (b) Number of retries as function of collision probability

In a congested ad hoc network operating under moderate mobility (1 m/s to 5 m/s), an assumption could be made that the probability of a successful transmission is roughly the same as the probability that a collision would not occur, in other words:

$$P_t = 1 - P_c \tag{3.6}$$

or in terms of collision probability

$$P_c = 1 - P_t \tag{3.7}$$

If the probability function of the geometric distribution from Figure-3.5(a) is changed from P_t to P_c according to Equation (3.7), then the number of retries can be represented as a function of collision probability as initially intended. That resulting dependence is shown in Figure-3.5(b). It describes how as congestion in an ad hoc network leads to increased collision probability, the number of retries increases as shown by Figure-3.5(b). That in turn affects the energy consumption efficiency by increasing the number of retransmitted data as described by Equation (3.2).

Figure-3.5(b) shows that at high collision probability the threshold of retries is reached more frequently. At that point the retransmission attempts are suspended. However, the routing protocol is triggered since it is assumed by the MAC protocol that the route to the destination is lost. In a congested ad hoc network of moderate mobility, this assumption is often wrong as the packet transmission is unsuccessful due to collisions rather than link breakages thus the routing action is unnecessary. As established by [65] route discovery has a profound effect on energy efficiency. Route request packets are never transmitted with channel reservation and collision avoidance procedures as they require broadcasting. Packets of this type are also disseminated through network flooding unless an heuristic is used for more efficient route discovery. Flooding turns all of the receivers of the broadcasted route request into sources for the following wave of broadcasting leading to a broadcast storm. To that end, routing significantly affects congestion by increasing the potential data sources represented by n in Equation (3.4) and therefore the number of retries k. Additionally the lack of collision avoidance procedures contributes to high collision probability as established in [64].

The routing activity and the high collision probability form an avalanche effect whereby, with higher collision probability, retry limits are reached more frequently and therefore more unnecessary route discovery is carried out. That in its turn increases congestion, collision probability and energy consumption. These analytical results are displayed in Section 3.6 and confirmed through simulation, the results of which are presented in Section 5. The results suggest that the reason for such unstable network behaviour is the MAC layer which does not minimise the occurrence of broadcast storms and does not differentiate between the two reasons for unsuccessful data transmission - packet collisions and link breakage.

3.6 Numerical Analysis

With the help of iterative numerical solutions of the expressions for *congestion*, *collision probability* and *energy consumption*, the relationship between the three is illustrated in this section in terms of graphs.

So far analysis has shown that high network load combined with frequent route discovery leads to congestion. Unrestrained routing procedures lead to broadcast storms with adverse effects on the network performance. According to [7], the effective coverage area of subsequent packet broadcasting, used with network flooding, depends on the location of the transmitting node with respect to the node from which the packet was received. Calculations in [7] of the effective area, represented by the grey semicircle in Figure-3.6, have suggested that on average it increases by 41 percent with every re-broadcast. Assuming

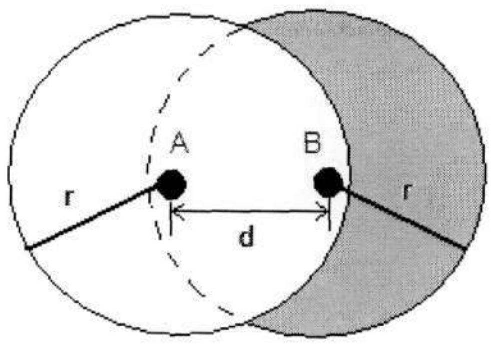


Figure 3.6: Effective area of subsequent broadcasts [7] (A-source node, B-forwarding node)

41 percent area coverage efficiency of the flooding mechanism, the increase of the number of contending nodes per broadcast is estimated in [7] by the following equation:

$$n = 0.41\rho\pi r^2 \tag{3.8}$$

where r is the transmission range of the nodes and the node density ρ is found as:

$$\rho = N/\pi R^2 \tag{3.9}$$

In Equation (3.9), N is the number of nodes and R is the radius of the area covered by the network.

Figure-3.7 shows how re-broadcasting increases the number of contending nodes in the network. It illustrates an example of flooding taking place in an area of 250 by 250 metres populated with uniformly distributed nodes. Their transmission range is set to r = 125 m. Numerical estimations are carried out for networks with populations in the range 50 to

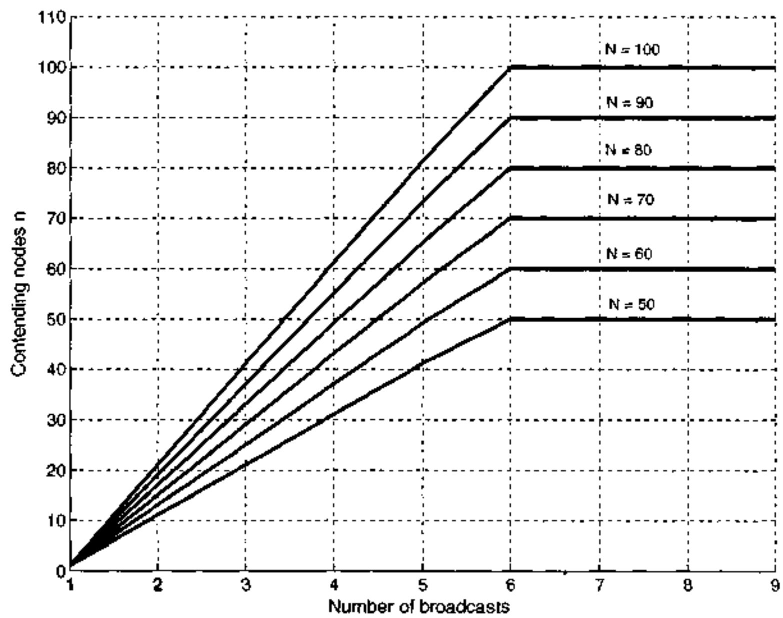


Figure 3.7: Increase of contending nodes as a function of broadcast retransmissions

100 nodes. The results show that under uniform node distribution, the size of the network (in terms of number of nodes) does not affect the number of broadcasts m required to turn all of the nodes to potential contenders ie. $\sum_1^m n = N$. Instead, that number is a function of the transmission range since the larger it is the fewer steps it would take to reach the $\sum_1^m n = N$ condition.

Numerical solutions of Equation (3.8), depicted in Figure-3.7, establish that only a small number of re-broadcasts of a data packet is required to turn all network nodes to channel access contenders. Under the $\sum_1^m n = N$ condition, collision probability is high. The relation between collision probability and the number of contending nodes is defined by Equation (3.4). Graphical representation of collision probability as a function of increasing number of contending nodes is provided by Figure-3.8(a). It shows that under the condition of $\sum_1^m n = N$, the only factor that could improve collision probability is increased contention window[1] W. However, such increase has a negative effect on performance characteristics such as throughput and end-to-end transmission delays.

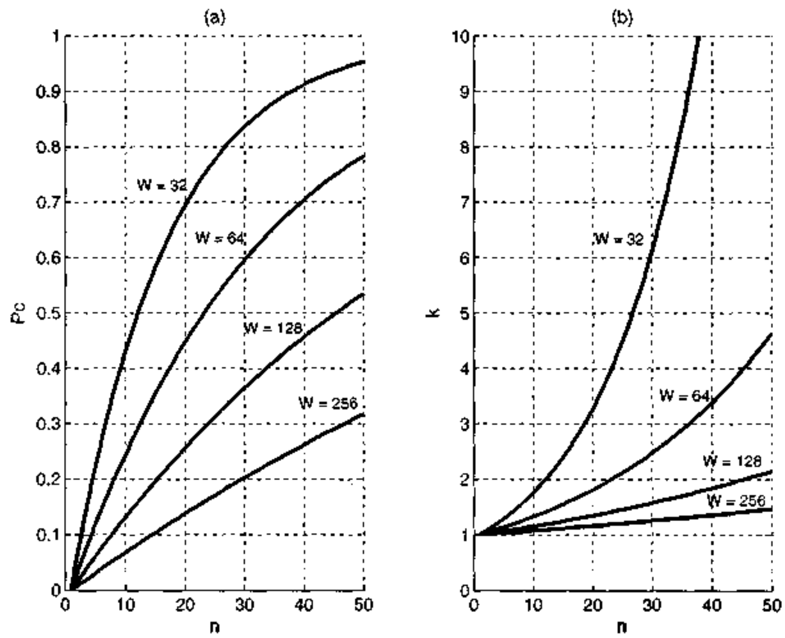


Figure 3.8: (a) Collision probability vs. number of active nodes; (b) Number of retransmissions vs. number of active nodes

Figure-3.8(b) represents the average number of retries as a function of contending nodes and W. The graphs are obtained with use of Equation (3.4) and the true relationship between k and Pc for the contention based $CSMA\backslash CA$ as derived in [64] which is:

$$k_{CSMA\backslash CA} = 1/(1 - P_c) \tag{3.10}$$

Finally, the relation between congestion, collision probability and energy consumption is established. Broadcast transmissions turn nodes into contenders for the channel, which increases the collision probability leading to a high number of data retransmissions. Figure-

[1]a contention window is the range from which random numbers are taken for the back-off procedure described in Section 2.3.1.6

3.8(b) proves that under inefficient medium access control, where the broadcast storm is allowed to occur, the number of retransmissions could rapidly reach the threshold values. The effect of high transmission attempts on energy efficiency is depicted by Figure-3.9. The graphs are based on numerical solutions of Equation (3.2) under the assumption of an ideal channel acquisition mechanism and therefore zero collision probability for payload data packets. Figure-3.9 shows the percentage increase of energy consumption as a function of

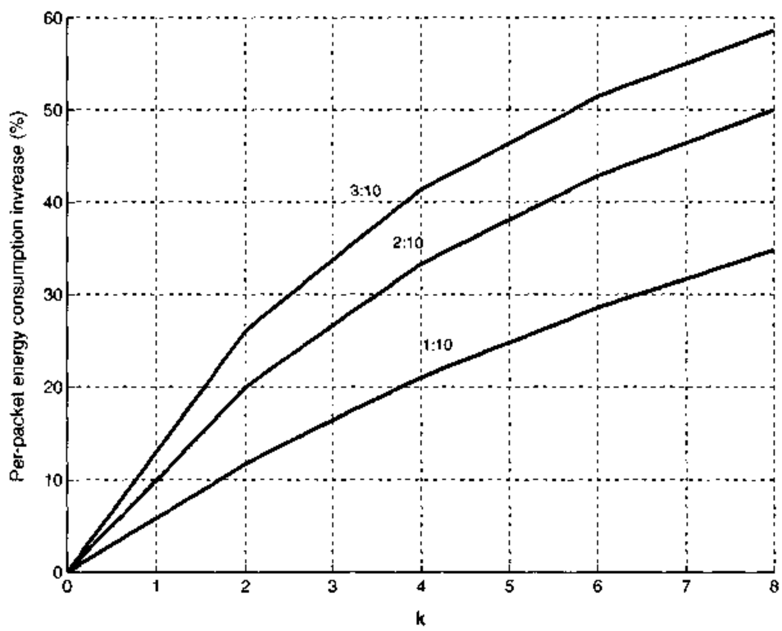


Figure 3.9: Point-to-point data transmission

control message retransmissions k. Three graphs are plotted for different ratios of control packet to data packet size. The analytical results suggest that the energy consumption could increase by up to 60% as a result of data retransmissions. These results account only for retransmissions of small MAC channel reservation control messages. It therefore can be concluded that in realistic cases where collision of payload data does occur, the energy consumption inefficiency could be expected to be higher.

3.7 Summary and conclusion

This chapter provided an analysis of the energy consumption of mobile ad hoc networks. It started giving a generic explanation of the hardware of mobile ad hoc network nodes.

It then described the typical states of operation with regards to communication. It was illustrated how, depending on the executed functions, the states of operation have different energy consumptions.

Further energy consumption analysis was carried out by investigating the network dynamics. The states of operation were combined with topology dependant factors to from a set of scenarios of operation. The energy consumption in each scenario was analysed with the help of the model presented in [62]. The analytical results suggested that the point-to-point data transmission process has the highest energy consumption in ad hoc networks. The reason for that being the transmit state which requires the the largest amount of energy compared to the other states of operation. In addition, the process includes the full channel reservation control communication overhead.

The presented energy model based on [62] led to a number of conclusions. It confirmed that energy consumption optimisation based solely on efficient use of transmit power does not provide optimal results. Such can only be achieved under the condition that excess communication overhead is addressed.

To investigate the effect of communication overhead on energy consumption efficiency, the model from [62] was extended to include congestion related overhead. With the help of the extension, the model accounts for transmission of redundant data in the form of packet retransmission attempts. The extended model describes how high collision probability relates to congestion which leads to a high number of transmission retries. An observation of the analytical results indicates that the MAC protocols trigger unnecessary route maintenance procedures once the number of retries reaches the threshold value. This results from the assumption that the route in use is invalid while often the actual reason in congested networks is the high collision probability. Such redundant routing procedures reinforce the congestion leading to an avalanche effect.

Numerical analysis of the relationship between congestion, collision probability, retransmission attempts and energy consumption has provided quantification of the energy consumption factors involved. It was shown in Section 3.6 that broadcast storms increase the number of contending nodes, leading to congestion associated with high collision probability. The resulting high number of packet retransmissions is responsible for the transmission of redundant data. That in turn could cause an increase in energy consumption by as much as 60%.

The analysis in this chapter has provided enough evidence to concluded that inefficient energy consumption in ad hoc networks can be attributed to inefficient medium access control. As investigated, in the case of the contention based MAC protocols, the determining factor is their inherent property of packet collisions and unstable behaviour. It has been discovered that in networks of moderate mobility the congestion could often be reinforced by the inability to differentiate between the two reason for transmission failures - link breakage and packet collision. Such differentiation by the medium access control is not a trivial task and therefore other techniques have to be applied to curb the generation of excessive communication overhead and inefficient energy consumption. This is addressed in the following chapter, Chapter 4, in which a hybrid between *contention* and *allocation* channel access control is investigated.

Chapter 4

PA-STDMA MAC Protocol

4.1 Introduction

The primary advantages of *contention* based medium access control algorithms are their mobility-independent operation and ease of deployment. However, a number of publications [66] [64] have shown that contention based MAC protocols operate efficiently only in networks with light to moderate load conditions. Under conditions of high to heavy network load, the protocols spend most of their time resolving collisions caused by factors discussed in Chapter 3. To avoid network instability and to curb congestion under high network load, the contention approach can be partially or fully replaced by an allocation based channel access mechanism such as TDMA.

In a pure TDMA network each node is assigned a unique transmission slot that repeats in time. This leads to guaranteed collision free and topology transparent communication. The collection of all transmission slots form a cyclic frame thus allowing for a deterministic time lapse between consecutive transmissions and bounded end-to-end transmission delay. The main disadvantages of the traditional TDMA approach is its poor *spatial reuse* and its inefficient operation under light traffic conditions. Spatial reuse is said to occur when two or more transmissions take place at the same time but do not interfere with each other as a result of spatial separation. In a pure TDMA network this does not take place due to the strict transmission slot policy. Strict scheduled transmission is also the reason for the inefficient operation of TDMA under light load conditions where, transmission slots are left unused by nodes that do not have data for transmission.

For the purpose of optimal TDMA performance in an ad hoc network environment, a number of publications [67] [68] [69] [70] [66] [71] [72] have proposed different forms of dynamic transmission slot scheduling. In some of these cases the process requires information about the schedules of neighbouring nodes located up to two hops away. Protocols that form schedules on the basis of such information are referred to as *topology dependent* protocols. The communication overhead related to the process of schedule formation and schedule maintenance in topology dependent protocols could have a negative effect on the operation of the network. Thus algorithms such as in [70] [66] [71] [72] attempt to achieve reliable and efficient transmission schedules in a topology transparent manner. The objective of the research work presented in this thesis is the achievement of power consumption efficiency through minimisation of energy expensive overhead, hence topology dependant TDMA protocols are not of interest.

Two popular topology transparent techniques are the *probabilistic* and the *deterministic* scheduling policies. Under the latter, each node transmits in the duration of slots determined by the roots of pre-distributed polynomials. Each polynomial has an unique set of roots so that collisions are avoided. The probabilistic approach is an attempt to improve on the performance of the deterministic policy. It allows the nodes to transmit with a certain probability in slots that are not designated by the roots of their polynomials in case they have data for transmission. These two methods are described in detail and investigated in [73].

A novel approach to optimised performance of ad hoc networks under a wide range of loading conditions and node densities, while maintaining low congestion and mobility independence, is the ADAPT MAC protocol proposed in [66]. The improvement is achieved by combining the contention policy with TDMA. Under high load most of the transmission slots are used by the nodes so the network operates under the rules of a traditional TDMA system. In cases of light network load, contention is allowed so nodes with data for transmission can make use of unused transmission slots. ADAPT operates with pre-distributed transmission schedules and thus mobility and topology changes do not affect its operation.

Pre-distribution of any information which is required for the network operation renders the network unscalable. While nodes can freely leave the network, no nodes other than the ones that have the required pre-distributed information can joint and use it. In other words, under ADAPT, nodes that do not have the pre-distributed transmission schedule would be unable to join the network. This does not hinder the operation of instances of ad hoc networks that require only predetermined number of nodes. However, it is against the ad hoc concept (introduced in Chapter 1). In addition to that, ADAPT inherits the broadcast storm problem described in Chapter 3 as it does not put a constraint on the number of contending nodes.

To address the adverse effects that congestion has on various network performance characteristics including energy consumption, the author has proposed a novel MAC protocol. It combines TDMA and contention in a way similar to ADAPT while addressing some of its shortcomings. The name of the protocol is Position-Aided Spatial-TDMA MAC and its detailed description and analysis is the focus of this chapter.

This chapter is organised as follows: Section 4.1.1 provides a brief conceptual discussion of the proposed protocol; Detailed description of PA-STDMA MAC is the focus of Sections 4.1.2, 4.1.3, 4.1.4; Performance analysis of the protocol is the topic of Section 4.2.

4.1.1 Overview of PA-STDMA MAC

PA-STDMA is a topology transparent medium access protocol that combines TDMA allocation with contention in a manner similar to ADAPT from [66]. It is designed for ad hoc network environments in which positional information is available. The difference between ADAPT and PA-STDMA is that PA-STDMA uses real-time transmission slot allocation and thus pre-distribution of schedules is not needed. This makes the network accessible at any time and the protocol compliant with the ad hoc idea. In comparison to a traditional TDMA MAC, PA-STDMA is scalable and is expected to have improved transmission latency in large ad hoc networks. The proposed protocol does not require dissemination of any global or local positional or scheduling information hence it is topology transparent. The nodes of the network use information regarding only their own location for the purpose of transmission scheduling. Positional information is obtained from GPS or any other positioning system. To achieve the position-aided scheduling, the protocol splits the network environment into geographical cells and assigns each cell a transmission slot. Similar to STDMA MAC from [70], the nodes of the network are allowed to transmit in the duration of the slot that corresponds to the cell in which they are located. In contrast to it, improved throughput efficiency is sought by allowing the nodes to transmit during a number of slots that belong to neighbouring cells. Such transmissions, however, assume lower priority and obey strict transmission rules in order for structured channel use and efficiency to be maintained. Through planned area partitioning and slot assignment, the protocol encourages spatial reuse where non-interfering simultaneous transmissions take place.

The cells, as defined by the protocol, are expected to contain more than one node at a time. The number of nodes in a cell is dependant on node density and cell size. To use the transmission slot that corresponds to a cell, co-located nodes undergo a contention process. This eliminates the need for transmission sequence lists as suggested in [70] and avoids the potential communication overhead associated with it. Contention, however, does not guarantee collision free communication. With the help of controlled cell size and thus controlled number of contending nodes, collisions are expected to be minimal.

The existence of position-based transmission schedules allows use of low power mode (sleep) without significant disruption of communications in the network. The use of sleep mode is determined as a trade-off between energy consumption efficiency and availability of the nodes.

The PA-STDMA MAC is designed for structured channel use and hence network operation characterised by low collision probability. This is expected to result in minimised transmission of redundant data and thus efficient bandwidth use and improved energy efficiency. Furthermore, the protocol minimises the use of RTS-CTS handshaking overhead by avoiding the hidden terminal phenomenon through constraints on the minimum distance between nodes which are allowed to use identical time slots as part of the spatial reuse planning. In the case of PA-STDMA, spatial reuse is a product of the design of grids that cover the area of the network. A detailed description of that design follows.

4.1.2 PA-STDMA - Cell Structure and Scheduling

Data transmission under PA-STDMA MAC takes place during a transmission slot. The slots are grouped in a frame which repeats over time. The frame structure is shown in Figure 4.1.

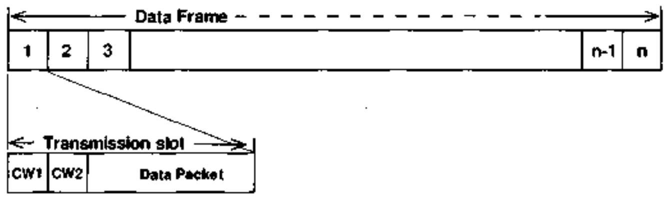
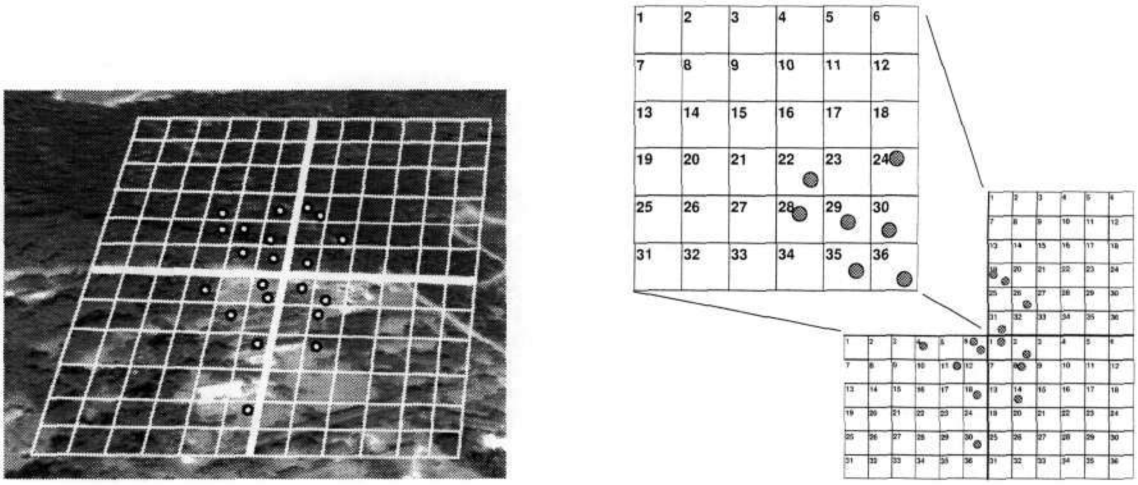


Figure 4.1: Data frame configuration (CW - contention window)

The main feature of the protocol is its transmission slot assignment technique which is based on spatial position. The protocol maps all transmission slots to geographical locations. To this end, the protocol divides the space at the location of the network into geographical regions referred to as partitions. The partitions are identically subdivided into cells. A unique transmission slot is assigned to each cell in a partition. Figure 4.2(a) shows the space partitioning over the area of the network while Figure 4.2(b) shows the slot assignment to each cell in the partitions. The number of cells across all partitions is the same. As a result, the number of transmission slots is also the same, forming a data frame of universal length. A geographical partition in the spatial domain directly maps to a transmission frame in the time domain. Hence a transmission frame has as many transmission slots as there are cells in the geographical partition.

The number of cells used in a partition depends on the choice of cell size and the hidden terminal avoidance rule. It stipulates that cells forming two different partitions with identical time slot allocations must be at a minimum distance of twice the maximum transmission range of the participating nodes as shown in Figure 4.3. The rule is expressed as the following equation:

$$S = \lceil \frac{2R}{M} \rceil \tag{4.1}$$



(a) (b)

Figure 4.2: (a) Spatial network partitioning; (b)Slot assignment

where M is the number of cells between two cells with the same slot assignment, R is the transmission range of the WNIA and S is the side length of the square cells in the grid. Equation (4.1) binds the cell size and the number of cells with the help of the transmission range characteristic of the nodes. If a reduction of the cell size is required Equation (4.1) gives the number of required cells between two cells with the same transmission slot so that simultaneous transmissions are kept non-interfering.

The full NxN grid of cells (ie the partition) is determined by the number of cells ,M, found between any two cells with the same slot assignment as follows:

$$N = M + 1 \tag{4.2}$$

Constructing the grids as defined by Equations (4.1) and (4.2) guarantees non-overlapping simultaneous transmissions and thus the hidden terminal interference problem is avoided.

Figure 4.3 shows two grids with cell configuration corresponding to a transmission range of $R = 250\ m$ and cell size of 100 x 100 m^2. Under these conditions, the configuration requires a space of at least five cells between any two cells of different partitions with identical slot assignment. The optimal solution for this particular case is a 6x6 cell grid. The choice of cell size is a trade-off between transmission delay and improved contention

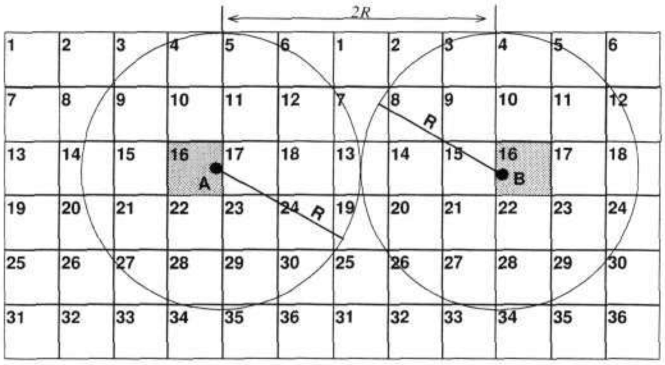


Figure 4.3: Hidden terminal avoidance rule

process. Reducing the cell size will reduce the number of contending nodes for a particular transmission slot. The distance between cells with identical slot assignments has to be kept to at least 2R hence, reduced cell size would necessitate more cells in the grid as defined by Equation (4.1). This will result in longer data frames and therefore longer transmission delays in moderately to heavily loaded networks.

Figure 4.4 shows how the design of the space partitions results in simultaneous transmissions taking place without interference. The depicted terminals positioned at three different locations occupy the same cell 1 in three different partitions. As a result they transmit during transmission slot 1. However, because they are spaced at least twice their transmission ranges apart, the transmissions do not collide.

Figure 4.4: Spatial reuse of bandwidth

If a node has data to transmit, it must be aware of its position which could be obtained using a GPS or any similar positioning system. Based on its position and with the help of placement procedures provided by the protocol, a node determines to which partition and cell it belongs and therefore which time slots it is allowed to use. Experimentation results in [74] reveal that GPS achieves absolute timing accuracy of about 100ns and relative timing resolution of about 10ns. Since the period of the time slots is in the range of milliseconds, it is concluded that the timing accuracy provided by GPS is more than adequate for time synchronisation purposes.

4.1.3 PA-STDMA - Data Transmission Rules

According to the basic transmission rule, nodes of the network transmit during the transmission slot that belongs to the cell in which they are located. That slot is referred to as a *primary* transmission slot. It is possible for a cell to have a single occupant. However, it is expected that most of the time the geographical cells will be populated with more than one node. This necessitates a mechanism for transmission slot use arbitration between co-located nodes. For this purpose, PA-STDMA uses a contention process similar to that used by IEEE 802.11 MAC [2], which is based on the popular *Backoff* technique. At the beginning of each time slot, all contending nodes choose a random number from a predetermined range and start decrementing it. The first node that reaches zero is the node that acquires the channel and initiates transmission. While decrementing, each node performs carrier sensing. If a transmission is detected, the nodes terminate the countdown and wait until the next transmission opportunity. The contention process takes place at the beginning of the transmission slot during the designated contention windows (CWs).

For efficient channel use, PA-STDMA allows the nodes to contend and transmit during a number of slots that belong to adjacent cells under the strict condition that these slots are not used. A slot that belongs to a cell adjacent to the one in which a given node is located is referred to as *secondary* transmission slot. All possible slot types and their relative cell mappings with respect to a given node are depicted in Figure 4.5.

The primary transmission slot of the node located in cell fifteen is transmission slot fifteen. The node has eight secondary slots that correspond to all adjacent cells and eleven idle slots which could be used for low power mode. The choice of eight secondary slots is based on the rule that a node has to cover the cells with its transmission range in order

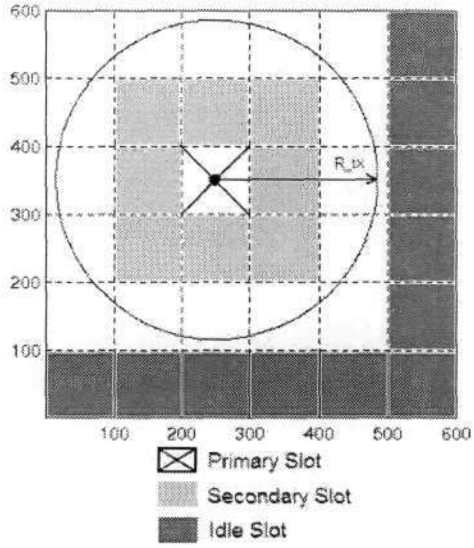


Figure 4.5: Slot types and their relative positions

to contend for the transmission slots that correspond to them.

A node attempting to transmit in a primary slot (in other words in its own slot) has higher priority over the channel than a node attempting to transmit in a secondary slot. To secure that priority, the contention process takes place over two contention windows. Contention for a priority slot takes place during $CW1$ while contention for transmission during a secondary slot takes place during the $CW2$ only if there is no winner in $CW1$.

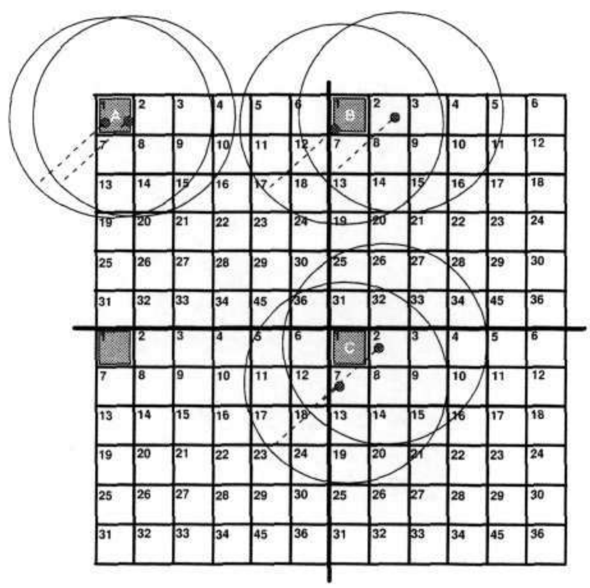


Figure 4.6: Transmission scenarios

With the help of Figures-4.6 and 4.1, the transmission rules can be clarified through an example. The scenario takes place during transmission slot 1 which is assigned to each cell 1 in all partitions. In the case of cell $1A$ in the first partition, the two nodes occupying the cell will contend for transmission in a primary slot. They will do so during $CW1$. In the case of cell $1B$, the node within the cell will acquire the right to transmit after contending during $CW1$. The node located in the cell adjacent to $1B$ will defer its transmission as it will hear the node located in cell $1B$ commencing transmission at the end of $CW1$. In the case of cell $1C$, the two nodes will contend for a secondary slot transmission during $CW2$.

As cells with identical time slots are spaced in such a way as to avoid hidden terminal interference, nodes that transmit during a primary slot do not use the RTS-CTS handshaking procedure. Nodes that transmit during secondary slots violate the distance stipulation as their location does not correspond to the location of the cell to which the time slot is assigned. In that case, RTS-CTS handshaking is used as there is a finite probability that the hidden terminal phenomenon would occur. Such is the case depicted in Figure-4.7.

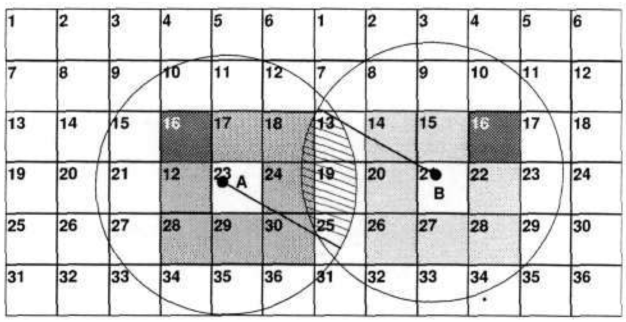


Figure 4.7: Overlapping secondary slot transmissions

Nodes A and B are attempting a secondary slot data transmission during slot 16 as the cells that correspond to that assignment are empty. The distance between the two nodes is less than 2R therefore their transmission ranges overlap producing an area where any data reception is potentially corrupted.

4.1.4 PA-STDMA - Power Management

The PA-STDMA MAC protocol is designed for efficient power consumption based on effective channel use. To that end, minimal control and configuration overhead per bit of successfully transmitted payload data and packet collision rate have been the main design criteria. The only source of overhead incurred by efficient MAC protocols is the RTS-CTS handshaking procedure. To reduce the generated overhead, PA-STDMA MAC reduces the use of the RTS-CTS handshaking by arranging the concurrent transmissions as stipulated by the spatial reuse and hidden terminal avoidance rules. Collision rate is another indication of effective channel use. The protocol is of a TDMA type so it has the inherited property of low packet collision rates achieved by the transmission scheduling approach. Further power efficiency is obtained through use of low power mode. The nodes of the network enter low power mode during slots which appear to them as neither primary nor secondary slots and the associated probability for incoming transmission is low. Slots of this type are referred to as *idle slots*. The extent to which low power mode is used could be determined dynamically. Figure-4.8 shows a node designated with a six-pointed star and dark cells[1] with assignments that appear as idle slots for that node. The

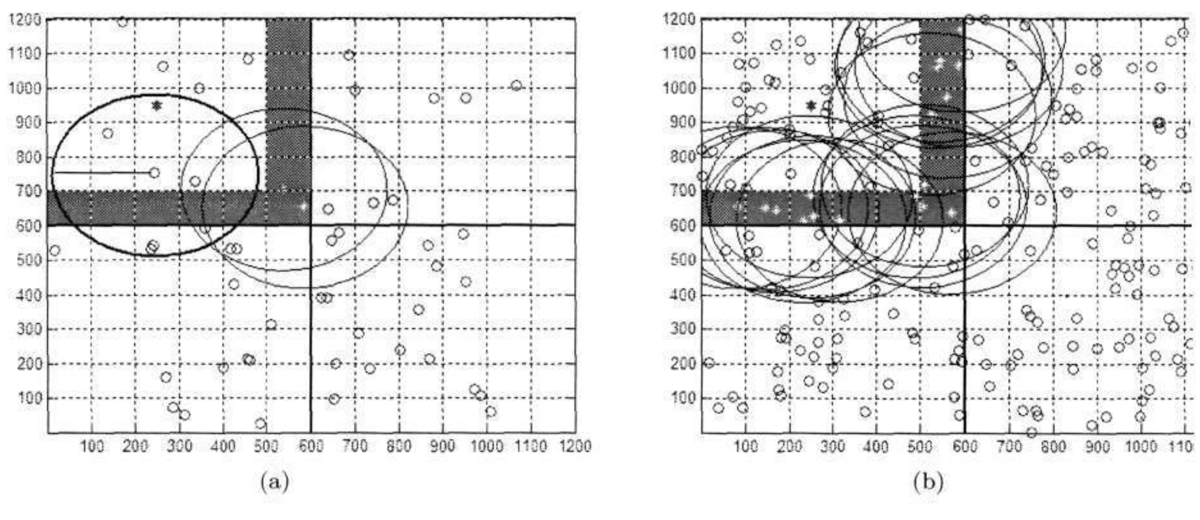


Figure 4.8: (a) Network topology - 50 nodes ; (b) Network topology - 200 nodes

diagrams depict two different node density cases. In Figure-4.8(a) there are 50 randomly distributed nodes whereas in Figure 4.8(b) they are 200. Figure-4.8(a) confirms that if a node is allowed to use its idle slots in an environment with lower node density it will have a greater chance of missing data transmissions than in the case of high node density.

[1] the nodes occupying the dark cells are drawn as asterisks and their transmission ranges are depicted as circles

This is a result of poor cell population leading to higher probability that a node which has the sleeping node in range (such as the node depicted with the thicker transmission radios circle in Figure-4.8(a)) will take over an adjacent slot. In contrast to that, in a network with higher node density the cells that correspond to the idle slots of the remote node will be populated as shown in Figure-4.8(b). In that case the slots assigned to these cells will be used as primary slots resulting in transmissions that do not reach the sleeping node (none of the transmission ranges of the nodes occupying the dark cells cover the six-pointed star node in Figure-4.8(b)). The investigation of the effect of sleep mode merely confirms the fact that in a less dense network the importance of each node, for the purpose of multi-hop communication, increases.

Low power mode is entered not only during slots designated as idle but in any of the remaining slots under the condition that after the contention period no communication (transmission or reception) has been initiated.

4.2 Performance Analysis PA-STDMA MAC

The PA-STDMA MAC protocol has two regions of operation characterised with different performance. The operating point of the protocol is determined by the density of the network. In moderate to high network density, the probability of a cell being occupied by a node is high. This results in predominantly primary slot data transmissions whereby the operation of the protocol closely resembles that of a traditional TDMA scheme. At high node densities the number of nodes in a cell is on average more than one. This leads to contention, but it is confined to a single cell, so the channel use, as stipulated by the TDMA schedules, is maintained. To this end, the protocol is said to operate in the *TDMA* region under moderate to high node density and high cell occupation probability.

On the contrary, under low network density the majority of the transmissions take place during secondary slots. This is always done through contention whereby the slot-cell structure merely defines the available number of secondary slots a node can contend for. The condition of predominantly secondary slot transmissions leads to higher collision probability as the contention process for slots that belong to adjacent cells is allowed over larger areas (ie nodes from a number of cells are allowed to contend). However, the condition of low node density which has placed the operation of the network in this region prevents undesirable high collision probability. The operating region of the protocol associated with low node density and low cell occupation probability is designated as the *contention* region. The two operating regions as a function of cell occupation probability P_{occ} are shown in Figure 4.9.

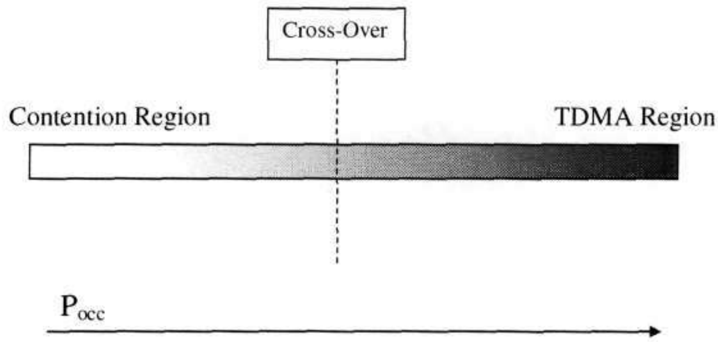


Figure 4.9: Operation Regions of PA-STDMA as function of P_{occ}

P_{occ} is derived as a function of number of cells - N_{cells}(and therefore cell size as defined in Equation (4.1)) and number of nodes in the network - M_{nodes}. For the purpose of the analysis, an assumption is made that the nodes are uniformly distributed in the network.

In the case of a network with only one node, the probability of a cell being occupied is given by :

$$P_1 = \frac{1}{N_{cells}} \tag{4.3}$$

The probability that a cell is empty is:

$$\hat{P}_1 = 1 - \frac{1}{N_{cells}} \tag{4.4}$$

The probability that a cell is empty when there are M_{nodes} number of nodes in the network is :

$$P_{empty} = \left(1 - \frac{1}{N_{cells}}\right)^{M_{nodes}} \tag{4.5}$$

and finally the probability that a cell is occupied is :

$$P_{occ} = 1 - \left(1 - \frac{1}{N_{cells}}\right)^{M_{nodes}} \tag{4.6}$$

or

$$P_{occ} = 1 - \left(\frac{N_{cells} - 1}{N_{cells}} \right)^{M_{nodes}} \tag{4.7}$$

Figures 4.10 and 4.12 depict two different node densities and the corresponding cell occupation probabilities.

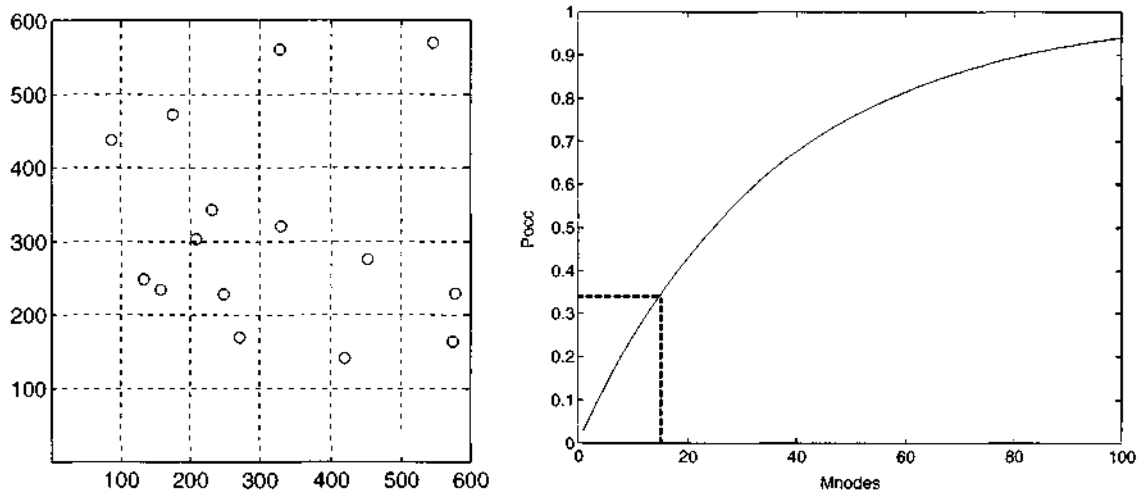


Figure 4.10: P_{occ} as function of node density in 600 x 600 m^2, 15 nodes random topology

In Figure 4.10 the probability of a cell being occupied is estimated by Equation (4.7) to be 0.36. The corresponding node density is 4.17×10^{-5} nodes per square metre or 0.417 nodes in a cell of size 100 x 100 m^2. In that case the network operates in the contention region with the majority of the transmissions taking place during secondary slots. Contention for a secondary slot is allowed over a larger region. However, the low network density keeps the collision probability at moderate levels. Figure-4.11 shows a typical contention region for a secondary transmission slot (provided by the empty dark cell under the node distribution from Figure-4.10) designated as a rectangle spanning nine cells. In the contention region of operation the RTS-CTS handshaking procedure is frequently used. The majority of the expected collisions are associated with the RTS control packets. Low density multi-hop networks require the highest number of nodes available. This is in order to secure connectivity hence low power mode is not advisable in this (contention) region of operation.

In Figure-4.12 cell occupation probability according to Equation (4.7) is 0.94. The node

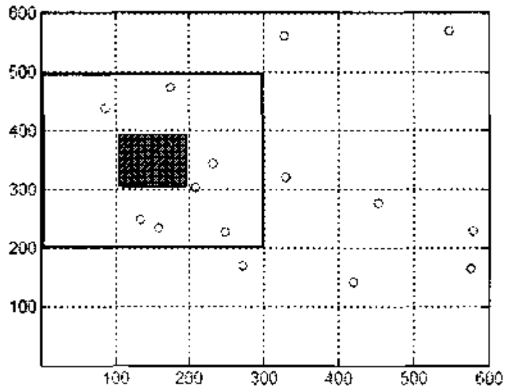


Figure 4.11: Contention area for a secondary transmission

density is 2.78×10^{-4} per square metre or in other words the expectation is for 2.78 nodes per cell. Clearly most of the cells are occupied and therefore transmissions take place during primary slots shifting the operation of the network into the TDMA region. The expected contentions are constrained to take place within the area of a cell. The majority of the transmissions being primary slot transmissions take place without the use of RTS-CTS procedures. In this region sleep mode could be used without affecting the operation of the network.

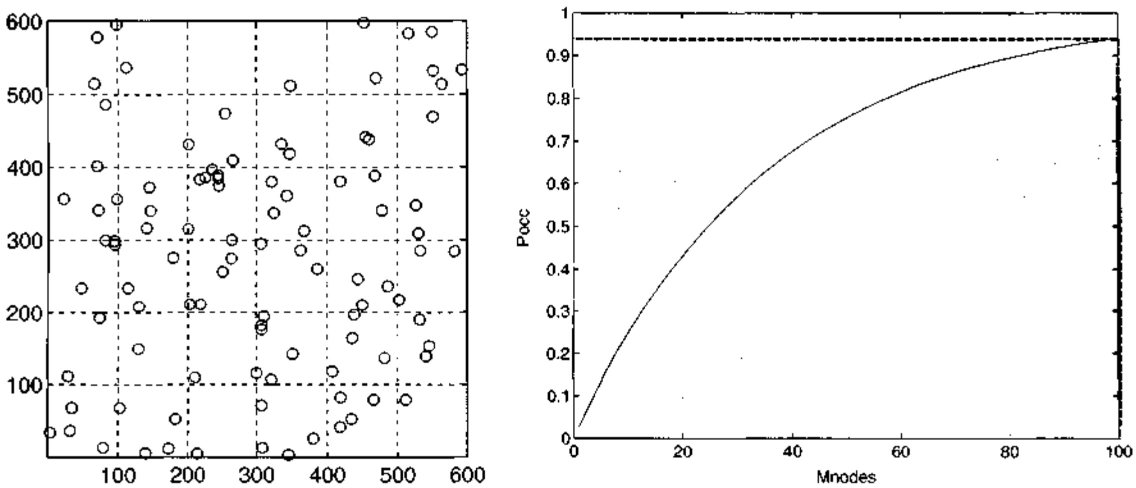


Figure 4.12: P_{occ} as function of node density in 600 x 600 m^2, 100 nodes random topology

What follows is an analysis of performance of the protocol in terms of energy efficiency, throughput and channel access delay. Where possible a comparison is made with the the performance of a pure contention medium access control protocol. Each of the characteristics is considered under the two regions of operation discussed in this section. For the purpose of the analysis an assumption is made that the protocol will be used by nodes with transmission range of 250 metres. Hence a single partition contains 36 cells each of size 100 by 100 metres. Under the assumed grid setup and uniform node distribution, the cross-over point between the two regions of operation is reached when the network size is 36 nodes. This according to Equation (4.7) corresponds to occupation probability of 0.64.

4.2.1 Energy Consumption Analysis

As mentioned earlier, under the contention region of operation, contention is likely to take place more often between nodes attempting to transmit during secondary slots. This includes nodes from all eight cells adjacent to the empty cell which provides the opportunity for a secondary slot transmission. As a result, the subarea over which contention is allowed in the grid is the sum of the areas of all adjacent cells plus the area of the empty cell. The expected number of contending nodes in that subarea can be approximated with the help of the node density of the network and the following equation:

$$N_{contend} = \beta\rho \times A \tag{4.8}$$

where ρ is the node density:

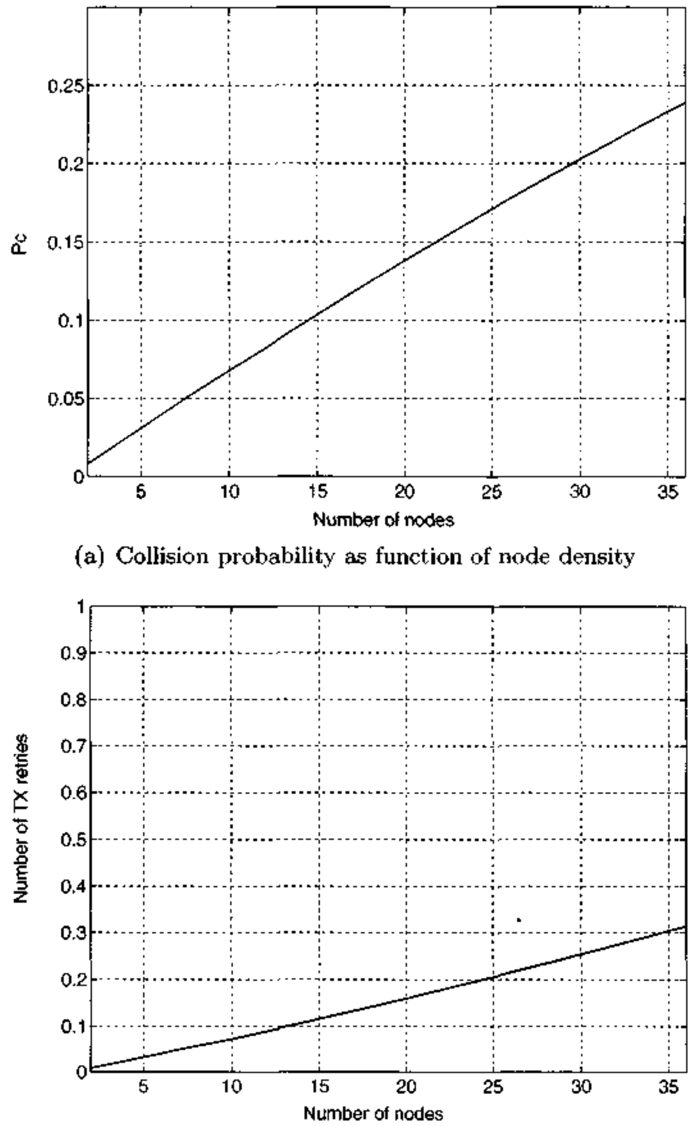
$$\rho = N/A \tag{4.9}$$

and β is the area reduction coefficient found as the ratio between the full grid area and the contention subarea A. Equation (4.8) could be rewritten as:

$$N_{contend} = \beta \times N \tag{4.10}$$

which contains only the area reduction coefficient and the total number of nodes N in the network.

In the case of the grid assumed for the purpose of the analysis, the contention area in the contention region of operation is nine cells which is 25 percent of the total area of the grid. Figure 4.13 shows the increase of collision probability and number of transmission attempts in the given region of contention as a function of node density. The relation between the network node density and P_c is obtained with use of Equation (3.4) and between node density and number of retries with use of Equation (3.10). The two figures



(a) Collision probability as function of node density

(b) Number of retries as function of node density

Figure 4.13: Contention region of operation with 600 x 600 m^2 grid

show that due to the low node density during operation in the contention region, the collision probability and therefore the number of retransmissions is kept to satisfyingly low levels. Figure-4.13(b) shows that retransmissions are highly unlikely so the per-packet energy consumption of the network can be estimated as:

$$E_{pkt} = E_{data} + E_{control} + E_{routing} \tag{4.11}$$

Equation (4.12) is a modified version of the extended per-packet energy model represented by Equation (3.2) from [62], discussed in Chapter 3 Section 3.5. It differs from the original by excluding the effect of packet retransmissions k as it is suggested by the analysis of this region of operation that such are unlikely.

The performance of the protocol under low node density conditions is expected to be very close to the performance of a pure contention protocol. This fact is based on the observation that the PA-STDMA MAC allows contention for secondary slots assigned to cells that are only within the transmissions range of the contending nodes. In other words, the nodes of the network would contend with the nodes within their reach just as they would do in the case of unconstrained contention where there is no space partitioning and slot assignments.

For the energy consumption analysis of the protocol in the TDMA operational region, the focus is set to a 300 by 300 metre subarea of the original partition as shown by Figure-4.14. The shown subarea and the nodes found therein could be a part of a very dense

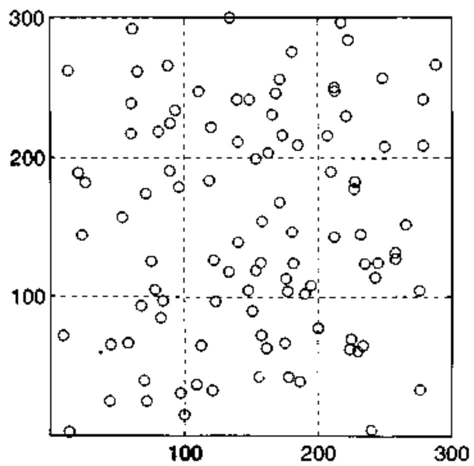
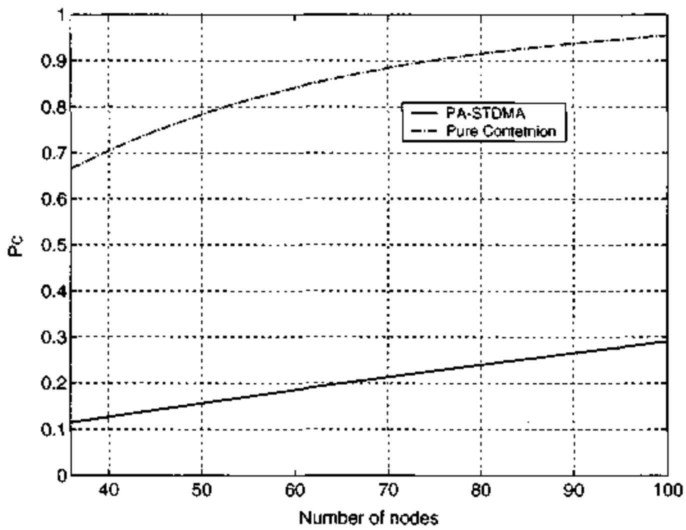
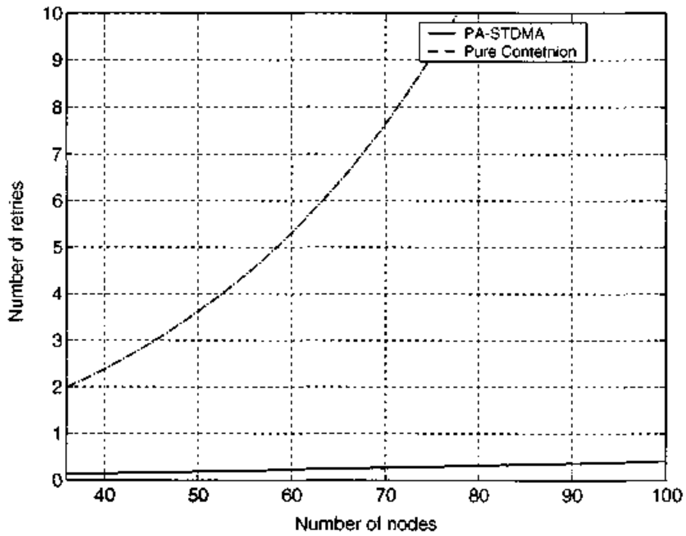


Figure 4.14: 300 x 300 m^2, 100 nodes subregion

network, a cluster of a network or it could be the entire network gathered around a hot spot. The region is chosen so that most of the nodes are within transmission range of each other and thus a direct comparison with pure contention could be made. All of the cells of the grid shown are occupied so transmission of data takes place only during primary slots. At any time, contention for the channel is allowed only within a single cell as opposed to pure contention under which most of the nodes in the whole 300 by 300 metres region would contend. This clearly leads to differences in the collision probability and the average number of data transmission retries associated with the two techniques. These two parameters as a function of number of nodes are illustrated in Figure-4.15. In



(a) Collision probability as function of node density



(b) Number of retries as function of node density

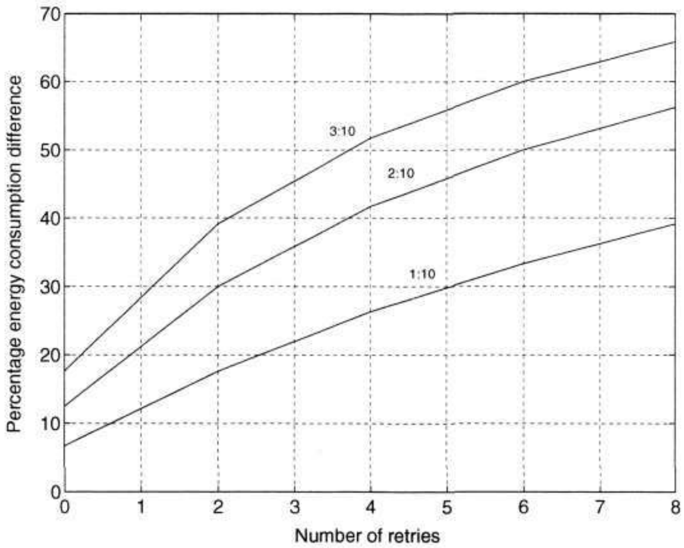
Figure 4.15: Contention region of operation with 300 x 300 m^2 sub-grid

Figure-4.15(a) it can be observed how the restriction on the number of contending nodes posed by the cell structure keeps the collision probability low. In contrast to that, under pure contention the collision probability is high and could only be affected by changes in the length of the contention window. The low collision probability under PA-STDMA MAC leads to operation with a low likelihood of data retransmission which is shown in Figure-4.15(b). As a result, the per-packet energy consumption of PA-STDMA MAC protocol in this region of operation could be estimated as:

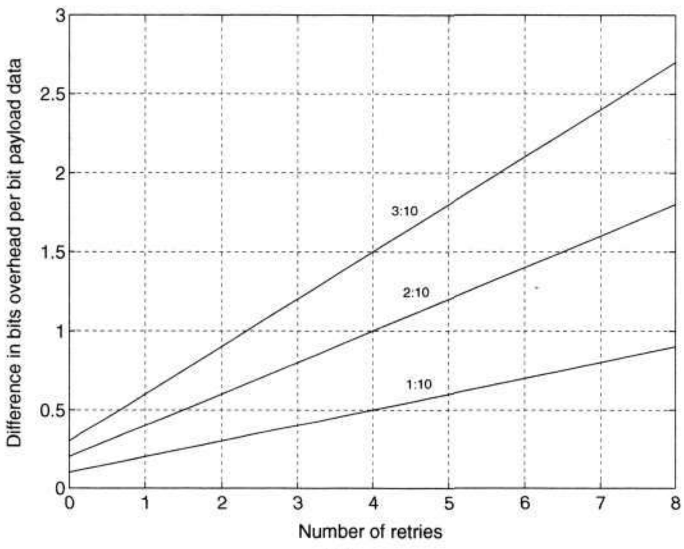
$$E_{pkt} = E_{data} + E_{routing} \tag{4.12}$$

Equation (4.12) is a reduced version of Equation (3.2) applicable to the TDMA region of operation. It does not include data retransmission since we have assumed this to be negligible. Transmission of control data is also excluded as primary slot transmissions are predominant in this region of operation.

The energy consumption analysis so far has shown that the proposed protocol is expected to operate with little or no data retransmissions. Based on this, the energy efficiency of PA-STDMA and a pure contention mechanism such as IEEE 802.11 can be compared on the basis of number of retries present under the operation of the former. The difference in energy consumption and overhead in terms of bits overhead per bit payload data between PA-STDMA and pure contention is shown by Figure-4.16. It illustrates the percentage increase of the two metrics in the case of pure contention with respect to the performance of the PA-TDMA protocol. Clearly pure contention becomes increasingly inefficient in comparison to PA-STDMA as the number of transmission retries increases. The inefficiency is based on the retransmissions of redundant data. The energy consumption difference and the incurred overhead is a function of the size of the control data in comparison to the payload data as well. When the ratio is 3:10 the per-packet energy consumption of pure contention is 60 percent higher than data of PA-STDMA MAC with an average difference of 3 bits more overhead for each transmitted bit of payload.



(a)



(b)

Figure 4.16: Per-packet energy consumption (a) and Overhead (b) performance difference between pure contention and PA-STDMA

4.2.2 Throughput Analysis

The maximum system throughput S is defined as the portion of a data frame which has been used for successful transmission of payload data. It is expressed as the following ratio:

$$S = \frac{T_{used}}{L} \tag{4.13}$$

The definition of S is based on a throughput analysis approach used in [64]. In the expression L is the time duration of the complete data frame. The numerator T_{used} is the overall time taken by successful payload data transmissions during a frame. It is equivalent to the portion of the frame which remains after subtracting the time taken by unused slots, slots involved in data collisions and contention periods from the total frame duration. The total time duration of unused slots in a frame is approximated as $L \times P_e$ where P_e is the probability that a slot is empty. The time spent in collisions is $L \times P_c$. The collision probability P_c is found with the help of Equation (3.4) form Chapter 3. The full expression for T_{used} is:

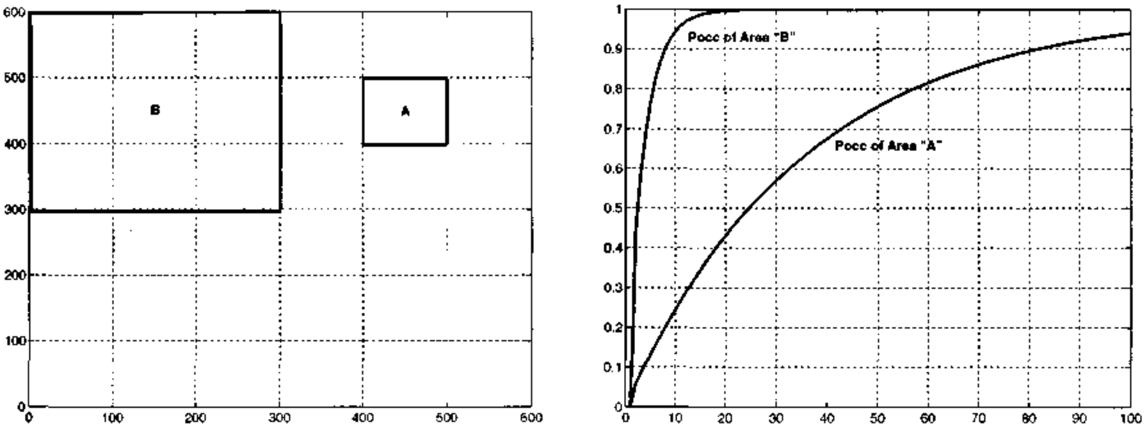
$$T_{used} = L - L \times P_e - L \times P_c - T_{cont} \tag{4.14}$$

where T_{cont} accounts for the time spent in contention and transmission of RTS-CTS control packets. Explanation of P_e follows.

A transmission slot could be empty either because of the absence of nodes in the vicinity of the corresponding cell or because of low data arrival rate. For the purpose of the maximum throughput analysis, an assumption is made that all network nodes have data for transmission at any time and hence P_e is found only as function of node density. A sufficient condition for a transmission slot to be assumed empty is that it has low associated probability of being used as a secondary transmission slot. Such low probability indicates low network density and infrequent primary slot transmissions. P_e is expressed as:

$$P_e = 1 - P_{sec} \tag{4.15}$$

The expression above states that if the transmission slot of a cell is not likely to be used as a secondary slot it is expected that the slot will be unused. The probability that a transmission slot is used as secondary P_{sec} is equal to the probability that there is a node in the region of interest as given by Equation (4.7) (ie P_{occ}). In the case of a secondary transmission, that region is a composite of all eight cells adjacent to the empty cell that provides the opportunity for a secondary transmission plus the area of the cell itself. It is designated with the letter B in Figure- 4.17. Area A is the area of a single cell.



(a) Regions of interest (b) P_{occ} for the regions of interest

Figure 4.17: Estimation of P_{occ}

The probabilities that regions A and B are occupied as a function of node density and calculated with use of Equation (4.7) are plotted in Figure- 4.17.

The throughput expression can be normalised to a frame length of one, so the maximum system throughput is expressed as:

$$S = T_{used} = 1 - P_e - P_c - T_{cont} \tag{4.16}$$

With the help of P_{sec}, P_e, P_c, T_{cont} and Equation (4.16), S can be estimated as a function of node density. Figure-4.18 shows the expected maximum throughput of the system as given by the analytical model and plotted as a solid line. It also shows throughput performance of the IEEE 802.11 MAC contention mechanism predicted by a similar analytical approach presented in [64]. It is plotted as a dashed line which represents the approximated behaviour of a hybrid system between the full RTS-CTS approach and the basic mechanism of IEEE 802.11 MAC. The numerical data from [64] corresponds to the same operation conditions as the ones assumed for the analysis in this section.

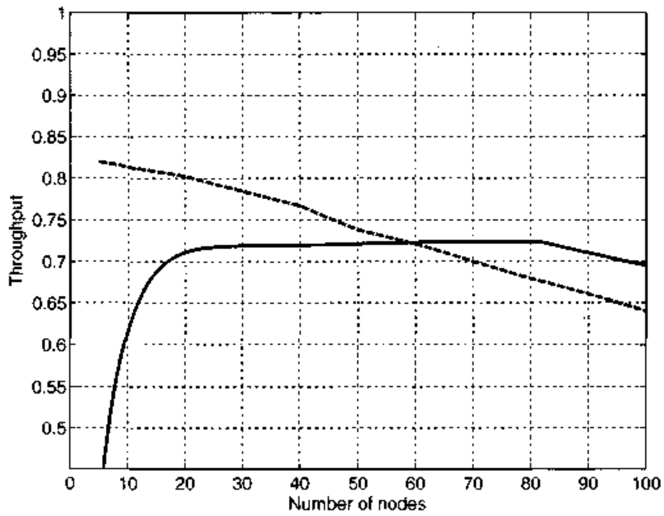


Figure 4.18: Approximated maximum throughput of PA-STDMA(solid) and pure contention(dashed) mechanisms

Numerical solutions of the throughput expression (Equation (4.16)) show suboptimal throughput performance in the contention operation region. This comes as a result of low node density and corresponding high probability of unused transmissions slots. The probability of a slot being unused decreases with increasing node density, hence an increase in throughput is observed. The analysis shows that the PA-STDMA MAC protocol should present maximum throughput performance at the beginning of the TDMA region of operation. At that point the channel is used optimally through primary slot transmissions. Thereafter, as the number of nodes increases so does the contention for a primary slot which in turn increases the collision probability and the throughput shows a decline. Comparison of the analytical results for the two medium access mechanisms suggests the existence of a cross over point at which the the PA-STDMA MAC should start outperforming the contention based approach as it handles congestion efficiently under increasing network load.

4.2.3 Channel access delay analysis

Channel access delay is defined as the average time interval between two successful channel acquisitions by a node in the network. For the purpose of the analysis it is estimated in the two operational regions of the protocol.

The best case channel access delay is expected in the contention region of operation whereby a node has a higher probability of transmitting more than once in the duration of a data frame. At the crossover point between the two regions of operation the delay is expected to assume a deterministic value corresponding to one transmission per frame. Thereafter, as the node density increases the nodes would have to contend for the use of a primary slot so it is expected that they would transmit once every few frames which would increase the channel access delay. Figure-4.19, in which the dark blocks designate a successful transmission slot, depicts these three cases.

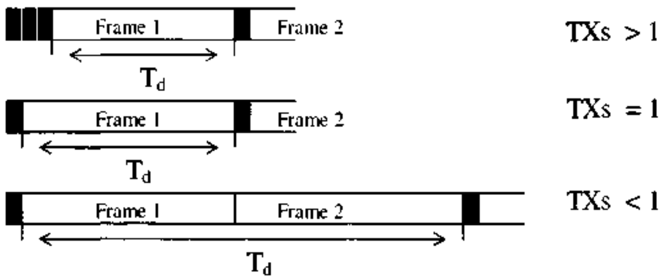


Figure 4.19: Formation of channel access delay

With the grid setup as established for the purpose of the analysis, a node can contend and transmit in at most nine transmission slots (eight adjacent slots plus its own). The likelihood of a node using nine transmission slots or less is dependant on node density. This likelihood can be expressed as the ratio between the probabilities of a secondary and primary slot transmission. The higher the latter the more likely it would be that a node transmits once or less during a frame (once every few frames). For the contention region the following applies:

$$T_d = M \times T_s - M_u \times T_s \tag{4.17}$$

where M is the number of slots in a data frame, M_u is the number of used transmission slots and T_s is the duration of a transmission slot. For the establish grid M_u is found as:

$$M_u = \left(9 - 9 \times \frac{P_{prim}}{P_{sec}}\right) \times P_t \tag{4.18}$$

The ratio $\frac{P_{prim}}{P_{sec}}$ determines how many of the nine optional slots a node would use under the given node density conditions. The estimated number of transmission slots is also dependent on the probability that there are no other transmissions on the channel or in other words that the transmission is successful. That probability is P_t and is given by Equation (3.5) from Section 3.5.

In the TDMA region of operation, as the node density increases the number of nodes found in a cell also rises. This together with the low probability of secondary transmissions leads to data transmission of once over a few data frames. As a result the delay rapidly increases with node density.

The expected channel access delay for a system that uses transmission slots of duration 0.006192 seconds (this corresponds to the duration used in the simulation of the protocol which is described in the following chapter) is depicted in Figure-4.20. As expected, the

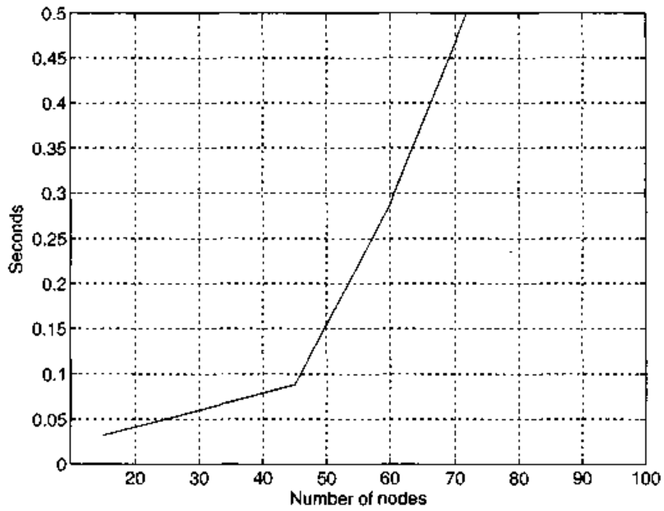


Figure 4.20: Expected channel access delay

lowest channel access delay is in the region of contention operation under low node density. The higher the node density, the smaller is the opportunity for a node to transmit more than once during a frame. This leads to a large channel access delay increase in the TDMA region of operation as the node density further increases.

4.3 Summary and conclusions

The analysis from Chapter 3 showed that excessive communication overhead, generated by inefficient medium access control, has an adverse effect on the performance of ad hoc networks. This has been the reason for the developed by the author PA-STDMA medium access control scheme. It was designed for improved energy consumption efficiency through optimised channel use in the form of reduced transmission of redundant data.

The proposed protocol is a hybrid between contention and channel allocation medium access mechanisms similar to the protocol proposed in [66]. Its main feature is the location based transmission scheduling achieved through partitioning the area of the network into cell grids. The expected reduction in communication overhead is based on its inherent low collision probability and the reduction of channel acquisition control overhead resulting from its cellular structure.

The chapter also provided analytical modelling which revealed some of the important performance properties of the proposed protocol. The first property to be investigated was network throughput. The reason for this analysis being that if the protocol performs poorly in terms of achieved throughput, the achieved energy reduction would be irrelevant. The throughput analysis, however, provided evidence that the proposed protocol should have satisfactory throughput characteristics. Analytical comparison with the contention based IEEE 802.11 MAC showed that the efficient channel use presented by PA-STDMA should lead to improved throughput under conditions of high network load.

The analytical model of the proposed protocol suggested that the occurrence of transmission retries is highly unlikely under its operation. Hence the energy savings resulting from reduced transmission of redundant data could be as high as 60%.

The channel access delay analysis of the PA-STDMA MAC suggests that the access delay will increase as a function of node density. This is as a result of the data transmission restrictions posed by the TDMA structure.

The results of the analytical model are verified by simulation of the PA-STDMA MAC protocol in Chapter 5.

Chapter 5

Protocol Simulation

5.1 Introduction

The performance of the proposed PA-STDMA MAC protocol was further studied and verified against the results of the analytical model of Chapter 4 with the help of simulations. Performance characteristics such as *power consumption efficiency*, *throughput* and *channel access delay* were monitored under varied network load and a range of node densities. The simulations of the PA-STDMA MAC were carried out concurrently with simulations of the IEEE 802.11 MAC[1]. Subjecting these two protocols to identical simulation conditions allowed for comparison between pure contention medium access control (IEEE 802.11 MAC) and a combined contention-TDMA approach in the case of PA-STDMA MAC.

The chosen simulation environment was NS2 [8]. It is a freely distributed, open-source simulator developed by Ernest Orlando Lawrence Berkeley National Laboratory for modelling and simulation of network protocols. Initially NS2 was focused at wired and satellite networking. However, after the *wireless node* extension contributed by the CMU Monarch project in 1999 it has become a standard framework for testing protocols developed for ad hoc networks. Its shared, community-based nature and the fact that it provides a wide range of existing network protocols (including IEEE 802.11) make it the simulator of choice in this research.

[1]description of the channel contention process used by IEEE 802.11 is provided in Chapter 2 Section 2.3.1.6. Full description is provided by [2]

This chapter is organised as follows: Section 5.2 provides a brief overview of the NS2 environment; Section 5.3 describes the choice of simulation parameters; Section 5.4 presents simulation results and interpretation.

5.2 Simulation Environment

NS2 is an object-oriented, discrete event driven simulator. It makes use of two different programming languages: object oriented *Tcl* (OTcl) as a frontend scripting tool and *C++* for pre-compiled objects. To that end, C++ is used for its code compilation efficiency while OTcl provides easy object initialisation and manipulation.

The first part of the simulator is a Tcl script interpreter. It allows the use of Tcl scripts to define simulation parameters prior to simulation. The parameters in the Tcl scripts are used to initialise the C++ implemented event scheduler and objects from the component library which form the second part of the simulator. Additional parameters in the Tcl script provide definitions for: protocols, network topology and mobility that are to be used in the simulation. The interfacing between the parameters and definitions from the Tcl script and the C++ objects is done through a process referred to as *shadowing*. It is the mapping of the interpreted Otcl objects to the corresponding C++ compiled objects.

In NS2 there are two C++ objects regarded as the main building block of the simulator. These are the *classifier* and the *connector* objects. Multiple instances of the classifier are used to associate different objects while the connector manipulates packet events. The integration of the two, together with the objects that they combine, form different network functional units such as the wireless node shown in Figure-5.1.

In Figure-5.1, below the data generation agents multiplexed by the classifier, is the *routing agent*(A) which represents the routing protocol used during a simulation. NS2 offers a variety of routing protocols both on-demand and table driven. Moving down through the node structure of Figure-5.1, the routing agent is followed by the *link layer*(B) which is based on the IEEE 802.2 link layer. Its role is to translate logical addresses to physical MAC addresses. This is done with the help of BSD and ARP (C) protocols. The next

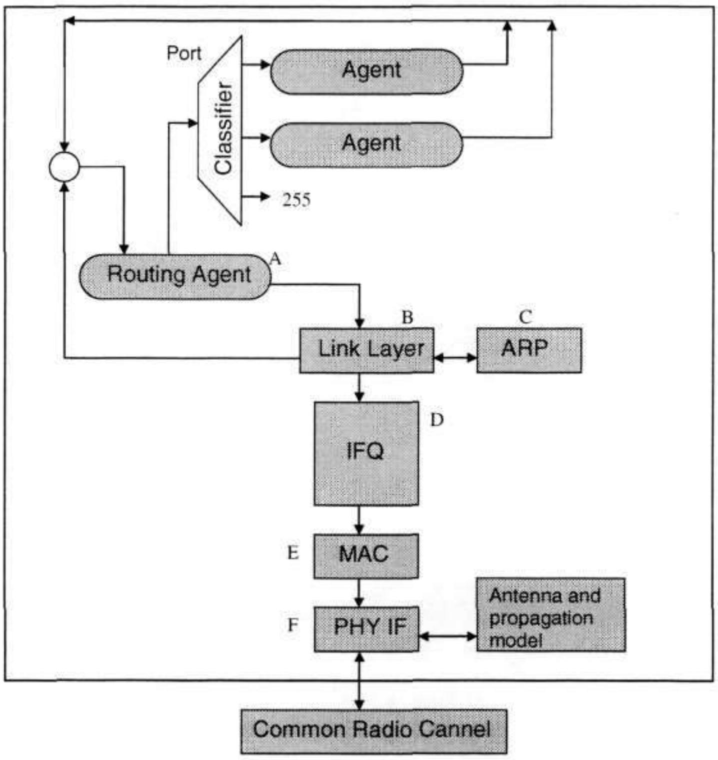


Figure 5.1: NS2 structure of a mobile node [8]

component from the structure is the *interface queue* (D). In general NS2 provides a variety of queuing disciplines. However, with the current implementation of the mobile node only Priority-based droptail buffering is fully supported.

The structure of NS2 and its components provides a way for simple substitution of network protocols and data generation agents. The development of custom protocols is done on the basis of the existing structures and objects. For example, PA-STDMA MAC stems from the MAC class defined in the simulator environment. As a subclass PA-STDMA MAC adheres to predefined structures and requirements of the simulator. In addition to that it defines and uses substructures relevant only to its operation. When selected in the initialising Tcl script, instances of PA-STDMA MAC are generated for each mobile node and placed at the position of the MAC box (E) in the node hierarchy shown in Figure-5.1. In addition to that, the standard NS2 package offers two other medium access control protocols: IEEE 802.11 MAC and a preamble based TDMA MAC protocol. The IEEE 802.11 MAC protocol uses the full RTS/CTS/DATA/ACK data transmission as defined by the standard. A basic channel access (DATA/ACK) is also possible through adjustment of the *RTSThreshold* parameter.

The last block of the structure of the wireless node is the *physical layer* (F) which represents the hardware interface used for channel access. The available model is based on the Lucent WaveLan DSSS radio interface. It takes into consideration packet collisions and radio propagation constraints. The available radio propagation models are *FreeSpace*, *TwoRayGround* and *Shadowing*.

The common radio *channel* is a class that simulates the actual transmission of the packet at the physical layer. It supports events such as carrier sensing, contention and collision detection and has a propagation delay property.

5.3 Simulation Setup

The simulation of the protocol serves two purposes: verification of the performance of PA-STDMA MAC protocol against the predictions of the analytical model and comparison with the performance of a contention based protocol such as IEEE 802.11 MAC. For comparison purposes the two protocols were subjected to the same network conditions over different simulation scenarios.

The radio model used was the default IEEE 802.11 DSSS Lucent WaveLan offered by the simulator. It was set for a transmission range of 250m, carrier sense range of 260m, channel capacity of 2 Mbps and initial battery energy of 500 Joules. The energy dissipated during *transmit*, *receive*, *idle* and *sleep* had the following ratios 1:0.75:0.65:0.0 respectively[2]. The propagation model used was the *TwoRayGround* model. The routing protocol used throughout the simulations was the Ad Hoc On-demand Distance Vector (AODV) which is provided by the simulator.

The chosen mobility model was the *random waypoint* model. It is frequently used in literature for the purpose of ad hoc network simulations. The mobility patterns were

[2] the value of 1 indicates the highest power level in the case of the transmit state, the rest of the levels are indicated as fractions of that value

generated by *mobgen-ss* which has been shown in [75, 76] to address and rectify some of the flaws of the original mobility generator *setdest* provided by the NS2 package.

The simulated data traffic was of a Constant Bit Rate (CBR) type with a payload data packet size of 1 kByte. The simulated network was confined within an area of 600 by 600 metres.

The performance of the two protocols was compared under four different scenarios:

Scenario 1: Static ad hoc network with varying node density

The throughput analytical model from Section 4.2.2 is limited to the performance of the MAC layer. Similarly the analytical results from [64], representing the throughput performance of the contention based protocol, also do not account for the presence of other network layers. To that end, for verification of the analytical results, a static ad hoc network was simulated as it presents conditions closest to the ones assumed by the analytical models. In a static ad hoc network the routing is minimal and takes place only at the start of the network operation. Therefore the performance of the MAC layer is lightly affected by routing activities and is shaped only by the transmission of payload data. For this particular scenario the offered load per node was fixed at a high rate of 15 packets per second[3]. It was chosen as such to approximate the condition of constantly available data for transmission, assumed in the throughput analysis of Section 4.2.2. The simulated number of nodes was varied between 15 and 90. The carrier sense threshold was set to 500m to prevent spatial reuse of bandwidth as it is also not considered in the analysis of both PA-STDMA MAC and IEEE 802.11 MAC.

Scenario 2: Mobile ad hoc network with varying node density

This simulation scenario helped to investigate and compare the performances of the PA-STDMA MAC and IEEE 802.11 MAC in the presence of node mobility. The nodes of the network were set to move with a moderate speed (resembling human walking/running) randomly chosen from the range of 3 m/s to 7 m/s. Once the nodes reach their destination

[3] the data transmission rate measured in bits/sec can be found as: (number of packets per second per node) x (packet size in bits) x (number of nodes)

they stop for 5 sec and then start moving toward a new randomly generated destination point. The number of nodes in the network was varied between 15 and 90. The offered load per node was normalised to the channel capacity[4], the worst case being at 90 nodes and 3 packets per second resulting in a requirement of 2.16 Mbps bandwidth. Offered network load at rates exceeding the available channel bandwidth were found to be responsible for poor packet delivery ratios(PDR)[5] (less than 10%) and therefore ware not considered.

Scenario 3: Mobile ad hoc network with fixed number of nodes

This simulation scenario serves to investigate the performance of the two protocols as a function of varying offered load while the number of nodes in the network is fixed. The number of simulated nodes was 50. The offered load was varied between 1 and 6 packets per second per node (resulting in a bandwidth requirement of 2.4 Mbps in the worst case of 50 nodes at 6 packets per second). As in the case of scenario 2, offered network load at rates exceeding the channel capacity were not considered as they were found to result in poor PDR. The speed of the nodes was randomly selected from a range between 3 m/s and 7 m/s.

Scenario 4: Network lifetime

As a result of energy conservation, the lifetime of the network operated by PA-STDMA MAC is expected to be extended. In order to quantify the achieved extension the two protocols, IEEE 802.11 MAC and PA-STDMA MAC, were simulated for a duration of 2000 seconds while the residual network energy (per node) was monitored. The network consisted of 50 nodes in an area of 600 by 600 metres. The offered traffic load was set to 5 data packets per second per node resulting in a bandwidth requirement equal to the available 2 Mbps. The speed of the nodes was randomly selected from the range 3 to 7 metres per second. The conditions under this scenario were chosen as such to resemble a realistic ad hoc network application.

During simulations run under scenarios 1,2 and 3 the PA-STDMA MAC protocol was

[4]the generated number of packets per second per node was inversely proportional to the number of nodes in the network, always resulting in a bandwidth requirement close to the available "raw" channel capacity

[5]PDRs are measured at the CBR application layer which serves as both data source and data sink

set for improved throughput performance. To that end, the nodes did not automatically enter low power mode during idle slots to ensure availability at all times. However, low power mode was available for use after the contention periods under the condition that no communication has been initiated. During simulations run under scenario 4, all possible combinations of power saving options were simulated. These included, use of idle slots, use of low power mode after contention periods, use of both and use of neither.

During all simulations the PA-STDMA MAC protocol was initialised to match the setup assumed for the analysis from Chapter 4. The geographic partitions were set to be 6 by 6 grids with square cells of 10000 m^2. The time duration of the transmission slots was set to 6.192 ms. The contention windows at the beginning of each transmission slot (see Figure-4.1, Section 4.1.2) were set to take a fifth of the slot duration.

All relevant simulation parameters are grouped in Table-5.1 according to the simulation scenarios.

Table 5.1: Simulation parameters

	Scenario 1	**Scenario 2**	**Scenario 3**	**Scenario 4**
PHY model	DSSS IEEE 802.11	DSSS IEEE 802.11	DSSS IEEE 802.11	DSSS IEEE 802.11
Channel bandwidth	2 Mb/s	2 Mb/s	2 Mb/s	2 Mb/s
Routing protocol	AODV	AODV	AODV	AODV
Simulation time	400 s	400 s	400 s	2000 s
Node number	15,30,45,60,75,90	15,30,45,60,75,90	50	50
Number of data connections	15,30,45,60,75,90	15,30,45,60,75,90	50	50
Mobility model	Random waypoint	Random waypoint	Random waypoint	Random waypoint
Node speed	0 m/s	5 m/s	5 m/s	5 m/s
Speed variance	0 m/s	2 m/s	2 m/s	2 m/s
Pause time	∞	5 sec	5 sec	5 sec
CBR data traffic rate	15 pkt/s	Normalised to channel bandwidth	1 to 6 pkt/s	5 pkt/s
Packet size	1 kByte	1 kByte	1 kByte	1 kByte

5.4 Simulation Results

The following section presents simulation results relating to the performance of the proposed PA-STDMA MAC protocol. In separate discussions, important performance metrics are compared to the findings of the analytical model from Chapter 4 and the corresponding performance of IEEE 802.11 contention based MAC protocol. The investigated metrics and the order of their discussion are as follows: *throughput*, *energy efficiency* and *channel access delay*

5.4.1 Throughput

The analysis of the proposed protocol suggested the existence of two regions of operation with different performance characteristics. The regions are namely the *contention* and the *TDMA* regions. According to the discussion of Section 4.2.2, at the beginning of the first region, the throughput performance is expected to be suboptimal as a function of low node density and high probability of unused transmission slots. With an increase in node density, the maximum throughput performance is expected to improve as a result of better utilisation of the transmission slots. The improvement is expected to be moderate in the late contention region (at the equivalent of 30 to 40 nodes in a 600 by 600 metres network area), as a result of elevated collision probability, and pronounced after the start of the TDMA region. It expected to peak in the early TDMA region where the channel is used through efficient (low communication overhead and low collisions probability) primary slot transmissions. At high node density, the throughput performance under the TDMA operation region is expected to reduce gradually as contention makes way for primary slot transmissions. The model also suggested a different cross-over point at which PA-STDMA MAC is expected to start performing better than the pure contention based access control (refer to Figure-4.18 from Section 4.2.2).

Figure-5.2 shows the throughput performance of both, PA-STDMA MAC and IEEE 802.11 MAC, subjected to identical conditions of increasing node density and high offered load at rate of 15 packets per second as defined under scenario 1. The throughput graphs are a measure of the average received payload data bits per second at the MAC layer.

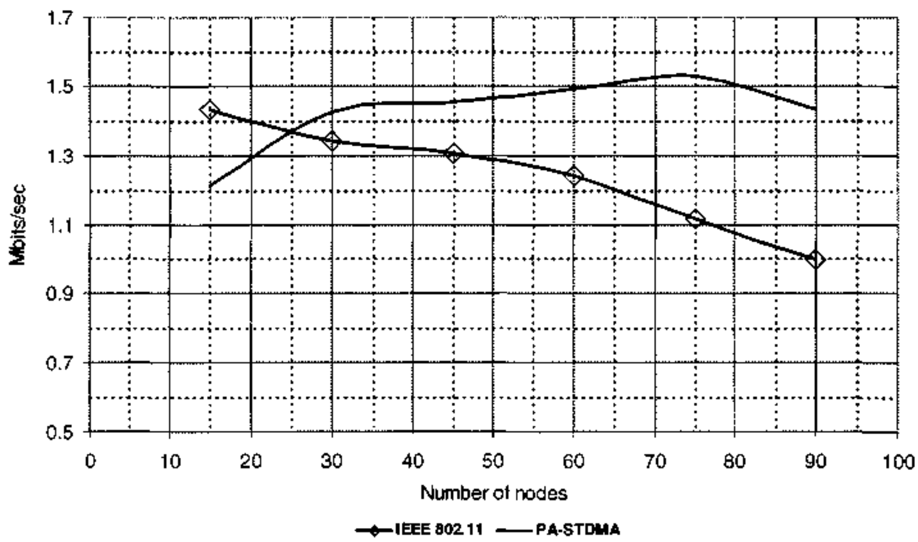


Figure 5.2: Maximum throughput at varied node density - scenario 1

The expected poor performance of PA-STDMA MAC in the low node density region and the performance cross-over point of the two protocols are confirmed by the simulation results. The shape of the throughput graph of PA-STDMA MAC protocol distinctly shows the two different regions of operation and their corresponding performance matching the predictions of the analytical model (the corresponding maximum throughput performance graph form the analysis is in Section 4.2.2, Figure-4.18). The performance of IEEE 802.11 MAC is worse than expected leading to a crossover point at node density $\rho = 6.94e^{-5}$ (25 nodes) rather than the anticipated by the analysis $\rho = 1.67e^{-4}$ (60 modes). The obtained simulation results presented in this section, regarding the maximum throughput performances of PA-STDMA MAC and IEEE 802.11, are consistent with results of similar independent investigations in [66] and [77][6].

Figure-5.3 shows the throughput of the two protocols under network conditions defined in scenario 2. The performance of the PA-STDMA MAC protocol shows a decrease of 15% which is attributed to the increasing (with node density) routing communication overhead in the presence of mobility. However, the hybrid protocol maintains a better throughput performance when compared to the contention scheme under high node density conditions. It presents as much as 10% higher throughput and 7% higher packet delivery

[6]these investigations are focused at performance comparison between hybrid contention-TDMA protocols and pure contention protocols

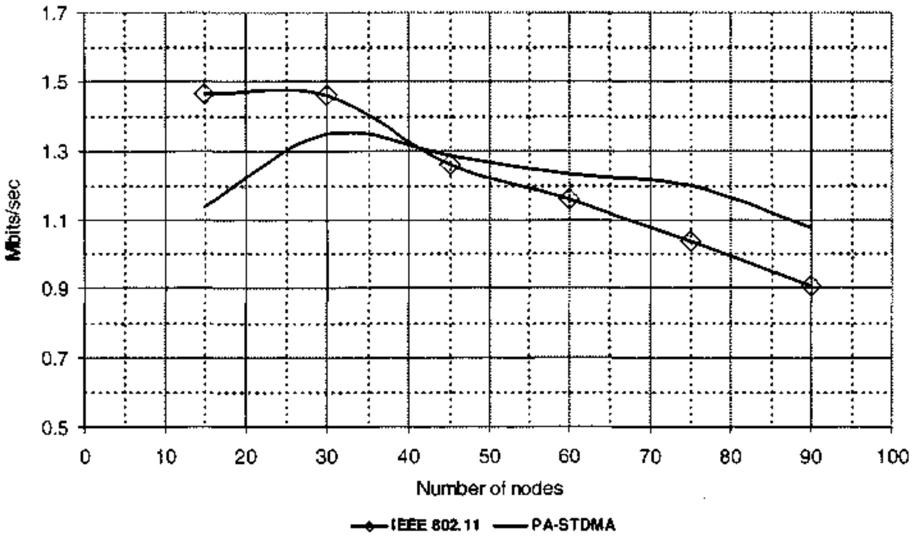


Figure 5.3: Maximum throughput at varied node density - scenario 2

ratio (measured at the application layer) in a mobile network of 90 nodes.

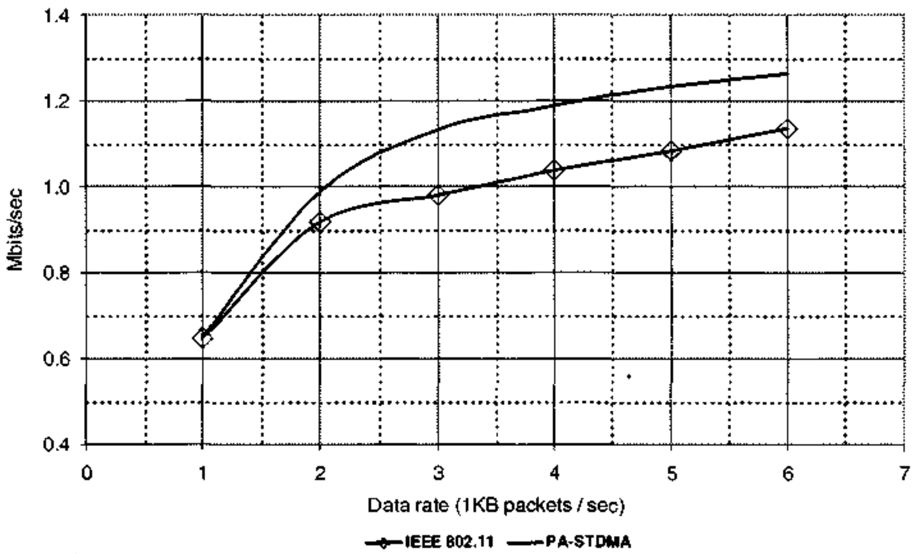


Figure 5.4: Maximum throughput at varied offered load - scenario 3

Figure-5.4 shows the throughput of the two protocols as a function of varying offered load with a fixed number of nodes as defined in scenario 3. Under the given conditions the achieved throughput of PA-STDMA MAC is up to 13% better than that of IEEE 802.11 MAC at most of the data rates. This is a clear manifestation of the improved throughput efficiency at higher node densities resulting from the structured channel access.

5.4.2 Energy efficiency

The energy consumption analysis from Chapter 3 suggested that in the case of the contention approach, network congestion would lead to increased communication overhead in terms of routing, RTS/CTS and retransmitted payload data packets. This in turn would lead to energy waste through transmission/reception of redundant data. The analysis from Sections 4.2.1 and 4.2.2 suggested that due to its TDMA nature and operation based on cell structure, PA-STDMA MAC should sustain throughput and energy efficient operation in the face of rising node density and offered network load. Verification of the throughput efficiency of the proposed protocol was discussed in the previous section. For verification of the energy efficiency claim, information was sought from the simulation results regarding *overhead*, *energy consumption rates* and *energy efficiency* in terms of payload data transmitted per Joule of energy.

The energy efficiency of the PA-STDMA MAC protocol was observed and compared with that of the contention based scheme in the context of scenario 2. That particular scenario was chosen as it represents a realistic mobile ad hoc network environment that includes routing.

Figure-5.5 investigates the congestion conditions of the simulated networks in terms of packet collisions, as a function of increasing node density. It clearly shows the increasing congestion under the operation of IEEE 802.11 MAC with a high number of collisions at high node densities. In comparison to that the number of collisions under the operation of PA-STDMA MAC is insignificant (99.8% less) and only slightly affected by changing the network node density. The graph in Figure-5.6 verifies the performance behaviour analysis related to the two regions of operation. With increasing node density the number of collisions increase as a result of higher number of nodes within the secondary slot contention area. This increase is observed until the operation region cross-over point is reached. Thereafter, a gradual decrease in collisions due to the predominant primary slot transmissions of the TDMA region of operation is witnessed. Although the existing collisions are responsible for data packet retransmissions, they do not appear to have significant effect on the performance of PA-STDMA MAC. This is confirmed by the simulation results

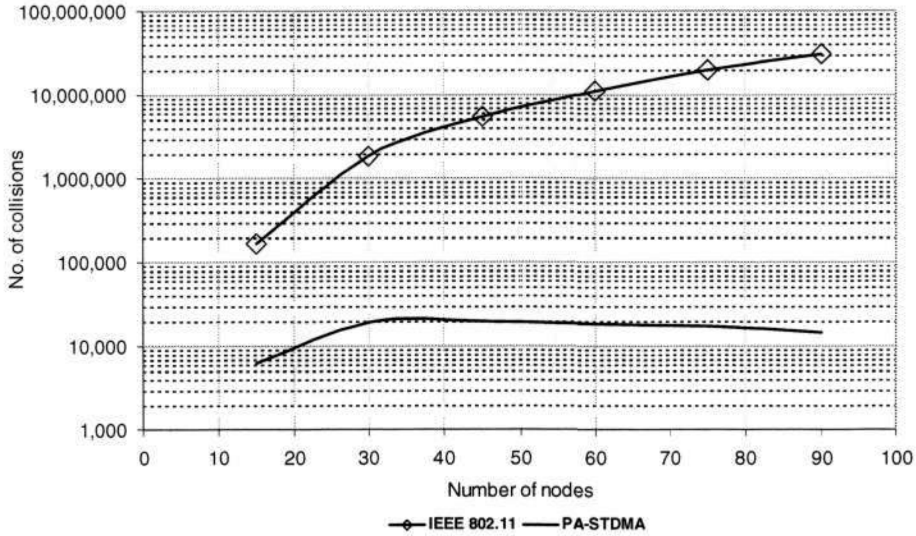


Figure 5.5: Collisions comparison

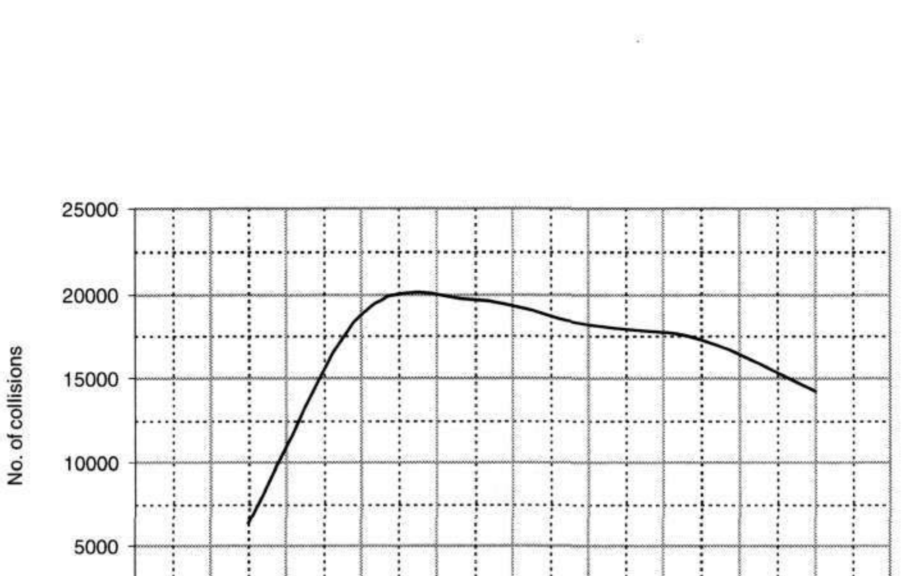


Figure 5.6: Collisions under PA-STDMA MAC

showing the generated communication overhead by the protocol, presented in Figure-5.7.

Analysis from both Section 3.6 and Section 4.2.1 have shown that efficient channel use, through combination of contention and TDMA, could achieve up to 60% active[7] energy consumption reduction as a result of the reduced communication overhead. Figure-5.7 shows the overhead incurred by the two protocols as node density increases. It includes routing, control and payload data retries.

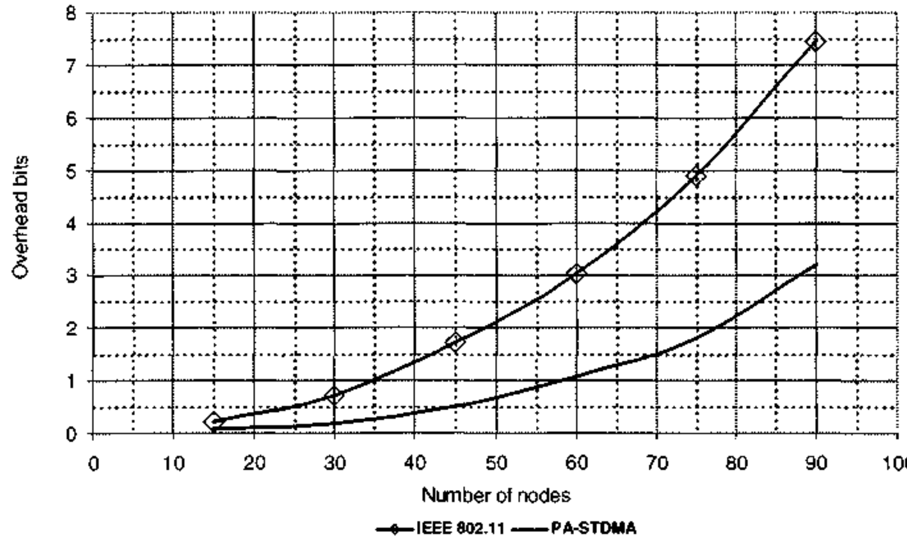


Figure 5.7: Number of bits of overhead for every bit of payload data

The simulation results show that PA-STDMA MAC achieves close to 60% reduction of communication overhead which is consistent with the analytical results from Section 4.2.1 (see Figure-4.16(b) in that section).

The communication overhead increase in the case of PA-STDMA MAC is that of necessary routing overhead which is clearly a function of network size. Investigation of the packet transmission logs shows that a network of 90 nodes generates three times more routing packets than that of a network with 45 nodes and less.

The energy consumption reduction produced by the optimised use of communication overhead is illustrated by Figure-5.8. It shows the active state energy consumption rates of the simulated protocols. The simulation results confirm those from the analysis[8] by show-

[7] active energy is the energy consumed for transmission and reception

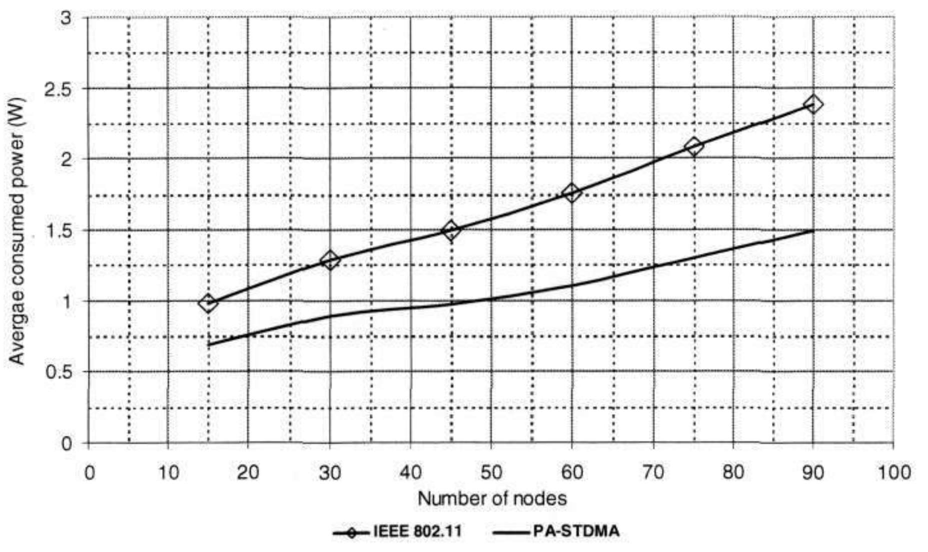


Figure 5.8: Average power consumption

ing that a PA-STDMA MAC operated network consumes up to 40% less energy for data transmission and reception than in the case of a pure contention based IEEE 802.11 MAC. The largest difference in energy consumption is under high node density where the effect of congestion variation between the protocols is most pronounced.

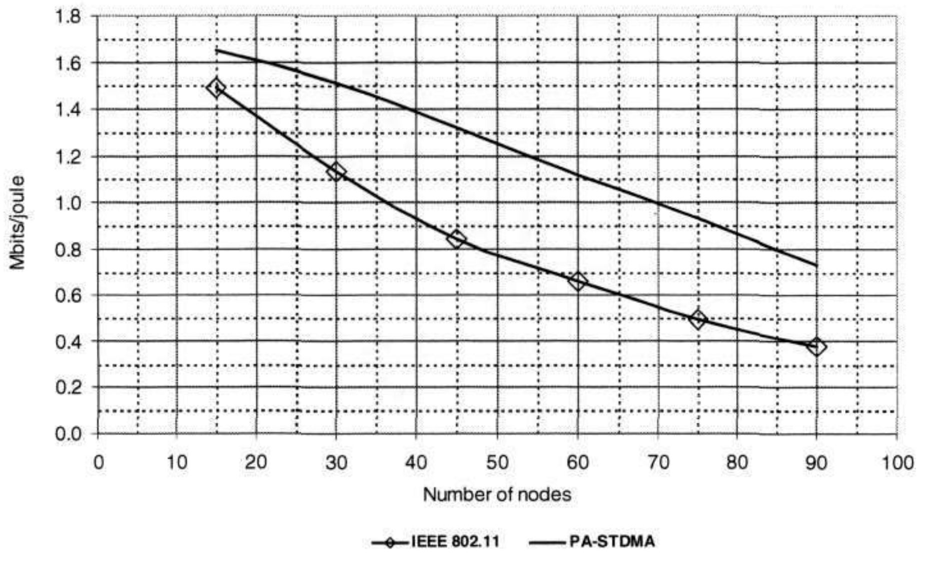


Figure 5.9: Energy use efficiency

Figure-5.9 provides further confirmation of the superior energy consumption efficiency of

[8]looking at the energy consumption reduction analytical results in the case of 1:10 ratio between control and payload data packet size which is shown in Figure-4.16(a). The 1:10 ratio is the closest to the simulated packet parameters

the proposed protocol. It shows the average amount of payload data transmitted by the two protocols per consumed Joule of energy. According to the simulation results PA-STDMA MAC transmits (on average) 380 Kb more data for every consumed Joule of energy in comparison to IEEE 802.11 MAC.

It was concluded in Chapter 3 that the routing activity and its broadcast nature has an adverse effect on network stability. It leads to an avalanche effect whereby unsuccessful transmissions are often falsely attributed to missing next hop nodes and hence initiate route maintenance procedures that contribute to network congestion. This in turn leads to further unnecessary routing activity. Ultimately this activity results in unstable network operation and inefficient energy consumption caused by the generated communication overhead. These analytical conclusions were confirmed via simulation. The two medium access control protocols were simulated with a fixed number of nodes and increasing offered traffic load (simulation scenario 3). Number of collisions recorded during simulation of the MAC protocols is presented in Figures-5.10 and 5.11.

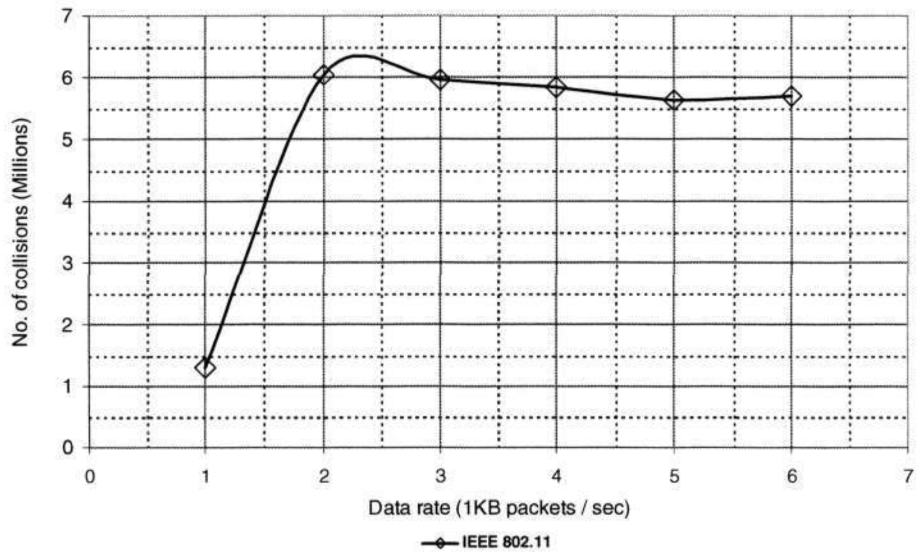


Figure 5.10: Congestion in terms of collisions under changing offered network load with IEEE 802.11 MAC

Investigation of the simulation output files confirmed a significant increase in the number of routing packets generated in the case of IEEE 802.11 MAC between the operation points of 1 and 2 pkt/sec/node network load. The packet collision logs indicated that 95% of

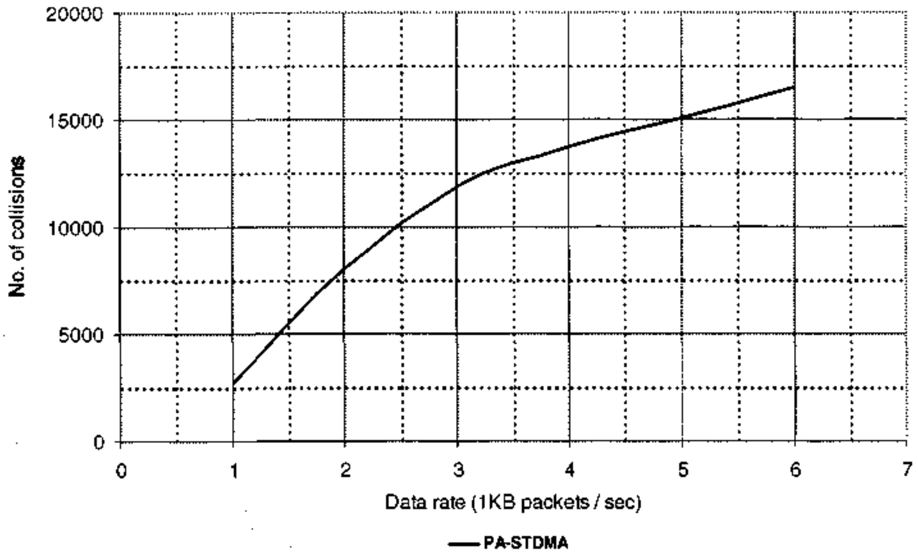


Figure 5.11: Collisions under PA-STDMA MAC

the packets involved in collisions are routing packets. Hence it can be concluded that the steep increase in the number of collisions shown in Figure-5.10 is a result of increased number of route requests. These simulation results confirm the negative effect that routing initiated broadcast storms have on congestion. As expected, under PA-STDMA MAC the broadcast storm effect is avoided and the network congestion is kept low.

The communication overhead achieved by the protocols is shown in Figure-5.12. Comparing

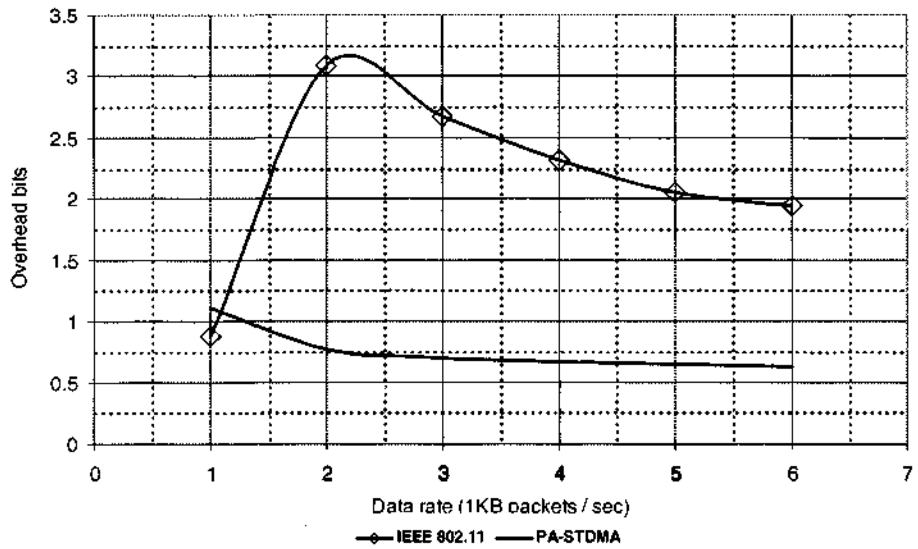


Figure 5.12: Number of bits of overhead for every bit of payload data

the collision graph of Figure-5.10 and the overhead graph, the effect that congestion has on overhead can be clearly observed. The same steep increase between 1 and 2 pkts/sec is observed in the energy consumption rate of the contention based protocol in Figure-5.13. The graphs of IEEE 802.11 from Figures-5.10, 5.12, 5.13 clearly demonstrate the congestion-overhead relation and the effect it has on energy consumption.

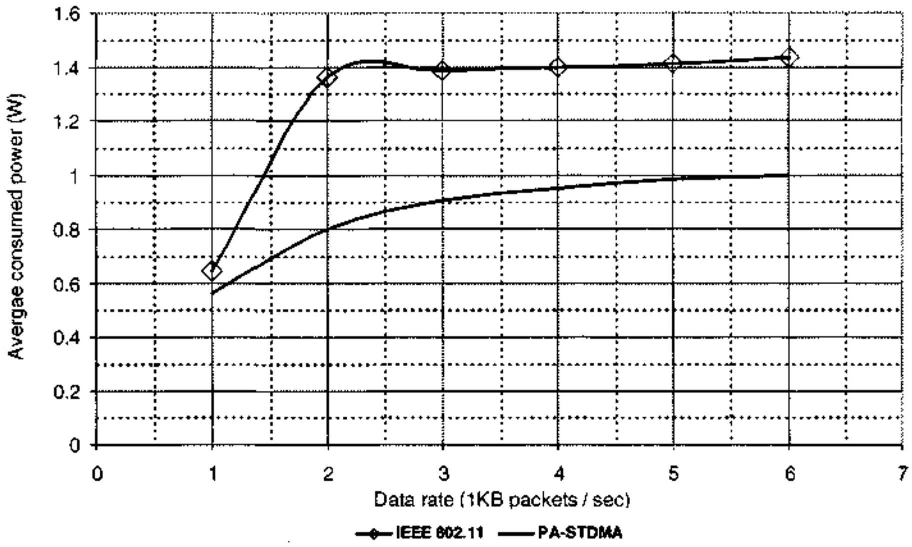


Figure 5.13: Average power consumption

Finally, Figure-5.14 shows the energy efficiency of the investigated protocols. As expected,

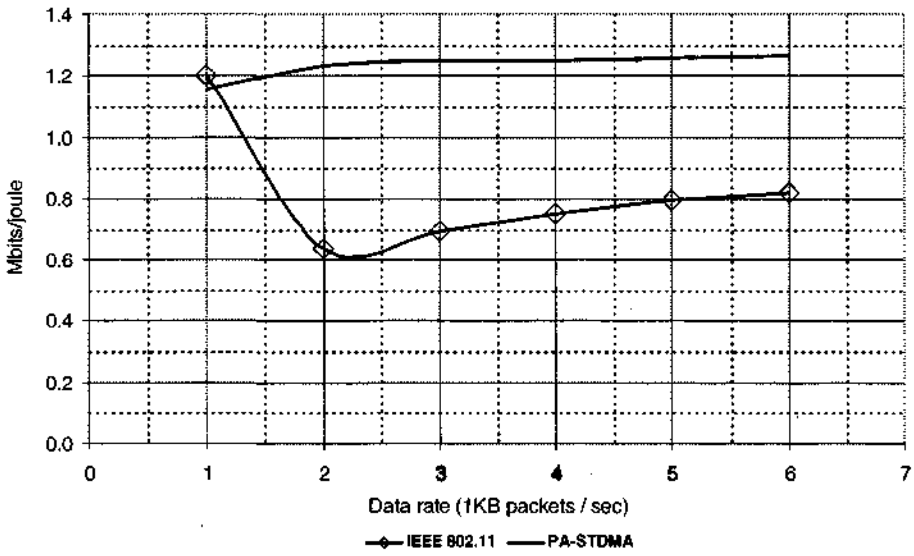


Figure 5.14: Energy efficiency

the graph shows that the performance of IEEE 802.11 MAC decays sharply suggesting smaller amounts of payload data successfully transmitted per joule of energy. This can be explained by the fact that energy is expended for transmission of information other than payload data. Therefore it can be concluded that increase in communication overhead rather than increase in offered payload data is the reason for the increased energy consumption in Figure-5.13.

5.4.3 Channel Access Delay

As in the case of the network throughput metric, it is expected that the channel access delay will be shaped by the operating regions of the PA-STDMA MAC protocol. The delay is expected to rise as the number of nodes contending for the channel access increases. With higher node density the delay is expected to increase sharply indicating that nodes do not have the chance to transmit during every data frame. This effect is expected to become prominent when most of the cells are occupied by more than one node.

Figure-5.15 shows the average channel access delay for the two protocols operating under network conditions defined in scenario 2. As expected the contention based protocol

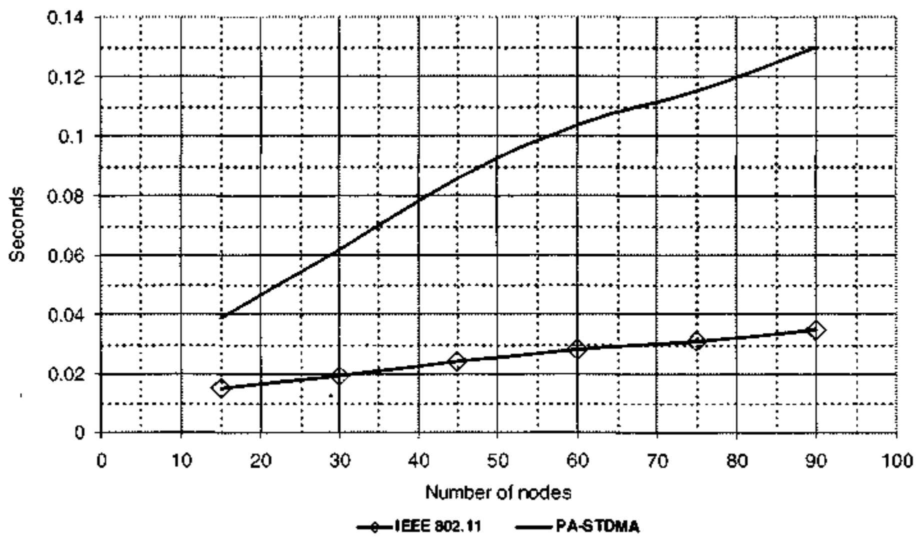


Figure 5.15: Channel access delay in scenario 2

achieves better access delay. It is as much as 70% lower than that achieved by PA-STDMA MAC in a network of 90 nodes. The high channel access delay in the case of the proposed protocol is attributed to the restrictions posed by the TDMA structure on data transmission. In a pure contention based channel access, the nodes have the freedom to attempt transmission as soon as data is available. In the case of PA-STDMA MAC, the nodes have to obey the transmission schedules which often prevents them from immediate transmission.

An observation is made that at the node densities corresponding to TDMA operation (over and above 40 nodes), the delay is lower than that predicted by the analysis (see Figure-4.20 in Section 4.2.3). The reason being that in addition to their primary slot transmission, the nodes manage to make use of one or more secondary slot transmissions hence reducing the time between consecutive transmissions. This is possible since the distribution of the simulated nodes was not strictly uniform (in contrast to the assumption of uniform node distribution made for the analysis). Such additional transmission opportunities were not accounted for by the channel access delay model from Section 4.2.3 resulting in higher channel access delay estimations.

5.4.4 Network lifetime

According to the analysis in Section 4.2.1 and the simulation results presented in this section, the use of PA-STDMA MAC leads to a 40% reduction in active energy consumption. To investigate how this energy reduction together with use of low power (sleep) mode contribute towards network lifetime extension, the two protocol (PA-STDMA MAC and IEEE 802.11 MAC) were simulated under conditions defined in scenario 4. The PA-STDMA MAC was simulated in three different modes of operation: (a) *power-unaware mode* i.e. low power mode is not used at all; (b) *optimal throughput performance mode* i.e. low power mode is not directly used during idle slots (it could, however, be entered at the end of the contention period under the condition that data transsmission/reception has not been initiated) and (c) *power-aware mode* i.e. low power mode is used during the designated idle slots regardless of impending data transmissions/receptions and conditionally used after the contention periods of the remaining transmission slots. All investigated modes are shown with and without the added energy cost presented by the required GPS[9]

[9]the use of *Rikaline* GPS device is assumed which has power consumption of 56 mW [78]

peripheral device in Figure-5.16. It depicts the residual node energy as a function of simulation time (i.e. duration of network operation).

The results show that the lifetime of a network operated by IEEE 802.11 MAC that does not use GPS and the same network operated by PA-STDMA MAC in power-unaware mode with the addition of GPS peripheral devices, is the same (cases *IEEE 802.11 MAC, no GPS* and *PA-STDMA power-unaware mode, GPS* in Figure-5.16). However, if the network requires position information for purposes other than the scheduling of PA-STDMA MAC, a change form IEEE 802.11 MAC to PA-STDMA MAC operated in power-unaware mode would extend the network life time by 11% (cases *IEEE 802.11 MAC, GPS* and *PA-STDMA power-unaware mode, GPS* in Figure-5.16).

Simulations show that the throughput and channel access delay performance of PA-STDMA MAC is the same for both power-unaware mode (a) and optimal throughput performance mode (b). However, if optimal throughput performance mode is chosen over the power-unaware mode, the lifetime of the network is extended by 30% (comparing it with that of IEEE 802.11 MAC in the presence of GPS for both cases).

Through extrapolation, the lifetime of the network operated by PA-STDMA MAC in power-aware mode was found to be 2639 seconds. This is an extension of 55%. The comparison is made under the assumption that both IEEE 802.11 MAC and PA-STDMA MAC operate in the presence of GPS peripheral devices (cases *IEEE 802.11 MAC, GPS* and *PA-STDMA power-aware mode, GPS* in Figure-5.16). In the power-aware mode of PA-STDMA MAC, all nodes enter low power mode in all idle slots. As this reflects on the availability of the nodes in the network the throughput has been found to decline by 14% networkwide. However, the successful packet delivery ratio (from source to destination) is found to be only 3% less than that in the case of improved throughput mode. Hence it can be concluded that PA-STDMA MAC operating in power-aware mode has minimal effect on the reliable operation of the ad hoc network.

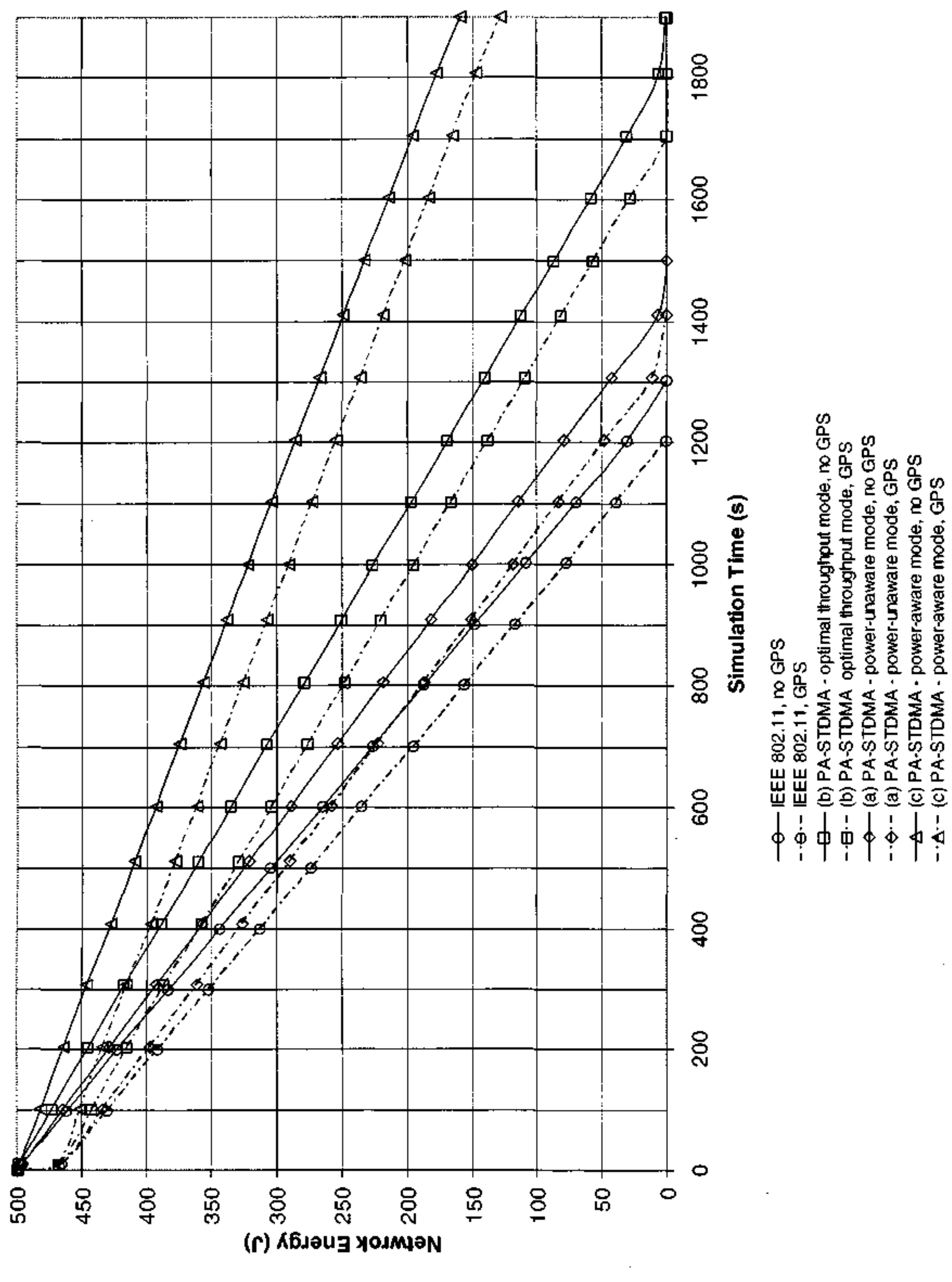


Figure 5.16: Network lifetime

5.5 Design Verification

This section compares the observed performance characteristics of the PA-STDMA MAC protocol against the design specifications from Chapter 4 as follows:

1. The PA-STDMA MAC protocol was designed for structured channel use through integration of the *contention* and *allocation* channel access methods. This was expected to lead to optimal use of the necessary communication overhead and hence low collision probability and reduced congestion. Simulation results confirmed the results from the analytical model showing 60% reduction of communication overhead in comparison to a pure contention channel access control method. Furthermore the reduced communication overhead and the structured use of the communication channel led to 99.8% decrease in packet collisions.

2. The optimal use of necessary communication overhead together with an efficient channel arbitration was expected to lead to improved energy consumption efficiency. The simulation results confirmed that expectation by showing active energy consumption efficiency improvement of 40%.

3. A separate goal of the design was the establishment of the MAC protocol such that little or no throughput efficiency is exchanged for energy efficient operation. The simulation results presented a satisfactory throughput performance of the proposed protocol and confirmed that the structured channel use achieves improved bandwidth utilisation in comparison to the pure contention channel access control. The improvement was observed under moderate to high network node density (over and above 45 nodes) where the PA-STDMA MAC protocol showed 10% higher throughput (in the context of scenario 2).

4. The use of PA-STDMA MAC in a network which makes use of position information for purposes other than the MAC protocol will result in 30% lifetime extension if the protocol is used in optimal throughput mode. If the protocol is set to power-aware mode the extension is 55% with only 3% difference in packet delivery ratio between the two modes;

5.6 On the Shortcomings of PA-STDMA MAC

The following have been considered as shortcomings of the proposed protocol and as such need further attention:

1. The operation of the PA-STDMA MAC protocol requires position information for the purpose of transmission scheduling. An assumption is made that the environment in which the protocol will be used such information is available. The most common source of position information is GPS which is capable of providing high accuracy. However, it is possible that such information is temporary or permanently (in buildings) unavailable. Therefore a mechanism is required that shifts the operation of the protocol from the hybrid TDMA-contention mode to pure contention as the latter does not require position information.

2. For the purpose of the analysis of the protocol an assumption was made that the processed data packets have a size that requires transmission time equal to the duration of the transmission slots. The transmission of packets of variable size could, however, lead to decline in throughput performance due to inefficient use of the transmission slots. This could be avoiding by allowing multiple packet transmissions in the duration of every slot.

3. The restrictions on data transmission posed by the transmission schedules lead to higher channel access delays. This is expected to result in high *end-to-end transmission delays* which could prohibit the use of time sensitive applications in large ad hoc networks characterised with high number of transmission hops. The increased channel access delay for improved channel access control is a trade-off analysed with some of the earliest networking protocols ALOHA and Slotted ALOHA in [9].

4. The restrictions on data transmission posed by the transmission schedules often prohibit immediate packet transmission. This could have a negative effect on on-demand routing in ad hoc networks operating under high mobility (10 m/s and more). Further investigation is required to quantify that effect and investigate to what extend it could be mitigated by traffic prioritisation.

Chapter 6

Conclusion

Infrastructure-free operation that provides *easy network deployment* and voids the *single point of vulnerability* problem are some of the attractive features of mobile ad hoc networks. They make MANETs suitable for use in harsh environments such as disaster stricken areas. battlefields, space etc. Providing these characteristics and at the same time ensuring reliable network operation have kept ad hoc networks developers busy for over a decade.

A popular research topic at the moment is energy efficiency optimisation in mobile ad hoc networks. Efficient energy consumption is needed as mobile ad hoc network nodes are powered by batteries with limited capacity. In addition, battery replacement in the field is often either impractical or impossible. In response to that challenge a number of power optimisation techniques have been proposed.

The first part of the thesis summarised existing power optimisation techniques and their performance. The techniques discussed were of two types: *power control algorithms* and *power aware routing*. It was concluded that power control algorithms are not directly applicable to the general ad hoc network environment. The distributed operation of the investigated algorithms was often associated with convergence to suboptimal solutions or questionable efficiency resulting from intense computational and communication overhead. The second class of power optimisation techniques - power aware routing, was regarded in its two forms found in literature: *unicast* and *multicast* routing. Power aware protocols from the former type presented as much as 95% improvement (in the case of the modified

DSR protocol from Section 2.3.1.3) in energy consumption efficiency in the case of static ad hoc networks. According to the investigated literature, however, as soon as mobility is introduced in the network, the energy consumption performance of these algorithms converges to that presented by the power-unaware routing protocols. The multicast power aware algorithms achieve power efficiency based on the construction of power efficient multicast trees. However, the construction and maintenance of such trees in mobile ad hoc network appears to be complex so the applicability of these algorithms is often limited to static ad hoc networks.

All of the existing power optimisation techniques have in common the fact that they rely on optimised use of transmit power as the only means of achieving energy efficient network operation. Some of the investigations presented in the survey indicated that such an approach could lead to suboptimal energy efficiency. Further evidence had to be obtained to confirm that fact. To that end, a full energy consumption analysis of mobile ad hoc networks was presented in Chapter 3. The analysis explained the typical design of mobile ad hoc network nodes and how the electronics of the mandatory wireless network interface adapter shapes their energy consumption. The energy analysis was performed by investigating the dynamics of ad hoc networks in terms of operation scenarios. The findings of the analysis included:

- broadcast storms present in contention based mediums access control environments and triggered by routing procedures have a significant effect on congestion;
- high collision probability increases the transmission of redundant data packets which generates additional communication overhead forming an avalanche effect;
- unrestrained communication overhead could lead to as much as 60% increase in energy consumption most of which is wasted through transmission of redundant data;
- optimisation of transmit power is applicable only to a few of the existing operation scenarios and hence results in suboptimal power efficiency if used on its own;
- optimal energy consumption efficiency is achieved when the following three condi-

tions are present: optimal use of the necessary control and configuration communication overhead, use of *sleep* mode during idle cycles and optimal use of transmit power.

To address the problem of excessive communication overhead the author of the thesis has proposed a medium access control protocol named Position Aided Spatial-TDMA MAC. It is a hybrid between *contention* and *allocation* channel use. The objective of the protocol is to provide efficient channel arbitration with the introduction of scheduled transmissions in the form of a TDMA system. The scheduling process relies on position information, the protocol is thus designed for environments in which such information is already in use. Contention for data transmission is allowed under certain circumstances in order to offset the shortcomings of the allocation approach. Full description of the protocol was provided in Chapter 4. The chapter also presented an analysis of the following three performance characteristics of the protocol: *energy efficiency* as a function of optimal use of control communication overhead, *throughput* and *channel access delay*. The results of the analysis indicated up to 60% increased energy efficiency (in comparison to a pure contention protocol such as IEEE 802.11 MAC) and a satisfactory throughput performance. The channel access delay was expected to rise as a function of network node density.

Verification of the analytical results was established with the help of simulations of the proposed protocol. The performance of the PA-STDMA MAC was also compared with that of IEEE 802.11. The used simulation environment was NS2. The results of the simulations confirmed most of the analytical results. According to the simulation results the PA-STDMA MAC achieves 40% increase in energy efficiency in comparison to the pure contention based mechanism. The throughput performance of the investigated protocols is comparable. It was confirmed that under moderate to high node densities (over and above 45 nodes in a 600m by 600m network area) the PA-STDMA MAC offers 10% better throughput performance than IEEE 802.11 MAC. The observed channel access delay presented by PA-STDMA MAC was smaller than that predicted by the analytical model but still sufficiently high to limit the use of some time sensitive applications in large ad hoc networks.

Simulations helped to investigate the expected network lifetime extension resulting from the energy efficient operation of the proposed protocol. It was established that if PA-STDMA MAC is used in a network that already processes position information the network lifetime extension could be either 30% if the MAC protocol is used in optimal throughput configuration or as much as 55% if it is used in power-aware mode. The use of power-aware mode was found to result in 3% decline in packet delivery ratio.

6.1 Summary of Contributions

The following is a list of the contributions of the thesis:

- a comprehensive survey of existing power consumption optimisation techniques applicable to mobile ad hoc networks;
- an extension to an existing analytical model allowing analysis of the effect of network congestion on energy consumption efficiency
- analysis of the effect network congestion has on energy consumption efficiency;
- a TDMA-contention hybrid medium access control protocol for mobile ad hoc networks that achieves optimal use of communication overhead and energy efficiency increase of 40%;
- performance comparison between two types of channel access control: contention based and TDMA-contention hybrid;

Bibliography

[1] J. C. Cano and P. Manzoni, "A performance comparison of energy consumption for mobile ad hoc networks routing protocols," in *Proc. MASCOTS*, San Francisco CA, 2000.

[2] ANSI/IEEE Std 802.11, "Ieee std 802.11-1999, wireless lan medium access control (mac) and physical layer (phy) specifications," 1999.

[3] P. Mavinkurve, H. Ngo, and H. Mehra, "Mip3s: Algorithms for power-conserving multicasting in static wireless ad hoc networks," in *Proc. 11th IEEE International Conference on Networks (ICON '03)*, Sydney, Australia, 2003.

[4] M. Cagalj, J.-P. Hubaux, and C. Enz, "Minimum-energy broadcast in all-wireless networks: Np-completeness and distribution issues," in *Proc. MobiCom*, Atlanta, USA, 2002.

[5] J. Wieselthier, G. Nguyen, and A. Ephremides, "On the construction of energy-efficient broadcast and multicast trees in wireless networks," *in Proc. IEEE INFOCOM*, 2000.

[6] S. Gupta and B. Wang, "S-remit: A distributed algorithm for source-based energy efficient multicasting in wireless ad hoc networks," in *Proc. IEEE 2003 Global Communications Conference (GLOBECOM)*, San Francisco, CA, 2003.

[7] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "Broadcast storm problem in a mobile ad hoc network," in *Proc. MOBICOM*, 1999.

[8] K. Fall and K. Varadhan, "The ns manual," http://www.isi.edu/nsnam/ns/doc/.

[9] D. Bertsekas and R. Gallager, *Data Networks.* Prentice Hall, 1992.

[10] P. Karn, "Maca: A new channel access protocol for packet radio," *In Proc. of ARRL/CRRL Amateur Radio 9th Conf.*, 1990.

[11] wikipedia, "OSI Model," http://en.wikipedia.org/wiki/OSI_model.

[12] N. Pate, "Power control algorithms in ad hoc networks," in *Proc. SATNAC*, 2001.

[13] S. Narayanaswamy, V. Kawadia, R. Sreenivas, and P. Kumar, "Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the compow protocol," in *Proc. European Wireless 2002. Next Generation Wireless Networks: Technologies, Protocols, Services and Applications*, 2002.

[14] S. Agarwal, R. Katz, S. Krishnamurthy, and S. Dao, "Distributed power control in ad-hoc wireless networks," *in Proc. PIMRC 2001*, 2001.

[15] R. Wattenhofer, L. Li, P. Bahl, and Y.-M. Wang, "Distributed topology control for power efficient operation in multihop wireless ad hoc networks," in *Proc. INFOCOM*, Alsaka, 2001.

[16] M. Petterson and P. Lehne, "An overview of smart antenna technology for mobile communications systems," *IEEE Communication Surveys*, vol. 2, 1999.

[17] P. Kumar and V. Kawadia, "Power control and clustering in ad hoc networks," in *Proc. INFOCOM*, San Francisco, 2003.

[18] P. Kumar and P. Gupta, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. IT-46, pp. 388–404, 2000.

[19] P. Bhagwat and C. Perkins, "Highly dynamic destination-sequenced distance-vector routing (dsdv) for mobile computers," *Computer Communication*, pp. 234–244, 1994.

[20] M. Corson and V. Park, "Temporally-ordered routing algorithm (tora) version 1 - functional specification," *MANET Working Group*, 1999.

[21] D. Maltz, J. Broch, and J. Johanson, "The dynamic source routing protocol for mobile ad hoc networks," *MANET Working Group*, 1999.

[22] E. Royer and C. Perkins, "Ad hoc on-demand distance vector (aodv) routing," *MANET Working Group*, 2000.

[23] M. Woo, S. Singh, and C. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. ACM/IEEE MOBICOM*, 1998.

[24] N. Dearham, T. Quazi, and S. McDonald, "A comparative assessment of ad-hoc routing protocols," in *Proc. SATNAC*, 2002.

[25] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark, "Scenario-based performance analysis of routing protocols for mobile ad-hoc networks," in *Proc. MobiCom*, Seattle, Washington USA, 1999.

[26] S. Lee, M. Gerla, and C. Toh, "A simulation study of table-driven and on-demand routing protocols."

[27] J. Broch, D. Maltz, D. Johanson, Y. Hu, and J. Jetcheva, "A performance comparison of muti-hop wireless ad hoc network routing protocols," in *Proc. MobiCom*, Dallas, Texas, 1998.

[28] D. Fulkerson and L. Ford, "Flows in networks," *Princeton University Press*, 1962.

[29] R. Bellman, "Dynamic programming," *Princeton University Press*, 1957.

[30] S. Doshi and T. Brown, "Minimum energy routing schemes for a wireless ad hoc network," in *Proc. IEEE INFOCOM*, 2002.

[31] S. Bhandare, S. Doshi, T. Brown, and S. Sanghani, "Comparison of two wireless ad hoc routing protocols on a hardware test-bed," in *Proc. IEEE WCNC*, 2003.

[32] H. Labiod and H. Moustafa, "Adaptive path energy conserving routing in manets," in *Proc. Mobile and Wireless Communication Networks (MWCN)*, Singapour, 2003.

[33] D. Maltz, J. Broch, J. Jetcheva, and J. Johanson, "The effects of on-demand behaviour in routing protocols for muti-hop wireless ad hoc networks," in *IEEE JSAC*, vol. 17, 1999, pp. 1439–1453.

[34] S. Sanghani, T. Brown, S. Bhandare, and S. Doshi, "Ewant: Emulated wireless ad hoc test-bed," in *Proc. IEEE WCNC*, 2003.

[35] G. Pujolle and S. Senouci, "Energy efficient routing in wireless ad hoc networks," *IEEE Communication Society*, 2004.

[36] H. Mouftah and B. Zhang, "Localized power-aware routing for wireless ad hoc networks," in *Proc. ICC*, 2004.

[37] J. Zhu, C. Qiao, and X. Wang, "A comprehensive minimum enregy routing scheme for wireless ad hoc networks," in *Proc. INFOCOM*, 2004.

[38] A. Misra and S. Banerjee, "Minimum energy paths for reliable communication in multi-hop wireless networks," in *Proc. MOBIHOC*, 2002.

[39] A. Ephremides, J. Wiesekthier, and G. Nguyen, "Algorithms for energy-efficient multicasting in static ad hoc wireless networks," *Mobile Networks and Applications*, vol. 6, no. 3, pp. 251–263, 2001.

[40] S.-J. Lee, W. Su, and M. Gerla, "On-demand multicast routing protocol (odmrp) for ad hoc networks," *IETF, Internet Draft*, 1999.

[41] N. H. Vaidya and Y.-B. Ko, "Location-based multicast in mobile ad hoc networks," Tech. Rep., 1998.

[42] M. Gerla, C.-C. Chiang, and L. Zhang, "Tree multicast strategies in mobile, multihop wireless networks," *ACM/Balzter Mobile Networks and Applications Journal*, 1998.

[43] C.-K. Toh, G. Guichala, and S. Bunchua, "Abam: On-demand associativity-based multicast routing for ad hoc mobile networks," in *Proc. IEEE International Conference on Vehicular Technology*, 2000.

[44] Z. Haas and S. Sajama, "Independent-tree ad hoc multicast routing (itamar)," *Mobile Networks and Applications*, vol. 8, no. 5, pp. 551–566, 2003.

[45] C. Perkins and E. Royer, "Multicast operation of the ad-hoc on-demand distance vector routing protocol," in *Proc. ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, 1999.

[46] M. Liu, R. R. Talpade, A. McAuley, and E. Bommaiah, "Amroute: Adhoc multicast routing protocol," University of Maryland and the Institute of Systems Research, Tech. Rep., 1999.

[47] C. W. Wu and Y. C. Tay, "Amris: A multicast protocol for ad hoc wireless networks," in *Proc. MILCOM*, 1999.

[48] J. J. Garcia-Luna-Aceves and E. L. Madruga, "The core-assisted mesh protocol," *Selected Areas of Communication*, pp. 1380–1394, 1999.

[49] S. Athanassopoulos, I. Caragiannis, C. Kaklamanis, and P. Kanellopoulos, "Experimental comparison of algorithms for energy-efficient multicasting in ad hoc networks," in *Proc. AdHocNow*, Vancouver,British Columbia, 2004.

[50] "Ip routing," 2001, http://www.hep.ucl.ac.uk/ỹtl/multi-cast/iprouting_01.html.

[51] P.-J. P.-J Wan, G. Calinescu, X.-Y. Li, and O. Frieder, "Minimum-energy broadcast routing in static ad hoc wireless networks," in *Proc. INFOCOM*, 2001.

[52] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*. McGraw-Hill Science/Engineering/Math, 2001.

[53] CISCO, "Open shortest path first protocol," 2002.

[54] P. Mavinkurve, H. Ngo, and H. Mehra, "Dmip3s: Algorithms for power-conserving multicasting in static wireless ad hoc networks," in *Proc. 11th IEEE International Conference on Networks (ICON '03)*, Sydney, Australia, 2003.

[55] I. Kang and R. Poovendran, "A comparison of power-efficient broadcast routing algorithms," in *Proc. IEEE Global Communications Conference (Globecom)*, San Francisco, CA, 2003.

[56] A. Rubini and J. Corbet, *Linux Device Drivers*. www.orelly.com, 2001.

[57] A. John, *Digital Transmission Engineering*. IEEE Press, Prentice Hall, 1999.

[58] Lekernel, "Prism45 chipset overview," 2005, http://lekernel.lya-fr.com/prism54.html.

[59] R. Katz and M. Stemm, "Measuring and reducing energy consumption of network interfaces in hand-held devices," *IEICE Transactions on Communications*, vol. E80-B, pp. 1125–37, 1997.

[60] J.-P. Ebert, S. Aier, B. B. Gunnar Kofahl, Alexander Becker, and A. Wolisz, "Measurement and simulation of energy consumption of an wlan interface," TKN, Tech. Rep., 2002.

[61] W. Qadeer, T. Rosing, J. Ankcorn, V. Krishnan, and G. De Micheli, "Heterogeneous wireless network management," in *Proc. PACS*, San Diego, California, 2003.

[62] M. Nilsson and L. Feeney, "Investigating the energy consumption of an wireless network interface in an ad hoc network environment," in *Proc. IEEE Infocom*, Anchorage, Alaska, 2001.

[63] Y. Tay and K. Chua, "A capacity analysis for ieee 802.11 mac protocol," *ACM Wireless Networks*, pp. 159–171, 2001.

[64] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," *IEEE Selected Areas in Communications*, 2000.

[65] Q. Zhao and L. Tong, "Energy efficiency of large-scale wireless networks: Proactive versus reactive networking," *IEEE/ACM Wireless Networks*, pp. 159–171, 2001.

[66] I. Chlamtac, A. Fargo, A. D. Myesr, V. R. Syrotiuk, and G. Zaruba, "Adapt: A dynamic self-adjusting media access control protocol," in *Proc. Globecom*, 1999.

[67] M. Corson and C. Zhu, "A five-phase reservation protocol (fprp) for mobile ad hoc network," in *Proc. IEEE Infocom*, 1998.

[68] J. Garcia-Luna-Aceves and L. Bao, "A new approach to channel access scheduling for ad hoc networks," in *Proc. ACM Mobicom*. 2001.

[69] L. Kleinrock and R. Nelson, "Spatial tdma: A collision-free multihop channel access control for packet radio networks," *IEEE Transactions on Communication*, vol. COM-33, 1985.

[70] K. G. Amouris, "Space-time division multiple access (stdma) and coordinated, power-aware maca for mobile ad hoc networks," in *Proc. IEEE Symposium on Ad Hoc Wireless Networks (SAWN2001)*, San Antonio, Texas, 2001.

[71] J.-H. Li and V. Ju, "An optimal topology-transparent scheduling method in multihop packet radio networks," *IEEE Transactions on Networking*, vol. 6, pp. 298–306, 1998.

[72] I. Stavrakakis and K. Oikonomou, "A probabilistic topology unaware tdma medium access control policy for ad-hoc environments," in *Proc. Personal Wireless Communications*, Venice, Italy, 2003.

[73] K. Oikonomou, N. Pronios, and I. Stavrakakis, "Performance analysis of tdma mac schemes for ad-hoc networks with topology control," in *Proc. Med-Hoc-Net*, 2004.

[74] D. J. Bird, "The use of gps for high relative timing accuracy for hires sites," in *Proc. XXIV Cosmic Ray Conference*, 1995.

[75] W. Navidi, T. Camp, and N. Bauer, "Improving the accuracy of random waypoint simulation throughput steady-state initialization," in *Proc. 15th International Conference on Modelling and Simulation (MS'04)*, 2004.

[76] W. Navidi and T. Camp, "Stationary distribution for the random waypoint mobility model," in *IEEE Transaction on Mobile Computing, vol. 3*, 2004.

[77] I. Chlamtac, A. Fargo, A. D. Myesr, V. R. Syrotiuk, and G. Zaruba, "A performance comparison of hybrid and conventional mac protocols for wireless networks," in *Proc. IEEE Vehicular Technology (VTC)*, 2000.

[78] Rikaline, *Rikaline GPS receiver module user's guide*. Rikaline.