The Dissertation Committee for John Andrew Stine

Certifies that this is the approved version of the following dissertation

Energy Conserving Protocols for Wireless Data Networks

Committee:

Gustavo de Veciana, Supervisor

Ross Baldick

Ross Baldick

Simon S. Lam

Hao Ling

Edward J. Powers

Wolfhard Vogel

Energy Conserving Protocols for Wireless Data Networks

by

John Andrew Stine, B.S, M.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2001

Dedication

To Carl Joseph Stine and what could have been.

Acknowledgements

Raising a family of six children, working full time, and completing a degree 750 miles from campus can only be done through the concessions of many people. First I would like to thank my advisor Dr. Gustavo de Veciana who willingly took me on as a graduate student knowing that this would be a long distance arrangement. His patience, responsive assistance, willingness to adjust his schedule to accommodate mine and encouragement were key to me completing this research. Second, I thank my bosses at work who allowed me time off to do much of my work, most specifically, COL C. Rundell Galles, LTC Donald Olson, LTC Dana Hurst, LTC John Hetzel, Jr., and LtCol Donald Walters. Third, I thank my wife's parents, COL(ret) Leon Byrd and Mary Louise Byrd for providing me a place to stay during my many visits to Austin to coordinate my research. Finally, and most significantly, I need to thank my wife and children for their support over the past five years. My wife Martha has borne the lion's share of raising our children and managing the household. She graciously accepted my many nights and weekends away doing research and remained a stalwart supporter of my completing my degree. My children, Andrew, Harry, Brian, Mazie, Bridget, and Christopher motivated me more than they will ever know.

Energy Conserving Protocols for Wireless Data Networks

Publication No. _________

John Andrew Stine, Ph.D.
The University of Texas at Austin, 2001

Supervisor: Gustavo de Veciana

In this thesis we study how wireless medium access control (MAC) protocols can be used to conserve energy in wireless portable communication systems. We start by evaluating protocols for centrally controlled networks separately evaluating the protocols used in the management of the exchange of data from those used by nodes to announce to the central controller that they have data to exchange. In the former, we compare different methods of scheduling, methods to announce schedules, and methods to recover from errors. In the latter, we compare various random access and polling techniques. We conclude that there are two significant energy conserving objectives. First, the primary goal of MAC protocols should be to put nodes not participating in data exchanges into a low energy state. Second, information that enables these nodes to enter these states must be made available to them as soon as they can use it. We then consider energy consumption in ad hoc networks. The objectives identified for centrally controlled networks are applicable but are not easily applied because most ad hoc MAC protocols rely on temporally random access techniques.

As a result, we created a new MAC protocol called Synchronous Collision Resolution (SCR). SCR assumes nodes are synchronized and thus can contend simultaneously. A collision resolution signaling mechanism is used to resolve contentions and request-to-send and clear-to-send exchanges are used to guard against hidden nodes. SCR has many other enhanced capabilities. In particular, the protocol includes a novel priority access scheme and a cooperative signaling approach to robustly support mutlihop stream based services. The proposed mechanisms for resolving contention are especially effective at identifying a spatially distributed set of transmitting nodes that can exchange data simultaneously. We show through analysis and simulation that the protocol is stable to spatial loads exceeding 40%, avoids congestion collapse, and has spatial capacity exceeding 60%. Moreover, we show that next hop routing strategies and the use of spread spectrum coding can more than double these performance measures. We conclude by arguing that the network can easily adapt to congestion (high load and/or node density) by varying the transmission power used during the contention process. Overall, SCR achieves performance levels in ad hoc networks that have heretofore only been possible for centrally controlled networks.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Minimizing energy consumption is a major concern in the design of wireless portable communication systems. Techniques for reducing energy consumption have penetrated all levels of the design hierarchy, including the algorithmic and architectural levels down to specific circuits and technologies. The objective of the research reported in this dissertation was to study how coordination among transmitters and receivers and knowledge of the state of the network could be used to reduce energy consumption. Specifically, we have provided a collection of energy conserving protocol options that designers of single channel wireless data networks can choose from in creating a network suitable for their applications. We also introduce a novel access protocol for ad hoc networks that was created to conserve energy but has many additional characteristics that make it very effective at achieving other ad hoc networking goals. The dissertation follows the order in which the research was conducted. It starts with an overview of the mechanisms available to protocols to conserve energy. It then looks at the use of these mechanisms first in centrally controlled networks and then in ad hoc networks.

Chapter 2 presents basic energy conserving mechanisms that protocols can employ. It describes how they conserve energy, reviews current research and proposed methods to use them, and, finally, proposes a hierarchy of which mechanisms offer the greatest potential to conserve energy. We claim that protocols should first attempt to use low energy transceiver states to conserve energy, then reduce overhead, leverage the use of nodes that do not need to conserve energy to the benefit of those that do, and finally attempt to use low energy routes.

In Chapters 3 and 4 we consider protocols that are used with centrally controlled networks. We assume in both chapters that centrally controlled networks alternate between two periods. In the first period, called the contention period (CP), nodes of the network contend to gain access to coordinate their transmission with the centrally controlling node called a point coordinator (PC). In the second period, called the contention free period (CFP), the PC manages the exchange of traffic amongst the nodes of the network. Chapter 3 presents our exploration of how a PC can best manage data exchange in the CFP to conserve energy. The primary mechanism we consider in conserving energy is the use of low energy states. The goal is to put all nodes not involved in a transmission into a doze state. However, in doing so, one must tradeoff the energy cost of coordinating dozing with the energy savings of putting nodes to sleep. In this chapter, we define three alternative directory protocols that may be used by the PC to coordinate the transmission of data and the dozing of nodes. We attempt to optimize their performance by using scheduling and protocol parameter tuning. In addition, we consider the impact of errors and error recovery methods on energy consumption. Although one can argue that carefully scheduling transmissions will improve performance, ultimately, appropriately tuning protocols reduces scheduling's significance. In most cases, scheduling transmissions between the same nodes contiguously and ordering such transmissions shortest processing time first results in good performance. However, the ability of our protocols to conserve energy is highly dependent on 1) network size, 2) traffic type (e.g. down/uplink, and peer-to-peer) and 3) channel bit error rate. In particular, we show that when protocols are faced with packet errors, more elaborate schemes to coordinate the dozing of nodes can pay-off. Our simulations show that while energy savings can vary by a factor of 10 over the class of protocols we considered, throughput varies by less than 20%. We find that the single most important factor in conserving energy is ensuring that nodes that can doze do so. This involves insuring that they receive information on when to doze as soon as they can use it.

In Chapter 4 we look at the how protocols should be designed for the contention period. In this chapter we propose and evaluate seven protocols that may be used by a PC to grant access while supporting energy conservation. We consider both random access protocols (i.e. p-Persistent Slotted Aloha, Time Slotted p-Persistent Carrier Sense Multiple Access, Elimination Yield Non-Preemptive Multiple Access, and Modified Random Addressing Protocol) and various polling protocols. We define and model each of these protocols, provide an optimization strategy to select operational parameters, and finally compare the protocols using numerical analysis. Our comparisons consider energy consumption, throughput, and their stability in changing loads. We find that on account of the short access packets used in the contention periods which allow many access attempts to be made in a short period of time, polling protocols are very attractive. Their very predictable nature is ideal for energy conservation since nodes can know exactly when they can doze and since the access packets are short, their normal shortcoming of throughput is not so significant. Polling achieves the least energy consumption and adapts best to changing loads. Additionally, the polling protocols are least affected by hidden node effects. For low load conditions some of the random access protocols offer exceptional performance. We note, however, that the scale of the energy that can be conserved during this period is not as large as that which can be conserved during the CFP since nodes that do not plan to contend can already doze throughout the CP without any assistance from the protocol and since the packets that are used are small. We also note that when using a polling protocol a third access period is required to allow new nodes to associate with the network. For these reasons one of the random access protocols may still be best for the application. Of the random access protocols, simple slotted aloha provides the best energy conserving performance.

Our efforts to find the best methods to first manage the scheduling and transmission of traffic in Chapter 3 and of gaining access in Chapter 4 for centrally controlled networks drive home two points. First, the primary goal is to put nodes not participating in data exchanges into a low energy state. Second, that to best put idle nodes into low energy states, information that enables these nodes to enter these states must be made available to them as soon as they can use it. Unfortunately, neither of these lessons are easily applied to the access protocols that are most often used for ad hoc networks since they are based on temporal randomness. (i.e. Access is gained since contending nodes attempt to gain access at different times that are randomly selected.) This randomness makes it difficult to predict which nodes need to be awake at any particular time. Energy conserving techniques that exist for these type of access protocols risk lower performance. Their optimization requires consideration of past traffic patterns to predict the viability of entering a low energy state. Stated another way, optimization of these protocols is highly dependent on whether past traffic patterns are a good indication of what will occur in the future. Since such a study departs from the goal of identifying the best mechanisms within protocols to conserve energy we decided to create a new protocol that offers the predictability that enables the main lessons of Chapter 3 and 4 to be applied to an ad hoc network.

In Chapter 5 we present our new protocol which we call Synchronous Collision Resolution (SCR). SCR still allows statistical multiplexing but is not temporally random. It assumes nodes are synchronized. Thus, nodes contend simultaneously for channel access. SCR uses a collision resolution signaling mechanism that is better than 98% effective at yielding a successful contender per contention within a given spatial area. This signaling is followed by request-to-send (RTS) and clear-to-send (CTS) exchanges to guard against hidden nodes. We prove that this exchange eliminates collisions in subsequent data exchanges. The synchronous nature of SCR provides the predictability that enables both of the energy conserving lessons of Chapters 3 and 4 to be applied. Nodes know exactly when contentions occur and when data is transmitted. Thus, nodes can enter a low energy state whenever they are not participating in a data exchange when one is scheduled. Since they know when contentions occur they know exactly when they can determine if they will be participating in a data exchange.

Perhaps of even greater significance are the other performance characteristics of SCR. The synchronous nature of SCR enables quality of service (QoS) differentiation. In particular, the protocol includes a novel priority access scheme and a cooperative signaling approach to robustly support mutlihop connection oriented services. The proposed mechanisms for resolving contention are especially effective at identifying a spatially distributed set of transmitting nodes that can exchange data simultaneously. We show through analysis and simulation that the protocol is stable to spatial loads exceeding 40%, avoids congestion collapse, and has spatial capacity exceeding 60%.

In Chapter 6 we extend our investigation of SCR to determine how its capacity can be improved. We consider two approaches, the use of next hop routing strategies and the use of spread spectrum (SS) coding. We show that next hop routing strategies more than double the range of the load for which the network remains stable and similarly more than double the spatial capacity. We show that when SS coding is used together with SCR the spatial capacity of SCR can also be more than doubled. The structured nature of SCR makes the solutions to the typical problems associated with using SS codes in distributed networks, assigning codes and knowing which codes to use when, trivial. It allows the use of separate codes for peer-to-peer traffic versus broadcast traffic without confusion as to which will be used. The mechanisms SCR uses for resolving contentions also prove to resolve a third challenge in using SS codes, distributing source destination pairs to reduce interference. We show through analysis and simulation that the combined use of these protocols is stable to spatial loads up to 1 arrival per transmission area per transmission slot, avoids congestion collapse, and has spatial capacity exceeding 1.2 transmission per transmission area. We also demonstrate through simulation that even in highly dynamic environments where there are fewer codes than there are nodes within range of each transmitter that the protocol remains effective. SCR and SS coding are perfectly complementary protocols that yield unprecedented performance in ad hoc networks. We conclude by arguing that the network can easily adapt to congestion (high load and/or node density) by varying the transmission power used during the contention process. Overall, SCR achieves performance levels in ad hoc networks that have heretofore only been possible for centrally controlled networks.

SCR has additional performance benefits. Although not included in this dissertation we have shown that all the energy conservation techniques described in Chapter 2 can be integrated with a routing protocol on account of the characteristics of SCR.[1]

This dissertation provides a comprehensive treatment of energy conservation techniques that are used in access and routing protocols but its most significant contribution is the introduction of a new protocol for ad hoc networking. This protocol,

---

[1] This research has not been published yet.

SCR, not only supports energy conservation but also results in great improvements in spatial capacity and is the first random access protocol for ad hoc networks that enables the creation of stream based connections. SCR is a very robust protocol that remains highly effective even in highly congested networks.

Chapter 2
Energy Conservation Mechanisms

2.1 Introduction

There are four potential sets of mechanisms that can be used to reduce energy consumption.

1. The first set of mechanisms attempts to reduce power by helping individual terminals to enter low energy states. For example, reducing the time terminals spend monitoring traffic. Terminals do not need to monitor all traffic. Processing traffic that is inconsequential to a terminal is wasteful. The objective of this mechanism is to allow mobile terminals to turn themselves off when not needed for data transmission. Other intermediate low energy states may also be used depending on the capabilities of the node.

2. The second set of mechanisms attempts to reduce energy consumption by routing traffic through the network. The power of a received signal is a power law function of the distance the receiver is separated from the transmitter. As a result, transmission across multiple hops along the same distance or even higher distance can consume less energy than the single hop transmission. The objective of these mechanisms is to find the optimum energy conserving route that meets delay con-

straints. Inherent in this analysis is defining an algorithm that considers energy consumption in choosing routes.

3. The third method involves the judicious use of member nodes of the network. Nodes that are not energy constrained or have large energy reserves can be leveraged to haul a greater portion of the network traffic or to perform more of the network's administrative tasks so energy constrained nodes can survive in the network longer.

4. The fourth method is to decrease the overhead to payload ratio.

In this chapter we look more closely at the potential these mechanisms have to conserve energy and review current efforts to employ them in existing and proposed protocols. Sections 2.2 through 2.5 each cover one of the first three mechanisms listed above. In Section 2.6 we briefly discuss the hierarchy of the energy conserving mechanisms. Section 2.7 concludes the chapter.

2.2 Conserving Energy Using Low Energy Transceiver States

2.2.1 Transceiver States

Energy consumption in a transceiver is hardware dependent. Generally, three energy consumption states are defined: transmitting, receiving, and dozing. Table 2.1 presents the energy consumption rates in these three states for some commercial and experimental transceivers. Except when transmitting the default state is receiving.[2]

---

[2] Some papers only account for energy consumed when data is transmitted ignoring energy consumed during idle periods evidently assuming the doze state is default except during data transmission. This assumption results in analysis that over emphasizes the energy cost of overhead and the selection of protocols that forfeit the opportunity to conserve energy during idle periods. Energy consumption during idle periods while in the receive state is empirically confirmed in [126], [8], and [9].

RADIO	TRANSMIT	RECEIVE	STAND-BY/ DOZE
Lucent WaveLAN/IEEE Turbo 11 Mb Card [1]	285mA	185mA	9mA
RoamAbout 915 MHz DS/ISA [2]	600mA	300mA	36mA
RoamAbout 2.4 GHz DS/ISA [2]	365mA	315mA	30mA
RoamAbout 2.4 GHz FH/ISA [2]	325mA	185mA	5mA
2.7V GSM RF Transceiver [3]	31mA	42mA	1μA
Nokia C020/C021 Wireless LAN Card [4]	1.7W	1.3W	0.2/0.1W
Aironet PC4800B In-Building Client Adapter [5]	350mA	250mA	<10mA

Table 2.1: Digital Radio Power States

MAC protocols conserve energy by either reducing transmission time required for data exchanges or, more significantly, by assisting terminals to enter a low energy doze state where the terminals neither transmit nor receive data.

An important issue in designing protocols that assist nodes to enter a low energy state is the time it takes for transceivers to enter and then to leave the state. Although protocols like the IEEE 802.11 wireless LAN standard [6] and HIPERLAN [7] specify the transition times between the transmit and receive states, 6 and 5 μsec respectively, they do not specify the transition times between these states and the doze state. A transition time of 100ms was empirically identified in [8] for an earlier model 802.11 LAN card. With this sort of delay there is little motivation to enter the low energy state unless the node is likely to doze for an extended period of time. Additional concerns for determining a general design for energy conserving techniques is that whatever transition times can be achieved they may not scale well with data transmission rates.

In this work we have assumed that transition times to low energy states are comparable to those between the transmit and the receive states. We believe that protocol design based on this assumption is more general. Long term dozing techniques are not compromised by this assumption. If transition rates cannot be improved for entering the doze state then these long transition times would motivate the creation of additional low energy states that do not conserve as much energy but can be entered and left more quickly. For example, the empirical results in [9] indicate that such a state is present in the WaveLAN 802.11 card when it discards a packet that another node is transmitting. The energy used in this state is about 25% less than that consumed in the standard receive mode. So protocols that effectively assist nodes to identify opportunities to enter low energy states for any length of time are the general solution to using low energy states to conserve energy.

2.2.2 Mechanisms for Centrally Controlled Networks

Centrally controlled networks leverage the relative omniscience of the base station or point coordinator (PC) and allow it to schedule the activities of nodes. Through the announcement of these schedules, nodes learn when they can doze. These networks alternate between periods where the PC learns of the pending traffic in the network and periods where it manages its exchange. There are two versions, one where the network uses two types of periods and one where it uses three types.

In the two period version, the network alternates between a contention period (CP) and a contention free period (CFP). Nodes that want access to the network to send packets in the CFP contend in the CP and inform the PC of their intent. The PC takes the information in their request and then schedules their transmissions in a subsequent CFP. The PCs conserve energy in the manner they announce and schedule transmissions during the CFP. The PC provides information at the beginning of the CFP that informs nodes whether they can doze. The 802.11 MAC uses a device called a traffic indication map (TIM). This is nothing more than a bitmap where every node in the network is assigned a bit. When a node's bit is set it is an indication that that node will participate in a data exchange in the current CFP. So, at the beginning of the CFP, the PC transmits the bitmap and nodes whose bits are not set may doze throughout the CFP. Another similar approach has the PC announcing the entire transmission schedule. With this greater quantity of detail, nodes can determine the exact times they must be awake. We compare these alternatives as well as some hybrid techniques of the same type in Chapter 3.

In the three period version, the network alternates between a polling period and a CFP and then at a longer period adds a CP for new nodes to associate with the network. The energy conserving medium access control protocol (ECMAC) is an example of this type of protocol [10-13]. The advantage of this approach is that polling is a more energy efficient way for PCs to learn of the pending traffic at the nodes it controls. Polling can be scheduled so nodes with traffic to send need only be awake during their scheduled poll. Currently, this is the only technique that has been proposed to support energy conservation in the process that the PC uses to learn of pending traffic. In Chapter 4 we consider polling and develop some new contention period techniques that conserve energy.

2.2.3 Mechanisms of Ad Hoc Networks

Conserving energy in ad hoc networks is much more challenging because of the temporally random nature of most access protocols. (i.e. Nodes gain access by contending at different times that are randomly selected.) On account of this randomness, other nodes can never be sure they will not be an intended destination of traffic and must remain awake or else compromise the throughput performance of the access protocol. There have been two general approaches to manage dozing nodes. The first allows individual nodes to choose whether they will doze but requires them to be awake at certain times. The second allows individual nodes to doze if they can coordinate for a surrogate to remain awake and collect traffic intended for them while they are dozing. The 802.11 and HIPERLAN standards provide examples of the two

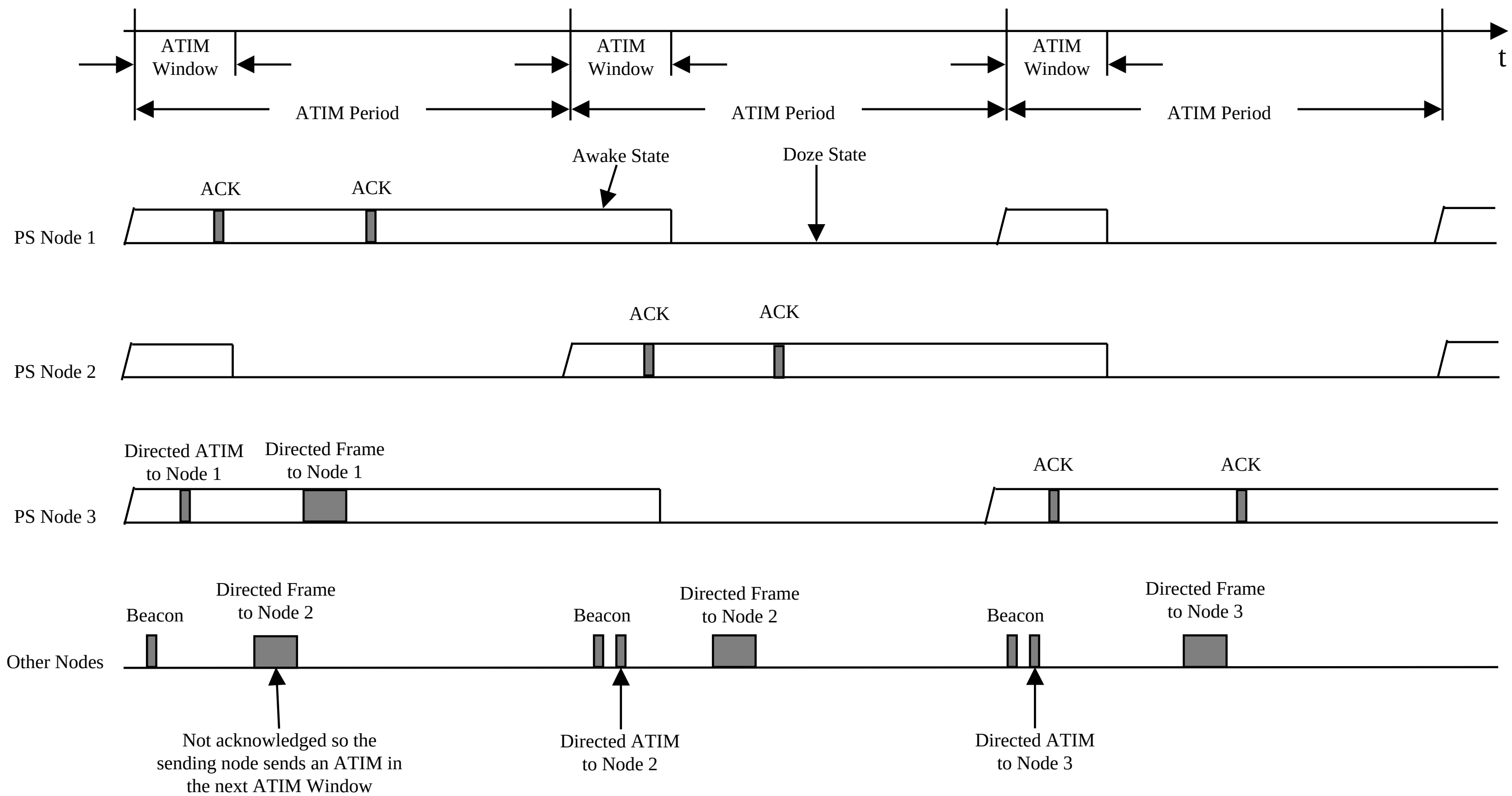


Figure 2.1: IEEE 802.11 Energy Conservation Mechanism for Ad Hoc Networks

different techniques. The 802.11 technique is an example of the first approach whereas the HIPERLAN technique is an example of the second. Both come at a cost of lower network performance.

In both the 802.11 and HIPERLAN protocols, the decision to doze is initiated by the individual nodes desiring to conserve energy. In the ad hoc version of an 802.11 network, an Independent Basic Service Set (IBSS), the node that first forms the network decides whether it permits energy conservation by establishing an "ATIM Period." If this value is greater than 0 then a node in the network that desires to conserve energy may doze so long as it wakes each ATIM Period to listen for ad hoc traffic indication messages (ATIM). ATIMs are transmitted during a short window at the beginning of each ATIM period, called an ATIM Window. If the node wakes and hears an ATIM directed to itself, it acknowledges the ATIM and remains awake for the rest of the ATIM period prepared to receive traffic. If it receives no ATIM directed to itself, the node returns to the doze state at the conclusion of the ATIM window. Note that there is no formal method for a node's intent to doze to be disseminated. Other nodes assume this state after failing to transfer data through regular contention. This process is not specified in the standard. Figure 2.1 illustrates an example of the IEEE 802.11 energy conservation mechanism. In this example, three nodes are in the power save mode. All wake up for the ATIM window and then remain awake if they received a directed ATIM or if they need to transmit data. If a node sends them data when they are dozing it is up to that node to determine whether it needs to send an ATIM or simply reattempt to send the data. Transmission errors cannot be distinguished from dozing.

The energy conserving mechanism in HIPERLAN requires a node desiring to doze, a "p-saver," to coordinate with another to serve as its surrogate, a "p-supporter".[3] As part of this coordination the two nodes agree to a period at which the p-saver will awaken to receive unicast messages and a period at which the p-supporter will transmit multicast messages. The p-supporter node collects transmissions intended for the p-saver and then attempts to forward them to the p-saver during the coordinated transmission periods. Neither protocol standard provides recommended parameters for their operation. In the case of the 802.11 protocol the ATIM period and ATIM window are selected at the creation of the network and they never change. Additionally, in both protocols, the decision of individual nodes to doze is made without benefit of knowing its effect on the overall network's performance.

The 802.11 wireless network protocol was studied in [14] and [15]. The authors determined that shorter ATIM periods result in higher energy savings and that the ATIM window should be about 1/4th the size of the ATIM period. The intuition the authors provide is that for longer ATIM periods it becomes more likely that nodes will have a need to transmit an ATIM and in turn to remain awake during the ATIM period. Additionally, the longer the ATIM window, the more ATIMs will be transmitted resulting in more nodes remaining awake during the ATIM period and then reduced throughput on account of a greater number of nodes contending with each other. These results were exacerbated when the network load was increased. These results were generated in a simulation that used a trace of an Ethernet as input. The total load was varied between 10 and 60 % of the network capacity. The best results[4] at each load decrease from an average of 70% of the time in the doze state for a 15% load to an average of 42% of the time in the doze state for a 30% load. This is more than a six time increase in the marginal energy consumption by all nodes for a mere

---

[3] We assume that the nodes that serve as p-supporters are not energy constrained and do not need to conserve energy themselves.

[4] Those results obtained with the best selection of an ATIM period and ATIM window.

15% increase in packets exchanged.[5] This study assumes all nodes could hear each other and does not model the effects of hidden nodes.

We are aware of no study of the HIPERLAN energy conserving mechanisms. Such a study would be difficult since the protocol depends on the availability of a node to serve as the p-supporter and the physical distribution of the nodes. Ad hoc networks depend on each other to route and distribute packets to each other. The arrangement of having a surrogate node collect data for another would defeat the objective of many routing protocols. The p-supporter node may not be in a location to collect data from a relaying node opposite in direction to the p-saver. Additionally, the p-saver may be a critical next hop in a route.

The critical deficiency of both the 802.11 and HIPERLAN techniques is that they do not account for the repercussions of a single node's decision to enter the doze state. These repercussions are more congestion as nodes attempt to send traffic to nodes that are dozing and complications for other protocols higher in the stack such as routing. To minimize these adverse effects, access protocols must be able to make dozing more predictable and to integrate the occurrence of dozing with the activities of other protocols, most specifically those of the routing protocol.

2.3 Conserving Energy in Routing

Routing protocols conserve energy by identifying routes based on energy consumption. From the protocol perspective, energy is consumed in transmission and in reception. The energy consumed in transmission can vary based on the range between a source and its destination.[6] The energy consumed in reception is constant.

---

[5] All nodes are awake 25% of the time because the ATIM window is 25% of the ATIM period. The marginal increase in energy consumption is from 5% to 33%.

[6] We assume that all nodes know each other's location and that a source can adjust its transmission power to the minimum required for a successful exchange with a destination.

Due to the power law relation of energy consumed to the distance transmitted a route with more shorter hops may consume less energy than a route with fewer longer hops. We start by presenting the propagation models that explain this benefit in Section 2.3.1 and then review the current research on how to exploit this capability in Section 2.3.2.

2.3.1 Propagation Models

The free space propagation model is used to predict the received power when the transmitter and receiver have a clear path between them. Friis' free space equation, Equation 2.1, shows that the received power is a power law function of the distance separating the transmitter and the receiver.

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \tag{2.1}$$

In this model, $P_r(d)$ is the received power, P_t is the transmitted power, G_t is the gain of the transmitting antenna, G_r is the gain of the receiving antenna, λ is the wavelength, L is a loss factor not related to propagation and d is the distance separating the transmitter and the receiver. In the first model, we assume all terms in the free space equation are constants with the exceptions of P_t and d. P_t is chosen to achieve a specified threshold reception power given d. We rewrite the equation as

$$P_t(d) = Kd^2$$

The free space propagation model has not been a good predictor of received power for land based systems. Modeling radio wave propagation is an active area of research. The motivation for this research has been to define models that would support the design of cellular radio systems. The factors that complicate the modeling are the propagation mechanisms of reflection, diffraction, and scattering. That is the propagation losses are highly dependent on the path between the transmitter and the receiver. Many models attempt to explicitly model these mechanisms but in practice

ENVIRONMENT	Path Loss Exponent, n
Free Space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5
In building line-of-sight	1.6 to 1.8
Obstructed in building	4 to 6
Obstructed in factories	2 to 3

Table 2.2: Path Loss Exponents for Different Environments [16]

the cellular system designs are normally completed in a statistical fashion. Measurements are made on the ground and then statistical models are used to estimate path loss over a coverage area. One model used for this purpose is the log-distance path loss model. This model is essentially the same as the free space model except the exponent of *d* may change. The equation is

$$P_t(d) = Kd^n \tag{2.2}$$

The variable *n* is referred to as the path loss exponent. Typical path loss exponents provided by [16] are found in Table 2.2.

The deficiency of the log-distance model is that it does not consider the fact that received powers can vary vastly between two locations with the same transmitter-receiver separation on account of different environments. Path loss will vary from the average value assumed in Equation 2.2. The log-normal shadowing model builds upon Equation 2.2 attempting to account for this variance. The model used can be written as shown below

$$P_t(d) = 10^{\left(\frac{Z\sigma}{10}\right)} Kd^n \tag{2.3}$$

This model assumes that Equation 2.2 defines a mean power requirement and that the log of the transmission power is normally distributed with a standard deviation of σ dB. Both *n* and σ are determined empirically by making measurements in a given environment. The random variable Z~N(0,1) is chosen by the user of the model to achieve a desired confidence that sufficient transmission power is used. In [17], a study of several German cities found that n had a value of 2.7 and σ had a value of 11.8 dB. This high deviation in required transmission power emphasizes the need for feedback to adjust transmission power levels.

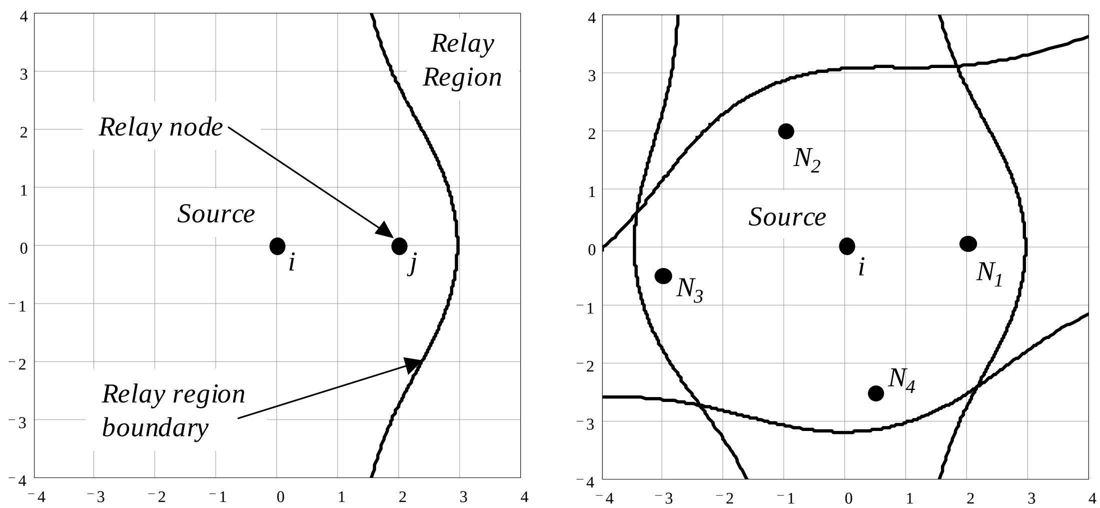
2.3.2 Low Energy Routing

A low energy route uses a series of hops that consume the least energy. A very concise method to select the next hop is described in [18]. This paper demonstrates that all traffic from a source should be forwarded through a subset of the neighbors that surround it. This subset includes all nodes for which a single hop exchange is the most energy efficient method of delivering a packet. It demonstrates that about these nodes a relay boundary can be drawn that defines the relay region to which each of these nodes could be used as an energy conserving intermediate hop. It then shows that the combination of these relay boundaries from these single hop neighbors forms an enclosure of the source. All next hop neighbors for low energy routing are included in this enclosure.

A node j is an energy conserving next hop to node k from node i if the following inequality is true.

$$d_{ik}^{n} > d_{ij}^{n} + d_{jk}^{n} + c \,. \tag{2.4}$$

The variables d_{ik}, d_{ij} and d_{jk} are the distances between nodes i and k, i and j, and j and k respectively, n is the power law exponent, and c accounts for the energy consumed by a node receiving a packet. In Figure 2.2a we illustrate a possible orientation of the nodes i and j and graph the boundary across which node k must be located for the inequality of Equation 2.4 to be true. Then in Figure 2.2b, we illustrate an enclosure formed by 4 nodes that surround a source node i.



a. Relay boundary of node j for source i

b. Enclosure of node i formed by nodes N_1, N_2, N_3, and N_4

Figure 2.2: Relay boundaries for energy conserving routing

2.4 Conserving Energy Through Selective Node Use

Not all nodes are energy constrained. Nodes that receive power from a vehicle, a utility system, or some other external source may not be concerned about the depletion of their energy supply. Additionally, those nodes that are energy constrained, may have different energy reserves available. Energy conserving protocols have two goals, reducing the consumption of energy by energy constrained nodes in the network and prolonging the lifetime of the network as a whole. The first goal can be achieved by giving preference to the use of energy unconstrained nodes in both access and routing protocols. The HIPERLAN energy conservation mechanism of using a p-supporter to accumulate traffic for nodes conserving energy is an example of an access protocol approach. A routing protocol approach is to give preference to routes that use nodes that are not constrained by energy. The second goal is achieved by considering the energy reserves that are available at each energy constrained node and giving preference to the use of those nodes with the largest reserves. Two approaches of this type have been proposed. In the first, multiple routes are used for each destination in an attempt to balance load across the network. The objective is to avoid taxing any single link too much. An example of this approach is seen in the sensor network routing algorithms described in [19] and [20]. These wireless networks consist of multiple low energy nodes all communicating to a single sink node. The proposed routing schemes build minimum spanning trees from each of the 1-hop neighbors of the central sink node. At the end of the tree building process most nodes have multiple paths to the central node. Two metrics are used in building the trees, an additive quality of service (QoS) metric where a higher metric represents a lower QoS, and an energy reserves metric that estimates the number of packets that can be transmitted on that path. The latter metric is constrained by the node with the least energy reserve. Nodes select one of the paths based on the priority of the packet. Lower priority packets get lower QoS despite the availability of a higher QoS route in favor of preserving the life of the network. The assignment of priority to packet types is based on a historical understanding of the types of packets the network has generated and the lifetime of the routes. The routing algorithm attempts to minimize the average weighted QoS metric where the weighting is correlated to the packet priority. Simulations show that this method of balancing the load conserves more energy than using a strict minimum energy metric. [19] This approach, however, is not applicable to ad hoc networks where nodes are mobile and the lifetime of routes would require many tree building events. Also, ad hoc networks do not communicate to a single sink node. Nevertheless, the objective of balancing load is also applicable to the ad hoc network. To insure connectivity, effort must be made to prevent early energy depletion of any node. In one approach to solve this problem, Power-Aware Routing [21], an energy cost metric is used in a minimum metric algorithm. This minimum cost per packet metric is obtained by weighting the energy consumption on the path by the energy reserve on each node of the path. This has a load balancing characteristic that steers traffic away from low energy nodes.

2.5 Energy Consumption: Issues of Overhead

Overhead is often an overlooked issue in conserving energy. Overhead is present in both access and routing protocols and may be especially wasteful in that the overhead is used to coordinate the activities amongst multiple nodes in the network. (i.e. More than just the source and destination listen to and consume energy during overhead transmissions.) Access protocols that require a lot of overhead to coordinate transmissions may be wasteful even though they are effective at putting idle nodes to sleep during the actual data transmissions. Routing protocols that require additional overhead to disseminate the information required to employ an energy conserving routing metric may actually consume more energy than a protocol that uses less overhead with a simple distance metric. Designers of energy conserving routing protocols must be sensitive to the energy consumption consequences of additional overhead.

2.6 The Hierarchy of Energy Conservation Mechanisms

In Sections 2.2 through 2.5 we have identified several mechanisms and issues in conserving energy. In the larger picture, all must be considered simultaneously. More significantly, some of the mechanisms have more potential than others to conserve energy. Protocols should be designed to first insure those mechanisms with the most energy conserving potential are implemented first. In this section we attempt to delineate the hierarchy of mechanisms based on their potential to conserve energy.

The primary characteristic that needs to be considered in prioritizing the mechanisms is scale. A mechanism that affects the energy consumed at just one node per packet exchanged would be lower on the hierarchy than a mechanism that affects all nodes in the network. From this simple observation we come to an intuitive hierarchy. Highest on the hierarchy is the use of energy conserving states. Ideally, all nodes except those exchanging data can be in an energy conserving state. On account of the low quantity of energy consumed in these states and the scale of its effect, using low energy states has the most potential to conserve energy. Second on the hierarchy is the quantity of overhead. Again, overhead affects multiple nodes simultaneously. Reducing overhead increases the time that these multiple nodes can be in a low energy state. Third on the hierarchy is the selective use of nodes. Leveraging the use of nodes that are not energy constrained eliminates the consumption of energy by energy constrained nodes. The scale of this mechanism is relatively small. At most it affects two nodes for each packet exchange (i.e. each hop along a route). Last on the list is energy conserving routing. At best, this technique conserves the energy consumed by just a single node per packet exchange and makes the tradeoff of adding exchanges for the end-to-end delivery of packets. The point is that protocols that conserve energy in routing at the expense of more overhead or requiring all nodes to be awake during the exchanges may not save energy.

2.7 Conclusion

In this chapter we have presented a comprehensive list of protocol mechanisms that are available to conserve energy. We explained how they work and reviewed current research to employ them. Finally, we attempted to prioritize the mechanisms based on the potential they have to conserve energy. Finding the best application of these mechanisms is the goal of this dissertation.

Chapter 3

Conserving Energy in Centrally Controlled Data Transmission

3.1 Introduction

Several papers have been written on energy conserving MAC protocols in networks with base stations. The approach in this chapter differs from most of these in that we isolate the management of packet transmission from both the arrival process of packets and the contention process amongst nodes to have these packets serviced. The justification is simple. Wireless networks that rely on central control generally alternate between contention and contention free periods. At the start of the contention free period the base station has already selected a set of packets that will be transmitted. This selection considers all the issues of achieving quality of service and of balancing the time allocated to the contention and contention free periods. The only remaining issues are how to schedule and how to direct the packet transmissions. Our objective is to answer this question from an energy perspective. We seek to determine how a central node managing packet transmission can reduce energy consumption.

Our goal is not to promote any particular protocol but to identify control mechanisms and then to seek their optimization. Failure to optimize protocols on their own merits can lead to false conclusions. We use as an example the paper [16]. The authors' seek an energy conserving wireless protocol for the transmission of ATM packets and espouse a protocol they have developed called Energy Conserving MAC (EC-MAC). In their efforts to promote the benefits of EC-MAC they compare its performance to that of the IEEE 802.11 protocol but they choose 802.11’s distributed coordination function rather than its point coordination function to manage the traffic. The effect of this choice is each ATM packet needs to contend for access. Contention consumes a lot of energy. If the authors had chosen the point coordination function the protocol could have been implemented to mimic many of the energy conserving features of EC-MAC, e.g. allowing a single contention to establish a connection on which multiple ATM packets can be transmitted. In this chapter we attempt to apply all applicable energy conserving features to all protocols we discuss. We find that the energy consumed in a single protocol can vary as much as 10 times depending on the parameters selected for its operation.

We attempt to address the effect the direction of traffic (i.e. uplink, downlink, peer-to-peer) has on a protocol’s design. A common practice is to imply a protocol’s performance for a general application can be extrapolated based on the analysis of a single traffic direction. For example, [22] discusses the application of energy conserving protocols in very large centrally controlled data networks. The introduction discusses their application to both downlink and uplink traffic. The analysis, however, although an enlightening one and serving as an inspiration for the models developed in this research, only considered the downlink traffic model. Our research shows that the direction of traffic affects the choices made in implementing protocols and in turn their performance.

A third distinction of our work is that we address recovery from errors. Although papers correctly identify the fact that collisions are a major cause of energy consumption in wireless protocols [10], they do not identify that nodes suffer the same penalty for transmission errors. The energy penalty comes from the requirement to re-contend and to retransmit. Protocols can save energy by providing mechanisms that allow nodes to retransmit packets without contention. We show that merely providing a recovery method is not sufficient. The choice of recovery method can greatly influence the energy consumed throughout a network.

Finally, we explore the role of scheduling in energy conservation. The order of exchanges between nodes affects the number of nodes that are awake and so the amount of energy consumed. We attempt to develop scheduling policies to minimize this quantity. Ultimately we show that applying the familiar shortest processing time first scheduling policy provides good results for most all protocols.

This chapter is organized as follows. Section 3.2 provides background on the environment in which we expect our protocols to work. It provides definitions that are used throughout the paper and describes the physical properties of the network that constrain our designs. Sections 3.3, 3.4, and 3.5 develop the details of how our suggested protocol mechanisms work as they support different directions of traffic. We also integrate our discussion of scheduling into these sections. In Section 3.6 we present the results of our simulations. We start by making protocol choices that favor throughput. We then use the results of these simulations to justify alternative operational strategies for each protocol mechanism that allow relaxing throughput in favor of least energy consumption. Finally, Section 3.7 concludes the chapter.

3.2 Background

3.2.1 Centrally Controlled Data Transmission

Centrally controlled data transmission uses a point coordinator (PC) to manage traffic transmission. For all types of traffic, the PC first learns what traffic in the network needs to be transmitted in a contention period (CP) and then directs its transmission in a contention free period (CFP). The PC manages the alternation between these periods in an attempt to balance consideration of delay, throughput and energy consumption. In this chapter we focus on the CFP only. We assume that the PC has learned of *k* transmissions in the CP. The purpose of our investigation is to determine the control, scheduling and error handling schemes that offer the best throughput and energy conservation characteristics for the delivery of these *k* packets.

The PC manages transmissions by broadcasting a directory. The purpose of these directories is two-fold, 1) manage who has access to the channel, and 2) help nodes not involved in the transmissions to doze. The content of the directory may vary. At the very least, it identifies all transmitters that will be active during the subsequent CFP.

At most, the directory will list the source and destination of all transmissions, their size, and the order in which they will take place. This information may either be implied or may be explicitly included in the directory. With this information, nodes can make informed decisions on when they may doze. Figure 1 illustrates the directory structures that we considered.

There are two broad classes of directories, traffic indication maps (TIM) and traffic lists. The TIM is a bitmap. The PC assigns every node a position in the bitmap when the node first associates itself with the network. Depending on the number of bits used in each position of the TIM, a TIM may indicate which nodes will participate in a data exchange, whether they will transmit or receive, or even how many exchanges they will participate in. The TIM does not normally indicate the order

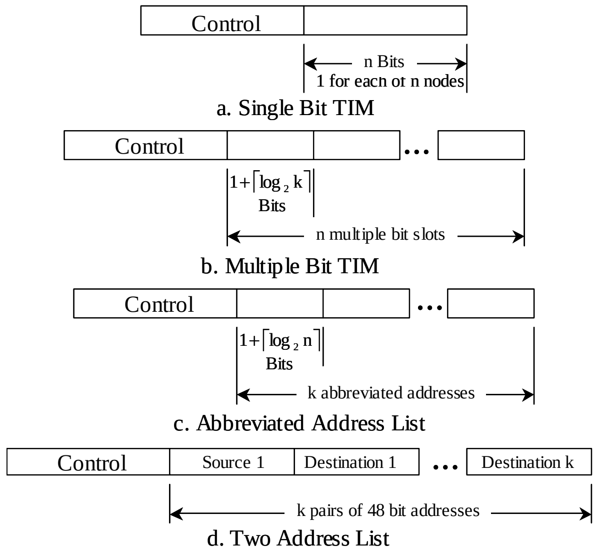


Figure 3.1: Directories for *k* packet exchanges in a network with *n* nodes

of transmissions or with whom the data will be exchanged.[7] This lack of information requires the PC to direct each transmission during the CFP using a poll. Traffic lists, on the other hand, provide all information about the CFP. Each transmission has an entry in the list. These entries specify the source and the destination of each transmission and their duration. After transmitting the traffic list, the PC has no other role during the CFP except to monitor the transmissions.

The tradeoffs between TIMs and traffic lists are the length of the directory itself, the overhead to control transmissions, the amount of information that is available to help nodes to doze, and the ability of the protocol to adapt to transmission errors. In this paper we explore these options and attempt to use all methods to improve performance. Before proceeding we describe the different traffic models considered in our analysis, describe key components of the physical layer, identify the tradeoffs found in the two error recovery options, and define our measures for performance.

[7] In this paper we provide an exception when we use an implied transmission schedule with the multiple bit TIM protocol. See Section 3.3.2.

3.2.2 Traffic Models

The effectiveness of an energy saving protocol depends on the predominant direction of traffic in the network. We consider four basic types of wireless traffic patterns and refer to them as Type I, II, III, and IV traffic. Type I traffic corresponds to nodes attempting peer-to-peer communication. Such networks are often referred to as ad hoc networks. Energy conserving MAC protocols supporting this traffic must synchronize each source's transmission with a period that the destination is awake. In Type II traffic, data is only transmitted to and from a base station. The base station may provide access to a larger wired network, it may be the central server for a group of mobile clients, or it may be the intermediary for intracellular traffic. There is no synchronization requirement between source and destination as the base station is always awake. In Type III traffic, the communications consist of downlink packet transmissions only as in a paging system. These are the simplest protocols to analyze as there is no requirement to manage the individual access of the mobile nodes, only the transmission of downlink messages and the scheduling of dozing periods. By contrast, Type IV traffic consists of a central node collecting data from mobile nodes such as in a telemetry application. In this chapter, we separately develop and compare protocols for each of these traffic patterns. Table 3.1 lists the various TIM and traffic list options that may be used for the different types of traffic.

3.2.3 Physical Layer

Our protocols are designed for a *single* channel network, i.e. there is no separate control channel such as that found in cellular telephone systems.[8] As a result, each transmission must be acknowledged on the same channel in order for the

[8] A two channel system would need to account for energy consumption on both channels.

DIRECTORY	APPLICATION	UNKNOWNS	DOZING BY ACTIVE NODES
TIM (1 bit per node) A 1 in a nodes bit position in the map indicates that node will participate in the data exchanges of the upcoming TIM period.	Type I Traffic	Transmission order, number of packets.	Nodes doze after receiving a negative TIM.
	Type II, III, and IV Traffic	Transmission order, number of packets per node.	Nodes identified in the TIM doze after they participate in a data exchange and then another node is polled.
TIM (Multiple bits identifying number of packets)	Type II, III, and IV Traffic (Exchanges in bitmap or other implicit order)	None	Nodes doze before and after data exchanges.
List - Single address for each data exchange. (Addresses may be abbreviated)	Type II Traffic (Uplink and downlink transmissions are not mixed)		
	Type III and IV Traffic		
List – Two addresses, source and destination, for each data exchange	Type I and II Traffic		

Table 3.1: Comparison of Alternative TIM and List Directories

transmitting node or the PC to learn of its reception. On account of access rules, detection of the acknowledgement is sufficient to determine reception. Reception of the acknowledgement may be necessary for other protocol decisions and will be described later.

Wireless networks have three timing considerations. First, time must be allowed for nodes to transition between transceiver states. Second, on account of propagation and transition delays, silent periods necessarily occur between transmissions. Third, the network uses silent periods as an indication that the channel is idle. To simplify our analysis we use a single time period, a slot time, to account for both the time it takes to transition among transceiver states and the time between transmissions. We use the slot time as the basic time unit in our analysis. We let S denote a slot time which for this work will be the time it takes to transmit 48 bits. Some protocols require more than a single slot silent period between transmissions to support prioritizing transmissions and error recovery. In those cases multiples of the slot time are used by the protocol. [6]

In order to receive a transmission, a receiver must be synchronized with the transmitter. To support synchronization and equalization the physical layer transmits a series of bits prior to any transmission of data. We assume that a receiver will not receive a transmission unless it is awake and receives these overhead bits first. In other words, a receiver cannot wake up in the middle of a transmission and then receive data. We label this overhead period as OH and define it as the time to transmit 192 bits[9], 4 slot times. A protocol can reduce such overheads by transmitting several protocol data units (PDU) contiguously after a single synchronization/equalization overhead transmission, but in order for a receiver to receive any one of the PDUs it must have been awake from the time this physical overhead was transmitted.

Our error model uses independent bit errors with a constant error rate. The objective is to correlate packet error probability with packet size. We chose a packet size of 103 slots, i.e. 618 bytes in our analysis. This size is just below the threshold where a bit error rate of 10^{-5} makes it beneficial for minimum delay to split the packet and send it in two transmissions.[10] The second error rate used in our analysis, 10^{-4}, yields significantly more packet errors. The protocols are thus compared in low and in high packet error environments.

Figure 3.2 illustrates the structures of polls, packets, and acknowledgments. As can be seen the physical layer overhead bits precede each transmission. The

---

[9] The overhead of the IEEE 802.11 physical layer is 196 bits. The overhead of the HIPERLAN physical layer prior to transmitting in the high bit rate mode is 450 bits.

[10] Large packets can be split in two and even with additional overhead be transmitted in less time, on average, than as a single packet due to reduced retransmission overheads.

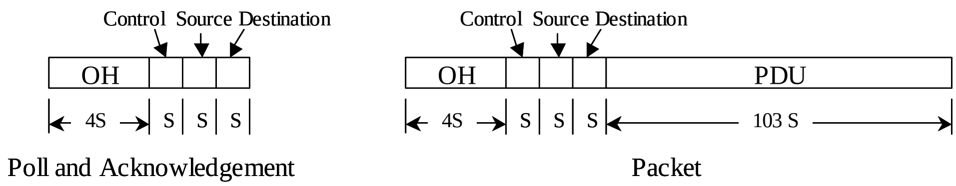


Figure 3.2: Poll, Acknowledgement and Packet Structures

content of the polls and acknowledgements is the same as the packet overhead, consisting of one slot for each address and one slot for any control information.

3.2.4 Error Recovery

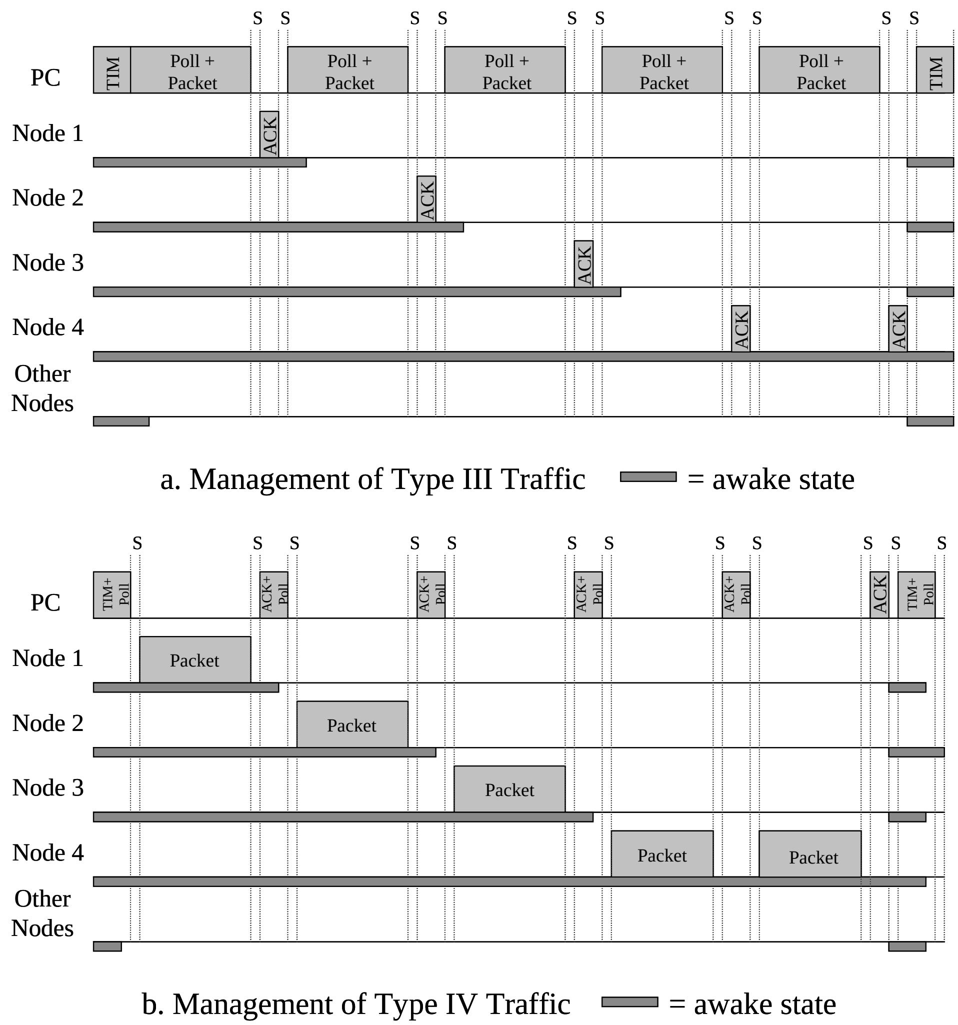
There are three alternative error recovery policies. In the first , which we call immediate retransmission, nodes attempt to retransmit a failed transmission immediately. Protocols provide a mechanism for the transmitting node to retain control of the medium after its transmission fails. In the second recovery method, which we call delayed retransmission, packet failures are rescheduled by the PC. This requires the PC to be omniscient of the status of each transmission. In the last recovery method, which we call recontention, nodes that are unsuccessful in sending their packets contend again starting in the next contention period. Recontending contributes to congestion and wastes energy so it is not considered any further in our research.

Immediate and delayed retransmission support energy conservation since both allow retransmission without contention. Immediate retransmission has an advantage over delayed retransmission in terms of throughput since the protocol does not allow silent time. However, if errors are spatially correlated, this advantage may switch. Delaying retransmission may allow spatial conditions to change thus improving the chances of the transmission succeeding. In the same circumstances immediate retransmission could tie up the network trying to make a failed link work. Nevertheless, our independent bit error model does not make this a factor in our analysis so the immediate retransmission option is the higher throughput option.

The relative difference in the energy consumed by the networks using these error recovery options depends on the protocol and the type of traffic. One observation that is made later on is that the delayed retransmission protocol has a timing advantage that makes it easier to keep idle nodes in the doze state. In most cases, delayed retransmission is more energy efficient.

3.2.5 Comparison Measures

The objective of our analysis of different protocols is to compare their performance. We assume that at the beginning of a period of centrally controlled data transmissions that the protocols that are being compared have identified the same k packets to send. Guarantees of quality of service and access fairness are managed at a higher level. The measures of performance for our comparison are total service time and network energy consumption. Service time is the total time required to deliver the k packets. Energy consumption is the total time nodes in the network are awake until the k packets are delivered. A third measure of performance is the energy consumed by a node per packet exchanged during the CFP. In this case nodes are categorized based on the number of packets they exchange during the CFP, so protocols can be evaluated in terms of their "fairness" and one can identify where energy is wasted, e.g. on idle nodes, on nodes that send only one packet in the CFP, etc.



a. Management of Type III Traffic = awake state

b. Management of Type IV Traffic = awake state

Figure 3.3: Type III and IV traffic management using single bit TIMs

3.3 Protocol and Scheduling Options for Type III and IV Traffic

We begin by considering traffic Types III and IV, that is uplink and downlink traffic to and from a PC. We consider how three protocols would service this traffic: single bit TIMs, multiple bit TIMs with the packet count, and single address lists using abbreviated addresses. We describe how the protocols implement immediate and delayed retransmission error recovery as well as how the scheduling of packet transmissions can achieve energy efficiency.

3.3.1 Single Bit TIMs

Description

The single bit TIM method of servicing Type III and IV traffic requires the PC to manage the traffic. Since the PC participates in every exchange of data, it is possible to gain some efficiency by piggybacking polls with transmitted packets and acknowledgements. This is illustrated in Figure 3.3. In this figure we assume that the PC directs transmissions to and from the same node in contiguous slots. On account of this assumption, nodes are awake from the transmission of the TIM until after they participate in an exchange and the PC polls a subsequent node. This figure illustrates when the nodes are awake.

At the cost of a small amount of delay some energy efficiency may be gained by dividing the k transmissions into multiple TIM periods each announced with a separate TIM. This division allows a greater number of nodes to doze while waiting to exchange data. The tradeoff is that all nodes must awaken to receive the additional TIMs.

Equations are derived in Appendix A for the transmission time and for the expected energy consumption for a network with uniformly distributed traffic and no transmission errors.

Error Recovery

The PC manages the recovery from a failed transmission. If the PC does not sense the transmission of a packet after a poll or the transmission of an ACK after a packet, it identifies an error condition. The PC then seizes control of the channel and directs the next data exchange according to the error recovery procedure. Since all nodes with pending traffic in the TIM period are awake, the PC can send the poll immediately for either recovery option. If the delayed retransmission option is used the packet is not retransmitted until a subsequent TIM period. In both cases, however, data exchanges are not allowed to interfere with the transmission of the subsequent TIM. The penalty of having all nodes in the network awake waiting for the delayed transmission of a TIM is considered too great. Even when the immediate retransmission recovery method is used, failed packets at the end of the TIM period whose retransmission would interfere with a scheduled TIM are deferred until after the transmission of the next TIM. Packets retransmitted after a subsequent TIM are rescheduled according to the scheduling policy in use, e.g. they are not given priority.

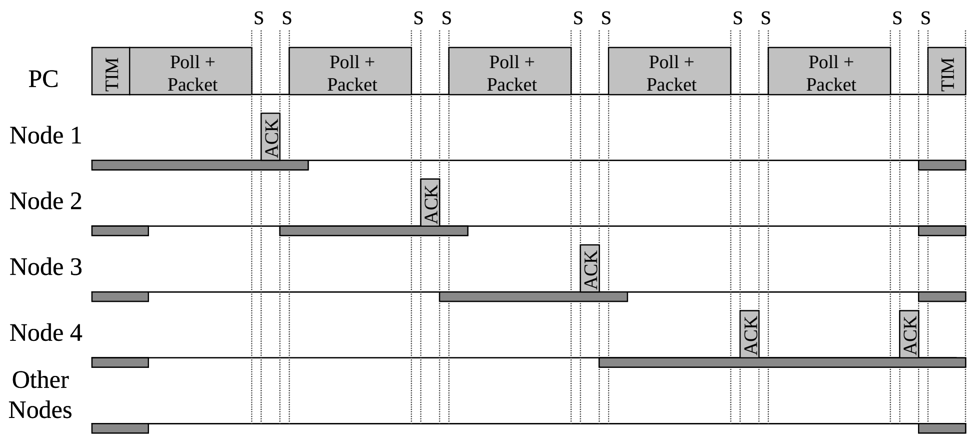
Scheduling

Since all nodes that participate in a TIM period are awake from the broadcast of the TIM until the next node is polled some efficiency can be gained by scheduling the order of packet transmissions. The scheduling problem of minimizing the average time spent awake is identical to that of minimizing the average delay of jobs that need to be serviced by a common resource. It is known that the optimal schedule in this case is to serve the shortest job first. [23] In this context, job size corresponds to the number of packets to or from a given node for Type III or IV traffic respectively.

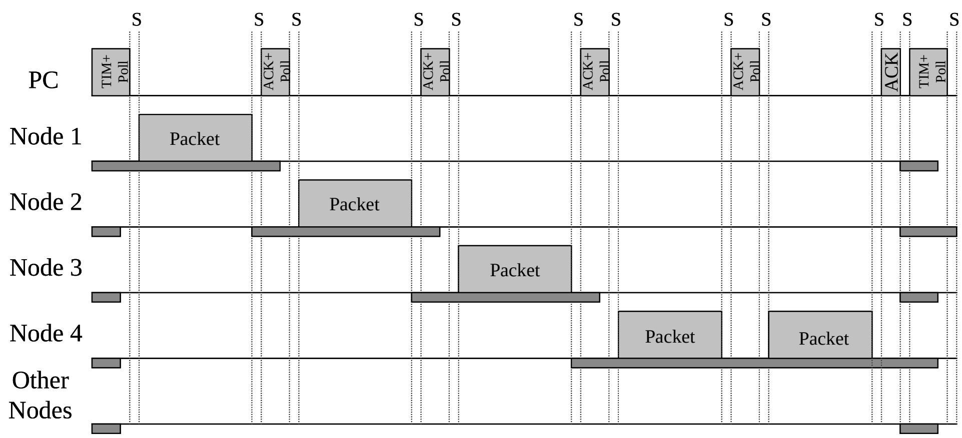
3.3.2 Multiple Bit TIMs

Description

Multiple bit TIMs consist of multiple bits in each node's map position. The multiple bits used in each node position of the TIM identify the number of packet



a. Management of Type III traffic = awake state



b. Management of Type IV traffic = awake state

Figure 3.4: Type III and IV traffic management using multiple bit TIMs

exchanges that will occur with that node. The PC still manages transmissions in the same manner as when using single bit TIMs. It polls each node prior to each exchange. The difference is that at the cost of having a longer TIM, nodes may doze during the time preceding their exchanges. This is possible only if there is an implicit order in which nodes identified in the TIM may transmit data. We assume the PC transmits the packets first in quantity, fewest packets first, and then in bitmap order. Figure 3.4 corresponds to the transmission of the same packets as in Figure 3.3 but

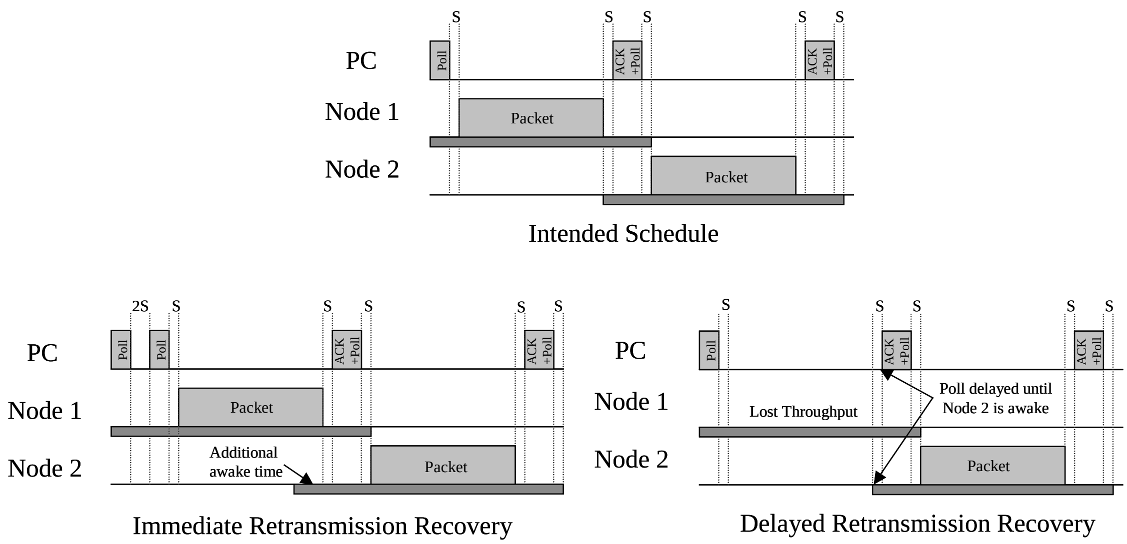


Figure 3.5: Multiple bit TIM protocol recovery from a failed poll for Type IV traffic

demonstrates that the multiple bit TIM protocol can reduce the number of nodes which are awake.

Equations are derived in Appendix 3.A for the transmission time and for expected energy consumption for a network with uniformly distributed traffic and no errors.

Error Recovery

The PC manages error recovery. The PC identifies errors if it does not sense the transmission of a packet after a poll or the transmission of an ACK after a packet. The PC may either reattempt transmission immediately or delay transmission until a subsequent TIM period. When the protocol uses immediate retransmissions all subsequent nodes scheduled to transmit or receive packets in the current TIM period will wake-up and have to wait for their exchange. We assume the control segment of polls include the packet number in the schedule and the target time for the transmission of the next TIM. Therefore, nodes that wake-up after a retransmission and that receive a poll can estimate when they would be rescheduled and can doze until that time.

When the protocol delays retransmission until the next TIM period, the PC must transmit the next poll according to the transmission schedule so that it is transmitted when the intended recipients are awake. If there is no response to a poll in Type IV traffic, the PC must still wait the duration of the planned packet transmission before it can poll the next node. Figure 3.5 illustrates the tradeoff between the two recovery options after a failed poll. Immediate retransmission results in additional power consumption by other nodes waking prematurely for their delayed exchanges. Delayed retransmission results in lost throughput and additional energy consumption by the node whose exchange failed.

Scheduling

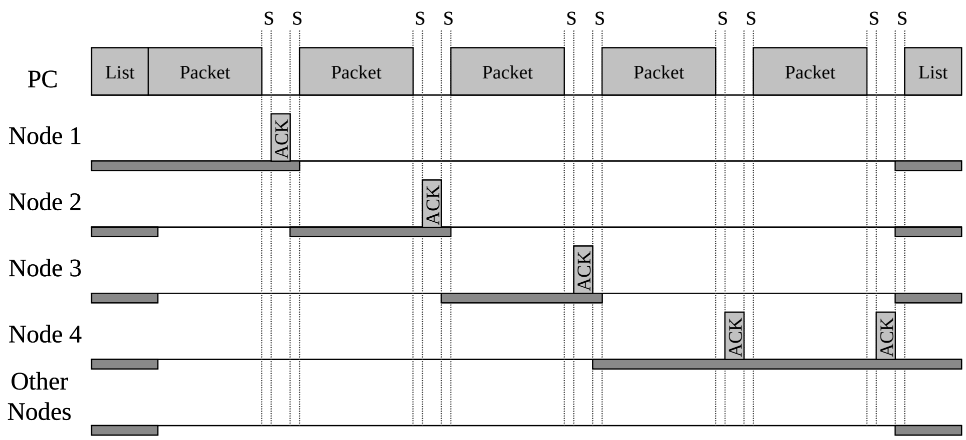
Immediate retransmission penalizes nodes scheduled to transmit at the end of a TIM period. These nodes will wake-up and have to wait for the delays caused by the retransmissions that occur earlier. Scheduling can reduce energy consumption by minimizing the number of nodes that wake-up at the end of a TIM period. Scheduling nodes with fewest packets to send first achieves this ordering. This transmission schedule corresponds to the implied schedule described above, again a shortest processing time first schedule.

Delayed retransmission does not penalize any of the nodes in the TIM period. Nodes are only awake when they transmit or receive data. Scheduling of packets offers no advantages. We continue to use the implied schedule in our analysis.

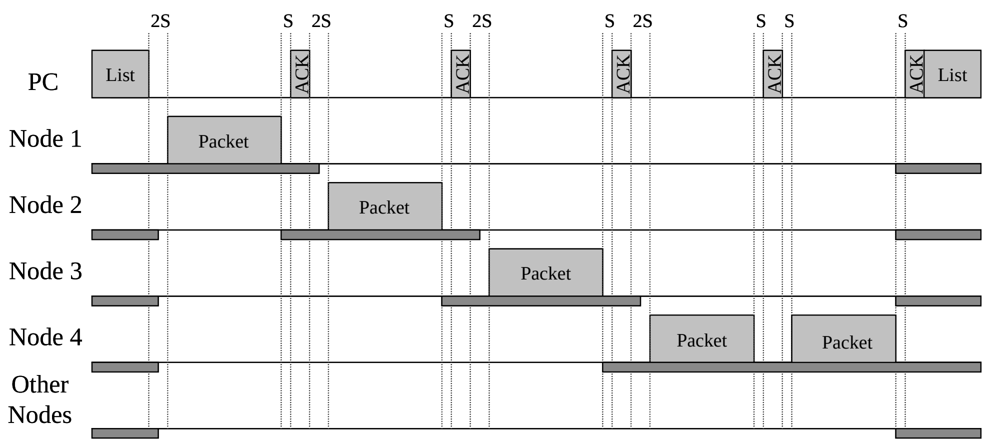
3.3.3 Single Address List

Description

The operation of the single address list protocol is very similar to that of the multiple bit TIM protocol. The differences are that the list is longer than the TIM and



a. Management of Type III Traffic = awake state



b. Management of Type IV Traffic = awake state

Figure 3.6: Type III and IV Traffic Management Using Single Address Lists

that the transmission of data is not preceded by a poll. Figure 3.6 illustrates the operation of the protocol. Note that there is a difference in the awake time per packet between Type III and IV traffic. Since with Type III traffic the data transmission originates at the PC, there is a greater level of control resulting in less awake time. Indeed, mobile nodes need only wake up prior to the scheduled data transmission. With Type IV traffic, the mobile nodes are the sources. To be sure they transmit at the appropriate time, they must wake-up early enough to monitor the preceding ACK to verify that it is their turn to transmit. We assume each packet transmission and ACK includes the packet number thus allowing subsequent nodes to determine the progress that has been made on the current transmission schedule. A node knows it can transmit when it monitors an ACK with the appropriate packet number.

Equations are derived in Appendix A for the transmission time and for expected energy consumption for a network with uniformly distributed traffic and no errors.

Error Recovery

This protocol can support both recovery options. With immediate retransmission, the node that is the source retransmits once it fails to sense an ACK. Similarly, if a node transmits an ACK and a packet transmission is not sensed thereafter that node will retransmit the ACK. As with the multiple bit TIM protocol this can cause subsequent nodes to wake too early since they are not aware of the retransmitted packets. Again we include the packet number in the control segments of all packet and ACK transmissions. Once awake nodes learn the packet number of the current exchange they can estimate a new wake-up time and return to the doze state. With delayed retransmission the PC is the critical player. With Type III traffic, the PC simply transmits packets at their scheduled times rescheduling those packets that are not acknowledged. With Type IV traffic, the PC transmits two types of ACKs, the standard ACK after it receives data and a negative ACK when it does not. Sending the ACK is critical since reception of the ACK is used by the mobile nodes as the signal when they may transmit. The PC reschedules packets that it does not receive.

Scheduling

This protocol schedules traffic such that the transmissions of each node are in contiguous slots and nodes with the fewest transmissions go first. Keeping transmissions to common nodes in contiguous slots is an energy saving arrangement that reduces the number of transitions to/from the doze state. Additionally, in the case of Type IV traffic, mobile nodes only need to wake-up to listen to the ACK of another node's transmission once. Immediate retransmission has the same effect as in the multiple bit TIM protocol of penalizing nodes scheduled to transmit late in the CFP. So again, scheduling the transmissions of nodes with the fewest exchanges first conserves energy. Similarly, as with the multiple bit TIM protocol, scheduling does not affect energy consumption when delayed retransmission is used.

3.4. Type II Traffic

The performance of protocols supporting Type II traffic is identical to that of Types III and IV traffic when the uplink and downlink are executed in separate cycles. When the uplink and downlink traffic is integrated the analysis changes by making the packet service time the average of an uplink and a downlink packet service time as each is assumed equally likely to occur. This analysis is equivalent to averaging the time to service k packets of Type III traffic and k packets of Type IV traffic. When lists are used, one additional modification is required. An additional bit would be added to each address to identify whether the mobile node is to send or receive data. We provide no analysis of protocols supporting Type II traffic assuming all performance information can be extrapolated from our analysis of protocols supporting Types III and IV traffic.

3.5. Protocol and Scheduling Options for Type I Traffic

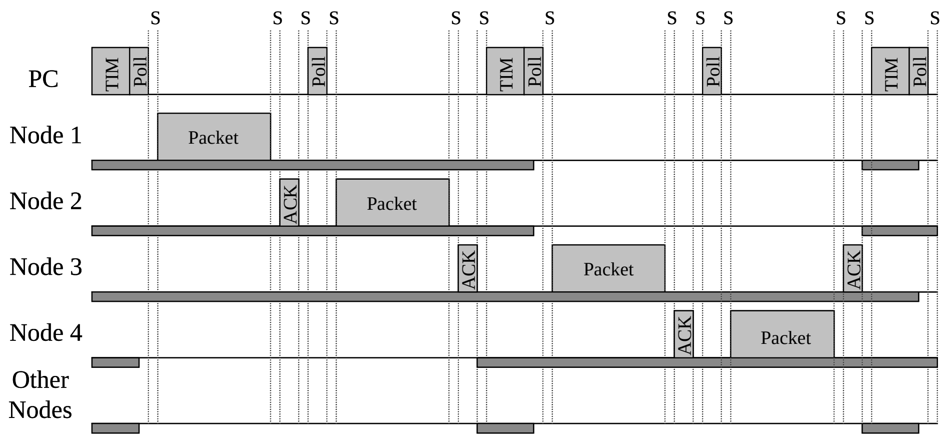
Two characteristics distinguish managing Type I traffic from managing Types II, III, and IV traffic. The first is data exchanges occur between pairs of mobile nodes. Scheduling approaches must consider the fact that to exhaust one node's exchanges multiple other nodes must be awakened. A node can no longer assume that it is finished participating in data exchanges when the PC stops polling it. The second is that the PC no longer participates in the data exchange so it cannot be certain whether traffic was successfully exchanged between two nodes. Retransmissions on account of errors must be initiated by the source. If nodes are not provided a means to immediately correct a failed transmission, then the source will be required to contend again to send the traffic.

We also assume that all nodes in the network are within range of each other. If the PC can determine that this is not the case it can redirect transmissions perhaps relaying the traffic between the nodes. These corrective actions are not discussed in this chapter but are essential, especially when lists are being used.

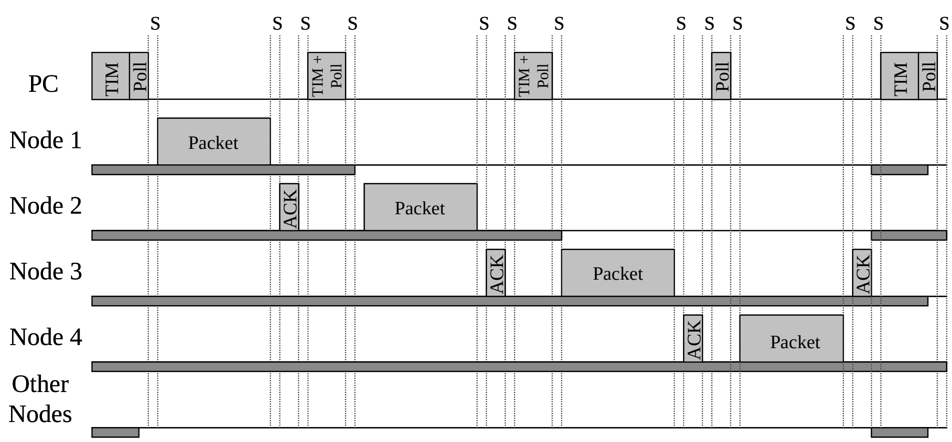
3.5.1 Single Bit TIMs

Description

Since with Type I traffic it is no longer possible to consolidate all communications of each node into contiguous slots, nodes can no longer assume that they can doze based on whom the PC polls. Nodes must be explicitly told that they may doze. So the single bit TIM method of servicing Type I traffic requires all nodes identified in the TIM to remain awake until a subsequent TIM puts them into the doze state. Thus, it is impractical to use just one TIM per transmission cycle. Energy can be conserved in one of two manners. In the first, the CFP is divided into smaller TIM periods so that only a subset of the nodes that are participating in the CFP are awake each TIM period. In the second, in addition to using multiple TIM periods, additional TIMs which we will refer to as "doze TIMs" are used within the TIM period for the express purpose of putting nodes into the doze state. Figure 3.7 illustrates the difference between these two methods of control. We made no attempt to model the performance of this protocol with this type of traffic on account of the expected complexity.



a. All nodes awaken for multiple TIMs = awake state



b. Doze TIMs put nodes into doze state = awake state

Figure 3.7: Alternative methods of single bit TIM control of Type I data transmission

Error Recovery

Portions of error recovery may be managed either by the PC or by each node transmitting a packet. The PC is the only entity that can manage recovery when a poll is not received. When a packet is not received, either the PC or the node transmitting the packet may manage the recovery. Both learn of the failure when an ACK is not received. If the PC manages the error recovery, it does so by polling the source a second time. If the transmitting node manages the error recovery then it attempts to retransmit the packet immediately. The advantage of letting the transmitting node manage retransmission is that the transmission of the poll is avoided. The disadvantage of node managed recovery is that it does not support delayed retransmission.

Scheduling

Schedules can be based on two intuitive observations. First, the optimum schedule when nodes remain awake throughout TIM periods is a schedule that minimizes the average number of nodes that are awake each TIM period. Second, the optimum schedule when doze TIMs are used is one that allows the doze TIMs to put the most nodes to sleep soonest. Schedules based on these observations are optimal only when there are no errors. Errors reduce the significance of achieving these goals.

The problem of minimizing the average number of nodes awake during each TIM period is a function not only of the number of packets scheduled per CFP but also the size of each TIM period. An exhaustive solution requires attempting all combinations of packet transmissions across the different TIM periods: there are $\frac{k!}{j!\left\lceil k/j \right\rceil!}$ such combinations where k is the number of packets per CFP and j is the number of TIM periods. As an alternative to an exhaustive solution we provide two heuristic algorithms that seek a minimum quantity of active nodes before each TIM period. The first, Algorithm A, selects active nodes for the TIM period in a greedy fashion. It starts by selecting the two nodes with the most exchanges and then adds nodes to the active set based on how many exchanges their addition provides giving preference to the node that adds the most unless the TIM period can be filled with fewer. The second, Algorithm B, seeks the minimum number of nodes that have just enough exchanges to fill

NEXT NODE TO SLEEP SOONEST			MOST NODES TO SLEEP SOONEST		
SOURCE	**DESTINATION**	**NODES AWAKE**	**SOURCE**	**DESTINATION**	**NODES AWAKE**
1	2	5	3	4	5
2	1	5	3	4	5
1	5	4	3	4	5
1	5	4	1	2	3
1	5	4	2	1	3
3	4	2	1	5	2
3	4	2	1	5	2
3	4	2	1	5	2
	TOTAL AWAKE:	**28**		**TOTAL AWAKE:**	**27**

Figure 3.8: Comparison of next node to sleep soonest to most nodes to sleep soonest scheduling

the TIM period. It does this search by trying all combinations of m of the n nodes participating in the CFP incrementing m until enough exchanges can be found to fill the TIM period. Algorithm B is computationally complex but was still considered in our analysis in order to determine its energy conservation potential. Details of Algorithms A and B are found in Appendix B.

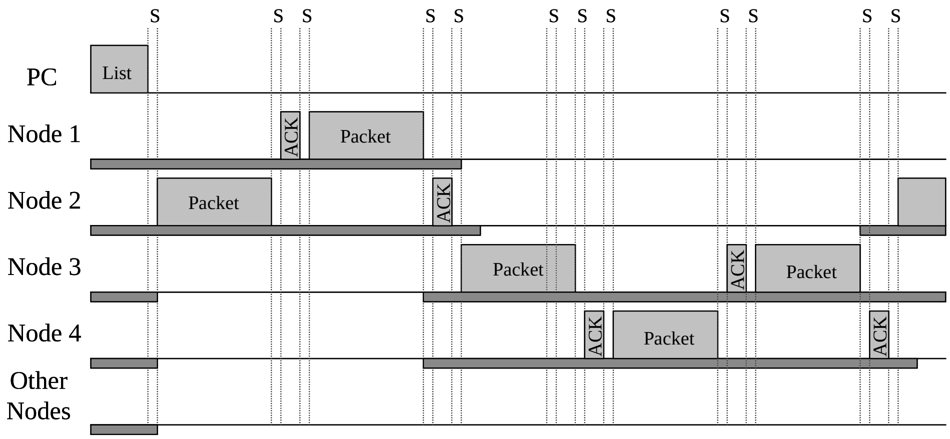
Since multiple nodes may have to awaken in order to exhaust any single nodes transmissions scheduling with the aim of putting the most nodes to sleep soonest is complex. We used a schedule that puts the next node to sleep soonest as an alternative. Although not optimum it provides good results. The difference between most nodes to sleep soonest and next node to sleep soonest is illustrated in Figure 3.8. In this figure, 8 exchanges need to be made amongst 5 nodes. Each exchange has a source node and a destination node. The 5 nodes cannot doze until all their exchanges for the TIM period have been completed. In the next node to sleep soonest schedule, the exchanges of node 2 are scheduled first since node 2 can enter the doze state soonest. In the most nodes to sleep soonest schedule, the exchanges between nodes 3 and 4 are scheduled first since both nodes can enter the doze state at the conclusion and since the final tally of node-awake times is reduced from 28 to 27. Although not optimum, putting the next node to sleep the soonest is a very simple algorithm. Our Algorithm C (see Appendix B) seeks this schedule. It looks at nodes individually and schedules the transmissions of the node with the fewest exchanges first.

In an attempt to come closer to the most nodes to sleep soonest schedule, we attempted to combine Algorithms A and B with C. The objective was to allow Algorithm A or B to select a reduced number of nodes to participate in a TIM period and then to use Algorithm C to schedule these exchanges to put the next node to sleep soonest. Reducing the number of nodes in a TIM period increases the likelihood that a next node to sleep soonest sort approaches a most nodes to sleep soonest schedule for that TIM period. We call these two hybrid Algorithms E and F.

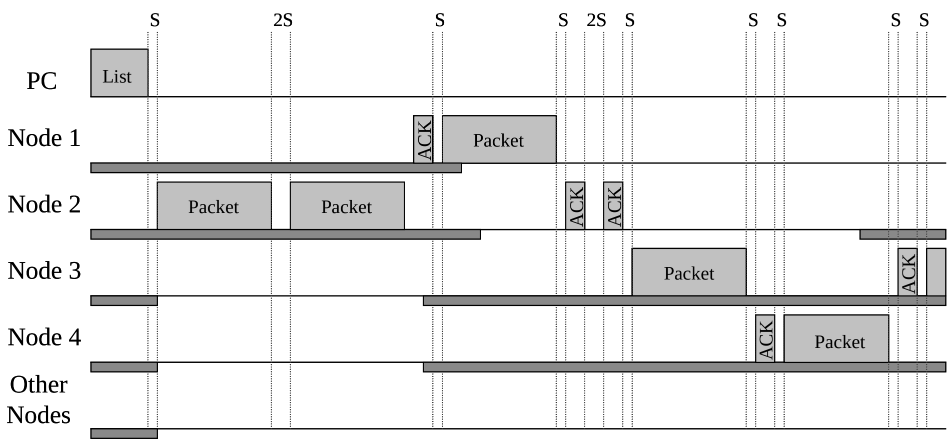
3.5.2 Two Address List

Description

The directory using a two address list consists of two addresses for each packet transmission. Abbreviated addresses are not used on account of the overhead and complexity of disseminating those addresses to all nodes. The addresses used are standard 48 bit MAC addresses that we assume are used for wireless nodes. Figure 3.9a illustrates the transmission of data using these lists. Pairs of nodes awaken for each transmission. To simplify identifying when each node has its turn to transmit and to minimize the time nodes spend transitioning, all exchanges between common pairs of nodes are executed in contiguous slots. Each packet transmission and ACK includes the packet exchange number for the CFP. A pair of nodes knows it is its turn to transmit when its packet exchange number immediately follows the number announced in the last ACK. For this reason all nodes except the first to transmit in the CFP will monitor an ACK before they transmit a packet.



a. Traffic Management Without Errors = awake state



b. Traffic Management With Errors = awake state

Figure 3.9: Type I traffic management using a two address list

Error Recovery

On account of the requirement for a node to monitor an ACK before transmitting, only immediate retransmission error recovery is attempted. If a destination does not receive a packet correctly, it will not send an ACK. If the source does not react by retransmitting the packet then the chain of transmissions identified in the list will be interrupted. It may be possible for the PC to detect the error and for either the PC

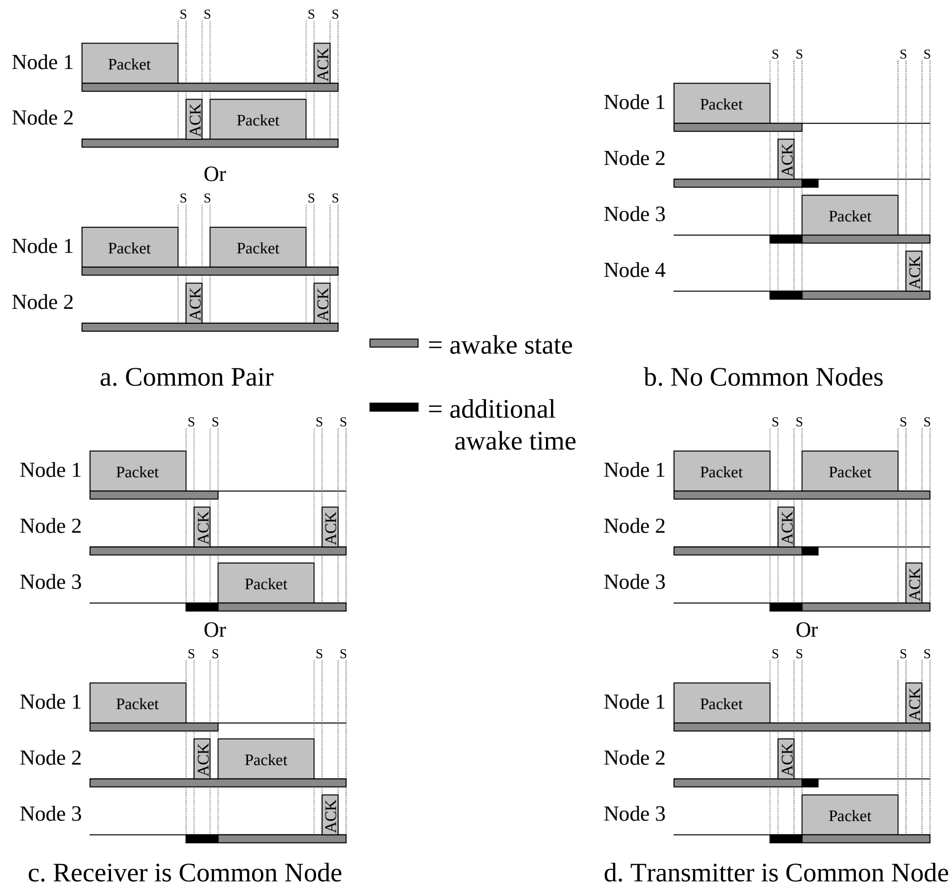


Figure 3.10: Energy consumption at pair transitions for Type I traffic using the list protocol

or the transmitting node to send a pseudo ACK to prompt progression in the transmission list but the timing would still be compromised. Similarly, if the next node to transmit does not respond to an ACK then the sending node must continue to resend the ACK until it does. This activity also compromises timing. Therefore, we chose to only support immediate recovery. Figure 3.9b illustrates the transmission of data when errors occur. If the timing gets too bad and a node wakes up early and monitors either an ACK or the beginning of a packet transmission, the packet exchange number in these transmissions allows the node to return to the doze state since it can estimate when its needs to awaken again.

Scheduling

Scheduling can improve energy performance in one of two ways. In the first, it can reduce the extra energy consumed at the transitions between packet transmissions. Figure 3.10 illustrates the possible transitions emphasizing the different amounts of energy consumed. As seen in this picture the best energy consumption occurs when the same nodes participate in the two packet exchanges followed by the transitions where at least one node participates in both exchanges. Algorithm D is a heuristic approach that attempts to create a schedule that optimizes according to this observation. Appendix B provides the details of Algorithm D. In the second method of scheduling for energy conservation the primary objective is to minimize the effects of failures. The schedule that minimizes the number of nodes that wake up prematurely after errors will best achieve this goal. This is identical to the objective of putting the most nodes to sleep soonest. So again we use Algorithm C to generate an energy-conserving schedule. Algorithm C also achieves some of the preferred transitions since it groups transmissions between common nodes together.

3.6. Model and Simulation Results

We developed a simulation that modeled the protocols and scheduling algorithms described above. Each node in the simulations was modeled using an independent random number generator for error states. The study methodology considered the performance of the protocols and algorithms for different size networks, different quantities of exchanges, different error conditions, and different TIM periods (when TIM protocols were used). We started our study using immediate retransmission as the error recovery option. Traffic was generated assuming a uniform distribution amongst the nodes.[11] The actual exchanges between pairs of nodes used in the simulations were the same so the effects of the protocols, scheduling algorithms, and error recovery methods could be isolated from the effects of the random clustering of transmissions among nodes. Total service time and total energy consumption (i.e. total time nodes are awake) was obtained for each simulation. Fairness statistics were obtained for each protocol with the best performing set of parameters (i.e. best TIM period measured in packets per TIM period (PPT)). Samples of the results are shown in Figures 11, 12, and 13. Panels a, b, and c exhibit total energy consumption versus total transmission time. The different points plotted for the TIM protocols correspond to different TIM periods. Panels d, e, and f show the energy consumed per packet transmitted parameterized by the total number of packets sent by each node in the CFP. Note the average energy consumed by nodes that sent no packets is identified by 0, the average energy per packet consumed by nodes sending just one packet is identified by 1, and the average energy per packet consumed by nodes sending x packets each is identified by the number x on the abscissa.

In the case of Types III and IV data we validated our simulation model by comparing the simulation results to the model predictions include in Appendix A. As illustrated in Figures 11a and 12a, the results match.

We simulated five different sized networks, 5, 10, 25, 50, and 100 nodes, under three different bit error rates, no errors, 10^{-5}, and 10^{-4} errors per bit, and for three traffic types, I, III, and IV. We ran 200 simulations for each set of conditions.

[11] Uniform distribution of traffic is worst case for energy conservation. Any clustering of exchanges between common nodes allows more nodes to doze for longer periods of time and is thus advantageous.

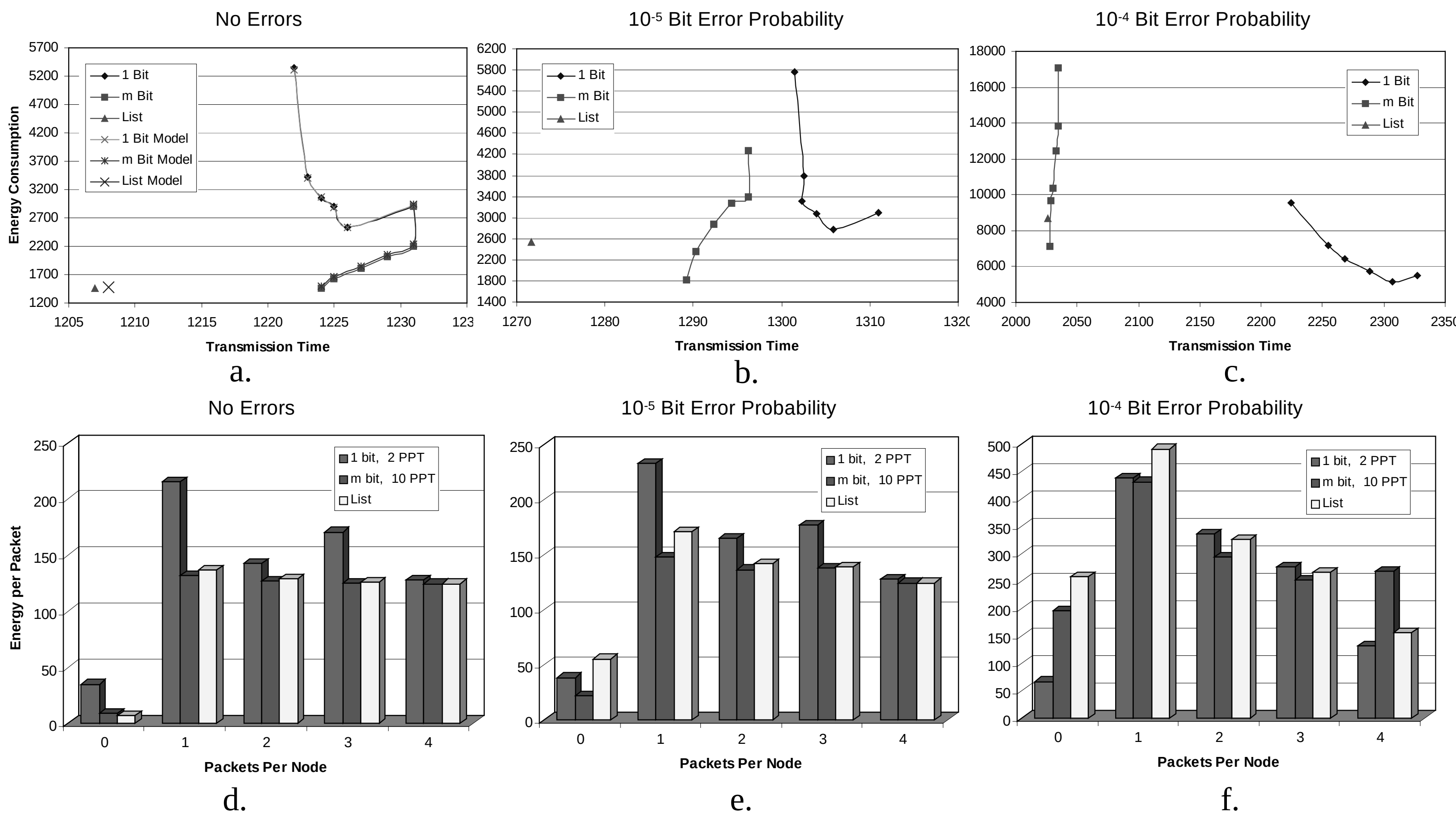


Figure 3.11: Performance of Type III traffic, 25 nodes, 10 packets

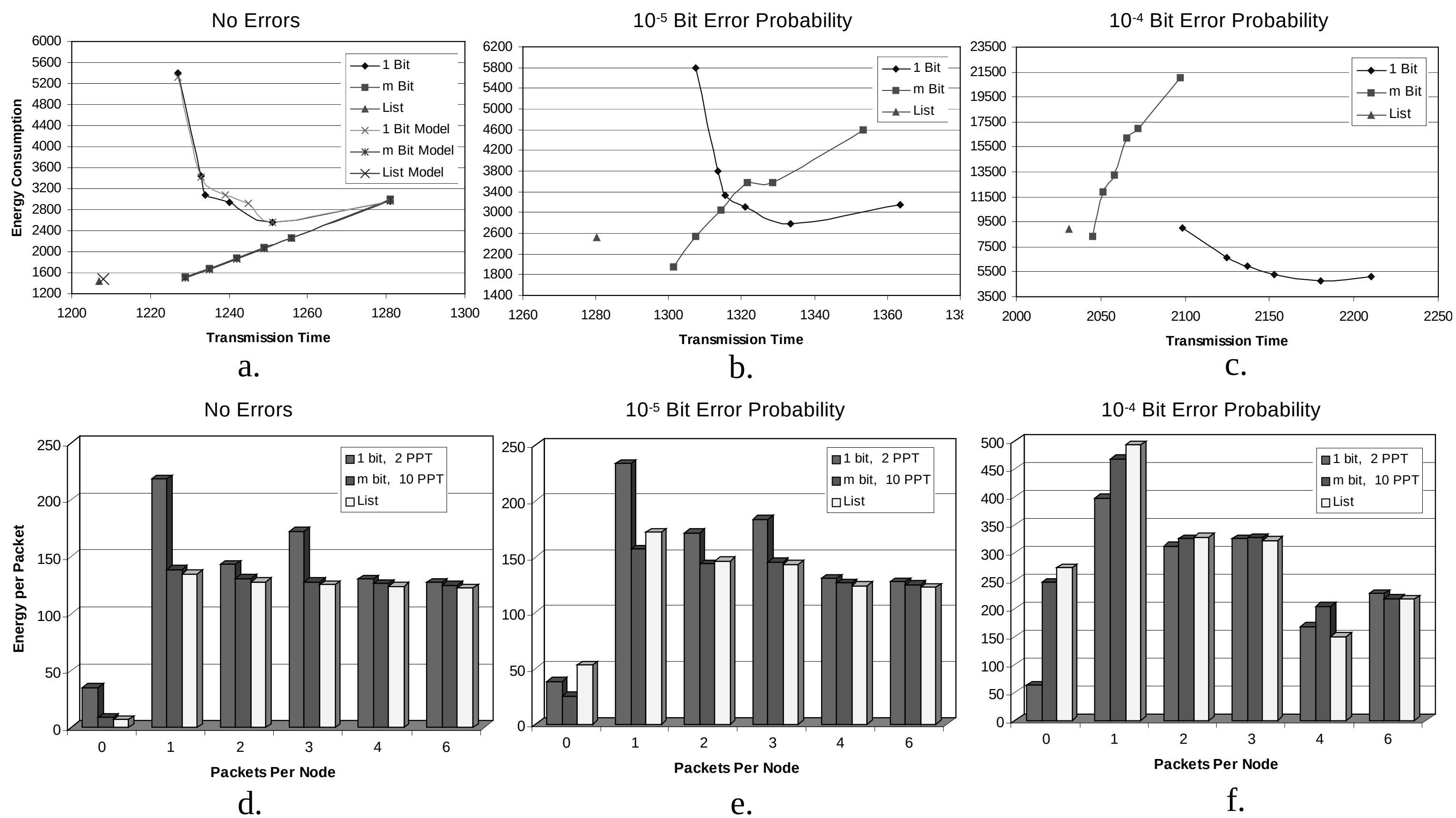


Figure 3.12: Performance of Type IV traffic, 25 nodes, 10 packets

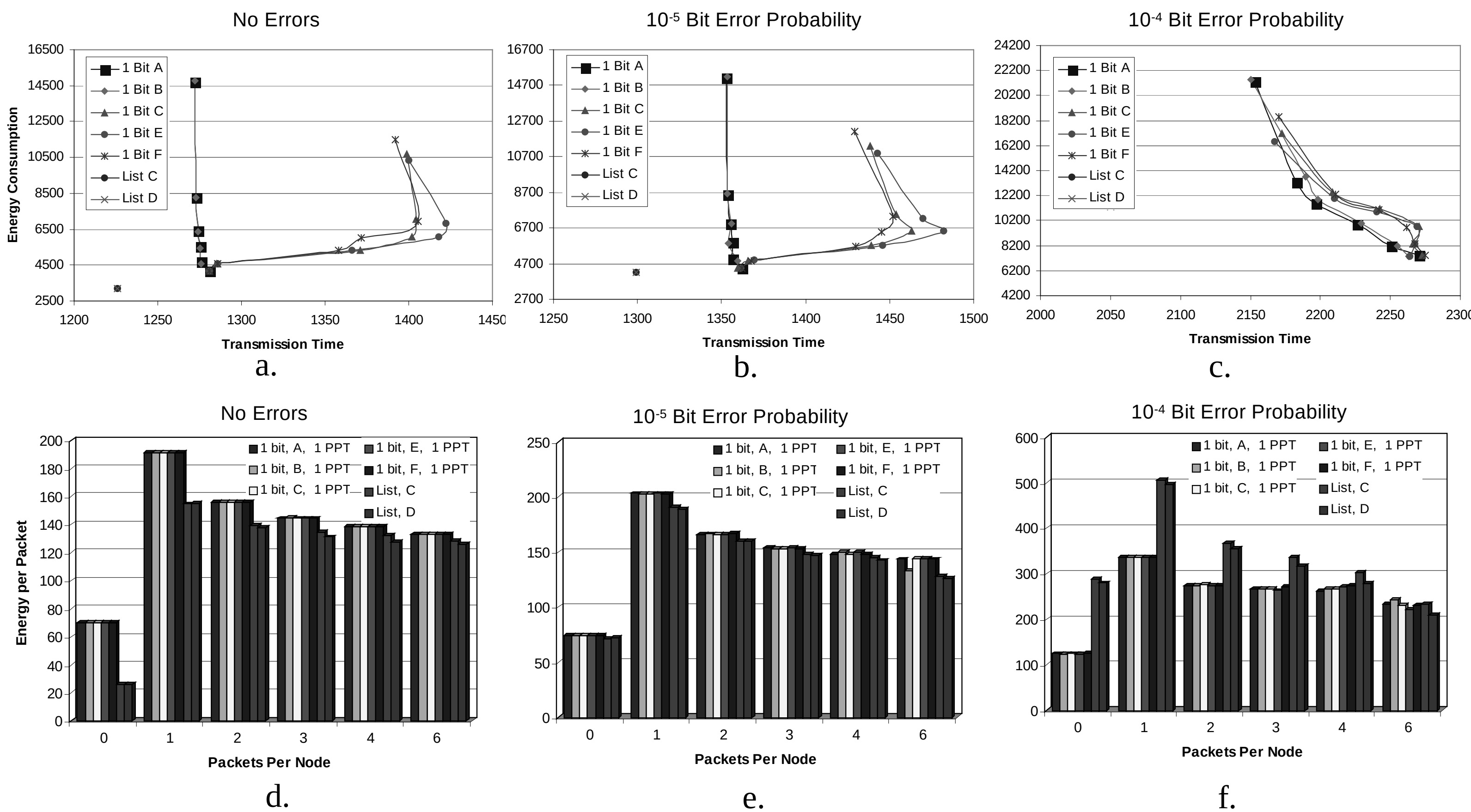


Figure 3.13: Performance of Type I traffic, 25 nodes, 10 packets

Figures 3.11, 3.12, and 3.13 are samples of our data. Panels a, b, and c illustrate the performance of the protocols for different error conditions and for different TIM periods (i.e. 10, 5, 4, 3, 2, and 1 packet per TIM) for a network size of 25 nodes and a CFP of 10 packets. Panels d, e, and f illustrate the energy consumed per node based on the number of packets the node sent in the cycle using the optimum TIM periods (i.e. least energy consuming) for the same networks. Of interest is that both the error condition and the type of traffic affect which protocol performs best.

3.6.1 Observations and Design Implications

The results of the simulations bring out the following four key design concepts.

1 Synchronize the waking up of dozing nodes with the broadcast of the status of the transmission schedule.

Networks with errors consume large amounts of energy on account of nodes waking early and having to wait for directory information before they can return to the doze state. Synchronizing the waking up of nodes to the transmission of the status of the transmission schedule minimizes this energy loss. The benefit of this type of synchronization is seen in the relative performance of the 1 bit TIM protocols. Observe Panels d, e, and f of Figures 3.11, 3.12, and 3.13. Note that the energy consumed per node for those nodes not transmitting or transmitting just one packet in a cycle increases rapidly with errors for the m bit TIM and list protocols but remains nearly constant for the 1 bit TIM protocol. The most significant difference between the 1 bit protocols and the other protocols is that the 1 bit protocol seeks this synchronization. The 1 bit protocol requires packet transmissions to be rescheduled into a subsequent TIM period if they would interfere with the broadcast of a TIM. As a result, dozing nodes are assured of receiving a TIM when they wake-up.

We did not require directories in the m bit and list protocols to be transmitted on a regular schedule since the nodes in these protocols could wake-up and estimate a

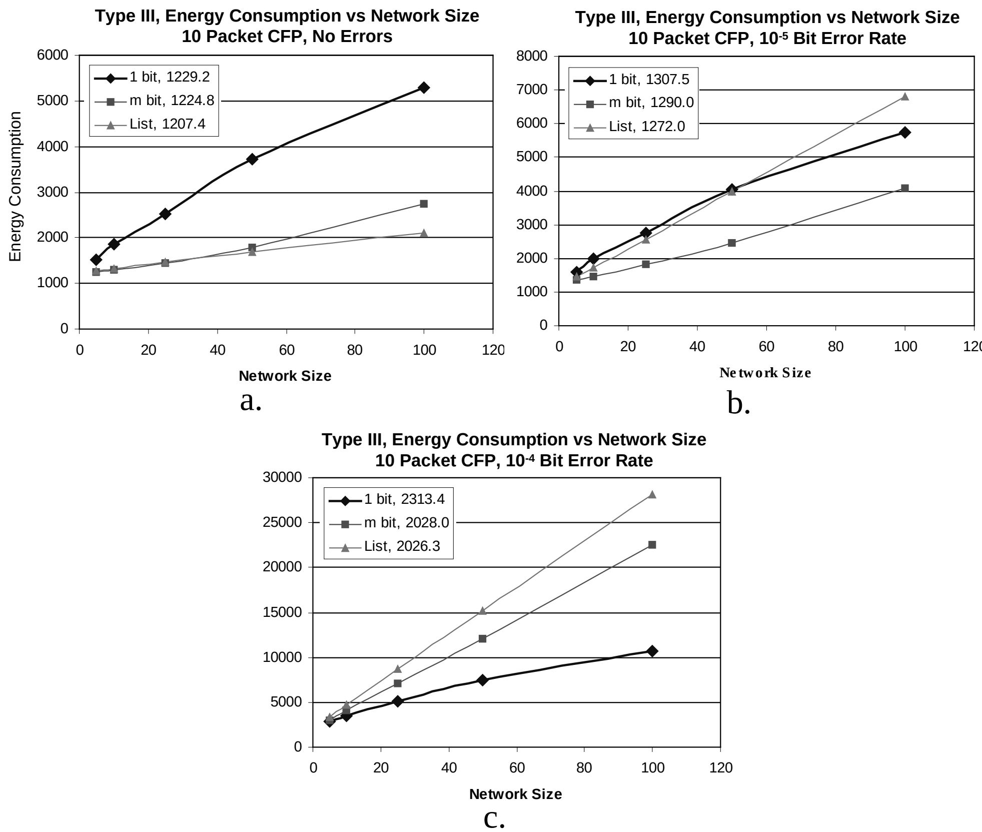


Figure 3.14: Protocol performance at optimum for Type III traffic with 10 packet CFPs[12]

new dozing period based on the known schedule for transmissions. The m bit protocol can learn the status of the network after the receipt of a poll and the list protocol learns the status after the receipt of the header information in a transmitted packet or ACK. These features allow the list and m bit protocols to achieve better throughput. The difference in energy consumption is caused by the cumulative effect of nodes waking up and having to wait for the beginning of one of these transmissions to determine what to do next. This occurs most frequently for nodes that have no transmissions or transmit their traffic early in the CFP since they are most likely to wake

---

[12] The legend provides the average transmission time for the protocols. Our results demonstrated that the network size had only a minor effect on the time it takes to transmit k packets despite the fact it affects the size of TIMs and lists.

up in the middle of the CFP expecting it to be over. These nodes estimate the time to wake up based on the assumption that there are no errors. In a high error environment they are likely to wake up several times expecting the CFP to be over. As illustrated in Figure 3.14, the cumulative effect of errors drastically affects the relative performance of the protocols.

2. The optimum period for TIM protocols is independent of error conditions.

We found that the optimum TIM period for 1 bit protocols is most affected by network size. The trend is that as the network gets larger a longer TIM period performs better. When networks are larger so too are the TIMs. Additionally, more nodes must listen to the TIMs. The penalty of more nodes listening to larger TIMs exceeds the penalty of nodes having to wait in the awake state for their turn to participate in a data exchange. Note that the TIM period is slightly lower for Type I traffic (see Figure 3.13). Since in Type I traffic two nodes participate in each data exchange, more nodes on average remain awake in a TIM period. Fortunately, since the optimum TIM period is not affected by error rate, it can be selected using the model found in Appendix A. Note, however, that the assumption in these models is that the traffic is uniformly distributed. Any clustering of traffic between a few nodes will tend to make even larger TIM periods more attractive.

The optimum TIM period for the m bit TIM protocol is the transmission cycle size. There is no benefit to using multiple TIM periods per cycle. Sufficient information is provided in the polls to allow the nodes to correct their dozing periods.

3. Scheduling Algorithm C provides excellent performance for all types of traffic and for all types of protocols.

Algorithm C is recommended for three reasons. It is the easiest algorithm to implement, it is the least complex of the algorithms listed in this paper, and it provides either optimum or near optimum schedules. When used with Type II, III, or IV traffic it is equivalent to the shortest processing time first scheduling policy and is therefore optimum. As demonstrated in the Type I traffic simulations all of the scheduling algorithms perform about the same for the 1-bit TIM protocols when the optimum TIM period is selected. The performance of the scheduling algorithms used with the list protocols supporting Type I traffic were also nearly identical for all network sizes, CFP sizes, and error rates.

4. **Scheduling all exchanges with a node in contiguous slots and reducing energy consumed by idle nodes promotes fairness.**

In energy conserving protocols, fairness might be interpreted as providing the best energy consumption to those nodes that take action to help the network conserve energy and achieve best throughput. A node contributes to these two goals by accumulating traffic and sending it contiguously. Such clustering of transmissions effectively reduces the active nodes in a CFP and in turn the number of TIMs required in the TIM protocols, the average number of nodes awake during the TIM periods, and the number of nodes penalized when there are errors. To encourage this behavior, the protocol should provide best energy consumption per packet to those nodes that send the most traffic in a CFP and minimize the energy consumed by the idle nodes in the CFP. As seen in the results, providing best energy efficiency to the nodes that send the most traffic appears to be the natural trend and is attributed to scheduling exchanges between common nodes in contiguous slots.

The second fairness objective is minimizing the energy consumed by idle nodes. Excessive consumption of energy by idle nodes eliminates the incentive to accumulate. This occurs when the energy consumed in idle periods exceeds the marginal difference in energy saved per packet with accumulation. For example, say the difference in energy consumed per packet transmission between sending one packet per CFP and sending two packets per CFP is 20 units but that the energy consumed when a node is idle is 50 units. Sending one packet in two CFPs would save the node 10 energy units. Methods to save idle nodes' energy such as synchronization help achieve fairness.

3.6.2 Improvements

The total energy consumption of each of the protocols described in this chapter can be improved by focusing on two characteristics, synchronization and better dozing periods. The protocols that can benefit most from improvements in synchronization are the list and m bit TIM protocols. The 1 bit TIM protocol benefits from better dozing periods.

For both the list and m bit protocols we considered two possible synchronization methods. In the first the base station monitors and announces the perceived error rate. Nodes that are not participating in data exchanges then use this error rate to estimate when the conclusion of the CFP will be and thus avoid waking up early. The risk with this procedure is that the base station cannot be assured that all nodes in the network are awake at the conclusion of the packet exchanges and must wait the worst case estimated time before directing the start of the CP. The second synchronization method attempted was to delay retransmission of failed packets until after a subsequent list or TIM. Nodes that wake up to receive a directory are assured to receive one. Also, nodes that wake up to participate in an exchange are assured of waking up on time.

The above recommendations were tested on both the multiple bit TIM and list protocols for Type III and IV traffic. In the case of the estimation versions the actual error rates (i.e. the ideal case) were used to estimate wake up times. Figure 3.15 illustrates the results for list and m bit protocols for Type III traffic and 10 packet CFPs. The results for Type IV traffic were similar. The estimation versions of the protocols are labeled with the letter E and the delayed retransmission versions are labeled with

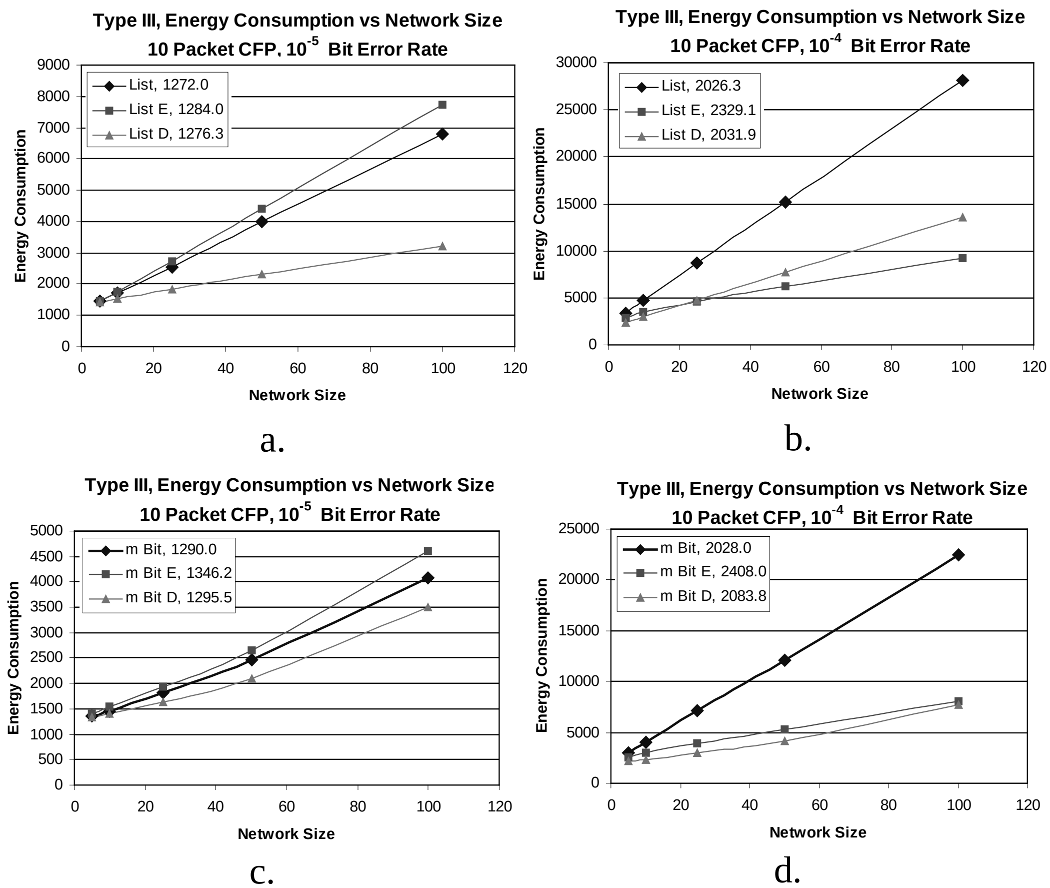


Figure 3.15: Comparison of protocols modified for improved performance with Type III traffic

the letter D. The average service time is shown in the legend. As expected, the advantage of these protocol variations increases with both the error rate and the network size. The whole purpose of these variations is to react to errors and the benefit increases as there are more nodes that can benefit. The performance of the estimation versions also improves as the CFP length increases. This is not the case for protocols using delayed transmission. The statistical nature of the estimation approach allows it to improve with larger numbers. The delayed retransmission, however, is likely to have more transmissions of lists or TIMs when there are larger CFPs resulting in more energy consumption by all nodes in the network. The delayed retransmission protocols cause an insignificant increase in transmission time e.g. less than 1%. The estimation versions, however, result in longer transmission times. They were as much as 20% longer in our simulations.

Three improvement techniques were attempted with the 1 bit TIM protocol, layered TIMs, local optimization of the TIM period, and delayed retransmission. Together, these techniques decrease the energy consumption of the 1 bit protocol despite the error conditions or the type of traffic.

The objective of the layered TIMs is to reduce the energy that is consumed by idle nodes listening to TIMs. The transmission cycle is layered into sequentially smaller TIM periods. A larger TIM period in the outer layer is subsequently divided into smaller TIM periods in the inner layer. The TIM for the outer layer reduces the set of nodes that listen to the TIMs of the inner layer. So at the beginning of the outer TIM period two TIMs are transmitted. The first TIM specifies a large TIM period and puts nodes to sleep that will not be participating in any data exchanges for that larger period. The second TIM then manages the reduced set of awake nodes.

Local optimization of the inner TIM periods is motivated by our observation that the TIM periods are affected by the distribution of traffic and not by error rates. Each inner TIM period is selected based on the next transmissions scheduled. It is at least as long as the number of transmissions of the next pair of nodes. A longer TIM period is selected if the penalty of having the active nodes of the inner TIM period listen to an additional TIM is more than the penalty of having the next pair of scheduled nodes listen to the preceding packet transmissions.

Finally, delaying retransmissions eliminates one of the penalties of immediate retransmission. Immediate retransmission error recovery results in every node waiting to transmit in a TIM period staying awake for each retransmission. Delayed retransmission only penalizes the nodes involved in failures. Retransmissions are only delayed until the next inner TIM period.

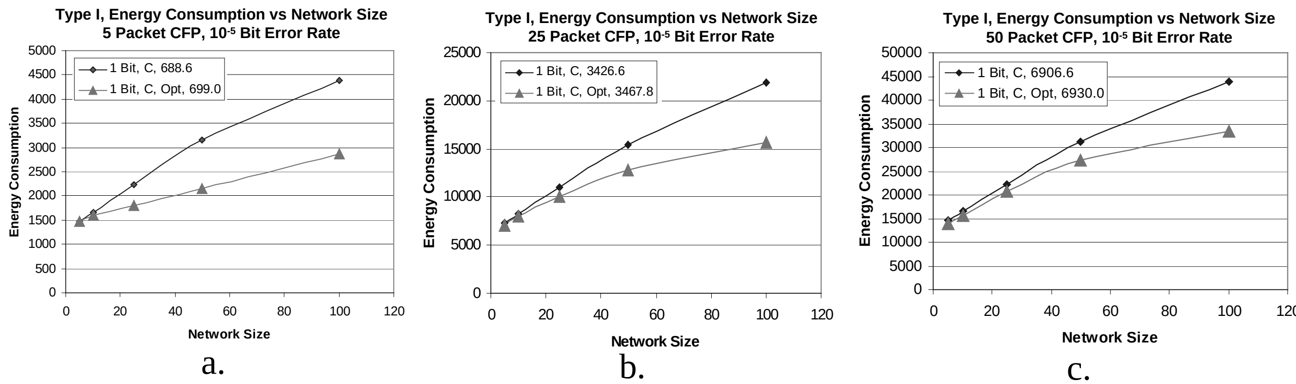


Figure 3.16: Comparison of the optimized and standard 1 bit TIM protocols with Type I traffic

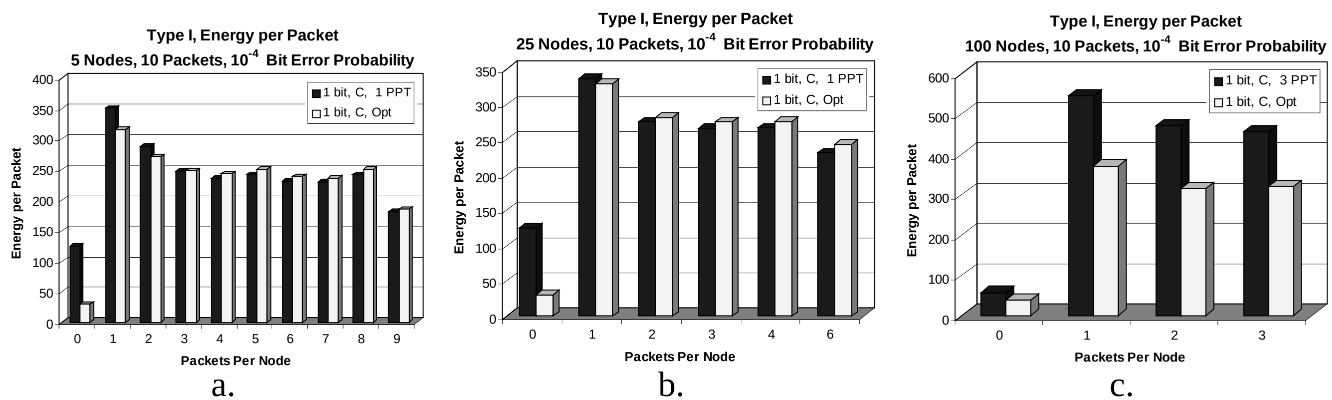


Figure 3.17: Comparison of the optimized and standard 1 bit TIM protocols with Type I traffic

The techniques above were attempted for all types of traffic. Figures 3.16 and 3.17 compare the performance of the standard 1 bit TIM protocol to that of the optimized 1 bit TIM protocol for Type I traffic. The optimized 1 bit TIM protocol always consumes the least energy. Figure 16 shows that the relative size of the improvement increases with network size but decreases as the CFP increases. These results illustrate that the most energy is conserved by the first improvement technique, i.e. layered TIMs. The objective of the outer TIM is to reduce the energy consumed by idle nodes. The more idle nodes there are the more effective it is. The number of idle nodes increases as the network size increase, and decreases as the CFP increases. Figure 3.17 not only illustrates that the idle nodes benefit the most from the optimized protocol but that the benefits for nodes that transmit traffic increases with network size. The dependence on the network size is caused by the change in the optimum TIM period of the standard 1 bit TIM protocol. The larger TIMs increase the optimum TIM period for the standard 1 bit TIM protocol reducing the penalty to the idle nodes but increasing it for nodes that transmit traffic. The optimized 1 bit TIM protocol has consistent performance for all network sizes.

3.6.3 Choosing the Best Traffic Management Protocol

Our results show that protocol performance needs to be compared in the high error environments as some protocols that perform very well when there are no errors quickly degrade when efforts are made to resend failed transmissions. In Figure 3.18 we compare the energy conservation of all the improvement techniques described in Section 3.6.2 for Type III traffic in a high error environment. These graphs demonstrate that the performance of all the protocols are fairly close to each other with the relative difference depending on network size. The consistent result for both Type III and IV traffic, all network sizes, and all transmission cycle sizes is that the delayed

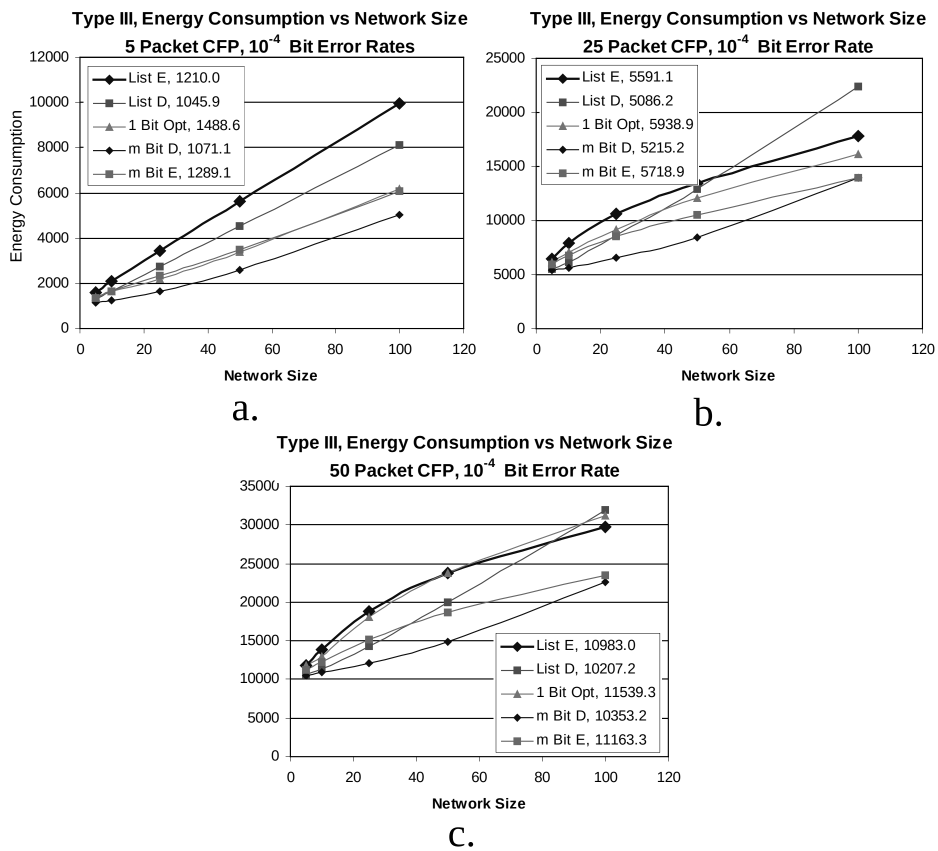


Figure 3.18: Comparison of Improved Protocols for Type III Traffic

retransmission version of the m bit TIM protocol is the best at conserving energy. Moreover, the delayed retransmission version of the m bit protocol was very competitive in throughput with only the delayed transmission version of the list protocol performing better.

From these results we recommend that the delayed retransmission version of the m bit protocol be used for Type II, III, and IV traffic. If the protocol is also to support Type I traffic, the m bit protocol cannot be used. In this case the optimized 1 bit TIM protocol provides the best energy conserving results for Type I traffic and competitive results for the other types of traffic. The optimized 1 bit TIM protocol achieves this energy conservation with a small sacrifice in throughput.

3.6.4 Sensitivity

Energy consumption is sensitive to three factors, the ratio of overhead to payload, transition times and the actual energy consumed in transmitting and receiving. The design of the protocols presented in this paper remain robust to all of these factors with the exception of transition times

Energy consumption is directly related to the ratio of overhead to payload. Our analysis attempted to choose representative numbers basing packet size on throughput performance for representative bit error rates and basing overhead, most specifically physical overhead, on current wireless networking standards. [6] [7] The ratio of overhead to payload for a packet was about 14%. This number considers the overhead of the packet transmission, the ACK and the transition time between them but not polls or directories. Popular in current literature is to design protocols to support the 53 byte payload of ATM packets. If packets are then sent individually, overhead accounts for nearly 65% of the transmission size. In turn, for a given payload, each participant in the data exchange consumes at least 40% more energy. This percentage increases when polls and directories are included in the overhead portion of the ratio. Reducing the overhead to payload ratio is an energy conservation objective and in the case of ATM supports the idea of sending several ATM cells per packet transmission.

Transition times between the energy states of commercially available transceivers are not presented in specification sheets. Both the HIPERLAN and IEEE 802.11 standards specify a maximum transition time between the transmit and receive states, 6 and 5 μsec respectively.[13] Neither standard provides any guidance on the transition time from the doze state to the awake state. It is this transition that is most critical to the design of energy conserving protocols. We have assumed in this paper that the transition between the doze and awake state occurs in a time comparable to that between the transmit and receive states. Longer transition times would make protocols that attempt to put the most nodes to sleep for longer periods of time the better performers. An objective in protocol design in this context would be to reduce the frequency of waking up to listen to directories. The 1 bit TIM protocol which relies on frequent short directories as opposed to the single long directories of the list and the m bit TIM protocols would be affected most by these long transition times. The optimum TIM period of the 1 bit TIM protocol would increase. If the transition time from the receive state to the doze is large, longer than the time to transmit a packet, protocols other than those presented in this dissertation may be best.[14]

The analysis in this dissertation does not make a distinction between the energy consumed in transmitting and the energy consumed in receiving. Nevertheless, we did keep statistics on transmission time and receiving time. The average transmission time was the same for all the protocols when the network size, transmission cycle size and error rate were the same. Differences in energy consumption were purely the result of the time spent by nodes receiving. Efforts to reduce the time spent transmitting may be worthwhile but there is a limited amount of this time to reduce, primarily overhead. These efforts would be equivalent to reducing the overhead to payload ratio. This effort is likely to result in more energy in the network being saved by multiple nodes receiving less than by individual nodes transmitting less. The ratio of energy consumed in transmitting versus that consumed in receiving has no effect on our results.

---

[13] With a transition time of 5 µsec, a 48 bit transition is equivalent to a data rate of 9.6 Mbps.

[14] The one reference we found that provided this transition time, [126], reported that it took 100 msec to transition from the doze to the awake state. This value was empirically measured. No information was provided on how it was measured. A 100 msec transition time is equivalent to 34 packet transmissions using our packet size and a transmission rate of 2 Mbps. None of the protocols in this dissedrtation nor protocols of most other papers published to date would be feasible.

3.7 Conclusion

This paper described various methods for a central controller to manage the transmission of fixed sized packets. The results show that protocols can vary greatly and that their performance depends on the traffic type, the channel characteristics in which they operate, as well as the parameters, i.e. TIM period and CFP size, chosen for the protocol. These results also show that the scheduling policy is an important choice in designing the protocol but that in most cases a single algorithm, servicing the nodes with fewest transmissions first, should be used. It provides optimum performance for Type II, III, and IV traffic and excellent performance for Type I traffic. Fortunately, this scheduling policy is one of the easiest algorithms to implement.

The most significant conclusion of this research is that the focus of energy conservation should be to reduce the set of nodes in the receive state. Two goals in protocol design are shown to be critical. First, the protocol should attempt to put as many nodes to sleep as possible as early as possible and for as long as possible. Second, the protocol should coordinate transmissions such that nodes wake up to hear a transmission that enables them to know the state of the transmission schedule. Although our work considers a half duplex channel these conclusions can be extrapolated to full duplex systems.

As a final comment we provide one additional result that emphasizes our conclusion that putting idle nodes to sleep is the critical objective of energy conserving protocols. We compare the energy savings using the best performing protocols of the smallest network and the smallest transmission cycle with that of the largest network and the largest transmission cycle for Type III traffic in high error conditions. The smaller network consumed 22% of the energy that would have been consumed if all nodes were awake for the entire transmission cycle. This equates to batteries used by the transmitters of this network lasting nearly 5 times longer on average. The larger network, however, consumed just 2.2% of the energy that would be consumed if all nodes were awake throughout the transmission cycle. Batteries in this network would last 45 times longer on average. Energy conserving protocols have much greater potential in larger networks since more nodes can doze.

Appendix 3.A

Models for Delay and Energy Consumption for Protocols Managing Types III and IV Traffic

We model the transmission time and energy consumption of the 1 bit TIM, multiple bit TIM, and list protocols for Type III and IV traffic when there are no errors. The transmission times are easily modeled since they are independent of any scheduling or traffic distribution factors. They are dependent on the size of the network, *n*, the number of packets being transmitted in the contention free period (CFP), *k*, and in the case of the TIM protocols, the number of TIM periods that are used, *j*. For these models we define five timing variables, the time to transmit a packet including its overhead, τ_{Pkt}, the time to transmit a poll including it's overhead, τ_{Poll}, the time to transmit an ACK including its overhead, τ_{ACK}, the time for the transmission overhead, τ_{OH}, and finally the duration of an interframe space, τ_S. The distinction between the equations for the three different types of traffic are the duration of the directories, the use of polls, the number of interframe spaces used, and some subtle end conditions.

We define $t_{1bit}(k,j)$ as the time to transmit *k* packets if *j* TIM periods are used with the 1 bit TIM protocol. This is given by:

$$t_{1bit}(k,j) = j\left\lceil \frac{n}{|S|} \right\rceil + k\left(2\tau_S + \tau_{Poll} + \tau_{Pkt} + \tau_{ACK} - \tau_{OH}\right) + c_1 . \qquad (3.A.1)$$

The first term of Equation (3.A.1) accounts for the transmission of the TIM. The size of the TIM is dependent on the number of nodes in the network adjusted to fit evenly into an integer number of time slots. Overhead is not included since the TIM is combined with the very next Poll thus overhead is included in the second term. We use $|S|$ to refer to the number of bits that can be transmitted in an interframe slot, in our case 48. The second term of (A.1) accounts for the time to transmit packets. Each packet transmission includes a poll, a packet, and an ACK. Since the polls are combined with either a packet or an ACK, in Type III or IV traffic respectively, one overhead transmission can be avoided. The interframe spaces occur between packet exchanges and between the packets and ACKs of each exchange. The third term accounts for the special end condition for Type IV traffic when the final ACK of a TIM period is not combined with a poll since a TIM is transmitted next. The constant c_1 has a value of 0 for Type III traffic and a value of $j \cdot (\tau_{OH} + \tau_S)$ for Type IV traffic.

We define $t_{mbit}(k, j)$ as the time to transmit k packets if j TIM periods are used with the multiple bit TIM protocol. It is given by:

$$t_{mbit}(k,j) = j\left(\tau_{OH+}\left\lceil \frac{k\left(1+\left\lfloor log_2\left\lceil k/j \right\rceil\right\rfloor\right)}{|S|} \right\rceil\right) + k\left(2\tau_S + \tau_{Poll} + \tau_{Pkt} + \tau_{ACK} - \tau_{OH}\right) + c_1 \quad (3.A.2)$$

It only differs from (A.1) in the first term since the TIMs are a different size. Each TIM includes n multiple bit positions. Each position use the number of bits required to specify the number of packets that can be transmitted in a TIM period.

We define $t_{list}(k)$ as the time to transmit k packets using the list protocol, which is given by:

$$t_{list}(k) = \left(\tau_{OH+}\left\lceil \frac{k\left(1+\left\lfloor log_2 n \right\rfloor\right)}{|S|} \right\rceil\right) + k\left(m \bullet \tau_S + \tau_{Pkt} + \tau_{ACK}\right). \quad (3.A.3)$$

The first term of (A.3) accounts for the time to transmit a list with abbreviated addresses. The size of the abbreviated addresses depends on the size of the network and the number of addresses in the list depends on the number of packets the list directs to be transmitted. The second term accounts for the time to transmit a packet. Each exchange includes the time to transmit a packet and an ACK. The number of interframe spaces, *m*, required depends on the traffic type. The PC requires precedence, i.e. priority, during the CFP for control purposes. To give this precedence to the PC we require that all mobile nodes wait two interframe spaces before attempting to transmit a new packet. The PC only needs to wait one and thus would have priority. Therefore, *m* is 3 for Type IV traffic and only 2 for Type III traffic.

Next we compute the average energy consumed in transmitting *k* packets. The energy consumed by a given collection of *k* packets depends on how these *k* packets are distributed among the network's *n* nodes and how they are scheduled to be transmitted. We determine the distribution of the packets by first conditioning on the number of nodes *i* spanned by the *k* packets to be transmitted. Let $p_{k,n}(i)$ denote the probability that the *k* packets are sent to/from *i* nodes in the network:

$$p_{k,n}(i) = \frac{\binom{n}{i}\sum_{j=0}^{i}(-1)^{j}\binom{i}{j}(i-j)^{k}}{n^{k}}. \tag{3.A.4}$$

Since our scheduling policy orders transmissions based on the number of packets exchanged with each mobile node, irrespective of which mobile node, our interest is in determining the likelihood of a given *type* of partition for *k* packets among *i* mobile nodes, where each mobile node participates in at *least* one exchange. Let $P_{i,k}$ denote one such partition and let $\mathcal{P}_{i,k}$ denote the set of all such partitions. We define a type of a partition $P_{i,k}$ as a vector $q(P_{i,k}) = (q_1,...q_i) \in \mathbb{N}^i$ with non-decreasing coordinates, where the j^{th} coordinate corresponds to the number of packets sent to a destination receiving the j^{th} smallest number of packets among the i nodes, thus $\sum_{j=1}^{i} q_j = k$. We shall let $\mathcal{Q}_{i,k} \subset \mathbb{N}^i$ denote the set of vectors corresponding to all possible partition types for k packets among i destinations. Finally, for any $t \in \mathcal{Q}_{i,k}$ we define a set of partitions $\mathcal{P}^t$ that have type t, i.e.

$$\mathcal{P}^t = \left\{ P_{i,k} \in \mathcal{P}_{i,k} \middle| q\left(P_{i,k}\right) = t \right\}.$$

and $\left|\mathcal{P}^t\right|$ as the number of partitions of type t.

We shall let $p_{k,n}\left(t \middle| i\right)$ denote the probability that a partition of Type $t \in \mathcal{Q}_{i,k}$ is obtained given the k exchanges are among i of the network's mobile nodes. Using a counting argument one can show that

$$p_{k,n}\left(t \middle| i\right) = \frac{\left|\mathcal{P}^{\mathrm{t}}\right|}{\sum_{r \in \mathcal{Q}_{\mathrm{i,k}}} \left|\mathcal{P}^{\mathrm{r}}\right|} \text{ where } \left|\mathcal{P}^t\right| = \binom{k}{t_1 \quad t_2 \quad \cdots \quad t_i}, \tag{3.A.5}$$

since the traffic is assumed to be uniformly distributed and so partitions are equally likely.

For example, suppose there are 10 mobile nodes in the network labeled d_1, d_2, …, d_{10}, and that three of these nodes are participating in five packet exchanges. The probability of this event is written $p_{5,10}(3)$ and is 0.18 by Eq. (A.4). Say the exchanges are with the following nodes, d_2, d_3, d_3, d_3, and d_4, then the partition would be written as $q\left(P_{3,5}\right) = (1,1,3) = t$ where $t_1 = 1$, $t_2 = 1$, and $t_3 = 3$. The only other possible partition with three mobile nodes and 5 exchanges would be $q\left(P_{3,5}\right) = (1,2,2)$ and we find that

$$p_{5,10}(t|3) = \frac{\binom{5}{1\ 1\ 3}}{\binom{5}{1\ 1\ 3} + \binom{5}{1\ 2\ 2}} = 0.4 ,$$

and the probability of this schedule over all possible schedules with $k = 5$ and $n = 10$ is $p_{5,10}(3)\, p_{5,10}(t|3) = 0.072$.

Now suppose the k packets to be transmitted correspond to a partition of type $t = (t_1, \cdots t_i)$. Recall that $t_1 \le t_2 \le \cdots \le t_i$ since jobs are scheduled shortest processing time first. Jobs are defined as the set of exchanges to a single node and its size is the number of exchanges in the set. Further suppose that these packets are equally distributed among j consecutive TIMs. We shall let $n_t(l)$ denote the number of nodes participating in exchanges during the l^{th} TIM period and $m_t(r,l)$ the number of packets sent by the r^{th} node participating in the l^{th} TIM period in the partition of type t. Clearly $n_t(l)$ and $m_t(r,l)$, $r = 1, \ldots, n_t(l)$ are directly determined from t. So in the above example when we use TIM periods of size 2, then there are three TIM periods required to send the five packets and $n_t(1) = 2$, $n_t(2) = 1$, and $n_t(3) = 1$ and $m_t(1,1) = 1$, $m_t(2,1) = 1$, $m_t(1,2) = 2$, and $m_t(1,3) = 1$.

We now have the basic definitions to build our models. The models for each protocol consider four different energy components of the CFP, energy consumed in transmitting directories, adjustments for end condition of directory transmissions, energy consumed in transmitting packets, and, finally, adjustments for end conditions of packet transmissions, TIM periods, and/or the CFP. The first component is the energy consumed in receiving directories. All nodes are assumed to be awake to listen to these directories. For each directory that is transmitted, nodes must awaken and then return to the doze state if not identified as needing to stay awake. In our models we assume that all nodes awaken and return to the doze state in the first component and then adjust in the second component for those nodes that actually remain awake after the directory transmissions. In the third component we account for the energy consumed in transmitting packets. This accounts for the number of nodes awake and the duration of the transmission. Finally, we make adjustments for the end conditions such as the time for a node in the 1 bit and multiple bit TIM protocols to identify that it can enter the doze state and the time in the multiple bit TIM and list protocols a node must be awake before the first exchange.

We start with the model for the 1 bit TIM protocol. The amount of energy consumed in sending k packets is dependent on the partition t. Our approach to determining the expected energy consumption is to consider the energy consumed by each possible partition weighted by the probability of the partition's occurrence when traffic is distributed uniformly. We define the energy consumed by a particular partition of type t to be

$$e_{1bit_{\mathbf{k,j}}}(t)=\begin{cases} j\cdot n\cdot\left(2\cdot\tau_S+\left(\tau_{OH}+\left\lceil\frac{n}{|S|}\right\rceil\right)\right)- \\ \sum_{l=1}^{j} n_t(l)\cdot(\tau_S+\tau_{OH})+ \\ \sum_{l=1}^{j}\sum_{r=1}^{n_t(l)}\left(n_t(l)+1-r\right)m_t(r,l)\left(2\cdot\tau_S+\tau_{Poll}+\tau_{Pkt}+\tau_{ACK}-\tau_{OH}\right)+ \\ c_1+\sum_{l=1}^{j}\left(n_t(l)-1\right)\left(\tau_S+\tau_{Poll}\right) \end{cases}, \qquad (3.A.6)$$

where c_1 is the constant defined earlier to account for differences in Type III and IV traffic transmission. To determine the expected energy consumption we consider all possible partitions, i.e.

$$\mathbf{E}\left[e_{1bit\boldsymbol{k,j}}(T)\right] = \sum_{i=1}^{min(n,k)} p_{k,n}(i) \sum_{t \in Q_{i,k}} p_{k,n}(t|i) e_{1bit\boldsymbol{k,j}}(t). \tag{3.A.7}$$

To evaluate A.7 we use a recursive algorithm to determine all $t \in Q_{i,k}$.

The expected energy consumption model for the multiple bit TIM protocol is formulated in much the same way. Again , energy consumption is dependent on the partition as follows:

$$e_{mbit\boldsymbol{k,j}}(t) = \begin{cases} j \cdot n \cdot \left(2 \cdot \tau_S + \left(\tau_{OH} + \left\lceil \dfrac{n\left(1 + \left\lceil log_2 \left\lceil {}^{k}\!/_{j} \right\rceil \right\rceil \right)}{|S|} \right\rceil \right)\right) - \\ j \cdot \tau_S + \\ k \cdot (2 \cdot \tau_S + \tau_{Poll} + \tau_{Pkt} + \tau_{ACK} - \tau_{OH}) + \\ c_1 + \sum_{l=1}^{j} \left(n_t(l) - 1\right) c_2 \end{cases} . \tag{3.A.8}$$

The constant c_1 is as defined before. The constant c_2 accounts for the differences between Type III and Type IV traffic in the time a node must be awake prior to the first packet exchange. It is $2 \cdot \tau_S$ for Type III traffic and $(\tau_{ACK} - \tau_{OH} + \tau_S)$ for Type IV traffic. And finally, to determine the expected energy consumption we average over all possible partitions:

$$\mathbf{E}\left[e_{mbit\boldsymbol{k,j}}(T)\right] = \sum_{i=1}^{min(n,k)} p_{k,n}(i) \sum_{t \in Q_{i,k}} p_{k,n}(t|i) e_{mbit\boldsymbol{k,j}}(t). \tag{3.A.9}$$

The expected energy consumption model for the list protocol is much simpler since it only depends on the number of nodes participating in the CFP rather than the partition type and the number of TIMs. The energy consumed when i nodes participate in the CFP is

$$e_{list\boldsymbol{k}}(i)=\begin{cases} n\cdot\left(2\cdot\tau_S+\left(\tau_{OH}+\left\lceil\dfrac{k\left(1+\lfloor log_2 n\rfloor\right)}{|S|}\right\rceil\right)\right)- \\ \tau_S+ \\ k\cdot\left(3\cdot\tau_S+\tau_{Poll}+\tau_{Pkt}+\tau_{ACK}\right)+ \\ (i-1)\left(3\cdot\tau_S+\tau_{ACK}\right) \end{cases}. \tag{3.A.10}$$

The expected energy consumption is

$$\mathbf{E}\left[e_{list\boldsymbol{k}}\right]=\sum_{i=1}^{min(n,k)} p_{k,n}(i)e_{list\boldsymbol{k}}(i). \tag{3.A.11}$$

Appendix 3.B

Scheduling Algorithms

Descriptions of the scheduling algorithms used in our study of the transmission of Type I traffic follow.

Algorithm A

1. Select the pair of addresses that occur most and use these transmissions to fill the first slots of the transmission period.
2. If there are more slots in the TIM period, add any transmissions of pairs of nodes that are already awake that are not already part of the transmission schedule. Add the pairs that occur least frequently first.[15]
3. If there are more slots in the TIM period, identify the pair with one address already in the transmission set that occurs most frequently. Use these transmissions to fill the remaining slots of the TIM period.
4. Repeat Steps 2 and 3 until all slots of the TIM period are filled or until no more pairs can be found to meet the criteria of these steps. If the latter occurs go to Step 1.
5. Repeat all steps for subsequent TIM periods until all traffic has been transmitted.

Algorithm B

1. Set m = 2. Let q = the number of possible transmissions in a TIM period.
2. Consider all nodes remaining with traffic. Count the occurrences of transmissions between all combinations of nodes taken m at a time.
3. If there are no combinations of m nodes with a quantity of transmissions greater than or equal to q, increase m by 1 and repeat Step 2.

[15] There is a benefit to exhausting as many of these pairs as possible since there is no penalty for participating in the given TIM period but there may be in a subsequent TIM period.

4. Choose the combination of m nodes that meet the criteria of Step 3 that has a quantity of transmissions closest to q. Schedule these in the next transmission slot.
5. Repeat starting at Step 2 for the next transmission cycle.

Algorithm C

1. Count the frequency that each address appears in the set of transmissions waiting to be scheduled.
2. Schedule those transmissions involving the node that has the lowest frequency.
3. Repeat Steps 1 and 2 until all traffic is scheduled.

Algorithm D

1. Count the frequency that each node participates in data exchanges and select the node with the most exchanges.
2. Place all data exchanges involving the node identified by Step 1 into a subgroup for transmission.
3. Repeat Steps 1 and 2 for the transmissions that have not been added to a subgroup. Advance to Step 4 if all exchanges have been added to a subgroup.
4. Group all exchanges within each subgroup between common pairs of nodes. If the identifying node of the subgroup is both transmitting and receiving in the set of a common pair, schedule the receptions last.
5. Go to the transition of the first two subgroups.
6. Reschedule the transmissions of the two adjacent subgroups such that at the transition there is a common node. Give preference to a node that receives in the forward subgroup. If necessary swap the order of the transmissions and receptions of the last pair in the forward subgroup.
7. If Step 6 is successful and there are more transitions, advance to the next transition and repeat Step 6. Otherwise, stop.

8. If common nodes cannot be found in the adjacent subgroups move the latter subgroup to the end of the transmission schedule and repeat Step 6 for the new adjacent pair.

Chapter 4

Conserving Energy in Accessing a Central Node

4.1 Introduction

In Chapter 3 we explored the energy conserving methods to manage the transmission of data. The model of the communications process was one that alternated between contention periods (CP) and contention free periods (CFP). The objective of Chapter 3 was to describe how a central node called a point controller (PC) should manage the transmission of data in the CFP to achieve best energy efficiency. It was assumed the PC became aware of pending traffic in the CP but there was no discussion on how. In this chapter we identify the access schemes that may be used in this type of network to gain access and we evaluate them for energy efficiency.

This alternating type of protocol removes the arrival process as an issue in the analysis. We assume at the start of the CP there is a set of nodes that have traffic to send and they are contending to notify the PC of their status.[16] Although the success of a protocol will be dependent on the number of nodes needing to gain access it is independent of when the packets arrive at those nodes. Classical access methods that assume a random arrival such as 1-persistent CSMA[24] and both Aloha and slotted

[16] The PC manages the actual transmission data in subsequent CFPs

Aloha [25] are not appropriate since they are assured of colliding. Similarly, the collision avoidance scheme of the 802.11 protocol [6] is not suitable. The 802.11 protocol uses a sliding contention probability that adjusts to congestion first assuming minimum congestion and then decreasing the contention probability as collisions occur. Since contending nodes accumulate in anticipation of the CP and they decrease in quantity as the CP progresses, we would expect the adjustment process of 802.11 to be counterproductive. The protocols that support access at the congested start of a CP will use either a collision avoidance or a collision resolution scheme. In the collision avoidance scheme, rather than ramping-up the collision avoidance with the occurrence of collisions, we choose the collision avoidance parameters for the expected congestion and then stick with it for the duration of the contention period.

In Chapter 3 we concluded that the most significant objective of an energy conserving traffic management protocol was to put idle nodes into the doze state. This goal is not as severe in the CP since nodes without any need to contend can doze throughout the CP without risk. The goal then is to help the contending nodes reduce their awake time. We answer several questions. 1) Given that there are k nodes waiting to gain access, how can a protocol be designed to minimize energy consumption and still achieve a certain percentage of successful contentions? 2) What happens to energy consumption and access success in these protocols if they are designed for k_1 nodes to gain access but there are actually k_2 nodes?

We conclude that the polling protocols are best suited for this type of network. The short access packets used in an alternating CP-CFP type protocol greatly reduce the time to poll all nodes in the network and the very predictable nature of polling makes it very easy for the PC to manage dozing. Additionally, the performance of polling protocols are not affected by the load and are least affected by hidden node effects. The random access protocols, however, remain options. They are still very effective when the load is small and in those cases have much better throughput than polling.

We start our presentation of our research in Section 4.2 which provides background on the environment in which we expect our protocols to work. It provides definitions that are used throughout the chapter and describes the physical properties of the network that constrain our designs. Section 4.3 develops the details and the models of seven different access protocols suitable for granting access while simultaneously conserving energy. Section 4.4 presents our analysis using these models. Finally, we conclude with Section 4.5.

4.2 Background

4.2.1 Centrally Controlled Data Transmission

Centrally controlled data transmission uses a point coordinator (PC) to manage traffic transmission. For all types of traffic, the PC first learns what traffic in the network needs to be transmitted in a contention period (CP) and then directs its transmission during a contention free period (CFP). The PC manages the alternation between these periods in an attempt to balance consideration of delay, throughput and energy consumption. In this chapter we focus on the CP only. We assume that at the beginning of the CP the PC transmits an announcement that the CP is starting and that it will last for some finite period. Nodes that do not need to contend can then enter the doze state for the entire CP. The remaining nodes then contend for access. In a single access event a node may inform the PC of one or more data packet transmissions or may attempt to obtain a connection with some quality of service (QoS) for an undetermined number of packets. The quantity of data that needs to be transmitted

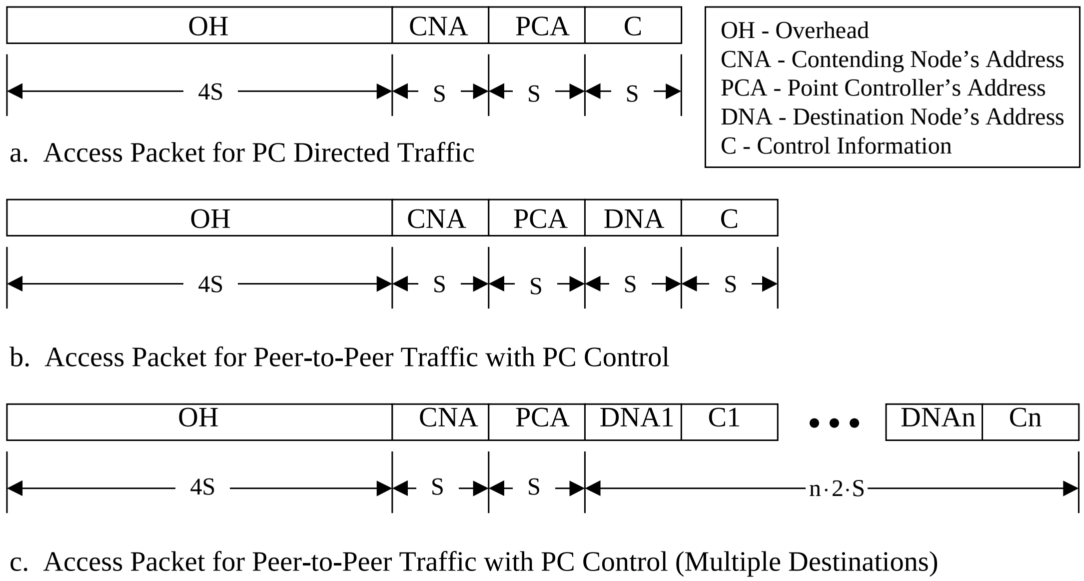


Figure 4.1: Access Packets

to a single destination does not affect the quantity of energy consumed to gain access in the CP.

4.2.2 Access Packets

Figure 4.1 illustrates different types of access packets. Data transmission in a centrally controlled network can be of two types. In the first, all packet transmissions are sent to the PC. In this case, the access protocol would use the fixed sized access packet illustrated in Figure 4.1a. In the second type, the network supports peer-to-peer communication allowing nodes to transmit packets directly to peer nodes in the network. In this case, the access protocol can either use fixed size packets of the type illustrated in Figure 4.1b and require the node to contend for each destination or use an expandable packet as illustrated in Figure 4.1c and contend just once. It is important to note that despite the destination of the data packet transmission, contention is for access to the PC only.

4.2.3 Physical Layer

These protocols work in the same environment that we described in Chapter 3 for the centrally controlled data transmission so we use the same assumptions about the physical layer as were described in Section 3.2.3. However, as illustrated in Figure 4.1, the packets that are exchanged are much shorter. Packet errors are rare. Additionally, the greater issue in contention is collisions. Therefore, we do not model errors nor hidden node effects in our analysis. We assume all nodes can receive each other's transmissions. However, late in Section 4.4 we describe the expected effects of these phenomena on each protocol's performance.

4.2.4 Comparison Measures

The objective of our analysis of different protocols is to compare their performance. We assume that at the beginning of the contention period that the protocols that are being compared have the same k nodes to contend and that no additional nodes arrive during the CP. The measures we seek for each protocol are the total network energy consumption, the number of successful contentions, and the duration of the contention period. Prior to comparing we seek the optimum parameters for each protocol. Since minimizing energy consumption and achieving maximum throughput may not be compatible goals, we specify the expected number of successful contentions as a constraint and choose the parameters for least network energy consumption. To compare the protocols we seek the same number of successful contentions and then compare the network energy consumption and the expected duration of the CP.

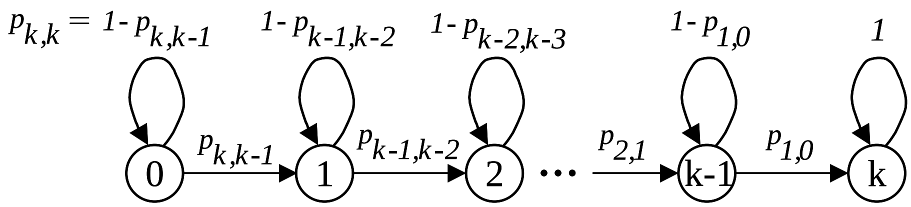


Figure 4.2: Markov chain of the p-Persistent Slotted Aloha access process

4.3 Access for PC Directed Data Packets

We consider seven protocols for access, p-persistent slotted aloha, time slotted p-persistent carrier sense multiple access (CSMA), elimination yield non-preemptive multiple access (EYNPMA) as found in HIPERLAN [7], polling, selective polling, modified random addressing protocol (RAP), and orthogonal addressing.

4.3.1 p-Persistent Slotted Aloha

In p-persistent slotted aloha, each of the *k* stations will contend in any given slot with some probability until it is successful contending. In our analysis we consider a contention successful if a single node contends in the slot. If two or more nodes contend in the same slot, there is a collision and none of the nodes succeed at gaining access. (We do not consider capture in our analysis.) We model two phenomena, the expected number of successful contentions and the expected energy consumption in the CP given a number of nodes contending, *k*, a probability that a node contends in a given slot, *p*, and a fixed number of slots, *n*. We define the transition probability that when there are *i* nodes contending in a slot that there are then *j* nodes that contend in the next slot as

$$p_{i,j} = \begin{cases} 1 - i \cdot p \cdot (1-p)^{i-1} & j = i \\ i \cdot p \cdot (1-p)^{i-1} & j = i-1 \\ 0 & Otherwise \end{cases} \tag{4.1}$$

The Markov chain illustrated in Figure 4.2 shows the state transition diagram associated with the process. The transition matrix is given by

$$\mathbf{P}=\begin{bmatrix} p_{k,k} & p_{k,k-1} & 0 & 0 & & 0 \\ 0 & p_{k-1,k-1} & p_{k-1,k-2} & 0 & \cdots & 0 \\ 0 & 0 & p_{k-2,k-2} & p_{k-2,k-3} & & 0 \\ & \vdots & & \ddots & & 0 \\ 0 & 0 & 0 & 0 & p_{1,1} & p_{1,0} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.2}$$

We define S_n as the number of successful contentions that occur in n slots. We can find the expected value of S_n using the transition matrix.

$$\mathbf{E}\left[S_n\right]=\sum_{i=1}^{k}\left[\mathbf{P}^n\right]_{0,i}\cdot i \tag{4.3}$$

We define the energy consumed by a node contending in a slot as $e_s = 17\tau_s$.[17] Therefore, the expected energy consumed by a node in the next slot conditioned on the fact it needs to contend is $p\cdot e_s$. If there are j nodes that need to contend then the total expected energy consumed in the next slot is $j\cdot p\cdot e_s$. Therefore the expected energy consumption by a network through n slots is

$$\mathbf{E}\left[E_n\right]=\sum_{j=1}^{n}\left[1 \quad 0 \quad \cdots \quad 0\right]\cdot\mathbf{P}^{j-1}\cdot\mathbf{d}\,, \tag{4.4}$$

where **d** is an energy consumption column vector: $\mathbf{d}_i=\left(k-i\right)\cdot p\cdot e_s \quad i\in 0,1,\cdots k$.

In Figures 4.3 and 4.4 we illustrate the expected number of successful contentions and the expected energy consumption for contention periods of five different lengths (i.e. 10, 50, 100, 150, and 200) while varying the contention probability, p. In

[17] There are 7 slots associated with the request, 7 slots associated with the acknowledgement, and 3 interframe slots.

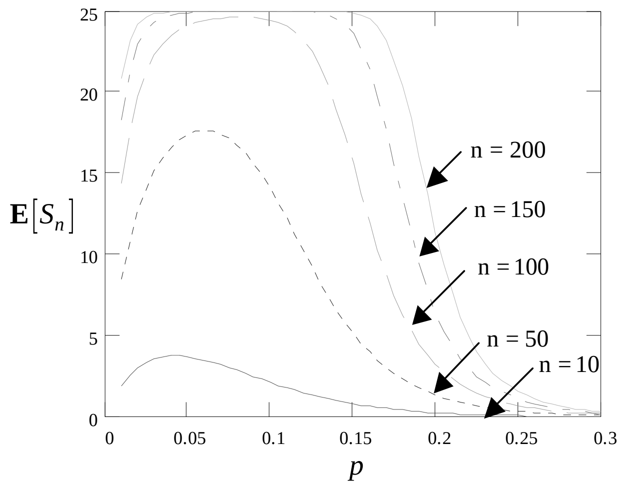


Figure 4.3: Expected number of successful contentions using slotted aloha in contention periods starting with 25 nodes, $k = 25$

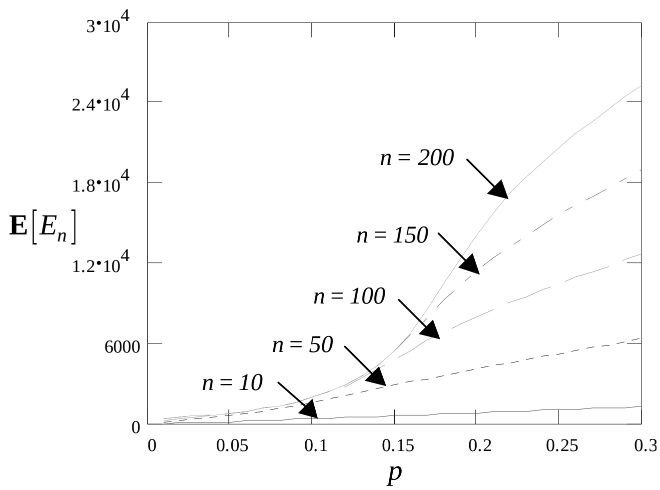


Figure 4.4: Expected energy consumed using slotted aloha in contention periods starting with 25 nodes, $k = 25$

Figure 4.3, we see that there is always a p that achieves the greatest expected number of successful contentions and that the expected number of successful contentions monotonically decreases from this point as p increases. In Figure 4.4, we see that energy consumption is monotonically increasing in p regardless of the size of the contention period and the expected number of successful contentions. This leads to a simple optimization strategy. Given a CP duration of n slots and a lower bound $\mathbf{E}[S_n]$ (i.e. the number of successful contentions) choose the smallest p that meets the bound so as to minimize energy consumption. If the threshold cannot be achieved, choose the p that yields the highest $\mathbf{E}[S_n]$.

4.3.2 Time Slotted p-Persistent Carrier Sense Multiple Access (CSMA)

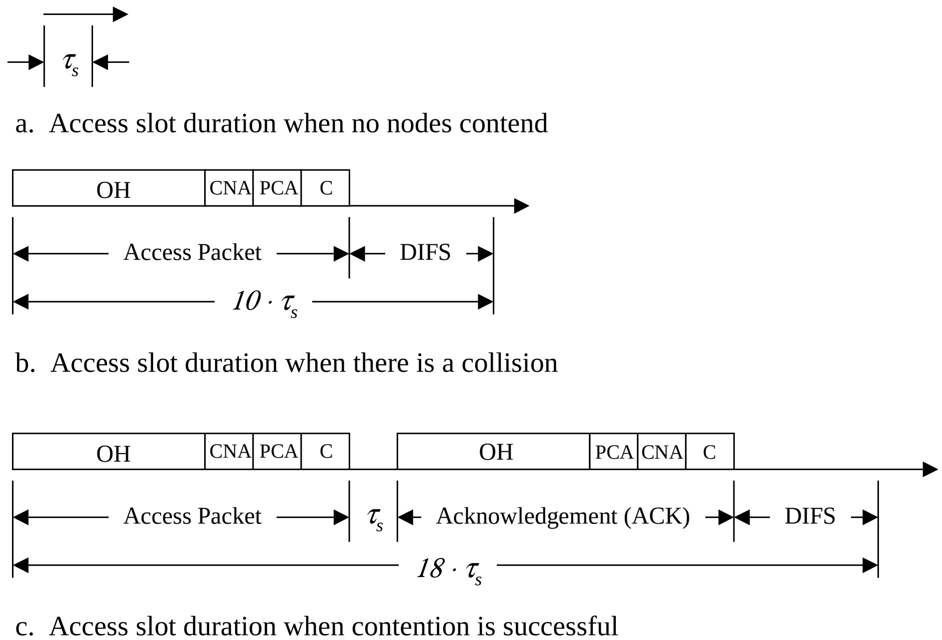
As mentioned earlier 1-persistent CSMA is not suitable for this type of network. It is necessary to distribute contention attempts during the CP in order to avoid collisions. These techniques are referred to as collision avoidance. Our collision avoidance scheme is a modification of the 802.11 collision avoidance scheme [6] that may be more accurately described as a p-persistent CSMA technique. We call our collision avoidance access scheme time slotted p-persistent CSMA. To describe this technique we will first describe the 802.11 collision avoidance scheme, and then will describe the modifications that convert the 802.11 scheme into our time slotted p-persistent CSMA.

In the 802.11 CSMA/CA scheme, before a node contends it first generates a random backoff interval to use in gaining access. This interval is measured in an integer number of slot times. It contends by first monitoring the channel in the CP until the channel is sensed idle for a specified period, in our case 3 slot times. At this point, the node begins to decrement its backoff counter as it senses the channel idle for additional slot times. If the node senses the channel busy at some time before the backoff expires, it stops decrementing the backoff counter. It then waits until the channel is sensed idle again for 3 slot times before continuing to decrement the backoff counter. If the backoff counter expires the node sends its access packet starting in the very next slot. At the conclusion of the transmission, the node waits for an acknowledgement (ACK). If one is received then the contention was successful and the node enters the doze state until the scheduled beginning of the CFP. Otherwise, it selects a new backoff interval and continues contending.

The standard procedure for selecting the backoff interval is to uniformly choose a backoff from a range (*0,w-1*) referred to as the contention window. The value of *w* depends on the transmission history of the node. In the IEEE 802.11 protocol there is a maximum and minimum *w*, CW_{min} and CW_{max} respectively. After each failed contention, *w* is increased according to $2^m CW_{min}$ where *m* is the number of consecutive collisions that have preceded the selection. This continues until *w* reaches CW_{max} at which time it is no longer increased for collisions. This method of increasing the contention window adapts to congestion.

The 802.11 CSMA/CA scheme is not appropriate for our applications since we assume congestion at the start of the CP and that this congestion can only decrease. A large contention window should be selected at the start. We keep the same contention window for the duration of the CP. We approximate this process in our modeling by specifying that all active nodes contend with probability *p* as done in the p-persistent slotted aloha protocol. Our analysis proceeds very much the same as that above for p-persistent slotted aloha. The distinction then between p-persistent slotted aloha and our time slotted p-persistent CSMA scheme is that the latter scheme has better throughput since a smaller slot is used when there is no contention attempt but consumes more energy since contending nodes must monitor the channel until they successfully contend.

Equations (4.1) and (4.2) remain valid for the time slotted p-persistent CSMA protocol except *n* is no longer the number of contention slots in the CP but the number of backoff slots. Unlike with p-persistent slotted aloha, the number of backoff slots in a fixed size CP can vary depending on the number of contentions that are attempted. We attempt to model the number of successful contentions and the energy



a. Access slot duration when no nodes contend

b. Access slot duration when there is a collision

c. Access slot duration when contention is successful

Figure 4.5: CSMA/CA Access Slot Duration for Different Access Events

consumed as a function of the CP size and to model the expected CP duration and energy consumption when the protocol uses n backoff slots. We start by defining the average time between backoff slots. There are three different times between backoff slots and are illustrated in Figure 4.5. When there are no contentions, the time between slots is τ_S, when there is a successful contention the time is $18 \cdot \tau_S$, and when there is a collision the time is $10 \cdot \tau_S$, or simply *1*, *18*, or *10* respectively in our equations. The probability of a successful contention is determined by Equation (4.1). The probability that no node contends is

$$p_{i,j;A} = \begin{cases} (1-p)^i & i = j \\ 0 & otherwise \end{cases}$$

where *A* is defined as the event that none of i nodes contend. Finally, the probability that a collision occurs is

$$p_{i,j;B} = \begin{cases} 1 - p_{i,j;B} - p_{i,i-1} & i = j \\ 0 & otherwise \end{cases}$$

where B is defined as the event that a collision occurs amongst the i nodes. Note that the probability that there is no transition in a particular contention slot with i nodes contending is

$$p_{i,i} = p_{i,i;A} + p_{i,i;B}$$

Using these probabilities we define a delay vector for which each entry corresponds to the expected delay when k nodes are remaining to gain access.

$$\mathbf{d} = \begin{bmatrix} p_{k,k;A} + 10 \cdot p_{k,k;B} + 18 \cdot p_{k,k-1} \\ p_{k-1,k-1;A} + 10 \cdot p_{k-1,k-1;B} + 18 \cdot p_{k-1,k-2} \\ \vdots \\ p_{1,1;A} + 10 \cdot p_{1,1;B} + 18 \cdot p_{1,0} \\ 1 \end{bmatrix} \tag{4.5}$$

With Equations (4.2) and (4.5) we can define the expected time it takes to backoff n slots as

$$\mathbf{E}\left[T_{BO}(n)\right] = \sum_{j=1}^{n} \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \cdot \mathbf{P}^{j-1} \cdot \mathbf{d}$$

We then approximate the expected number of backoffs that will occur in a window of size τ_{CP} as

$$\bar{n} \approx \max_{n} \left\{ \mathbf{E}\left[T_{BO}(n)\right] < \tau_{CP} \right\}$$

Then we can find the expected number of successful contentions with

$$\mathbf{E}\left[S_{\tau_{CP}}\right] \cong \sum_{i=1}^{k} \left[\mathbf{P}^{\bar{n}}\right]_{0,i} i \,.$$

The expected energy consumption in n backoff slots is

$$\mathbf{E}\left[E_n\right]=\sum_{j=1}^{\bar{n}}\left[1 \quad 0 \quad \cdots \quad 0\right]\cdot\left[\mathbf{P}^{j-1}\right]\cdot\begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k-1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}\cdot\mathbf{d}. \qquad (4.6)$$

where Equation (4.6) differs from Equation (4.4) in that the addition of a diagonal matrix accounts for the fact that all nodes with traffic to send stay awake until they successfully contend.

In Figures 4.6 and 4.7 we illustrate the approximate values for the expected number of successful contentions and the expected energy consumption for contention periods of five different lengths (i.e. 10, 50, 100, 150, and 200 p-persistent slotted aloha slots[18]) while varying the contention probability, *p*. The choice of this size contention period allows Figures 4.6 and 4.7 to be directly compared to Figures 4.3 and 4.4. As can be seen the time slotted p-persistent CSMA protocol consumes much more energy than the p-persistent slotted aloha protocol but unlike the p-persistent slotted aloha protocol, the expected energy consumption does not monotonically increase with p. Rather, as illustrated in Figure 4.8, there is a clear optimum *p*. The optimization strategy for this protocol depends on the starting parameters. Given a CP duration of *n* slotted aloha slots and a threshold $\mathbf{E}\left[S_{\tau_{CP}}\right]$, if the threshold $\mathbf{E}\left[S_{\tau_{CP}}\right]$ can be achieved choose the *p* that results in least energy consumption. If the threshold cannot be achieved, choose the *p* that yields the highest $\mathbf{E}\left[S_{\tau_{CP}}\right]$. Of interest, note that when the CP is sufficiently large the optimum energy consumption is not affected by the CP size. (See Figure 4.8.) This indicates that the strategy to choose the optimum set of

[18] For the sake of doing a comparison we use time units defined by slotted aloha slots. We convert this time unit to a number of backoff events using Equation (10).

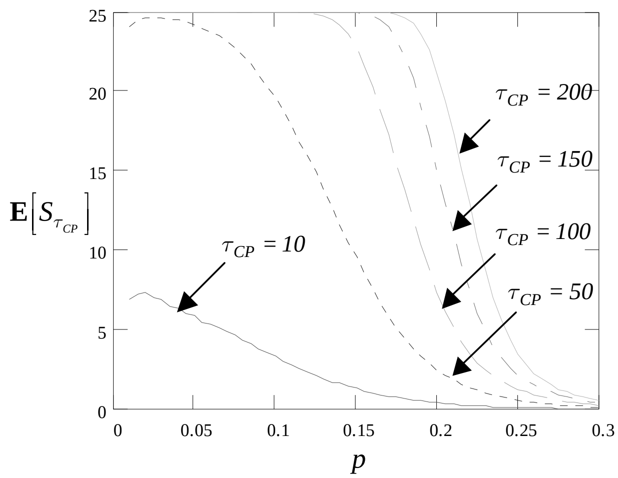


Figure 4.6: Approximation of the expected number of successful contentions using time slotted p-persistant CSMA in contention periods starting with 25 contending nodes, $k = 25$, as in Figure 4.3 (τ_{CP} is measured in equivalent slotted aloha contention slots)

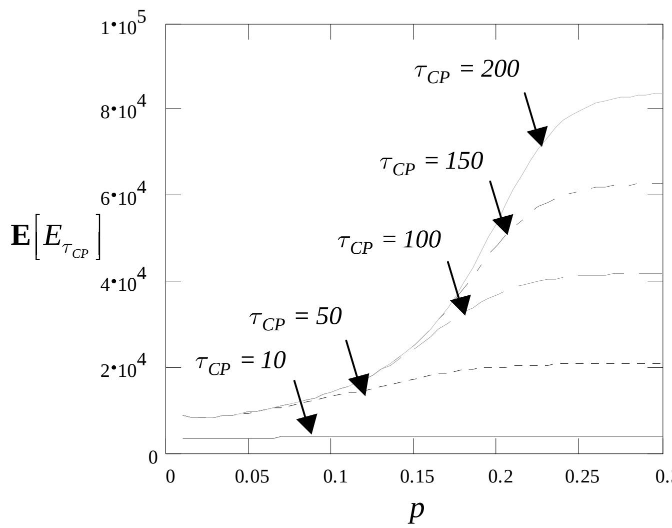


Figure 4.7: Approximation of the expected energy consumed in a contention period using time slotted p-persistent CSMA starting with 25 nodes, $k = 25$ as in Figure 4.4 (τ_{CP} is measured in equivalent slotted aloha contention slots.)

parameters is to choose the smallest CP size that can achieve the threshold $\mathbf{E}\left[S_{\tau_{CP}}\right]$. Optimizing under this strategy is identical to that used for p-persistent slotted aloha.

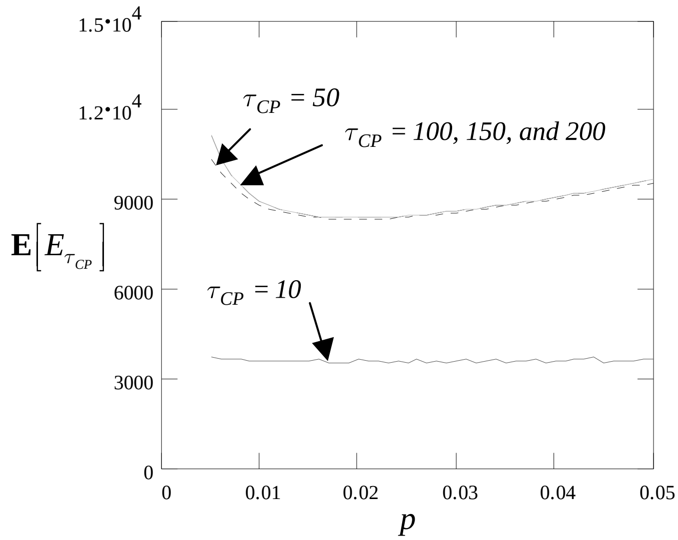


Figure 4.8: Zoomed-in view of Figure 4.7 where the optimum p is found

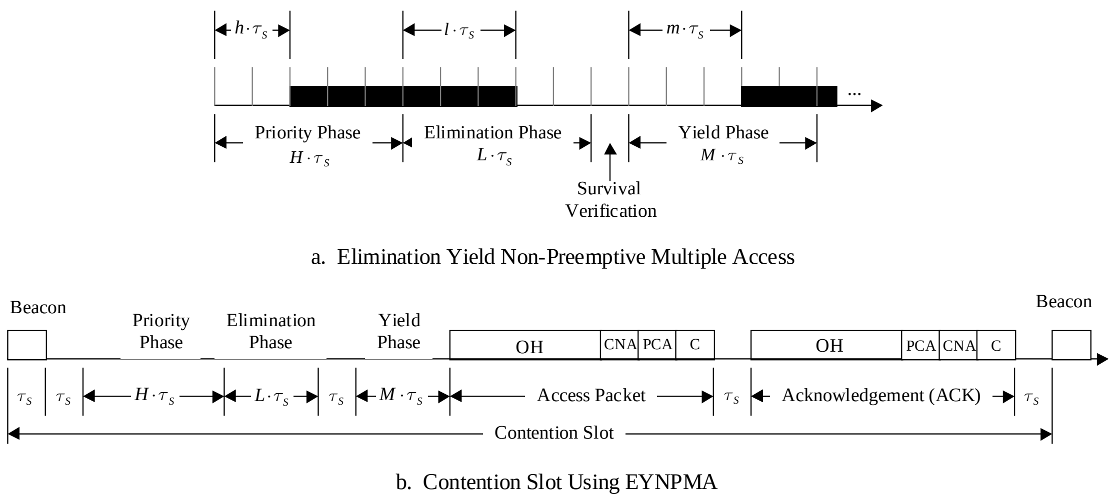


Figure 4.9: Modified Elimination Yield Non Preemptive Multiple Access

4.3.3 Elimination Yield Non Preemptive Multiple Access (EYNPMA)

EYNPMA is a protocol in the class described as collision avoidance resolution trees in [26-30]. These protocols have the benefit of increasing the probability of access in each contention slot at the expense of more overhead. EYNPMA is specifically used by HIPERLAN. [7] We describe it here as it was intended to be implemented according to the standard and then we modify it to suit the requirements of our application. According to the HIPERLAN standard, nodes wait until they can sense the channel is idle for 1700 bit-periods before attempting access. The subsequent channel access attempts occur in four phases, prioritization, elimination, yield, and finally transmission phases. Figure 4.9a illustrates the access cycle. Nodes gain access by sending a signal and by when it starts to send the packet. The signal begins in the priority phase and ends in the elimination phase, and the packet transmission starts in the yield phase. A node wins the contention by being among the first to start transmitting in the first phase, among the last to stop transmitting in the second phase and the first to start transmitting a packet in the third phase. Nodes that recognize that they have lost the contention in any one of the phases will defer from attempting to gain access.

Note that the prioritization, elimination, and yield phases are all slotted. These three phases have H, L, and M slots respectively. In HIPERLAN, each access transmission is assigned a priority level h which is numbered from 0 to $H\text{-}1$ with 0 being the highest priority. The priority level is assigned to a transmission based on its residual lifetime. A node with a transmission of priority level h will transmit an access burst for the duration of the priority phase starting in the $h+1^{\text{th}}$ slot provided the channel is sensed idle up until that time. Contending nodes that hear a burst before the designated priority slot for their transmission will stop contending. Each node that transmits an access burst is granted admission into the elimination phase. These nodes continue to transmit their access burst into the elimination phase with probability q of transmitting in each next elimination slot until it does not transmit or the L slots have been used. At this point the node listens. If it hears another node still transmitting then it loses the contention and returns to the doze state. Otherwise it listens through the remaining slots of the elimination phase, through a survival verification interval, and into the yield phase. A node surviving the elimination phase listens to each slot of the yield phase with probability p and begins to transmit the packet with probability (1-p) except for the last slot when the transmission probability is 1. If a node senses that another node begins to transmit a packet then it stops contending. Performance models for access using the HIPERLAN access protocol are found in [31]. The values specified for H, L, M, q, and p in HIPERLAN are 5, 12, 14, 0.5, and 0.9 respectively.

We modify EYNPMA for our application in the following manner. First, contentions occur in slots each large enough to handle all phases of the contention. These slots are initiated by a beacon transmitted by the PC so nodes do not have to wait for 1700 bits of idle time. (See Figure 9b.) The second modification in our application, since we do not define a priority for access, is to use the priority phase as one additional level in the collision avoidance resolution tree. Similar to the elimination phase each contending node listens in each slot of the priority phase with probability r and begins to transmit the access burst in a slot with probability (1-r) except for the last slot when the transmission probability is 1.

The slots in which a single independent node transmits in the priority phase, B, stop transmitting in the elimination phase, C, and start in the yield phase, D, are random variables with truncated geometric distributions as follows[19]:

$$Pr\left(B=b\right)=\begin{cases}(1-r)r^{b-1} & 1\le b<h \\ r^{h-1} & b=h\end{cases};$$

$$Pr\left(C=c\right)=\begin{cases}(1-q)q^{c-1} & 1\le c<l \\ q^{l-1} & b=l\end{cases};$$

$$Pr\left(D=d\right)=\begin{cases}(1-p)p^{b-1} & 1\le d<m \\ p^{m-1} & b=m\end{cases}.$$

When k nodes contend for transmission they interact by listening to each other during each phase reducing the number of the original contenders surviving each phase. In order to compute the probability that a single node survives the last phase, we make several intermediate calculations. We denote the number of nodes surviving each of the three phases by V, W, and X, whose distributions are given by:

$$Pr\left(V=v\middle|k\right)=\binom{k}{v}\sum_{b=1}^{h}Pr\left(B=b\right)^{v}Pr\left(B>b\right)^{k-v}\quad 0<v\le k\quad ;$$

$$Pr\left(W=w\middle|V=v\right)=\binom{v}{w}\sum_{c=1}^{l}Pr\left(C=c\right)^{w}Pr\left(C<c\right)^{v-w}\quad 0<w\le v\quad ;$$

$$Pr\left(X=x\middle|W=w\right)=\binom{w}{x}\sum_{d=1}^{m}Pr\left(D=d\right)^{x}Pr\left(D>d\right)^{w-x}\quad 0<x\le w\quad .$$

The probability that only a single node survives when there are k contenders is given by

$$Pr\left(X=1\middle|k\right)=\sum_{v=1}^{k}\sum_{w=1}^{v}Pr\left(X=1\middle|W=w\right)\cdot Pr\left(W=w\middle|V=v\right)\cdot Pr\left(V=v\middle|k\right)\ .\quad (4.7)$$

This simple model provides the necessary intuition to guide the selection of the signaling parameters. With this we define our transition probabilities

$$p_{i,j}=\begin{cases}Pr\left(X=1\middle|i\right) & j=i-1\\ 1-Pr\left(X=1\middle|i\right) & j=i\\ 0 & otherwise\end{cases}\quad (4.8)$$

We can then build the transition matrix in (4.2) and can use equation (4.3) to determine the expected number of successful contentions. We now define the expected energy consumed in a slot when k nodes are contending.

[19] This analysis is similar to that found in [31].

$$\mathbf{E}\left[E\left(k\right)\right]=\begin{cases}\sum_{v=1}^{k}\sum_{h=1}^{H}\binom{k}{v}\cdot\left(Pr\left(B=h\right)\right)^{v}\left(Pr\left(B>h\right)\right)^{k-v}\left(\left(v\cdot H\right)+\left(j-v\right)\left(h+4\right)\right)+\\ \sum_{v=1}^{k}Pr\left(V=v\right)\sum_{w=1}^{v}\sum_{l=1}^{L}\binom{v}{w}\left(Pr\left(C=l\right)\right)^{w}\left(\left(Pr\left(C<l\right)\right)^{v-w}\cdot w\cdot\left(l+1\right)+\sum_{i=1}^{l-1}Pr\left(C=i\right)\cdot\left(v-w\right)\cdot\left(i+4\right)\right)+\\ \sum_{v=1}^{k}\sum_{w=1}^{v}Pr\left(W=w\middle|V=v\right)Pr\left(V=v\right)\sum_{x=1}^{w}\sum_{m=1}^{M}\binom{w}{x}\left(Pr\left(D=m\right)\right)^{x}\left(\left(Pr\left(D>m\right)\right)^{w-x}\left(x\cdot\left(X+ap\right)+\left(w-x\right)\left(m+ap\right)\right)\right)\end{cases}\tag{4.9}$$

This equation assumes that once nodes identify that they fail in a contention, they enter the doze state and do not wake-up until the next contention slot. The constant *ap* is the duration of one of the contention slots as illustrated in Figure 4.9b. less the time for the collision avoidance tree $\left(H+L+M+1\right)$ plus a slot for the time to wake-up. We can now determine the expected energy consumed in a contention period using Equation (4.4) where the transition matrix is built with the transition probabilities defined by Equation (4.8) and where we define the vector **d** with values determined using Equation (4.9), $\mathbf{d}_i=\mathbf{E}\left[E\left(k-i\right)\right]$.

Identifying the parameters for best energy conservation is a much more complex problem than that for the p-persistent slotted aloha and the time slotted p-persistent CSMA protocols since it is an optimization problem with 6 variables, *r*, *p*, *q*, *H*, *L*, and *M* rather than just 1 variable. We applied a brute force search for best energy consumption varying *H*, *L*, and *M* in increments of 1 and varying *r*, *p*, and *q* by 0.01 with a *k* slot access period for 6 different values of *k*. Note that since there is at least one node that finally transmits its access packet in each slot minimizing energy consumption also reduces collisions since the energy consumption penalty for a collision is so great. Table 4.1 illustrates the optimum parameters that we found and the expected number of successes in a *k* contention slot CP, $\mathbf{E}\left[S_k\right]$. Of interest in this table is that the size of each phase is much smaller than those recommended by the HIPERLAN standard and that it is the priority phase that varies the most in size. This phenomenon is clearly understood. It is advantageous to reduce the number of

k	H	L	M	r	p	q	$\mathbf{E}[S_k]$
5	4	3	2	0.77	0.58	0.52	4.699
10	5	4	2	0.86	0.61	0.52	9.432
15	6	4	2	0.90	0.60	0.52	14.106
20	6	4	2	0.92	0.58	0.53	18.694
25	6	5	2	0.93	0.63	0.52	23.552
50	9	5	2	0.96	0.64	0.52	47.031

Table 4.1: Results of a brute force search of optimum parameters for EYNPMA based on number of contending nodes, *k*

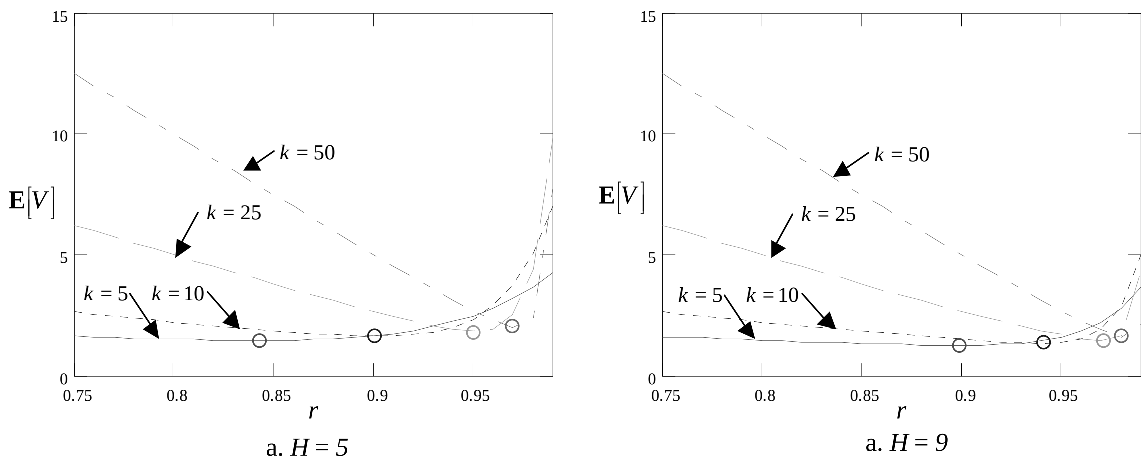


Figure 4.10: Expected number of surviving nodes after the priority phase

contending nodes in a contention slot as soon as possible. When *k* is large, a larger number of slots, *H*, with a large listening probability, *r*, results in a higher probability that when the first access burst starts a smaller number of nodes will be transmitting. Thus, a larger number of nodes can enter the doze state early in the contention slot. We also note that the multiple layers have a leveling affect where the number of nodes surviving each layer has less variance. For example we see in Figure 4.10 the plot of the expected number of survivors, $\mathbf{E}[V]$, for different values of *r* for two different priority phase sizes. We see that as the size of *H* increases that the optimum *r* increases and that the range of optimum *r*'s (designated by circles) decreases despite the large difference in the *k* values. At the settings found in Table 4.1 the expected number of survivors of the priority phase is consistently less than 3. We found that increasing the number of contention slots in the contention period from those specified in Table 4.1 has an insignificant effect on the optimum parameter values chosen for the protocol.

As we see in the results in Table 4.1, our variation of EYNPMA is especially effective at gaining access at the optimum energy consumption parameters. As a result the selection of the optimum set of parameters is not constrained by the length of the contention period but rather the threshold $\mathbf{E}\left[S\right]$. It is possible to choose a near optimum set of parameters for a protocol by selecting a contention period equivalent to the number of nodes that begin contending and searching for the best set of parameters exhaustively. Table 4.1 can be used to constrain the search space. Once these parameters are found the last parameter, number of slots required, can be selected such that the threshold $\mathbf{E}\left[S_n\right]$ is achieved.

4.3.4 Polling

In the standard polling strategy nodes do not contend with each other. Rather, the PC polls nodes based on its knowledge of the number of nodes in the network that could potentially have traffic to send.[20] Energy consumption in this network is no longer based on the number of nodes contending but rather on the size of the network and the status of each node. The assumption the PC must make is that every node is a potential contender and should be polled. Polling strategies to conserve energy would follow those guidelines found in Chapter 3. A traffic indication map (TIM) indicating

[20] Another contention period would be required for new nodes to associate with the network.

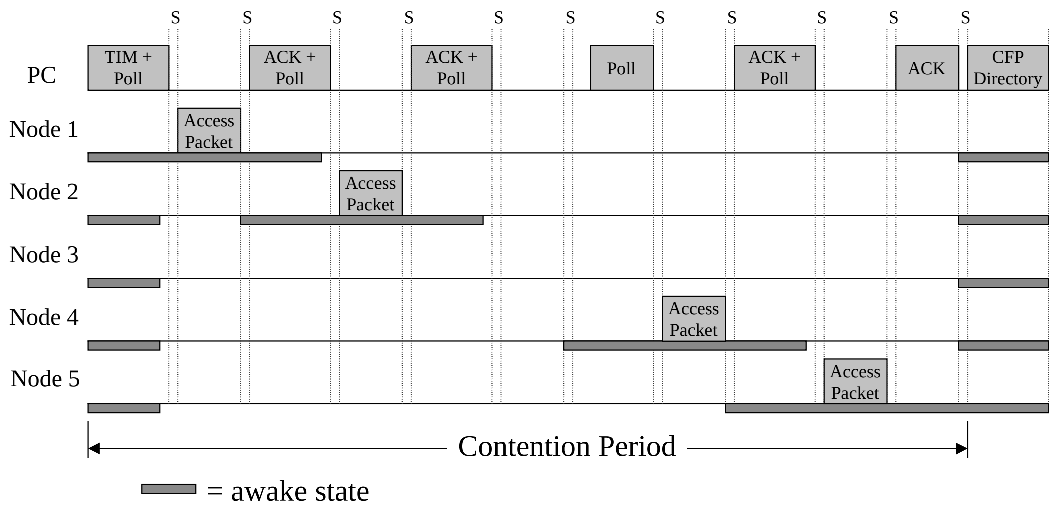


Figure 4.11: The polling access process

each node that would be polled is broadcast.[21] This TIM provides each contending node information on which nodes would be polled and the position in the bitmap tells when since we will assume the PC polls the active nodes in bitmap order. Polls would occur at a regular interval, each interval being large enough to accommodate a poll, an access packet, and an acknowledgement. With these assumptions each node can determine when it would be polled and can doze until that time. Energy consumption by mobile nodes would be limited to that consumed by each contending node listening to the polling directory and the time it is awake in its polling slot. Energy wise this is extremely efficient achieving a constant rate of energy consumption per node gaining access but throughput-wise it may be an inefficient scheme, especially if the network is large and only a few nodes will in fact contend during a given polling round.

[21] Some nodes may not be polled. For example some protocols allow nodes to enter an extended doze state on their own initiative by announcing that intent to the PC. A node in this state would not be polled. Additionally some nodes may have been assigned a position in the bit map of the TIM but then subsequently left the network. These nodes would not be polled. Finally, a protocol may not poll nodes that are active already in the CFP.

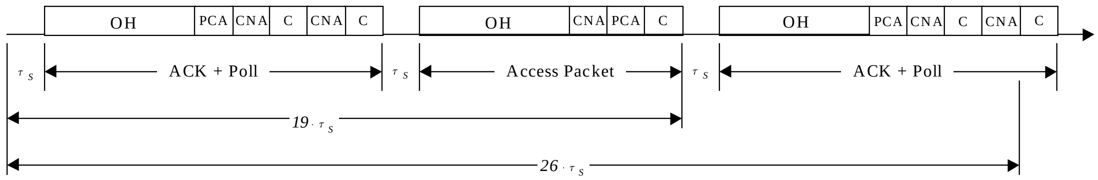


Figure 4.12: Detailed illustration of an access slot

Figure 4.11 illustrates the operation of the polling process. All nodes in the network listen to the first TIM at the start of the CP to determine when they would be polled if they have traffic to send or to determine the length of the CP so they can wake in time for the first directory of the CFP. In the illustration only Node 3 has no traffic, so it dozes until the completion of the CP and the time that was reserved for it to send an access packet goes unused. Note that nodes with traffic to send awaken prior to their access slot and then return to the doze state when the next node is polled. A node is awake for at most the polling slot and the directory announcing the CP.

Figure 4.12 illustrates the detail of an access slot providing information on its duration and the time a node spends awake for each slot. The length of a CP using polling is a function of the number of nodes in the network, *n*, which is also the number of access slots in the CP, is

$$\tau_{CP} = 8 \cdot \tau_S + n \cdot 19 \cdot \tau_S \, , \tag{4.10}$$

The number of nodes in the network affects the size of the TIM. In equation (4.10) and all subsequent equations we simplify our results by assuming a TIM will accommodate 96 nodes. The energy consumed by the network is also a function of both the number of nodes in the network and the number of nodes that have traffic to send, *k*.

$$e_{CP} = k \cdot 10 \cdot \tau_S + k \cdot 26 \cdot \tau_S - 9 \cdot \tau_S \, . \tag{4.11}$$

The final negative term in equation (4.11) accounts for the end conditions of the contention period.

The Polling protocol has a fixed result with no optimization parameters. The CP duration is fixed and energy consumption is only dependent on the number of nodes contending.

4.3.5 Selective Polling

In selective polling the CP is split into two parts. In the first, abbreviated polling is used to identify the nodes that have traffic and then in the second part only those nodes with traffic are polled. The objective of this method of access is to achieve the energy efficiency of polling with an improved throughput. In the abbreviated polling the PC allocates a number of short access slots during which nodes with traffic transmit a short access burst. One slot is assigned to each node in the network. The PC then knows which nodes to poll by detecting in which slots there were bursts. We assume those nodes that are contending remain awake for all the access slots. The length and the total energy consumption of the contention period are then:

$$\tau_{CP} = 8 \cdot \tau_S + (n+2) \cdot \tau_S + k \cdot 19 \cdot \tau_S \,,$$

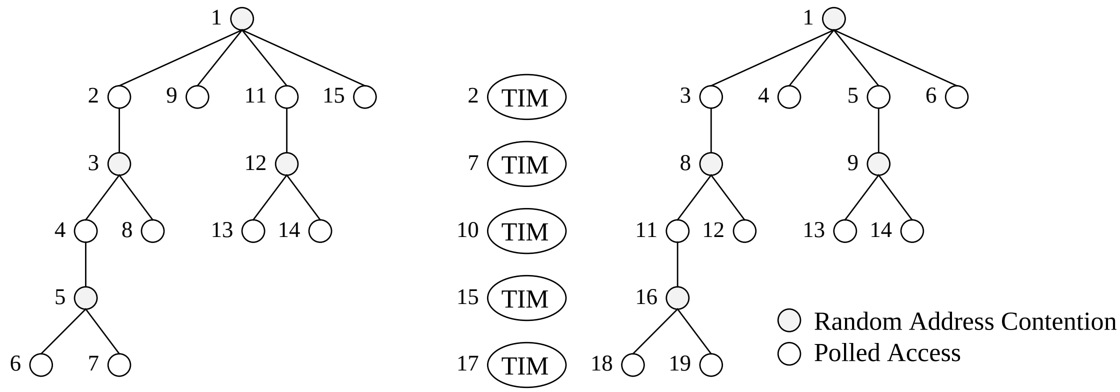
and

$$e_{CP} = k \cdot 10 \cdot \tau_S + k \cdot (n+26) \cdot \tau_S - 9 \cdot \tau_S \,.$$

Selective polling's performance is also deterministic where the CP length and the energy consumption is linearly related to the number of nodes contending.

4.3.6 Modified Random Addressing Protocol (RAP)

Modified RAP is a polling protocol that uses RAP [32] to identify which nodes have traffic to send. We assume that the PC is capable of simultaneously detecting some finite number of orthogonal signals which are referred to as addresses. At the beginning of the contention period, all contending nodes transmit an access burst using one of these addresses that each has selected randomly. The PC then polls all the addresses it has detected. In the case that two or more nodes have selected the



a. Random Address Polling (RAP)   b. Modified RAP for Energy Conservation

Figure 4.13: Access using random addressing

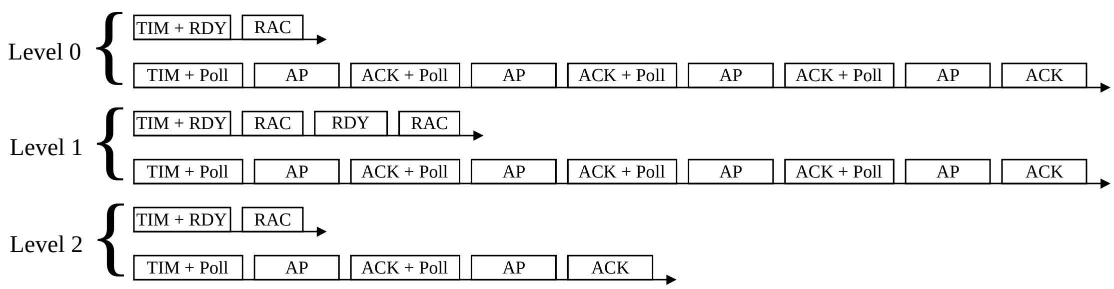


Figure 4.14: Contention period transmissions that correspond to Figure 4.13b

same address, then those nodes will collide when polled. The PC resolves this collision by repeating the process described above. As originally envisioned, RAP proceeds down a tree-like structure until there are no more collisions. This is illustrated in Figure 4.13a. Each node in this tree corresponds to an exchange with the PC. In the beginning the PC directs a random address contention. In this contention the PC distinguishes four addresses and then polls the first. The response to the poll is a collision so the PC then directs a random address contention for those nodes that collided. The protocol proceeds through the tree following the left most edge. This protocol is not conducive to energy conservation since nodes have no way of predicting when they will be polled so they cannot doze until they have successfully contended. In Figure 4.13b we show how we modify this approach to provide a more predictable wake-up schedule for the nodes. Rather than proceeding down the left edge of the tree the protocol traverses each layer first. In this figure after the first random address period the PC transmits a TIM with a schedule for four polling periods. As illustrated, two result in successful contentions and two result in collisions. This results in a second TIM that then schedules two random addressing periods. Each of the random addressing periods result in two distinguishable addresses so the PC then transmits a third TIM this time scheduling 4 polling periods. And the protocol proceeds as illustrated. This modification supports energy conservation since the information in the TIMs allow nodes to doze until their polling or random address contention periods. Figure 4.14 illustrates the events associated with each level of Figure 4.13b.

We modeled and analyzed this protocol bootstrapping off a simulation. Although presumably infinitesimal, there is always the possibility that the tree will continue to experience collisions ad infinitum. Simulation of the tree splitting process results in a finite set of attempts. We define two different probabilities that we obtain from the simulation. The first is for the number of surviving nodes at a particular level. We define $ps\left(j \middle| k,l,L\right)$ as the probability that there are j survivors after L levels of contention when starting with k nodes contending at the beginning, and l random addresses. In Figure 4.13b the CP starts with 7 nodes contending. There are 5 survivors at level 1, two survivors at level 2, and no survivors at level 3. The number of addresses is not specified but there were at least four in our example. The second probability we define is for the number of random address contentions at a particular level. We define $pc\left(j \middle| k,l,L\right)$ as the probability that there are j contentions at the L^{th} level of contention when starting with k nodes contending at the beginning, and l random addresses. In Figure 4.13b there is one contention at level 0, two contentions at level 1, one contention at level 2, and no contentions at level 3. We use these probabilities to determine three intermediate probabilities and expectations. First,

$$\mathbf{E}\left(TIM\left|k,l,L\right.\right)=\sum_{j=1}^{\left\lceil k/2\right\rceil} pc\left(j\left|k,l,L\right.\right)+\sum_{j=1}^{k} ps\left(j\left|k,l,L\right.\right),$$

is the expected number of TIMs that are required at level L when the CP starts with k nodes contending and the network uses L random addresses. The second,

$$\mathbf{E}\left(RAC\left|k,l,L\right.\right)=\sum_{j=1}^{\left\lceil k/2\right\rceil} pc\left(j\left|k,l,L\right.\right)\cdot j\ ,$$

is the expected number of random address contentions at level L. The third,

$$\mathbf{E}\left(Poll\left|k,l,L\right.\right)=\sum_{j=1}^{k} ps\left(j\left|k,l,L\right.\right)\cdot j-\sum_{j=1}^{k} ps\left(j\left|k,l,L+1\right.\right)\cdot j+\sum_{j=1}^{\left\lceil k/2\right\rceil} pc\left(j\left|k,l,L+1\right.\right)\cdot j\ ,$$

is the expected number of polls at level L. We use these expectations to determine the expected duration of the CP.

$$\mathbf{E}\left(\tau_{CP}\left|k,l\right.\right)=\sum_{L}\left(\begin{array}{l}\mathbf{E}\left(TIM\left|k,l,L\right.\right)\cdot\tau_{TIM}+\mathbf{E}\left(RAC\left|k,l,L\right.\right)\cdot\tau_{RAC}+\\ \mathbf{E}\left(Poll\left|k,l,L\right.\right)\cdot\tau_{AP}+\sum_{j=1}^{k} ps\left(j\left|k,l,L\right.\right)\cdot\tau_{OH}\end{array}\right) \tag{4.12}$$

where $\tau_{TIM}=2\cdot\tau_S$, $\tau_{RAC}=13\cdot\tau_S$, $\tau_{AP}=20\cdot\tau_S$, and $\tau_{OH}=4\cdot\tau_S$ are the lengths of time for a TIM to be transmitted, for the ready command and RAC to be executed, for a poll and access packet transmission, and for the additional overhead associated with the last ACK of the access periods at each level. The expected energy consumption is dependent only on the number of survivors at each level since each survivor at each level will monitor two TIMs and participate in a RAC and an AP. We define the expected energy consumption as:

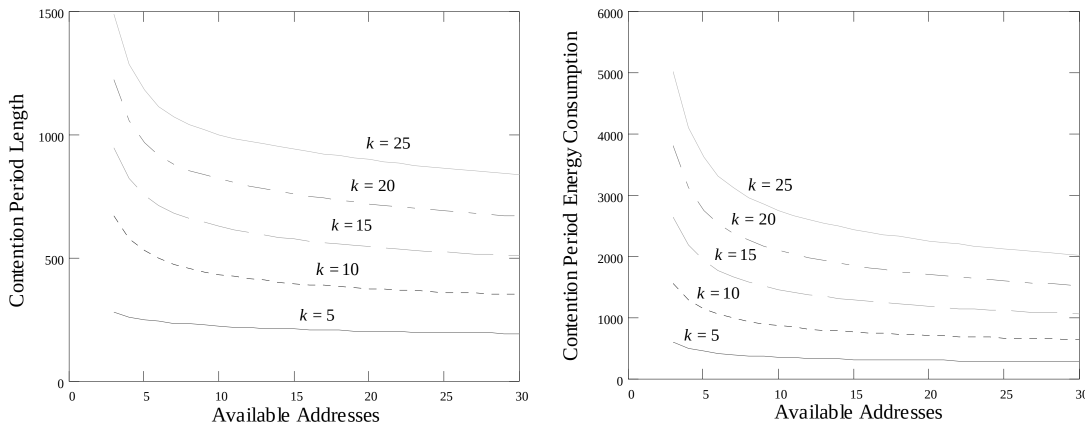


Figure 4.15: Contention period duration and energy consumption of energy conserving RAP

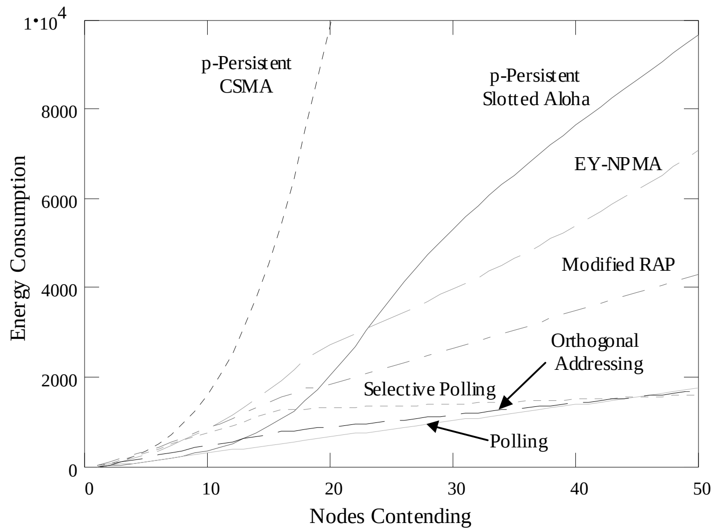
$$\mathbf{E}\left(e_{CP}\left|k,l\right.\right)=\sum_{L}\left(\sum_{j=1}^{k}ps\left(j\left|k,l,L\right.\right)\cdot j\cdot\left(2\cdot\left(\tau_{TIM}+6\cdot\tau_{S}\right)+\left(\tau_{RAC}+\tau_{S}\right)+\left(\tau_{AP}+\tau_{S}\right)\right)\right). \tag{4.13}$$

We obtained statistics on the splitting process for each (k,l) combination using 3 x 10^4 simulations and then determined the energy consumption using Equations (4.12) and (4.13). These results are shown in Figure 4.15. These results show that the performance of the protocol improves with the increase in the number of addresses that are available. As a possible optimization strategy, the protocol should use the maximum number of random addresses that the PC can support.

4.3.7 Orthogonal Addressing

If the PC can distinguish a sufficient number of addresses it may be feasible to assign each node in the network a unique address. This would allow accesses to occur in one level of contention. The CP would consist of just one RAC and a polled access slot for each node that requires access. The execution of this type of protocol would provide exceptional throughput. The length and the total energy consumption of the contention period are then:

$$\tau_{CP}=8\cdot\tau_{S}+16\cdot\tau_{S}+k\cdot19\cdot\tau_{S},$$



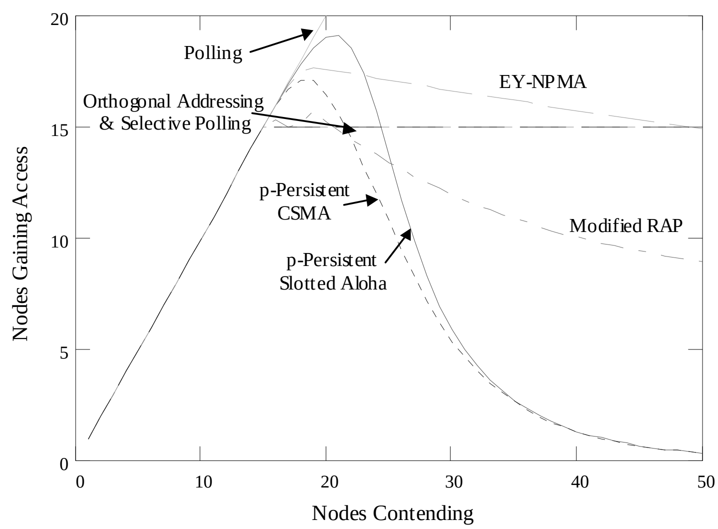


Figure 4.16: Effect of the number of nodes contending on the performance of protocols optimized for 15 contending nodes

and

$$e_{CP} = k \cdot 10 \cdot \tau_S + k \cdot (17 + 26) \cdot \tau_S - 9 \cdot \tau_S \,.$$

There are no parameters that can be adjusted to affect performance.

4.4 Protocol Comparison

Each protocol has three performance measures, the expected number of successful contentions, **E**[19], the expected energy consumption, e_{CP} or $\mathbf{E}[E_{CP}]$, and the expected duration of the contention period, τ_{CP} or $\mathbf{E}[\tau_{CP}]$. We compare the protocols presented in Section 3 by comparing the optimum performance of each for different numbers of contending nodes. We seek the parameters of each protocol that achieve the least energy consumption while simultaneously achieving some threshold expected number of successful contentions. In Table 4.2 we set the threshold for the expected number of successful contentions to be 99.9% of the number of nodes that contend and determine the results for when 5, 10, 15, 20, and 25 nodes contend. Since the PC may select the set of parameters for the protocol based on the expected number of nodes that will contend we check the robustness of each protocol when the number of nodes contending differs from the expected number of nodes contending. In Figure 4.16 we compare the performance of the protocols when the parameters selected for each protocol are those for optimum performance when 15 nodes contend. These can be taken directly from Table 4.2. The major constraint on the protocol is duration of the CP.

We find that the most robust protocols are the polling protocols. They are always the most energy efficient and will achieve maximum throughput when the number of nodes contending increases. Changing the number of nodes contending has no affect on the length of the contention period and the designed throughput will always be achieved.[22] At the other end of the performance spectrum is Time Slotted p-Persistent CSMA which has worst energy performance and is very sensitive to increasing numbers of nodes contending. We also note that the relative performance of

[22] The number of nodes in the network determines the length of the CP.

PROTOCOL	SIZE	PARAMETERS	**E**[τ_{CP}]	**E**[E_{CP}]
p-Persistent Slotted Aloha	5	$n = 30, p = 0.32$	480	213
	10	$n = 55, p = 0.193$	880	536
	15	$n = 81, p = 0.141$	1296	909
	20	$n = 106, p = 0.111$	1696	1296
	25	$n = 131, p = 0.091$	2096	1674
Time Slotted p-persistent CSMA	5	$n = 30, p = 0.32$	227	430
	10	$n = 55, p = 0.193$	515	1917
	15	$n = 81, p = 0.141$	813	4553
	20	$n = 106, p = 0.111$	1102	8330
	25	$n = 131, p = 0.091$	1387	13190
EYNPMA	5	$n = 8, (H, L, M, r, p, q) = (4, 3, 2, 0.77, 0.58, 0.52)$	224	262
	10	$n = 13, (H, L, M, r, p, q) = (5, 4, 2, 0.86, 0.61, 0.52)$	390	848
	15	$n = 19, (H, L, M, r, p, q) = (6, 4, 2, 0.90, 0.60, 0.52)$	589	1736
	20	$n = 25, (H, L, M, r, p, q) = (6, 4, 2, 0.92, 0.58, 0.53)$	775	2925
	25	$n = 30, (H, L, M, r, p, q) = (6, 5, 2, 0.93, 0.63, 0.52)$	960	4371
Polling	5	None (Assumes a 50 node network)	958	171
	10	None (Assumes a 50 node network)	958	351
	15	None (Assumes a 50 node network)	958	531
	20	None (Assumes a 50 node network)	958	711
	25	None (Assumes a 50 node network)	958	891
Selective Polling	5	None (Assumes a 50 node network)	155	421
	10	None (Assumes a 50 node network)	250	851
	15	None (Assumes a 50 node network)	345	1281
	20	None (Assumes a 50 node network)	440	1711
	25	None (Assumes a 50 node network)	535	2141
Modified RAP	5	(Assumes $l = 10$)	224	354
	10	(Assumes $l = 10$)	434	873
	15	(Assumes $l = 10$)	629	1467
	20	(Assumes $l = 10$)	821	2101
	25	(Assumes $l = 10$)	1002	2755
Orthogonal Addressing	5	NA	119	256
	10	NA	214	521
	15	NA	309	786
	20	NA	404	1051
	25	NA	499	1316

1. p-persistent slotted aloha and the time slotted p-persistent CSMA parameters were chosen by seeking the fewest number of slots and the corresponding smallest access probability that could achieve a 99.9% access rate for the specified number of contenders.
2. EYNPMA parameters were selected based on the exhaustive search described earlier.
3. The length of the contention period when polling and selective polling are the medium access protocols is dependent on the number of nodes in the network. We assume 50 nodes for this table.
4. The modified RAP protocol has improved performance as the number of available addresses increases. (See Figure 4.15.) We assume that there are 10 addresses available throughout.

Table 4.2: Protocol performance comparison

these protocols is further magnified when hidden nodes exist in the network. Polling is immune to hidden node effects. On the other hand, time slotted p-persistent CSMA, which relies on all nodes monitoring each other, quickly degrades to pure Aloha when nodes do not monitor each other. Of the random access protocols that do not use polling, EYNPMA is the best. It retains a throughput close to the designed throughput even when the number of nodes contending varies. Additionally its energy performance is very good when the number of nodes contending is small.

4.5 Conclusion

In this chapter we have presented several access protocols that may be used in the contention period of medium access protocols that alternate between contention and contention free periods. These protocols enable the PC to provide information to the mobile nodes that allow them to enter the doze state during the CP. We have provided models that may be used to predict not only the expected number of successful contentions but also the energy that is consumed by these protocols as they support mobile nodes gaining access. We find that although the random access methods can provide efficient access they have a hard time competing with the polling protocols in conserving energy, especially as the number of nodes contending increases. But considering the facts that nodes not contending for access can doze during the CP, that the packets used to gain access are relatively small, and that when polling is used a third period must be provided to give new nodes the opportunity to associate with a network, the random access protocols are still most likely to be used. Of the random access protocols, orthogonal addressing followed by p-persistent slotted aloha result in the least energy consumption.

Chapter 5

Synchronous Collision Resolution: An Energy Efficient Access Protocol for Ad Hoc Networking[23]

5.1 Introduction

Two very significant results are revealed in Chapters 3 and 4 on what makes an access protocol effective at conserving energy. First, the access protocol must be efficient in its distribution of the information to nodes that they need to know in order to enter low energy states. Second, it must disseminate this information in a timely manner so nodes do not consume energy while waiting for it. Unlike with the protocols that use a central controller, there appeared to be no candidates among the existing protocols most often used in ad hoc networks since virtually all are based on temporally random access techniques. That is, access is granted on account of when nodes do things and these activities occur randomly. Indeed, there is no predictability that can be used to help nodes conserve energy. Any efforts to add predictability result in a corresponding decrease in other performance objectives such as throughput and capacity of the network. For this reason we did not seek to compare different ac-

[23] Patent Pending

cess mechanism already in existence since none were very attractive. Rather, from the beginning, we sought to create an access protocol that offers the predictability that can be exploited for energy conservation. The result is a novel protocol that we call Synchronous Collision Resolution (SCR). This protocol not only supports energy conservation but it also achieves exceptional performance in most every measure. It is the only protocol that we are aware of that provides a mechanism for multihop stream based services in an ad hoc network. It achieves stable throughputs that exceed the maximum throughputs for such familiar protocols as slotted aloha and CSMA. It does not suffer congestion collapse. And it can achieve high spatial reuse. On account of this breadth of abilities, this chapter does not merely focus on the protocol's energy conservation mechanisms but attempts to fully define a protocol that can meet the myriad of challenges that confront ad hoc networking access protocols.

We start our discussion by with background information on the challenges that confront ad hoc networks in Section 5.2. Then in Section 5.3 we describe SCR. We conclude the chapter with Section 5.4. Chapter 6 then continues the discussion with an investigation of the spatial capacity of SCR and presents techniques that may be used to improve it.

5.2 Background and Challenges

Ad hoc networks have been proposed as a solution to wireless networking where nodes are mobile, the range of their mobility exceeds the transmission range of any single transceiver, and there is no existing network infrastructure. Typical proposed applications include military command and control nets, emergency networks for disaster relief, networks for search and rescue operations, and sensor networks. However, ad hoc networking may find broader application as a predominant networking approach in offices but also across college campuses, in homes, and even possibly as the architecture of a future generation of mobile wireless data networks. Indeed, the increasing availability of high frequency transceivers (i.e. high bandwidth) will allow simultaneous support of multiple connections at different qualities of service. However, the increased attenuation associated with such high frequencies will make cellular architectures centered on nodes communicating directly with base stations impractical. The natural alternative is for nodes to communicate with their peers and to cooperate to support connectivity as is envisaged by the ad hoc networking approach.

Unfortunately, ad hoc networking must overcome several challenges to realize its full potential. Below we briefly discuss these challenges, identify current efforts to address them, (specifically by the IEEE 802.11, [6], and the ETSI HIPERLAN, [7], medium access control (MAC) protocols), and succinctly describe how the Synchronous Collision Resolution (SCR) protocol proposed in this chapter tackles the problems and offers superior performance.

5.2.1 Single Channel Access

The primary objective of access protocols is to provide uninterrupted access to a radio channel. This is typically achieved by either scheduling accesses or by introducing temporal randomness so that nodes may gain access at different times. Due to the distributed nature of ad hoc networks, temporally random access techniques improving on the principles underlying the Aloha and CSMA protocols have dominated.

The 802.11 and the HIPERLAN MAC protocols exemplify the state of the art. The 802.11 MAC is based on a distributed CSMA like protocol that seeks to avoid collisions when the channel is congested. When a packet arrives, a node schedules the transmission for a randomly selected time slot within a contention window. It then senses the channel until it is idle for that selected number of time slots, after which it transmits the packet. If the packet is not received successfully, a backoff algorithm randomly schedules retransmission in a larger contention window. The contention window increases exponentially with each failure. Although this reduces the probability that nodes which collide will subsequently collide, the protocol's performance still degrades with increased density of nodes in the network. [33]

The HIPERLAN protocol takes a different approach using *collision resolution signaling* to resolve congestion. Nodes first attempt to send packets as they arrive. If a packet arrives when the channel is busy, the node waits until the channel is free again and then contends immediately using the collision resolution signaling protocol EYNPMA described in Chapter 5. When multiple nodes within range of each other contend simultaneously, this protocol is better than 96% effective at isolating a single winner. [34], [31] By contrast with the 802.11 protocol, the performance of this access mechanism is robust to the density of nodes

Our Synchronous Collision Resolution (SCR) access method leverages the effectiveness of signaling at resolving simultaneous access attempts. Rather than using a temporally random access mechanism; however, SCR requires all nodes to contend simultaneously and synchronously and then relies on the collision resolution signaling to give a single node the right to transmit data on the channel. We will show that applying collision resolution signaling in a synchronous manner results in multiple benefits not yet achieved using temporally random access schemes.

5.2.2 Spatial Reuse

Ad hoc networks consist of spatially distributed nodes with no guarantees that all nodes are within range of each other. This allows spatial reuse of the access channel but also makes access control subject to spatial interference more challenging. Three different phenomena associated with signal strengths at receivers play a role:

hidden nodes, capture, and dominant nodes. The hidden node phenomenon occurs when a destination node is between two transmitting nodes that are too far apart to hear each other. The destination is unable to receive its traffic because these nodes interfere with each other. The capture effect occurs when a receiver can distinguish one of several signals since it is stronger. The capture effect may mitigate the hidden node problem; however, capture may also cause the dominant node problem. The dominant node problem occurs when two nodes are perennial competitors in sending data to a third node. The node that has the stronger signal may, perhaps unfairly, always have precedence in gaining access and is referred to as the dominant node. Access protocols in ad hoc networks attempt to eliminate the hidden node problem while simultaneously attempting to achieve fair spatial reuse of the access channel.

The deleterious impact of the hidden node problem is a significant problem for CSMA protocols. In fact, when CSMA was presented as a wireless access mechanism it was noted that CSMA's performance would degrade to that of Aloha if the problem were severe and so an out-of-band signaling solution was suggested [35].[24] Subsequently, [36] demonstrated that a two-way handshake prior to transmission of data could suppress collisions. The principle is for the source to first transmit a short "request-to-send" (RTS) packet to the destination implicitly directing all nodes within range to back off while the data exchange takes place. Upon receiving the RTS, the destination replies with a "clear-to-send" (CTS) packet, which simultaneously directs its neighbors to back off while giving the source the go-ahead to transmit the data packet. Collisions may still occur during RTS-CTS exchanges but they only interfere with these short transmissions rather than the longer data transmis-

[24] The performance of Aloha is not affected by the hidden node problem since it does not use carrier sensing.

sions. In fact, [37] shows that a successful RTS-CTS exchange is a sufficient condition to assure no collision in the subsequent data transmission. The RTS-CTS exchange is currently used to suppress hidden node interference by the 802.11 protocol. By contrast, HIPERLAN takes a different approach in dealing with hidden nodes. A node assumes a collision has occurred when it fails to gain access but then does not detect a data transmission. When this is the case the node foregoes attempting to gain access for at least 500 ms. This response is called the “hidden elimination” condition. Unfortunately, studies in [38] and [34] reveal HIPERLAN’s performance degrades significantly with an increase of hidden nodes.

The mechanism used in SCR to suppress hidden nodes is conceptually different. SCR first uses the collision resolution signaling protocol to resolve the contending nodes to a reduced set of spatially distributed nodes. These nodes then *simultaneously* transmit RTSs. The destinations that successfully receive an RTS then *simultaneously* transmit a CTS in response. Rather than suppressing subsequent access attempts by nodes within range, the RTS-CTS exchanges serve to ensure that in an environment with capture the subsequent data transmissions will not collide. Only contending nodes that receive a CTS will transmit a data packet. We will show that this technique prevents collisions in data exchanges.

A second issue in ad hoc networks with spatially distributed nodes is channel reuse. We will demonstrate that SCR’s collision resolution signaling is not only effective at isolating successful contenders but that it also resolves to an efficient distribution of nodes that can successfully exchange data simultaneously.

5.2.3 Quality of Service (QoS)

The distributed nature of access protocols for ad hoc networks makes it difficult to obtain per packet service differentiation or to guarantee compliance with any bandwidth or delay constraints, particularly if a temporally random access scheme is used. Although mechanisms can be designed to give contention priority to a particular node, it is difficult to ensure bandwidth guarantees especially in congested mobile networks where there may be an increased and unpredictable number of nodes contending at high priority. The typical approach to supporting these services in networks is to enable nodes to reserve resources. However, we are aware of no current access protocol for ad hoc networks that provides this capability.

The 802.11 protocol supports two types of coordination functions built upon its distributed access mechanisms, a distributed coordination function (DCF) and a point coordination function (PCF). The DCF is the default coordination function and the PCF is present to support QoS. PCF enables a central node, the point coordinator (PC), to schedule transmissions; however, the standard provides no guidance on how this should be done. We have shown that the PCF can be used to support multimedia traffic in [39], but this assumes that all nodes are within range of the PC. In a widely distributed network, this approach would require multiple PCs working together to manage access. We are aware of no work attempting such a solution for an ad hoc network. In fact, the DCF remains the predominant coordination function used in ad hoc networking. In [40] and [41], attempts are made to add signaling to the DCF so as to enable it to support constant bit rate services; however, this work again requires *all* nodes to be within range of each other.

HIPERLAN's method of supporting QoS is to give priority to packets based on their residual lifetime. The EYNPMA protocol has five levels of priority, all based on the residual lifetime of a packet. Unfortunately, experiments in [42] demonstrate that this priority scheme is not effective at supporting QoS since there are no mechanisms to distinguish between packets associated with real time services and old data packets. The problem is, of course, more pronounced in congested networks where more nodes would contend to transmit old, i.e. higher priority, data packets.

SCR's supports per packet service differentiation and connection oriented services, all in a completely distributed manner. The proposed protocol uses a novel priority scheme in the signaling protocol that allows individual nodes to contend for and maintain slots based on a "use it or lose it principle." Moreover, the mechanism can be used to set up a multihop connection oriented service within a dynamic ad hoc network.

5.2.4 Energy Consumption

We have found in Chapters 3 and 4 that the critical design issue in energy conserving protocols is the simultaneous scheduling of wake up periods with the dissemination of information that allows those nodes to return to the doze state. Access protocols that are based on temporally random access have difficulty achieving such simultaneous scheduling. In particular, waking nodes are required to stay awake for some period of time due to unforeseen contentions. By contrast, SCR is synchronous and thus can schedule these events simultaneously.

5.2.5 Timing

Establishing a common clock in a distributed network is quite challenging due to the propagation delays among nodes. In a distributed network, there is no clear choice of a node to provide the clock. In 802.11, every node has the burden of providing the clock. On a periodic basis, all nodes contend to transmit beacons. The node that successfully transmits the beacon provides a timestamp. All other nodes in the network that monitor this beacon adjust their clocks to this timestamp. Propagation delays result in clocks varying by 1 μs for every 300 meters of separation. In widely distributed networks such synchronization may not be based on the same beacon. The 802.11 protocol tries to mitigate these effects using time slotting and interframe spacing. These spaces create specified delays between a transmission and a subsequent response that prevents other nodes from interfering with ongoing exchanges. In HIPERLAN, the synchronization of the EYNPMA channel access protocol is always keyed to the previous data transmission. If nodes do not contend at this time then the network enters the “channel free condition” where there is no particular timing mechanism and data is not sent using EYNPMA. Since in a widely distributed network it is likely many nodes do not hear every transmission, this timing mechanism has only a local effect.

Timing is a critical component of SCR. Synchronization of the network directly affects the amount of overhead associated with signaling. We provide no specific mechanism as part of the protocol to achieve it. However, we envision the timing to be achieved by other mechanisms in the network. Several techniques for providing a universal clock exist and can be managed hand-in-hand with location awareness technology. We believe that location awareness is a significant requirement not only in the operation of ad hoc networks but also will become integral to applications on the mobile platforms. The Global Positioning Satellite (GPS) system is the obvious source for this location information. If used the GPS satellites can also provide a universal clock to the distributed nodes of an ad hoc network. Complete integrated GPS receivers are inexpensive and boast a timing accuracy of 5 nsec. [43] Timing can be achieved locally using the same principles as GPS. So long as a mobile node can receive signals from four synchronized clocks it can calculate the three dimensions of its location and synchronize its own clock.

5.3. Synchronous Collision Resolution (SCR)

To our knowledge, SCR is the first protocol for ad hoc networks to simultaneously address the problems of congestion and hidden nodes while providing mechanisms for connection oriented QoS. In addition, it does this in a manner that allows spatial reuse of the channel. It uses signaling to address congestion and then uses an

RTS-CTS exchange to insure that there are no collisions with hidden nodes during the transmission of data. It features three additional characteristics. First, we synchronize all contentions in the network. This timing provides three significant advantages. The simultaneous attempts to contend allow collision resolution signaling to resolve a spatial distribution of transmitting nodes that will not interfere with each other during the transmission of data. The timing also results in the periodic transmission of data that is easily exploited for QoS and energy conservation. Second, the protocol uses time slotting. Time slotting minimizes the effect of propagation delays and transceiver state transitions. Finally, signaling is specifically designed to support QoS. Unique to its signaling approach is a cooperative effort between the source and destination to seize and retain control of the channel for stream based traffic. This is of particular importance for connection-oriented traffic in mobile networks. The synergism of these wireless access techniques results in a remarkable protocol that outperforms currently proposed protocols for ad hoc networking.

SCR requires all nodes to synchronize their attempts to gain access to the network channel. This approach has traditionally been ignored on account of the high probability of a collision, the difficulty in isolating a single node, and the difficulty in synchronizing a distributed set of nodes. We start this section by describing the signaling protocol that we will use to isolate one node from multiple simultaneous contenders. Then, we address how we solve the hidden node problem, describe how a refinement to the signaling scheme enables quality of service (QoS), and finally, discuss how the protocol supports energy conservation and how we enhance this capability through additional functionality in the signaling scheme.

Two concepts that are important to distinguish in this description are that of the transmission slot and the signaling slot. The transmission slots occur at a regular interval and accommodate the transmission of a single packet. At the beginning of each transmission slot is a signaling period used to determine which nodes get access to the channel. This signaling period is also slotted. These slots are referred to as signaling slots. Figure 5.4 shows the relation of the signaling slots to the transmission slots.

5.3.1 Collision Resolution Signaling

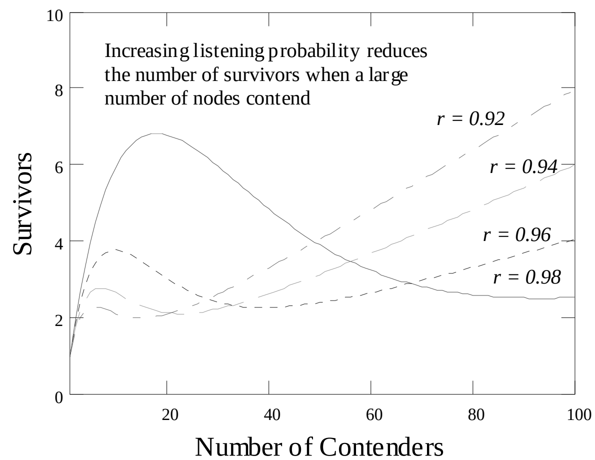
Collision resolution signaling uses signaling to select a single node to transmit data among multiple contenders. Our version of this channel access scheme consists of three signaling phases and two access signals. The first signal starts at some random point in the first phase and ends in the second. The second access signal starts at some time within the third phase and ends at the phase's end when a node starts to transmit a packet. A node wins the contention by being among the first to start transmitting in the first phase, among the last to stop transmitting in the second phase and the first to start transmitting in the third phase. Nodes that recognize that they have lost the contention in any one of the phases will defer from attempting to gain access.

Signaling Organization

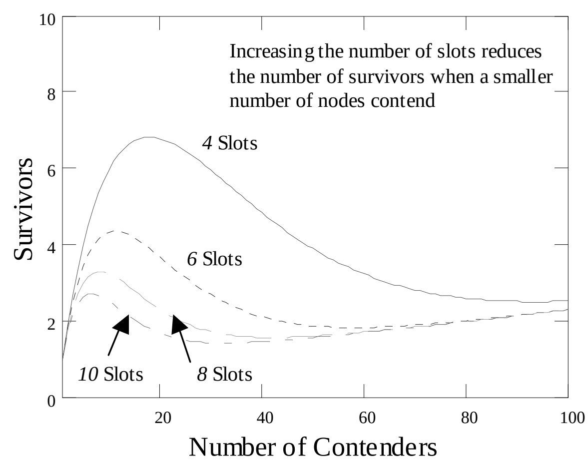
This protocol is similar to the EYNPMA channel access control protocol used in the HIPERLAN standard so we use the standard's names for the three phases in the sequel: priority, elimination, and yield phases. Each phase consists of an integer number of signaling slots which we denote by H, L, and M respectively. A node with a packet to transmit will choose to start transmitting in one of the priority phase slots in the following manner. For each of the first $h - 1$ slots, if a node has not sensed another node's access signal, it will chose to start its own transmission with probability $1 - r$. (r is referred to as the listening probability) If the node has not started to transmit prior to the h^{th} slot and it has not sensed another node's access signal, it will start in the last slot with probability 1. Nodes that successfully start transmitting an access signal continue throughout the priority phase and into the elimination phase. The same technique is used for selecting a slot to end the access signal. A node stops transmitting in any of the first $l - 1$ slots of the second phase with probability $1 - q$ or stops on the l^{th} slot with probability 1. (q is referred to as the transmission probability.) After the contending node stops transmitting, it listens to the channel and will defer its contention if it hears another node still transmitting an access signal. A slot is reserved at the end of the elimination phase in which no signaling takes place. This slot allows nodes that transmit through the last elimination phase slot to verify their survival. It is called a survival verification slot. Finally, if a node is still contending, it will repeat the process used in the priority phase on the m slots of the yield phase but using a listening probability of p. Figure 4.9a illustrates this collision resolution signaling process. Note that this protocol is similar to that described in Section 4.3.3 but differs in that contending nodes do not start transmitting the packets in the yield phase. Nevertheless, Equation 4.7 remains valid as the probability that a contention is successful when k nodes are contending. Note that the analysis performed in Section 4.3.3 considered nodes to have the option to doze as soon as they discerned that they would not win the contention. In ad hoc networks nodes must remain awake through the signaling as they may be the destination of the contention winner. The energy consumed in signaling is essentially fixed depending on the number of signaling slots used. Energy consumption does not enter the equation used to select the listening and transmission probabilities. We now discuss how to select the parameters of the signaling protocol.

Selecting the Signaling Parameters

We classify the signaling phases as being one of two types, “first to assert” or “last to assert.” The priority and yield phases are first to assert phases while the elimination phase is a last to assert phase. In first to assert phases the goal is for a small number of nodes to assert themselves first thus excluding the remaining nodes. In last to assert phases the goal is to gracefully allow nodes to stop transmitting such that there are a small number of survivors in the end. The characteristic performance of the two types of phases are illustrated in Figures 5.1 and 5.2. The issue is how the number of contenders, the listening or transmission probability, and the number of slots affect the number of survivors of each phase. The first to assert and last to assert phases have different behavior. For a given set of parameters, i.e. number of slots and listening probability, first to assert phase survivor quantities have a relative maximum, followed by a relative minimum and then monotonically increase with the number of contenders. Increasing the listening probability while keeping the number of slots constant moves the relative minimum to when a large number of nodes contend with the tradeoff of increasing the relative maximum. For a given listening probability, increasing the number of slots reduces the relative maximum. In last to assert phases, the survivor quantity increases monotonically; however, it has a region where the performance is nearly flat. Increasing the transmission probability reduces the initial survivor quantity at the expense of reducing the range of the flat region. Increasing the number of slots extends the flat region.



a. First to Assert Phase with *4* Slots



b. First to Assert Phase with *0.98* Listen Probability

Figure 5.1: The effect of parameters on the expected number of survivors in first to assert signaling phases

The goal of using three signaling phases is to use the first two phases to thin out the contenders and then to use the final phase to isolate a single winner. Thinning out can be accomplished using just a few slots in the priority phase. Note that in Figure 5.2a that 4 slots and a listening probability of 0.96 will result in an expected number of survivors that varies by just 2 for a range of 4 to 100 nodes contending. The

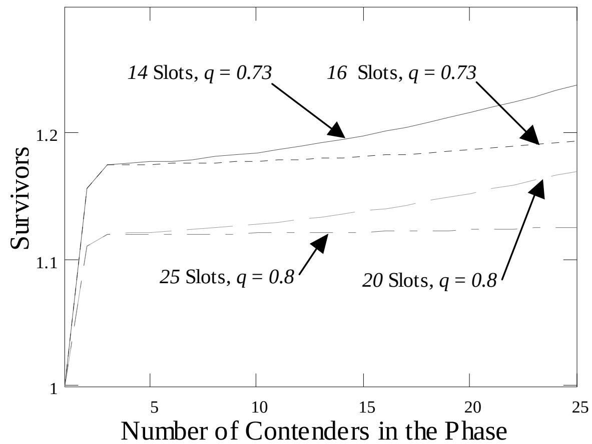


Figure 5.2: Example of the performance of a last to assert phase

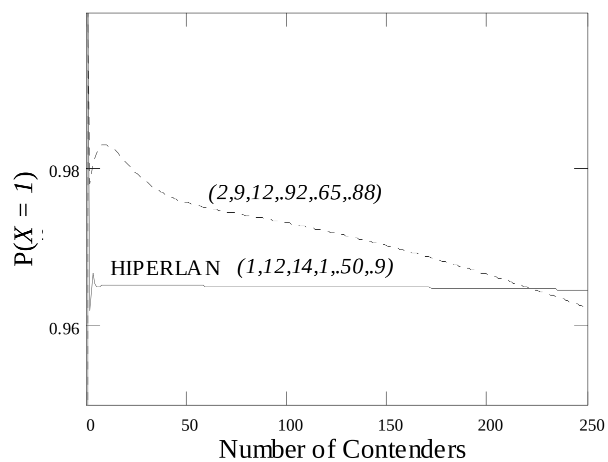


Figure 5.3: Comparison of aggressive versus consistent signaling parameters, (H, L, M, r, q, p)

elimination phase then seeks a low survivor quantity for the expected range of survivors from the priority phase. In the end, the yield phase uses a large number of slots and a high listening probability to isolate just one node. The effectiveness of collision resolution signaling can always be improved by adding more slots to any of the phases. Also for any number of slots used, the transmission and listening probabilities can be chosen either for consistent success rates for a wide range of contending

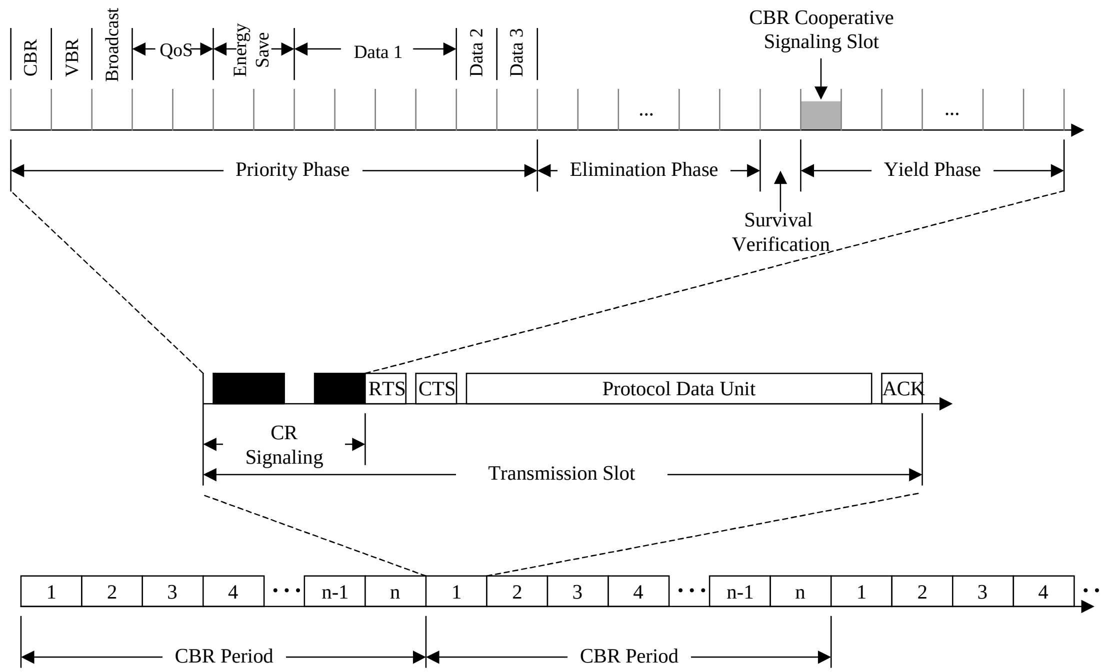


Figure 5.4: The organization of the Synchronous Collision Resolution protocol

nodes or to aggressively seek higher success rates at the expense of less consistency. Figure 5.3 exhibits an example. The flat graph corresponds to the performance of HIPERLAN's EYNPMA parameter settings. The more aggressive performance is achieved using the same number of signaling slots, 31.[25] Nevertheless, in both cases, the performance of collision resolution signaling is very robust.

5.3.2 The RTS-CTS Exchange

Hidden nodes can interfere with transmissions in four different ways. First, two sources may interfere at a destination. Second, a source may interfere with the reception of an acknowledgement at another source. Third, the acknowledgement of one destination may interfere with the reception of data at a second destination.

Fourth, the acknowledgement of a destination may interfere with the reception of an acknowledgement from another destination at its source.

SCR approaches the hidden node problem using an RTS-CTS exchange, but the underlying principle is different. Indeed, rather than relying on the timing of RTS-CTS exchanges to suppress other contending nodes, the protocol relies on collision resolution signaling to suppress other contenders. RTSs and then CTSs in SCR are transmitted *simultaneously* to test whether current capture conditions will support successful reception of subsequent packets in an environment with hidden nodes. Subject to the following assumptions, one can show that this approach prevents collisions while enabling efficient use of capture to promote spatial reuse.

ASSUMPTIONS

i) RTS packets are transmitted simultaneously.

ii) CTS packets are transmitted simultaneously.

iii) The network uses fixed sized data packets and fixed sized intervals between RTS, CTS, data packet, and acknowledgement transmissions, so packets and acknowledgements are transmitted simultaneously.

iv) A node will never transmit data using a higher power than it uses in transmitting the RTS or the CTS.[26]

v) Channel characteristics remain constant throughout the transmission slot.

THEOREM: The synchronous transmission of RTS and CTS packets in SCR prevents all collisions during the transmission of data.

---

[25] Note that we assume that the contenders all contend at the same priority level. EYNPMA has five priority levels. In the second set of parameters we used 2 slots rather than 1in the priority phase. This equates to moving 5 slots from the elimination and yield phases.

[26] A node may use a lower power during transmission of a packet or an ACK based on feedback from the destination that confirms that the lower power would support a successful exchange under the transmission conditions that existed during the transmission of the RTS or the CTS.

PROOF: Source nodes transmit packets if there is a successful RTS-CTS handshake. Since RTS and CTS transmissions are sent simultaneously, so too are packets and acknowledgements. Collisions cannot occur between acknowledgements and data packets. A destination's successful reception of a RTS transmission from a source indicates it will also successfully receive a data packet transmitted from that source. A source's successful reception of a CTS from a destination indicates that it will also successfully receive an acknowledgement from that destination. Therefore, a successful RTS-CTS handshake indicates that the subsequent data packet and acknowledgement transmissions will not fail on account of collisions. ■

The protocol can further exploit capture if the transceivers can support pairwise coding and power control. The RTS-CTS exchange can be used to coordinate the codes and power levels used in the subsequent data transmission. Power levels will only be decreased since RTS-CTS exchanges are executed at the maximum power level and it is these signals that are used as the basis of any adjustment. The use of lower power levels by all nodes will reduce background noise. Unique codes allow pairs of nodes to capture their exchange even through fades. With these techniques, data transmissions become more robust to interference and fades validating Assumption (v) above. Since all nodes monitor the RTS and CTS transmissions they can be leveraged to disseminate network state and/or coordination information to support routing, power control, and energy conservation.

5.3.3 Synchronization and Time Slotting

Time slotting is a convenient way to minimize the impact of synchronization discrepancies, propagation delays, transceiver transition delays, and channel sensing times. We demonstrate how to select the duration of signaling slots and interframe spaces to overcome lack of synchronization in Appendix 5.A.

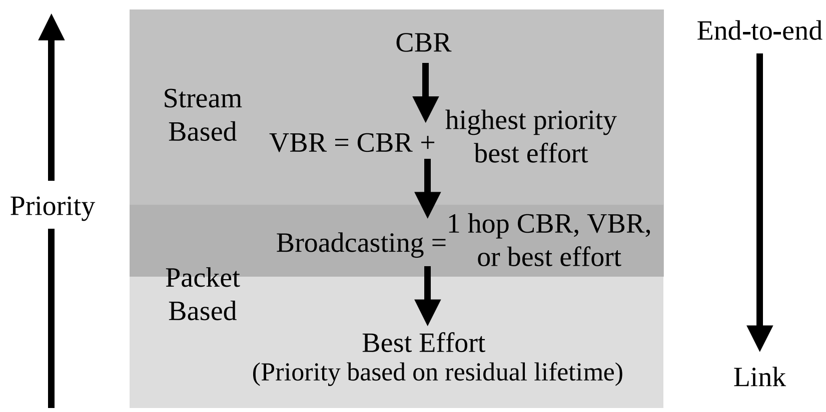


Figure 5.5: Hierarchy of service types

5.3.4 Quality of Service (QoS)

Our traffic management objectives are to enable the reliable transmission of stream based traffic, both constant bit rate (CBR) and the more bursty variable bit rate (VBR) traffic, to give priority to best effort packets based on their residual lifetime, and to support broadcast transmissions. Figure 5.5 illustrates their relationship. CBR and VBR are both stream based services that can be provided end-to-end while best effort is a packet based service that occurs across a link. The broadcast service is in the middle being either stream based or packet based but only being available to destinations within the transmission range of the source.

These QoS goals are achieved using periodic frames, a priority access scheme, and a specialized signaling mechanism. Figure 5.4 illustrates their organization. The periodic frame is called a Constant Bit Rate (CBR) frame since it is repeated at a rate that allows a single transmission slot in each frame to support the lowest desired CBR rate. A node can achieve a higher CBR rate by using more than one transmission slot per frame. The CBR frame is repeated on an interval we call the CBR period. The priority phase is used to differentiate between the service types. As illustrated, the highest priority signaling slots occur first and are associated with services having higher priority. Note that the priority phase also has two additional slots to support energy conservation. Their use will be described in Section 5.3.5. Signaling follows the general description presented in Section 5.3.1 but note that an additional signaling slot called the cooperative signaling slot has been added to the yield phase. Signaling has been modified to allow destinations to also signal using this slot to support transmission slot reservation for CBR services. The description of the integrated use of these components follows.

Contending nodes make two decisions in using the priority phase. First they decide which group of slots to use and then which slot within the group to begin transmitting. Each node selects the group of slots based on the service type needed and then selects the signaling slot within the group to begin transmitting its assertion burst according to the procedures described in Section 5.3.1. If the group has only one slot, the node begins transmitting with probability 1 otherwise it uses a listening probability r selected based on the number of slots in the priority group. Priority in access is achieved by contending early in the priority phase. Since the groups are ordered by the priority of the service type, the higher priority services will contend first and thus have precedence. Nodes use listening and transmission probabilities (q and p) in the elimination and yield phases that are optimized for the number of slots used in the priority group selected.

The priority phase provides three different priority levels for best effort services, one slot each for the lower data priorities and four slots for the highest priority. Nodes contending for best effort service decide which group of slots to use based on the residual lifetime of the packet they are trying to send. This partition of slots assumes that when a node can gain access contending with a low priority there are just a few nodes that it will compete against so there is little need for multiple signaling slots. However, when the highest priority is required it is very likely that multiple nodes will be contending (i.e. the network is overloaded) so multiple slots are included in the priority group.

Three groups of signaling slots are associated with CBR and VBR stream based services. A node first contends for such services using the lowest priority QoS group. If the node is successful it can then assert priority in gaining access to the same transmission slot in the next CBR frame using the CBR signaling slot. Since a node can only use a CBR signaling slot to gain priority access if it had accessed the same transmission slot in the previous frame it is assured that it is the only node in its transmission area that can contend using this signaling slot.[27] In this manner a source node can effectively reserve a specific slot in each CBR frame. Depending on the CBR rate it requires, this source node can repeat the process and reserve additional transmission slots in the CBR frame.

We assume that a VBR stream can be serviced with a combination of a CBR stream and a variable number of additional transmission slots corresponding to the bursty nature of the stream. Therefore, nodes requiring VBR service first contend and reserve transmission slots in the same manner as the CBR streams and then use the VBR slot to access transmission slots to send additional bursts. The right to use the VBR signaling slot is reserved to those nodes that already have CBR access within the frame. The node is not guaranteed access but has greater priority than any other nodes except those sending CBR or those also contending to send VBR traffic. Note that CBR slot reservations require a node to use the slot or lose it while VBR access allows nodes to transmit packets in addition to those sent on a CBR stream on an as needed basis.

---

[27] It is possible in mobile networks for a second node that is using the same CBR frame slot for CBR service to move within range of the first. In the case of this rare event, the protocol relies on the elimination phase to distinguish which node wins the contention.

A source node can only reserve a transmission slot in its transmission range, thus there is a risk that a contender outside a source's transmission range but within that of the destination may interfere. To assure both source and destination to have priority in CBR contentions we provide a mechanism that allows the destination to also clear contenders from the area within its range. A destination node recognizes when it is the recipient of CBR traffic, i.e. it knows that the CBR signaling slot was used to gain access and that it was the destination of the traffic sent in the same transmission slot of the previous CBR frame. Under these conditions the destination also participates in the signaling protocol. We add an additional slot to the yield phase, called the CBR cooperative signaling slot, which is used by both the source and the destination of CBR traffic. Starting with the CBR cooperative signaling slot of the yield phase both the source and destination transmit an assertion signal. Since only CBR traffic can use this signaling slot, the CBR traffic receives priority over all other contentions. The destination, however, stops transmitting the signal early so that it can transition to receive the RTS from the source at the usual time in the protocol. The rest of the exchanges in the transmission slot are the same. By having the destination cooperate in this way one guarantees that no nodes interfere at the destination.

A mainstay of ad hoc networking is broadcasting packets to multiple nodes simultaneously. This presents a challenge to SCR since SCR relies on the RTS-CTS exchange to determine whether a packet should be transmitted. When a packet is broadcasted there is no destination that has the clear responsibility to transmit a CTS. If none of the destinations transmit a CTS then contenders outside the range of the broadcasting node may gain access and interfere at the broadcasting node's destinations. If all destinations respond with a CTS, the source, following our earlier rules for the RTS-CTS exchange, will identify a collision even though there may not have been one. Therefore we provide a special priority for broadcasting and an optional CTS procedure, either all potential destinations send CTSs or none do.. First, we give broadcasts a higher priority than data and new QoS attempts. This higher priority allows potential destinations to identify that a broadcast is being attempted. Second, if the first CTS option is used, we require all destinations that identify the use of the broadcast priority slot to respond by broadcasting a CTS despite successfully receiving another transmitting node's RTS. The exception is if the destination also hears the use of cooperative signaling in the yield phase in which case the destination should not transmit a CTS at all. Despite the fact that the broadcasting node cannot decipher the CTS, it can assume that if it survives the signaling phase that it can broadcast without interfering with another exchange within its transmission region since no node should be a destination of another source. Nodes that receive a broadcast do not respond with an acknowledgement so there is no risk that a broadcast will interfere with a distant CBR exchange. There is, however, a small risk that some destinations may not receive the broadcast if there is either a CBR or another broadcast source that is in closer proximity. The second CTS option is for none of the destinations to respond with a CTS. This option reduces the occurrence of broadcasts overwhelming the network and suppressing point-to-point packet transmissions.[28] The cost of this option with respect to the first is that more destinations will not receive the broadcast on account of interference. We allow broadcasting nodes that successfully contend to then use CBR signaling slots in the same transmission slots of subsequent CBR frames and all previous destinations then use the cooperative signaling slot in response. In this way the protocol supports single hop stream-based broadcast service.

[28] When the CTS option is used a single broadcast can suppress other exchanges in an area up to four times as large as would a successful contention for a point-to-point packet exchange.

The CBR mechanisms for priority access provide the capability to establish a connection-oriented service in ad hoc networks. Nodes along a path may sequentially reserve transmission slots to establish the connection. Say a source wants to establish a CBR connection to a destination two hops away. It starts by reserving transmission slots to the intermediate node. The intermediate node cooperates with the source while simultaneously attempting to reserve an equivalent number of slots to the destination. Once these slots are reserved the connection is established and the CBR stream may begin. Call admission is self administered and cooperatively enforced. Connections are maintained so long as they are used. In this manner, ad hoc networking may become a commercially viable method to provide telephony services. This connection oriented capability may also enable ad hoc networks to increase capacity through a type of flow routing control. Multihop CBR links can be set up between distant nodes across uncongested regions and then advertised as a single low cost terminal to the routing protocols, thereby diverting traffic away from congested regions.

We conclude this section by emphasizing that the number of slots used in each of the signaling phases including the number of slots used at different priority levels should be chosen based on the projected demands for services and the density of nodes that the network is expected to support. When different quantities of slots are used with different priority levels choosing the parameters of the collision resolution signaling protocol becomes a bit more complicated. In our example there are three different numbers of slots used in the priority phase signaling, 4, 2, and 1. To obtain the best performance it is practical to use different transmission and listening probabilities for the elimination and yield phases based on the number of slots used to resolve contention in the priority phase. Only nodes that use the same access priority survive to contend during the elimination phase within a transmission region so it is

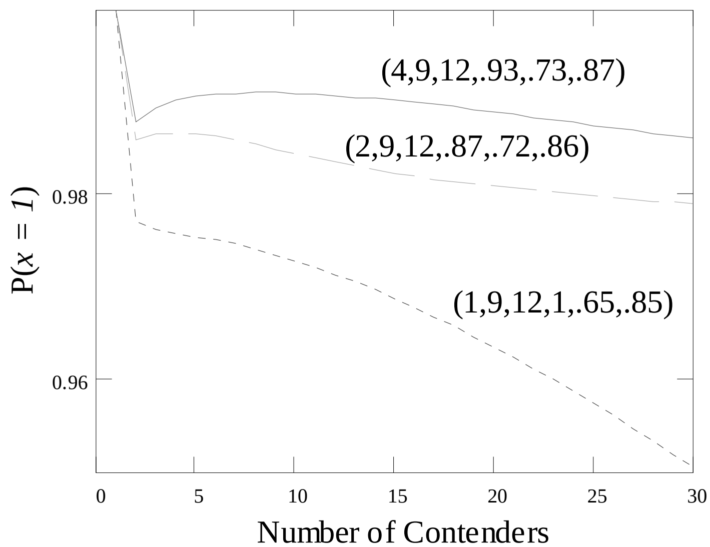


Figure 5.6: Contention success rates for signaling parameters (H, L, M, r, q, p).

reasonable to expect that these nodes will use the same transmission and listening probabilities. We illustrate several possible sets of parameters that may be used in this protocol in Figure 5.6. As expected, the more slots used to resolve contentions in the priority phase the greater the probability that a contention will result in a single survivor. Nevertheless, in a network where up to 30 nodes may contend at the same priority, the expected success rate varies by less than 4% for the different numbers of contention signaling slots used in the priority phase.

5.3.5 Energy Conservation

In Chapters 3 and 4 we explored energy conservation in wireless networks that have a central controller to manage the transceiver states of the mobile nodes in the network. This work argues and provides evidence that the most significant characteristic of an energy conserving protocol is its ability to promptly assist nodes not participating in data exchanges to enter the doze state. The key feature of wireless protocols that enables nodes to promptly enter the doze state is their ability to schedule the disseminate network state information when energy conserving nodes first wake up from dozing periods. We are aware of no distributed access protocol that achieves this goal. SCR meets these requirements since contentions are synchronous and take a finite amount of time. Nodes can wake-up prior to the contention signaling and then immediately return to the doze state after the contention if they will not participate in a data exchange. This is the default energy conservation mode and it has no effect on any other performance measure of the access protocol. The energy conserved in this protocol is a function of the ratio of the duration of the data transmission portion of a transmission slot to the total duration of the transmission slot.[29] Increasing the length of the data transmission portion or decreasing the length of the contention portion both reduce energy consumption. Suppose participating in the contention period requires less than 25% of the total energy that a node would consume if it also exchanged data in the transmission slot. Then the energy consumed by the network would be better than the best performance reported in [15] for the IEEE 802.11 protocol. On top of this base energy conservation protocol one can build additional energy conservation mechanisms that emulating either of those used by the 802.11 and HIPERLAN protocols. Next we briefly describe how energy conservation is achieved in those standards and then we describe how a better implementation of the same methods can be employed in SCR.

In the 802.11 protocol the decision to doze is initiated by individual nodes desiring to conserve energy. In the ad hoc version of an 802.11 network[30] the node that first forms the network decides whether it permits energy conservation by establish-

---

[29] The viability of this type of dozing is dependent on the transition time of the transceiver to and from the dozing state. The work presented in [126] and [8] demonstrates that the transition time to the doze state in some transceivers is as much as 100ms. This energy conserving approach would not be possible with this long of a transition time. Our assumption that transmission times can be shorter, however, is consistent with most other papers written on energy conservation protocols.

[30] The ad hoc version of an 802.11 network is referred to as an independent basic service set.

ing an "ATIM Period." If permitted, a node in the network that desires to conserve energy may doze as long as it wakes up each ATIM Period to listen for ad hoc traffic indication messages (ATIM). ATIMs are transmitted during a short window at the beginning of each ATIM period. If the node wakes and hears an ATIM directed to itself, it acknowledges the ATIM and remains awake for the rest of the ATIM period prepared to receive traffic; otherwise, the node returns to the doze state.

The energy conserving mechanism in HIPERLAN requires a node desiring to doze, a "p-saver," to coordinate with another to serve as its surrogate, a "p-supporter".[31] The two nodes agree to a period at which the p-saver will awaken to receive unicast messages and a period at which the p-supporter will transmit broadcast messages. The p-supporter node collects transmissions intended for the p-saver and then attempts to forward them to the p-saver during the coordinated transmission periods.

Both the 802.11 and the HIPERLAN methods of energy conservation can be applied on top of the default energy conservation approach described initially in this section. In the case of a small network where there is little data exchanged among nodes, the 802.11 approach may be applied. The idea is for there to be a period of transmission slots that a node might doze through with all nodes subsequently waking up prior to a designated transmission slot. If a node contends at that time then all nodes revert to the default power save mode waking prior to each transmission slot. They stay in this mode until the first transmission slot where there is no contention attempt. At that time all nodes return to using extended dozing periods following the original dozing schedule. In the case of a network where there are nodes that are predominantly stationary and have utility connections, the HIPERLAN methodology is appropriate. Mobile nodes may use these "p-supporter" nodes to help with energy conservation by transmitting to these nodes their dozing schedule. The p-supporter nodes that receive packets for the p-saver nodes would hold onto them until the slots upon which the p-saver nodes will awaken. To enhance the exchange of data to these p-saver nodes, the p-supporter nodes use the energy conservation slots of the priority phase to gain access. The use of these slots or higher priority slots for gaining access is an indication to the dozing nodes that they should remain awake. These energy conserving nodes then use the default energy conserving state until the energy conservation and higher priority slots are no longer used. At that time they return to using the original dozing schedule.

Further enhancements to the energy conservation functions may include rules that designate when nodes should contend to send stream-based traffic so these transmissions do not inadvertently keep energy constrained nodes awake. For example, if we require nodes contending for a stream-based service to delay contention until the start of the next CBR frame we will reduce the probability that there will be contention for stream-based services in the latter part of the CBR frame. If we in turn require dozing nodes to awaken in the latter part of the CBR period, we will reduce the occurrence of dozing nodes waking and being kept awake by the stream-based and broadcast contentions.

5.4 Conclusion

In this chapter we have introduced a novel MAC protocol for ad hoc networks that uses synchronous rather than a temporally random contention attempts. This approach results in many gains. Our synchronous protocol, Synchronous Collision

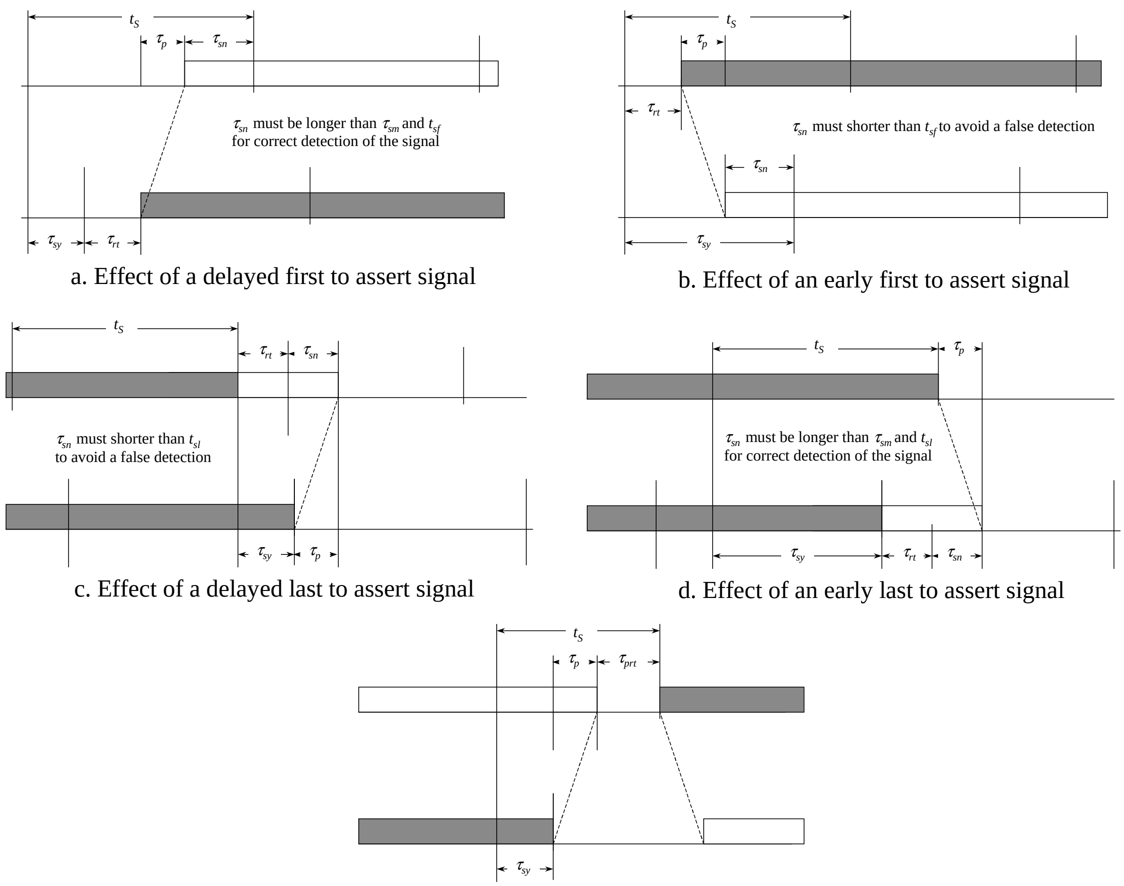
---

[31] We assume that the nodes that serve as p-supporters have a more reliable power supply than just batteries.

Resolution, provides mechanisms for differentiating service requirements in contentions, for reserving resources, and for orchestrating the transitions of nodes to low energy transceiver states. In Chapter 6 we investigate the performance of SCR, specifically its ability to support spatial reuse of a channel.

Appendix 5.A

Selecting Signaling Slot and Interframe Space Sizes



a. Effect of a delayed first to assert signal

b. Effect of an early first to assert signal

c. Effect of a delayed last to assert signal

d. Effect of an early last to assert signal

e. Effect of a delayed packet transmission

Figure 5.A1: Timing in first to assert, last to assert and interframe spaces

In this appendix we look at timing considerations for sizing the time slots used in collision resolution signaling, and used to separate RTS, CTS, data packet, and acknowledgement transmissions. For the purpose of this discussion we define the timing constraints in Table 6.A1 and design parameters in Table 6.A2.

τ_p	Propagation delay between nodes displaced the maximum receiving distance from each other
τ_{rt}	Minimum time required by a transceiver to transition from the transmit to the receive state or vice versa
τ_{prt}	Minimum time to process a signal and then to transition from the receive to the transmit state
τ_{sy}	Maximum difference in the synchronization of two nodes
τ_{sm}	Minimum time to sense a signal in order to detect its presence
τ_{sn}	Time a node senses a signal in a particular slot as a result of constraints and chosen slot size

Table 6.A1: Timing constraints and results that affect signal slot size

t_S	Duration of a slot
t_{sf}:	Selected minimum time to sense a signal in a first to assert slot to detect it.
t_{sl}	Selected minimum time to sense a signal in a last to assert slot to detect it.

Table 6.A2: Design parameters

The choice of t_S, is constrained by the minimum time required to sense a signal to detect whether a signal is present, t_{sf}. In Figure 6.A1a, where the first to assert signal is sent late, we see that t_{sn} must be longer than τ_{sf} in order for the signal to be correctly sensed. But in Figure 6.A1b, where the first to assert signal is sent early, we see that if τ_{sn} is longer than t_{sf} then there would be a false detection of the signal. From these two illustrations we derive the following equations for selecting the size of a first to assert signaling slot.

$$t_{sf} > \max\left(\tau_{sy} - \left(\tau_{rt} + \tau_p\right), \tau_{sm}\right)$$

$$t_S > \tau_{sy} + \tau_{rt} + \tau_p + t_{sf}$$

By selecting a large t_{sf} and t_S we can account for differences in the synchronization of nodes.

We find a similar result for the selection of the slot size of the last to assert phases. In Figure 6.A1c, where the a last to assert signal finishes late, τ_{sl} must be longer than t_{sn} in order to avoid a false detection. In Figure 6.A1d, where the last to assert signal finishes early, τ_{sl} must be shorter than t_{sn} in order to avoid a false detection. We provide the following equations to size the last to assert slots.

$$t_{sl} > \max\left(\tau_{sy} + \tau_{p} - \tau_{rt}, \tau_{sm}\right)$$

$$t_{S} > \tau_{sy} + \tau_{rt} + t_{sl} - \tau_{p}$$

Again we see that by selecting a large t_{sl} and t_S we can account for differences in the synchronization of nodes. We also see that the last to assert slots can be shorter then the first to assert slots.

Finally, we use Figure 6.A1e to size the interframe space between two packet transmissions. The next equation follows from the illustration.

$$t_{S} > \tau_{sy} + \tau_{p} + \tau_{prt}$$

We see from these results that the better the synchronization in the network the more efficient the protocol will be but that the failure to achieve perfect synchronization can be compensated for by using larger slot times and detection periods.

Chapter 6

Getting the Best Performance from SCR

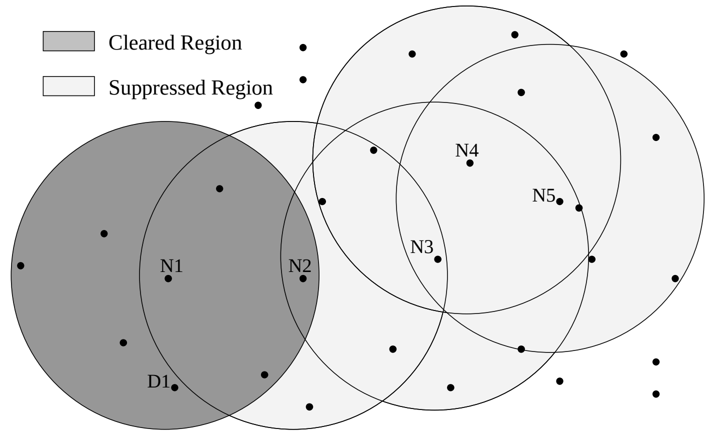
6.1 Introduction

In recent research, one measure of access protocol performance in ad hoc networks that has been woefully missing is spatial capacity. Although authors address the problem of hidden nodes, the analysis that follows only attempts to measure throughput in a single transmission region with an assumed number or percentage of hidden nodes. The more interesting and representative problem of protocol performance in an ad hoc wireless environment is the interaction of nodes in a spatial sense that allows more than one data exchange to occur at a time. In this chapter, we describe how nodes interact spatially when SCR is used, define a model to represent the spatial capacity, and then use a simulation to validate the spatial performance of SCR. Based on the intuition provided by this study we go further and explore techniques that may be used to enhance the spatial capacity of SCR, specifically next hop routing policies and the use of spread spectrum coding. We find that SCR has an inherently stable performance over a wide range of loads and is robust to congestion both in terms of load and node density. We also learn that we can more than double the range of stable performance and spatial capacity of a network using next hop routing policies and spread spectrum coding techniques. Ultimately, SCR can achieve spatial capacity that exceeds that of Aloha and CSMA based protocols by more than two to three times.

We start our presentation in Section 6.2 discussing how nodes might interact using synchronous collision resolution to gain access. Then in Section 6.3 we give a description of how spatial capacity of ad hoc networks may be modeled for SCR. In Section 6.4 we present the results of a simulation of SCR and compare this performance to that of the model of Section 6.3. In Section 6.5 we attempt to improve the performance of SCR using next hop routing policies. This section describes three policies and compares their performance. In Section 6.6 we then attempt to improve the performance of SCR using spread spectrum methods. We show that SCR and spread spectrum coding perfectly complement each other and that together they can more than double the performance of SCR. In Section 6.7 we briefly discuss the role of selecting a transmission radius to achieve optimum capacity. In Section 6.8 we briefly discuss the potential to get further capacity using advanced multiple access techniques such as code division multiple access, space division multiple access, and other smart antenna approaches. Finally, in Section 6.9 we conclude the chapter.

6.2 Spatial Interaction of Nodes Using SCR

Nodes contending using collision resolution signaling interact to determine a spatially distributed set of surviving nodes. This interaction can result in the suppression of transmissions outside the transmission range of the ultimate winner of the contention. We define two regions, the region within the transmission range of the ultimate winner of a contention called the *cleared region*, and the additional region beyond this cleared region also cleared of transmitters called the *suppressed region*. Suppressed regions are created by nodes that survive through the priority and/or



NODE	PRIORITY	h	l	m
N1	Data 1	1	4	2
N2	Data 1	1	4	4
N3	Data 1	1	3	3
N4	Data 1	1	2	2
N5	Data 1	1	1	1

Figure 6.1: Example of cleared and suppressed regions created by collision resolution signaling

elimination phases of the access protocol. Even though these nodes ultimately do not gain access, they suppress access by other nodes within their range. In turn, nodes suppressed by these nodes may have suppressed additional nodes. Figure 6.1 exhibits an example. The contention parameters are shown for all nodes transmitting packets at priority level 1. Note that nodes N1 though N5 all survive through the priority phase and in so doing suppress all other nodes in there transmission regions. Then in turn N1 suppresses N2 in the yield phase, N2 suppresses N3 in the elimination phase, and N3 suppresses both N4 and N5 in the elimination phase. N1 wins the contention while suppressing transmissions in the total shaded area. The darker shaded area is the region cleared by the signaling of the contention winner and the rest is the area cleared by the chain interaction of other nodes suppressing contenders. The discrepancy between the cleared region and the suppressed region is a measure of how effective the protocol is at spatially distributing contention winners.

6.3 Modeling Spatial Capacity

One difficulty in studying spatial capacity is defining an appropriate measure for it. A simple definition is throughput per unit area, *packets/sec/m*^{2}. It can be argued that this is not an effective metric since it can be increased by reducing nodes' transmission radii with the consequence that packets do not progress as rapidly toward their destination in a multihop network. Therefore, a second measure, the rate of forward progress of packets in a network, has also been proposed, (*packets*m*)/*sec*/*m*^{2}. This second measure is particularly relevant when next hop routing policies are applied as will be discussed in subsequent sections. In this section we focus our discussion on throughput per unit area. We first define it more carefully. To address the issue of adjusting radii, we normalize our throughput to the area covered by the transmission range of a single node and then use the average density of nodes in this transmission area as an additional parameter in the analysis. We start by assuming that the maximum transmission radius used by nodes in contending is 1 unit so that the transmission area is π square units. We use the variable σ_A to denote the density of nodes in a network and define it as the average number of nodes per transmission area. We use the variable λ_A to denote the spatial arrival rate of packets and define it as the arrival rate of packets to all the nodes in a transmission area. So, the packet arrival rate per node within a transmission area is $\lambda = \lambda_A / \sigma_A$. Finally, we denote the spatial use of the network as U_A and define it as the average number of successful transmissions in a transmission slot per transmission area, *packets*/*slot*/(π *units*^{2}).

The ability to spatially reuse a channel has been explored by several authors, especially in the early days of packet radio research. The modeling techniques focused on temporally random access protocols, e.g. Aloha and CSMA. These models required the simultaneous selection of the transmission/sensing probability[32], transmission range[33], and routing strategy[34]. Achieving the performance these models suggest, requires the independent control of the spatially offered load and the transmission range of nodes. These parameters, however, are obviously coupled. Since in collision resolution signaling the probability of having a successful contention is nearly independent of the number of nodes contending, our protocol does not have the same sensitivity to offered load. Rather, spatial capacity is dependent on the ability of collision resolution signaling to resolve a set of source-destination (S-D) pairs that do not interfere with each other. The maximum capacity would then be achieved through an optimum packing of S-D pairs such that source transmissions do not interfere with each other at their destinations. Our goal is to determine how well collision resolution signaling packs these S-D pairs.

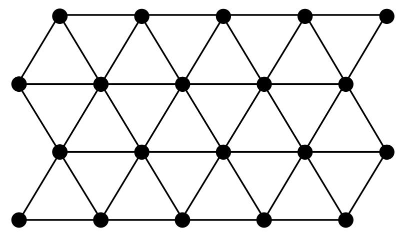
The approach of packing transmitting nodes to identify an optimum distribution of transmitting nodes was used in [44]. The authors identified that the triangular tessellation illustrated in Figure 6.2a, where transmitting nodes occur on each vertex of repeated equilateral triangles, offers the greatest density of transmitting nodes. Of course such a distribution of transmitting nodes is unlikely since nodes are randomly distributed and mobile. We will, however, adopt this approach to develop insights and to obtain a model against which our simulation results can be compared.

---

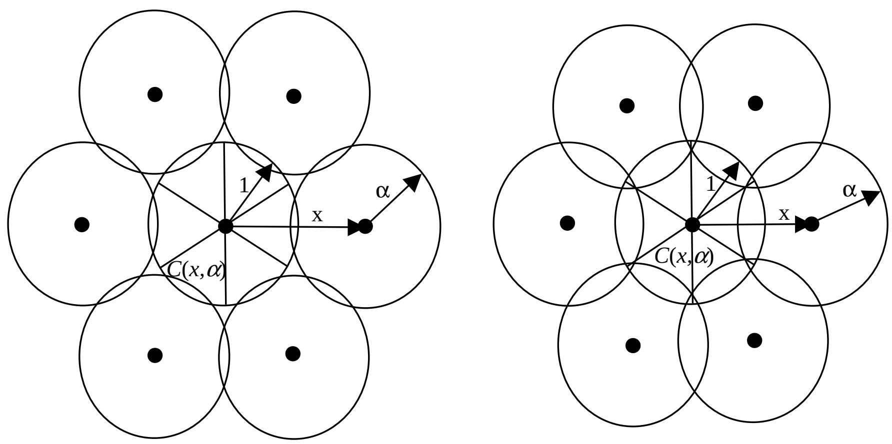
[32] Choosing the transmission/sensing probability is an attempt to choose an optimum offered load. Such a selection is not very realistic but a necessary assumption to evaluate the performance of a temporally random access protocol.

[33] Transmission range is adjusted to get an ideal density of nodes contending in a spatial region. For example [48] suggests a range that includes 6 nodes on average and [49] suggests a range that includes 8 nodes on average.

[34] Routing strategy is used when transmission range is adjustable and can be chosen to just meet the next hop distance. Such adjustment reduces interference and in turn increases throughput. Research in this area can be found in the works of Hou and Li, [21], [22], [47], and [24].



a. Triangular tessellation



b. Intermediate node compaction

c. Tight node compaction

Figure 6.2: Source node distribution for optimal spatial reuse

We seek the optimum density of transmitting nodes. As long as the transmission of each source does not interfere with the destination of another source the tessellation can be compacted. We assume that a source node will choose to transmit with equal probability to any node within its transmission radius so the probability of success of a contention will depend on the area within its transmission range that does not overlap the interference range of another transmitting node. We illustrate two different interference free areas in Figures 6.2b and 6.2c. A third configuration would have the source nodes clearly out of range of each other. We denote the separation distance between transmitting nodes in this model as *x*. We denote the area that is within the transmission region of a node but not within the interference region as $C(x,\alpha)$ and call it the interference free zone.

In [44] the authors consider the transmission areas to be isolated. As long as they do not intersect they will not interfere with each other. Our model, however, considers the effects of interference and capture. We assume that a destination is able to detect its source's transmission if it is closer to the source than the transmission radius and if no other transmitting node is any closer than α times the separation distance between the source and destination. We refer to α as the interference radius and its value is a function of the path loss conditions and the required power differential to be able to detect one signal over another. The parameters x and α then determine the area about a transmitting node where there is no interference, $C(x,\alpha)$. According to this model if a destination is in this region then it will receive the source's transmission. The method used to calculate the area of this interference free zone is described in Appendix 6.B.

Next we calculate the density of successful exchanges defined as the number of successful exchanges that occur in the hexagon formed by the six closest transmitting nodes to a given source on the tessellation. This density of successful exchanges is given by

$$S(x) = \frac{C(x,\alpha)}{\pi} \cdot \frac{3}{6 \cdot T(x)} = \frac{C(x,\alpha)}{\pi} \cdot \frac{2}{\sqrt{3} \cdot x^2} \tag{6.1}$$

The first term is the ratio of the area of the interference free zone to the total area covered by a transmission and thus corresponds to the probability of a successful contention. The second term is the density of transmitting nodes. There are three

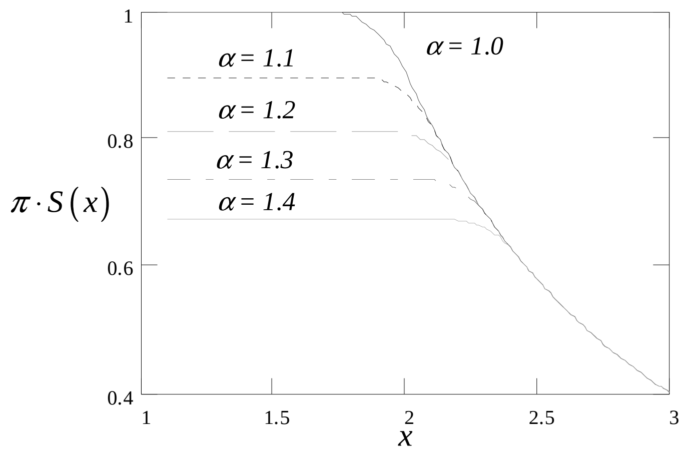


Figure 6.3: Success rate per transmission slot per transmission area based on transmitting node density

α	$\pi \cdot S(x^*)$	**Received Power Ratio Based on Distance Power Law Exponent**				
		2	**2.5**	**3**	**3.5**	**4**
1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.1	0.898	1.21	1.27	1.33	1.40	1.46
1.2	0.812	1.44	1.58	1.73	1.89	2.07
1.3	0.738	1.69	1.93	2.20	2.50	2.86
1.4	0.674	1.96	2.32	2.74	3.25	3.84

Table 6.1: Equivalent power ratios for the interference radius, α

transmitting nodes in the area covered by a hexagon formed by six equilateral triangles of area $T(x)$, 1 node in the center and then $\frac{1}{3}$ of a node contributed by each node at each of the six outside vertices. In Figure 6.3 we plot $\pi S(x)$, the expected number of successful exchanges per transmission slot per transmission area based on the separation distance between transmitting nodes. This value represents the spatial usage of the network, $U_A = \pi S(x)$. We have plotted this quantity for various interference radii, α. Table 6.1 translates this interference radius to relative received power ratios as a function of the power law exponents.[35] Note that in each case as the distance between transmitters decreases the expected number of successful exchanges per slot per transmission area converges to a constant. Eventually the increased density of transmitting nodes exactly counters the reduction in the area of the interference free zone. This value is the optimum spatial capacity that can be achieved with the corresponding interference ratio. Note that this model assumes the random selection of a destination. A tighter compaction is possible if there is some control over the S-D placement. In the limit the source and destination would be coincident and the transmitting nodes would be separated on the tessellation by the transmission radius. (We note that this would not be a practical arrangement of S-D pairs.) This distribution of S-D pairs yields the maximum spatial use and is defined as

$$U_A^* = \frac{2\pi}{\sqrt{3}} = 3.63 \ .$$

6.4 Simulation of Spatial Capacity

We simulated the SCR protocol to study the effects of node density, traffic load, and capture on performance. We normalize all metrics to a transmission radius of 1 unit such that the transmission area is π units squared. The simulation randomly places nodes on a 7×7 unit square grid at the density specified for the simulation run using a uniform distribution. Our simulation model considers the grid to be wrapped "spherically" so there are no boundaries. (A square grid is spherically wrapped around by considering all corners and opposite edges to be neighbors as illustrated in Figure 6.4.) The parameters of our simulation include the parameters of the collision

[35] Received power from a transmitter decreases as a power law function of distance. It varies from an exponent of 2 for free space propagation to more than 4 for a congested urban area. [16]

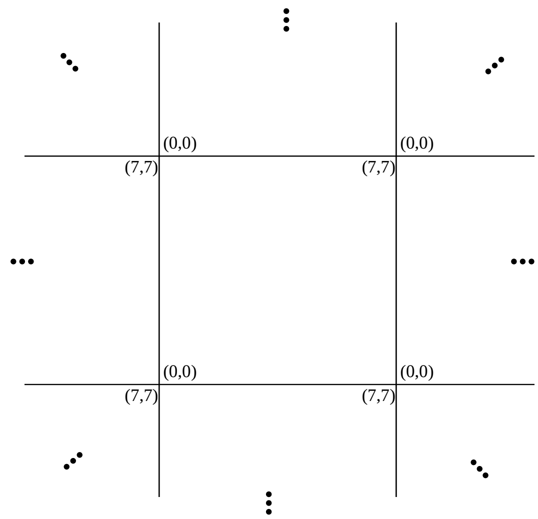


Figure 6.4: Spherically Wrapping a Square Grid

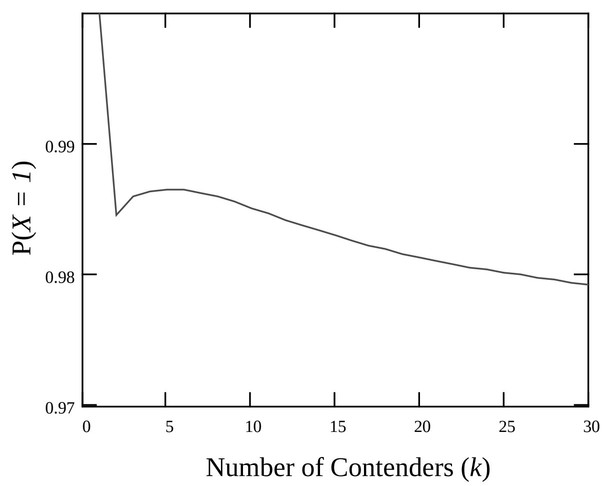


Figure 6.5: Access Probability for Simulation EYNPMA Parameters

5 priority levels, $H = 2$, $L = 10$, $M = 12$, $r = 0.87$, $q = 0.72$, $p = 0.86$

resolution signaling protocol, σ_A, λ_A, and the relative interference radius, α. We modeled arrivals as Poisson processes. The collision resolution signaling parameters that were used are: 5 priority levels (data priority levels), $H = 2$ (for each priority level), L

level), $L = 10$, $M = 12$, $r = 0.87$, $q = 0.72$, and $p = 0.86$.[36] These parameters for collision resolution signaling provide an expected success rate of better than 98% when the number of nodes contending for access is 25 or less, see Figure 6.5. We model each node as having a 6 packet buffer and set the packet lifetime to 48 transmission slots. Packets that are not transmitted before the end of their lifetime are dropped and packets that arrive when the buffers are full are dropped. Priority level in contention is assigned based on the age of the packet that will be sent where the oldest packets have highest priority. All simulation results are based on ten 1000 transmission slot simulation runs where each simulation uses a different random node placement.

We performed simulations at various node densities and spatial arrival rates for when $\alpha = 1.3$. Figure 6.6 provides the results of these simulations. We see from these results that the protocol's performance is excellent. In particular, as seen in Figure 6.6a the throughput remains stable[37] up to a load of $\lambda_A = 0.5$. Moreover and by contrast with other random access protocols, the throughput does not decrease with increased load or a larger number of contending nodes. The throughput, far exceeds the maximum spatial capacity of slotted Aloha and CSMA obtained in [44] and given by $U_A{}^*{}_{Slotted\ Aloha} = 0.36$, $U_A{}^*{}_{CSMA} = 0.45$.

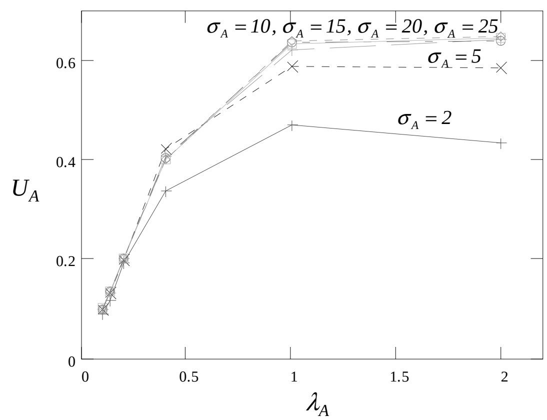
Figure 6.6b exhibits SCR's lack of sensitivity to increased node density. Initially, when node densities are small, there is lower performance. Investigation of our simulations revealed that the cause of the lower performance at smaller densities was the increased likelihood of nodes being disconnected[38] and the greater distances between adjacent nodes. At larger densities, however, the capacity is limited by the

---

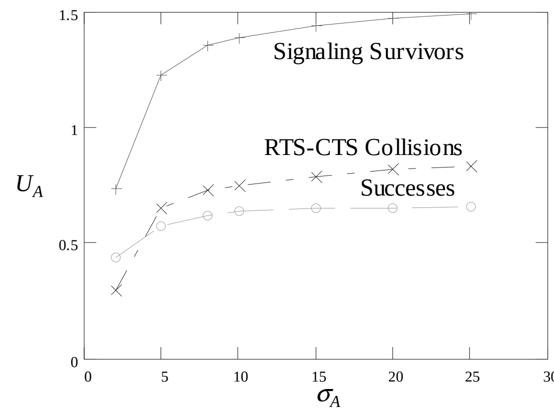
[36] Our simulation considered only best effort traffic. We broke the priority phase into 5 priority phases as done by HIPERLAN but assigned two slots to each priority group.

[37] Stable in this sense means that the spatial use was equal to the spatial arrival rate.

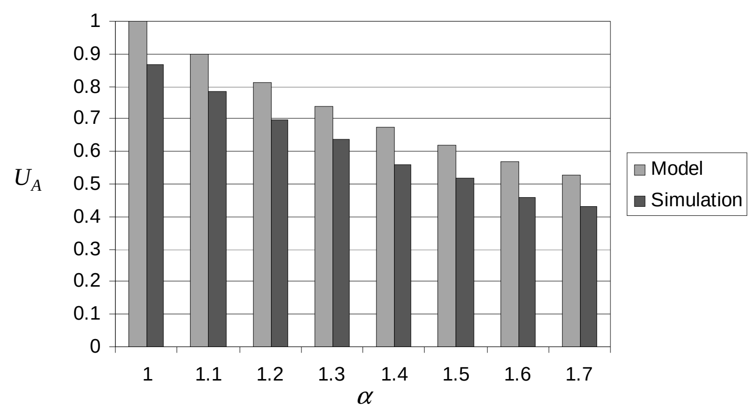
[38] A node is disconnected if it is further than one transmission radius from its nearest neighbor.



a. Spatial usage versus load, $\alpha = 1.3$



b. Spatial usage versus density, $\lambda_A = 2,\ \alpha = 1.3$



c. Spatial usage versus interference radius, $\lambda_A = 2,\ \sigma_A = 10$

Figure 6.6: Simulation results

interaction of nodes in the signaling process where the separation of contention winners is a function of transmission radius not node dispersion. As exhibited in Figure 6.6b, the knee of the signaling survivors graph occurs at a density of about 8 nodes per transmission area. This is the threshold where transmission radius not dispersion has the dominant effect. However, we see in Figure 6.6a, that even a smaller density of 5 nodes will achieve stable performance when the spatial arrival rate, λ_A, is 0.5 or less.

In Figure 6.6c we show the spatial throughput obtained via simulation to the optimum given by Equation (2). We note that the model is, in fact, fairly accurate at predicting the trend. The simulation results are consistently about 87% of the predicted model value. This result is quite impressive considering the model is based on the best possible distribution of transmitting nodes and is an indication of how well the collision resolution signaling resolves to an efficient set of S-D pairs. It also suggest strategies that may be used to improve the spatial capacity of SCR. Simply, any strategy that results in either increasing the probability a destination sends a packet to the interference free zone or technique that increases the size of the interference free zone (i.e. decreases the interference ratio) will increase the spatial capacity of the network. In the next few sections we investigate the use of such strategies, specifically the use of next hop policies to increase the probability of selecting a destination in the inference free zone and the use of spread spectrum coding to effectively increase the size of the interference free zone.

6.5 Improving Spatial Capacity Using Next Hop Routing Strategies

6.5.1 Nearest with Forward Progress (NFP) Next Hop Routing Strategy

Early work on spatial capacities focused on the simultaneous use of adjustable transmission radii and next hop strategies, see e.g. [45], [46], [47], [48], and [49]. It is clear from our simulation results that the number of nodes in a transmission area does not affect the probability that there is a successful contention in that region. But we also note in Figure 6.6b that a substantial number of nodes survive the signaling phase of the protocol but fail to successfully gain access on account of collisions during the transmission of the RTS and CTS. The objective of next hop strategies is to reduce collisions by increasing the probability that destinations are in the clear regions about each node. In particular, if the next hop is closer to the source then there is greater probability that the contention will not result in a collision. We now consider two next hop strategies. The first, nearest with forward progress (NFP), chooses the next hop based

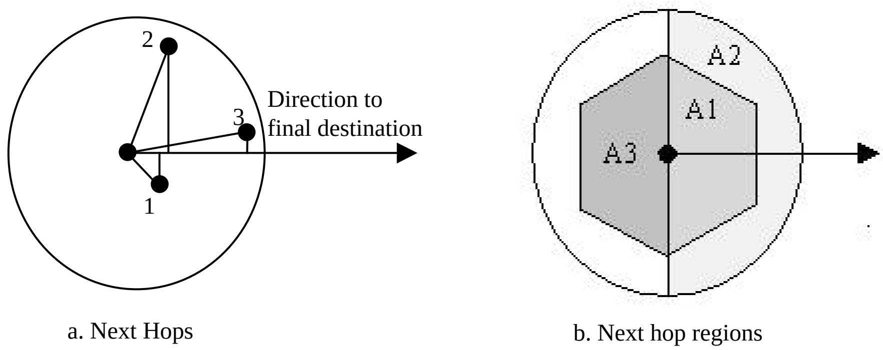


Figure 6.7: Next hops using the "Closest with Forward Progress" strategy

on the direction to the packet's final destination. The nearest node to the source that results in forward progress toward the final destination is selected as the next hop destination. Figure 6.7a provides an example of some next hop options. In this example, the next hop using NFP would be node 1 since it is the nearest node and results in forward progress toward the destination. Note, if there is no node with forward progress toward the destination then the next hop would simply be to the closest node. We develop an approximate model for the capacity of this strategy using Figure 6.7b. A node with a packet to transmit will first look for a destination in area A1, then area A2, and then third in area A3.[39] The transmission is successful if the destination chosen is in A1 or A3. We have assumed that nodes are distributed as a Poisson point process with spatial density σ_A so the probability that there is at least one node is in a given area is

$$P\left(N\left(A\right)>1\right)=1-e^{-|A|\sigma_A}$$

and no nodes in an area

$$P\left(N\left(A\right)=0\right)=e^{-|A|\sigma_A},$$

[39] There is a small probability that at the interface of these two regions that a node will be selected from area A2 before another node in A1 since it will be closer to the source. Since we are already making a gross assumption that transmitting nodes are placed on a triangular tessellation we see no benefit in making a more complex model that takes this into consideration.

$M \equiv \textit{Set of all nodes in node i's transmission area}$
$E \equiv \textit{Set of enclosing nodes}$

$E = \varnothing$
Sort M in ascending order based on distance from node i
$F = M$
$\textit{while } (F \neq \varnothing)\{$
$\quad h = f_1$
$\quad E = E \cup h$
$\quad F = F - h$
$\quad \textit{for each}(f \in F)\{$
$\quad\quad \textit{if } \left(d_{if}^{n} > d_{ih}^{n} + d_{hf}^{n} + c\right)$
$\quad\quad\quad F = F - f$
$\quad \}$
$\}$

Figure 6.8: Algorithm to determine a transmitting node's enclosing neighbors

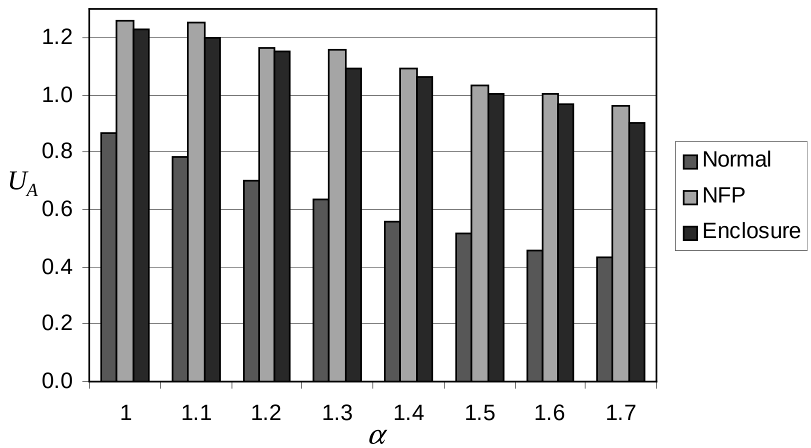
where $N(A)$ is the number of nodes in a region A where $|A|$ denotes the area of region A. We modify Equation (6.1) to account for the NFP next hop selection process

$$S(x) = \Big[\big(P(N(A1) > 0) + P(N(A1 \cup A2) = 0) \cdot P(N(A3) > 0)\big)\Big] \cdot \frac{2}{\sqrt{3} \cdot x^2} \;=\; \left(1 + e^{\frac{-\sigma_A}{2}}\right) \cdot \left(1 - e^{\frac{-C(x,\alpha)\cdot\sigma_A}{2\pi}}\right) \cdot \frac{2}{\sqrt{3} \cdot x^2}, \tag{6.2}$$

where a successful contention occurs when there is a node in area A1or there are no nodes in areas A1and A2 but then there is one in area A3.

6.5.2 Enclosure Hopping Next Hop Routing Strategy

A more efficient selection of a next hop might be based on the work in [18]. In this approach, the next hop is selected based on the energy consumed to transmit and receive a packet. The basic premise is that less energy is consumed in transmitting a packet to its final destination using multiple short hops than in using longer hops. Using this observation [18] shows that each node can be enclosed by a subset of nodes through which all packets should be forwarded if energy conservation is to be achieved. The criterion for an intermediate node to be selected as a next hop



6.9: Comparison of next hop policies,

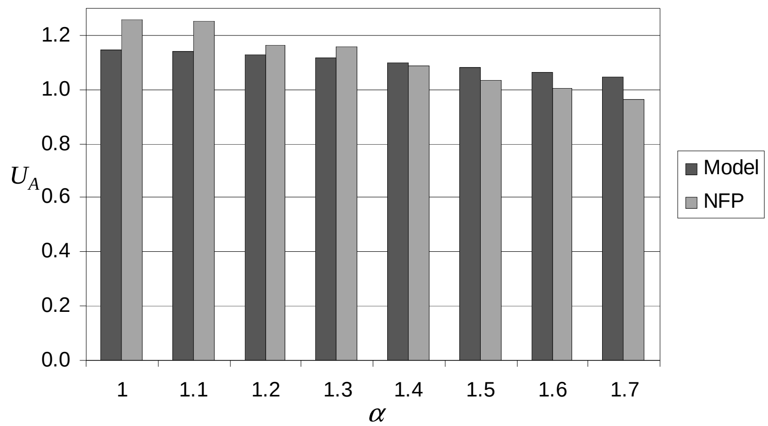
(spatial usage versus interference radius, $\lambda_A = 2$, $\sigma_A = 10$)candidate for another destination is whether it consumes less energy to relay the packet through the intermediate node than to transmit it directly to the final destination. Required transmission energy can be modeled as a function of distance. The intermediate node j is a potential intermediate hop for a packet destined for node k from node i if the following inequality is true

$$d_{ik}^n > d_{ij}^n + d_{jk}^n + c\,. \tag{6.3}$$

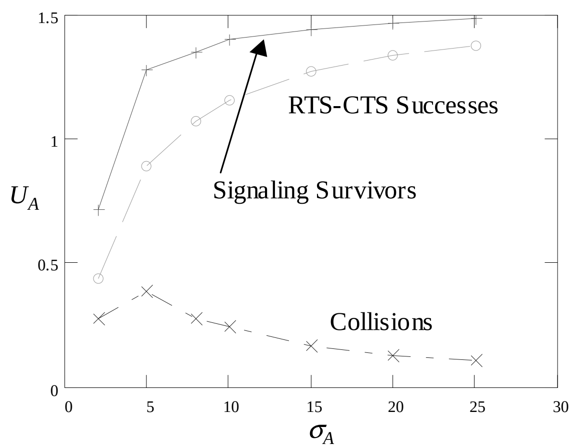
In this equation n is the exponent of the path loss model which may range from 2 for free space to greater than 4 in urban areas, [16], d_{ik} is the distance between nodes i and k, and c is a constant representing the power used by the intermediate node to receive and process a packet. The subset of nodes in a transmission area through which all packets are sent can be determined using the algorithm in Figure 6.8. We refer to this type of next hop selection as enclosure hopping.

6.5.3 Next Hop Routing Strategy Simulations and Results

Simulations were run using both policies. In the NFP hopping simulation, a direction was selected at random and then the node meeting the NFP criteria was



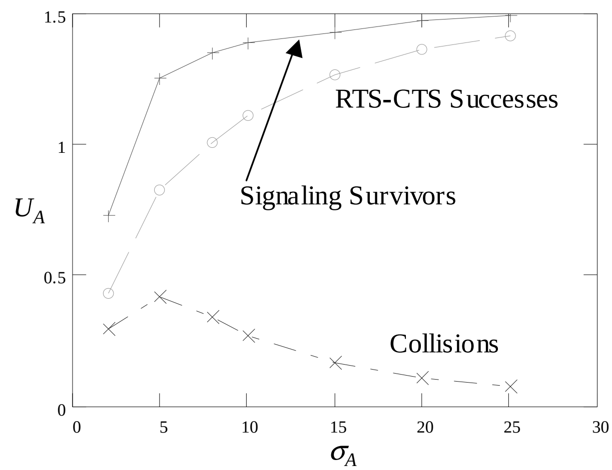
a. Spatial usage versus interference radius, $\lambda_A = 2$, $\sigma_A = 10$



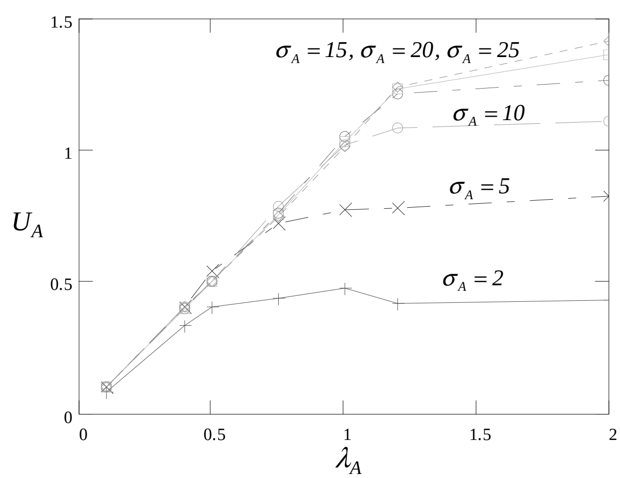
b. NFP spatial usage versus density, $\lambda_A = 2$, $\alpha = 1.3$

Figure 6.10: Performance of the NFP next hop policy

selected as the destination. In the enclosure hopping, a node within a transmission area is selected at random and then the node forming the enclosure that is closest to the destination is selected as the next hop. The enclosure nodes were determined using the algorithm in Figure 6.8 and Equation (6.3) with a path loss exponent of 4 and a reception constant equivalent to the energy required to transmit a packet a tenth of the transmission radius. Figures 6.9 through 6.11 illustrate the simulation results of the two next hop policies described in this section. We see that both next hop policies



a. Spatial usage versus density, $\lambda_A = 2$, $\alpha = 1.3$



b. Spatial usage versus load, $\alpha = 1.3$

Figure 6.11: Enclosure hopping spatial usage versus density, $\lambda_A = 2$, $\alpha = 1.3$

dramatically improve the spatial capacity of SCR. Figure 6.9 compares the performance of the next hop strategies to that of when normal hopping[40] is used. Both the NFP and enclosure hopping strategies dramatically increase the spatial capacity. Remarkably, their spatial usage, U_A, can be better than 1. Their performance is nearly twice as good as that of normal hopping when the interference radius is large. In

Figure 6.10 we compare the simulation performance of the NFP next hop policy to that of the model defined in Equation (6.2) when the separation distance, x, is 1.[41] Figures 6.10a and 6.10b illustrate the spatial usage as a function of node density. Comparing these results to those for normal hopping illustrated in Figure 6.6b we see that the number of survivors of the signaling portion of SCR is the same for all the next hop strategies but that the NFP and enclosure next hop strategies are exceptionally effective at reducing collisions. Their effectiveness at decreasing collisions increases with the density of nodes because of the greater probability that the next hop will be closer to the source when the density is higher.

In Figure 6.11b we illustrate the limits of stability when enclosure hopping is used. For densities $\sigma_A \geq 10$ the network remains stable to an offered spatial load exceeding 1.

As mentioned in Section 6.3, the rate packets make forward progress is often used as a measure of spatial capacity. We define the quantity d_{FPA} as the average total forward progress of all the packets sent in a transmission area per transmission slot. To determine this measure we determine the average progress each transmitted packet makes towards its final destination, d_{FP}, and multiply it by the usage,

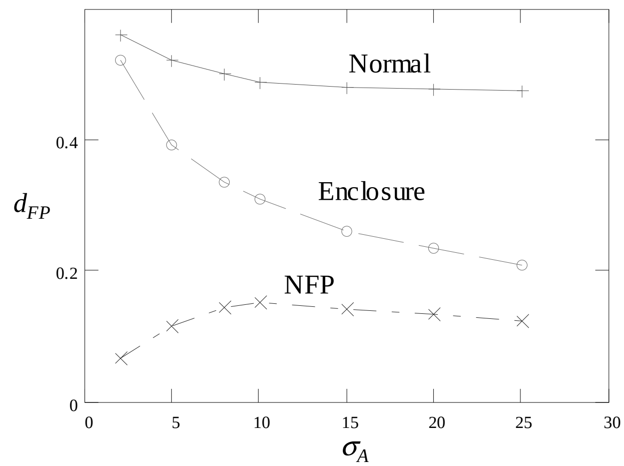
$$d_{FPA} = d_{FP} \cdot U_A \,.$$

Figure 6.12 illustrates that there is not much difference in the forward progress measure of the normal next hop and enclosure next hop strategies in a congested network. The benefit of using the enclosure next hop strategy is not truly an increased throughput. The enclosure next hop strategy only slightly outperforms the normal next hop strategy for node densities, σ_A, in the range of 3 to 20 nodes per transmission area.

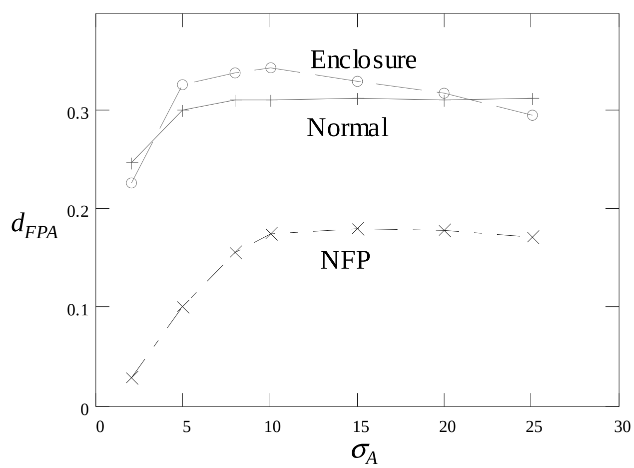
---

[40] Normal hopping is simply the random selection of a destination within range of the source.

[41] This separation distance is significant since it is further evidence that the collision resolution signaling is especially effective at resolving to a distribution of non-interfering nodes. A separation distance of 1 on a triangular tessellation is an ideal result for the signaling protocol.



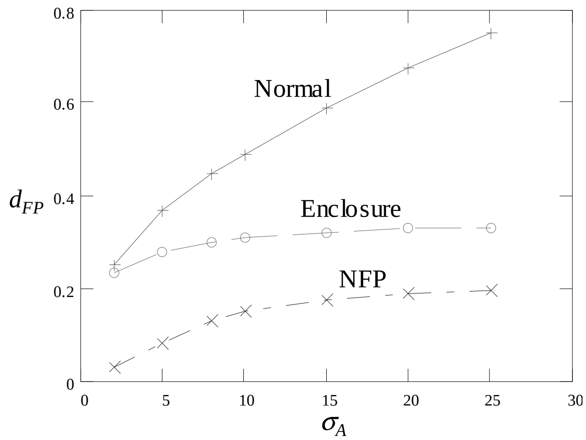
a. Forward progress per transmission



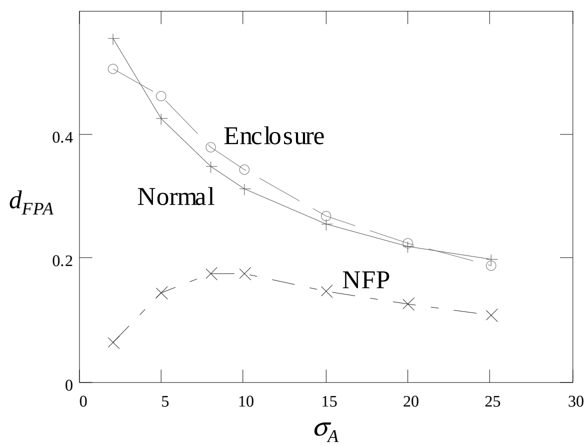
b. Forward progress per slot per transmission area

Figure 6.12; Packet progress as a function of packet density in a transmission area, $\alpha = 1.3$, $\lambda_A = 2$ (Transmission radius remains constant and density increases)

The comforting observation is that using enclosure hopping to help conserve energy does not compromise throughput. Figure 6.13 provides another view of the same data. In Figure 6.13 we have assumed that the spatial distribution of nodes is fixed and that we adjust the transmission radius to achieve the desired density, σ_A. The expected result is that with larger transmission radii the forward progress per transmission will increase but the total progress in the network will decrease since fewer



a. Forward progress per transmission



b. Forward progress per slot per π square units area

Figure 6.13: Packet progress as a function of adjusting transmission radii, $\alpha = 1.3$, $\lambda_A = 2$

(Density remains constant and transmission radius changes, $\sigma_A = 10$ at $radius = 1$)

nodes will transmit during each slot. This is indeed the case for normal hopping. Figure 6.13a, however, reveals that for enclosure hopping forward progress per successful contention remains nearly constant starting at a density of about 10 nodes per contention transmission area. This is an indication that most of the enclosing nodes are included within a radius that achieves this density of nodes in a contention transmission area.

6.6 Improving Spatial Capacity Using Spread Spectrum Coding

6.6.1 Related Work in Using Spread Spectrum Codes

We are proposing that spread spectrum (SS) coding can be used to improve the spatial capacity of SCR. The use of SS codes is not new to the discussion of improving the spatial capacity of wireless networks. SS technology has frequently been proposed as a method to multiplex signals within the same geographic region.[42] [50] The basic approach is to use either unique SS codes or time shifted sequences for different links so that multiple transmissions can coexist spatially. [51] If successful, this effort results in removing the spatial dependence of node-to-node links on each other. Unfortunately, several complications are introduced by the distributed nature of ad hoc networks. The first question is which code to use and when? Synchronizing the codes transmitters use to those that receivers are using to receive is not trivial. Second, in a channel with a finite number of codes available the network needs to assign codes such that they are spatially distributed. In a dynamic network, keeping track of codes can be as complicated as keeping track of routes. Third, unlike CDMA cell phone systems there is a likelihood that an interfering transmitter will be closer to a receiver than its source transmitter. Even with a different spreading sequence the relative power of the interfering signals can raise the effective noise level such that the desired signal is lost. Finally, when codes are used, networks become packet sensing as opposed to carrier sensing. Receivers must synchronize with the transmitter from the beginning of the packet transmission or else the transmission appears as noise. To avoid interfering with an ongoing reception at a destination a source must also receive the same packet the destination is receiving.

[42] Spread spectrum offers other very attractive features such as resistance to detection, resistance to multipath induced fades, and resistance to jamming.

Most research in the use of SS in ad hoc networks has been directed toward the selection of codes. There are four basic approaches to selecting codes: transmitter-based, receiver-based, pairwise oriented, and common code shifted in time. In transmitter-based schemes a unique code is assigned to each transmitter. The resultant problem is obvious. Which transmitter's code should a potential receiver use at a given time. In receiver-based schemes a unique code is assigned to each receiver. In this approach, there is no confusion at either the receiver or the transmitter as to which code to use but there is the question as to when the transmitter should start transmitting. Transmitting to a destination that is already receiving a packet from another source can interfere with that reception. We note that with transmitter-based codes the problem associated with an adjacent node interfering with a destinations reception is not as severe since it is precluded from using the same code. In pairwise oriented schemes unique codes are assigned to pairs of nodes. Pairwise coding presents the same challenge to destinations as transmitter-based codes. Each potential destination needs to determine which code it should use in receiving. Common code systems rely on the offset of the phase of codes used in transmissions that occur concurrently. The offsets may either occur randomly or be deliberate. In the latter case, the selection of the offset is no different than the selection of the SS code in the first three approaches. Additionally, despite the presence of an offset when using common codes there is still an increased likelihood that signals may interfere with each other. However, using a common code has an advantage in that it is the only approach that can support broadcast transmissions.

Proposed solutions to the problem of selecting codes normally involve the use of hybrid protocols. Two hybrids are considered in [52]. In the first, all nodes monitor and contend using a common code but after the addresses of the source and destination are transmitted the transmitter uses a transmitter-based code to send the data packet. The destination knows which code to use since it receives the source address before the transition to the use of the transmitter-based code. In the second protocol, every node is assigned two codes, one to receive on and one to transmit on. When a node is idle it listens using its receive code. A source will use the receive code of the destination of a packet to send the source and destination address portion of the packet but will again transition to the source's transmit code to send the data portion. Both approaches attempt to limit the time that another node may interfere with the exchange. In the first, other nodes will only interfere during the initial transmission of the addresses and in the second the potential interference occurs during the same period but is further reduced on account that interfering nodes must be transmitting to the same destination. The disadvantage of the second protocol over the first is its inability to broadcast packets.

The assignment of codes is also a problem in SS ad hoc networks. Normally there are a finite number of codes that must be distributed to a larger number of nodes. Random selection of codes may not assure the required physical separation of nodes using the same codes. Various algorithms have been developed for the assignment of codes. A review of several algorithms is presented in [53]. Interestingly, this paper concludes that pairwise code assignment requires the fewest codes. Other methods of assigning and distributing codes are associated with the hierarchical organization of networks. In [54], the authors present a method of organizing a network into clusters with unique SS codes assigned to each cluster.

Code assignment does nothing to prevent the unwanted interference from transmitters in close proximity of the destination end of other exchanges. We are aware of no work that attempts to schedule exchanges in an effort to avoid this interference. In [55] the authors address this problem in a completely different manner. Rather than orchestrate where transmitting nodes are located they attempted to determine the transmission range of nodes that limits the cumulative effect of this type of interference. The practical application of this type of model is limited by the relevance of its modeling approach of the physical distribution of nodes and of the arrival rates of packets. Ideally, an access protocol would deliberately select transmitting nodes that do not interfere at the destinations.

Although there has been much work on the use of SS in ad hoc networks there are no protocols that completely integrate within a medium access control protocol the assignment of codes, the scheduling of a spatially distributed set of transmissions, and a method used by nodes to select codes for transmission and reception. In this section we introduce the use of SS codes with SCR. The two protocols are perfect complements for each other. SCR is very effective at resolving a set of transmitting nodes that are spatially separated from each other thus mitigating interference. Its use of signaling allows nodes to identify which codes they should use prior to the transmission of data. Then, the use of these codes enhances SCR's already high spatial capacity. To complete the picture we provide an additional protocol for the selection and dissemination of SS codes.

In Section 6.6.2 we discuss the issues of interference in SS packet radio networks and explain how SCR and SS are mutually beneficial to each other. In Section 6.6.3 we discuss the mechanisms used to disseminate and to identify the codes for data transmission. In Section 6.6.4 we present our simulation results of the concurrent use of the two protocols in a static network and demonstrate the superior performance that is achieved. Then in Section 6.6.5 we present our simulation results of the concurrent use of the two protocols in a dynamic network demonstrating the protocol's ability to adapt to changing codes.

6.6.2 Capture In Spread Spectrum SCR Networks

Capture is a phenomenon that can be used to enable multiple destinations to receive different signals in the same geographic region using the same spectrum simultaneously. As a simple model of the effect we say that a destination receives a signal if the energy of the signal exceeds the background noise by some level. This signal to noise level is a function of the energy used to transmit the signal, the distance that separates the source from the destination, and the background noise level. In a network with a single transmitting node the ratio can be expressed as

$$SNR = \frac{P_t / d^n}{N} \tag{6.4}$$

where P_t is the effective radiated power from a transmitter, N is the background noise power, d is the distance that separates the source from the destination, and n is the path loss exponent. As demonstrated, the power of a received signal is a power law function of the distance that separates the source from the destination. The path loss exponent n is a function of the environment ranging from 2 for free space to over 4 in urban areas.[16][43] In networks where multiple transmissions can occur concurrently the same model is appropriate but is modified as follows:

$$SIR_j = \frac{P_t / {d_j}^n}{N + \sum_{i \neq j} P_t / {d_i}^n} \, . \tag{6.5}$$

In this model we assume that the signals of interfering nodes are seen by the destination as noise and that the energy of the interfering signals is also a function of the separation distance between the receiving destination and the interfering sources.

[43] In practice, the value of 4 is most often used in analysis using the path loss model with ad hoc networks.

Note that we also assume that all sources transmit using the same power.[44] A destination will receive its intended packet if it is closer to the packet's source than it is to the interfering nodes and if the threshold signal noise ratio is exceeded. The advantage of using SS is the apparent reduction of the energy of the interfering signals. This energy reduction is a function of the processing gain, *PG*, and Equation (6.5) becomes

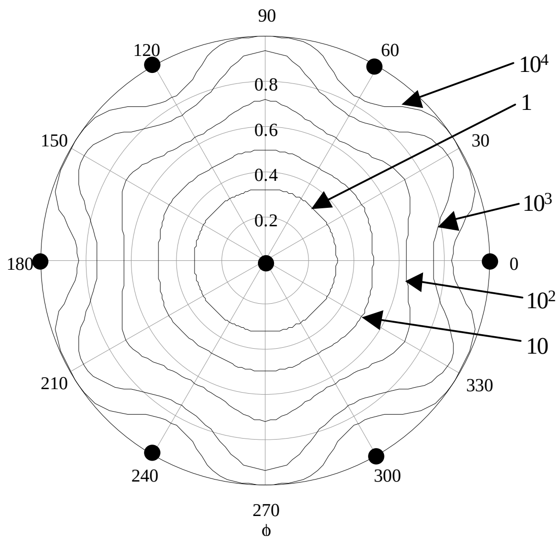
$$SIR_j = \frac{P_t / d_j^{\,n}}{N + \frac{1}{PG} \sum_{i \neq j} P_t / d_i^{\,n}} \; . \tag{6.6}$$

As seen from this model, the use of SS codes does not insure correct reception of transmissions, it only improves conditions. The capacity of the network is dependent on the distribution of transmitting nodes with respect to destinations.

The collision resolution signaling mechanisms in SCR achieves a spatial distribution of nodes that will transmit. Ideally, the winner will suppress all contenders within its transmission range, the range at which the threshold SNR is achieved using Equation (6.3). We call the geographic area covered by this transmission range the transmission area. The tightest compaction of transmitting nodes occurs with all sources on the vertices of a triangular tessellation made up of equilateral triangles with sides equal to this transmission range. See Figure 6.2a. In this configuration, all sources are separated by exactly the transmission range. The number of source nodes per transmission area is 3.63. This is the theoretical limit to the density of signaling survivors. We use this distribution of transmitting nodes and show the required processing gain to assure successful transmission to a destination in Figure 6.14a. The

---

[44] We assume that all nodes using SCR contend using the same power level but may adjust after contention based on the measured energy levels of the RTS-CTS exchange.



a. Interference contours about a node in a triangular tessellation of transmitting nodes

Processing Gain	Coverage
1	0.0996
10	0.2333
100	0.4379
1000	0.6855
10000	0.8617

b. Contour coverage of the transmission area

Figure 6.14: The effect of processing gain on an ideal compaction of transmitters

contours of Figure 6.14a are associated with different processing gains used in Equation (6.6) where the required SNR to capture a signal is 10. We assume that so long as the destination is within the area encircled by the contour of the processing gain that is used, it will capture the signal and the data exchange will be successful. The percentage of the transmission area in which destinations will capture a data exchange as a function of processing gain is listed in the table of Figure 6.14b.

6.6.3 Selecting and Using Spread Spectrum Codes

Two types of code assignments are used in SCR, a common code for signaling and broadcasting and receiver based codes for peer-to-peer exchanges. Nodes can distinguish between broadcast and peer-to-peer transmissions by whether the broadcast priority is used in gaining access. A node that fails to win a contention will either start listening using the broadcast code if it detects that the broadcast priority slot had been used to gain access or will use its own SS code otherwise. A node that wins the contention will use either the broadcast SS code if it won the contention using the broadcast priority or will use the SS code of the receiver to which it is transmitting the packet.

This method of selecting and distributing codes also allows individual nodes to multiplex multiple packets in a single transmission (i.e. concurrently transmit) to multiple destinations. Nodes that survive the contention signaling using a peer-to-peer priority may multiplex the signals using the SS codes of the destinations. The multiplexed RTS transmissions that follow the signaling would direct each destination as to which code to use subsequently when sending a CTS or an ACK.

Spread spectrum codes are self-assigned. Nodes make periodic broadcasts to advertise their codes. In these broadcasts, nodes advertise their codes and their neighbors' codes. A node selects a code that is not being used by any of its neighbors or any of its neighbors' neighbors. In this way, transmissions to a destination will use a code that is the only code used within the transmission region of the destination. If a node has selected a code and then detects that another node has moved within the two-hop region that is using the same code, then the node that detects the conflict will select another code that is not being used and then broadcast it to its neighbors. A node does not start using the new code until it has successfully broadcasted the change to its neighbors. This method of assigning codes will allow a node to select a unique code so long as there are more codes than nodes within the node's two hop region. If there are not enough codes and a node detects another node using its code the detecting node still changes its code by randomly selecting a code from those used. We evaluate the effect of this option on blocking in subsequent simulations.

6.6.4 Simulation of the Combined Use of SCR and Spread Spectrum Coding in a Static Network

We simulated the SCR protocol to study the effects of node density and processing gain on the density of transmissions. Our simulation environment and SCR

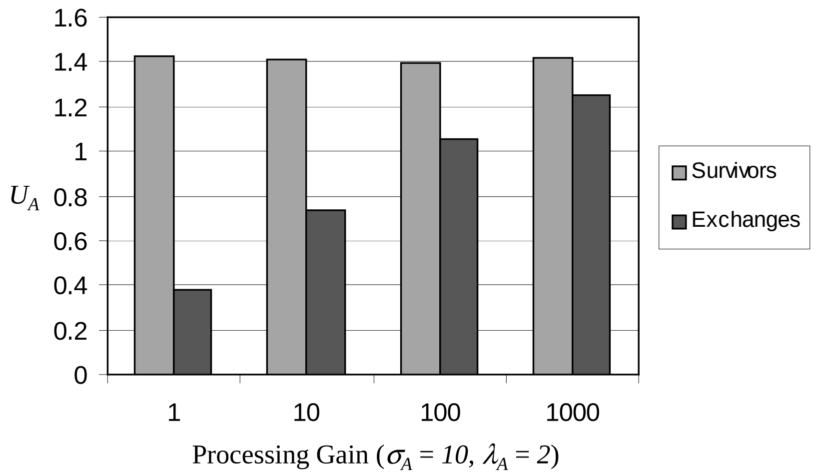
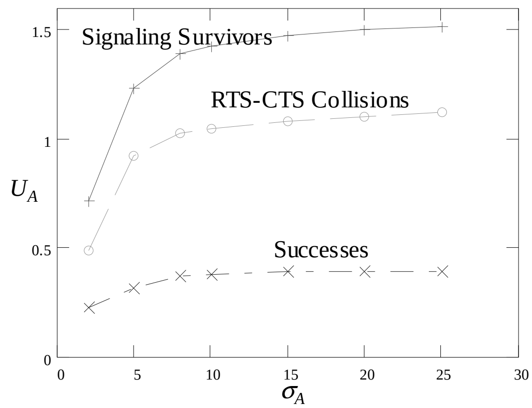


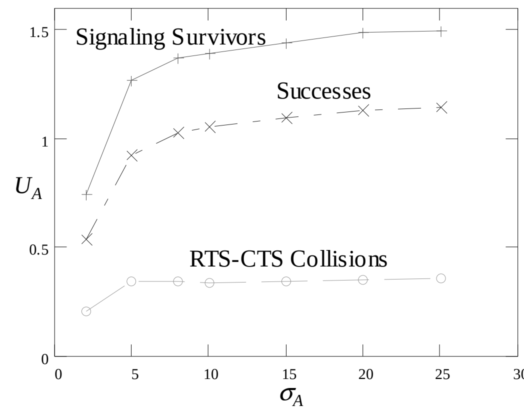
Figure 6.15: Spatial capacity as a function of processing gain

parameters are identical to those used in the simulations of Section 6.4. However, we use the interference models of Equations (6.4) through (6.6). Our metrics for a transmission radius of 1 unit corresponds to the threshold separation distance where nodes meet the SNR requirements as defined by Equation (6.4).

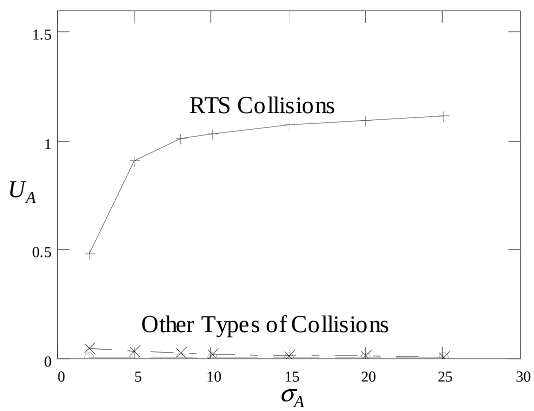
The simulation operates as follows. Nodes with traffic to send attempt to contend every transmission slot until there is no more traffic in their queues. Each contention begins with signaling and if a node is a survivor, it then confirms success with the RTS-CTS exchange as described in Section 5.3.2. The destinations of these transmissions are selected randomly from among all nodes that are within range. The range of a signaling node is defined by Equation (6.4). We assume all nodes that receive the signaling nodes signal with a SNR of 10 or more will be affected and will react according to the rules of the signaling protocol. We assume that the capture that occurs during the RTS transmissions is governed by Equation (6.5) unless multiple sources are attempting to transmit traffic to the same destination in which case Equation (6.4) is used to determine the SNR since common spreading codes would be used. Finally, we assume that the capture that occurs during the CTS transmissions is governed by Equation (6.5). In addition to the RTS and CTS failures our simulation



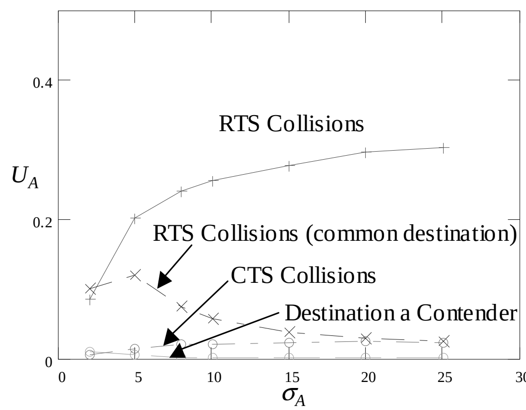
a. Contention results with *PG = 1*



b. Contention results with *PG = 100*



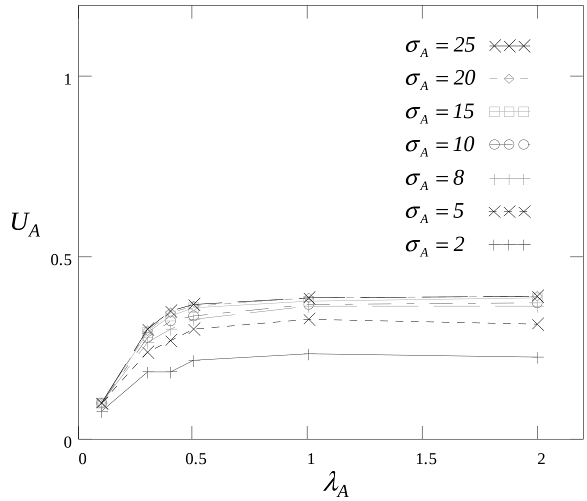
c. Collision breakout *PG = 1*



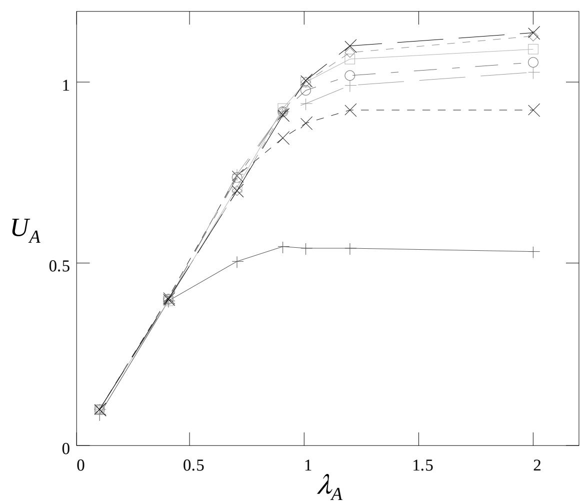
d. Collision breakout *PG = 100*

Figure 6.16: Comparison of contention results for different processing gains and different node densities ($\lambda_A = 2$)

also identified the rare occurrences (<0.3%) when sources select destinations that also survive the contention signaling. (i.e. The destination also transmits an RTS.) Figures 6.15 through 6.17 illustrate the results of the simulations. First, Figure 6.15 illustrates the effect of processing gain on the spatial capacity. (For these results we assumed a spatial density of 10 nodes per transmission area and a spatial arrival rate of 2 packets per transmission slot per transmission area, an overloading condition.) For all simulations the number of survivors of the signaling phases of the contention are approximately equal but the number of survivors of the RTS-CTS handshake is



a. Spatial capacity with *PG = 1*



b. Spatial capacity with *PG = 100*

Figure 6.17: Comparison of spatial capacities for different loads, node densities, and processing gains

greatly affected by the processing gain used. Increasing the processing gain from 1 to 10 nearly doubles the spatial capacity while increasing the processing gain to 100 almost triples the spatial capacity. Figure 6.16 illustrates the effect of node density on the performance of the SCR and compares the performance of SCR with a processing gain of 1 to that of SCR with a processing gain of 100. As is dramatically illustrated, improving the processing gain improves the total number of successful contentions. We note also that the density of nodes tends to improve the spatial capacity up to a point when this improvement levels off. At low densities the throughput is much less. The cause of this decrease is the increased tendency of nodes to suppress each other in contention. Figures 6.16c and 6.16d break out the causes of failures of signaling survivors to gain access. When spread spectrum coding is not used, the dominant cause of contention failures is RTS collisions. When SS is used, the RTS collisions still cause the largest number of failures but a much greater percentage of these failures occur since signaling winners frequently choose the same destination. The greater significance of these failures results from the fact that they occur despite the use of SS coding since receiver directed codes are used. Finally, Figure 6.17 illustrates the simultaneous effect of load and node density on spatial capacity and again compares the performance of a network that does not use SS to one that does. Again SS coding is shown to dramatically improve the spatial capacity of the network.

6.6.5 Simulation of the Combined Use of SCR and Spread Spectrum Coding in a Dynamic Network

Next we simulated the combined use of SCR and SS coding in a dynamic network where there are a finite number of codes and all nodes move. Our objective was to determine the effect of code availability and concurrent use of codes on the occurrence of common code failures. (Common code failures occur when two nodes transmit to two different destinations using the same SS code and as a result one or both of the exchanges fail.) Our simulation environment is much different than that in Section 6.6.4. Here we consider a square network region that is 8 transmission lengths on a side. Nodes in this network constantly move at the same speed but in different directions. The simulation starts by randomly placing the nodes on the network region. Each node then selects another point in the network region and moves to that location at the designated speed for the simulation. Once a node reaches a destination point, it immediately chooses another point at random in the network region and then moves toward it. On account of node movement and the finite number of codes used in this simulation nodes are required to keep track of their neighbors codes and to change their own as the need arises. The dissemination of codes can be integrated with the routing protocol. Nodes can disseminate code information together with routing information. Assuming all nodes advertise their connectivity to their next hop neighbors, all nodes ultimately learn the codes of all nodes within two hops of themselves. On account of these updates, each individual node learns the codes used by nodes up to two hops away. Due to the fixed size of packets and the size of the networks we used in our simulations nodes cannot advertise their full routing tables in a single transmission. Rather, we limited each routing table transmission to only nine entries. The rate at which tables are advertised is a function of when they were last advertised, how far they are away from the advertising node, and how critical the information is. In this case, the identification of a changed SS code is considered critical and is advertised immediately. Otherwise, nodes advertise their SS codes and those of their neighbors at a fixed rate. Nodes that identify a conflict with a code that they are using attempt to find a new code that is not in use. If successful, the node first advertises the new code in a routing table transmission and only then uses the new code as its own. In the case that all available codes are being used the node randomly selects a new code and starts using that new code after it is advertised despite its simultaneous use by another node in the two hop region. The situation where there are too few codes for the density of nodes can result in the repetitive transmission of routing table broadcasts precluding the transmission of payload packets. To curtail this type of congestion collapse we defined a parameter called the “minimum time between updates” (MTBU) that prevents nodes from sending updates all the time.

Bit Rate	Radio Range (meters) 200	250	400	1000
1000000	1.30	1.63	2.60	6.51
2000000	2.60	3.26	5.21	13.02
11000000	14.32	17.90	28.05	71.61

Table 6.2a: Node velocities in kilometers/hour

(0.00001 of transmission range per transmission slot)

Bit Rate	Radio Range (meters) 200	250	400	1000
1000000	6.51	8.14	13.02	32.55
2000000	13.02	16.28	26.04	65.10
11000000	71.61	89.52	143.23	358.07

Table 6.2b: Node velocities in kilometers/hour

(0.00005 of transmission range per transmission slot)

We used a speed parameter for the nodes that is normalized to the duration of a transmission slot. We assume that the fixed size packet that is transmitted is 512 bytes long and that the signaling and acknowledgement together take 35% of the time that it takes to transmit a packet. The speed is then defined as the fraction of a transmission length that a node moves in a single transmission slot. We considered two speeds in our simulations, 10^{-5} and 5×10^{-5} transmission lengths. Tables 6.2a and 6.2b list the possible land speeds these normalized speeds correspond to depending on the bit rate and the range of the transceivers.

Tables 6.3a and 6.3b display the frequency at which common code use caused the blocking of a transmission as a percentage of the number of data packets that are successfully exchanged.[45] These results are for simulations that each lasted 50,000 transmission slots. These are presented for different node state update rates, different

[45] The number of data packets used for this ratio does not include the node state broadcasts. The number of contending nodes that were not successful in sending a packet on account of simple interference is also not included in this ratio.

	Update Rate / MTBU (Transmission Slots)	2000 / 32	3000 / 50	5000 / 50	2000 / 32	3000 / 50	5000 / 50
	Codes	**10**	**10**	**10**	**25**	**25**	**25**
Speed	**Spatial Arrival Rate (Packets/Slot/Transmission Region)**						
0.00001	**0.05**	0.485%	0.543%	0.428%	0.611%	0.517%	0.391%
0.00001	**0.1**	0.508%	0.447%	0.524%	0.687%	0.576%	0.423%
0.00001	**0.2**	0.633%	0.634%	0.633%	0.978%	0.704%	0.728%
0.00005	**0.05**	0.855%	0.787%	0.783%	0.423%	0.442%	0.262%
0.00005	**0.1**	0.847%	0.821%	0.937%	0.564%	0.426%	0.343%
0.00005	**0.2**	0.928%	0.774%	0.973%	0.623%	0.477%	0.370%

Table 6.3a: Percentage of contentions that fail on account of nodes using common codes ($\sigma_A = 5$)

	Update Rate / MTBU (Transmission Slots)	2000 / 32	3000 / 50	5000 / 50	2000 / 32	3000 / 50	5000 / 50
	Codes	**10**	**10**	**10**	**25**	**25**	**25**
Speed	**Spatial Arrival Rate (Packets/Slot/Transmission Region)**						
0.00001	**0.05**	0.446%	0.445%	0.449%	0.228%	0.197%	0.288%
0.00001	**0.1**	0.463%	0.462%	0.485%	0.251%	0.253%	0.284%
0.00001	**0.2**	0.568%	0.551%	0.534%	0.270%	0.315%	0.317%
0.00005	**0.05**	0.549%	0.526%	0.448%	0.333%	0.232%	0.235%
0.00005	**0.1**	0.522%	0.539%	0.471%	0.366%	0.514%	0.279%
0.00005	**0.2**	0.535%	0.517%	0.601%	0.352%	0.296%	0.450%

Table 6.3b: Percentage of contentions that fail on account of nodes using common codes ($\sigma_A = 10$)

speeds and different loads. Four different experimental parameters were considered in this comparison, speed, spatial arrival rate, number of codes, and the rate that routing table updates are distributed throughout the network. We use independent random number generators for movement, arrival of packets and the selection of packet destinations. As a result, simulations with common speeds have nodes that move in the exact same way. For simulations with common arrival rates, packets arrive at nodes at the exact same times and are directed to the exact same destinations. Note that when a packet arrives at an individual node its destination may be any other node in the network. Thus the arrival of a single packet may require multiple transmissions to traverse the network. The results reveal that only a fraction of one percent of the contentions fail on account of two transmitters sending packets to destinations that use the same code. This is true even when there is a great shortage of codes. These tables consider two different quantities of codes, 10 and 25. When the node density is 5 the average number of nodes in a two-hop region is 20 and when the density is 10 the average number of nodes in a two hop region is 40.[46] Nevertheless, the SCR protocol remains effective at distributing transmitting nodes and the probability that two transmitting nodes select two different destinations with the same code remains small even with just a few codes available. In fact, the density of nodes has almost as much effect on the percentage of common code failures as does the availability of codes. Our explanation for this observation is that with a higher density of nodes, more nodes contend with each other and so it is less likely that a set of contentions that failed on account of common code use would occur simultaneously again. So, common code interference is rare, even in highly dynamic networks using a small number of codes.

6.7 Adaptive Networking

SCR has all the characteristics that network designers should like. Adjusting the transmission radius used for contention easily modifies the protocol's performance. As demonstrated in Sections 6.4 though 6.6, the density of nodes and the spatial arrival rate determine the performance of SCR. The contention transmission radius affects both of these factors. Reducing the radius decreases the spatial density of nodes

[46] The node densities are not uniform in the network on account of the movement model. Since nodes move to other randomly selected points in the simulation region by a direct path there will tend to be a greater density of nodes near the center.

and, in turn, the spatial arrival rate. The objective in adjusting the radius is to achieve a stable spatial arrival rate while simultaneously maintaining the connectivity required to keep the network connected. Figure 6.13b illustrates that there is an additional incentive to keeping the contention transmission radius small. The smaller the radius, the greater the forward progress per transmission slot when the network is congested. But our simulations also reveal that at low node densities there are many nodes that are disconnected resulting in dropped packets. So to select a transmission radius in a network using SCR we offer the following heuristic. Start by selecting a contention transmission radius that contains the neighboring nodes that would enclose each node. This is the minimum radius that should be used. Then increase the radius if it can be increased while keeping the spatial arrival rate in the stable regime. Adding the use of SS coding to SCR will enhance the capacity of any radius selected.

6.8 Other Capacity Enhancement Techniques

Recall that the signaling of the collision resolution access mechanism that SCR uses to resolve contentions also resolves a set of surviving nodes that are spatially separated. On account of this separation of the transmitting nodes, each transmitting node may assume that any node within its range could receive traffic. These transmitting nodes may then use code division multiple access (CDMA) or space division multiple access (SDMA) techniques to allow the simultaneous transmission of data to multiple destinations. The methods used to select SS codes for data transmission perfectly complement the CDMA approach. The SDMA approaches, however, require some coordination, at least to identify the directions to destinations. Assuming the routing protocol disseminated location information, no additional coordination may be required.[47] For both options the RTS-CTS handshake validates the ability of destinations to receive the traffic. Subsequent interference during data transmissions can only decrease since nodes can only drop not add destinations.

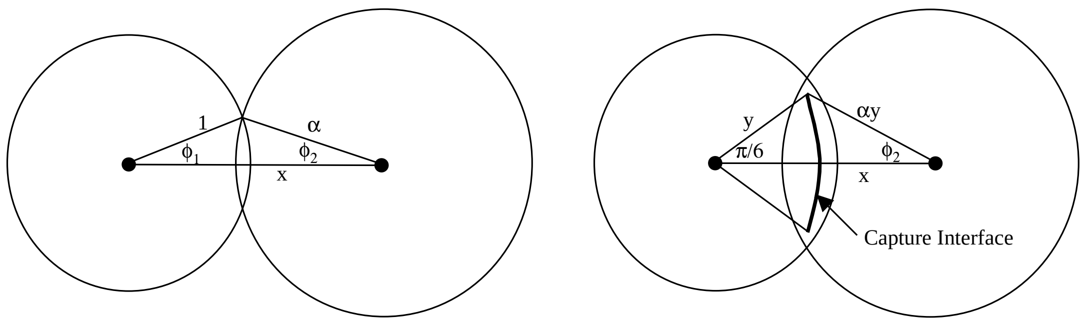
6.9 Conclusion

In this chapter we developed a model for the spatial capacity of SCR that provided an intuitive view of the factors that affect the spatial capacity. We used this intuition to consider two techniques to improve spatial capacity, the use of next hop routing policies and the use of spread spectrum coding. We demonstrate that both are easily employed with SCR and that they demonstratively increase the capacity of the protocol. We further describe how the adjustment of transmission ranges and the use of multiplexing techniques can also improve performance. Simply put, SCR has tremendous capacity that is easily exploited.

[47] Smart antenna approaches where signals are used to identify spatial signatures could not be employed in this manner. Acquiring spatial signatures requires the separate transmission of a signal from each of the destinations.

Appendix 6A

Calculating the Interference Free Zone



a. $\phi_1 < \frac{\pi}{6}$ b. $\phi_1 \geq \frac{\pi}{6}$

Figure 6A.1: Calculating the Capture Interface for the Interference Free Zone

In this appendix we derive the calculation for the area of the interference free (IF) zone of a transmitting node in a triangular tessellation, $C(x,\alpha)$. We define the clear zone as the region that a destination can receive a transmission from a source node without interference from an adjacent node that is also transmitting. We assume FM capture such that if the destination node is further than α times the distance to the source node from all other transmitting nodes then it is in the IF zone. We also assume that the maximum transmission radius for a node is 1 such that if the transmitting nodes are separated by at least 1 + α, then they will not interfere with each other. There are three different regions with different equations for calculating the area of the IF zone. The first region is when the nodes are further than $1 + \alpha$ apart. In this case, the area of the IF zone is the same as the transmission zone, π. The second region is when the arc formed by the intersection of the edge of the transmission zone (i.e. a circle of radius *1*) and the maximum interference zone (i.e. a circle of radius α) with the line connecting the two transmitting nodes is less than $\frac{\pi}{6}$ radians. This angle is denoted as ϕ_1 in Figure B1a. The third region occurs when this intersection occurs at an arc greater than $\frac{\pi}{6}$ radians. The angle of this intersection can be determined using the law of cosine. With the law of cosine we have

$$\alpha^2 = x^2 + 1 + 2x\alpha \cdot \cos(\phi_1).$$

Applying the quadratic formula and some algebra we determine that the separation between nodes when $\phi_1 = \frac{\pi}{6}$ is

$$x_1 = \frac{\sqrt{3} + \sqrt{3 - 4(1 - \alpha^2)}}{2}. \tag{6A.1}$$

To determine the clear area we derive an equation for the clear area in the arc segment from 0 to $\frac{\pi}{6}$ radians and multiple it by 12. In both regions 2 and 3 we are interested in the distance to the arc formed by the meeting of the end points of two lines exactly y and αy from the transmitting nodes as illustrated in Figure 6A.1b. We call this arc the capture interface We again use the law of cosines to determine the value of y. We start with

$$\alpha^2 y^2 = x^2 + y^2 - 2xy \cdot \cos(\phi)$$

and then solve for y and obtain

$$y(\phi, x) = \frac{-2x \cdot \cos(\phi) + \sqrt{(2x \cdot \cos(\phi))^2 + 4(\alpha^2 - 1)x^2}}{2(\alpha^2 - 1)}. \tag{6A.2}$$

Applying the law of cosines in a similar manner we determine ϕ_1 to be

$$\phi_1 = \arccos\left(\frac{x^2 + 1 - \alpha^2}{2x}\right). \tag{6A.3}$$

The area of the IF zone is then

$$C(x,\alpha)=\begin{cases}\pi & x\geq(1+\alpha)\\ 12\left(\int_{0}^{\phi_1}\frac{\left(y(\phi,x)\right)^2}{2}d\phi+\int_{\phi_1}^{\pi/6}\frac{1}{2}d\phi\right) & x_1<x<(1+\alpha)\\ 12\int_{0}^{\pi/6}\frac{\left(y(\phi,x)\right)^2}{2}d\phi & x\leq x_1\end{cases}.$$

where x_1, y, and ϕ_1 are determined using equations (6A.1), (6A.2), and (6A.3) respectively.

Chapter 7
Conclusion

In this thesis we have attempted to conduct a comprehensive review of the protocols and techniques that may be used by single channel wireless data networks to conserve energy. We have considered networks that are centrally controlled and networks that are distributed following the ad hoc model. Unique to our work is the concerted effort to optimize the performance of each energy conserving mechanism. Our primary focus has been to maximize the time nodes that are not involved in data exchanges can spend in low energy states.

In our study of centrally controlled networks we identify two very simple rules that should be used in selecting or in designing protocols that conserve energy. First, the primary goal of protocols should be to put nodes not participating in data exchanges into a low energy state. Second, information that enables these nodes to enter these states must be made available to them as soon as they can use it. We found that the design alternatives that accomplish these objectives in centrally controlled networks minimize the impact that other efforts such as packet accumulation, scheduling, and reducing overhead have on energy consumption. We also considered the impact of transmission errors and error recovery procedures on the performance of the protocols. Invariably we found that immediate retransmissions are very costly in energy consumption since they violate the two rules above. The additional time associated with immediate retransmissions and their random occurrence makes it difficult to predict how long nodes should remain in low energy states. This inability frequently results in nodes waking and having to wait for information in order to return to a low energy state. The solution is for retransmissions to be rescheduled. Since in some protocols this requires the source nodes to contend again and thus consume additional energy, we found that those protocols where the point coordinator is cognizant of these failures and can automatically reschedule them can also conserve the most energy. Consistent with this theme we found that access protocols used by nodes to inform the PC that they have traffic to send conserve the most energy when either the PC manages the access attempts such as through polling or when the nodes are awake only when they attempt to gain access.

The most significant contribution of this dissertation is the introduction of a new access protocol for ad hoc networks. The lessons that were learned in our study of centrally controlled networks cannot be applied to most access protocols currently in use for ad hoc networking because of their temporally random nature. In these protocols, nodes gain access by attempting to gain access at different times. This is accomplished by nodes randomly selecting when to attempt to gain access. On account of this type of access strategy, nodes cannot predict when they need to be awake. This deficiency in existing protocols motivated us to try to create a new protocol to remedy the randomness. Our solution was to synchronize access attempts and to use signaling techniques to resolve these contentions. In this sort of scheme nodes know when they must listen to the network (i.e. during access attempts and when they are exchanging data) and when they can doze (i.e. after other nodes win the contention and are exchanging data or when no nodes contend.) Surprisingly, the more significant contributions of this protocol are its other performance characteristics.

Synchronous Collision Resolution (SCR) is a high capacity access protocol. We demonstrated through analysis and simulation that this protocol can achieve exceptional capacity, easily exceeding that which is considered optimum for random access protocols such as CSMA and Slotted Aloha. We proved that our use of request-to-send and clear-to-send exchanges prevented hidden node effects in the exchange of data. We further designed our protocol to provide service differentiation. This service differentiation supports both best effort and stream based services. The stream based service capability allows the protocol to be leveraged to provide both bandwidth and delay guarantees. We also differentiate between peer-to-peer and broadcast transmissions. This differentiation gives broadcast transmissions priority over best effort peer-to-peer and enables the use of receiver directed codes. Finally, we made a special priority for energy conservation. This priority enhances the protocol's ability to support extended dozing periods.

The thesis concludes with our effort to enhance the capacity of SCR using next hop policies and spread spectrum coding techniques. We demonstrate that each of these techniques more than double the number of transmissions that can occur simultaneously. Of importance, we demonstrate that it is the unique features of SCR that make the use of spread spectrum coding feasible. SCR and the use of spread spectrum coding are complementary protocols. Additionally, SCR has the potential to use CDMA and SDMA techniques to further increase its spatial capacity. Finally, we show that simple adjustment of the maximum allowable transmission range of nodes can also be used to enhance capacity.

The synchronous collision resolution protocol forms the first half of the set of protocols required to make ad hoc networks function. The second half is routing. In the sequel to this research we developed a routing protocol that complements SCR and uses the dissemination of routing information as a vehicle to support the use of spread spectrum codes and the use long term dozing for energy conservation. Our research in this area demonstrates that SCR's resource reservation features can be used to create multihop connections and that these paths can be assembled so as to route traffic away from congested areas.

The ability of protocols to support energy conservation is likely to remain a key objective of protocols used in wireless data networks. In this dissertation, we have provided a comprehensive study of how access protocols best support energy conservation. Two simple rules dominate the design of energy conservation protocols. Our application of these rules to ad hoc networks led to the creation of a new protocol that could increase the significance of ad hoc networking as a commercial networking architecture.

Appendix A

An Annotated List of References by Topic

Wireless Medium Access Control Protocols

L. Kleinrock and S Lam, “Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation,” *IEEE Transaction on Communications*, Vol. 24, No. 5, April 1975, pp. 410 – 423.

L. Kleinrock and F. A. Tobagi, “Packet Switching in Radio Channels: Part I - Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics,” *IEEE Transactions on Communications*, December 1975, pp. 1400-1416.

F. A. Tobagi and L. Kleinrock, “Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy Tone Solution,” *IEEE Transactions on Communications*, December 1975, pp. 1417-1433.

C. Chang and J. Chang, “Optimal Design Parameters in a Multihop Packet Radio Network Using Random Access Techniques,” *IEEE Global Communication Conference ’84*, Atlanta, GA, November 1984, pp. 493-497.

D. Goodman, R. Valenzuela, K. Gayliard, and B. Ramamurthi, “Packet Reservation Multiple Access for Local Wireless Communications,” *IEEE Transactions on Communications*, Vol. 37, No. 8, August 1989, pp. 885 – 890.

P. Karn, “MACA – A New Channel Access Method for Packet Radio” *ARRL/CRRL Amateur Radio 9th Computer Networking Conference*, 1990, pp. 134-140.

R. Prasad, “Performance Analysis of Mobile Radio Networks in Real Channels With Inhibit-Sense Multiple Access,” *IEE Proceedings-I*, Vol. 138, No. 5, October 1991.

J. Sobrinho and J. Brazio, “C-MCMA: A New Multiple Access Protocol for Centralized Wireless Local Networks,” *1993 Global Communications Conference*, 1993, pp. 64 – 69.

J. Sobrinho and J. Brazio, “D-MCMA: A New Multiple Access Protocol for Distributed Wireless Local Networks,” *Global Communications Conference*, 1993, pp. 70 - 75.

J. Sobrinho and J. Brazio, “Analysis of a Multiple Access Protocol for Wireless Data Networks with Bi-directional Traffic,” *1994 IEEE Symposium on Information Theory*, 1994, p. 412.

V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, “MACAW: A Media Access Protocol for Wireless LAN's,” *Proceedings of the Conference on Communications Architectures, Protocols, and Applications*, 1994, pp. 212 – 225.

G.A. Halls, “HIPERLAN: The High Performance Radio Local Area Network Standard,” *Electronics and Communication Engineering Journal*, December 1994, pp. 289-296.

T. Phipps, "HIPERLAN Medium Access Control," *5th IEEE Symposium on Personal, Indoor and Mobile Radio Communications*, Vol 3, 1994, pp. 870-874.

N. Abramson, "Multiple Access in Wireless Digital Networks," ," *Proceeding of the IEEE*, V82, N9, Sep 1994, pp 1360-1369.

K. C. Chen, "Medium Access Control of Wireless LANs for Mobile Computing," *IEEE Network Magazine*, September/October 1994, pp 50 – 63.

D. Jeong, C. Choi, and W. Jeon, "Design and Performance Evaluation of a New Medium Access Control Protocol for Local Wireless Data Communications," *IEEE/ACM Transactions on Networking*, Vol. 3, No. 6. December 1995, pp. 742 – 752.

R. Davies, R. Watson, A. Munro, and M. Barton, "Ad-Hoc Wireless Networking: Contention Free Multiple Access," *Fifth IEE Conference on Telecommunications*, 1995, pp. 73-77.

S. Nor and I. Ismail, "Performance of CSMA-CA MAC Protocol for Distributed Radio Local Area Networks," *Sixth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Vol. 2, 1995, 912-916.

N. Muriithi and A. Burr, "A Robust Token Passing Protocol for Peer-to-Peer Radio LANS," *IEE Colloquium on Radio LANs and MANs*, 1995, pp 6/1 – 6/6.

T. Ho and K. Chen, “Performance Analysis of IEEE 802.11 CSMA/CA Medium Access Control Protocol,” *Seventh International Symposium on Personal, Indoor and Mobile Radio Communications*, Vol. 2, 1996, pp. 407 – 411.

G. Bianchi, L. Fratta, and M. Oliveri, “ Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs,” *Seventh International Symposium on Personal, Indoor and Mobile Radio Communications,* Vol. 2, 1996, pp. 392 - 396.

K.C.Chen and Y.K Sun, “A General Multi-layer Collision Resolution Multiple Access Protocol for Wireless Networks,” *46th Vehicular Technology Conference*, 1996, pp. 116-120.

H. Chhaya and S. Gupta, “Performance of Asynchronous Data Transfer Methods of IEEE 802.11 MAC Protocol,” *IEEE Personal Communications Magazine*, October 1996, pp. 8 – 15.

P. Alajajian, “An Enhanced Wireless Local Are Network Architecture,” *Seventh International Symposium on Personal, Indoor, and Mobile Radio Communications*, Vol. 3, 1996, pp. 1173 – 1177.

J. Meierhofer and U. Bernhard, “Medium Access Control for Wireless Extensions of ATM Networks,” *1996 Global Communications Conference*, 1996, pp. 29 – 33.

I.R. Johnson and S.K. Barton, “Standards for Wireless LANs,” *IEE Colloquium on Wireless Technology*, (Digest No. 1996/199) pp. 5/1-5/5.

L. Nenonen and J. Mikkonen, “Wireless ATM MAC Performance Evaluation, Case Study: HIPERLAN vs Modified MDR,” *Second Annual International Conference on Mobile Communications and Networking*, 1996, pp. 165 – 175.

D. Gillies, Y. Guo, S. Barton, “Synchronisation Techniques for HIPERLAN,” *Wireless Personal Communications*, Vol. 4, No. 1, Kluwer Academic Publishers, January 1996, pp. 1-10.

A. Wittneben and W. Liu, “The European Wireless LAN Standard HIPERLAN: Key Concepts and Testbed Results,” *IEEE 47th Vehicular Technology Conference*, Vol. 3, 1997, pp. 1317-1321.

K. Fu, Y. Guo, and S. Barton, “Performance of the EY-NPMA Protocol,” *Wireless Personal Communications*, Vol. 4, No. 1, Kluwer Academic Publishers, January 1997, pp. 41-50.

P. Jacquet, P. Minet, P. Mühlethaler, and N.Rivierre, “Data Transfer for HIPERLAN,” *Wireless Personal Communications*, Vol. 4, No. 1, Kluwer Academic Publishers, January 1997, pp. 65-80.

Y.K. Sun, K.C. Chen, and D. C. Twu, “Generalized Tree Multiple Access Protocols in Packet Switching Networks,” *Eighth International Symposium on Personal, Indoor and Mobile Radio Communications*, Vol. 3, 1997, pp. 918-922.

B. Mullins, N. Davis IV, and S. Midkiff, “A Wireless Local Area Protocol That Improves Throughput Via Adaptive Control,” *1997 International Conference on Communications*, Vol. 5, 1997, pp. 1427 – 1431.

G.Bianchi, F. Borgonovo, L. Fratta, L. Musumeci. and M. Zorzi, “C-PRMA: A Centralized Packet Reservation Multiple Access for Local Wireless Communications,” *IEEE Transactions on Vehicular Technology*, Vol. 46, No. 2, May 1997, pp. 422 – 436.

P802.11, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Press, May 1997.

H. Chhaya and S. Gupta, “Performance Modeling of Asynchronous Data Transfer Methods of IEEE 802.11 MAC Protocol,” *ACM/Baltzer Wireless Networks*, Vol. 3, No. 3, Aug 1997, pp. 217 – 234.

B. Bing and R. Subramanian, “A New Multiaccess Technique for Multimedia Wireless LANs,” *IEEE Global Telecommunications Conference*, 1997, pp 1318-1322.

J. Cai, and D. Goodman, “General Packet Radio Service in GSM,” *IEEE Communications Magazine*, October 1997, pp. 122 – 131.

P. Jacquet, P. Minet, P. Mühlethaler, and N.Rivierre, “Increasing Reliability in Cable-Free Radio LANs, Low Level Forwarding in HIPERLAN,” *Wireless Personal Communications*, Vol. 4, No. 1, Kluwer Academic Publishers, January 1997, pp. 51 - 63.

F. Ayo, B. Jafarian, and A. Aghvami, “A Medium Access Control Protocol for Wireless ATM LAN Based Polling,” *IEE Colloquium on ATM Traffic in Personal Mobile Communication Environments*, 1997, pp. 2/1 – 2/6.

P. Jacquet, P. Minet, P. Mühlethaler, and N.Rivierre, "Priority and Collision Detection with Active Signaling – The Channel Access Mechanism of HIPERLAN," *Wireless Personal Communications*, Vol. 4, No. 1, Kluwer Academic Publishers, January 1997, pp. 11-26.

J. Weinmiller, M. Schlager, A. Festag, and A. Wolisz, "Performance Study of Access Control in Wireless LANs – IEEE 802.11 DFWMAC and ETSI RES 10 Hiperlan," *ACM/Baltzer Mobile Networks and Applications*, Vol. 2, No. 1, 1997, pp. 55 – 67.

B. Crow, et al., "IEEE 802.11 Wireless Local Area Networks," *IEEE Communications Magazine*, September 1997, pp 116-126.

I. Vukovic, "HIPERLAN Type 1: Performance Analysis of the Channel Access Control Protocol," *48th IEEE Vehicular Technology Conference*, Vol. 1, 1998, pp. 377-381.

C. Wu and G. Ma, "Performance of Packet Reservation MAC Protocols for Wireless Networks," *48th IEEE Vehicular Technology Conference*, Vol. 3, 1998, pp. 2537 – 2541.

Y.K. Sun and K.C. Chen, "Optimization of Generalized Tree Protocols with Capture," *48th IEEE Vehicular Technology Conference*, Vol. 2, 1998 pp. 801 – 805.

G. Anastasi, L. Lenzini, and E. Mingozzi, "Stability and Performance Analysis of HIPERLAN," *INFOCOM '98*, 1998, pp. 134 – 141.

Y.K. Sun and K.C. Chen, “Dynamics of Generalized Tree Protocols and Its Application to Delay Sensitive Traffic Control,” *1998 International Conference on Communications*, Vol. 2, 1998, pp. 849 – 853.

K. Leung, J. Ho, and H. Chien, “A New Digital Sense Multiple Access (DSMA) Protocol for High-Speed Wireless Networks,” *The Ninth IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications*, Vol. 3,1998, pp. 1360 – 1366.

S. Miura, H. Nakamura, M. Kamienoo, and K. Araki, “Radio Link Control Method by Using Radio Token in High Speed Wireless LAN System,” *1998 Global Communications Conference*, Vol. 3, 1998 pp. 1811 – 1816.

J. Lai, Y. Lai, and S. Lee, “A Medium Access Control Protocol for Wireless Networks,” *1998 IEEE International Conference on Communications*, 1998, pp. 146 – 150.

F. Callegati, M. Casoni, M. Merani, and C. Raffaelli, “A Wireless ATM LAN with an 802.11-Like Data Link Layer,” *48th IEEE Vehicular Technology Conference*, Vol. 2, 1998, pp. 1135 – 1139.

R. Garces and J.J. Garcia-Luna-Aceves, “ A Near-Optimum Access Protocol Based on Incremental Collision Resolution and Distributed Transmission Queues,” *INFOCOM 98*, March 1998, pp 158 – 165

L. Nenonen and J. Mikkonen, “Wireless ATM MAC Performance Evaluation, A Case Study: HIPERLAN Type 1 vs Modified MDR,” *ACM/Baltzer Mobile Networks and Applications*, Vol. 3, No. 3, 1998, pp. 247-259.

G. Binachi, “IEEE 802.11 – Saturation Throughput Analysis,” *IEEE Communications Letters*, Vol. 2, No. 12, December 1998, pp. 318 – 320.

P. Omiyi and T. O’Farrell, “A New Throughput Analysis of a Novel MAC Protocol for CDMA-Based Wireless LANs,” *IEEE Radio and Wireless Conference*, 1998, pp. 23 – 26.

M. Naraghi-Pour, M. Hegde, and R. Pallapotu, “Peer-to-Peer Communication in Wireless Local Area Networks,” *IEEE International Conference on Computer Communications and Networks*, 1998, pp. 432 – 439.

A. Muir and J. Garcia-Luna-Aceves, “An Efficient Packet Sensing MAC Protocol for Wireless Networks,” *ACM/Baltzer Mobile Networks and Applications*, Vol. 3, No. 2, 1998, pp. 221 – 234.

C. L. Fulmer, Collision Avoidance Techniques for Packet-Radio Networks, PhD Dissertation, University of California, Santa Cruz, 1998.

ETSI, EN300 652 V 1.2.1, “Broadband Radio Access Networks (BRAN); High Performance Radio Local Area Network (HIPERLAN) Type 1; Functional Specification, July 1998.

W.Luo and A. Ehphremides, “Power Levels and Packet Lengths in Random Multiple Access,” Submitted to *IEEE Transactions on Information Theory*, 1999.

I. Vokovic, “Analysis of the Hidden Node Problem in HIPERLAN,” *1999 IEEE Conference on Communications*, 1999, pp. 1234-1238.

H. Lee, S. Djoko, H. Jiang, S. Subramanian, and K. Basu, “Wireless Networks Self Engineering Engine,” *IEEE Symposium on Application-Specific Systems and Software Engineering and Technology*, 1999, pp. 258 – 263.

M. Zorzi and R. Rao, “Perspectives on the Impact of Error Statistics on Protocols for Wireless Networks,” *IEEE Personal Communications Magazine*, October 1999, pp. 32-40.

R. van Nee, G. Awater, M. Morikura, H Takanashi, M. Webster, and K. Halford, “New High-Rate Wireless LAN Standards,” *IEEE Communications Magazine*, December 1999, pp. 82-88.

G. Bianchi, “Performance Analysis of the IEEE 802.11 Distributed Coordination Function,” *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 3, March 2000, pp 535 – 547.

L. Wing and M. Hamdi, “A Hybrid Adaptive Wireless Channel Access Protocol for Multimedia Personal Communications Systems,” *Wireless Personal Communications*, Vol. 13, No. 1/2, Kluwer Academic Publishers, May 2000, pp. 79-96.

J. Dai, “The Scheduling to Achieve Optimized Performance of Randomly Addressed Polling Protocol,” *Wireless Personal Communications*, Vol. 15, No. 2, Kluwer Academic Publishers, November 2000, pp. 161-179.

Energy Conserving Wireless Medium Access Control Protocols

K. M. Sivalingam, M. Srivastava and P. Agrawal, "Low power link and access protocols for wireless multimedia networks", *IEEE Vehicular Technology Conference* , (Phoenix, AZ), pp. 1331-1335, May 1997.

Y.Leung, “Mean Power Consumption of Artificial Power Capture in Wireless Networks,” *IEEE Transactions on Communications*, Vol. 45, No. 8, August 1997, pp. 957-964.

K. M. Sivalingam , Mani Srivastava, Prathima Agrawal and Jyh-Cheng Chen, "Low-power access protocols based on Scheduling for Wireless and Mobile ATM networks," *IEEE International Conference on Universal Personal Communications (ICUPC)*, (San Diego, CA), Oct. 1997.

A. Salkinzis and C. Chamzas, “An In-Band Power-Saving Protocol for Mobile Data Networks,” *IEEE Transactions on Communications*, Vol. 46, No. 9, September 1998, pp. 1194-1205.

A. Chockalingam and Michele Zorzi, “Energy Efficiency of Media Access Protocols for Mobile Data Networks,” *IEEE Transactions on Communications*, Vol. 46, No. 11, November 1998, pp. 1418-1421.

I. Chlamatac, C. Petrioli, and J. Redi, “Energy-conserving Go-Back-N ARQ Protocols for Wireless Data Networks,” *1998 International Conference on Universal Personal Communications*, Vol. 2, 1998, pp. 1259-1263.

I. Chlamatac, C. Petrioli, and J. Redi, “Energy-conserving Selective Repeat ARQ Protocols for Wireless Data Networks,” *The Ninth International Symposium on Personal, Indoor, and Mobile Radio Communications*, Vol. 2, 1998, pp. 836-840.

P. Agrawal, “Energy Efficient Protocols for Wireless Systems,” *Personal, Indoor and Mobile Radio Communications Conference*, 1998, pp. 564 – 569.

C. Rohl, H. Woesner, A Wolisz, “A Short Look on Power Saving Mechanisms in the Wireless LAN Standard Draft IEEE 802.11,” Source publication not known.

J-C. Chen, K. M. Sivalingam , P. Agrawal, and S. Kishore, “A Comparison of MAC Protocols for Wireless Local Networks Based on Battery Power Consumption,” *INFOCOM 98*, March 1998, pp. 150 - 157.

H. Woesner, J. Ebert, M. Schlager, and A. Wolisz, “Power-Saving Mechanisms in Emerging Standards for Wireless LANs: The MAC Level Perspective,” *IEEE Personal Communications*, June 1998, pp. 40-48.

S. Singh and C. Raghavendra, “Power Efficient MAC Protocol for Multihop Radio Networks,” *Personal, Indoor and Mobile Radio Communications Conference*, 1998, pp. 153 – 157.

K. M. Sivalingam, J.-C. Chen and P. Agrawal, "Design and Analysis of Low-Power Access Protocols for Wireless and Mobile ATM Networks", *ACM/Baltzer Journal on Special Topics in Mobile Networking and Applications (MONET)*, June 1998.

Shalinee Kishore, Jyh-Cheng Chen, K. M. Sivalingam, and P. Agrawal, “Adaptive Power Control and Scheduling Algorithms Based on Battery Power for CDMA Wireless Networks", *IEEE International Conference on Universal Personal Communications (ICUPC)*, (Florence, Italy), Oct. 1998.

Y.K. Sun and K.C. Chen, “ Energy Efficient Multiple Access Protocol Design,” *IEEE Communications Letters*, Vol. 2 No. 12, December 1998, pp. 334 – 336.

I. Chlamtac, C. Petrioli, and J. Redi, “Energy-Conserving Access Protocols for Identification Networks,” *IEEE/ACM Transactions on Networking*, Vol. 7, No. 1, February 1999, pp. 51-59.

Y. Lin and Y. Chuang, “Modeling the Sleep Mode for Cellular Digital Packet Data,” *IEEE Communications Letters*, Vol. 3, No. 3, March 1999, pp. 63 – 65.

J. Redi and D. Avresky, “Performance of Energy-Conserving Access Protocols Under Self Similar Traffic,” *1999 IEEE Wireless Communications and Networking Conference*, Vol. 2, 1999, pp. 626 – 630.

P. Bahl, “ARMAP – An Energy Conserving Protocol for Wireless Multimedia Communications,” *Ninth IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, 1998, pp. 575 – 580

J.A. Stine and G. de Veciana, “Improving Energy Efficiency of Centrally Controlled Wireless Data Networks,” Submitted to the ACM/Baltzer Wireless Networking Journal.

Multimedia Access Protocols

J. Sobrinho and A.S. Krishnakumar, “Real-Time Traffic over the IEEE 802.11 Medium Access Control Layer,” *Bell Labs Technical J*ournal, Autumn 1996, pp 172-187.

J. Wu and R Kohno, “A Wireless Multimedia CDMA System Based on Transmission Power Control,” *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 4, May 1996, pp. 683-691.

C. Andren, “IEEE 802.11 Wireless LAN: Can We Use It for Multimedia?,” *IEEE Multimedia*, Vol 52, 1998, pp. 84-89.

J. Sobrinho and A. Krishnakumar, “Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Wireless Networks,” *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 8, August 1999, pp. 1553 - 1568.

A. Zahedi, K. Pahlavan, and J. Rulnick, “Traffic Engineering for the Wireless Access Point,” *Ninth International Symposium of Personal, Indoor and Mobile Radio Communications*, 1998, pp. 1170-1175.

J. Stine and G. de Veciana, "Tactical Communications Using the 802.11 MAC Protocol," *MILCOM 98*, October 1998.

I. Akyildiz, J. McNair, L. Martorell, R. Puigjaner, and Y. Yesha, "Medium Access Control Protocols for Multimedia Traffic in Wireless Networks," *IEEE Network Magazine*, July/August 1999, pp. 39 – 47.

N. Figueira and J. Pasquale, "Providing Quality of Service for Wireless Links: Wireless/Wired Networks," *IEEE Personal Communications*, October 1999, pp. 42-51.

M. Johnsson, "HiperLAN/2 – The Broadband Radio Transmission Technology Operating in the 5 GHz Frequency Band," HiperLAN/2 Global Forum, 1999.

G. Kuo and P. Ko, "A Collision-Free Medium Access Control Protocol for Flow-Oriented Ad Hoc Wireless LAN," *IEEE 49th Vehicular Technology Conference*, Vol. 1, 1999 pp. 325-331.

Wireless Network Implementation

B. Leiner, D. Nielson, and F. Tobagi, "Issues in Packet Radio Network Design," *Proceedings of the IEEE*, V75, N1, pp 6- 20, Jan 1987.

J. Jurbin and J. Tornow, "The DARPA Packet Radio Network Protocols," *Proceeding of the IEEE*, V75, N1, Jan 1987, pp 21-32.

J. Rustad, R. Skaug, and A. Aasen, "New Radio Networks for Tactical Communications," *IEEE Journal on Selected Areas in Communications*, Vol. 8 No.5, June 1990, pp. 713 – 727.

D. Beyer, "Accomplishments of the DARPA SURAN Program," *MILCOM 1990*, pp. 855-862.

G.J. Pottie and W.J. Kaiser, "Wireless Integrated Network Sensors," *Communications of the ACM*, Vol. 43, No. 5, pp.51 – 58.

A. Fasbender, F. Reichert, E. Geulen, J. Hjelm, and T. Wierlemann, "Any Network, Any Teminal, Anywhere," *IEEE Personal Communications*, April 1999, pp. 22 – 30.

J. Evans, G. Minden, K. Shanmugan, G. Plescott, V. Frost, B. Ewy, R. Sanchez, C. Sparks, K. Malinimohan, J. Roberts, R. Plumb, and D. Petr, "The Rapidly Deployable Radio Network," *IEEE Journal on Selected Areas in Communications,* Vol. 17 No.4, June 1999, pp. 689 - 703.

K. Sohrabi, J. Gao, V. Ailawadhi, and G. Pottie, "A Self Organizing Wireless Sensor Network," *Proceedings of the 39th Annual Allerton Conference on Communications, Control, and Computers*, October 1999.

K. Sohrabi and G. Pottie, "Performance of a Novel Self-Organizing Protocol for Wireless Ad-Hoc Sensor Networks," *Proceedings of the 50th Vehicular Technology Conference*, 1999, pp.1222-1226.

P. Sass, “Communications Networks for the Force XXI Digitized Battlefield,” *Mobile Networks and Applications*, Vol. 4, No. 3, 1999, pp. 139 – 155.

A. Salkintzis, “A Survey of Mobile Data Networks,” *IEEE Communications Surveys*, Vol 2, No. 3, Third Quarter 1999, pp 2-18.

K. Sohrabi, J. Gao, V. Ailawadhi, and G. Pottie, “Protocols for Self-Organization of a Wireless Sensor Network,” *IEEE Personal Communications*, October 2000, pp. 16 – 27.

Other Energy Conservation Papers

A. Datta, A. Celik, J. Kim, and D. VanderMeer, “Adaptive Broadcast Protocols to Support Power Conservant Retrieval by Mobile Users,” *13th International Conference on Data Engineering*,1997, pp. 124-133.

M. Stemm, P. Gauthier, D. Harada, and R. Katz, “Reducing Power Consumption of Network Interfaces in Hand-Held Devices,” Presented at the 3rd International Workshop on Mobile Multimedia Communications but not in proceedings. Found at http://www.cs.berkley.edu/~stemm/pubs.html.

M. Stemm and R. Katz, “Measuring and Reducing Energy Consumption of Network Interfaces in Hand-Held Devices,” *IEICE Transactions on Communications*, Vol. E80-B, no. 8, Aug 1997 pp.1125-1131.

J. Lorch and A. Smith, “Software Strategies for Portable Computer Energy Management,” *IEEE Personal Communications*, June 1998, pp 60 – 73

S. Dolev, E. Korach, and D. Yukelson, “The Sound of Silence: Guessing Games for Saving Energy in Mobile Environments,” *INFOCOM ’99*, pp 768-775.

S. Tragoudas, “Power Dissipation Component of a Management Protocol for Ad-Hoc Networks,” *Proceedings of the WCNC*, Sep 2000, pp

L.M. Feeney and M. Nilsson, “Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment,” *INFOCOM 2001*, April 2001.

Wireless LAN Modems

W. Fifer and F. Bruno, “The Low-Cost Packet Radio,” *Proceeding of the IEEE*, Vol. 75, No. 1, January 1987, pp 33 – 42.

T. Yamawaki, et al, “A 2.7-V GSM RF Transceiver IC,” *IEEE Journal of Solid State Circuits*, Vol. 32, No. 12, December 1997, pp. 2089–2095.

J. Min, A. Rofougran, H. Samueli, and A. Abidi, “An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread Spectrum Transceiver,” *IEEE 1994 Custom Integrated Circuits Conference*, 1994, pp. 379-382.

Nokia C020/C021 Wireless Lan Card Technical Specifications, 1998 http://www.aironet.com/products/in_building/PC4800B.asp.

P. Lettieri and M. Srivastava, “Advances in Wireless Terminals,” *IEEE Personal Communications*, February 1999, pp. 6-19.

Lucent Technologies WaveLAN High-Speed Wireless Network Interface Card IEEE Turbo 11 Mb Specification, 1999.

Mobile Ad Hoc Network Routing Protocols

J. Hahn and D. Stolle, “Packet Radio Network Routing Algorithms: A Survey,” *IEEE Communications Magazine*, Vol. 22, No. 11, November 1984, pp. 41 – 47.

J. Garcia Luna Aceves and N. Schacham, “Analysis of Routing Strategies for Packet Radio Networks,” *Proceeding of the IEEE INFOCOM ’85*, Washington DC, March 1985, pp.292-302.

T. Hou and V. Li, “Position Updates and Sensitivity Analysis for Routing Protocols in Multihop Mobile Packet Radio Networks,” *IEEE Global Communication Conference ’85,* New Orleans, LA, Dec 1985, pp. 243-249.

C. Perkins and P. Bhagwat, “Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers,” *Proceedings of the Conference on Communications Architectures, Protocols, and Applications*, August 1994, pp. 234 – 244.

G. Lauer, “Packet-Radio Routing,” *Routing in Communications Networks*, Ed. M. Steenstrup, Prentice-Hall, Inc, Englewood Cliffs, NJ, 1995, pp. 351-396.

M. Gerla and J. Tsai, “Multicluster, Mobile, Multimedia Radio Networks,” *ACM Wireless Networks*, V.1, N3, Oct. 1995, pp. 255-265.

S. Murthy and J. Garcia-Luna-Aceves, “A Routing Protocol for Packet Radio Networks,” *MOBICOM 95*, pp 86 – 95.

R. Dube, C. Rais, K-Y. Wang, and S. Tripathi, “Signal Stability-Based Adaptive Routing (SSA) for Ad Hoc Mobile Networks,” *IEEE Personal Communications*, February 1997, pp. 36 – 45.

S. Murthy and J. Garcia-Luna-Aceves, “An Efficient Routing Protocol for Wireless Networks,” *ACM Mobile Networks and Applications Journal*, Oct 1997, pp 183-197.

V. Park and M. Corson, “A Highly Adaptive Routing Algorithm for Mobile Wireless Networks,” *Proceedings IEEE INFOCOM 1997*, April 1997, pp. 1405 – 1413.

C-K Toh, “Associativity-Based Routing for Ad-Hoc Mobile Networks,” Wireless Personal Comunications, Vol. 4, No. 2, Kluwer Academic Publishers, March 1997, pp. 1 – 36.

B. Das and V. Bharghavan, “Routing in Ad-Hoc Networks Using Minimum Connected Dominating Sets,” *IEEE International Conference on Communications*, June

1997, pp.376 – 380.

B. Das, R. Sivakumar, and V. Bharghavan, "Routing in Ad Hoc Networks Using a Spine," *Sixth International Conference on Computer Communications and Networks*, 1997, pp. 34 – 39.

Y-B Ko and N. Vaidya, "Location-Aided Routing (LAR) in Mobile Ad Hoc Networks," *Proceedings of the 4th Annual IEEE/ACM Conference on Mobile Computing and Networking*, 1998, pp. 66 – 75.

R. Sivakumar, B. Das, and V. Bharghavan, "The Claude Vertebrata: Spines and Routing in Ad Hoc Networks," *IEEE Symposium on Computers and Communications*, 1998, pp. 599 – 605.

S. Basagni, I. Chlamatac, and V. Syrotiuk, "A Distance Routing Effect Algorithm for Mobility (DREAM)," *Proceedings of the 4th Annual IEEE/ACM Conference on Mobile Computing and Networking*, 1998, pp. 76 – 84.

J. Broch, D. Maltz, D. Johnson, Y-C Hu, J. Jetcheva, "A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols," *The Fourth Annual ACM/IEEE International Conference in Mobile Computing and Networking*, 1998, pp. 85 – 97.

C. Perkins and E. Royer, "Ad-hoc On-Demand Distance Vector Routing," Proceedings *IEEE Workshop on Mobile Computing Systems and Applications*, February 1999, pp. 90-100.

E. Royer, and C-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," *IEEE Personal Communications*, April 1999, pp. 46 – 55.

S-J. Lee, M. Gerla, and C-K Toh, "A Simulation Study of Table-Driven and On-Demand Routing Protocols for Mobile Ad Hoc Networks," *IEEE Network*, July/August 1999, pp. 48 – 54.

S. Chen and K. Nahrstedt, Distributed Quality-of-Service Routing in Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 8, August 1999, pp. 1488 – 1505.

M. Joa-Ng and I. Lu, "A Peer-to-Peer Zone-Based Two-Level Link State Routing for Mobile Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 8, August 1999, pp. 1415- 1425.

A.B. McDonald and T. Znati, "A Mobility-Based Framework for Adaptive Clustering in Wireless Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 8, August 1999, pp. 1466– 1487.

C. Lin and J. Liu, "QoS Routing in Ad Hoc Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 8, August 1999, pp. 1426 – 1458.

Y-B Ko and N. Vaidya, "Geocasting in Mobile Ad Hoc Networks: Location-Based Multicast Algorithms," *Proceedings Second IEEE Workshop on WMCSA '99*, 1999,pp. 101 – 110.

Y-B Ko and N. Vaidya, "Using Location Information in Wireless Ad Hoc Networks," *1999 IEEE 49th Vehicular Technology Conference*, 1999, pp. 1952 – 1956.

M. Shah and P. Flikkema, "Power-Based Leader Selection in Ad-Hoc Wireless Networks," *IEEE International Performance, Computing, and Communications Conference*, 1999, pp. 134 – 139.

J. Raju and J. Garcia-Luna-Aceves, "A Comparison of On-Demand and Table Driven Routing for Ad-Hoc Wireless Networks," *International Conference on Communications*, 2000, pp. 1702-1706.

G. Pei, M. Gerla, X. Hong, and C-C. Chiang, "A Wireless Hierarchical Routing Protocol with Group Mobility," *1999 IEEE Wireless Communications and Networking Conference*, 1999, pp. 1538 – 1542.

C. Tschudin, H. Lundgren, and H. Gulbrandsen, "Active Routing for Ad Hoc Networks," *IEEE Communications,* April 2000, pp. 122 – 127.

S. Das, C. Perkins, and E. Royer, "Performance Comparison of Two On-demand Routing Protocols for Ad Hoc Networks," *Proceedings IEEE INFOCOM 2000,* April 2000, pp. 3 – 12.

G. Pei, M. Gerla, and T-W. Chen, "Fisheye State Routing: A Routing Scheme for Ad Hoc Wireless Networks," *Proceedings of the International Communications Conference*, June 2000, pp. 70 – 74.

Y-B. Ko and N. Viadya, “Location-Aided Routing (LAR) in Mobile Ad Hoc Networks,” *ACM/Baltzer Wireless Networks*, Vol. 6, No. 4, July 2000, pp. 307 – 321.

C. Perkins, E. Royer, and S. Das, “Ad Hoc On-Demand Distance (AODV) Routing,” http://www.ietf.org/internet-drafts/draft-ietf-manet-aodv-06.txt, 14 July 2000.

S. Woo and S. Singh, “Longest Life Routing Protocol (LLRP) for Ad Hoc Networks with Highly Mobile Nodes,” *Proceeding of the IEEE Wireless Communications and Networking Conference*, September 2000.

G. Pei, M. Gerla, and X. Hong, “LANMAR: Landmark Routing for Large Scale Wireless Ad Hoc Networks with Group Mobility,” *Proceeding IEEE GLOBECOM 2000*, Nov 2000, pp. 11 – 18.

M. Gerla, G. Pei, X. Hong, and T-W. Chen, Fisheye State Routing Protocol (FSR) for Ad Hoc Networks (Work in progress), draft-ietf-manet-fsr-00.txt, 17 November 2000.

D. Johnson, D. Maltz, Y-C Hu, and J. Jetcheva, The Dynamic Sour Routing Protocol for Mobile Ad Hoc Networks, , (Work in progress), draft-ietf-manet-dsr-04.txt, 17 November 2000.

M. Gerla, X. Hong, L. Ma, and G. Pei, Landmark Routing Protocol (LANMAR) for Large Scale Ad Hoc Networks, (Work in progress), draft-ietf-manet-lanmar-00.txt, 17 November 2000.

P. Jacquet, P. Muhlethaler, A. Qayyum, A. Laouiti, L. Viennot, and T. Clausen, Optimized Link State Routing Protocol (Work in progress), draft-ietf-manet-olsr-03.txt, 24 November 2000.

Clustering Protocols for Routing Efficieny

C. Lin and M. Gerla, "Adaptive Clustering for Mobile Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 15, No. 7, September 1997, pp. 1265 – 1275.

G.Pei and M. Gerla, "Mobility Management in Heirarchical Multi-hop Mobile Wireless Networks," *Eighth International Conference on Computer Communications and Networks*, 1999, pp. 324 – 329.

M. Pearlman and Z. Haas, "Determining the Optimal Configurations for the Zone Routing Protocol," *IEEE Journal on Selected Areas in Communication*, Vol. 17, No. 8, August 1999, pp. 1395 – 1414.

Y-L. Chang and C-C Hsu, "Routing in Wireless/Mobile Ad-Hoc Networks Via Dynamic Group Construction," *ACM/Baltzer Mobile Networks and Applications,* Vol. 5, No. 1, May 2000, pp. 27 – 37.

Energy Conserving Routing Protocols for Mobile Ad Hoc Network

T. Meng and V. Rodolpho, "Distributed Network Protocols for Wireless Communications," *IEEE International Symposium on Circuits and Systems*,1998, pp. IV-600

– IV-603.

S. Singh, M. Woo, and C. Raghavendra, “Power-Aware Routing in Mobile Ad Hoc Networks,” *The Fourth Annual ACM/IEEE International Conference in Mobile Computing and Networking*, 1998, pp. 181 – 190.

V. Rodoplu and T. Meng, “Minimum Energy Mobile Wireless Networks,” *IEEE Journal on Selected Areas in Communication*, Vol. 17, No. 8, August 1999, pp. 1333-1344.

J. Wieselthier, G. Nguyen, and A. Ephremides, “On the Construction of Energy-Efficient Broadcast and Multicast Trees in Wireless Networks,” Submitted to *INFOCOM ‘2000*, pp. 585 - 594.

S. Tragoudas and S. Dimitrova, “Routing with Energy Considerations in Mobile Ad-Hoc Networks,” *Proceedings of the WCNC*, 2000, Sep 2000 pp.

J. Gao, *Energy Efficient Routing For Wireless Sensor Networks*, PhD Dissertation, University of California, Los Angeles, 2000.

Wireless Network Capacity

L Kleinrock and J. Silvester, “Optimum Transmission Radii for Packet Radio Networks or Why Six Is a Magic Number,” *IEEE National Telecommunications Conference*, December 1978, pp. 4.3.1-4.3.5.

R. Nelson and L. Kleinrock, “Maximum Probability of Successful Transmission in a Random Planar Packet Radio Network,” *Proceedings of the IEEE INFOCOM ’83*, 1983, pp. 365 – 370.

J. Silvester and L. Kleinrock, “On the Capacity of Multihop Slotted ALOHA Networks with Regular Structure,” *IEEE Transactions on Communications*, Vol. 31, No. 8, August 1983, pp. 974 – 982.

J. Silvester and L. Kleinrock, “On the Capacity of Single-Hop Slotted ALOHA Networks for Various Traffic Matrices and Transmission Strategies,” *IEEE Transactions on Communications*, Vol. 31, No. 8, August 1983, pp. 983 – 991.

R. Nelson and L. Kleinrock, “The Spatial Capacity of a Slotted ALOHA Multihop Packet Radio Network with Capture,” *IEEE Transactions on Communications*, Vol. 32, No. 6, June 1984, pp. 684 – 694.

H. Takagi and L. Kleinrock, “Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals,” *IEEE Transactions on Communications*, Vol. 32, No. 3, March 1984, pp. 246-257.

T. Hou and V. Li, “Performance Analysis of Routing Strategies in Multihop Packet Radio Network,” *IEEE Global Communication Conference ’84*, Atlanta, GA, November 1984, pp. 487-492.

T. Hou, “Routing Strategies and Power Control in Multihop Packet Radio Networks,” PhD Dissertation, University of Southern California, 1985.

T. Hou and V. Li, “Routing Strategies of Multihop Packet Radio Networks with a More Realistic Packet Reception Probability,” *Proceeding of the IEEE INFOCOM ’85*, Washington DC, March 1985, pp 303-309.

T. Hou and V. Li, “Performance Analysis of Multihop Packet Radio Networks with Capture,” *IEEE International Conference on Communications, ’85*, Chicago, IL, June 1985, pp. 1311-1317.

T. Hou and V. Li, “Transmission Range Control in Multihop Packet Radio Networks,” *IEEE Transactions on Communications*, Vol. 34, No. 1, January 1986, pp. 38-44.

L. Kleinrock and J. Silvester, “Spatial Reuse in Multihop Packet Radio Networks,” *Proceeding of the IEEE*, V75, N1, Jan 1987, pp 156-167.

J. Stevens, “Spatial Reuse Through Dynamic Power and Routing Control in Common-Channel Random-Access Packet Radio Networks,” *SURAN Program Technical Note (SRNTN) 59*, Richardson, TX: Rockwell Inc., August, 1988. DTIC AD 197898

E. Sousa and J. Silvester, “Optimum Transmission Ranges in a Direct-Sequence Spread-Spectrum Multihop Packet Radio Network,” *IEEE Journal on Selected Areas in Communications*, Vol. 8, No. 5, June 1990, pp. 762 – 771.

F. Shiao and J. Yee, “On Determining the Transmission Range for Multi-Hop Slotted Aloha Packet Radio Networks,” *Military Communications Conference*, 1990, pp.

851-855.

L. Hu, “Local throughput Performance of Packet Radio Networks with Transmitting Power Control,” *IEEE International Conference on Communications*, Denver, CO, June 1991, pp.1621-1626.

Networking (General)

F. Tobagi, “Modeling and Performance Analysis of Multihop Packet Radio Networks,” Proceeding of the IEEE, Vol. 75, No. 1, January 1987, pp.135 – 155.

D. Mills, “Internet Time Synchronization: The Network Time Protocol,” *IEEE Transactions on Communications*, Vol. 39, No. 10, October 1991, pp. 1482 – 1493.

D. Bertsekas and R. Gallagher, *Data Networks*, Prentice Hall, Inc., Upper Saddle River, NJ, 1992.

M. Molle, M. Kalkunte, and J. Kadambi, “Frame Bursting: A Technique for Scaling CSMA/CD to Gigabit Speeds,” *IEEE Network*, July/August 1997, pp.

Wireless Communications (Miscellaneous)

R. Kahn, “The Organization of Computer Resources into a Packet Radio Network,” *IEEE Transactions on Communications*, Vol. 25, No. 1, January 1977, pp. 169 – 177.

R. Kahn, S Gronemeyer, J. Burchfiel, and R. Kunzelman, "Advances in Packet Radio Technology," *Proceeding of the IEEE*, V66, N11, pp1468-1495, Nov 1978.

M. Pursley, "The Role of Spread Spectrum in Packet Radio Networks," *Proceedings of the IEEE*, Vol. 75, No. 1, January 1987, pp. 116 – 134.

E. Sousa and J. Silvester, "Spread Spectrum Protocols for Distributed Spread-Spectrum Packet Radio Networks," *IEEE Transactions on Communications*, Vol. 36, No. 3, March 1988, pp. 272 – 281.

B. Van Veen and K. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, April 1988, pp. 4-22.

D. Goodman, "Cellular Packet Communications," *IEEE Transactions on Communications*, Vol. 38 No. 8, August 1990, pp 1272-1280.

J. Escobar, "Radio Parameter Selection Algorithm for Receiver-Directed Packet-Radio Networks," *SURAN Program Technical Note (SRNTN) 73*, Cambridge, MA, BBN Systems and Technology Corporation, 1990, DTIC AD 224866

R. Esmailzadeh, M Nakagawa, and A. Kajiwara, "Power Control in Packet Switched Time Division Duplex Direct Sequence Spread Spectrum Communications," *Vehicular Technology Society 42nd VTS Conference*, 1992, pp 989-992.

L. Hu, "Distributed Code Assignment for CDMA Packet Radio Networks," *IEEE/ACM Transactions on Networking*, Vol. 1, No. 6, December 1993, pp. 668 – 677.

H. Xue, A. Nix, M. Beach, M. Li, J. Mcgeehan, T. Wilkinson, I. Johnson, and S. Barton, “Air Interface Considerations for Wireless LANS,” *Fourth European Conference on Radio Relay Systems*, October 1993, pp. 45 – 50.

K. Pahlaven and A. Levesque, “Wireless Data Communications,” *Proceeding of the IEEE*, Vol. 82, No. 9, September 1994, pp. 1398 – 1430.

A. Viterbi and A. Viterbi, “Other Cell Interference in Cellular Power-Controlled CDMA,” *IEEE Transactions on Communications*, Vol. 42, No. 2/3/4, February/March/April 1994, pp. 1501-1504.

A. Sheikh, Y. Yao, and S. Cheng, “Throughput Enhancement of Direct-Sequence Spread Spectrum Packet Radio Networks by Adaptive power Control,” *IEEE Transactions on Communications*, Vol. 42, No. 2/3/4, February/March/April 1994, pp. 884-890.

R. Padovani, “Reverse Link Performance of IS-95 Based Cellular Systems,” *IEEE Personal Communications*, Third Quarter 1994, pp. 28 – 34.

J. Andersen, T. Rappaport, and S. Yoshida, “Propagation Measurements and Models for Wireless Communications Channels,” *IEEE Communications Magazine*, Jan 1995, pp 42-49.

A. Viterbi, *CDMA, Principles of Spread Spectrum Communications*, Addison Wesley, Reading MA, 1995.

R. Yates, “A Framework for Uplink Power Control in Cellular Radio Systems,” *IEEE Journal on Selected Areas in Communications*, Vol. 13, No`. 7, September 1995, pp. 1341-1348.

R. Yates and C. Huang, “Integrated Power Control and Base Station Assignment,” *IEEE Transactions on Vehicular Technology*, Vol 44, No. 3 PP. 638-644, Aug 1995.

Y. Sanada and M. Nakagaw, “Power Control Techniques in a Multihop CDMA Packet Radio Network,” *IEICE Transactions on Communications*, Vol E79-B, No. 9, September 1996, pp. 1287-1294.

T. Rappaport, *Wireless Communications, Principles and Practice*, Prentice-Hall, Inc, Upper Saddle River, NJ, 1996

V. Garg and J Wilkes, *Wireless and Personal Communications Systems*, Prentice-Hall, Inc, Upper Saddle River, NJ, 1996

J. Wu and R. Kohno, “Performance Evaluation of Wireless Multimedia CDMA Networks Using Adaptive Transmission Control,” *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 9, December 1996, pp 1688-1697.

E. Kudow, I. Okazaki, and S. Ogose, “The Lower Bound of Open Loop Transmission Power Control Error in DS-CDMA System,” *IEICE Transactions on Communications*, Vol E80-B, No. 12, December 1997, pp. 1805-1809.

S. Lal and E. Sousa, "Distributed Resource Allocation for DS-CDMA Based Multi-Media Wireless LANs," *MILCOM 9*8, October 1998.

R. Kravets and P. Krishnan, "Power Management Techniques for Mobile Communication," *Proceedings of the 4th Annual IEEE/ACM Conference on Mobile Computing and Networking*, 1998, pp. 157 – 168.

G. Okamoto, *Smart Antenna Systems and Wireless LANS*, Klewer Academic Publishing, Boston, MA, 1999.

R. van Nobelen, N. Seshardi, J. Whitehead, and S. Timiri, "An Adaptive Radio Link Protocol with Enhanced Data Rates For GSM Evolution," *IEEE Personal Communications*, February 1999, pp 54-64.

S. Kim, "Adaptive Rate and Power DS/CDMA Communications in Fading Channels," *IEEE Communications Letters*, Vol. 3, No. 4, April 1999, pp. 85-87.

L. Larson, P. Asbeck, G. Hanington, E. Chen, A. Jayamaran, R. Langridge, and X. Zhang, "Device and Circuit Approaches for Improved Wireless Communications Transmitters," *IEEE Personal Communications*, October 1999, pp. 18-23.

Y. Lu, L. Benini, and G. de Micheli, "Low-Power Task Scheduling for Multiple Devices," ???? pp. 39 – 43.

MAX2740 Integrated GPS Receiver and Synthesizer Data Sheet, MAXIM Integrated Products 2000.

W. Su, S-J. Lee, and M. Gerla, "Mobility Prediction in Wireless Networks," *Proceedings IEEE Military Communications Conference 2000*, October 2000, pp. 491-495.

V. Wong and C. Leung, "A Transmit Power Control Scheme for Improving Performance in a Mobile Packet Radio System," *IEEE Transactions on Vehicular Technology*, Vol. 43, No. 1, Feb 1994, pp. 174-180.

Bibliography

[1] C. Wu and G. Ma, "Performance of Packet Reservation MAC Protocols for Wireless Networks," *48th IEEE Vehicular Technology Conference*, Vol. 3, 1998, pp. 2537 – 2541.

[2] http://www.networks.digital.com/dr/npg/deina-ch.html, February 1999.

[3] T. Yamawaki, et al, "A 2.7-V GSM RF Transceiver IC," *IEEE Journal of Solid State Circuits*, Vol. 32, No. 12, December 1997, pp. 2089–2095.

[4] Nokia C020/C021 Wireless Lan Card Technical Specifications, 1998

[5] http://www.aironet.com/products/in_building/PC4800B.asp.

[6] P802.11, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Press, May 1997.

[7] ETSI, EN300 652 V 1.2.1, "Broadband Radio Access Networks (BRAN); High Performance Radio Local Area Network (HIPERLAN) Type 1; Functional Specification, July 1998.

[8] M. Stemm and R. Katz, "Measuring and Reducing Energy Consumption of Network Interfaces in Hand-Held Devices," *IEICE Transactions on Communications*, Vol. E80-B, no. 8, Aug 1997 pp.1125-1131.

[9] L.M. Feeney and M. Nilsson, “Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment,” *INFOCOM 2001*, April 2001.

[10] J-C. Chen, K. M. Sivalingam , P. Agrawal, and S. Kishore, “A Comparison of MAC Protocols for Wireless Local Networks Based on Battery Power Consumption,” *INFOCOM 98*, March 1998, pp. 150 - 157.

[11] K. M. Sivalingam, M. Srivastava and P. Agrawal, "Low power link and access protocols for wireless multimedia networks", *IEEE Vehicular Technology Conference* , (Phoenix, AZ), pp. 1331-1335, May 1997.

[12] K. M. Sivalingam, J.-C. Chen and P. Agrawal, "Design and Analysis of Low-Power Access Protocols for Wireless and Mobile ATM Networks", *ACM/Baltzer Journal on Special Topics in Mobile Networking and Applications (MONET)*, June 1998.

[13] K. M. Sivalingam , Mani Srivastava, Prathima Agrawal and Jyh-Cheng Chen, "Low-power access protocols based on Scheduling for Wireless and Mobile ATM networks," *IEEE International Conference on Universal Personal Communications (ICUPC)*, (San Diego, CA), Oct. 1997.

[14] C. Rohl, H. Woesner, A Wolisz, “A Short Look on Power Saving Mechanisms in the Wireless LAN Standard Draft IEEE 802.11,” Source publication not known.

[15] H. Woesner, J. Ebert, M. Schlager, and A. Wolisz, “Power-Saving Mechanisms in Emerging Standards for Wireless LANs: The MAC Level Perspective,” *IEEE Personal Communications*, June 1998, pp. 40-48.

[16] T. Rappaport, *Wireless Communications, Principles and Practice*, Prentice-Hall, Inc, Upper Saddle River, NJ, 1996

[17] J. Andersen, T. Rappaport, and S. Yoshida, “Propagation Measurements and Models for Wireless Communications Channels,” *IEEE Communications Magazine*, Jan 1995, pp 42-49.

[18] V. Rodoplu and T. Meng, “Minimum Energy Mobile Wireless Networks,” *IEEE Journal on Selected Areas in Communication*, Vol. 17, No. 8, August 1999, pp. 1333-1344

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

[20] J. Gao, *Energy Efficient Routing For Wireless Sensor Networks*, PhD Dissertation, University of California, Los Angeles, 2000.

[21] S. Singh, M. Woo, and C. Raghavendra, “Power-Aware Routing in Mobile Ad Hoc Networks,” *The Fourth Annual ACM/IEEE International Conference in Mobile Computing and Networking*, 1998, pp. 181 – 190.

[22] I. Chlamtac, C. Petrioli, and J. Redi, “Energy-Conserving Access Protocols for Identification Networks,” *IEEE/ACM Transactions on Networking*, Vol. 7, No. 1, February 1999, pp. 51-59.

[23] E. Buffa and R. Saring, *Modern Production/Operations Management*, Eighth Edition, John Wiley & Sons, Inc., New York, NY, 1987, p. 290.

[24] L. Kleinrock and F. A. Tobagi, “Packet Switching in Radio Channels: Part I - Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics,” *IEEE Transactions on Communications*, December 1975, pp. 1400-1416.

[25] D. Bertsekas and R. Gallagher, *Data Networks*, Prentice Hall, Inc., Upper Saddle River, NJ, 1992.

[26] Y.K. Sun, K.C. Chen, and D. C. Twu, “Generalized Tree Multiple Access Protocols in Packet Switching Networks,” *Eighth International Symposium on Personal, Indoor and Mobile Radio Communications*, Vol. 3, 1997, pp. 918-922.

[27] Y.K. Sun and K.C. Chen, “Dynamics of Generalized Tree Protocols and Its Application to Delay Sensitive Traffic Control,” *1998 International Conference on Communications*, Vol. 2, 1998, pp. 849 – 853.

[28] Y.K. Sun and K.C. Chen, “Optimization of Generalized Tree Protocols with Capture,” *48th IEEE Vehicular Technology Conference*, Vol. 2, 1998 pp. 801 – 805.

[29] Y.K. Sun and K.C. Chen, " Energy Efficient Multiple Access Protocol Design," *IEEE Communications Letters*, Vol. 2 No. 12, December 1998, pp. 334 – 336.

[30] K.C.Chen and Y.K Sun, "A General Multi-layer Collision Resolution Multiple Access Protocol for Wireless Networks," *46th Vehicular Technology Conference*, 1996, pp. 116-120.

[31] G. Anastasi, L. Lenzini, and E. Mingozzi, "Stability and Performance Analysis of HIPERLAN," *INFOCOM '98*, 1998, pp. 134 – 141.

[32] J. Dai, "The Scheduling to Achieve Optimized Performance of Randomly Addressed Polling Protocol," *Wireless Personal Communications*, Vol. 15, No. 2, Kluwer Academic Publishers, November 2000, pp. 161-179.

[33] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 3, March 2000, pp 535 – 547.

[34] P. Jacquet, P. Minet, P. Mühlethaler, and N.Rivierre, "Priority and Collision Detection with Active Signaling – The Channel Access Mechanism of HIPERLAN," *Wireless Personal Communications*, Vol. 4, No. 1, Kluwer Academic Publishers, January 1997, pp. 11-26.

[35] F. A. Tobagi and L. Kleinrock, "Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy

Tone Solution," *IEEE Transactions on Communications*, December 1975, pp. 1417-1433.

[36] P. Karn, "MACA – A New Channel Access Method for Packet Radio" ARRL/CRRL Amateur Radio 9th Computer Networking Conference, 1990, pp. 134-140.

[37] C. L. Fulmer, Collision Avoidance Techniques for Packet-Radio Networks, PhD Dissertation, University of California, Santa Cruz, 1998.

[38] I. Vokovic, "Analysis of the Hidden Node Problem in HIPERLAN," *1999 IEEE Conference on Communications*, 1999, pp. 1234-1238.

[39] J. Stine and G. de Veciana, "Tactical Communications Using the 802.11 MAC Protocol," *MILCOM 98*, October 1998.

[40] J. Sobrinho and A.S. Krishnakumar, "Real-Time Traffic over the IEEE 802.11 Medium Access Control Layer," *Bell Labs Technical J*ournal, Autumn 1996, pp 172-187.

[41] J. Sobrinho and A. Krishnakumar, "Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 8, August 1999, pp. 1553 - 1568.

[42] J. Weinmiller, M. Schlager, A. Festag, and A. Wolisz, "Performance Study of Access Control in Wireless LANs – IEEE 802.11 DFWMAC and ETSI RES 10

Hiperlan," *ACM/Baltzer Mobile Networks and Applications,* Vol. 2, No. 1, 1997, pp. 55 – 67.

[43] MAX2740 Integrated GPS Receiver and Synthesizer Data Sheet, MAXIM Integrated Products 2000.

[44] R. Nelson and L. Kleinrock, "Maximum Probability of Successful Transmission in a Random Planar Packet Radio Network," *Proceedings of the IEEE INFOCOM '83,* 1983, pp. 365 – 370.

[45] T. Hou and V. Li, "Transmission Range Control in Multihop Packet Radio Networks," *IEEE Transactions on Communications*, Vol. 34, No. 1, January 1986, pp. 38-44.

[46] T. Hou and V. Li, "Performance Analysis of Routing Strategies in Multihop Packet Radio Network," *IEEE Global Communication Conference '84*, Atlanta, GA, November 1984, pp. 487-492.

[47] T. Hou and V. Li, "Performance Analysis of Multihop Packet Radio Networks with Capture," *IEEE International Conference on Communications, '85*, Chicago, IL, June 1985, pp. 1311-1317.

[48] L Kleinrock and J. Sylvester, "Optimum Transmission Radii for Packet Radio Networks or Why Six Is a Magic Number," *IEEE National Telecommunications Conference*, December 1978, pp. 4.3.1-4.3.5.

[49] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals," *IEEE Transactions on Communications*, Vol. 32, No. 3, March 1984, pp. 246-257.

[50] R. Kahn, "The Organization of Computer Resources into a Packet Radio Network," *IEEE Transactions on Communications*, Vol. 25, No. 1, January 1977, pp. 169 – 177.

[51] M. Pursley, "The Role of Spread Spectrum in Packet Radio Networks," *Proceedings of the IEEE*, Vol. 75, No. 1, January 1987, pp. 116 – 134.

[52] E. Sousa and J. Silvester, "Spread Spectrum Protocols for Distributed Spread-Spectrum Packet Radio Networks," *IEEE Transactions on Communications*, Vol. 36, No. 3, March 1988, pp. 272 – 281.

[53] L. Hu, "Distributed Code Assignment for CDMA Packet Radio Networks," *IEEE/ACM Transactions on Networking*, Vol. 1, No. 6, December 1993, pp. 668 – 677.

[54] C. Lin and M. Gerla, "Adaptive Clustering for Mobile Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 15, No. 7, September 1997, pp. 1265 – 1275.

[55] E. Sousa and J. Silvester, "Optimum Transmission Ranges in a Direct-Sequence Spread-Spectrum Multihop Packet Radio Network," *IEEE Journal on Selected Areas in Communications*, Vol. 8, No. 5, June 1990, pp. 762 – 771.

VITA

John Andrew Stine was born in Miami, Florida on October 4, 1959, the son of Mary Biggers Stine and Carl William Stine. After completing his work at Christopher Columbus Catholic Boys High School, Miami, Florida, in 1977, he entered the United State Military Academy at West Point, New York. He received the degree of Bachelor of Science from the Academy and was commissioned as an engineer officer in the United States Army in 1981. In 1990 he earned two degrees of Masters of Science in Engineering from the University of Texas at Austin. In 1992 he earned his registration as a professional engineer from the state of Virginia. In the summer of 1996 he again entered the Graduate School at the University of Texas.

Simultaneously, John has served in the Army for 20 years. Highlights of his military career include 7 years in troop leadership positions in combat engineer units, 3 years as an assistant professor of electrical engineering at the United States Military Academy, 3 years as an operations research systems analyst for the Army's Test and Experimentation Command, two years as a construction manager and then one year as chief of engineering for the Technology Development Test Division of the Defense Threat Reduction Agency. He served one tour overseas in the Federal Republic of Germany from 1982-1985. Also, in 1985 he married the former Martha J. Byrd of San Marcos, Texas and together they now have six children.

Permanent Address: 402 Lamar Avenue, San Marcos, Texas 78666

This dissertation was typed by the author.