MEDIUM ACCESS CONTROL AND ADAPTIVE TRANSMISSION TECHNIQUES IN WIRELESS NETWORKS

by

Alaa Hilal Essa Muqattash

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2 0 0 5

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Alaa Hilal Essa Muqattash
entitled Medium Access Control and Adaptive Transmission Techniques in Wireless Networks
and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 23 September 2005
Marwan Krunz

Date: 23 September 2005
William E. Ryan

Date: 23 September 2005
Srinivasan Ramasubramanian

Date: 23 September 2005

Date: 23 September 2005

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 23 September 2005
Dissertation Director: Marwan Krunz

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: _____

Alaa Hilal Essa Muqattash

ACKNOWLEDGEMENTS

I would like to thank many people, without whom I would not have been able to finish this dissertation. First, I would like to thank my advisor, Professor Marwan Krunz, for his guidance, support, and assistance in developing and refining my research. His technical advice and encouragement have provided me with the skills needed to carry me through the completion of this dissertation. I would also like to extend my thanks to Professor William E. Ryan and Professor Srinivasan Ramasubramanian who generously agreed to participate in my dissertation committee, and provided me with valuable feedback. I also wish to thank all my colleagues at the Broadband Networking Laboratory at The University of Arizona. Special thanks to Dr. Tao Shu who introduced me to the area of mathematical optimizations. I would also like to thank Dr. Ghobad Heidari and the whole group at Olympus Communication Technology of America with whom I started my career while finishing this work.

Far too many friends to mention individually have assisted in so many ways during my work. A huge thank-you goes to all of my friends.

Finally, I am forever indebted to my parents, brothers, and sisters for their unbounded support and encouragement and for always being there for me. My father's great advice, my mother's prayers, and their love and dedication provided the foundations for this work.

DEDICATION

I would like to dedicate this dissertation to my parents

Hilal Muqattash and ***Seham El-Nimri.***

Their constant love and caring are the reasons for where I am and what I am. My gratitude and my love to them are beyond words.

I would also like to dedicate this dissertation to my home country

Jordan,

which I will serve until the last day of my life.

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF TABLES

ABSTRACT

Efficient utilization of the limited wireless spectrum while satisfying applications' quality of service requirements is an essential design goal of forthcoming wireless networks and a key to their successful deployment. The need for spectrally efficient systems has motivated the development of adaptive transmission techniques. Enabling this adaptation requires protocols for information exchange as well as mathematical tools to optimize the controllable parameters. In this dissertation, we provide insights into such protocols and mathematical tools that target efficient utilization of the wireless spectrum. First, we propose a distributed CDMA-based medium access protocol for mobile ad hoc networks (MANETs). Our approach accounts for multiple access interference at the protocol level, thereby addressing the notorious near-far problem that undermines the throughput performance in MANETs. Second, we present a novel power-controlled MAC protocol, called POWMAC, which enjoys the same single-channel, single-transceiver design of the IEEE 802.11 Ad Hoc MAC protocol, but which achieves a significant throughput improvement over the 802.11 protocol. Third, we consider joint power/rate optimization in the context of orthogonal modulation (OM) and investigate the performance gains achieved through adaptation of the OM order using recently developed optimization techniques. We show that such adaptation can significantly increase network throughput while simultaneously reducing the per-bit energy consumption relative to fixed-order modulation systems. Finally, we determine the maximum achievable "performance" of a wireless CDMA network that employs a conventional matched filter receiver and that operates under *optimal* link-layer adaptation where each user individually achieves the Shannon capacity. The derived bounds serve as benchmarks against which adaptive CDMA systems can be compared.

CHAPTER 1

INTRODUCTION

1.1 Motivation and General Scope

Over the last decade, we have witnessed a rapid increase in the demand for wireless communication access. While most of this growth has been in basic voice services, there is a clear trend towards multimedia communications, which require higher data rates and more stringent quality of service (QoS) than voice services. To address this demand for more reliability and "bits/sec," researchers have been looking into novel solutions, including flexible network paradigms, new transmission technologies, and efficient resource allocation policies.

Wireless networks have traditionally been designed according to a cellular structure, where a geographical area is divided into regions called cells. A mobile node located in a given cell communicates directly with a stationary base station located at the center of that cell. Cellular networks rely heavily on an existing infrastructure, and mobile users can communicate only if such infrastructure is in place. The available network resources are limited in cellular networks and can only be increased by adding more infrastructure and upgrading existing ones, which may incur significant cost.

An alternative approach to the cellular paradigm is the *ad hoc paradigm*. A mobile ad hoc network (MANET) is an autonomous system of mobile nodes connected by wireless links. Nodes are free to move randomly and organize themselves arbitrarily. Also known as peer-to-peer or infrastructureless networks, MANETs do not require a pre-existing infrastructure or centralized control. Figure 1.1 shows an example of an ad hoc and a cellular network. MANETs are typically used in situations where network connectivity is temporarily needed or where it is infeasible (or expensive) to install a fixed infrastructure network. Examples of such situations

are meetings, lectures, crowd control, search-and-rescue missions, disaster recovery, and battlefield scenarios. The emerging wireless personal area networks (WPAN) [6] is an example of an indoor MANET that will be used to connect home appliances. Ad-hoc-at-the-last-mile (i.e., within each cell) is another example of a MANET that will likely increase in popularity, and may become the dominant solution for wireless Internet access. The solutions proposed in the first two chapters of this dissertation are completely distributed, and so they are primarily targeted to ad hoc networks, while the rest of the dissertation is aimed at both ad hoc as well as cellular networks.

(a) An ad hoc network. (b) A cellular network.

Figure 1.1: Network topologies considered in the paper.

With about two billion current wireless customers worldwide [40] requiring a mixture of real-time traffic (e.g., voice, video, etc.) and data traffic (e.g., email, messaging, web browsing, file transfer, etc.), efficient utilization of the limited wireless spectrum became an essential design goal of fourth-generation (4G) wireless networks and a key to their successful deployment [131]. The need for spectrally efficient systems has motivated the development of *adaptive transmission techniques*, several of which are in the process of being standardized. Such techniques adapt users' transmission parameters (e.g., power, rate, modulation order, etc.) according to the time-varying channel conditions, interference levels, rate requirements, bit error rate (BER) needs, and energy constraints [83]. Enabling this adaptation requires protocols for information exchange as well as mathematical tools or algorithms that use this information to optimize the users' parameters by maximizing certain utility

functions. In this dissertation, we develop protocols and mathematical tools that target efficient utilization of the wireless spectrum.

With regard to protocols, several medium access control (MAC) schemes for code division multiple access (CDMA) based MANETs have been proposed in the literature (e.g., [108, 54, 73, 38, 60]). These schemes, in general, are based on random channel access, whereby a terminal with a packet to transmit can proceed immediately with its transmission (starting, possibly, with the exchange of control packets), irrespective of the state of the channel. We refer to such schemes as *random access CDMA* (RA-CDMA). Under appropriate code assignment and spreading-code schemes, RA-CDMA protocols are guaranteed to be free of primary collisions[1]. However, as explained in Chapter 2, the nonzero cross-correlations between different CDMA codes can induce *multi-access interference* (MAI), resulting in secondary collisions at a receiver (collisions between two or more transmissions that use different CDMA codes). As shown later in the dissertation, this MAI problem can cause a significant reduction in network throughput, and hence cannot be overlooked when designing CDMA-based MAC protocols for MANETs. Accordingly, one of the main goals of this dissertation is to provide a CDMA-based MAC solution for MANETs that addresses the MAI problem at the protocol level.

The Ad Hoc mode of the IEEE 802.11 standard [2] is largely considered the *de facto* MAC protocol for MANETs. Despite its appealing simplicity, this protocol can be overly conservative, leading to an unnecessary reduction in network throughput. To address this problem, many transmission power control (TPC) protocols have emerged to improve network throughput by means of increasing the channel spatial reuse. Theoretical studies [45] and simulation results [80, 81] have demonstrated that TPC can provide significant gains in capacity and energy consumption, not to mention its benefits in providing admission control and in QoS provisioning [14]. In the third chapter of this dissertation, we propose a TPC scheme that is based on a single-channel, single-transceiver approach, and that is shown to provide a significantly higher network throughput than the IEEE 802.11 scheme while yet

[1] A primary collision involves two or more transmissions that are spread using the same code.

preserving the collision avoidance properties of the IEEE 802.11 scheme.

Next, we focus on CDMA systems for which coherent reception is not possible and where orthogonal modulation (OM) is used (e.g., uplink IS-95 [92]). For such systems, classical (i.e., fixed OM order) variable processing gain (VPG) techniques have been the main adaptation approach because of their performance benefits, flexibility, and practicality (e.g., low peak-to-mean envelope power, fixed chip rate, etc. [57]). The extensive work on VPG has clearly quantified the performance advantages of combined adaptive rate/power control over power control alone [105, 57]. However, adapting the modulation order for variable-rate OM-based systems remains an unexplored area of research, and one for which joint rate/power control has not yet been investigated. In this dissertation, we investigate the theoretical performance limits of *joint* rate/power control for adaptive orthogonal modulation (AOM) CDMA networks and gain insights into the technique itself. To this end, we purse an algorithmic approach, making use of recent development in the area of convex optimization techniques [18].

Given the growing interest in adaptation techniques, there is a need to answer a fundamental question: For given channel conditions, users' requirements (e.g., minimum rate, maximum BER), and system constraints (e.g., maximum power), what is the maximum instantaneous achievable performance[2] of a CDMA network that employs a matched filter (MF) receiver[3] and that operates under "optimal" link-layer adaptation? In this context, "optimal adaptation" refers to any combination of transmission powers, user rates, modulation orders, and coding rates, optimized for given instantaneous channel conditions. In this dissertation, we answer the above question using recently developed nonlinear optimization tools.

[2]Throughout this dissertation, the term "performance" is used to refer to network throughput and/or per-bit energy consumption.

[3]MF receivers, popular for their low complexity and cost, treat users interference as Gaussian noise.

1.2 Main Contributions and Dissertation Overview

The main contributions of this dissertation are as follows:

- In Chapter 2, we propose a distributed CDMA-based MAC protocol for MANETs. By using power control, our approach accounts for the MAI at the protocol level. To the best of our knowledge, this is the first attempt to address the MAI problem in the design of MAC protocols for MANETs. Simulation results indicate that the proposed protocol achieves a significant increase in network throughput relative to the 802.11 approach, at no additional cost in energy consumption.

- Many TPC schemes for MANETs have been proposed in the literature. However, these schemes suffer from one or more of the following deficiencies: (1) the TPC approach may yield energy reduction but not throughput gain, (2) the MAC design may not support collision avoidance, resulting in the well-known hidden terminal problem, (3) the TPC approach requires extra hardware (e.g., multiple transceivers), (4) lack of link-layer reliability, i.e., ACK packets are not protected, and (5) many of the assumptions made in the MAC design are unrealistic (e.g., channel gain is assumed to be the same for both the control and data channels). In Chapter 3, we introduce a new TPC scheme, called POWMAC (Power Controlled MAC), that ameliorates these deficiencies. Our scheme is based on a single-channel, single-transceiver approach, and is shown to provide a significantly higher network throughput than the IEEE 802.11 scheme while preserving the collision avoidance properties of the IEEE 802.11 scheme. To the best of our knowledge, this is the first TPC solution that is based on a single-channel, single-transceiver design, that can increase the throughput of a MANET relative to the IEEE 802.11 scheme, and that supports link-layer reliability. Simulation results are used to demonstrate the throughput and energy gains that can be obtained under the POWMAC protocol.

- In Chapter 4, we consider joint power/rate optimization in the context of OM and investigate the performance gains achieved through adaptation of the *OM order*. We show that such adaptation can significantly increase network throughput while simultaneously reducing the per-bit energy consumption relative to fixed-order modulation systems. The optimization is carried out under two different objective functions: minimizing the maximum service time and maximizing the sum of user rates. For the first objective function, we prove that the optimization problem can be formulated as a generalized geometric program (GGP). We then show how this GGP can be transformed into a nonlinear convex program, which can be solved optimally and efficiently. For the second objective function, we obtain a lower bound on the performance gain of AOM over fixed-modulation systems. Numerical results indicate that relative to an optimal joint rate/power control fixed-order modulation scheme, the proposed AOM scheme achieves significant throughput and energy gains.

- In Chapter 5, we determine the maximum achievable "performance" of a wireless CDMA network that employs a conventional MF receiver and that operates under *optimal* link-layer adaptation where each user individually achieves the Shannon capacity. The derived bounds serve as benchmarks against which adaptive CDMA systems can be compared. We focus on two throughput-related optimization criteria: minimizing the maximum service time and maximizing the sum of users' rates. We show that the problem of joint optimization of the transmission powers and rates so as to minimize the maximum service time can be formulated as a GGP, which can be transformed into a nonlinear convex problem and solved optimally and efficiently. When the goal is to maximize the sum of the rates, we show that the problem can be *approximated* as a GGP.

The conclusions and some directions for future research are drawn in Chapter 6.

CHAPTER 2

CDMA-BASED MAC PROTOCOL FOR WIRELESS AD HOC NETWORKS

2.1 Introduction

2.1.1 Motivation

One of the fundamental challenges in MANETs research is how to increase the overall network throughput while maintaining low energy consumption for packet processing and communications. The low throughput is attributed to the harsh characteristics of the radio channel combined with the contention-based nature of medium access control MAC protocols commonly used in MANETs. The focus of this chapter is on improving the network throughput of a MANET by means of a CDMA-based MAC protocol. Compared to the DCF (Distributed Coordination Function) mode of the IEEE 802.11 standard [2], which is currently the *de facto* MAC protocol for MANETs, the proposed protocol is shown to achieve a significant increase in network throughput for the same or less energy consumption per delivered packet.

CDMA is based on spread spectrum (SS) techniques, in which each user occupies the entire available bandwidth. At the transmitter, a digital signal of bandwidth, say, B_1 bits/sec is *spread* using (i.e., multiplied by) a pseudo-random noise (PN) code of bandwidth, say, B_2 bits/sec ($B_2/B_1 \gg 1$ is called the *processing gain*). The PN code is a binary sequence that statistically satisfies the requirement of a random sequence, but that can be exactly reproduced at the intended receiver. Using a locally generated PN code, the receiver *de-spreads* the received signal, recovering from it the original information. The enhancement in performance obtained from spreading the signal makes it possible for several, independently coded signals to occupy the same channel bandwidth, provided that each signal has a distinct PN code. This type of communication in which each transmitter-receiver pair has a

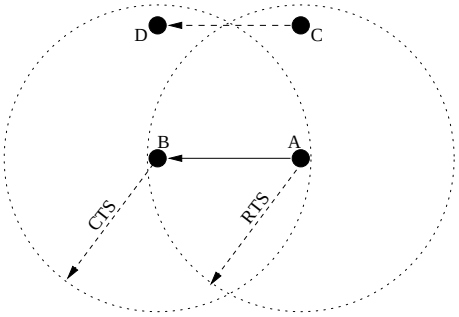


Figure 2.1: Example showing the low throughput of the 802.11 scheme (only one transmission can take place at a time).

distinct PN code for transmitting over a common channel is called code division multiple access (CDMA) [92].

Due to its superior characteristics, CDMA has been the access technology of choice in cellular systems, including the recently adopted 3G systems [87]. In such systems, CDMA has been shown to provide up to six times the capacity of TDMA- or FDMA-based solutions [39]. This throughput gain comes along with other desirable features, including graceful signal degradation, multipath resistance, inherent frequency diversity, and interference rejection. It is, therefore, of no surprise that CDMA is being considered for ad hoc networks. Interestingly, the IEEE 802.11 standard uses SS techniques at the physical layer[1], but only to mitigate the impact of the harsh wireless channel. More specifically, in the 802.11 protocol all transmitted signals are spread using a *common* PN code, precluding the possibility of multiple concurrent transmissions in the vicinity of a receiver. This situation is exemplified in Figure 2.1, where the transmissions $A \rightarrow B$ and $C \rightarrow D$ cannot take place at the same time.

[1]Both direct sequence spread spectrum (DSSS) and frequency hopping spread spectrum (FHSS) are included in the IEEE 802.11 specifications. However, since DSSS has more desirable properties than FHSS, it has been favored in recent wireless standards, including IS-95 [92]. Accordingly, our focus in this chapter is on DSSS techniques.

2.1.2 Code Assignment and Spreading Protocol Issues

Enabling CDMA-based solutions for MANETs is fraught with challenges, which are essentially related to the absence of centralized control (i.e., a base station). First, a *code assignment protocol* is needed to assign distinct codes to different terminals. This problem is trivial in small networks, but becomes acute in large networks where the number of PN codes is smaller than the number of terminals[2], necessitating spatial reuse of the PN codes. Several code assignment protocols have been previously proposed (e.g., [53, 16, 38]). In general, these protocols attempt to assign codes to nodes with the constraint that all neighbors of a node have different PN codes [53].

Besides the code assignment protocol, a *spreading-code protocol* is also needed to decide which codes to use for packet transmission and for monitoring the channel in anticipation of a packet reception [108]. Such a protocol can be receiver-based, transmitter-based, or a hybrid. In a receiver-based protocol, the transmitter uses the code of the intended receiver to spread the packet, while an idle terminal constantly monitors its own code. This approach simplifies the receiver's circuitry because the receiver does not have to monitor the whole code set. Unfortunately, *primary collisions* are still possible, even under a correct code assignment (a primary collision involves two or more transmissions that are spread using the same code). For example, consider two non-neighboring nodes A and C that have two different codes. These nodes may have a common neighbor, say B, with its own code. A primary collision may occur if nodes A and C simultaneously attempt to transmit to node B using B's code[3]. The only way to guarantee that primary collisions cannot happen is to use different codes for different, concurrently transmitted *signals* (not nodes). Another disadvantage of the receiver-based approach is that a broadcast message requires that the transmitter unicasts the message to each receiver.

[2]The number of codes is usually constrained by the available spectrum and the required information data rate [107].

[3]Note, however, that if the received power of one signal is much greater than the other, then capture is still possible and the stronger signal can still be received correctly.

In a transmitter-based spreading protocol, a transmission code is assigned to each terminal, and receivers must be able to monitor the activity on the whole set of PN codes. The advantage of this approach is that primary collisions cannot happen. In addition, broadcast is inherently supported. However, the drawback is that the receiver circuitry is very complex and expensive.

Various hybrids of the above two approaches are also possible. For example, the authors in [108] proposed two hybrid schemes: the *common-transmitter-based protocol* and the *receiver-transmitter-based protocol*. In the first protocol, the fields in the packet header that contain the source and destination addresses are spread using a common code, while the rest of the packet is spread using the transmitter's code. An idle terminal constantly monitors the common code. Upon recognizing its address in the destination field, the listening terminal switches to the code of the transmitting node to receive the rest of the packet. The receiver-transmitter-based protocol works similarly, but with the common code replaced with the receiver's code.

2.1.3 Goals and Contributions

Several CDMA-based MAC protocols for MANETs have been proposed in the literature (e.g., [108, 54, 73, 38, 60]). These protocols are, in general, based on random channel access, whereby a terminal with a packet to transmit can proceed immediately with its transmission (starting, possibly, with an request-to-send (RTS), clear-to-send (CTS) exchange), irrespective of the state of the channel. We refer to such schemes as *random access CDMA* (RA-CDMA). Under appropriate code assignment and spreading-code schemes, RA-CDMA protocols are guaranteed to be free of primary collisions. However, as explained in detail in Section 2.2, the nonzero cross-correlations between different CDMA codes can induce MAI, resulting in *secondary collisions* at a receiver (collisions between two or more transmissions that use different CDMA codes). In the literature, this problem is known as the near-far problem [91]. As shown in Section 2.2, the near-far problem can cause a significant reduction in network throughput, and hence cannot be overlooked when designing

CDMA-based MAC protocols for MANETs. Accordingly, the main goal of this chapter is to provide a CDMA-based MAC solution for MANETs that addresses the near-far problem. In our protocol, the transmission powers are dynamically adjusted such that the MAI at any receiver is not strong enough to cause a secondary collision. As indicated in our simulations, this results in a significant improvement in network throughput at no additional cost in energy consumption. In fact, the proposed protocol is shown to achieve some energy saving compared to the 802.11 scheme. We further investigate the applicability of variable processing gain as a means to increase the capacity of the network.

The rest of the chapter is organized as follows. In Section 2.2, we explain in details the near-far problem in MANETs and show its adverse effect on the throughput performance. Section 2.3 provides an overview of related CDMA-based protocols for MANETs. The proposed protocol is presented in Section 2.4, followed by simulation results and discussion in Section 2.5.

2.2 Near-Far Problem in RA-CDMA

2.2.1 Imperfect Orthogonality of CDMA Codes

The roots of the near-far problem lies in the fact that, unlike FDMA and TDMA channels which can be completely orthogonal, CDMA codes suffer from nonzero cross-correlation between codes. When a CDMA receiver de-spreads a signal, it effectively computes the cross-correlation between the signal and a locally generated PN sequence. If this PN sequence is identical to the one used to spread the signal at the transmitter (i.e., the message is intended for *this* receiver), cross-correlation computations restore the original information data. Otherwise, such computations result in either a zero or a nonzero value, depending on whether the system is *synchronous* or *asynchronous*.

A system is called time-synchronous if all signals originate from the same transmitter, as in the case of the downlink of a cellular CDMA network[4]. Here, synchrony

[4]In theory, it is possible to have multiple transmitters in a synchronous system. However, in

is manifested in two ways. First, different transmissions that are intended for different receivers will have a common time reference. Second, from the viewpoint of a given mobile terminal, all signals (intended or not) propagate through the same paths, and thus suffer the same time delays. In synchronous systems, it is possible to design *completely orthogonal* spreading codes. In fact, in the IS-95 standard for cellular CDMA networks [92], each user is assigned a Hadamard (or Walsh) code. These codes are orthogonal and are used to "channelize" the available bandwidth.

On the other hand, a system is called time-asynchronous if signals originate from multiple transmitters, as in the case of the uplink of cellular networks and also in MANETs. The reason behind this terminology is twofold. First, since signals originate from *different* transmitters, it is generally not feasible to have a common time reference for all the transmissions that arrive at a receiver. Second, these transmissions propagate through different paths; thus, they suffer different time delays [93]. In an asynchronous system, it is *not* possible to design spreading codes that are orthogonal for all time offsets [92]. In this case, the cross-correlation between codes cannot be neglected. In fact, codes that are orthogonal in synchronous systems (e.g., Hadamard codes) exhibit high cross-correlation when not perfectly synchronized. Instead, PN codes that are designed specifically to have low cross-correlation are used.

While the code design problem is crucial in determining the system performance, of greater importance is the problem of nonzero cross-correlation of the PN codes [91]. Unintended transmissions add nonzero MAI during the de-spreading at a receiver. The near-far problem is a severe instance of MAI, whereby a receiver that is trying to detect the signal of the ith transmitter may be much closer in distance to, say, the jth transmitter than the ith transmitter. Under equal transmission powers, the signal from the jth transmitter will arrive at the receiver in question with a substantially larger power than that of the ith transmitter, causing incorrect decoding of the ith transmission (i.e., a secondary collision).

practice, it is difficult to achieve perfect synchronization between those transmitters.

2.2.2 Impact of the MAI Problem

We now elaborate on the performance implications of the MAI problem. Consider the reception of a packet at terminal i. Let $P_0^{(i)}$ be the average received power of the desired signal at the ith terminal. Suppose that there are K interfering transmissions with received powers $P_j^{(i)}$, $j = 1, \ldots, K$. The quality of the intended reception is adequately measured by the *effective bit energy-to-noise spectral density ratio* at the detector, denoted by $\mu^{(i)}$. For an asynchronous direct-sequence BPSK system, $\mu^{(i)}$ is given by [100, 109][5]:

$$\mu^{(i)} \triangleq \frac{E_b}{N_{0\text{eff}}} = \left(\frac{2 \sum_{j=1}^{K} P_j^{(i)}}{3W P_0^{(i)}} + \frac{1}{\mu_0} \right)^{-1} . \tag{2.1}$$

where W is the processing gain and μ_0 is the $E_b/N_{0\text{eff}}$ ratio at the detector in the absence of interference. As the interfering power increases, $\mu^{(i)}$ decreases, and the bit error probability increases.

As an example, consider a CDMA system that uses BPSK modulation and a convolutional code with rate 1/2, constraint length 7, and soft decision Viterbi decoding. Let $W = 100$. To achieve a bit error probability of 10^{-6}, the required $E_b/N_{0\text{eff}}$ is 5.0 dB [92]. Ignoring the thermal noise and using (2.1), the total interference power must satisfy:

$$\frac{\sum_{j=1}^{K} P_j^{(i)}}{P_0^{(i)}} \leq 47.43 \tag{2.2}$$

Transmitters are, in general, situated at different distances from the receiver. Suppose that the transmission powers are fixed and equal. Consider the case of one interferer ($K = 1$) at distance d_1 from the receiver. Let d_0 be the distance between this receiver and its intended transmitter. Using the two-ray propagation model for terrestrial communications (power loss $\sim 1/d^4$), it is easy to show that to satisfy the

[5] Assuming truly random sequences of rectangular chip pulses and using a Gaussian approximation with constant but unequal powers.

required bit error rate, we must have $d_1 \geq 0.38 d_0$. So if there is only one interferer that is at distance less than $0.38\, d_0$ from the receiver, reliable communication will not be possible (i.e., a secondary collision will occur).

The above example shows that the near-far problem can severely affect packet reception, and consequently, network throughput. A good measure of network throughput is given by the *expected forward progress* (EFP) per transmission, defined as the product of the *local* throughput of a terminal and the distance between the transmitter and the receiver [109]. The EFP was derived in [109] for multihop RA-CDMA networks, assuming a slotted system and Poisson distributed terminals in 2D space. Let p be the probability that a terminal is transmitting a packet in a given time slot (i.e., the per-node load) and let L be the number of nodes that are within a circle centered at the transmitter and of radius that equals the transmitter-receiver separation distance. A scaled version of the EFP is plotted in Figure 2.2 as a function of p for various values of L. The figure shows that the EFP initially increases with p up to some point, say p^*, beyond which the EFP starts to decrease rapidly with p. This says that the channel becomes unstable when the load exceeds p^*, which is caused by the increase in the number of transmitted packets beyond the multiple access capability of the system. Our goal is to design a CDMA-based MAC protocol that prevents this rapid degradation in network throughput and that increases the throughput relative to the 802.11 approach.

2.3 Related Work

In [108] the address field in each packet is spread using a common code, while the rest of the packet is spread using the transmitter-based approach. A receiver notes the address of the source terminal and uses this address to switch to the corresponding code. In [54] the authors proposed the coded tone sense protocol, in which K busy tones are associated with K spreading codes. During packet reception on a certain code, the receiving station broadcasts the corresponding busy

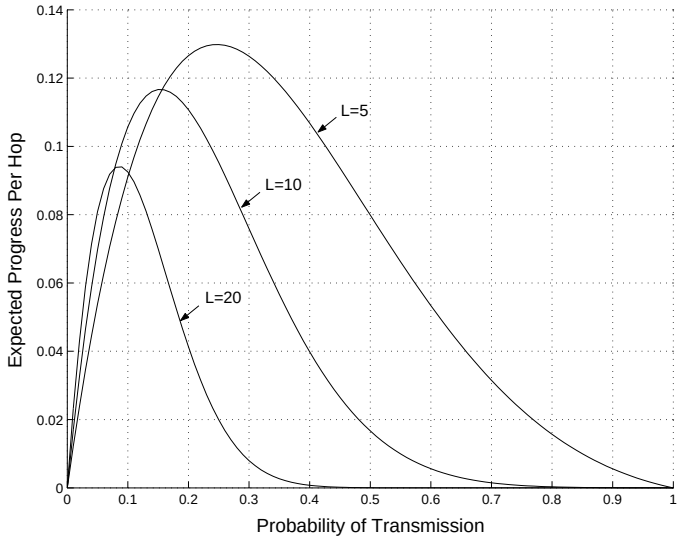


Figure 2.2: Throughput performance versus load in RA-CDMA networks.

tone. In [38] all terminals send the RTS-CTS packets on a common code, while the data packets are sent using a transmitter- or a receiver-based approach. Somewhat similar approaches were proposed in [60] and [125]. In all the above protocols, the authors assume perfect orthogonality among spreading codes, i.e., they ignore the near-far problem.

A reservation-based scheme was proposed in [124], whereby small control packets are used to request slot assignments for data packets. The authors investigated the use of FHSS to avoid MAI. Their approach, however, cannot be used for DSSS, which is the method of choice in recent wireless standards (e.g. IS-95).

In [27] and [35] the authors proposed distributed channel assignment algorithms for SS multihop networks. Those protocols, however, do not allow for any MAI, and hence cannot support concurrent transmissions of signals with different codes. Clustering, as proposed in [72], is another interesting approach for power control in CDMA networks. It simplifies the forwarding function for most terminals, but at the expense of reducing network utilization (since all communications have to go through the cluster heads). This can also lead to the creation of bottlenecks.

In [104] the authors proposed the use of a multiuser detection circuit at the

receiver to mitigate the near-far problem in MANETs. The proposed scheme also requires the use of GPS receivers to provide accurate position and timing information. Such a scheme relies heavily on physical layer techniques to mitigate MAI, and makes no effort to account for MAI at the MAC layer. Moreover, although it is feasible to deploy multiuser GPS receivers at the base station, presently it is impractical (and expensive) to implement such receivers within the mobile terminal. Recently, an interesting approach for joint scheduling and power control in ad hoc networks was proposed [33]. This approach, however, requires a central controller for executing the scheduling algorithm, i.e., it is not a truly distributed solution. Furthermore, it assumes the existence of a separate feedback channel that enables receivers to send their SNR measurements to their respective transmitters in a contention free manner.

In [24] and [32] the authors analyzed RA-CDMA protocols for MANETs in the presence of MAI. They assumed that transmissions of all neighbors produce the same noise effect, and therefore, the SNR threshold can be converted into a threshold on the number of transmissions (n) in the receiver's neighborhood. A packet is correctly received when that number is less than the predetermined threshold n. Hence, the protocol was called CDMA/n. Although such an approximation may not be accurate in topologies where nodes are not equally spaced, it shows that MAI can significantly degrade network performance.

2.4 Proposed Protocol

2.4.1 Protocol Intuition and Design Goals

Before presenting the operational details of the protocol, it is instructive to first discuss how the near-far problem is being addressed in cellular networks and why the same solution cannot be extended to MANETs. In the uplink of a cellular CDMA system, the near-far problem is combated through a combination of open- and closed-loop power control, which ensures that each mobile terminal generates the *same* signal power at the base station. The base station monitors the received

Figure 2.3: Example that demonstrates that power control alone is not enough to combat the near-far problem in MANETs.

signal power from each terminal and instructs distant terminals to increase their signal powers and nearby terminals to decrease theirs.

Unfortunately, the same solution cannot be used in MANETs. To see why, consider the situation in Figure 2.3. Let d_{ij} denote the distance between nodes i and j. Suppose that A wants to communicate with B using a given code and C wants to communicate with D using a different code. Suppose that $d_{AB} \approx d_{CD}$, $d_{CB} \ll d_{AB}$, and $d_{AD} \ll d_{CD}$. Then, the MAI caused by C makes it impossible for B to receive A's transmission. Similarly, the MAI caused by A makes it impossible for D to receive C's transmission. It is important to note that the two transmissions cannot take place simultaneously, irrespective of what transmission powers are selected (e.g., if A increases its power to combat the MAI at B, then this increased power will destroy the reception at D).

The above example reveals two issues. First, it may not be possible for two transmissions that use two different spreading codes to occur simultaneously. Obviously, this is a medium access problem. Second, the two transmissions can occur simultaneously if the terminals adjust their signal powers so that the interference caused by one transmission is not large enough to destroy packet reception at other terminals. Obviously, this is a power control problem. So the solution to the near-far problem has to have both elements: power control and medium access.

It is important here to differentiate between the spreading code protocol and the MAC protocol. The former decides which PN code is used to spread the signal, but does not solve the contention on the medium. On the other hand, the MAC protocol is responsible for minimizing or eliminating collisions, thereby achieving

good utilization of the available bandwidth. The use of the MAC protocol implies that even if a terminal has an available spreading code, it may not be allowed to transmit.

The design of our MAC protocol, described in detail in subsequent sections, is guided by the following objectives:

- The protocol must be asynchronous, distributed, and scalable for large networks. It must also involve minimal exchange of information and must be suitable for real-time implementation.
- The receiver circuitry should not be overly complex in the sense that it should not be required to monitor the whole code set.
- The protocol should adapt to channel changes and mobility patterns.
- Finally, although we assume that a code assignment protocol is running at a higher layer, the MAC protocol must minimize (or eliminate) collisions even if the code assignment is not "correct". This is important because it is usually difficult to guarantee correct code assignment at all times when network topology is continuously changing.

2.4.2 Architecture

In our design, we use two *frequency* channels, one for data and one for control (i.e., FDMA-like partitioning). A common spreading code is used by all nodes to communicate over the control channel, while several terminal-specific codes can be used over the data channel. This architecture is shown in Figure 2.4. Note that the different codes used over the data channel are not perfectly orthogonal. However, because of the frequency separation, a signal over the control channel is *completely* orthogonal to any signal (or code) over the data channel. The splitting of the available bandwidth into two non-overlapping frequency bands is fundamentally needed to allow a terminal to transmit and receive simultaneously over the control and data channels, *irrespective of the signal power*. As we explain shortly, our

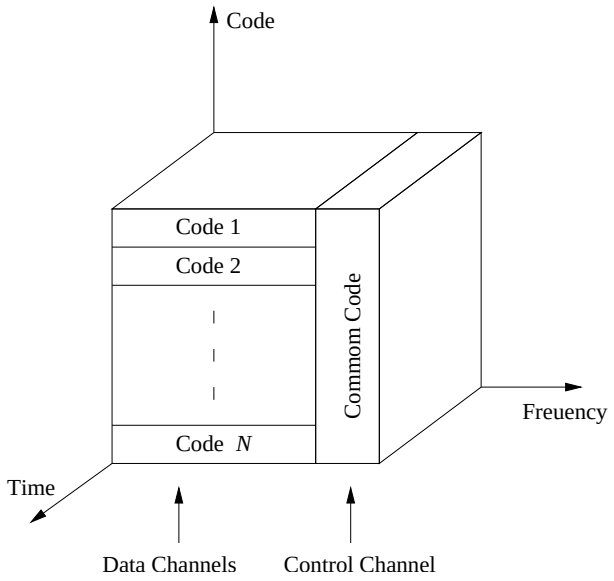


Figure 2.4: Data and control codes in the proposed protocol.

protocol utilizes this fact to allow interference-limited transmissions that use (quasi-orthogonal) data channel codes to proceed concurrently.

2.4.3 Channel Model and Protocol Assumptions

In designing our protocol, we assume that: (1) the channel gain is stationary for the duration of the control and the ensuing data packet transmission periods; (2) the gain between two terminals is the same in both directions; and (3) data and control packets between a pair of terminals observe similar channel gains.

In addition to the above assumptions, we assume that the radio interface can provide the MAC layer with the average power of a received control signal as well as the average interference power. Each terminal is equipped with two transceivers and carrier-sense hardware that senses the control channel for any carrier signal. No carrier-sense is needed for the data channel. The carrier frequency spacing between the control and data channels is enough to ensure that the outgoing signal on one channel does not interfere with the incoming signal on the other channel.

2.4.4 Controlled Access CDMA (CA-CDMA) Protocol

Our CA-CDMA protocol is contention based and uses a modified RTS-CTS reservation mechanism. RTS and CTS packets are transmitted over the control channel (using a common code) at a fixed (maximum) power $P_{\max}$. These packets are received by all potentially interfering nodes, as in the IEEE 802.11 scheme. However, in contrast to the IEEE 802.11 scheme and RA-CDMA protocols, interfering nodes *may* be allowed to transmit concurrently, depending on some criteria that will be discussed later. For the ensuing data packet, the receiver and the transmitter must agree on two parameters: the spreading code and the transmission power. Code selection can be done according to any code assignment scheme. As explained later, even if the code assignment scheme is not correct, our protocol will still function properly. The choice of the power level is critical and represents a tradeoff between link quality and MAI. More specifically, as the transmission power increases, the bit error rate at the intended receiver decreases (i.e., link quality improves), but the MAI added to other ongoing receptions increases (i.e., the quality of these receptions deteriorates). In addition to accounting for these two factors, our protocol also incorporates an interference margin in the power computations. This margin allows terminals at some interfering distance from the intended receiver to start new transmissions in the future. The computation of this margin is discussed in Section 2.4.5.

In the CA-CDMA protocol, terminals exploit knowledge of the power levels of the overheard RTS and CTS messages to determine the power that they can use without disturbing the ongoing receptions. In Section 2.4.6 we develop a distributed admission control strategy that decides when terminals at some distance can proceed concurrently with their transmissions.

We note here that the CA-CDMA protocol is, to some extent, similar to Qualcomm's CDMA protocol [69], adopted by the US Telecommunication Industry Association as the IS-95 standard for cellular networks. In both protocols, users contend on a control channel to request "network resources." However, the interpretation of

"resources" is different in the two protocols; in the Qualcomm protocol, it refers to connection availability, while in the CA-CDMA protocol it refers to a " transmission floor." The similarity is important since the Qualcomm system has proven to be successful.

2.4.5 Interference Margin

An interference margin is needed to allow terminals at some distance from a receiver to start new transmissions in the future. In this section, we describe how this margin is computed. Consider an arbitrary receiver i. Let μ^* be the $E_b/N_{0\text{eff}}$ ratio that is needed to achieve the target bit error rate at that receiver. It follows from (2.1) that to achieve the target error rate, we must have

$$\frac{P_0^{(i)}}{P_{\text{thermal}} + P_{\text{MAI}}^{(i)}} \geq \mu^* \tag{2.3}$$

where $P_0^{(i)}$ was defined before, P_{thermal} is the thermal noise power and $P_{\text{MAI}}^{(i)}$ is the total MAI at receiver i (in (2.1) $P_{\text{MAI}}^{(i)} = 2\sum_{j=1}^{K} P_j/3W$). So the minimum required received power is $(P_0^{(i)})_{min} = \mu^*(P_{\text{thermal}} + P_{\text{MAI}}^{(i)})$.

The interference margin strongly depends on the network load, which itself can be conveyed in terms of the so-called *noise rise* ($\xi^{(i)}$), defined as follows:[6]

$$\xi^{(i)} \stackrel{\text{def}}{=} \frac{(\frac{E_b}{N_0})_{\text{unloaded}}}{(\frac{E_b}{N_0})_{\text{loaded}}} = \frac{P_{\text{thermal}} + P_{\text{MAI}}^{(i)}}{P_{\text{thermal}}} \tag{2.4}$$

Note that $(P_0^{(i)})_{min} = \xi^{(i)}\mu^* P_{\text{thermal}}$ is also dependent on the noise rise. While more capacity can be achieved by increasing the noise rise (i.e., allowing larger $P_{\text{MAI}}^{(i)}$), the maximum allowable noise rise is constrained by two factors. First, Federal Communications Commission (FCC) regulations limit the power to some fixed value (e.g., 1 Watt for 802.11 devices). Given this maximum transmission power, as the noise rise is increased, the received power $(P_0^{(i)})_{min}$ must increase (μ^* and P_{thermal} are constants) and hence, the maximum range (or coverage) for reliable communication will decrease. Second, increasing the noise rise increases the power used to transmit

[6]This definition is similar but not exactly equal to the definition used in [87] for cellular systems.

the packet, which in turn increases energy consumption. Energy is a scarce resource in MANETs, so it is undesirable to trade off energy for throughput.

We set the interference margin used by a transmitter to the maximum *planned* noise rise ($\xi_{\max}$), which is obtained by taking into account the above two restrictions on $\xi^{(i)}$. The computations are performed as follows. First, we require that the maximum range, say $d_{\max}$, of our protocol be the same as the maximum range of the 802.11 scheme. For the maximum range, the power used in our protocol equals $\xi^{(i)}$ times the power used in the 802.11 standard. Thus, $\xi_{\max}$ cannot be greater than the ratio of the power limit set by the FCC and the power used in the 802.11 scheme. To account for the second constraint, we choose the interference margin in a manner that maintains the same *energy per bit* consumed in the 802.11 scheme. The value of the interference margin that achieves the above goals can be derived as follows. We assume that the transmission power attenuates with the distance d as k/d^n (k is a constant and $n \geq 2$ is the loss factor). The minimum required transmit power in CA-CDMA is:

$$P_{\text{CA-CDMA}} = \frac{\xi_{\max}\mu^* P_{\text{thermal}} d^n}{k} \tag{2.5}$$

Assuming that the distance d is uniformly distributed from zero to $d_{\max}$, we compute the expectation of $P_{\text{CA-CDMA}}$ with respect to d:

$$E[P_{\text{CA-CDMA}}] = \frac{\xi_{\max}\mu^* \ P_{\text{thermal}} d_{\max}^n}{k(n+1)} \tag{2.6}$$

As for the 802.11 protocol, its corresponding transmission power is:

$$P_{802.11} = \frac{\mu^* \ P_{\text{thermal}} d_{\max}^n}{k} \tag{2.7}$$

Note that $P_{802.11}$ does not depend on d since the 802.11 standard uses a fixed transmission power.

Accordingly, to achieve equal average energy per bit consumption, we must have:

$$\frac{E\left[P_{\text{CA-CDMA}}\right]}{R_{\text{CA-CDMA}}} = \frac{P_{802.11}}{R_{802.11}} \tag{2.8}$$

where $R_{\text{CA-CDMA}}$ and $R_{802.11}$ are the bit rates for the transmitted data packets in the CA-CDMA and 802.11 protocols, respectively. The reason why these rates can

be different is that in our protocol we use two distinct frequency bands, one for control packets and one for data packets, while the standard uses only one band for all packets. Hence, for a fair comparison, data packets in the CA-CDMA protocol must be transmitted at a slower rate.

From (2.6), (2.7), and (2.8), the interference margin is given by:

$$\xi_{\max} = (n+1)\frac{R_{\text{CA-CDMA}}}{R_{802.11}} \tag{2.9}$$

As an example, consider the two-ray propagation model with $n = 4$, and let the control channel occupy 20% of the total available bandwidth. Then $\xi_{\max} = 6$ dB. It is worth noting that 6 dB lies within the range of values used in already deployed cellular systems [87].

2.4.6 Channel Access Mechanism

We now describe the admission control and channel access strategy in the CA-CDMA protocol. The admission scheme allows only transmissions that cause neither primary nor secondary collisions to proceed concurrently. RTS and CTS packets are used to provide three functions. First, these packets allow nodes to estimate the channel gains between transmitter-receiver pairs. Second, a receiver i uses the CTS packet to notify its neighbors of the additional noise power (denoted by $P_{\text{noise}}^{(i)}$) that each of the neighbors can add to terminal i without impacting i's current reception. These neighbors constitute the set of *potentially interfering* terminals. Finally, each terminal keeps listening to the control channel regardless of the signal destination in order to keep track of the average number of active users in their neighborhoods. These functions are now explained in detail.

If terminal j has a packet to transmit, it sends a RTS packet over the control channel at $P_{\max}$, and includes in this packet the maximum *allowable* power level ($P_{\text{map}}^{(j)}$) that terminal j can use that will not disturb any ongoing reception in j's neighborhood. The computation of this power will be discussed shortly. The format of the RTS packet is similar to that of the IEEE 802.11, except for an additional

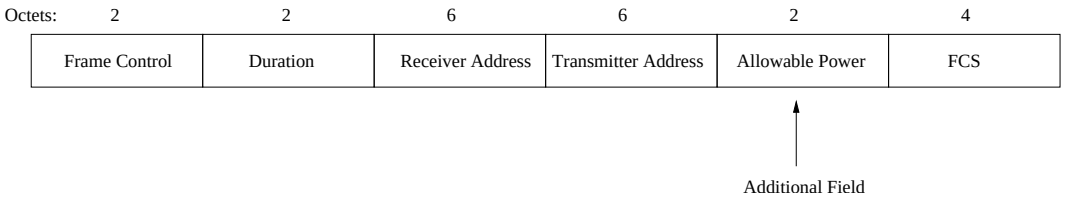


Figure 2.5: Format of the RTS packet in the CA-CDMA protocol.

two-byte field that contains the $P_{\text{map}}^{(j)}$ value as shown in Figure 2.5.

Upon receiving the RTS packet, the intended receiver, say terminal i, uses the predetermined $P_{\max}$ value and the power of the received signal $P_{\text{received}}^{(ji)}$ to estimate the channel gain $G_{ji} = P_{\text{received}}^{(ji)}/P_{\max}$ between terminals i and j at that time (note that we assume channel reciprocity, and so $G_{ij} = G_{ji}$) . Terminal i will be able to correctly decode the data packet if transmitted at a power $P_{\min}^{(ji)}$ given by:

$$P_{\min}^{(ji)} = \frac{\mu^*(P_{\text{thermal}} + P_{\text{MAI-current}}^{(i)})}{G_{ji}} \tag{2.10}$$

where $P_{\text{MAI-current}}^{(i)}$ is the effective *current* MAI from all ongoing (interfering) transmissions.

Note that because of the assumed stationarity in the channel gain over small time intervals, G_{ji} is approximately constant throughout the transmissions of the control packet and the ensuing data packet. Now, $P_{\min}^{(ji)}$ is the minimum power that terminal j must use for data transmission in order for terminal i to correctly decode the data packet *at the current level of interference*. This $P_{\min}^{(ji)}$, however, does not allow for any interference tolerance at terminal i, and thus, all neighbors of terminal i will have to defer their transmissions during terminal i's ongoing reception (i.e., no simultaneous transmissions can take place in the neighborhood of i).

Now, according to the link budget calculations in Section 2.4.5, the power that terminal j is allowed to use to send to i is given by:

$$P_{\text{allowed}}^{(ji)} = \frac{\xi_{\max}\mu^* P_{\text{thermal}}}{G_{ji}} \tag{2.11}$$

If $P_{\text{allowed}}^{(ji)} < P_{\min}^{(ji)}$, then the MAI in the vicinity of terminal i is greater than the one allowed by the link budget. In this case, i responds with a negative CTS, informing j that it cannot proceed with its transmission (the negative CTS is used to prevent multiple RTS retransmissions from j). The philosophy behind this design is to prevent transmissions from taking place over links that perceive high MAI. This consequently increases the number of active links in the network (subject to the available power constraints).

On the other hand, if $P_{\text{allowed}}^{(ji)} > P_{\min}^{(ji)}$, then it is possible for terminal i to receive j's signal but only if $P_{\text{allowed}}^{(ji)}$ is less than $P_{\text{map}}^{(j)}$ (included in the RTS). This last condition is necessary so that transmitter j does not disturb any of the ongoing transmissions in its vicinity. In this case, terminal i calculates the *interference power tolerance* $P_{\text{MAI-future}}^{(i)}$ that it can endure from *future* unintended transmitters. This power is given by:

$$P_{\text{MAI-future}}^{(i)} = \frac{3W\ G_{ji}}{2\ \mu^*}(P_{\text{allowed}}^{(ji)} - P_{\min}^{(ji)}) \tag{2.12}$$

Note that the factor $3W/2$ comes from the spreading gain (see (2.1)).

The next step is to equitably distribute this power tolerance among future potentially interfering users in the vicinity of i. The rational behind this distribution is to prevent one neighbor from consuming the entire $P_{\text{MAI-future}}^{(i)}$. In other words, we think of $P_{\text{MAI-future}}^{(i)}$ as a network resource that should be shared among various terminals. Let $K^{(i)}$ be the number of terminals in the vicinity of i that are to share $P_{\text{MAI-future}}^{(i)}$. This number is determined as follows. Terminal i keeps track of the number of simultaneous transmissions (i.e., load) in its neighborhood, which we donate by $K_{\text{inst}}^{(i)}$. This can be easily achieved by monitoring the RTS/CTS exchanges over the control channel. In addition, i keeps an average $K_{\text{avg}}^{(i)}$ of $K_{\text{inst}}^{(i)}$ over a specified window. Then, $K^{(i)}$ is calculated as:

$$K^{(i)} = \begin{cases} \beta(K_{\text{avg}}^{(i)} - K_{\text{inst}}^{(i)}), & \text{if } K_{\text{avg}}^{(i)} > K_{\text{inst}}^{(i)} \\ \beta, & \text{otherwise} \end{cases} \tag{2.13}$$

where $\beta > 1$ is a safety margin.

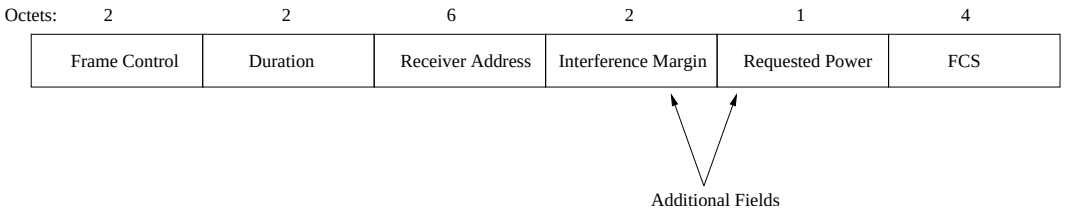


Figure 2.6: Format of the CTS packet in the proposed protocol.

Now, the MAI at terminal i can be split into two components: one that is attributed to terminals that are within the range of i (denoted by $P^{(i)}_{\text{MAI-within}}$), and one that is caused by terminals outside that range (denoted by $P^{(i)}_{\text{MAI-other}}$) [63]. While terminal i can have some control over $P^{(i)}_{\text{MAI-within}}$, it cannot influence $P^{(i)}_{\text{MAI-other}}$. We account for this fact in the value of $P^{(i)}_{\text{noise}}$ as follows. In line with cellular systems, we assume that $P^{(i)}_{\text{MAI-other}} = \alpha P^{(i)}_{\text{MAI-within}}$, where $\alpha < 1$ and depends mainly on the propagation path loss factor (practical values for α are ≈ 0.5 for the two-ray model [87]). Accordingly, the interference tolerance $P^{(i)}_{\text{noise}}$ that each *future* neighbor can add to terminal i is given by

$$P^{(i)}_{\text{noise}} = \frac{P^{(i)}_{\text{MAI-future}}}{(1+\alpha)K^{(i)}} \tag{2.14}$$

When responding to j's RTS, terminal i indicates in its CTS the power level $P^{(ji)}_{\text{allowed}}$ that j must use. In addition, terminal i inserts $P^{(i)}_{\text{noise}}$ in the CTS packet and sends this packet back to terminal j at $P_{\max}$ over the control channel using the common code. The format of the CTS packet is shown in Figure 2.6.

A potentially interfering terminal, say s, that hears the CTS message uses the signal strength of the received CTS to compute the channel gain G_{si} between itself and terminal i. The channel gain along with the broadcasted $P^{(i)}_{\text{noise}}$ values are used to compute the maximum power $P^{(s)}_{\text{map}}$ that s can use in its future transmissions. More specifically, $P^{(s)}_{\text{map}}$ is taken as the minimum of the $P^{(k)}_{\text{noise}}/G_{sk}$ values, for all neighbors k of s (i.e., $P^{(s)}_{\text{map}}$ is updated dynamically whenever s overhears a new CTS). Note that it is possible for more than $K^{(i)}$ terminals to start transmitting

during i's reception and this may result in MAI at i that is greater than $P^{(i)}_{\text{MAI-future}}$. We address this issue in Section 2.4.7.

The approach we discussed in this section provides a distributed mechanism for admission control. In contrast to cellular systems where the base station makes the admission decision, in here each terminal, and depending on previously heard RTS and CTS packets, decides whether its transmission can proceed or not.

Following a successful reception of a data packet, receiver i responds with an ACK packet, which is transmitted over the data channel using the same power level that would have been used if i were to send a data packet to j. We assume that enough FEC code is used to protect ACK packets from most types of collisions (given the small size of the ACK packets, the FEC overhead is not significant). A similar argument has been used in other, previously proposed protocols (e.g., [80]).

2.4.7 Protocol Recovery

In [31] the authors observed that when the transmission and propagation times of control packets are long, the likelihood of a collision between a CTS packet and a RTS packet of another contending terminal increases dramatically; the vulnerable period being twice the transmission duration of a control packet. At high loads, such a collision can lead to collisions with data packets. This is illustrated in Figure 2.7, where we assume that terminal D starts sending a RTS to terminal C while C is receiving B's CTS that is intended to A. A collision happens at C, and hence, C is unaware of B's subsequent data reception. Afterwards, if C decides to transmit a CTS to D, it may destroy B's reception.

Another problem that was mentioned earlier is if the interference goes above $P^{(.)}_{\text{MAI-future}}$. In CA-CDMA, we avoid the above two problems as follows. Suppose that while receiving a data packet, terminal i hears a RTS message (destined to any terminal) that contains an allowable power $P^{(.)}_{\text{map}}$ value that if used could cause an unacceptable interference with i's ongoing reception. Then terminal i shall respond

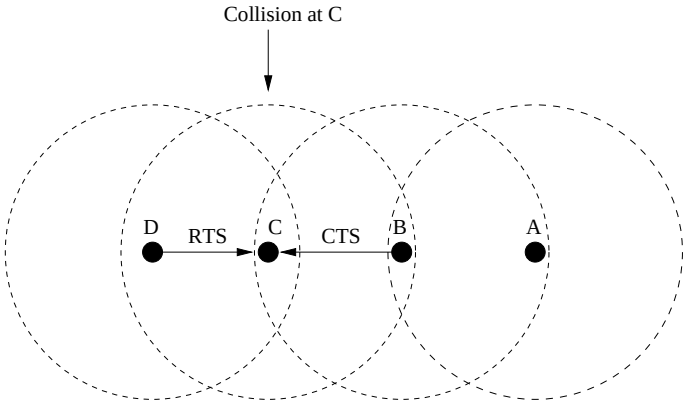


Figure 2.7: Example of a collision between control packets that eventually leads to a collision with a data packet.

immediately with a special CTS packet over the control channel, preventing the RTS sender from commencing its transmission. The duration field of the CTS packet contains the time left for terminal i to finish its ongoing reception.

To see how this solution helps in reducing the likelihood of collisions with data packets, consider the situation in Figure 2.7. Suppose that terminal A sends a RTS to terminal B, and B responds back with a CTS that collides at C with a RTS from D. Now, C does not know about B's ongoing reception. Two scenarios can happen. In the first, terminal C may later wish to send a packet to, say, terminal D. It sends a RTS, which will be heard by terminal B. B responds back with a special CTS. Note that there is a good chance that B's special CTS will collide with the CTS reply from D; however, this is desirable since C will fail to recover D's CTS packet, and will therefore defer its transmission and invoke its backoff procedure. In essence, B's special CTS acts as a jamming signal to prevent C from proceedings with its transmission. The second possible scenario is that D (or any other terminal that is out of the maximum range of B) may send a new RTS to C. C will respond to D with a CTS, and D will start sending data to C. Simultaneously, A may be sending to B, without any collision. This is possible because in CA-CDMA, data and RTS/CTS packets are sent over orthogonal channels.

Note that in CA-CDMA we try to avoid likely collision scenarios such as the one

mentioned in [31]. However, there are still few complicated (and definitely much less probable) scenarios where data packets may collide; recovery from such collisions is left to the upper layers.

2.4.8 Variable Processing Gain

In this section, we show how variable processing gain (VPG) [55] can be used to increase the capacity of the proposed protocol. As explained earlier, in CDMA systems, the low-rate information signal is spread by a high-rate PN code such that each user occupies the total available bandwidth. As long as there are enough concurrent transmissions to achieve the capacity at the required $E_b/N_{0\text{eff}}$, the bandwidth is fully utilized. However, if the number of transmissions is small, then we are actually using more spreading than needed. As an example, suppose that a terminal is transmitting while all of its neighbors are silent. Then, only a small fraction $1/W$ (recall that W is the processing gain) of the capacity is actually being used, thus, bandwidth is wasted and the spectral efficiency (i.e., data rate per channel bandwidth) is very low.

One way to improve this situation is to increase the information rate. This, in turn, decreases the processing gain, and thus, the immunity of the desired signal against MAI, which is small by assumption. However, varying the processing gain (or the information rate) allows the developed protocol to adapt to different loads. Therefore, VPG allows for near "peak capacity" access by a few users at low loads, thus increasing the spectrum efficiency and the overall network throughput. Another motivation for VPG is to accommodate users with different or even time-varying rate requirements.

We now present a method to enhance the proposed CA-CDMA protocol with VPG. Note here that our goal is to demonstrate the usefulness of VPG, rather than to investigate it completely. Therefore, for simplicity, we consider transmitting at only two rates: R and $2R$. More rates can be easily included by studying further extensions of the method. The proposed simple two-rate CA-CDMA protocol is as follows. Suppose that terminal i has just received an RTS from terminal j. Terminal

i computes $P_{\text{allowed}}^{(ji)}$ according to (2.11) at rate R, and checks for the following:

- $P_{\text{allowed}}^{(ji)} \leq P_{\text{map}}^{(j)}/2$ (recall that $P_{\text{map}}^{(j)}$ was included in j's RTS).
- $P_{\text{allowed}}^{(ji)} \leq \lambda P_{\max}$, where $\lambda \in [0, 0.5]$.

If both conditions are satisfied, then terminal i instructs transmitter j (via the CTS packet) to use rate $2R$ at twice the calculated required power $P_{\text{allowed}}^{(ji)}$. Note that $P_{\text{allowed}}^{(ji)}$ increases in this case, since by (2.11), $P_{\text{allowed}}^{(ji)}$ is a function of P_{thermal}, which itself is directly proportional to the information rate.

The first condition is necessary since doubling the rate requires doubling the required power to achieve the same $E_b/N_{0\text{eff}}$. Note that it is much more probable for this condition to hold at low load since $P_{\text{map}}^{(j)}$ would be higher. On the other hand, the second condition favors higher-rate links between nodes that are near to each other, i.e., nodes that see higher channel gains; λ is left to the discretion of the system designer. In Section 2.5 we demonstrate via simulation the throughput advantage of VPG.

2.4.9 Code Assignment

Because of the continuously changing network topology, it is difficult to guarantee correct code assignment at all time. Moreover, since not every node is active at all times, it may be desirable to oversubscribe the medium by assigning the same code to two neighboring terminals, thus violating the assignment goal. In this situation, it is the function of the MAC layer to reduce (or eliminate) contention on the medium. In CA-CDMA, this problem is addressed as follows. When terminal j sends a RTS, it inserts in that RTS the identity of the code that j intends to use for the ensuing data packet. A neighboring terminal that is receiving a packet on the *same* code can then respond back with the "special" CTS (explained in Section 2.4.7), which prevents j from commencing its data transmission. Note here the advantage of our architecture, which allows terminals to be informed about all neighborhood activities.

Another possible implementation is to combine the code assignment and access schemes [125]. In such an implementation, the RTS/CTS handshake over the common channel serves to reserve codes so that while the reception is ongoing, no other neighboring terminal can use any of the reserved codes. Although these two problems have been studied separately and dealt with at different layers in the protocol stack, there are two main motivations for combining them. The first is to reduce the overhead of exchanged information sharing. That is, information distributed to solve one problem (e.g., RTS and CTS) can be used to solve the other one (e.g., code assignment). Second, the MAC layer represents the most dynamic and mobility-transparent layer of the protocol stack. Thus, it is beneficial to do code assignment at the MAC layer. On the other hand, separating the two problems has its own advantages, including fairness. It is generally difficult to provide fairness in a contention-based MAC protocol. Thus, an upper layer code assignment can account for that.

2.5 Protocol Evaluation

2.5.1 Simulation Setup

We now evaluate the performance of the CA-CDMA protocol and contrast it with the IEEE 802.11 scheme. Our results are based on simulation experiments conducted using CSIM programs (CSIM is a C-based process-oriented discrete-event simulation package). In our simulations, we investigate both the network throughput as well as the energy consumption. For simplicity, data packets are assumed to have a fixed size. Each node generates packets according to a Poisson process with rate λ (same for all nodes). The routing overhead is ignored since the goal here is to evaluate the performance improvements due to the MAC protocol. Furthermore, because the interference margin is chosen so that the maximum transmission range under the CA-CDMA and 802.11 protocols is the same, it is safe to assume that both protocols achieve the same forward progress per hop. Consequently, we can focus on the one hop throughput, i.e., the packet destination is restricted to one

Data packet size	2 KB
802.11 data rate	2 Mbps
CA-CDMA data rate	1.6 Mbps
Control channel rate	400 Kbps
Processing gain	11
SNR threshold	10 dB
Reception threshold	-94 dBm
Carrier-sense threshold	-108 dBm
Thermal+receiver noise	-169 dBm/Hz
802.11 power	20 dBm
$\xi_{\max}$	6 dB

Table 2.1: Parameters used in the simulations.

hop from the source. The Random Waypoint model is used for mobility, with a host speed that is uniformly between 0 and 2 meters/sec. Note, however, that mobility has a little effect on our protocol, since an RTS-CTS exchange precedes every packet transmission. The transmission periods for the RTS, CTS, data, and ACK packets are all in tens of milliseconds, so no significant changes in topology take place within these periods. The capture model is similar to the one in [121]. Other parameters used in the simulations are given in Table 2.1. These paraments correspond to realistic hardware settings [4]. Note that the information rate is kept fixed for the first set of simulations. VPG advantages will be demonstrated later on in this section.

2.5.2 Simulation Results

We consider two types of topologies: *random grid* and *clustered*. In the random grid topology, M mobile hosts are placed across a square area of side length 3000 meters. The square is split into M smaller squares. The location of a mobile user is selected randomly within each of these squares. For each generated packet, the destination node is randomly selected from the one-hop neighbors.

The performance for random grid topologies is demonstrated in Figure 2.8. In parts (a) and (b), we set $M = 36$ and vary the packet generation rate (λ). Part (a)

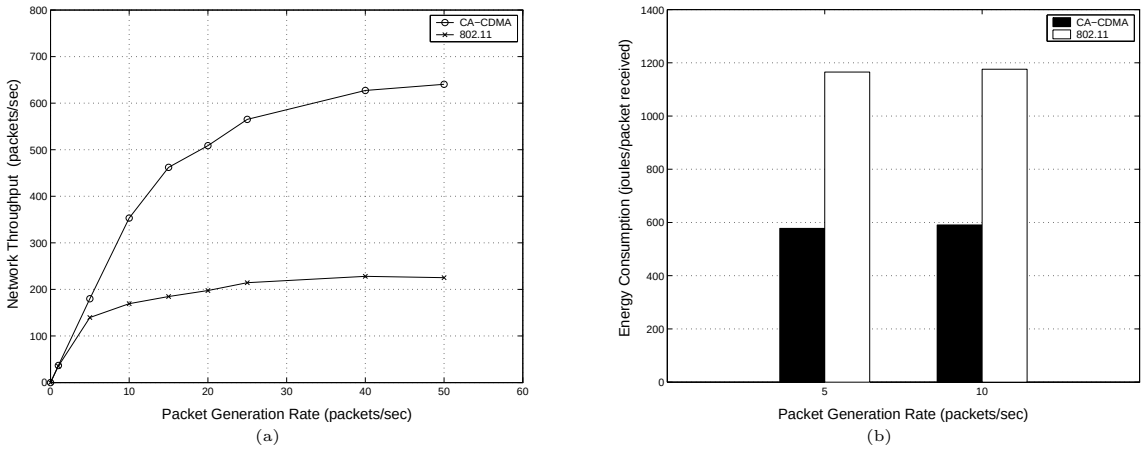


Figure 2.8: Performance of the CA-CDMA and the 802.11 protocols as a function of λ (random grid topologies).

of the figure depicts the network throughput. It is shown that CA-CDMA achieves up to a 280% increase over the throughput of the IEEE 802.11 scheme. This increase is attributed to the increase in the number of simultaneous transmissions. Furthermore, CA-CDMA saturates at about twice the load at which the 802.11 scheme saturates.

Part (b) of Figure 2.8 depicts the energy consumption versus λ. Energy consumption is the total energy used to *successfully transmit* a packet. It includes the energy of the control packets and the lost energy in retransmitting data and control packets in case of collisions. For all cases, CA-CDMA requires less than 50% of the energy required under the 802.11 scheme. This may, at first, seem to counterintuitive, since in Section 2.4.5 the interference margin was chosen so that both protocols consume the same energy per packet. However, according to the topology we examine here, the transmitter-receiver separation distance is not uniform. More links are formed with neighbors that are much closer than the maximum transmission range (1061 meters in our simulations). Unlike the 802.11 scheme, CA-CDMA makes use of shorter links to save energy. Note that in both protocols, the required energy increases with the load. The reason for this is that as λ increases, the probability of

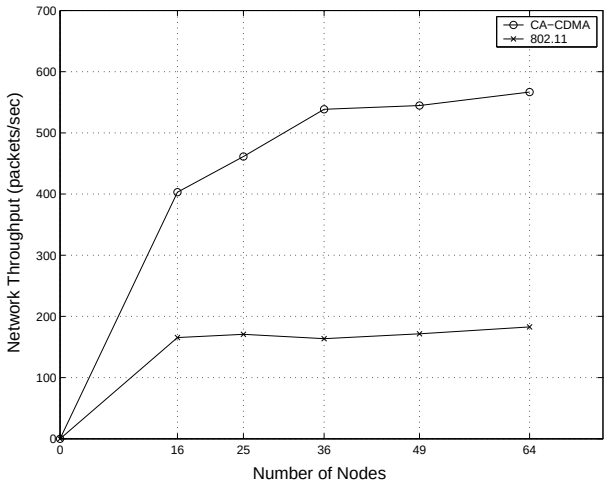


Figure 2.9: Performance of the CA-CDMA and the 802.11 protocols as a function of the number of nodes (random grid topologies).

collisions also increases, and hence, more energy has to be spent on retransmissions.

In Figure 2.9 we investigate the effect of varying the number of nodes while the dimensions of the region are kept fixed (3000m×3000m). A persistent load is used in this experiment, i.e., nodes always have packets to send. As shown in the figure, the throughput enhancement due to CA-CDMA increases with node density. This can be explained by noting that CA-CDMA bounds the transmission power rather that prevents simultaneous transmissions. Therefore, as the density of nodes increases, more concurrent links are formed and the network throughput increases. The 802.11 scheme reserves a fixed floor, and thus, all nodes within that floor have to defer their transmissions. Therefore, the density of the nodes has little effect on the 802.11 throughput.

The authors in [75] argued that traffic locality is the key factor in determining the feasibility of large ad hoc networks. This motivates studying the performance of CA-CDMA under *clustered* topologies. In such topologies, a node communicates mostly with nodes within its own cluster, and rarely with neighboring cluster nodes. These topologies are common in practice (e.g., a historical site where users of wireless

devices move in groups). To generate a clustered topology, we consider an area of dimensions 1000×1000 (in meters). We let $M = 24$ nodes, which are split into 4 equal groups, each occupying a 100×100 square in one of the corners of the complete area. For a given source node, the destination is selected from the same cluster with probability $1 - p$ or from a different cluster with probability p. In each case, the selection from within the given cluster(s) is done randomly.

Part (a) of Figure 2.10 depicts the network throughput versus λ for $p = 0.25$. According to the 802.11 scheme, only one transmission can proceed at a time since all nodes are within the carrier-sense range of each other. However, according to CA-CDMA, three to four transmissions can proceed simultaneously, resulting in a significant improvement in network throughput. In Part (b) of the figure, we further investigate the locality of the traffic by fixing λ and varying p. Indeed, as the figure shows, the locality of the traffic can highly impact the network throughput of CA-CDMA, while the 802.11 performance is almost unchanged. As the traffic locality increases (i.e., p decreases) the enhancement of CA-CDMA increases.

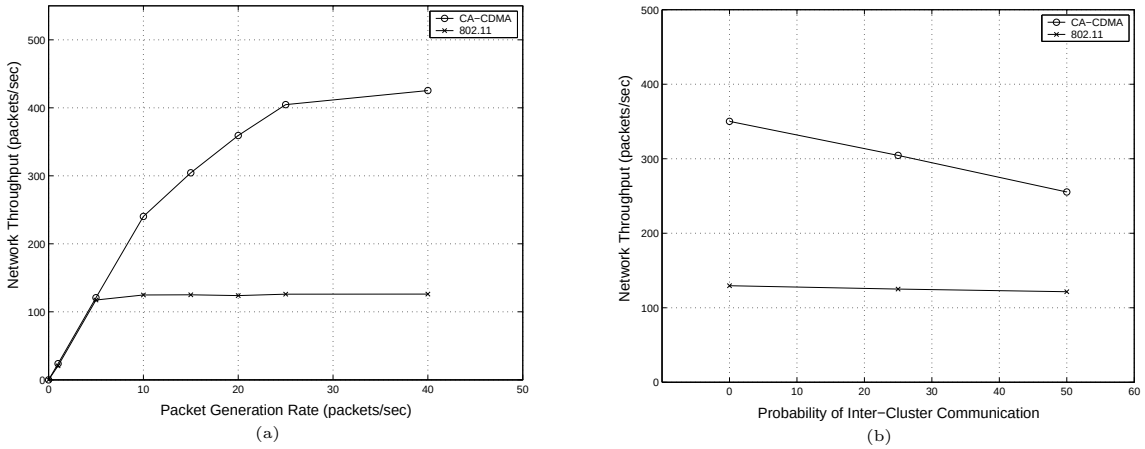


Figure 2.10: Performance of the CA-CDMA and 802.11 protocols (clustered topologies).

Next, we study the throughput advantage of VPG. We simulate the method proposed in Section 2.4.8, where λ is set to 0.1. We also simulate a variable-rate 802.11-based scheme, where terminals are allowed to transmit at twice the fixed

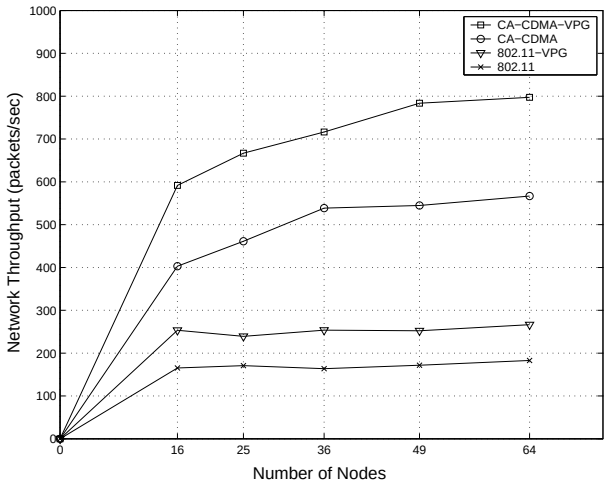


Figure 2.11: Performance of two-rate CA-CDMA and two-rate 802.11 protocols as compared to fixed-rate approaches (random topologies).

rate (i.e., at 4Mbps) if the measured SNR at the receiver is greater than twice the SNR threshold. Figure 2.11 demonstrates the advantage of VPG. Here, we vary the number of nodes while the dimensions of the region are kept fixed (3000m×3000m). We assume that, had the receiver indicated a rate $2R$ via its CTS packet, then the transmitter sends two packets back-to-back to that receiver. As shown in that figure, the (two-rate) CA-CDMA-VPG protocol achieves on average about 41% increase over the throughput of the fixed-rate CA-CDMA. The (two-rate) 802.11-VPG scheme, while still way below the fixed-rate CA-CDMA, achieves on average about 49% increase over the throughput of the fixed-rate 802.11. The improvement in both cases is because the VPG approach uses the "good" link conditions to send multiple packets to the destination, thus, improving the overall network throughput.

CHAPTER 3

THE POWMAC PROTOCOL

3.1 Introduction

So far, the Ad Hoc mode of the IEEE 802.11 standard [2] has been used as the *de facto* MAC protocol for MANETs. This protocol uses a 4-way handshake to resolve channel contention; when a terminal, say A, wants to send data to another terminal, say B, it first sends an RTS packet to B, which replies back using a CTS packet. The data transmission[1] $A \rightarrow B$ can now proceed, and once completed, terminal B sends back an acknowledgement (Ack) packet to A. The RTS and CTS packets include the duration of the ensuing data packet and are needed to reserve a *transmission floor* for the subsequent data packet. Any other terminal that hears the RTS or the CTS message defers its transmission until the ongoing transmission is over. The CTS message prevents collisions with the data packet at the destination terminal B, while the RTS message prevents collisions with the Ack packet at the source terminal A. Terminals transmit their control and data packets at a *fixed (maximum) power level*.

Despite its appealing simplicity, the 802.11 MAC approach can be overly conservative, leading to an unnecessary reduction in network throughput. To illustrate, consider the situation in Figure 3.1, where terminal A uses its maximum transmission power (TP) to send packets to terminal B (we assume omnidirectional antennas, so a terminal's reserved floor is represented by a circle in 2D space). According to the IEEE 802.11 scheme:

[1] Throughout this chapter, the notation $j \rightarrow i$ indicates a data transmission from j to i and an Ack transmission from i to j. We also refer to the data transmitter (the Ack recipient) as the source, and to the data receiver (the Ack transmitter) as the sink. Finally, we use the term "activity" to mean either a transmission or a reception.

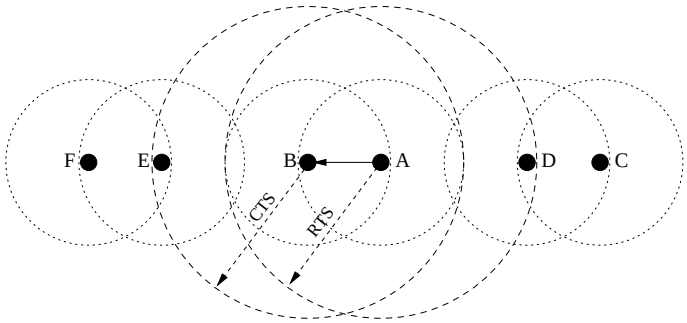


Figure 3.1: Inefficiency of the classic RTS/CTS approach. Terminals A and B are allowed to communicate, but terminals D and E are not. Dashed circles indicate the maximum transmission ranges, while dotted ones indicate the ranges of the minimum transmission powers needed for coherent reception at the respective receivers.

1. When terminal D hears A's RTS, it refrains from transmitting to terminal C to avoid corrupting A's reception of B's Ack packet. The inability of terminal D to transmit while A is transmitting its data packet is the well-known *exposed terminal problem*.

2. Terminal D also refrains from receiving from terminal C to avoid having its reception corrupted by A's data transmission.

3. Terminal E hears B's CTS and, therefore, refrains from transmitting to terminal F to avoid corrupting B's reception of A's data packet.

4. Terminal E also refrains from receiving from terminal F to avoid having its reception corrupted by B's Ack transmission.

However, it is not hard to show that the three transmissions $A \rightarrow B$, $C \rightarrow D$, and $E \rightarrow F$ can, in principle, proceed simultaneously if terminals are able to select their TPs appropriately. Enabling multiple transmissions to take place within the same neighborhood leads to an increase in network throughput and possibly a reduction in the overall energy consumption. The scheme proposed in this chapter is intended to allow for such transmissions to take place.

The roots of the problem with the IEEE 802.11 scheme lie in its overstated definition of a collision; if a terminal i is currently receiving a packet from a terminal j, then all other terminals in i's and j's transmission ranges[2] must defer their transmissions to avoid colliding with i's ongoing reception of the data packet and j's subsequent reception of the Ack packet. Furthermore, the use of a *fixed* common TP in the 802.11 scheme leads to reduced channel utilization and increased energy consumption. To explain the inefficiency of this design principle, consider a network with a fixed TP. Let P_j be the TP used by terminal j. Let G_{ji} be the channel gain from terminal j to terminal i. Then the signal-to-interference-plus-noise ratio (SINR) at terminal i for the desired signal from terminal j (SINR(j,i)) is given by:

$$\text{SINR}(j,i) = \frac{P_j G_{ji}}{\sum_{k \neq j} P_k G_{ki} + P_{\text{thermal}}} \tag{3.1}$$

where P_{thermal} is the thermal noise power. When the TP is fixed and common among all terminals, the above equation is a function of only the channel gains. Terminal j's packet can be correctly received at i if SINR(j,i) is above a certain threshold (denoted by μ^*) that reflects the QoS of the link[3]. So even if there is an interfering transmitter, say v, that is within the transmission range of i, it may still be possible for i to correctly receive j's packet. The simplified analysis in [128], where interference is attributed to only one terminal v, shows that under a path loss exponent of 4, i can correctly receive the desired packet as long as v is at distance $1.78d$ or more from i[4], where d is the distance between i and j (assuming a common TP). Therefore, in many cases v can be allowed to transmit and cause interference at i without necessarily colliding with i's reception of j's packet. Moreover, if terminals are able to control their TPs , then as shown in Figure 3.1, a greater number of concurrent transmissions can proceed simultaneously, leading to a increase in network throughput.

[2]The transmission range of terminal i is the largest region around i over which i's transmission can be successfully received with probability near 1 in the absence of interference from other terminals.

[3]μ^* includes the effect of any employed forward error correction scheme.

[4]The analysis in that chapter assumes no spread spectrum, i.e., no spreading gain.

3.1.1 Goals and Contributions

The previous discussion motivates the need for an interference-aware transmission power control (TPC) protocol to improve network throughput by means of increasing the channel spatial reuse. Theoretical studies [45] and simulation results [80, 81] have demonstrated that TPC can provide significant gains in capacity and energy consumption, not to mention its benefits in providing admission control and in QoS provisioning [14].

Many TPC schemes for MANETs have been proposed in the literature. However, as explained in Section 3.2, these schemes suffer from one or more of the following deficiencies: (1) the TPC approach may yield energy reduction but not throughput gain, (2) the MAC design may not support collision avoidance, resulting in the well-known hidden terminal problem, (3) the TPC approach requires extra hardware (e.g., multiple transceivers), (4) lack of link-layer reliability, i.e., Ack packets are not protected, and (5) many of the assumptions made in the MAC design are unrealistic. Accordingly, we introduce a new TPC scheme for MANETs that ameliorates these deficiencies. Our scheme is based on a single-channel, single-transceiver approach, and is shown to provide a significantly higher network throughput than the IEEE 802.11 scheme while yet preserving the collision avoidance properties of the IEEE 802.11 scheme. To the best of our knowledge, this is the first TPC solution that is based on a single-channel, single-transceiver design, that can increase the throughput of a MANET relative to the IEEE 802.11 scheme, and that supports link-layer reliability.

The rest of the chapter is organized as follows. In Section 3.2, we present related TPC schemes for MANETs and show their limitations. The proposed POWMAC protocol is presented in Section 3.3, followed by simulation results and discussion in Section 3.4.

3.2 Related Work

TPC schemes for MANETs can be generally classified into two classes. In the first class (e.g., [34, 99, 103, 122]), TPC is used to control the network topology, indirectly impacting the set of next-hop neighbors of a terminal and the subsequent routing decisions taken by that terminal. The *same* TP is used by a terminal to transmit its packets to any of its neighbors. This TP is updated following a mobility-related topological change. For pedestrian speeds, such a change occurs at a time scale of hundreds of milliseconds to seconds (in contrast, packet transmission times are, at most, in the order of few milliseconds). The main design issue here is how to determine the minimum TP for a given terminal such that some topological properties (e.g., connectivity, node degree, etc.) are guaranteed. One limitation of this class of protocols is its reliance solely on CSMA for accessing/reserving the shared wireless channel. It is known that using CSMA alone for accessing the channel can significantly degrade network performance (throughput, delay, and power consumption) because of the hidden terminal problem [111]. Unfortunately, this issue cannot be addressed by simply using a standard RTS/CTS-like channel reservation approach as explained in the example in Figure 3.2. Here, terminal A has just started a transmission to terminal B at a power level that is just enough to ensure coherent reception at B. Suppose that B uses the same power level to communicate with A. Terminals C and D are outside the floors of A and B, so they do not hear the RTS/CTS exchange between A and B. For terminals C and D to be able to communicate, they have to use a power level that is reflected by the transmission floors in Figure 3.2 (the two circles centered at C and D). However, the transmission $C \rightarrow D$ will interfere with $A \rightarrow B$ transmission, causing a collision at B. In essence, the problem is caused by the asymmetry in the transmission floors (i.e., B can hear C's transmission to D but C cannot hear B's transmission to A).

In the second class of TPC schemes, power control is applied on a per-packet basis, with the TP being dependent on both the transmitting and receiving terminals. The TP in this case is not directly tied to the routing layer or the topological

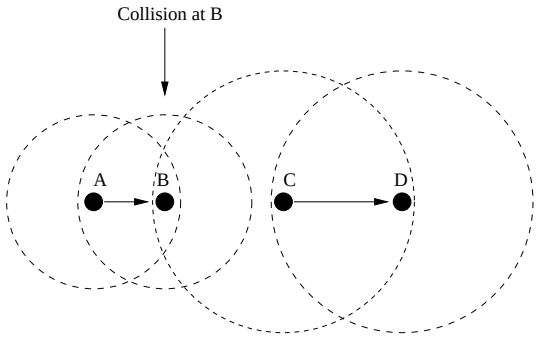


Figure 3.2: Challenge in implementing power control in a distributed fashion. Terminal C is unaware of the ongoing transmission $A \rightarrow B$, and hence it starts transmitting to terminal D at a power that destroys B' s reception.

properties of the network (although some schemes in this class indirectly influence the decisions taken by the routing layer). For a given next hop that is provided by the routing layer, the main question here is what TP to use for sending a given data packet to that next hop. This class of TPC schemes can be further divided into two subclasses: energy- and throughput-oriented schemes. The former subclass (e.g., [44, 63, 67, 94]) aims primarily at reducing energy consumption, with network throughput being a secondary factor. Terminals exchange their RTS and CTS packets at a maximum power ($P_{\max}$), but send their data and Ack packets at the minimum power needed for reliable communication ($P_{\min}$). The value of $P_{\min}$ is determined based on the required QoS (i.e., the signal-to-interference-plus-noise ratio (SINR)), the interference level at the receiver, and the channel gain between the transmitter and the receiver. In [63] the authors enhanced the performance of this approach by periodically increasing the TP of the data packet to $P_{\max}$ for enough time to protect the reception of the Ack at the source terminal. While this class of TPC protocols achieves good reduction in energy consumption (relative to the 802.11 MAC protocol), at best it gives comparable throughput to that of the 802.11 scheme. The main reason is that, as in the 802.11 approach, RTS and CTS messages are used to *silence* neighboring terminals, preventing concurrent transmissions from

taking place over the maximum transmission range[5].

Throughput-oriented TPC schemes (e.g., [80, 126, 81]) use per-packet power control to increase the spatial channel reuse. These schemes allow for concurrent transmissions in the same vicinity of a receiver by locally broadcasting *collision avoidance information* (CAI) over a separate control channel. In the PCMA protocol [80], the receiver advertises its *interference margin* by sending busytone pulses over a separate control channel. The use of a separate control channel in conjunction with a busytone scheme was proposed in [126], where the sender transmits data packets and busytones at reduced power, while the receiver transmits its busytones at the maximum power. A terminal estimates the channel gain from the busytones and is allowed to transmit if its transmission is not expected to add more than a fixed interference to the ongoing receptions. The authors, however, make strong assumptions about the interference power. Specifically, they assume that the antenna is able to reject any interfering power that is less than the power of the "desired" signal (i.e., they assume perfect capture). Furthermore, the schemes in [80, 126] do not address the issue of contention among busytones or their energy overhead. The PCDC protocol [81] uses two frequency-separated channels for data and control. RTS and CTS packets are transmitted over the control channel, providing CAI that facilitates interference-limited concurrent transmissions in the same vicinity.

Although the simulations in [80, 81, 126] indicate impressive improvement in throughput over the 802.11 scheme, we see five major design problems with these schemes that make their practicality questionable:

- In [80, 81, 126], the channel gain is assumed to be the same for both the control (or busytone) and data channels, and that terminals are able to transmit on one channel and, simultaneously, receive on the other. It is very difficult to achieve these two assumptions simultaneously. For the first assumption to hold, the control channel must be within the coherence bandwidth of the data

[5]The maximum transmission range of terminal i is the largest region around i over which i's maximum power transmission can be successfully received in the absence of interference from other terminals.

channel, which places an upper bound on the allowable frequency separation between the two channels. However, for the second assumption to hold, there must be some minimal frequency spacing between the two channels. Typically, a guard-band of about 5% of the nominal RF frequency is needed to keep the price and complexity of the transceiver at a reasonable level [100]. However, spacing the control and data channels by that much can make the first assumption invalid! Another issue is that the spectrums allocated to the control and data channels should not be equal; otherwise, considerable bandwidth would be wasted on signaling/control. However, fading is spectrum dependent, and thus the two channels experience different amounts of fading.

- To be able to receive/transmit and simultaneously receive/transmit over two channels, the mobile terminal must be equipped with two transceivers. The complexity and cost of the additional hardware may not justify the increase in throughput. Furthermore, it is unfair to compare the performance of these protocols to the single-channel, single-transceiver IEEE 802.11 scheme.

- Currently, most wireless devices implement the IEEE 802.11b standard. The class of two-channel protocols is not backward-compatible with the IEEE 802.11 standard, which makes it difficult to deploy such protocols in real networks.

- The above schemes do not provide reliability, i.e., they do not protect the reception of the Ack packet. Providing link-layer error control is important not only because it provides faster recovery than transport-layer error control, but also because the performance of traditional transport layer protocols (such as TCP) degrades significantly over wireless links, resulting in a large number of unnecessary retransmissions [21]. The 802.11 scheme is a reliable link-layer protocol; any terminal that hears the RTS or the CTS message must defer its transmission until the ongoing transmission (data and Ack) is completed. Thus, power controlled MAC protocols must be designed with the goal of protecting both the data and Ack receptions. Furthermore, it is unfair to

compare the performance of a reliable protocol (i.e., 802.11) with unreliable ones (i.e., [80, 81, 126]).

- Finally, the optimal allocation of the total spectrum between the data and control channels is load dependent. So for the allocation to be optimal under a varying traffic load, it has to be adjusted adaptively, which is not feasible in practice.

The protocol proposed in this chapter addresses all the above issues and provides a comprehensive, throughput-oriented MAC solution for MANETs using a single-transceiver, single-channel approach.

Before closing, we mention few other schemes in the literature that tackle the problem of power control from a completely different perspective. The COMPOW protocol [85] relies on routing-layer agents to converge to a *common* power level for all network terminals. However, for constantly moving terminals, the scheme (like any other routing-protocol-based scheme) incurs significant overhead, and convergence to a common power may not be possible. Moreover, in situations where network density varies widely (i.e., terminals are clustered), restricting all terminals to converge to a common power is a conservative approach. A clustering approach was proposed in [72], which simplifies the forwarding function for most terminals but at the expense of reducing network utilization (since all communications have to go through an elected terminal). This can also lead to the creation of bottlenecks.

A joint clustering/TPC protocol was proposed in [68], where clustering is implicit and is based on TP levels rather than on addresses or geographical locations. No CHs or gateways are needed. Each terminal runs several routing-layer agents that correspond to different power levels. These agents build their own routing tables by communicating with their peer routing agents at other terminals. Each terminal along the packet route determines the lowest-power routing table in which the destination is reachable. The routing overhead in this protocol grows in proportion to the number of routing agents, and can be significant even for simple mobility patterns (note that for the DSR routing protocol, for example, routing packets account for

approximately 38% of the total received bytes [61]). The protocol in [8] is energy-oriented and is basically a mechanism to learn the minimum TP level required for a terminal to successfully transmit to a neighboring terminal. This approach, however, suffers from the hidden terminal problem (see [81] for more details). Another novel approach for TPC is based on joint scheduling and power control [33]. This approach requires a central controller to execute the scheduling algorithm, i.e., it is not a truly distributed solution. Furthermore, it makes a number of strong assumptions, including synchronization, that each terminal knows the geographical location of all other terminals, and that the SINR measurement at each receiver is known to *all* transmitters. The Medium Access via Collision Avoidance with Enhanced Parallelism (MACA-P) proposed in [7] allows for parallel transmissions in situations only when two neighboring nodes are either both receivers or both transmitters, but a receiver and a transmitter are not neighbors. In addition, TPC was not considered in that work.

3.3 Proposed Protocol

3.3.1 Assumptions

In designing POWMAC, we assume that the channel gain is stationary for the duration of a few control and one data packet transmission periods. As discussed in Section 3.3.11, this assumption holds for typical mobility patterns and transmission rates. We also assume that the gain between two terminals is the same in both directions. This is the underlying assumption in any RTS/CTS-based protocol, including the IEEE 802.11 scheme. Finally, we assume that the radio interface can provide the MAC layer with the average power of a received control signal as well as the average interference power. Off-the-shelf wireless cards (e.g., [4]) readily provide such measured values using SINR estimators like the ones discussed in [90]. In POWMAC, each terminal is equipped with one transceiver that has standard carrier-sense hardware (i.e., a basic IEEE 802.11-compliant transceiver).

3.3.2 Overview of POWMAC

POWMAC is distributed, asynchronous, and adaptive to channel changes. Its key features are as follows. First, unlike the IEEE 802.11 approach (and the schemes in [8, 44, 63, 67, 94]), POWMAC does not use the control packets (i.e., RTS/CTS) to silence neighboring terminals. Instead, CAI is inserted in the control packets and is used in conjunction with the received signal strength of these packets to dynamically *bound* the TP of potentially interfering terminals in the vicinity of a receiving terminal. The details of this mechanism are presented in Section 3.3.4. The second main feature of POWMAC is that the required TP of a data packet is computed at the packet's intended receiver, say terminal i, according to a predetermined *maximum load factor*. The rationale behind this approach is to allow for some interference tolerance at receiver i, so that multiple interference-limited transmissions can simultaneously take place in the neighborhood of i. The tradeoffs involved in determining this load factor are discussed in Section 3.3.3.

The third feature of POWMAC is that some control packets (CTS packets and newly defined Decide-to-Send (DTS) packets) are transmitted at an adjustable power level so that they reach all and only potentially interfering terminals. This improves the spatial reuse for the control packets themselves and reduces their collisions. Section 3.3.8 presents the details of this aspect of power control.

Finally, in POWMAC, after terminals exchange their control packets, they refrain from transmitting their data packets for a certain duration, referred to as the *access window* (AW). The AW allows several pairs of neighboring terminals to exchange their control packets such that (interfering) data transmissions can proceed simultaneously as long as collisions are prevented. The AW consists of an adjustable number of fixed-duration *access slots*. As explained later, this number is adaptively varied, depending on network load. The AW is needed for two reasons. First, it reduces the likelihood of collisions between control and data packets. Even when power controlled, control packets will, in general, be transmitted at a higher power than data packets, so that they can reach many potential interferers. So allowing

these control packets to overlap in time with data packets (to enable concurrent RTS/CTS-based transmissions in the same neighborhood) would increase the likelihood of collisions. We remedy this situation by using an AW, whereby a receiving terminal i allows its neighbors to exchange their RTS/CTS packets before i's data reception starts, and when possible, to have these neighbors' own data transmissions proceed simultaneously with i's reception. Note that data packets are transmitted at a reduced power level to reach only the intended receiver, and so multiple data packets can be transmitted concurrently and still be received correctly.

The second purpose of the AW is to inform terminals that are currently transmitting or receiving of the ensuing data transmission. Because POWMAC uses a single-channel architecture, terminals can either transmit or receive at a given time, but not both. As a result, a terminal, say i, is basically "deaf" while transmitting, so it cannot hear any transmitted control packets in its vicinity. Consequently, when i becomes idle, its information about the ongoing receptions in its vicinity can be outdated, which can lead to collisions (if i decides to transmit again). The protocols in [80, 81, 126] alleviate this problem by using a two-channel, two-transceiver architecture; terminals are able to transmit/receive their data packets and still hear the control signals. However, as we discussed in Section 3.2, these approaches are not desirable for several reasons.

We note here that allowing several RTS/CTS exchanges to take place prior to data-packet transmissions was also used in the MACA-P protocol [7]. However, in that work the objective was not to address TPC, but rather to prevent collisions between control and data packets.

We conclude this section with an example that illustrates the basic operation of POWMAC (see Figure 3.3). The network topology is the one shown in Figure 3.1. Terminal A transmits an RTS to B at a maximum (known) power ($P_{\max}$). Terminal B replies back with a CTS packet that is sent at an adjustable power level to reach all and only potentially interfering terminals. The RTS/CTS exchange allows terminals A and B to agree on the TP of the ensuing data packet. It also provides a way to inform potentially interfering terminals (e.g., terminal E) of the power that

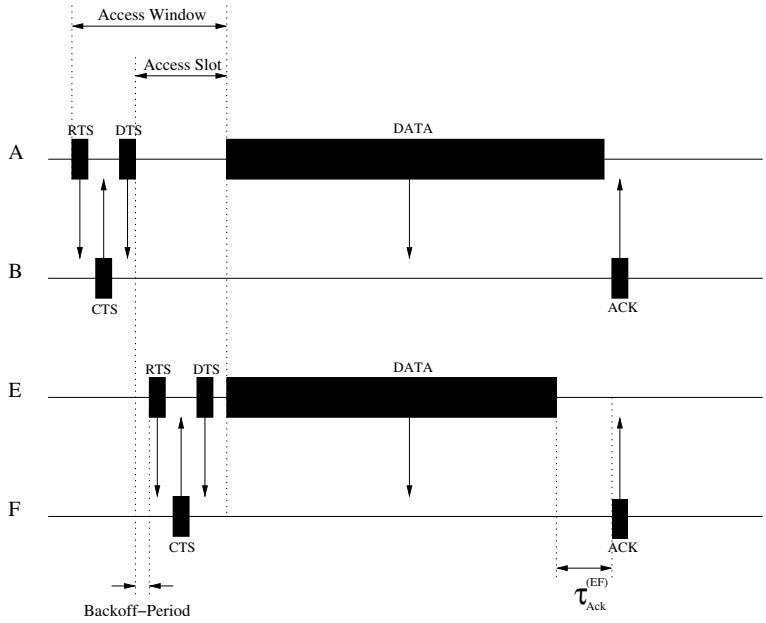


Figure 3.3: Basic operation of POWMAC.

they can use without disturbing the scheduled reception of the data packet at B. Terminal A confirms that the transmission $A \rightarrow B$ can proceed using the newly defined DTS control packet. Besides other reasons mentioned in Section 3.3.4, the DTS packet is used to inform A's neighbors of the power level that A intends to use for its data transmission. As explained later, this information is needed so that A's neighbors (i.e., terminal D) can determine whether or not they can receive a data packet from some other terminal (e.g., C) simultaneously while A is transmitting to B. In addition, the DTS provides a way to inform potentially interfering terminals (e.g., terminal D) of the power that they can use without disturbing the reception of the Ack packet at A. After the RTS/CTS/DTS exchange, terminal A refrains from sending its data packet for the remaining of the AW duration. During this duration, E and F can exchange control packets and decide if they can start the transmission $E \rightarrow F$ depending on whether or not this transmission will disturb the scheduled transmission $A \rightarrow B$.

3.3.3 Load Control

Load control is a concept that allows a prospective receiver to determine the appropriate TP for its upcoming data reception and the impact of this TP on ongoing as well as scheduled receptions of both data and Ack packets. If the power used to transmit a data packet to a terminal, say i, is just enough to overcome the current interference at i, then none of i's neighbors should be allowed to start new transmissions during i's reception. This silencing of neighboring terminals negatively impacts the aggregate throughput. On the other hand, if the TP is too high, it may induce high interference on other terminals in the vicinity of the transmitter, preventing them from receiving.

The load factor at terminal i, denoted by $\xi^{(i)}$, is a measure of the activity in terminal i's neighborhood. Formally, it is defined as[6]:

$$\xi^{(i)} \stackrel{\text{def}}{=} \frac{P_{\text{thermal}} + P_{\text{MAI}}^{(i)}}{P_{\text{thermal}}}, \tag{3.2}$$

where $P_{\text{MAI}}^{(i)}$ is the current multi-access interference (MAI) at receiver i[7]. Now, consider the transmission of a packet from j to i. Let d_{ij} be the distance between i and j, and let μ^* be the SINR threshold required to achieve a target bit error rate (BER) at receiver i. We assume that the TP attenuates with d_{ij} as k/d_{ij}^n, where k is a constant and $n \geq 2$ is the loss factor. Then, the minimum TP that is needed to achieve the target BER is

$$P_{\text{min}}^{(ji)} = \frac{\mu^*(P_{\text{thermal}} + P_{\text{MAI}}^{(i)})}{G_{ji}} = \frac{\mu^*\xi^{(i)}P_{\text{thermal}}}{G_{ji}} = \frac{\mu^*\xi^{(i)}P_{\text{thermal}}d_{ij}^n}{k}, \tag{3.3}$$

[6]This definition is somewhat similar but not quite identical to the definition used in [87] for cellular systems.

[7]Traditionally, MAI has been used to refer to the interference between signals that are spread using different CDMA codes. Since terminals in the IEEE 802.11 scheme use the same spreading code, in this chapter the term MAI will be used to refer to interference from unintended signals that are spread using the same code.

where $G_{ji} = k/d_{ij}^n$ is the channel gain from terminal j to terminal i ($G_{ji} \ll 1$). While more capacity can be achieved by increasing $\xi^{(i)}$ (i.e., allowing larger $P_{\text{MAI}}^{(i)}$), this also increases the power needed to transmit the packet, which in turn increases energy consumption. Energy is a scarce resource in MANETs, so it is undesirable to trade it off for throughput. Moreover, the Federal Communications Commission (FCC) regulations put a limit on the maximum power that can be used by terminals in the 2.4 GHz spectrum (e.g., 1 Watt for 802.11 devices). Given this limit, as the load is increased, the channel gain must be increased (with μ^* and P_{thermal} held constant), and so the maximum range (or coverage) for reliable communication will decrease.

Collectively, the above factors necessitate *load planning*, i.e., imposing a maximum load factor (MLF), denoted by $\xi_{\max}$, that terminals are not allowed to exceed. This $\xi_{\max}$ is set at the design phase to reflect several goals, including throughput, network lifetime, etc. One possible choice is as follows. First, to increase the spatial channel reuse, terminal j uses a TP that results in the MLF at terminal i. This TP is given by (see Equation 3.3):

$$P_{\text{POWMAC}}^{(ji)} = \frac{\mu^* \xi_{\max} P_{\text{thermal}}}{G_{ji}} = \frac{\mu^* \xi_{\max} P_{\text{thermal}} d_{ij}^n}{k}. \tag{3.4}$$

Second, we require that the (interference-free) maximum transmission range for both POWMAC and the 802.11 scheme, denoted by $d_{\max}$, to be the same. Then, assuming that d_{ij} is uniformly distributed between zero and $d_{\max}$ (other distance distributions, which could depend on the routing protocol, may also be used), we have

$$E[P_{\text{POWMAC}}^{(ji)}] = \frac{\mu^* \ \xi_{\max} \ P_{\text{thermal}} \ d_{\max}^n}{k \ (n+1)}. \tag{3.5}$$

As for the 802.11 protocol, its corresponding TP is:

$$P_{802.11} = \frac{\mu^* \ P_{\text{thermal}} \ d_{\max}^n}{k}. \tag{3.6}$$

Note that $P_{802.11}$ does not depend on d_{ij} since the 802.11 scheme uses a fixed TP. To account for the energy-consumption factor, we require that $\xi_{\max}$ be chosen such that the two protocols consume the same average energy per bit. Equating (3.5)

and (3.6), we end up with $\xi_{\max} = n + 1$. As an example, consider the two-ray propagation model with $n = 4$. Then $\xi_{\max} = 7$ dB, which lies within the range of values used in already deployed cellular systems [87]. Finally, we require that the *maximum* TP used in POWMAC be constrained by the FCC limit (from (3.4) and (3.6), this maximum power is given by $\xi_{\max} P_{802.11}$).

3.3.4 Channel Access Mechanism

Given a predetermined MLF, the purpose of the channel access mechanism is to allow the source and the sink to agree on the required TP such that the MLF is not exceeded at the source (Ack recipient) and at the sink (data recipient) during the reception periods. The access protocol should also ensure that the ensuing data transmission does not disturb any of the scheduled data/Ack receptions in the vicinities of the source and sink terminals. We now describe the details of the POWMAC access mechanism. In contrast to cellular systems where the base station makes the admission decision, in our case each terminal decides whether its transmission can proceed or not, depending on previously heard RTS, CTS, and DTS packets.

Each terminal i maintains a Power Constrained List denoted by PCL(i). This list is an extension of the Network Allocation Vector (NAV) used in the IEEE 802.11 scheme. Basically, PCL(i) encodes i's knowledge about other *active* terminals, i.e., terminals that are receiving, transmitting, or scheduled to do either function in i's vicinity. For every active terminal u in i's vicinity, PCL(i) contains the following entries (as explained shortly, these entries are computed using some information advertised by terminal u in its CTS or DTS control packets, and by measuring the signal strength of these control packets):

- The address of terminal u.
- The channel gain between terminals i and u (G_{iu}), computed using the received signal strength of u's control packet.

- The start time and duration of u's activities (data-reception/Ack-transmission or data-transmission/Ack-reception), as advertised by terminal u in its CTS or DTS packet.

- The maximum tolerable interference (MTI) of terminal u, denoted by $P_{\mathrm{MTI}}^{(u)}$ during u's data or Ack reception. This is the maximum additional interference that terminal u can tolerate from an interfering terminal such as terminal i. As will be explained shortly, this information is advertised by terminal u.

- The TP that terminal u will use during its scheduled data or Ack transmission, advertised in terminal u's CTS or DTS packet.

Let $\pi_i(u)$ be the maximum TP that terminal i can use without disturbing u's reception. Using G_{iu} and $P_{\mathrm{MTI}}^{(u)}$, terminal i computes $\pi_i(u)$ as:

$$\pi_i(u) = \min\left\{\frac{P_{\mathrm{MTI}}^{(u)}}{G_{iu}}\ ,\ \xi_{\max} P_{\max}\right\}. \tag{3.7}$$

Let $\Psi(i)$ be the set of terminals in i's vicinity whose receptions overlap with i's transmission ($\Psi(i) \subset \mathrm{PCL}(i)$). Then the maximum allowable TP that terminal i can use without disturbing any of its neighbors, denoted by $P_{\mathrm{MAP}}(i)$, is given by:

$$P_{\mathrm{MAP}}(i) = \min_{u \in \Psi(i)} \{\pi_i(u)\}. \tag{3.8}$$

Depending on the order in which terminals initiate their RTS messages in a given AW, we classify them into *master* and *slave* terminals. Terminal j is a master if it has a packet to send, its PCL is empty, and it does not sense any carrier signal. In this case, j's RTS packet announces the start of an AW (the size of this AW is also set by terminal j). On the other hand, a terminal, say k, is a slave terminal if it is in the vicinity of a master terminal, say j. In this case, terminal k may send an RTS message in any, but not the first, slot of the AW initiated by terminal j. Clearly, the master-slave designation is time-varying. We now explain the access rules for both master and slave terminals.

D.1 Master Terminals

Consider a master terminal, say j, that has a data packet to transmit to another terminal, say i. If j does not sense a carrier (for a random duration of time), it sends an RTS message at $P_{\max}$, and includes in this packet the values of $P_{\text{MAP}}^{(j)}$ and $N_{AW}^{(j)}$; the remaining number of slots in j's AW (how terminal j determines $N_{AW}^{(j)}$ will be explained shortly).

Upon receiving the RTS packet, receiver i uses the predetermined $P_{\max}$ value and the power of the received signal to estimate the channel gain G_{ji} between terminals j and i (note that we assume channel reciprocity, and so $G_{ij} = G_{ji}$) . The minimum TP that is needed so that i can decode the packet was given in (3.3). In that equation, $P_{\text{MAI}}^{(i)}$ represents the total MAI from *already ongoing* interfering transmissions, and it does not account for any interference tolerance[8]. Now, according to the load planning calculations in Section 3.3.3, the power that terminal j is allowed to use to send to i was given by $P_{\text{POWMAC}}^{(ji)}$ in (3.4). If $P_{\text{POWMAC}}^{(ji)} < P_{\min}^{(ji)}$ (i.e., $\xi^{(i)} >$ MLF), then the MAI in the vicinity of terminal i is greater than the one allowed by the planned loading. In this case, i responds with a negative CTS, informing j that it cannot proceed with its transmission (the negative CTS is used to prevent multiple RTS retransmissions from j). The philosophy behind this design is to prevent transmissions from taking place over links that perceive high MAI. This consequently increases the number of active links in the network, subject to the available power constraints, and limits the energy consumed in the $j \rightarrow i$ communication.

On the other hand, if $P_{\text{POWMAC}}^{(ji)} > P_{\min}^{(ji)}$, then it is possible for i to receive j's signal. In that case, i calculates the *maximum additional interference power* ($P_{\text{MAI-add}}^{(i)}$) that it can endure from future unintended transmitters so that the SINR at i does not drop below μ^*. This $P_{\text{MAI-add}}^{(i)}$ is given by:

$$P_{\text{MAI-add}}^{(i)} = \frac{G_{ji}}{\mu^*}(P_{\text{POWMAC}}^{(ji)} - P_{\min}^{(ji)}) = \left(\xi_{\max} - \xi^{(i)}\right) P_{\text{thermal}}. \tag{3.9}$$

The next step is to equitably distribute $P_{\text{MAI-add}}^{(i)}$ among future potential inter-

[8]In [95] the authors derived a finite value for the interference range in the case of minimum TP. However, the thermal noise power was not taken into account in that derivation.

ferers in the vicinity of i. The rational behind this distribution is to prevent one neighbor from consuming the entire $P^{(i)}_{\text{MAI-add}}$. In other words, we think of $P^{(i)}_{\text{MAI-add}}$ as a network resource that should be shared among various neighboring terminals. Recall that j's RTS contains $N^{(j)}_{AW}$; the remaining number of access slots in the current AW. Obviously, the number of concurrent transmissions should not exceed $N^{(j)}_{AW}$. Thus, terminal j uses $N^{(j)}_{AW}$ as the number of future potential interferers in its neighborhood.

Future interference at terminal i comes from interferers within the maximum range of i and interferers outside that range. The interference margin $P^{(i)}_{\text{MAI-add}}$ has to account for both types of interferers (if $P^{(i)}_{\text{MAI-add}}$ is distributed among within-range interferers only, an increase in the interference from outside the range of i could cause at packet collision at i). Let $P^{(i)}_{\text{MAI-within}}$ and $P^{(i)}_{\text{MAI-other}}$ be the two components of $P^{(i)}_{\text{MAI-add}}$. While terminal i can predict the number of within-range interferers, it cannot do the same for outside-range interferers. To estimate $P^{(i)}_{\text{MAI-other}}$, we follow a similar approach to the one used in cellular networks for an analogous problem. In cellular networks, the base station has control over in-cell interference (using open- and closed-loop power control), but it cannot influence out-of-cell interference. This problem has been thoroughly investigated in [102], and a practical (widely adopted) solution for it is to assume that the out-of-cell interference is a certain fraction of the in-cell interference. Considering the similarity between the role of a receiver in a power-controlled MAC protocol for MANETs and the role of a base station in cellular systems, we let $P^{(i)}_{\text{MAI-other}} = \varrho P^{(i)}_{\text{MAI-within}}$, where $\varrho \approx 0.5$ for the two-ray propagation model and uniformly distributed terminals. A simple weighting factor can be used to account for other distributions [102].

Based on the above, the maximum tolerable interference $P^{(i)}_{\text{MTI}}$ that a *single* future interferer can add to terminal i is set to:

$$P^{(i)}_{\text{MTI}} = \frac{P^{(i)}_{\text{MAI-add}}}{(1+\varrho)N^{(j)}_{AW}}. \tag{3.10}$$

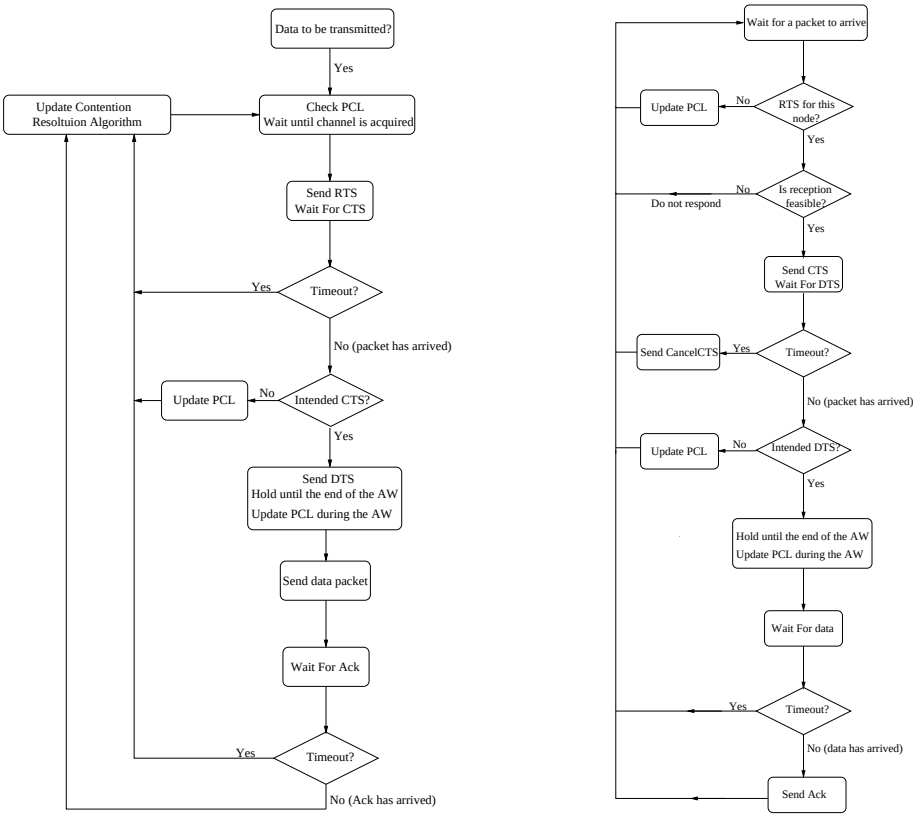
When responding to j's RTS, terminal i indicates in its CTS the power level $P^{(ji)}_{\text{POWMAC}}$ that j must use for the data transmission. In addition, terminal i inserts

$P_{\text{MTI}}^{(i)}$ in the CTS message to inform its neighbors of the maximum power they can use such that i's reception is not disturbed. The CTS is sent at an adjustable power ($P_{CTS}^{(i)}$) whose value is included in the CTS packet, as explained in Section 3.3.8.

Upon receiving i's CTS, terminal j replies back with a DTS packet that includes the value of $P_{\text{POWMAC}}^{(ji)}$. The DTS is needed to inform j's neighbors that may have not heard i's CTS about $P_{\text{POWMAC}}^{(ji)}$. Using $P_{\text{POWMAC}}^{(ji)}$ and the channel gain information, j's neighbors can compute the amount of *expected* MAI due to the scheduled transmission $j \rightarrow i$. The total expected MAI due to scheduled transmissions in the neighborhood of a terminal, say u, allows u to determine if it can receive a packet (data or Ack) following the current AW. If this MAI exceeds $\xi_{\max} P_{\text{thermal}}$, then u is expected to perceive high MAI, and therefore, should refrain from scheduling a reception; otherwise, u is free to receive a packet.

Similar to the CTS packet, the DTS packet contains the amount of additional interference $P_{\text{MTI}}^{(j)}$ that node j can tolerate during its Ack reception. As in [17], the DTS packet in POWMAC also provides a mechanism to announce the success of the RTS/CTS exchange between j and i to those neighbors of j who have not heard i's CTS. The IEEE 802.11 scheme uses carrier sensing for this purpose; if the neighbors of j do not sense a carrier after hearing the RTS for some time, they assume that the RTS/CTS exchange was not successful. This same mechanism, however, cannot be used in POWMAC since the data packet is transmitted at a power less than the RTS power, and thus the carrier sense range of the data packet is much smaller than that of the RTS (or CTS) packet. The DTS is also sent at an adjustable power as explained in Section 3.3.8. A flowchart that describes the operation of the POWMAC protocol is shown in Figure 3.4.

Once the RTS/CTS/DTS exchange is completed, no further negotiations are made for the corresponding data/Ack transmission. This makes TPC schemes in MANETs fundamentally different from their cellular counterparts. In cellular systems, every time a new session is started or terminated, the powers of ongoing transmissions are renegotiated. In contrast, power in MANETs is allocated only once at the start of the session, i.e., the whole data packet is transmitted at one



(a) When a terminal has data to send.

(b) When a terminal receives an RTS packet.

Figure 3.4: Flowchart for the operation of the POWMAC protocol.

power level, regardless of what follows the start of that packet transmission. The cellular approach requires that the entire state of the system (power used by every terminal in the network) be known whenever a new session is to be admitted, which cannot be achieved in a distributed MANET.

D.2 Slave Terminals

Slave terminals are terminals that are within the transmission range of a master terminal. In addition to the computations that master terminals perform (e.g., computing $P^{(.)}_{\text{POWMAC}}$, $P^{(.)}_{\text{min}}$, etc.), there are two "feasibility conditions" (FCs) that each slave terminal, say k, must fulfill for its activity (transmission or reception) to proceed simultaneously with each scheduled activity in k's vicinity. The FCs are:

- FC_1 *(Effect of terminal k's transmission on the receptions in k's neighborhood)*: Terminal k's data or Ack transmission should not disturb already scheduled receptions in k's vicinity.

- FC_2 *(Effect of k's neighbors' transmissions on k's reception)*: The additional interference due to already scheduled transmissions should not increase the load factor at terminal k above $\xi_{\max}$ during terminal k's data or Ack reception.

The two FCs must be satisfied with respect to *all* scheduled activities in k's vicinity that are known to terminal k. As will become clear shortly, the chances for terminals to fulfill their FCs can be improved by allowing pairs of communicating terminals to move forward the transmission times of their Ack packets. In other words, POWMAC allows for a delay lag between the reception of a data packet and the transmission of its corresponding Ack packet. Thus, a recipient, say i, of a data packet may wait for a certain period, denoted by $\tau_{Ack}^{(ji)}$ (see Figure 3.3), before sending the Ack to terminal j. This lag allows i to avoid overlapping its Ack transmission (reception) with other data or Ack receptions (transmissions) in i's or j's vicinities. $\tau_{Ack}^{(ji)}$ is communicated using a 1-byte field in the RTS/CTS/DTS packets. Note that it is not useful to change the transmission time of a data packet to avoid overlapping data packets since the main goal of POWMAC is for data packets to proceed simultaneously.

Delaying the transmission time of an Ack packet must be carefully coordinated between the the source and sink terminals; otherwise, conflicts may arise and may result in collisions. For example, the source may choose to delay the Ack by Δ_1 seconds, and later on the sink terminal chooses to delay the same Ack packet by $\Delta_2 < \Delta_1$ seconds, thus violating the source's FCs. Another issue is how to compute $\tau_{Ack}^{(ji)}$ when there are multiple scheduled activities in terminal j's neighborhood (each activity calls for a different value of $\tau_{Ack}^{(ji)}$). To address these issues, we establish two "viability rules" (VRs) for changing the Ack transmission time:

- VR_1: Each terminal that wishes to fulfill its FCs (with respect to a certain neighboring activity) is allowed to increase the present value of $\tau_{Ack}^{(.)}$, but not

decrease it.

- VR_2: Each terminal computes $\tau_{Ack}^{(.)}$ that fulfills its FCs with respect to a given neighboring activity in such a way that if $\tau_{Ack}^{(.)}$ is later *increased* by the same terminal to accommodate another neighboring activity or is increased by the communication peer, then that terminal's FCs are not violated. An example that explains this rule will be given shortly.

The significance of the VRs is that they allow each terminal to *independently* consider its interaction with its active neighbors (i.e., fulfill its FCs by choosing an appropriate $\tau_{Ack}^{(.)}$) on a *per-terminal* basis. To illustrate, consider Figure 3.5, where four terminals are in the same vicinity, i.e., control packets of any terminal are heard by the other three terminals. Terminal j has already scheduled a data packet transmission to terminal i. Terminal v wishes to schedule a transmission to terminal n simultaneously with the transmission $j \rightarrow i$. The VRs allow terminal v to evaluate its future interaction with terminal j and accordingly choose a value for $\tau_{Ack}^{(vn)}$, and independently to consider its interaction with terminal i and accordingly choose a possibly different value for $\tau_{Ack}^{(vn)}$. Furthermore, the VRs also allow the receiving terminal n to independently change the value of $\tau_{Ack}^{(vn)}$ to fulfill its own FCs without worrying that this new value could affect the FCs at terminal j. To demonstrate how terminals operate to fulfill their FCs, we examine four basic scenarios shown in Figures 3.6, 3.11, 3.12, and 3.13. In these scenarios, terminals j and i have just completed an RTS/CTS/DTS exchange. The (slave) terminal v has a data packet that it wishes to transmit to terminal n. We now examine what terminals v and n have to do in each scenario to fulfill the FCs. Other scenarios are basically a combination of one or more of these four scenarios.

1) Source-Source Interaction:

The first scenario represents source-source interaction. An example of this scenario is shown in Figure 3.6. Here, source terminal v can potentially interfere with source terminal j, and vise versa. After v hears j's RTS and DTS messages, it uses the

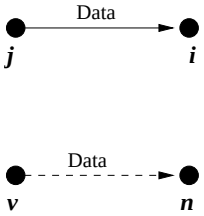


Figure 3.5: Example of a network topology where POWMAC allows for two simultaneous transmissions in the same vicinity.

Figure 3.6: Scenario that describes a source-source interaction.

signal strength of the received RTS message and the value of the RTS transmission power ($P_{\max}$) to estimate the channel gain G_{vj}. The channel gain and the value of $P_{\text{MTI}}^{(j)}$ (included in the DTS message) are used to update the maximum power $P_{\text{MAP}}^{(v)}$ that v can use in its future transmissions, according to (3.8), *during j's Ack reception*. Terminal v also records the transmission times and the TP of the $j \rightarrow i$ data/Ack packets (recall that $P_{\text{POWMAC}}^{(ji)}$ is used for both data and Ack). This information is part of the DTS; the exact format of the control packets will be given later.

In order for terminal v to fulfill its FCs, it compares its data packet length with j's data packet length. Note that terminals that contend in the same AW schedule their data transmissions to start at the same time but may complete them at different times. If v's data packet is shorter than j's data packet (see Figures 3.7 and 3.8), and the additional interference due to j's data transmission (i.e., $P_{\text{POWMAC}}^{(ji)} G_{vj}$) would *not* increase the load factor at terminal v beyond $\xi_{\max}$ during terminal v's Ack reception, then v does not do any more computations. Else, v delays the Ack transmission time until j finishes its data transmission, i.e., the Ack packet is moved from Position 1 to Position 2 in both figures. This way, terminal v satisfies FC_2, while FC_1 is also satisfied (with respect to the interaction $v \leftrightarrow j$) since v's

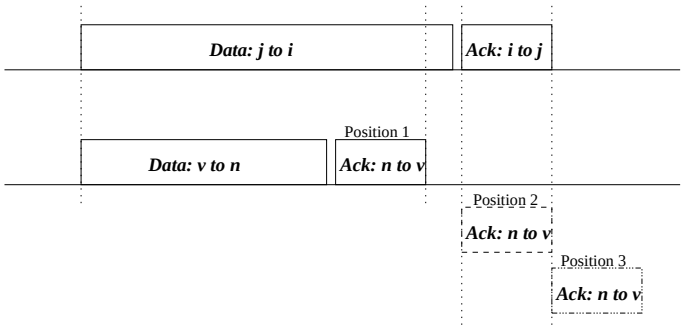


Figure 3.7: Slave terminal's Ack packet transmission completes before master terminal's Ack transmission starts.

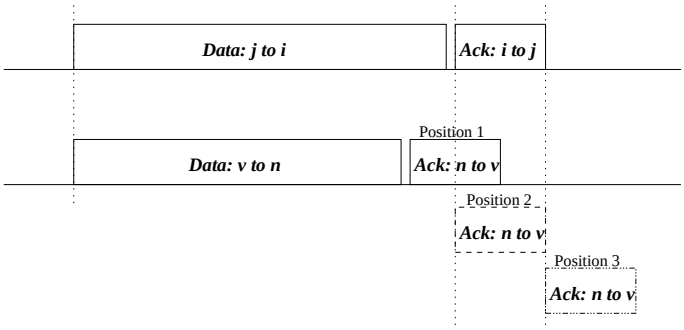


Figure 3.8: Slave terminal's Ack packet transmission overlaps with master terminal's Ack packet transmission.

transmission does not overlap with j's reception.

In case v's data packet is equal to j's data packet, then v does not do any more computations. If v's data packet is longer than j's data packet (see Figures 3.9 and 3.10), then the maximum TP used by v for its data transmission must not exceed the new value of $P_{\mathrm{MAP}}^{(v)}$ updated from j's DTS message. Terminal v cannot decide in advance how much TP the communication $v \rightarrow n$ requires. Therefore, v includes the value of $P_{\mathrm{MAP}}^{(v)}$ in its RTS message and leaves the decision of the TP determination to receiver n. This way, terminal v satisfies FC_1, while FC_2 is also satisfied (with respect to the interaction $v \leftrightarrow j$) since v's reception does not overlap with j's transmission. Note that both VR_1 and VR_2 are satisfied in all the above cases when considering the interaction $v \leftrightarrow j$.

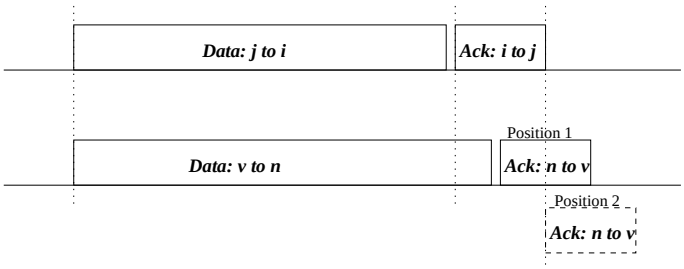


Figure 3.9: Another case where slave terminal's Ack packet transmission overlaps with master terminal's Ack packet transmission.

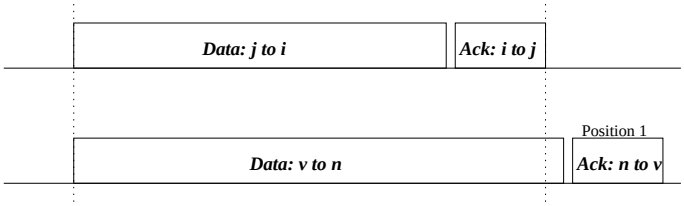


Figure 3.10: Slave terminal's Ack packet transmission starts after master terminal's Ack packet transmission completes.

2) Source-Sink Interaction:

The second scenario is shown in Figure 3.11, where the source terminal v can potentially interfere with an already scheduled reception at sink i, and vise versa. When v hears i's CTS, it uses the signal strength of the received message and the value of the CTS transmission power (included in the CTS) to estimate the channel gain G_{vi} between itself and terminal i. The channel gain and the broadcasted $P_{\mathrm{MTI}}^{(i)}$ value are used to update the maximum power $P_{\mathrm{MAP}}^{(v)}$ that v can use in its future transmissions, according to (3.8). Terminal v also records the transmission times and the TP $P_{\mathrm{POWMAC}}^{(ji)}$ of the $j \rightarrow i$ data/Ack packets.

In order for v to fulfill FC_1, its maximum TP must not exceed the new value of $P_{\mathrm{MAP}}^{(v)}$. Terminal v cannot decide in advance how much TP the communication $v \rightarrow n$ requires. Therefore, v includes the value of $P_{\mathrm{MAP}}^{(v)}$ in its RTS message and leaves the decision to the receiver n.

Now, in order for v to fulfill FC_2, it checks whether the additional interference due

Figure 3.11: Scenario that describes a source-sink interaction.

to i's Ack transmission (i.e., $P^{(ji)}_{\text{POWMAC}} G_{vi}$) would increase the load factor at terminal v beyond $\xi_{\max}$. If it would not, then v does not do any more computations; else, v checks if there is an overlap between its Ack reception and i's Ack transmission. There are three possibilities to consider:

- If there is no overlap and v's Ack reception starts *after* i finishes its Ack transmission (see Figure 3.10), then v does not perform any more computation to satisfy FC_2 with respect to the $v \leftrightarrow i$ interaction.

- If there is an overlap (see Figures 3.8 and 3.9), then terminal v delays the Ack until i finishes its Ack transmission, i.e., the Ack packet is moved from Position 1 (or 2) to Position 3 in Figure 3.8 and from Position 1 to Position 2 in Figure 3.9. This way, terminal v satisfies FC_2.

- The last case is the one shown in Figure 3.7 where there is no overlap and v's Ack reception finishes *before* i starts its Ack transmission. This case requires special attention. Recall that to increase the chances for terminals to fulfill their FCs, we allow pairs of communicating terminals to move forward the transmission times of their Ack packets. This means that the receiver, terminal n in this case, may actually delay the Ack transmission time to fulfill its own FCs, which could violate v's FCs (for example, if terminal n delays the Ack transmission time such that the new schedule results in an overlap between v's Ack reception and i's Ack transmission). Therefore, terminal v delays the Ack reception time until i finishes its Ack transmission, i.e., the Ack packet is moved from Position 1 to Position 3 in Figure 3.7. This example shows the importance of VR_2.

Figure 3.12: Scenario that describes a sink-source interaction.

3) Sink-Source Interaction:

The third scenario is shown in Figure 3.12, where the sink terminal n can potentially interfere with the source terminal j, and vise versa. When n hears j's RTS and DTS packets, it uses the signal strength of the received RTS message and $P_{\max}$ to estimate the channel gain G_{nj}. The channel gain and the value of $P_{\text{MTI}}^{(j)}$ (included in the DTS message) are used to update the maximum power $P_{\text{MAP}}^{(n)}$ that n can use in its future transmissions *during j's Ack reception.* Terminal n also records the transmission times and the TP $P_{\text{POWMAC}}^{(ji)}$ of the $j \rightarrow i$ data/Ack packets.

Next, terminal n checks if the additional interference due to j's data transmission (i.e., $P_{\text{POWMAC}}^{(ji)} G_{nj}$) would increase the load factor at terminal n beyond $\xi_{\max}$. If it would, then n decides that it cannot receive v's data packet and sends back a negative CTS. Else, the interference due to j's transmission is small enough to allow n to receive v's data packet (i.e., FC_2 is fulfilled).

Now, in order for terminal n to fulfill FC_1, it checks if the TP of its Ack transmission ($P_{\text{POWMAC}}^{(vn)}$) exceeds the new value of $P_{\text{MAP}}^{(n)}$, updated from j's DTS message. If it does not, then n does not do any more computations; else, n checks if there is an overlap between its Ack transmission and j's Ack reception. If so, there are three possibilities to consider:

- If there is no overlap and n's Ack transmission starts *after* j finishes its Ack reception (see Figure 3.10), then n does not do any further computations.

- If there is an overlap (see Figures 3.8 and 3.9), then terminal n delays the Ack transmission time until j finishes its Ack reception, i.e., the Ack packet is moved from Position 1 to Position 3 in Figure 3.8 and from Position 1 to Position 2 in Figure 3.9. This way, terminal n satisfies FC_1.

Figure 3.13: Scenario that describes a sink-sink interaction.

- The last case is the one shown in Figure 3.7 where there is no overlap and n's Ack transmission finishes *before* j starts its Ack reception. As before, and to satisfy FC_2, terminal n delays the Ack transmission time until j finishes its Ack reception, i.e., the Ack packet is moved from Position 1 to Position 3 in Figure 3.7.

4) Sink-Sink Interaction:

The last scenario is the one shown in Figure 3.13, in which the sink terminal n can potentially interfere with the already scheduled sink i, and vice versa. When n hears i's CTS, it uses the signal strength of the received message and the value of the CTS transmission power to estimate G_{ni}. The channel gain and the broadcasted $P_{\mathrm{MTI}}^{(i)}$ value are used to update $P_{\mathrm{MAP}}^{(n)}$ according to (3.8). Terminal n also records of the transmission times and the TP $P_{\mathrm{POWMAC}}^{(ji)}$ of the $j \rightarrow i$ data/Ack packets.

In order for terminal n to fulfill its FCs, it compares its data packet length with i's data packet length (obtained from a field in the CTS). In case n's data packet is shorter than i's data packet (see Figures 3.7 and 3.8) and the TP of n's packet is less than $P_{\mathrm{MAP}}^{(n)}$, then n does not do any more computations. Else, n delays the Ack transmission time until i finishes its data reception, i.e., the Ack packet is moved from Position 1 to Position 2 in both figures. This way, terminal n satisfies FC_1, with FC_2 being also satisfied with respect to the interaction $n \leftrightarrow i$, since n's reception does not overlap with i's transmission.

In case n's data packet is equal to i's data packet, then n does not do any more computations. In case n's data packet is longer than i's data packet (see Figures 3.9 and 3.10) and the additional interference due to i's Ack transmission (i.e., $P_{\mathrm{POWMAC}}^{(ji)} G_{ni}$) would *not* increase the load factor at terminal n beyond $\xi_{\max}$

during terminal n's data reception, then n does not do any more computations. Else, terminal n decides that it cannot receive v's data packet and a negative CTS message is sent back to terminal v.

Figure 3.14 summarizes the sequence of events that take place at the four terminals for the network in Figure 3.5. In this example, we let all terminals be within the transmission range of each other. We assume that v's data packet is shorter than j's data packet. The interference from j on v would increase the load factor at v beyond $\xi_{\max}$, so terminal v delays its Ack reception until j finishes its data packet transmission. Finally, the interference from n's Ack transmission is less that the tolerable interference for j's Ack reception, so terminal n does not change $\tau_{Ack}^{(vn)}$.

3.3.5 Contention Resolution

For contention resolution, we follow the work in [84], which, unlike the IEEE 802.11 scheme, performs contention resolution in the persistent domain instead of the backoff domain. As shown in [84], if the access probability (x_r) of terminal r is adapted according to

$$\dot{x_r} = \alpha - \beta p_r x_r \ , \tag{3.11}$$

where α and β are system parameters, and p_r is the loss probability experienced by terminal r, then the system converges to an optimal point that maximizes the network throughput under a proportional fairness model.

If a terminal, say r, wants to transmit a data packet, it first verifies that its FCs are satisfied. If so, then with probability x_r, r contends for the channel in the next access slot of j's AW (j is a neighboring master terminal). If successful, terminal r chooses a wait time B_r that is uniformly distributed in the interval [0,B]; B is a system-wide backoff counter. After this waiting time, terminal r senses the channel. If the channel is free, terminal r transmits its RTS in the current access slot. Note that B is in the order of few microseconds while a time slot is in milliseconds, so the backoff mainly serves to prevent synchronized RTS attempts. Figure 3.15 shows the state diagram of the contention resolution algorithm. Note that x_r is increased

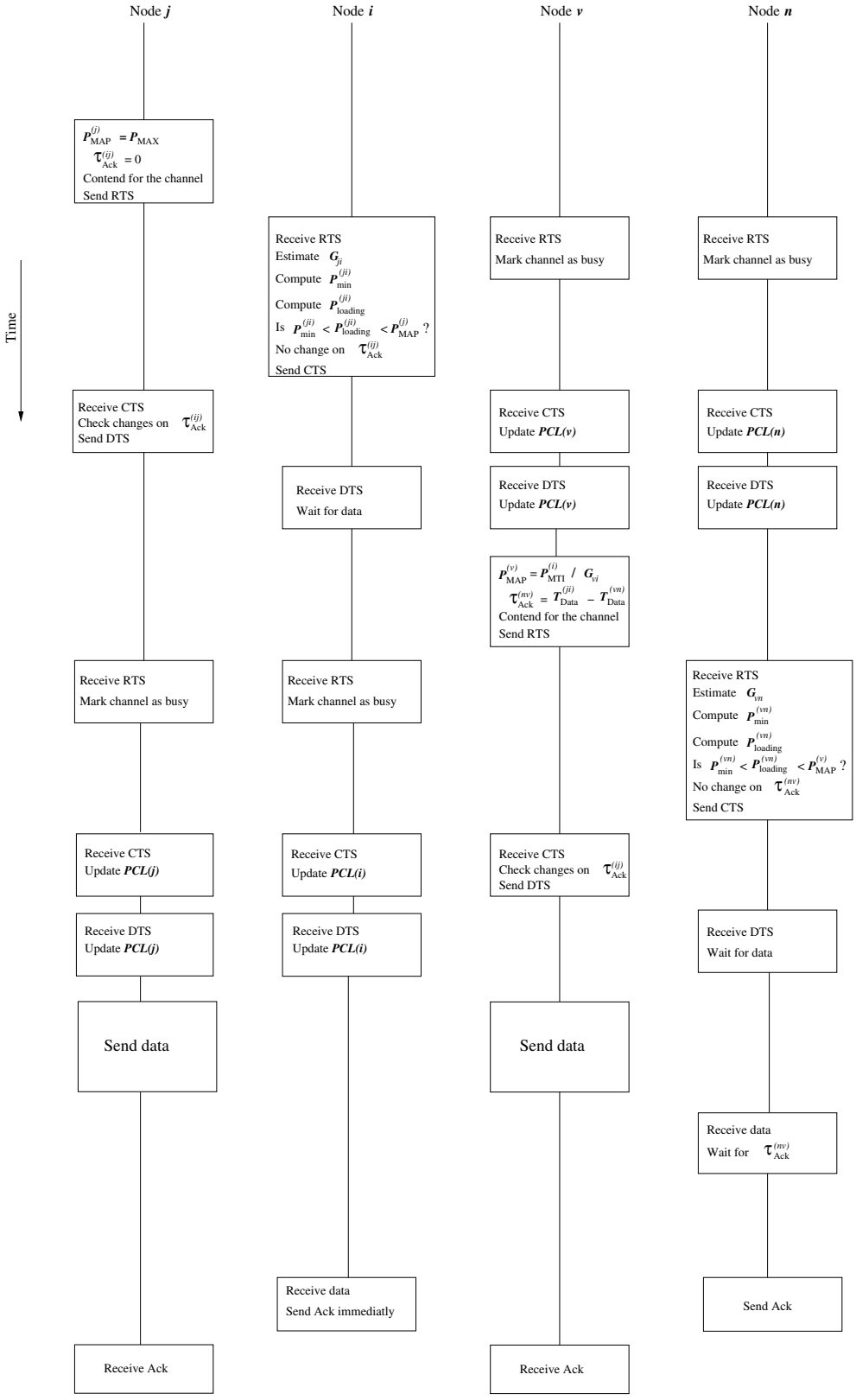


Figure 3.14: Sequence of events that allow terminals in POWMAC to schedule simultaneous transmissions in the same vicinity ($T^{(ji)}_{Data}$ and $T^{(vn)}_{Data}$ refer to the lengths of j's and v's data packets, respectively).

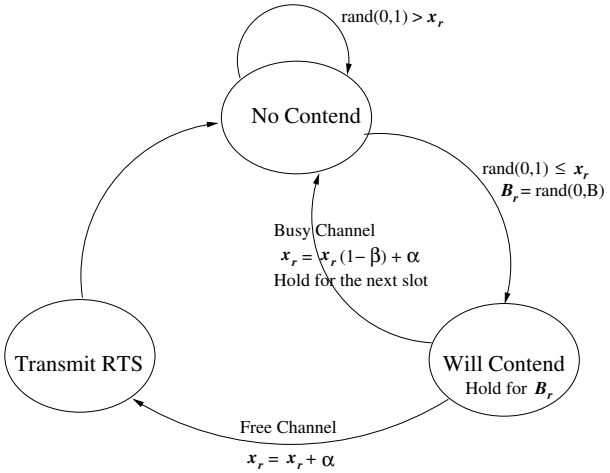


Figure 3.15: State diagram of the contention resolution algorithm used in the POWMAC protocol.

by α at the end of each access slot, but decreased by β only when the contention is not successful (i.e., with probability p_r). Hence, (3.11) is satisfied. Note also that when using this mechanism for POWMAC, we do *not* require any synchronization. Basically, once terminal r receives j's RTS, it divides its time access into $N_{AW}^{(j)}$ slots of predetermined length, regardless of the absolute time at terminals r and j. This issue is explored further in the next section.

3.3.6 Synchronization of the Access Window

So far, we have assumed that terminals can synchronize with a neighboring master terminal. We now explain the mechanism underlaying this process. *Note that by synchronization, we do not mean that terminals have the same clock; rather, they can determine the boundaries of the AW slots.* Consider the scenario in Figure 3.16 where master terminal j has scheduled a data transmission to terminal i, and (slave) terminal v has synchronized with j's AW (as we will explain shortly) and has scheduled a data transmission to terminal n. Suppose now that terminal k wishes to transmit to terminal l. We now explain how k synchronizes with j's AW. Note that v is in i's but not j's vicinity, and likewise, k is in n's but not v's vicinity.

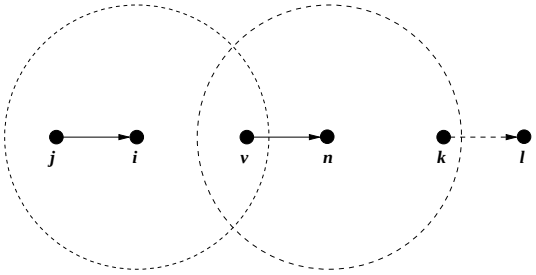


Figure 3.16: Example that illustrates how slave terminals synchronize with the master terminal's AW. The two circles represent the maximum transmission ranges of terminals i and n.

First, we design the duration of the AW slot (AWS) to be fixed and common to all terminals. Specifically, an AWS consists of the sum of the transmission durations of the RTS, CTS, and DTS packets, the maximum backoff interval, plus two fixed short interframe spacing (SIFS) periods[9]. However, fixing the AWS duration is not enough for terminal k to synchronize with j's AW; the reason is that when v transmits its RTS message, it chooses a random wait time B_v that is uniformly distributed in the interval [0,B]. Since k hears only n's CTS, it is not possible for k to synchronize with j's AW. The situation is exemplified in Figure 3.17. The main problem is that k cannot determine the value of B_v, and so it cannot determine the end of that AWS. To remedy this situation, the value of B_v is announced in both the RTS and CTS control packets, allowing terminal k to synchronize with j's AW.

Finally, when the master terminal j sends its RTS message, it sets the value of B_j in the RTS message to the maximum backoff duration B. Thus, the following slot in the AW (i.e., the slot where v and n exchange their control messages) starts immediately after the reception of the j's DTS message, as shown in Figure 3.17.

3.3.7 Updating the Access Window Size

The AW size at a terminal, say j, is updated adaptively as a function of the load in the vicinity of j. The goal is to choose an AW size that maximizes the chances

[9]As defined in the IEEE 802.11b standard [2], a SIFS period consists of the processing delay for a received packet plus the turnaround time.

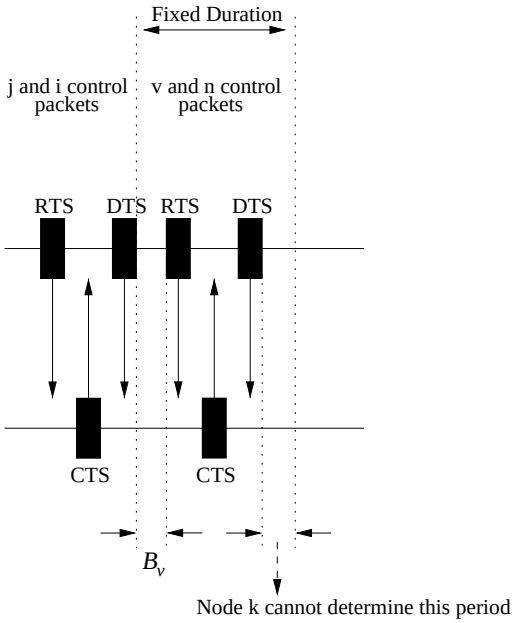


Figure 3.17: Example that illustrates the challenge in synchronizing with the AW of a master terminal.

of concurrent data transmissions. To achieve that, terminal j examines two history values: the actual interference perceived by terminal j during its reception, and the number of concurrent data transmissions and receptions in j's vicinity.

At the end of the data reception at terminal j, if the actual interference perceived by terminal j is higher than a given fraction (e.g., 75%) of the planned interference $\xi_{\max} P_{\text{thermal}}$, then the AW size need not be changed, since the allocated additional power to combat MAI was efficiently utilized to allow for concurrent transmissions.

On the other hand, if less than that threshold was used, then terminal j should adapt (either increase or decrease) the AW size so that the allocated power is not wasted. To this end, terminal j checks the number of concurrent transmissions that actually took place in the previous AW (based on the numbers of CTS and DTS packets). If this number is less than, say η%, of the AW size, then either the load is low or the value of the AW size is too big to the extent that $P_{\text{MTI}}^{(j)}$ is too small (see (3.10)), i.e., $P_{\text{MTI}}^{(j)}$ is not large enough to allow for other nearby terminals to transmit. In both of these cases, terminal j decreases its AW size. In contrast, if the

number of concurrent transmissions that actually took place in that AW is greater than $\eta\%$ of the AW size, then there is room for increasing the number of concurrent transmissions in the vicinity of terminal j. Hence, the AW size is increased. In case of a data-packet collision, the AW size is kept constant. Note that a collision may happen if the control messages were not successfully heard by a neighboring station. Finally, to prevent unstable fluctuations in the AW values, the AW size is incremented or decremented in steps of 1.

3.3.8 Adaptive Reservation Mechanism

In the IEEE 802.11 scheme, the RTS and CTS packets are transmitted at a fixed power $P_{\max}$. As discussed in Section 3.1, this approach can be overly conservative. Recall that in POWMAC, a receiver, say i, sends a CTS packet that contains CAI, namely $P_{\text{MTI}}^{(i)}$, to bound the TP of potentially interfering neighbors. A terminal, say n, that hears this packet sets its $P_{\text{MAP}}^{(n)}$ according to (3.8). If $\xi_{\max}P_{\max}$[10] is less than $P_{\text{MTI}}^{(i)}/G_{ni}$, the CAI is actually irrelevant to terminal n, and the CTS packet has reached farther than necessary. In POWMAC, this issue is not harmful as in the IEEE 802.11 scheme, simply because control packets in POWMAC do not prevent neighbors from transmitting. Nonetheless, one way to further enhance the operation of POWMAC is to transmit control packets only to those terminals who can actually make use of the CAI. This has the added advantage of reduced contention among control packets, leading to an increase in the spatial reuse. POWMAC uses the following adaptive TP approach for the control packets.

The farthest neighbor from terminal i that can actually make use of the CAI contained in i's CTS (node s in Figure 3.18) is the one with channel gain of $P_{\text{MTI}}^{(i)}/(\xi_{\max}P_{\max})$. For any other terminal n that is more than G_{is} away from i, $\xi_{\max}P_{\max}$ is less than $P_{\text{MTI}}^{(i)}/G_{iv}$, and thus the CAI that is contained in i's CTS is irrelevant to terminal n. Accordingly, we set the range of the CTS of terminal i to

[10]Recall that the maximum TP in POWMAC is $\xi_{\max}P_{\max}$.

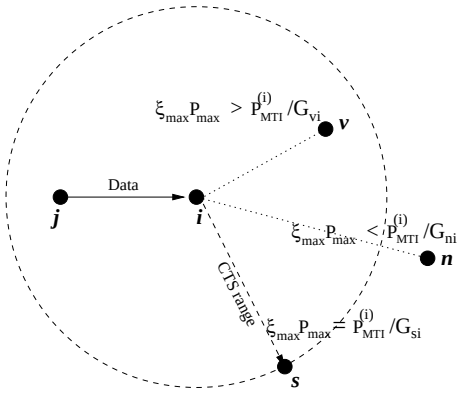


Figure 3.18: Range of the CTS message is limited to neighbors that can make use of the CAI conveyed in the CTS message.

$P^{(i)}_{\mathrm{MTI}}/(\xi_{\max}P_{\max})$. Thus, the TP for the CTS packet of terminal i is:

$$P^{(i)}_{CTS} = \min\left\{\mu^* P_{\mathrm{thermal}}\frac{\xi_{\max}P_{\max}}{P^{(i)}_{\mathrm{MTI}}}\ ,\ \xi_{\max}P_{\max}\right\}, \tag{3.12}$$

where the minimum is taken because of the hardware constraints of the wireless interface. A similar computation is also applied to find the TP of the DTS packet at the transmitter. Note that the CTS (or DTS) packet may not be heard by all potential interferers (because of the hardware constraints of the wireless interface, i.e., the second term in the right hand side of (3.12) is less than the first). Such a limitation also exists in the IEEE 802.11 scheme, as it does not prevent nodes in the interference region from causing collisions with the data packet at the destination node (see [63] for details). Thus, this problem is not introduced by the proposed protocol. Moreover, POWMAC already takes into account future MAI due to terminals that do not hear the control packets by using $\varrho = 0.5$ in (3.10). Note also that in (3.12), we assume no interference at the CTS receiver. This is because in the design of wireless systems, the maximum range is typically calculated using only the thermal noise value [87], since there is no way of predicting all potential interferers beforehand.

Before concluding this section, we give the formats of the various control packets in POWMAC. For a source terminal j and a sink terminal i, the format of the RTS

packet is:

$$\text{RTS}(j \rightarrow i) = \left\{ j,\ i,\ P^{(j)}_{\text{MAP}},\ N^{(j)}_{AW},\ T^{(ji)}_{\text{data}},\ \tau^{(ji)}_{Ack},\ B_j \right\}. \tag{3.13}$$

The format of the CTS packet is:

$$\text{CTS}(i \rightarrow j) = \left\{ i,\ j,\ P^{(ji)}_{\text{POWMAC}},\ P^{(i)}_{\text{MTI}},\ P^{(i)}_{CTS},\ N^{(j)}_{AW},\ T^{(ji)}_{\text{data}}, B_j,\ \tau^{(ji)}_{Ack} \right\}. \tag{3.14}$$

Finally, the format of the DTS packet is:

$$\text{DTS}(j \rightarrow i) = \left\{ j,\ i,\ P^{(ji)}_{\text{POWMAC}},\ P^{(j)}_{\text{MTI}},\ \tau^{(ji)}_{Ack} \right\}. \tag{3.15}$$

3.3.9 POWMAC Limitations

In this section, we discuss some of the limitations of POWMAC and outline possible remedies for them. Specifically, we present two scenarios where concurrent transmissions in the same vicinity are, in principle, possible but may not be allowed under POWMAC.

So far, we have assumed that slave terminals are in the transmission range of only one master terminal. However, this may not be true; the example shown in Figure 3.19 presents a scenario where slave terminal v is within the transmission range of the two (unsynchronized) master terminals j and l. According to POWMAC, terminal v may send its RTS packet (or respond with a CTS to terminal n) only if the two master terminals' AWs are misaligned by less than the maximum backoff window B, since otherwise, the control/data packets sent by terminal v will not be synchronized with at least one of its masters.

A second scenario is shown in Figure 3.20, where terminal n has synchronized with the master j as a result of hearing j's RTS packet, while terminal v is out of j's transmission range, and is thus unaware of the j's AW. According to POWMAC, if n receives an RTS packet from v, then it responds with a CTS only if v's proposed AW is misaligned with j's AW by less than the maximum backoff window B.

A close look at the above two scenarios reveals that they both occur when a terminal, say v, that is two hops away (see Figure 3.20) from a master terminal (j in that figure), is unaware of j's AW slots alignment. If v starts its own AW,

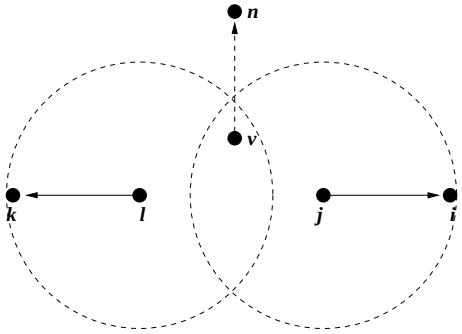


Figure 3.19: Example of a slave terminal v that falls in the transmission ranges of two (unsynchronized) master terminals j and l.

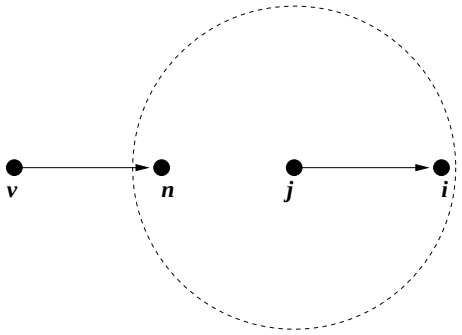


Figure 3.20: Terminal n, which is in j's vicinity, receives an RTS message from terminal v, which is not in j's vicinity. Terminal n may or may not be able to respond to v's RTS message.

then there is a good chance that the AWs of j and v are not synchronized. One possible approach that can reduce the chances of such scenarios to occur is to allow terminals that *overhear* any RTS/CTS/DTS messages (e.g., terminal n), to send their own RTS messages before terminals that are outside j's range (e.g., terminal v) send theirs. The idea here is to allow more terminals to synchronize with the same master. We cannot actually guarantee that n sends its RTS before v, because of the randomness in the contention resolution mechanism; however, what we can do is to increase the access probability x_n of terminal n (see (3.11)) beyond that of x_v, thus reducing the probability that the above two scenarios will occur.

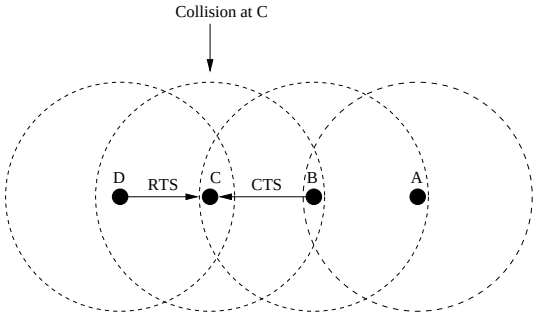


Figure 3.21: Example of a collision between control packets that eventually leads to a collision with a data packet.

3.3.10 Protocol Recovery

In [31] the authors observed that when the transmission and propagation times of control packets are long, the likelihood of a collision between a CTS packet and an RTS packet of another contending terminal increases dramatically; the vulnerable period being twice the transmission duration of a control packet. At high loads, such a collision can lead to collisions with data packets, as illustrated in Figure 3.21. In this figure, terminal D starts sending a RTS to terminal C while C is receiving B's CTS that is intended to A. A collision happens at C, hence C is unaware of B's subsequent data reception. Afterwards, if C receives a retransmitted RTS packet from node D and decides to reply back with a CTS, it may destroy B's reception.

Another problem is if the interference goes above the planned interference tolerance $P^{(.)}_{\text{MAI-add}}$. In POWMAC, we rely on two mechanisms to solve the above two problems. First, we require the carrier-sense range to be at least twice the maximum transmission range[11]. This makes the vulnerable period twice the propagation delay (less than 1 microsecond) instead of twice the transmission duration of a control packet (in the order of 100s of microseconds) and thus, the chances of control packets collisions will decrease significantly in the case of no channel shadowing effect. The second mechanism is to send a control packet preventing a potential interferer

[11]In fact, typical values for the carrier-sense range are more than twice the transmission range [63].

from commencing its transmission. In other word, suppose that while waiting in an AW to receive a data packet, terminal i hears an RTS message (destined to any terminal) that contains an allowable power $P_{\text{map}}^{(.)}$ value that if used could cause an unacceptable interference with i's scheduled reception. Then terminal i shall respond *immediately* with a special CTS, preventing the RTS sender from commencing its transmission. This method is similar to the use of the Object-to-Send (OTS) control packet proposed in [133, 130]. To see how this solution helps in reducing the likelihood of collisions with data packets, consider the situation in Figure 3.21. Suppose that terminal A sends a RTS to terminal B, and B responds back with a CTS that collides at C with a RTS from D. Now, C does not know about B's ongoing reception. Two scenarios can happen. In the first, terminal C may later wish to send a packet to, say, terminal D. It sends a RTS, which will be heard by terminal B. B responds back with a special CTS. Note that there is a good chance that B's special CTS will collide with the CTS reply from D; however, this is desirable since C will fail to recover D's CTS packet, and will therefore defer its transmission and invoke its backoff procedure. In essence, B's special CTS acts as a jamming signal to prevent C from proceedings with its transmission.

Note that in POWMAC we try to avoid likely collision scenarios such as the one mentioned in [31]. However, there are still few complicated (and definitely much less probable) scenarios where data packets may collide; recovery from such collisions is left to the upper layers.

3.3.11 Mobility and POWMAC

To determine the TP for data packets, POWMAC relies on the assumption that the channel gain determined at the time of the RTS/CTS/DTS exchange is stationary for the duration of the current AW and the ensuing data packets. The channel gain can change as a result of mobility. However, as we now explain, such a change has no impact on the assumptions used in POWMAC.

For *large-scale* channel variations (e.g., mean channel gain), mobility has negligible impact on POWMAC since packet transmission times occur on the scale of few

milliseconds while mobility occurs on the time scale of seconds. So the time between a control packet and an ensuing data packet is small enough to make the estimation sufficiently accurate. As for *small-scale* channel variations, although their impact can be mitigated through diversity techniques at the physical layer (e.g., RAKE receivers [100]), we argue that even if such techniques are not available, the channel stationarity assumption in POWMAC is still valid. Consider a multipath environment, where multiple versions of the transmitted signal arrive at the receiver at slightly different times and combine to give a resultant signal that can vary widely in amplitude and phase. The spectral broadening caused by this variation is measured by the Doppler spread, which is a function of the relative velocity (v) of the mobile and the angle between the direction of motion and the directions of arrival of the multipath waves [100]. This variation can be equivalently measured in the time domain using the *coherence time* (T_c), which is basically a statistical measure of the time duration over which the channel can be assumed time invariant. As a rule of thumb in modern communication system, $T_c \approx 0.423/f_m$, where $f_m = v/\Lambda$ is the maximum Doppler shift and Λ is the wavelength of the carrier signal.

Now, at a mobile speed of $v = 1$ meter/sec and 2.4 GHz carrier frequency, $T_c \approx 52.89$ msec. This time reduces to 10.56 msec when $v = 5$ meters/sec. For the channel stationarity assumption in POWMAC to be valid, the access window and the data packet duration must not exceed T_c. At a channel rate of 2 Mbps, it takes 4 msec to transmit a 1000-byte packet. This duration of time becomes even less at higher data rates. The propagation delay and the turnaround time (time it takes a terminal to switch from a receiving mode to a transmitting mode) are in the order of microseconds, and so they can be safely ignored. Thus, the assumption about channel stationarity is valid for moderate values of mobility (e.g., pedestrian speeds). The IEEE 802.11 was designed for such mobility scenarios [13]. In cases when terminals move faster, the packet size can be shortened so that the stationarity assumption still holds.

3.3.12 POWMAC in Rate-Controlled Environments

In this section, we explain how rate control can be combined with the POWMAC protocol. The IEEE 802.11b specifications provide a physical-layer multi-rate capability. All control packets are transmitted at the lowest rate (1 Mbps) to achieve the maximum range, while data packets can be transmitted at rates 1, 2, 5.5, and 11 Mbps. These different rates are achieved using multiple modulation schemes; BPSK, QPSK, and two variants of CCK. The higher is the rate, the higher is the SINR threshold (i.e., μ^*) that is needed to achieve the target BER.

Several schemes have been proposed for rate adaptation (e.g., [51]). The main idea in such schemes is to use the measured SINR of the received RTS packets to set the transmission rate for each data packet according to the highest feasible value allowed by the channel condition[12]. These schemes use a fixed TP, and a higher rate if the measured SINR is more than μ^*, i.e., these approaches utilize the additional available power in the received signal to allow for a higher rate. POWMAC, on the other hand, utilizes that additional signal power to allow for interference-limited transmissions in the neighborhood of a receiver. This, however, does not mean that a TP scheme and a rate control scheme cannot be combined together. In fact, it was shown in [42] that adapting the transmit power, data rate, and coding scheme achieves maximum spectral efficiency. For example, one way to integrate the protocol proposed in [51] with POWMAC is as follows. First, the maximum feasible rate is chosen according to the scheme in [51]. Second, the POWMAC protocol is used with the required μ^* for that chosen rate being used in (3.4). The message here is that POWMAC and rate-control schemes are complementary for maximizing network throughput.

[12]Note that in the above schemes, the RTS and CTS packets are still transmitted at the lowest rate so that neighboring terminals can overhear these packets and are informed of the ensuing data transmission

3.3.13 Protocol Overhead

We now explore, using a simplified analysis, the potential throughput improvement of a multi-rate POWMAC protocol over a multi-rate 802.11 scheme. Let L_c be the total length (in bits) of the IEEE 802.11 RTS plus CTS packets. The total length of the POWMAC RTS, CTS, plus DTS packets is $\approx 1.68L_c$. Hence, the length of the AW slot is $1.68L_c + B$ (recall that B is the maximum backoff duration). Let L_d be the average data packet length. Let R_c and R_d be the transmission rates of control and data packets, respectively. Suppose that there are N feasible simultaneous in the same vicinity. The duration of time it takes to send N data packets according to POWMAC is $T_{POWMAC} = N\left(\frac{1.68L_c}{R_c} + B\right) + \frac{L_d}{R_d}$. The duration of time it takes to send the same N packet according to the IEEE 802.11 is $T_{802.11} = N\left(\frac{L_c}{R_c} + \frac{L_d}{R_d}\right)$ [13]. Computing T_{POWMAC} and $T_{802.11}$ in this way is quite optimistic since we are assuming that for POWMAC, all AW slots result in successful RTS/CTS/DTS exchanges, and that for the 802.11 scheme, an RTS/CTS exchange follows immediately the transmission of the previous data packet[14].

For POWMAC to outperform the 802.11 scheme, we must have $T_{POWMAC} < T_{802.11}$. With some manipulations, this condition can be written as $(0.68)\frac{R_d}{R_c} + \frac{BR_d}{L_c} < \frac{N-1}{N}\frac{L_d}{L_c}$. Clearly, the larger the ratio $\frac{R_d}{R_c}$, the lesser is the improvement of POWMAC over the 802.11. Furthermore, the greater the value of N, the more is the improvement of POWMAC over the 802.11. For example, according to the IEEE 802.11b specifications, the maximum value of R_d/R_c is 11 ($R_d = 11$ Mbps). Furthermore, L_d is typically in the order of tens of L_c. For example, for 2-KB data packets, $L_d/L_c \approx 59$. Using these values, it can be shown that as long as $N > 1$, POWMAC will outperform the 802.11 scheme. Even for N as small as 2, T_{POWMAC} is only 73% of $T_{802.11}$.

[13] For simplicity, the Ack packet overhead is not considered.

[14] The IEEE 802.11 scheme requires terminals to backoff after the end of a data transmission even if the channel is idle.

Data packet size	2 KB
Data rate	1 Mbps
SINR threshold	6 dB
Transmission range	750 meters
Carrier-sense range	1500 meters
Path loss factor	4
ξ_{max}	7 dB

Table 3.1: Parameters used in the simulations.

3.4 Performance Evaluation

3.4.1 Simulation Setup

We now evaluate the performance of the POWMAC protocol and contrast it with the IEEE 802.11 scheme. Note that we do not compare POWMAC to energy-oriented protocols (e.g., [44, 63, 67, 94]), since at best these protocols give comparable throughput to that of the 802.11 scheme. Furthermore, since POWMAC uses a single-channel, single-transceiver design, it is unfair to compare it with two-channels, two-transceivers based protocols (e.g., [80, 81, 126]). Our results are based on simulation experiments conducted using CSIM programs (CSIM is a C-based process-oriented discrete-event simulation package [3]). For simplicity, data packets are assumed to be of a fixed size. The routing overhead is ignored since the goal here is to evaluate the performance improvements due to the MAC protocol. Furthermore, because the interference margin is chosen so that the maximum transmission range under the POWMAC and 802.11 protocols is the same, it is safe to assume that both protocols achieve the same forward progress per hop. Consequently, we can focus on the one hop throughput, i.e., the packet destination is restricted to one hop from the source. The two-ray propagation model is used, and the capture model is similar to the one in [121]. Other parameters used in the simulations are given in Table 3.1. These parameters correspond to realistic hardware settings [4]. According to these parameters, each node has, on average, ten neighbors.

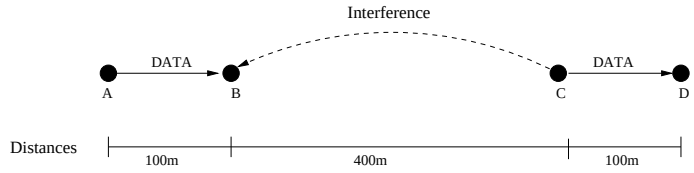


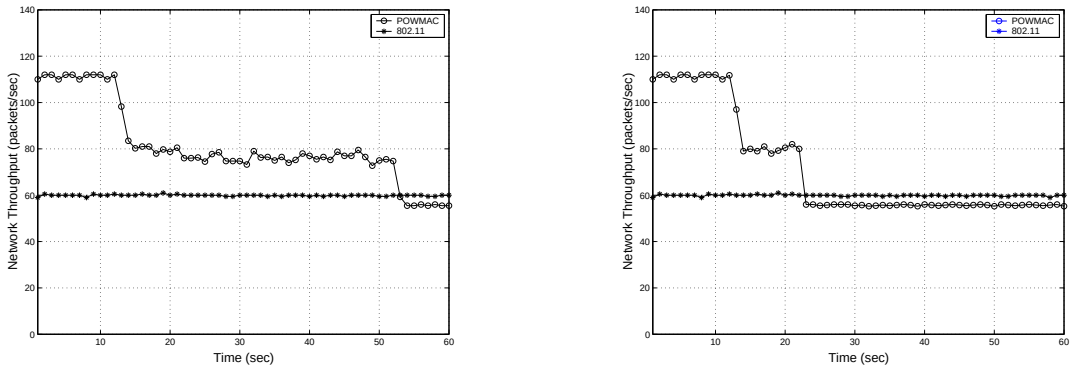
Figure 3.22: Toy topology where the two interfering transmissions $A \rightarrow B$ and $C \rightarrow D$ can proceed simultaneously if A's and C's transmission powers are appropriately chosen.

3.4.2 Macroscopic Results

We first simulate a set of basic scenarios for the purpose of highlighting the advantages and operational details of POWMAC. Consider the line topology in Figure 3.22. The distances between the terminals are also shown in the figure. Terminal A is transmitting to node B, and node C is transmitting to node D. Persistent load is used in this experiment, i.e., terminals A and C always have packets to send. The transmissions from A and C interfere with the data reception at D and B, respectively. However, the interference from A to D is much smaller than the one from C to B, and so in the following discussion, we focus on the latter one.

In the first scenario, node B starts moving in the direction of node C at speed of 5 m/s. Figure 3.23(a) depicts the throughput of the network as a function of time. According to the 802.11 scheme, only one transmission can proceed at a time since all terminals are within the carrier-sense range of each other. However, according to POWMAC, for the first 12 seconds, the two transmissions $A \rightarrow B$ and $C \rightarrow D$ can proceed simultaneously, resulting in about 84% improvement in network throughput. For the next 40 seconds, as node B gets closer to node C, the channel gain G_{BC} increases and so $P_{\text{MAP}}^{(C)}$ decreases until it becomes less than the one required by node D to achieve its SINR threshold. Therefore, once node A exchanges RTS/CTS/DTS with B, node C cannot transmit to D[15]. On the other hand, if node C exchanges RTS/CTS/DTS packets with D before A does that with B, then node A increases

[15]When C sends an RTS to D, D replies with a negative CTS since $P_{\text{MAP}}^{(C)}$ is less than $P_{\text{POWMAC}}^{(CD)}$ as computed by node D.



(a) Scenario 1: B is moving towards C.

(b) Scenario 2: C is moving towards B.

Figure 3.23: Performance of the POWMAC and the 802.11 protocols (line topologies).

its TP to overcome the interference induced from C at node B. Hence, the two transmissions $A \rightarrow B$ and $C \rightarrow D$ can proceed simultaneously. Roughly, half of the time A starts before C and half of the time C starts before A, so the throughput enhancement is about 34% during the period between 12 and 52 seconds. After 52 seconds, the interference at B due to C becomes larger than the one allowed by the planned loading, so either $A \rightarrow B$ or $C \rightarrow D$ can proceed, but not both. The small degradation in throughput after 52 seconds is attributed to the overhead of the AW when no simultaneous transmissions are taking place.

In the second scenario, terminal C moves in the direction of B at a speed of 5 m/s, while all other terminals are stationary. Figure 3.23(b) shows the throughput of the network as a function of time. The difference between this scenario and the previous one is that this time, not only is $P_{\text{MAP}}^{(C)}$ decreasing (as a result of G_{BC} increasing), but $P_{\text{POWMAC}}^{(CD)}$ is also increasing as a result of the decrease in G_{CD}. In the first 12 seconds, the two transmissions $A \rightarrow B$ and $C \rightarrow D$ can proceed simultaneously. Between 12 and 22 seconds, the throughput enhancement is 34% for the same reason given in the previous scenario. After that, only one transmission proceeds, and the throughput becomes comparable to that of the 802.11 scheme.

3.4.3 Random Grid Topologies

We now study the performance under more realistic network topologies. First, we consider a *random-grid* topology, where 25 mobile terminals are placed within a square area of length 1500 meters. The square is split into 25 smaller squares, one for each terminal. The location of a mobile terminal within the small square is selected randomly. For each generated packet, the destination terminal is selected randomly from the one-hop neighbors. Each terminal generates packets according to a Poisson process with rate λ (same for all terminals). The Random Waypoint model [20] is used for mobility, with a terminal speed that is uniformly distributed between 0 and 2 meters/sec.

The performance is demonstrated in Figure 3.24. Part (a) of the figure depicts the throughput versus λ. It can be shown that at high loads, POWMAC achieves about 50% increase in throughput over the IEEE 802.11 scheme. This increase is attributed to the increase in the number of simultaneous transmissions. Part (b) of Figure 3.24 depicts the energy consumption versus λ. This is the total energy used to *successfully transmit* a packet. It includes the energy used to transmit control packets and the lost energy in retransmitting data and control packets in case of collisions[16]. For all cases, POWMAC requires roughly the same energy required by the 802.11 scheme. These results are in line with the analysis in Section 3.3.3, where the interference margin was chosen so that both protocols consume the same energy per bit.

3.4.4 Clustered Topologies

Next, we study the performance of POWMAC under *clustered* topologies. In such topologies, a terminal communicates mostly with terminals within its own cluster, and rarely with neighboring clusters. These topologies are common in practice (e.g., a historical site where users of wireless devices move in groups). To generate a

[16]The *processing* power consumption in the transmitter and the receiver circuitry is the same for both POWMAC and the IEEE 802.11 scheme. Furthermore, this power depends on the wireless card used, unlike the TP consumption. Our energy model accounts only for the TP.

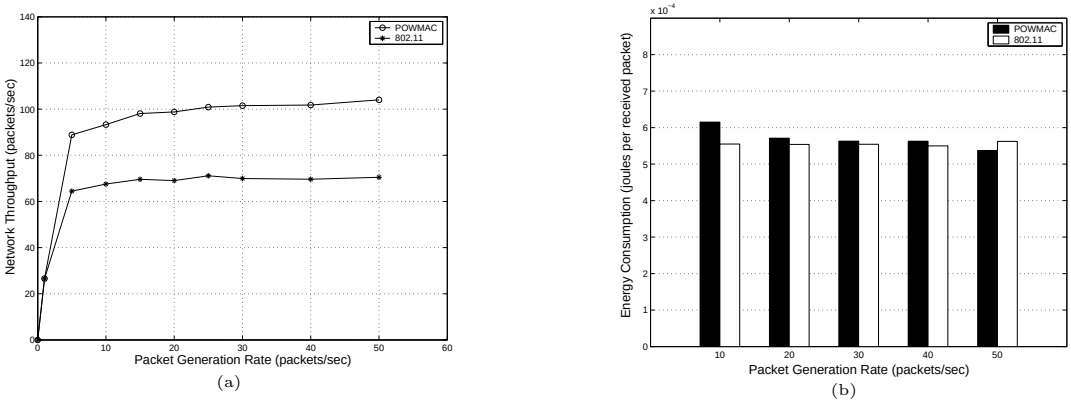


Figure 3.24: Performance of the POWMAC and 802.11 protocols as a function of λ (random-grid topologies).

clustered topology, we consider an area of dimensions 600×600 (in meters). Sixteen terminals are split into 4 equal groups, each occupying a 100×100 square in one of the corners of the complete area. For a given source terminal, the destination is selected from the same cluster with probability $1-p$ or from a different cluster with probability p. In each case, the selection from within the given cluster(s) is done randomly.

Part (a) of Figure 3.25 depicts the network throughput versus λ for $p = 0.25$ and $p = 0.5$. According to the 802.11 scheme, only one transmission can proceed at a time since all terminals are within the carrier-sense range of each other. Furthermore, its throughout performance is approximately the same regardless of the value of p. In other words, the 802.11 scheme does not benefit from the locality of the traffic. On the other hand, according to POWMAC, two to three transmissions can proceed simultaneously, resulting in a significant improvement in network throughput. Moreover, it is clear that POWMAC utilizes traffic locality to increase network throughput; its performance is better for smaller values of p.

Part (b) of the figure shows that POWMAC saves a significant amount of energy relative to the 802.11 scheme. Since a terminal communicates mostly with terminals within its own cluster, the destination terminal is within 100 meters of the source terminal, thus requiring much less TP than $P_{\max}$. This is the reason why the figure

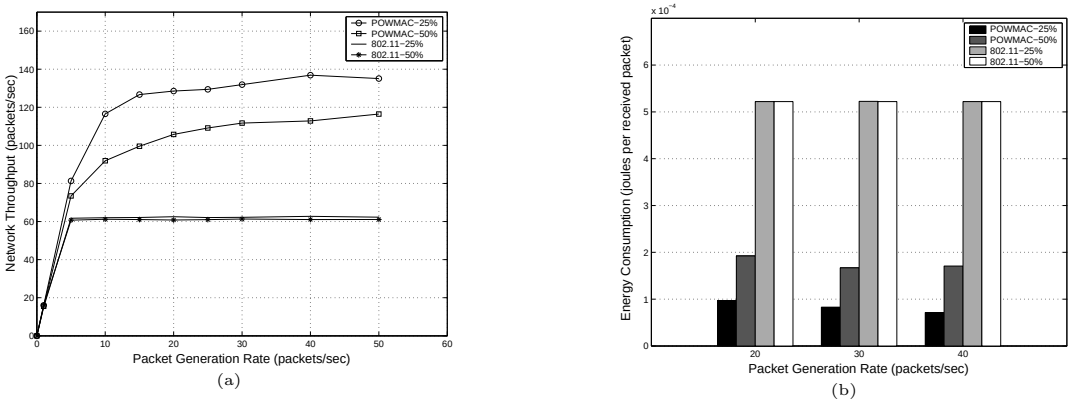


Figure 3.25: Performance of the POWMAC and 802.11 protocols as a function of λ (clustered topologies).

shows a huge advantage of POWMAC over the 802.11 scheme in terms of energy consumption. So, although POWMAC was designed with the goal of increasing throughput, significant energy may be saved as a consequence of reducing the TP. Understandably, energy saving increases as traffic becomes more localized (i.e., when $p = 0.25$).

Next, we show the strong parallelism that is achieved by POWMAC. To this end, we study the percentage of time during which N transmissions take place simultaneously in the *same* neighborhood. Figure 3.26 depicts this measure of performance. The 802.11 scheme allows for only one transmission in a neighborhood, and so for both cluster and grid topologies, $N = 1$ for all the time. In contrast, POWMAC allows for up to 5 and 7 simultaneous transmissions in the same neighborhood in the random-grid and clustered topologies, respectively.

Figure 3.27 shows the time evolution of the AW for a typical terminal. To produce this figure, we look at the AWS of a terminal that is located roughly in the center of the random-grid topology. The initial value of the AW was 4. The terminal changes its AWS every time a data packet has been received by monitoring the measured interference during that packet reception versus $\xi_{\max}$, as explained in Section 3.3.7. Over several runs, *the average size of the AW was found to be approximately 3.*

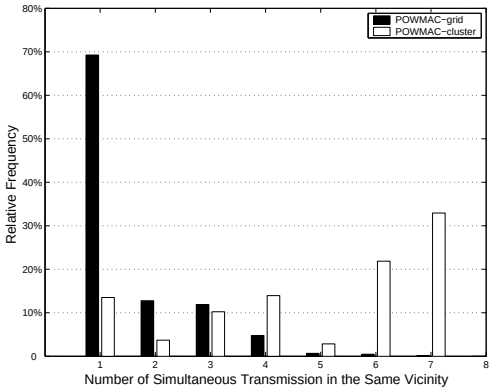


Figure 3.26: Relative frequency with which N transmissions take place simultaneously in the same neighborhood.

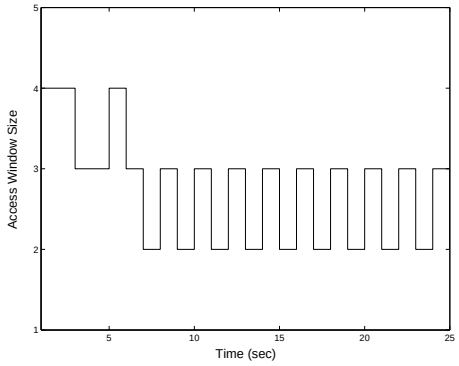


Figure 3.27: Time evolution of the AW for a typical terminal.

Table 3.2 shows the percentage of data collisions under different load conditions for both POWMAC and the IEEE 802.11 scheme. Because of its conservative design, the 802.11 scheme results in fewer collisions. The price, however, is loss in throughput. On the other hand, POWMAC takes an avant-garde approach of allowing concurrent interference-limited transmissions. Although POWMAC results in more collisions, it is able to significantly improve the overall network performance (i.e., throughput).

Next, recall that in Section 3.3.8, we pointed out that both POWMAC and the 802.11 scheme cannot completely eliminate collisions due to interference. The reason is that the interference range is typically larger than the transmission range.

λ (packets/sec)	802.11 (%)	POWMAC (%)
1	1	1.5
5	4.7	12.5
10	6.2	17.4
15	9.2	19.1
20	9.0	17.9

Table 3.2: Percentage of data-packet collisions as a function of λ.

ϱ	0.1	0.5	0.9	5.0
Throughput (packets/sec)	104	106	102	90

Table 3.3: Performance of POWMAC as a function of ϱ.

To address this limitation in the 802.11 scheme, the authors in [128] proposed to limit the communication range to a value that is below the maximum one. In other words, if the minimum power required to receive a packet at a terminal, say i, is P_{th}, then i responds with a CTS packet only if the reception power is at least γP_{th}, where $\gamma > 1$. While it does not solve the problem completely, this solution reduces the severity of it; the price being decreasing the transmission range and affecting the topological properties of the network. For similar reasons, POWMAC would also benefit from this solution. Figure 3.28 depicts the throughput of POWMAC and the 802.11 scheme versus the transmission range. The experiment was conducted for a random-grid topology with each terminal generating packets according to a Poisson process at a rate of 20 packets/sec. It is shown that the throughput performance of both POWMAC and the 802.11 scheme increases by up to 26% and 22% respectively, as the transmission range is decreased from 100% to 70% of the maximum range.

To study the impact of ϱ in (3.10) on the performance of POWMAC, we run simulations for the random-grid topology using different values of ϱ. The throughput is shown in Table 3.3. It is clear that the throughput does not vary much with ϱ, when ϱ is between 0.1 and 0.9. However, as ϱ increases to 3.0, the throughput decreases noticeably, since receivers' interference tolerances become too small for the receivers' neighbors to start their own transmissions.

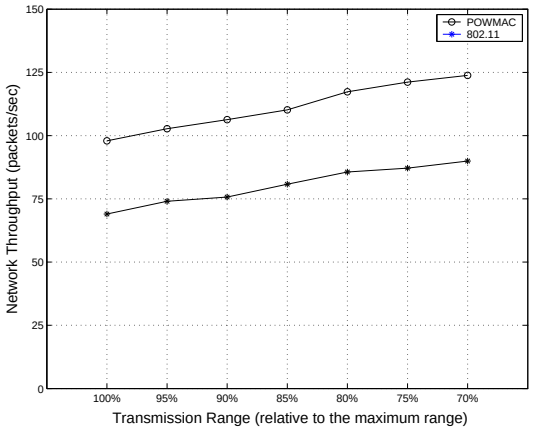


Figure 3.28: Performance of the POWMAC and the 802.11 protocols as a function of the transmission range.

Next, we study the impact of the data packet size on the performance. We run simulations for the random-grid topology, where each terminal generates packets according to a Poisson process with $\lambda = 20$ packets per second. The performance versus the packet size is shown in Figure 3.29. Part (a) of the figure shows that the throughput enhancement of the proposed protocol is lesser for shorter packets, which agrees with the analysis given in Section III-M. Part (b) of that figure shows that the energy consumption decreases as we increase data packets' sizes. This is again not surprising, since the fraction of energy consumed on control packet compared to data packets is smaller when data packets are larger.

Finally, we would like to compare the benefits of power control with rate control. But before doing that, we should mention that the simulations of POWMAC and the IEEE 802.11 do not account for the effect of spread spectrum; meaning that, there is no spreading gain at the physical layer [17]. Both POWMAC and the 802.11 schemes would benefit significantly from this spreading gain, but in different ways. POWMAC would allow for more simultaneous transmissions in the same neighborhood, since the interference is now scaled down by the processing gain, while the IEEE 802.11 scheme would be able to do rate control without using more spectrum,

[17]This is also the case with other widely used simulators such as ns-2 [5]

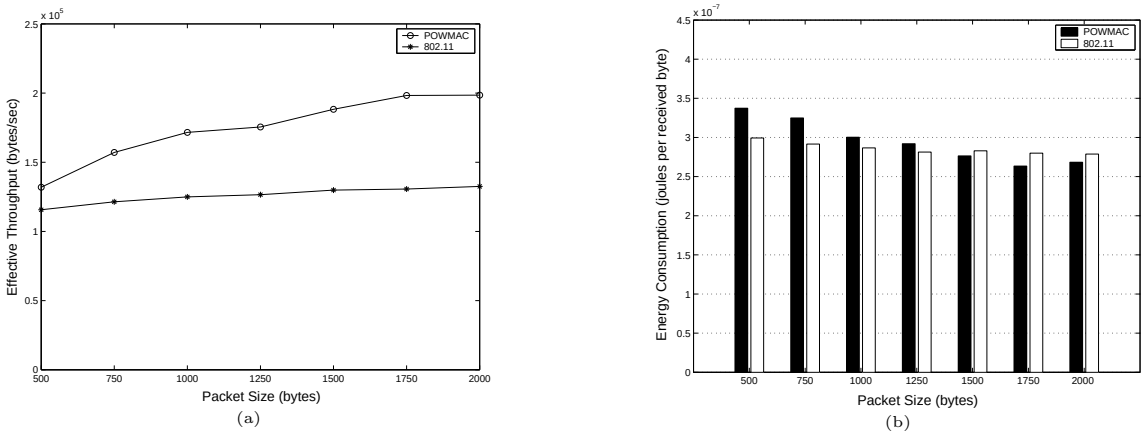


Figure 3.29: Performance of the POWMAC and 802.11 protocols as a function of the data packet size (random-grid topologies).

as the signal will be spread using a high chip rate.

However, since our model does not include spread spectrum, the only fair comparison is between POWMAC and a BPSK/QPSK two-rate 802.11 scheme. The reason is that QPSK allows for twice the rate as BPSK without requiring additional energy per bit (E_b/N_0), or more spectrum [92]. This is not the case for other modulation schemes. For example, 8-PSK requires more E_b/N_0 than BPSK for the same BER, while CCK as implemented in the 802.11 scheme requires more spectrum.

Motivated by above discussion, we have run simulations for a two-rate (1 Mbps and 2 Mbps) 802.11 scheme. Figure 3.30 depicts the throughput of the two-rate scheme and that of POWMAC versus the packet generation rate. As shown in that figure, both rate and power control achieve quite similar throughput enhancement. This is an interesting result, as it motivates further research to identify other advantages of both approaches.

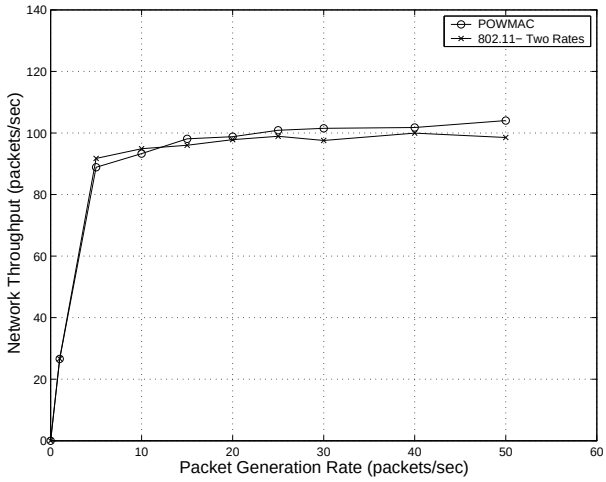


Figure 3.30: Performance of POWMAC and a rate-controlled 802.11 scheme (with two rates).

CHAPTER 4

PERFORMANCE ENHANCEMENT OF ADAPTIVE ORTHOGONAL MODULATION IN WIRELESS CDMA SYSTEMS

4.1 Introduction

The need for spectrally efficient systems has motivated the development of adaptive transmission techniques, several of which are in the process of being standardized. These techniques adapt users' parameters according to the time-varying channel conditions, interference levels, rate requirements, bit error rate (BER) needs, and energy constraints [83].

In *narrow-band* (i.e., non-spread spectrum) systems, adaptation includes varying the transmission power [49], modulation order [43], symbol rate [22], coding rate [118], or any combination of these parameters [10, 42, 41, 79]. In particular, it is well known that adaptive modulation is a promising technique for increasing the user data rate in narrow-band systems. This was demonstrated in [43] for the single-user case, where it was shown that adaptive modulation can provide up to 10 dB gain over a fixed-rate system that uses only power control. In [96], the authors studied the multiuser case and showed that even without power control, adaptive modulation has a significant throughput advantage over fixed-rate power control schemes. Much of the work on adaptive modulation in narrow-band systems (e.g., [11, 43, 64, 65, 123]) has been motivated by recent advances in designing low-complexity adaptive modulation circuitry and channel estimation techniques [43].

In the context of (wide-band) CDMA networks, power control has traditionally been the single most important adaptation parameter [39], and has been thoroughly studied (see [110] and the references therein). Recent efforts on adaptation in CDMA networks have also focused on adapting the transmission rate using multiple codes [56, 106], parallel combinatory spread spectrum [134], multiple chip-rate [127],

adaptive modulation and coding (AMC) [1, 15, 50], and "classical" variable processing gain (VPG) techniques [29, 30, 57, 62, 66, 71, 86, 105, 119, 129] in which both the transmission power and data rate are adapted, but the modulation and coding are kept fixed.

For CDMA systems that require coherent reception, a pilot signal must usually be transmitted for each user. This is the case, for example, in WCDMA systems [1], where a high-rate coherent two-dimensional modulation[1] such as 16QAM [1, 50] is used. Alternatively, to reduce the implementation complexity associated with coherent reception (e.g., recovering the pilot signals from users) and to potentially improve energy efficiency (a pilot signal consumes a considerable amount of the mobile user's energy), noncoherent reception can be used [58]. M-ary orthogonal modulation (OM) is a spectrally-efficient modulation technique that is well suited for this application [39]. Although differential phase shift keying (DPSK) can also be used for noncoherent reception, it has been shown that OM outperforms DPSK for $M > 8$ in additive white Gaussian noise (AWGN) and in Rayleigh fading channels [92]. OM has been used successfully in the uplink of IS-95 and is also part of the radio configurations of the cdma2000 standard [56].

This chapter focuses on CDMA systems for which coherent reception is not possible and where OM is used (e.g., uplink IS-95). For such systems, classical (i.e., fixed OM order) VPG has been the focus of research because of its performance benefits, flexibility, and practicality (e.g., low peak-to-mean envelope power, fixed chip rate, etc. [57]). The extensive work on VPG has clearly quantified the performance[2] advantages of combined rate/power control over power control alone (e.g., see [105, 57]). However, to the best of our knowledge, adapting the modulation order for variable-rate OM-based systems remains an unexplored area of research, and one for which joint rate/power control has not yet been investigated. Our first contribu-

[1]By two-dimensional modulation, we mean modulation schemes for which the modulation symbol can be represented by a 2-dimensional vector, i.e., by a point in the 2-dimensional signal space (or constellation).

[2]Throughout the chapter, the term "performance" is used to refer to network throughput and/or per-bit energy consumption.

tion (Section 4.2) is to show that when OM is used, the performance of variable-rate CDMA networks can be improved by using higher OM orders at lower data rates. We then use these results to show that, in the single link case, variable-rate systems with adaptive orthogonal modulation (AOM) significantly outperforms VPG systems with a fixed OM order[3]. Thus, similar to adaptive modulation in narrowband systems, AOM in CDMA systems is shown to be a promising technique for increasing the user data rate. Note that the processing gain and transmission power are varied in both AOM and VPG. However, in AOM the OM order is also varied depending on the data rate, whereas VPG uses the *same* OM order for all data rates.

The main goal of our study is to investigate the theoretical performance limits of *joint* rate/power control for AOM-based CDMA networks and to gain insights into the technique itself. We consider both point-to-point (PTP) as well as multipoint-to-point (MultiPTP) networks (see Figure 4.1). PTP networks is the more general communication paradigm. It can represent a completely distributed mobile ad hoc network, or a microcellular network in which mobile-base station pairs compete for the same frequency spectrum. In MultiPTP networks, multiple nodes transmit to one node, as in the case of a cluster-based ad hoc or sensor network [88] or in the case of the uplink of a single cell in a CDMA-based cellular network (e.g., IS-95 [92]). With very few exceptions, previous work has mainly considered MultiPTP networks.

To jointly optimize the powers and rates, we consider two *throughput-related* objective functions: (1) minimizing the maximum service time, and (2) maximizing the sum of users' transmission rates. Both functions are optimized subject to constraints on the maximum transmission power, on the minimum and maximum transmission rates, and on the BER. The first function is novel in our context and has not received much attention; previous research has primarily focused on the second objective function. However, as we argue in Section 4.3, there are important *practical* advantages of the first objective function.

[3]For brevity, we use the acronym AOM to refer to a variable-rate system with adaptive OM, while the acronym VPG refers to a variable-rate system with a fixed OM order.

(a) PTP Networks.

(b) MultiPTP Networks.

Figure 4.1: Network topologies considered in the chapter.

We obtain the optimum solution to the problem of minimizing the maximum service time in both PTP and MultiPTP networks by formulating the problem as a generalized geometric program (GGP) [19]. We then transform this GGP into a geometric program (GP), which itself can be transformed into a nonlinear convex program. The advantage of these transformations is that a convex program has a global optimum that can be found very efficiently [19]. Furthermore, in the case of MultiPTP networks, we derive a simple expression for computing the optimal powers and rates that minimize the maximum service time. Our solutions are computationally efficient. They can also be used to determine the feasibility of a set of rate and BER requirements under certain constraints, thus, allowing for the use of admission control policies.

Although the second objective function (i.e., maximizing the sum of rates) has the advantage of being in the exact form of throughput, it has the limitation of having several local maxima. As a result, there are no computationally efficient algorithms to solve this problem[4]. Hence, for PTP networks, although we do not know the optimal rate/power solution for VPG and AOM, we provide some numerical results that demonstrate the performance advantages of AOM over VPG. For MultiPTP networks, we start from theorems proved in [57], and we analytically de-

[4]This may be one reason why previous studies that pursued an algorithmic approach to this problem considered other objectives, such as minimizing the power or even *minimizing* the sum of rates [66].

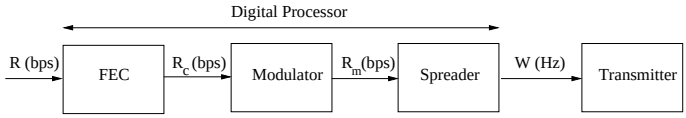


Figure 4.2: Simplified block diagram of the transmitter circuit.

rive a simple procedure for maximizing the sum of rates for VPG systems. Then, we show how this solution, which is optimal in VPG systems, can be used *heuristically* in AOM MultiPTP networks. Using these results, we derive a lower bound on the achievable gain of AOM over VPG schemes. As shown in Section 4.4, this gain is substantial.

Note that our goal in this chapter is *not* to promote OM as a modulation scheme, but rather to advocate *adapting* the order of OM for CDMA systems that already use OM (e.g., the uplink of IS-95). The rest of the chapter is organized as follows. In the next section, we take a system-level approach to the analysis of AOM in CDMA multimedia networks and show its performance advantages over VPG. In section 4.3, we present the objective functions, formulate the optimization problems, and present their solutions. The performance of AOM is presented and contrasted with VPG in Section 4.4.

4.2 Orthogonal Modulation in CDMA Networks

4.2.1 Motivation for Higher Orthogonal Modulation Orders

The main goal of this section is to show that for any data rate, increasing the OM order improves the performance of a CDMA system. The maximum OM order that can be used, however, is constrained by the chip rate. We first start with a system-level analysis of CDMA systems. The benefits of a higher OM order is then established using this analysis and through an analogy between OM and FEC. The message we will try to convey is that, in CDMA systems, it is always advantageous to use an FEC or an OM order that reduces the bit-energy-to-noise spectral density ratio (E_b/N_0) required for a given BER.

The transmitter circuit of the system under study is shown in Figure 4.2. It consists of (digital) FEC encoder, modulator, direct-sequence spreader, and (analog) amplifier and transmitter [39]. Consider packet reception for link i. Let I be the set of active links in the network, $P_t^{(i)}$ be the transmission power of link i, and h_{ji} be the channel gain between the receiver of link i and the transmitter of link j. Then the signal-to-noise (and interference) ratio at i is:

$$\mathrm{SNR}^{(i)} = \frac{h_{ii}P_t^{(i)}}{\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\mathrm{thermal}}} \tag{4.1}$$

where P_{thermal} is the thermal noise, which is modeled as a white Gaussian noise process. The interference from other users is also assumed to be Gaussian. This assumption has been shown to produce throughput results that are reasonably accurate [98]. For reliable communication, a more relevant metric than $\mathrm{SNR}^{(i)}$ is the effective bit energy-to-noise spectral density ratio at the detector, denoted by $\mu^{(i)}$ and given by [39]:

$$\mu^{(i)} \stackrel{\mathrm{def}}{=} \frac{E_b}{N_0} = \frac{W}{R^{(i)}} \frac{h_{ii}P_t^{(i)}}{\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\mathrm{thermal}}} \tag{4.2}$$

where W is the Fourier bandwidth occupied by the signal (i.e., chip rate) and $R^{(i)}$ is the data rate of i's intended signal. Let μ_{req} be the required $\mu^{(i)}$ for a certain BER. Then, the maximum achievable data rate at i is:

$$R^{(i)} = W\frac{\mathrm{SNR}^{(i)}}{\mu_{\mathrm{req}}}. \tag{4.3}$$

Both (4.2) and (4.3), which hold for any CDMA system, do not explicitly indicate the effects of FEC and modulation on the achievable data rate. However, these effect appear indirectly through the value of μ_{req}. For example, the stronger the FEC code (i.e., the lower the code rate), the lesser is μ_{req} and the higher is the achievable data rate. This analysis is inline with the findings of Viterbi [116], in which he showed that the *jamming margin* is actually increased by coding; the idea is that with coding, μ_{req} is lower, and so more interference is allowed for the same rate (i.e.,

$\text{SNR}^{(i)}$ in (4.3) can be decreased). In other words, for CDMA systems it is always preferable to use schemes that enable operation at a lower μ_{req}.

In the case of M-ary OM, the modulator takes $k = \log_2 M$ FEC-coded bits and maps them into one of the M Walsh (or Hadamard) orthogonal sequences [92] of length M bits. So the resulting *modulated bit rate* R_m is equal to $R_c\, M/k$, where R_c is the coded bit rate (see Figure 4.2). At the receiver, the signal is first despread and then noncoherently detected, generating k soft output bits for each transmitted Walsh symbol, which are fed to the Viterbi decoder (see [114] for further details). A tight upper bound on the probability of bit error in OM is given by [92]:

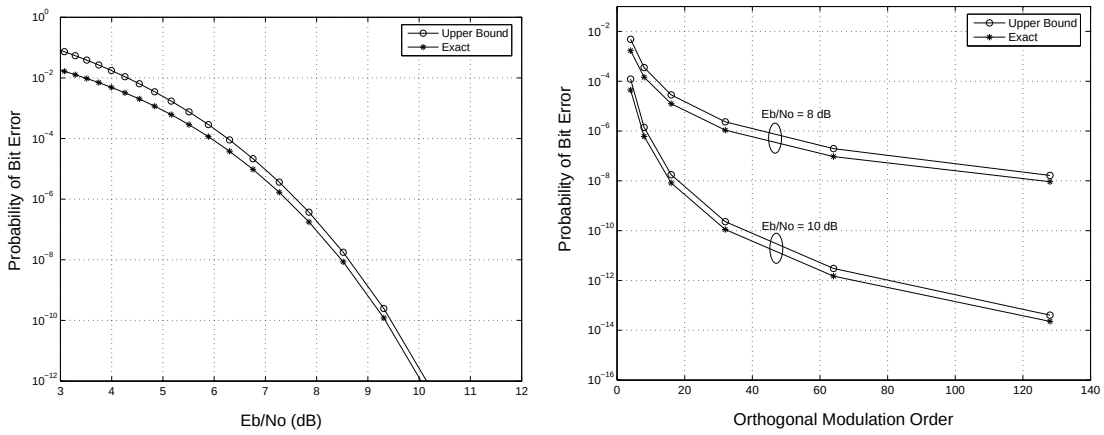
$$P_b < \frac{1}{2} e^{-k(\mu^{(i)} - 2\ln 2)/2}. \tag{4.4}$$

It is clear from (4.4) that the higher the value of k, the lower is the BER. *Therefore, the higher the OM order M, the better is the BER performance for the same E_b/N_0 value.* OM in this sense works as an FEC code; the higher the value of M, the lower is the modulation rate k/M, but the better is the BER performance. Note that the higher the OM order, the higher is R_m; however, this has no impact on the system bandwidth as long as $R_m \leq W$, since the signal is spread by a high-rate CDMA code.

4.2.2 Performance Advantages of Adaptive Orthogonal Modulation

In the previous section, we showed that increasing the OM order is beneficial for the performance of a CDMA network. However, the higher the user data rate R, the lower must be the maximum allowable M to ensure that $R_m \leq W$. Thus, in AOM, M must be adapted according to R. Our goal in this section is to quantify the performance gains of adapting M according to R. To do this, we derive the relationship between the user's SNR and the achievable data rate for AOM and for non-adaptive OM (i.e., VPG).

First, we claim that it is sufficiently accurate to use (4.2) and the upper bound in (4.4) to analyze OM in CDMA systems. To substantiate our claim, we compare



(a) (b)

Figure 4.3: Probability of bit error in an OM-based CDMA system.

the performance obtained from these two simple equations with the results reported in [70], which were obtained using rigorous analysis. We simulate the same setup of [70]: a MultiPTP network that uses 64-ary OM with equal received powers at the common receiver. The number of transmitters is varied to obtain different E_b/N_0. Part (a) of Figure 4.3 shows the probability of bit error versus E_b/N_0. The "exact" plot is the same one that was obtained in [70], while the upper-bound curve is the one obtained using (4.2) and (4.4). This figure demonstrates that the bound is sufficiently tight for all practical purposes. To verify the tightness of the bound for other values of M, we show in Part (b) of Figure 4.3 the probability of bit error versus M for $E_b/N_0 = 8$ dB and $E_b/N_0 = 10$ dB. As can be seen, the bound is tight, and hence will be used in our subsequent analysis.

Next, we use (4.2) and (4.4) to derive the relationship between the user's SNR and the achievable rate with and without adapting M. From this relationship, we demonstrate the performance advantages of AOM over VPG for the single-link case. Without loss of generality, we assume that the system under study does not use any FEC (i.e., $R_c = R$). VPG uses the *same* modulation order M for all data rates. This M is chosen such that for a given R, $R_m \leq Z \leq W$, where Z is a threshold that is often determined by regulatory laws. For example, the Federal Commission

Commission (FCC) calls for at least a ratio of 10 (i.e., 10 dB) of spreading rate to modulation bit rate in the 2.4 GHz ISM band [12], so in this case $Z = W/10$. Accordingly, the modulation order for VPG is decided based on W, Z, and the maximum desired data rate ($R_{\max}$). If $Z = W$ and $R_{\max} = W/2$, then the (fixed) modulation order $M = 2$. If the required BER is 10^{-6}, then for this VPG system, μ_{req} is about 14.8, and so using (4.3), the required SNR at $R_{\max}$ is 7.4. Note that whereas μ_{req} is fixed, the required SNR is a function of R.

AOM, on the other hand, uses a variable M that depends on R. The higher the value of M, the smaller is the value of μ_{req}, but also the higher is R_m. For $Z = W$ and $R_{\max} = W/2$, the value of M at $R_{\max}$ cannot exceed 2 (to ensure that $R_m \leq Z$), implying that there is no performance advantage of AOM over VPG at $R_{\max}$. However, for $R < R_{\max}$, AOM uses a higher value for M, enabling operation at a lower μ_{req}, or equivalently, resulting in a higher data rate (see (4.3)). For each data rate R, the corresponding value of M is the largest value such that R_m, which in the absence of FEC is equal to $R\,M/k$, does not exceed Z. Assuming M is continuous (more on this assumption shortly), R can be expressed as:

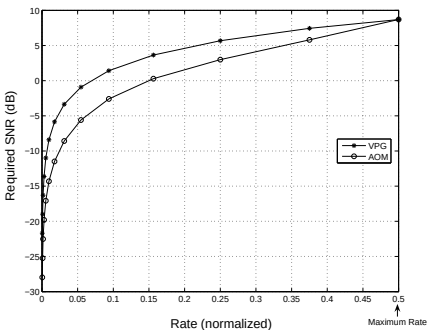
$$R = Z\ k\ 2^{-k}. \tag{4.5}$$

For a given target P_b, we use (4.4) as an equality, replace $\mu^{(i)}$ with μ_{req}, and derive μ_{req} as a function of k. This function along with (4.5) is used to approximate μ_{req} as a function of R, say $g(R)$. The approximation can be done by simple curve fitting. Finally, using $\mu_{\text{req}} = g(R)$ and (4.3), one can express the required SNR as a function of R:

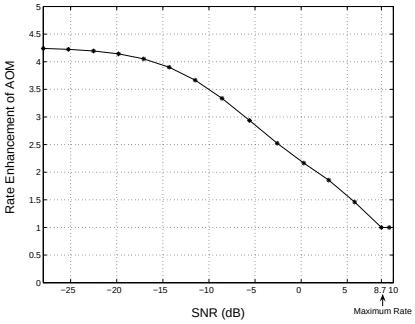
$$\text{SNR}_{\text{req}} = \frac{R}{W} g(R) \stackrel{\text{def}}{=} \frac{f(R)}{W}. \tag{4.6}$$

In the case of AOM, $f(R)$ can be well-approximated (less than 1% fitting error) by the posynomial function[5] aR_i^b, for some real-valued coefficients $a > 0$ and $b > 1$. On the other hand, in the case of VPG, $g(R)$ is a constant that is equal to μ_{req} (e.g., $g(R) = 14.8$ for $M = 2$), and therefore, SNR_{req} is simply a linear function of R. This linearity between R and SNR_{req} has been the underlying assumption in all

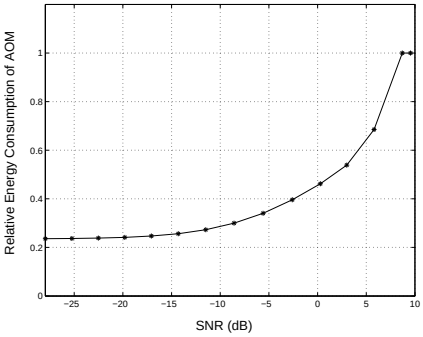
[5]The definition of a posynomial can be found in Appendix A.



(a) Required SNR versus data rate.



(b) Rate enhancement of AOM over VPG.



(c) Energy consumption of AOM relative to VPG.

Figure 4.4: Performance of AOM and VPG for a single link.

previous adaptive rate/power control schemes for OM-based CDMA networks. We now know that this assumption does not hold for AOM.

Using the relationships between R and $\mathrm{SNR}_{\mathrm{req}}$, we are now in a position to compare the performance of AOM with VPG for the single-link case. Figure 4.4 demonstrates several performance metrics obtained using $Z = W$ and $R_{\max} = W/2$. Part (a) of the figure depicts $\mathrm{SNR}_{\mathrm{req}}$ versus the normalized rate R/W. It is clear that for all $R < R_{\max}$, AOM requires a significantly less SNR than VPG to achieve a certain data rate. Such an improvement essentially reflects a *power gain*. Equivalently, AOM achieves a much higher rate than VPG for the same $\mathrm{SNR}_{\mathrm{req}}$ (i.e., rate gain). Part (b) of the figure shows the relative rate enhancement of AOM over VPG versus the SNR. It is shown that the rate advantage of AOM over VPG increases as the SNR decreases, and is very significant in the low SNR regime. Note that when

SNR $\geq$ 8.7 dB, the link operates at $R_{\max}$, and AOM uses the same modulation order as VPG, i.e., there is no rate improvement. Part (c) of the figure shows the energy-per-data bit (E_b) consumption of AOM relative to that of VPG versus the SNR. E_b is defined as the transmission power divided by R. The figure shows that AOM consumes much less E_b than VPG in the low SNR regime. The E_b consumption of AOM increases as the SNR increases until the maximum rate is reached, at which AOM consumes the same E_b as VPG.

In the above discussion, we permitted the modulation order M to take any real positive value; however, in real life, M is restricted to a finite set [6]. Nonetheless, we evaluate the potential gains without this additional constraint to serve as an upper bound on the performance of AOM in practice.

4.3 Joint Rate/Power Optimization for AOM Systems

The analysis presented in the previous section focused on the single-user case. For a network of users, increasing one user's power increases that user's SNR, and consequently its rate. However, this comes at the expense of the SNR for other users, whose data rates must now be reduced to combat the added interference. Determining the best powers and rates that optimize a given objective function (e.g., network throughput) is not straightforward. The goal of this section is to define objective functions and derive policies that optimize them for the case of a network of users (i.e., multiuser case).

We study two throughput-oriented objective functions: (1) minimizing the maximum service time, and (2) maximizing the sum of users' transmissions rates. The two functions differ in two aspects: the time scale at which rate adaption is carried out and the required hardware.

[6] The burden of demodulation for high values of M can be alleviated by using the Fast Walsh Transform method [9], which requires only $M\log_2 M$ real additions and subtractions.

4.3.1 Minimizing the Maximum Service Time

Let L_i be the load (in bits) to be transmitted over link i, $i \in I$, where I is the set of active links in the network. Recall that R_i is the data rate (in bits/sec) for link i. The service time for link i, denoted by S_i, is L_i/R_i. A scheme that minimizes the maximum service time $S_{\max} = \max\{S_i,\ i \in I\}$ has the advantage of being easy to integrate in many current wireless network standards. For example, the access point (AP) of an IEEE 802.11 WLAN (or the Piconet controller of an IEEE 802.15.3 WPAN) can utilize its polling medium access mechanism to measure the channel gains between the AP and each mobile node, and to probe nodes about their loads. Using channel gains and load values, the AP can compute the optimum powers and rates that minimize $S_{\max}$. A scheme that minimizes $S_{\max}$ does not require users to receive any feedback from the AP while transmitting, i.e., only one transceiver is required at a node. Furthermore, rate adaptation is carried out on a per-packet basis (i.e., the whole packet is transmitted at one rate), which is practical for current wireless networks standards [83].

Given the channel gains and the loads $L_i\ \forall i \in I$, the goal is to find the transmission powers and rates (i.e., $P_t^{(i)}$ and R_i, $\forall i \in I$) so as to minimize $S_{\max}$. Formally, this problem is stated as follows:

$$\begin{cases} \underset{\{R_i, P_t^{(i)},\ i \in I\}}{\text{minimize}} \left\{ \max_{i \in I} \frac{L_i}{R_i} \right\} & \\ \text{subject to:} & \\ \frac{h_{ii} P_t^{(i)}}{\sum\limits_{j \in I - \{i\}} h_{ji} P_t^{(j)} + P_{\text{thermal}}} \geq \frac{f(R_i)}{W}, & \forall i \in I \\ 0 \leq P_t^{(i)} \leq P_{\max}, & \forall i \in I \\ R_{\min} \leq R_i \leq R_{\max}, & \forall i \in I \end{cases} \tag{4.7}$$

The first constraint reflects the BER requirement of link i, since it mandates that i's SNR be greater than or equal to $\frac{f(R_i)}{W} = \text{SNR}_{\text{req}}$ (see (4.6)). $\frac{f(R_i)}{W}$ is equal to $\frac{R_i}{W}\mu_{\text{req}}$ for VPG and is approximated by $a(R_i/Z)^b(Z/W)$ for AOM, where a and

b are two constants whose values are obtained from the fitting of $f(R)$. In our simulations, $a \approx 9.8$ and $b \approx 1.2$, with less than 1% fitting error. Although the formulation in (4.7) assumes the same minimum rate, maximum rate, and maximum power constraints for all nodes, this can be easily extended to handle the case of node-specific constraints. Note that this formulation is applicable to both PTP and MultiPTP networks.

Proposition 1 *The optimization problem in (4.7) is a generalized geometric program (GGP). This GGP can be transformed into a geometric program (GP), which itself can be transformed into a nonlinear convex program*[7]*.*

Proof: With simple algebraic manipulations, (4.7) can be expressed as:

$$\begin{cases} \underset{\{R_i, P_t^{(i)},\ i\in I\}}{\text{minimize}} \left\{ \max_{i\in I} \{L_i R_i^{-1}\} \right\} \\ \text{subject to:} \\ \left[\sum_{j\in I-\{i\}} h_{ji} P_t^{(j)} + P_{\text{thermal}} \right] \left[h_{ii} P_t^{(i)} \right]^{-1} \frac{f(R_i)}{W} \leq 1 \\ P_t^{(i)} P_{\max}^{-1} \leq 1 \\ R_i R_{\max}^{-1} \leq 1 \\ R_i^{-1} R_{\min} \leq 1 \end{cases} \tag{4.8}$$

where the constraints in (4.8) are to be satisfied for all $i \in I$. If $f(R)$ is a posynomial (see Appendix A), which is the case for both VPG and AOM, (4.8) is a GGP. In its current form, this GPP cannot be solve optimally and efficiently. Therefore, we make two transformations. The first one transforms the above GGP into a GP. To this end, we introduce a new auxiliary variable t such that:

$$t \geq \frac{L_i}{R_i}, \quad \forall i \in I. \tag{4.9}$$

[7]See Appendix A for a brief description of GGP and GP.

With the introduction of t, (4.8) becomes:

$$\begin{cases} \underset{\{t,R_i,P_t^{(i)},\ i\in I\}}{\text{minimize}} \quad t \\ \text{subject to:} \\ L_i R_i^{-1} t^{-1} \leq 1 \\ \left[\sum_{j\in I-\{i\}} h_{ji} P_t^{(j)} + P_{\text{thermal}}\right] \left[h_{ii} P_t^{(i)}\right]^{-1} \frac{f(R_i)}{W} \leq 1 \\ P_t^{(i)} P_{\max}^{-1} \leq 1 \\ R_i R_{\max}^{-1} \leq 1 \\ R_i^{-1} R_{\min} \leq 1 \end{cases} \tag{4.10}$$

It is obvious that (4.8) and (4.10) are equivalent forms, meaning that the powers and rates that minimize t also minimize the objective function in (4.8). Formulation (4.10) is an example of a GP, which can be easily transformed into a nonlinear convex program using a logarithmic change of variables [19]. Formally, let $z \stackrel{\text{def}}{=} \log t$, $x_i \stackrel{\text{def}}{=} \log P_t^{(i)}$, and $y_i \stackrel{\text{def}}{=} \log R_i\ \forall i \in I$ (so that $t = e^z$, $P_t^{(i)} = e^{x_i}$, and $R_i = e^{y_i}$). Instead of minimizing the objective function t, we now minimize $\log t$. Also, each constraint of the form $f \leq 1$ is changed to $\log f < 0$. This results in the following (equivalent) optimization problem:

$$\begin{cases} \underset{\{z,x_i,y_i,\ i\in I\}}{\text{minimize}} \quad z \\ \text{subject to:} \\ \log\ L_i e^{-y_i} e^{-z} \leq 0 \\ \log\ \left[\sum_{j\in I-\{i\}} h_{ji} e^{x_j} + P_{\text{thermal}}\right] h_{ii}^{-1} e^{-x_i} \frac{f(e^{y_i})}{W} \leq 0 \\ \log\ e^{x_i} P_{\max}^{-1} \leq 0 \\ \log\ e^{y_i} R_{\max}^{-1} \leq 0 \\ \log\ e^{-y_i} R_{\min} \leq 0 \end{cases} \tag{4.11}$$

At first, the above formulation may look more complicated than (4.10). However,

unlike (4.10), (4.11) is a *convex* optimization problem that can be solved efficiently (see [19] for more details). Once (4.11) is solved for x_i and y_i, $\forall i \in I$, the optimal power and rate allocation is simply given by $P_t^{(i)} = e^{x_i}$ and $R_i = e^{y_i}$ $\forall i \in I$.

Proposition 1 applies to both PTP and MultiPTP networks, and also for VPG as well as AOM schemes. In the case of MultiPTP networks, the structure of the problem can be further simplified to allow for even a faster computation of the optimal solution. The following proposition enables the subsequent derivation of this solution.

Proposition 2 *The powers and rates that optimize (4.7) are such that the first constraint is satisfied with equality.*

Proof: See Appendix B.

In MultiPTP networks, the receiver is common to all transmitters, and so the channel gains h_{ji} and h_{ii} can be simply written as h_j and h_i, respectively. Hence, utilizing Proposition 2, the optimal power and rate allocation in the case of MultiPTP networks must satisfy the following set of linear equations:

$$\frac{h_i P_t^{(i)}}{\sum\limits_{j \in I - \{i\}} h_j P_t^{(j)} + P_{\text{thermal}}} = \frac{f(R_i)}{W}, \qquad \forall i \in I. \tag{4.12}$$

Using the same derivation methodology as in [105], (4.12) can be reduced to:

$$\sum_{j \in I} \frac{1}{\left(\frac{W}{f(R_j)} + 1\right)} = 1 - \frac{P_{\text{thermal}}}{P_t^{(i)} h_i \left(\frac{W}{f(R_i)} + 1\right)}, \forall i \in I. \tag{4.13}$$

By imposing the constraint $P_t^{(i)} < P_{\max}$ and noting that (4.13) is valid $\forall i \in I$, the following inequality can be obtained:

$$\sum_{j \in I} \frac{1}{\left(\frac{W}{f(R_j)} + 1\right)} \leq 1 - \frac{P_{\text{thermal}}}{\min\limits_{i \in I} \left[P_{\max} h_i \left(\frac{W}{f(R_i)} + 1\right)\right]}. \tag{4.14}$$

This equation determines the feasibility of a set of rates, BER requirements, and maximum power constraints. Next, we use (4.14) to derive the optimal solution for (4.7). Consider the following proposition:

Proposition 3 *The powers and rates that optimize (4.7) are such that* $\frac{L_i}{R_i} = \frac{L_j}{R_j}$ $\forall\ i, j \in I$.

Proof: See Appendix C.

This proposition says that, at the optimal solution to (4.7), all users have the same service time (S). Hence, $R_i = L_i/S\ \forall i \in I$. Accordingly, (4.14) can be written as:

$$\sum_{j \in I} \frac{1}{\left(\frac{W}{f(L_j/S)} + 1\right)} \leq 1 - \frac{P_{\text{thermal}}}{\min\limits_{i \in I}\left[P_{\max} h_i \left(\frac{W}{f(L_i/S)} + 1\right)\right]}. \tag{4.15}$$

The only unknown in this equation is S, and so it can be easily solved for the minimum S. Note that a unique solution always exist, since the left-hand side (LHS) of (4.15) is 0 at $S = \infty$, and it increases as S decreases, while the RHS is 1 at $S = \infty$, and it decreases as S decreases. In Section 4.4, we use (4.15) to show the significant performance improvement of AOM over VPG.

4.3.2 Maximizing the Sum of Users Rates

The goal of this objective function is to maximize network throughput, subject to constraints on the BER, the maximum transmission power, and the minimum and maximum transmission rates. This function, which has been the focus of much previous research, requires fast rate adaptation; for the network to operate at the optimal point, whenever a user completes the transmission of a packet, all other transmitters must update their rates in the midst of transmitting their packets. This means that users must use intra-packet rate adaptation (i.e., different portions of the same packet must be transmitted at different rates). Furthermore, maximizing the sum of rates requires users to be able to receive feedback about their new rates while transmitting, which may necessitate the use of a multiple-channel multiple-transceiver architecture. Note that the minimum-rate constraint, which has been overlooked in most previous studies, is crucial for multimedia networks; without this constraint, some users may never be allowed to transmit, particularly if they

experience a "bad" channel relative to other users (i.e., their channel gains are relatively small).

The power/rate optimization problem for both AOM and VPG can be formulated as follows:

$$\begin{cases} \underset{\{R_i, P_t^{(i)},\ i\in I\}}{\text{maximize}} \quad \sum_{i\in I} R_i \\ \text{subject to:} \\ \frac{h_{ii}P_t^{(i)}}{\sum_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\text{thermal}}} \geq \frac{f(R_i)}{W}, \quad \forall i \in I \\ 0 \leq P_t^{(i)} \leq P_{\max}, \quad \forall i \in I \\ R_{\min} \leq R_i \leq R_{\max}, \quad \forall i \in I. \end{cases} \tag{4.16}$$

Unfortunately, this objective function cannot be transformed into the minimization of a posynomial as was done in the previous section. So it is not possible to formulate this problem as a GGP, a GP, or a nonlinear convex program. In fact, the problem exhibits an unknown number of local maxima, and there are no efficient algorithms to solve it optimally for the general case (i.e., PTP networks). However, in order to get a feeling of how much improvement AOM can provide over VPG, we fix one dimension of the problem, namely, the transmission powers, and limit our attention to rate optimization. Specifically, for PTP networks, we examine the case when nodes use the maximum power ($P_{\max}$). First, consider the following result.

Proposition 4 *The powers and rates that optimize (4.16) are such that the first constraint is satisfied with equality.*

Proof: The proof is similar to the one for Proposition 2, and is omitted for brevity.

If all users operate at $P_{\max}$, then from Proposition 4, it is easy to compute the users rates for both AOM and VPG by solving the following set of equations:

$$R_i = f^{-1}\left(\frac{W\ h_{ii}P_{\max}}{\sum_{j\in I-\{i\}} h_{ji}P_{\max} + P_{\text{thermal}}}\right), \quad \forall i \in I. \tag{4.17}$$

For MultiPTP networks, we follow a different approach that allows us to obtain a lower bound on the achievable gain of AOM over VPG schemes. Without loss of generality, let the users in the set I be ordered according to their link-channel gains, i.e., $i < j \Rightarrow h_i \geq h_j$. It has been shown in [57] that in the case of VPG [8], the optimal solution for (4.16) has the following structure:

- The set of best v_1 users (I_{v_1}) operate at rate $R_{\max}$ (i.e., at the maximum-rate boundary) and their powers satisfy $h_i P_t^{(i)} = h_j P_t^{(j)} \ \ \forall i, j \in I_{v_1}$, i.e., they have equal *received* powers.

- The set of next v_2 best users (I_{v_2}) operate at power $P_{\max}$ (i.e., at the maximum-power boundary) and rates $R_i < R_{\max} \ \ \forall i \in I_{v_2}$. Note that $h_i P_t^{(i)} < h_j P_t^{(j)}$ $\forall i \in I_{v_2}$ and $\forall j \in I_{v_1}$ (see [57] for more details).

- At most, there is one user U (whose order in I is $v_1 + v_2 + 1$) that operates at rate R_U and power $P_t^{(U)}$ such that $R_{\min} < R_U < R_{\max}$ and $P_t^{(U)} < P_{\max}$. Furthermore, $h_U P_t^{(U)} < h_i P_t^{(i)}$ and $R_U < R_i \ \ \forall i \in \{I_{v_1} \bigcup I_{v_2}\}$.

- The remaining users, $I_{v_3} = I - I_{v_1} - I_{v_2} - \{U\}$, operate at rate $R_{\min}$ (i.e., at the minimum-rate boundary) and power $h_i P_t^{(i)} = h_j P_t^{(j)} \ \ \forall i, j \in I_{v_3}$, i.e., they have equal *received* powers. Furthermore, $h_i P_t^{(i)} < h_j P_t^{(j)} \ \ \forall i \in I_{v_3}$ and $\forall j \in \{I_{v_1} \bigcup I_{v_2} \bigcup U\}$.

Using this solution structure, we now present a proposition that will enable us to derive a novel algorithm for finding the optimal solution for VPG networks. We then show how this algorithm can be used as a heuristic for AOM networks.

Proposition 5 *For VPG MultiPTP networks, the optimal solution to (4.16) is such that there is only one element in the set $\{I_{v_2} \bigcup U\}$, i.e., some users operate at the maximum-rate boundary, others operate at the minimum-rate boundary, and only one user operates at a rate in between these two boundaries.*

[8] The authors in [57] did not consider a minimum-rate constraint; however, their results extend to the case when $R_{\min} > 0$.

Proof: See Appendix D.

This optimal solution is intuitive and agrees with previously reported information theoretic results [112]; if there is no constraint on the maximum rate, the system throughput is maximized while simultaneously satisfying each user's minimum-rate constraint only when the best-channel user is allowed to transmit at a power larger than the one required for it to achieve $R_{\min}$. If there is a constraint on the maximum rate, allowing only the best user to increase his power may not achieve the maximum network throughput. The reason is that the best user cannot utilize any extra power beyond the one required to achieve $R_{\max}$. Hence, the optimal policy will then be that some best-channel users operate at $R_{\max}$ (without using $P_{\max}$), some bad-channel users operate at $R_{\min}$, and at most one user operates at a rate that is between $R_{\max}$ and $R_{\min}$.

Based on Proposition 5, the optimal solution for VPG networks can be found by assigning rate $R_{\max}$ to the maximum possible number of users such that the feasibility condition (4.14) is not violated, and then assigning to the next best user the maximum power at which (4.14) is satisfied with equality. The details of the algorithm are as follows:

1. Assign rate $R_{\min}$ $\forall i \in I$ and check the feasibility condition in (4.14); if this condition is not satisfied, then there is no solution to this problem; otherwise go to the next step.

2. Assign rate $R_{\max}$ to the best user in I, say user j, and check the feasibility condition in (4.14); if satisfied, then set $I = I - \{j\}$ and repeat this step; otherwise, go to the next step.

3. Find the maximum power ($P_{allowed}$) that j can use such that (4.14) is satisfied with equality; the transmission power of j is then given by $P_j = \min\{P_{allowed}, P_{\max}\}$.

This rate/power assignment (RPA) algorithm gives an optimal solution for VPG. We now explain the intuition behind using the same algorithm as a *heuristic* for AOM.

The main idea is to replace the objective function in (4.16) by a slightly different but related objective function, and then measure the actual throughput under this new function. First, note that in the case of AOM, $f(R_i)$, which was shown in Part (a) of Figure 4.4, can be well-approximated by a second-degree polynomial in R_i, say $a_1R_i^2 + b_1R_i + c_1$, which can be written as $(a_1(R_i + b_2)^2 + c_2)$ for some coefficients a_1, b_1, c_1, b_2, and c_2. For the set of data we have, the polynomial approximation has a higher fitting error than the posynomial fitting chosen earlier $\left(\text{i.e., } aR_i^b\right)$. Let $O_i \stackrel{\text{def}}{=} (R_i + b_2)^2$, $O_{\min} \stackrel{\text{def}}{=} (R_{\min} + b_2)^2$, and $O_{\max} \stackrel{\text{def}}{=} (R_{\max} + b_2)^2$, and replace the objective function in (4.16) by $\sum_{i\in I} O_i$. Then, the optimization problem becomes:

$$\begin{cases} \underset{\{O_i, P_t^{(i)},\ i\in I\}}{\text{maximize}} & \sum_{i\in I} O_i \\ \text{subject to:} & \\ \frac{h_{ii}P_t^{(i)}}{\sum_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\text{thermal}}} \geq \frac{a_1O_i + c_2}{W}, & \forall i \in I \\ 0 \leq P_t^{(i)} \leq P_{\max}, & \forall i \in I \\ O_{\min} \leq O_i \leq O_{\max}, & \forall i \in I \end{cases} \tag{4.18}$$

This formulation has a similar structure to the one of VPG. Since RPA finds the optimal solution to VPG, it can also find the optimal solution to (4.18) (a_1 and c_2 are constants that do not affect the optimization algorithm). This means that RPA can be used to maximize $\sum_{i\in I}(R_i + b_2)^2$. The powers and rates that maximize $\sum_{i\in I}(R_i + b_2)^2$ are not necessarily equal to the ones that maximize $\sum_{i\in I} R_i$. However, we expect them to be close. In this sense, RPA can be used as a heuristic method to maximize $\sum_{i\in I} R_i$. The simulation results in Section 4.4 show that based on this heuristic, AOM provides significant performance advantages over VPG.

4.4 Performance Evaluation

4.4.1 Simulation Setup

In this section, we evaluate the performance of AOM and contrast it with that of VPG [57]. Our results are based on numerical experiments conducted using MATLAB. Our performance metrics include the service time (S)[9], the sum of users rates, and the average energy consumption per bit (E_b), defined as $\frac{\sum_{i=1}^{N} P_i}{\sum_{i=1}^{N} R_i}$. Note that E_b is a more significant measure than the average transmission power. In fact, it is misleading to compare the average transmission power of two systems that transmit at different data rates, as the cost of transmitting a certain number of bits depends on both the transmission power and the rate. In some cases, we also study the throughput and energy fairness indexes $I_R = \frac{\left(\sum_{i=1}^{N} R_i\right)^2}{N \sum_{i=1}^{N} R_i^2}$ and $I_E = \frac{\left(\sum_{i=1}^{N} (P_i/R_i)\right)^2}{N \sum_{i=1}^{N} (P_i/R_i)^2}$, respectively [97]. The fairer the system, the higher are the values of I_R and I_E. I_R measures the "equality" of users allocation of throughput. If all users get the same amount of throughput, then the fairness index is 1, and so the system is 100% fair. As the discrepancy in throughput increases, I_R decreases. A scheme that favors only few users has a fairness index close to zero. I_E measures the discrepancy in the amount of energy each user invests in delivering one bit of information. Typically, a system designer would like this per-bit energy to be equal for all users to extend the lifetime of users' batteries. To simulate the channel gains, we assume the two-ray propagation model with a path loss factor of 4. Note, however, that the problem formulation does not depend on *how* the channel attenuation matrix is generated, i.e., any other fading model can be used. The total bandwidth of the system (i.e., the chip rate) is $W = 1$ MHz. We let $P_{\max} = 20$ dBm.

4.4.2 Point-to-Point Networks

In this scenario, N transmitting nodes are randomly placed across a square area of length 600 meters. For each transmitter j, the receiving node i is placed randomly

[9] At the optimal solution for the first objective function, all users have the *same* service time S (see Proposition 3).

Number of Node Pairs (N)	Scheme	Minimize Max S_i		Maximize $\sum R_i$			
		S (sec)	E_b (microjoules/bit)	$\sum R_i$ (Mbps)	E_b (microjoules/bit)	I_R	I_E
20	VPG	73.6	1.86	1.83	1.09	0.50	0.21
	AOM	25.5	0.45	2.08	0.96	0.58	0.30
30	VPG	137.5	0.58	2.57	1.17	0.47	0.51
	AOM	69.3	0.34	2.94	1.02	0.55	0.58

Table 4.1: Performance Comparison between AOM and VPG in PTP networks.

within a circle of radius 100 meters that is centered at j. Given the location of the N receivers and N transmitters, the channel attenuation the channel attenuation between any pairs of nodes i and j is computed using the the two-ray propagation model with attenuation factor equals to 4. The matrix H is then formed with entries h_{ij}. Whenever the solution set is empty for the generated H (i.e., $R_{\min}$ cannot be achieved for all users), a new set of transmitters and receivers are randomly generated. The maximum-rate constraint $R_{\max}$ is chosen such that the modulation order M used in VPG is equal to 16, which is the *minimum* M used in AOM. For this experiment, we let $R_{\min} = R_{\max}/100$.

The performance of AOM and VPG is shown in Table 4.1. The results are reported for $N = 20$ and $N = 30$ based on the average of 100 independent realizations of the matrix H. For the first objective function, (i.e., minimizing the maximum service time), all nodes are assumed to have 1 Mbits of data. Although a randomly generated workload is more practical, the choice of equal workloads is meant to facilitate the discussion. For $N = 20$ and $N = 30$, AOM achieves a reduction in S by 65.4% and 50%, respectively, while simultaneously achieving about 75% and 42% energy savings, respectively. The reason for this considerable improvement can be explained as follows. From Proposition 3, we know that at the optimal powers and rates, all users have the same S. Since users have the same load (1 Mbits), the optimal solution is when all users transmit at the *same* rate. This rate must be chosen to accommodate the worst-channel user (i.e., lowest SNR). AOM, as Figure 4.4 shows, has a significant performance advantage over VPG at *low* SNR values; thus

providing a smaller service time and a much lower energy consumption than VPG.

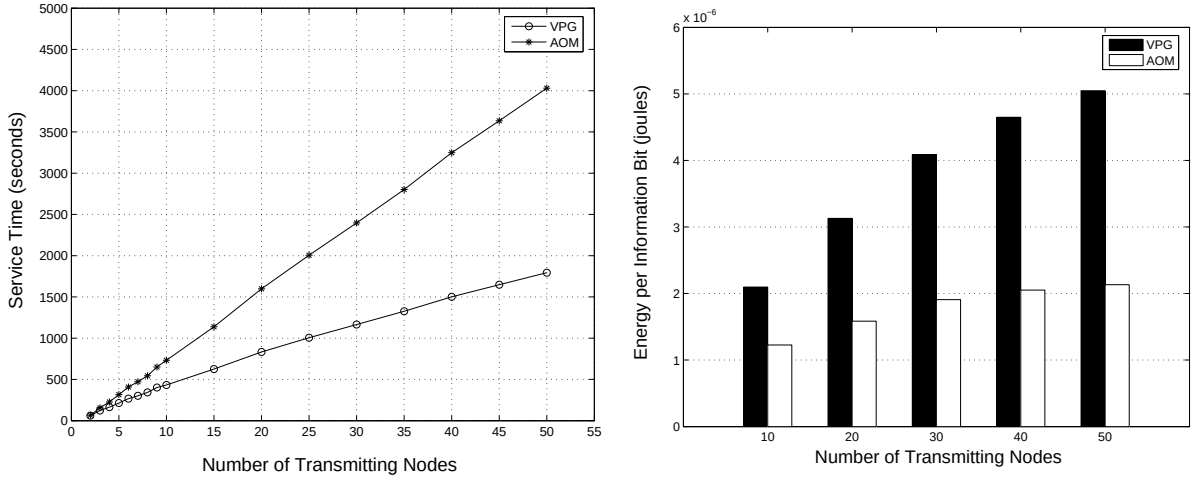
For the second objective function (i.e., maximizing the sum of users rates), the optimal solution in PTP networks is unknown; however, to provide a feeling of what the AOM improvement is, we let all users transmit at $P_{\max}$, and compute the corresponding optimal rates using (4.17). For both $N = 20$ and $N = 30$, AOM achieves about 15% increase in throughput, 13% saving in energy, and about 16% improvement in I_R relative to VPG. The improvement in I_E for $N = 20$ is particulary significant (about 42%). Such an improvement is justified by noting that AOM achieves a significant throughput gain for low-rate (low-SNR) links, sometimes twice that of VPG, but provides little gain for high-rate links. This has a negligible impact on throughput, but has a significant impact on I_E.

4.4.3 Multipoint-to-Point Networks

In this section, we consider N transmitting nodes that are randomly placed within a square area of length 200 meters. The common receiver is placed at the center of the square. Given the location of the nodes, the channel attenuation matrix H. Similar to the PTP case, whenever the solution set is empty for the generated H, a new set of transmitters is randomly generated. The results are obtained based on the average of 100 independent realizations of the matrix H.

For the first objective function, the workload at each transmitter is selected randomly between 1 and 20 Mbits. As before, $R_{\max}$ is chosen such that the modulation order M used in VPG is equal to 16. Figure 4.5 depicts the performance of AOM and VPG for the first objective function. Part (a) of the figure depicts the service time S versus N. It is shown that as N exceeds 10, AOM achieves considerably lower S than VPG. For example, when $N = 50$, S under AOM is only 45% of S under VPG. It is also shown that S under both AOM and VPG increases with N. This is expected since as N increases, the multiple access interference (MAI) also increases, and users are forced to transmit at lower rates, which increases their service times.

Note that for VPG, the S-versus-N curve is approximately linear, while for



(a) (b)

Figure 4.5: Performance of AOM and VPG based on the minimization of the maximum service time in MultiPTP networks.

AOM, the slope of that curve decreases slightly with N. This can be explained by examining (4.15). The RHS of (4.15) is close to 1, as P_{thermal} is typically very small. In the case of VPG, $f(L_i/S) = \mu_{\text{req}} L_i/S$, $WS/L_i\mu_{\text{req}} \gg 1$, and so the LHS of (4.15) can be well-approximated by $\frac{\mu_{\text{req}}}{S}\sum_{i\in I} L_i \approx \frac{\mu_{\text{req}}}{S} N L_{ave}$, where L_{ave} is the expected value of L_i. This explains why S increases almost linearly with N. For the AOM case, $f(Li/S) = a(L_i/S)^b$ for some coefficients $a > 0$ and $b > 1$. It is easy to show that the S-versus-N curve can be approximated by $S \approx cN^{1/b}$, for some coefficients $c > 0$. Thus, its derivative (or slope) decreases with N.

Part (b) of Figure 4.5 depicts E_b versus N. It shows that in addition to reducing the service time, AOM achieves a significant energy saving over VPG. For example, for $N = 50$, AOM energy expenditure is less than 40% that of VPG.

Next, we study the impact of increasing P_{thermal} on the service time S. Figure 4.6 shows S as a function of P_{thermal} for $N = 30$. The workload is generated as in Figure 4.5. For all values of P_{thermal}, AOM consistently shows a good improvement over VPG. For both AOM and VPG, however, S starts to increase exponentially when P_{thermal} exceeds -60 dBm. The reason is that at this value, P_{thermal} becomes comparable to the maximum received powers for bad-channel users. Hence, the

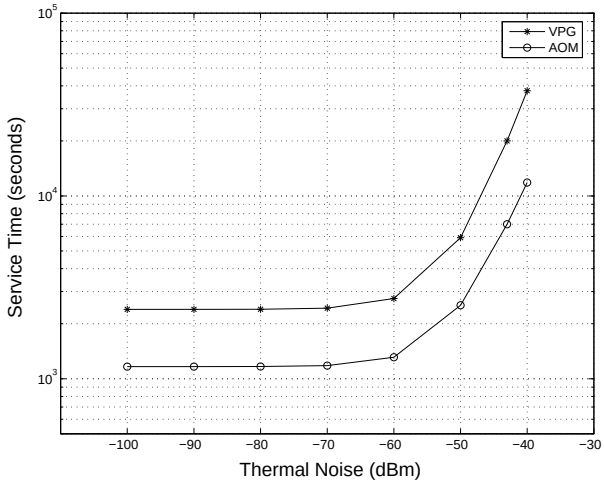


Figure 4.6: Service time S of AOM and VPG in MultiPTP networks as a function of P_{thermal}.

SNR of the users deteriorates significantly, causing a fast drop in their rates, and a corresponding dramatic increase in S.

In the case of the second objective function (i.e., maximizing the sum of users rates), we set the maximum modulated bit rate Z to $W/5$. As before, $R_{\max}$ is chosen such that the modulation order M used in VPG is equal to 16 (i.e., $R_m/R = 4$), so $R_{\max} = W/20$. Part (a) of Figure 4.7 depicts the throughput performance versus N for three different values of $R_{\min}$ ($R_{\max}/50$, $R_{\max}/100$, and zero). Recall that the used RPA algorithm is optimal for VPG , but is only a heuristic for AOM, so the results in Figure 4.7 represent a lower bound on the achievable gain of AOM over VPG. Several observations can be made based on this figure.

First, AOM achieves considerably more throughput than VPG; e.g., for $R_{\min} = R_{\max}/50$ and $N = 50$, AOM achieves about 30% more throughput than VPG. This is because for any power allocation vector, AOM enables higher rates than VPG. Second, in the cases of $R_{\min} = R_{\max}/50$ and $R_{\min} = R_{\max}/100$, as N increases, the throughput for AOM increases, while the throughput for VPG decreases. This can be explained as follows. For VPG, as N increases, more bad-channel users are required to operate at $R_{\min}$. To enable this, other (good-channel) users must

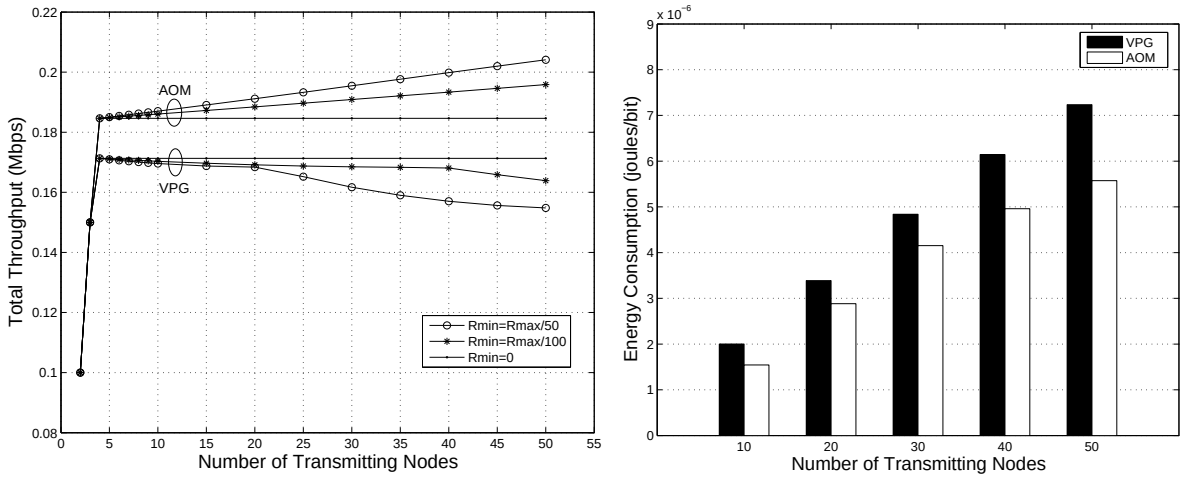


Figure 4.7: Performance versus N under the throughput maximization criterion for MultiPTP networks.

decrease their powers (and consequently their rates) to reduce the MAI. The increase in the total throughput due to a higher number of bad-channel users does *not* offset the decrease in the throughput of the good-channel users. Therefore, the overall effect is a slight reduction in network throughput. This is not the case, however, for AOM. Simulation results indicate that the increase in the total throughput due to more bad-channel users is higher than the decrease in the throughput of the good-channel users. This can be justified as follows. Unlike VPG, AOM uses higher OM orders at low data rates and thus requires much less SNR than VPG to achieve $R_{\min}$. Thus, good-channel users do *not* need to reduce their powers (and their rates) considerably to accommodate the new users, and so the reduction in the throughput of the good-channel users is not considerable (when compared to the VPG case). The overall effect is a slight increase in network throughput. The throughput of AOM increases with N until the RPA is unable to find a feasible solution.

Another observation is that as $R_{\min}$ increases, the throughput for VPG decreases, while the throughput for AOM increases. So the throughput gain of AOM over VPG goes up with $R_{\min}$. This can be explained as follows. Increasing $R_{\min}$ tightens the constraints (i.e., reduces the solution space), and this results in a lower throughput

whenever RPA is optimal. This is exactly what happens in the VPG case since RPA is optimal for VPG. But since RPA is *heuristic* for AOM, we conjecture that its performance becomes closer to the optimal one as $R_{\min}$ increases, and so the throughput increases.

The last point to note about Figure 4.7-(a) is that for $R_{\min} = 0$, both AOM and VPG are almost linear. The reason is that when $R_{\min} = 0$, RPA allocates powers only to good-channel users until the network "saturates," i.e., v_1 users are assigned $R_{\max}$ and only one user is assigned the rest of the power such that (4.14) is satisfied. Adding more users has no impact once (4.14) is satisfied.

As in the PTP case, the throughput advantage of AOM over VPG comes with energy savings. Part (b) of Figure 4.7 depicts the energy consumption of AOM and VPG as a function of N for $R_{\min} = R_{\max}/50$. This figure shows that AOM achieves a significant energy saving over VPG (up to 25%).

Next, we study the fairness properties of AOM and VPG. Part (a) and (b) of Figure 4.8 depict I_R and I_E, respectively, as a function of N (recall that the fairer the system, the higher are the values of I_R and I_E). The results are for $R_{\min} = R_{\max}/50$. It can be observed that relative to VPG, AOM can improve I_R and I_E up to 21% and 30%, respectively.

Finally, we study the effect of varying the minimum processing gain (PG) by varying $R_{\max}$. We fix $R_{\min}$ in this experiment at $W/500$. Figure 4.9 shows the performance of AOM and VPG as a function of the minimum PG. It can be observed that the sum of rates decreases as the PG increases for both AOM and VPG. This agrees with the previous intuition that reducing $R_{\max}$ tightens the solution space, and so decreases the achieved maximum. Furthermore, it not difficult to notice that RPA favors higher values of $R_{\max}$.

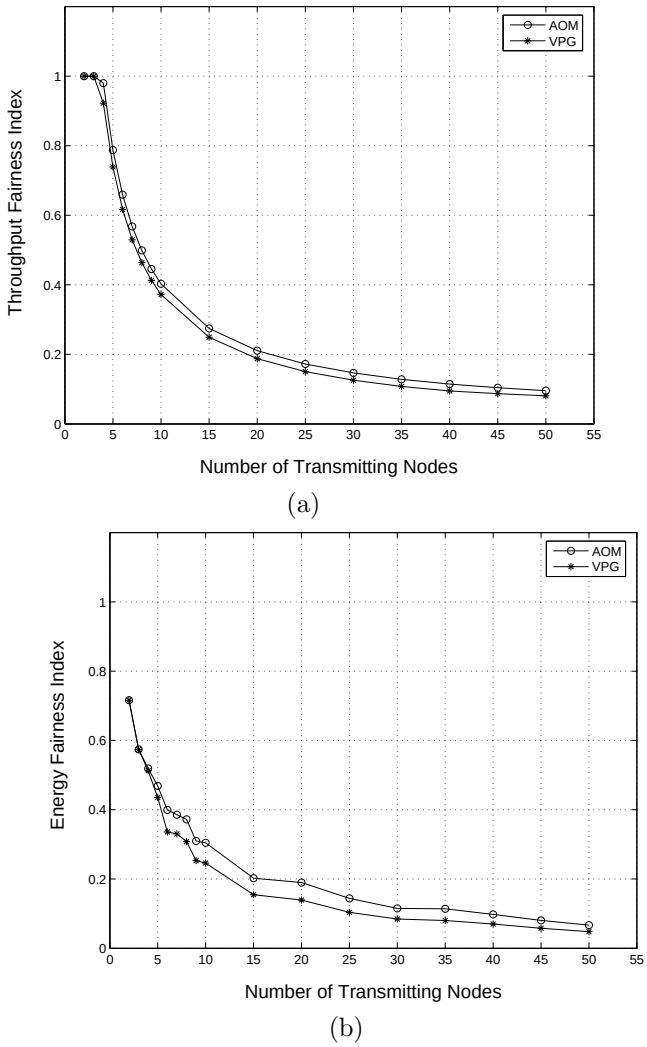


Figure 4.8: Fairness versus N under the throughput maximization criterion for MultiPTP networks.

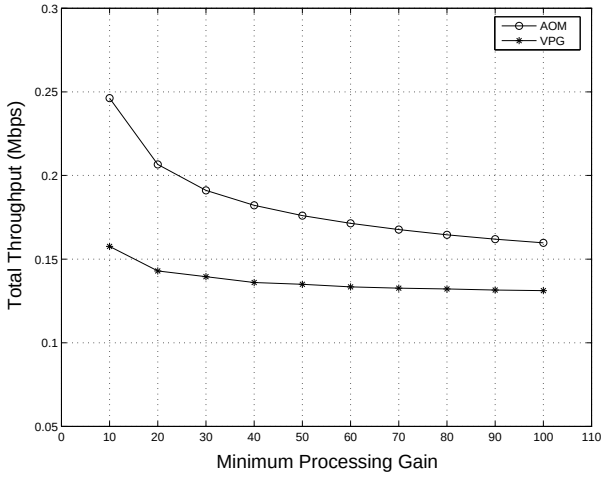


Figure 4.9: Throughput in MultiPTP networks as a function of the minimum processing gain (varied through $R_{\max}$).

4.5 Adaptive Channel Coding

Adaptive convolutional coding has been the topic of extensive research (e.g., [47, 46, 37]). The basic idea is to change the code rate, and hence the correction capability of the code, to match the source and channel needs. For practical purposes and to avoid high complexity, it is important *not* to require switching between a set of encoders and decoders, but rather use one encoder and one decoder that can be modified without changing their basic structure [47]. There are several ways to achieve this including puncturing [47], repetition codes [23], and nested codes [74]. While puncturing is a widely used technique for high rates, repetition and nested codes are used to achieve very-low rates, low-complexity coding [37]. The system we examine in this dissertation focuses on very-low rates capacity-achieving coding schemes, which have been the focus of many previous works (e.g., Viterbi [115] and Haccoun [46]).

Repetition code cannot provide any coding gain [23], so lowering the code rate by simply repeating all the code bits shows no improvement in the performance in AWGN channel (but can provide some improvement in Rayleigh fading channels [23]). Nested convolutional codes, on the other hand, are a simple low-

complexity coding method that is used to provide very-low rates with some coding gain. Such codes are obtained by extending a code rate of k/n to a rate $k/(n+1)$ code by searching for the best additional generator polynomial [46]. In order to compare adaptive orthogonal modulation (AOM) with adaptive nest convolution codes (ANCC), we pick one good nested code and examine its performance. The nested code we use here is the (561,753) convolutional code of constraint-length 9 that was found in [78] with rates $2/(5+n)$, where $n = 0, \ldots, 35$. For this code, and assuming binary orthogonal modulation[10], the required E_b/N_0 (and hence SNR) to achieve a certain BER can be found using the union bound [92]:

$$P_b < \sum_{d=d_f}^{\infty} \frac{c_d}{2} \text{erfc} \left(\frac{dE_c}{2N_0} \right)^{1/2}, \tag{4.19}$$

where c_d of the total number of bit errors for all different starting points of error events of distance d given in [78], and E_c/N_0 is the signal energy-to-noise ratio per coded bit.

Figure 4.10 shows the probability of error for orthogonal modulation (OM) and for the nested code given in [78] at BER 10^{-6}. The highest OM order ($M = 128$) is chosen such that the coding rate ($\log_2 M/M$) corresponding to that order is roughly equal to the lowest coding rate of the nested code given in [78]. It can be seen that for moderate to low rates, nested codes significantly outperform OM. However, as the coding rate is decreased, OM performance comes very close to that of nested codes. To investigate the performance at very low rates, we used curve-fitting to extend the previous results to very low rates. Figure 4.11 shows the performance for such rates. It is shown that OM outperforms nested codes for very low rates. To better understand these results, the performance of OM *relative* to nested codes is shown in Figure 4.12. For high rates, say 0.25, OM requires 40% more SNR than nested codes. However, for very low rates, say 0.001, OM requires about 35% less SNR relative to nested codes. Hence, at very low code rates, OM become an attractive coding scheme for CDMA systems.

[10]It is important to emphasize that the results we derive here are limited to binary orthogonal modulation.

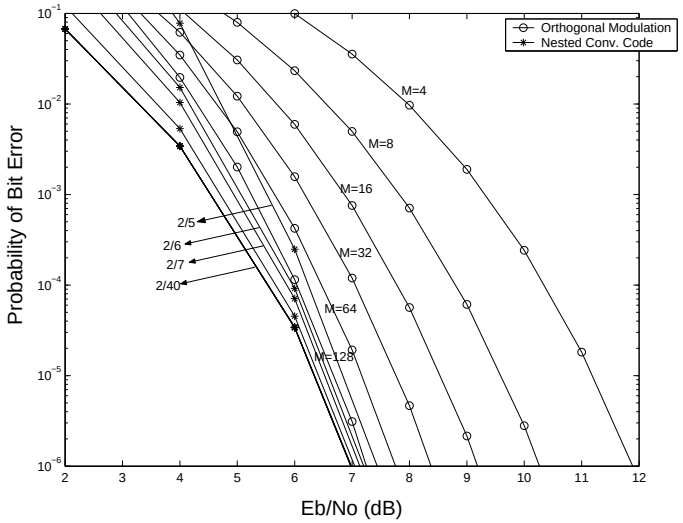


Figure 4.10: Probability of bit error for AOM and ANCC.

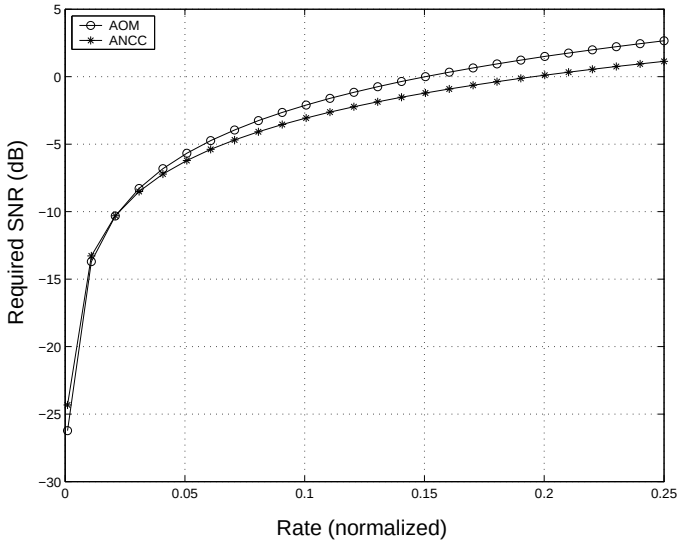


Figure 4.11: Performance of AOM and ANCC for the single link case.

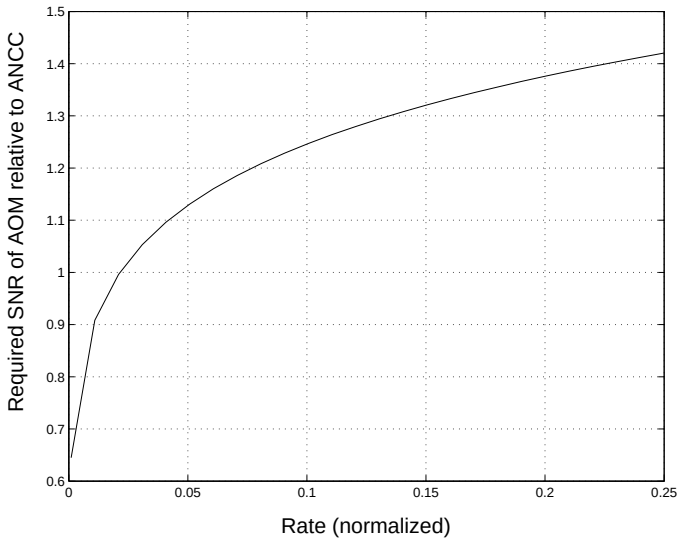


Figure 4.12: Relative performance of AOM to ANCC for the single link case.

Using the relationship between the required SNR and the code rate, we are now in a position to compare ANCC with AOM (and no coding). For this purpose, we focus on minimizing the maximum service time since we know how to find the optimum solution for that problem. In the simulation, the maximum and minimum data rate were set to $W/4$ and $W/400$, respectively, where W is the Fourier bandwidth of the channel. Figure 4.13 shows the service time of AOM relative to ANCC. It is shown that for a low number of users, the service time of AOM is 30% more than that of ANCC. The reason is that for a small number of users, the resulting data rates are high. For such high data rates, the coding gain of convolutional codes is much higher than that of OM. On the other hand, as the number of users is increased, users' rates decrease, and the coding gain provided by OM outperforms that of convolutional codes. Hence, the maximum service time of OM is less than that of convolutional codes. Note that as the code rate goes to infinity, AOM achieves the Shannon bound [92]. The same cannot be said about ANCC because the derived lower rate codes do not achieve the same coding gain as the corresponding (same rate) optimal convolutional codes. The actual service time can be seen in Figure 4.14.

Note that M-ary OM can be looked at as a (M,$\log_2 M$) block code of distance

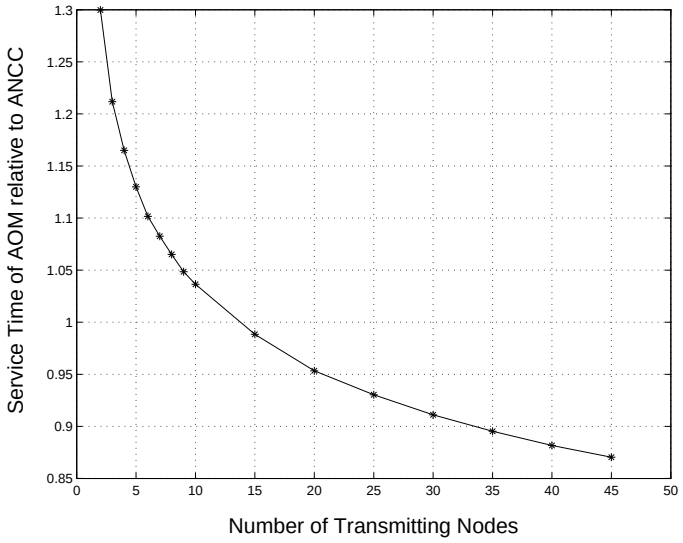


Figure 4.13: Performance of AOM relative to ANCC for minimizing the maximum service time metric.

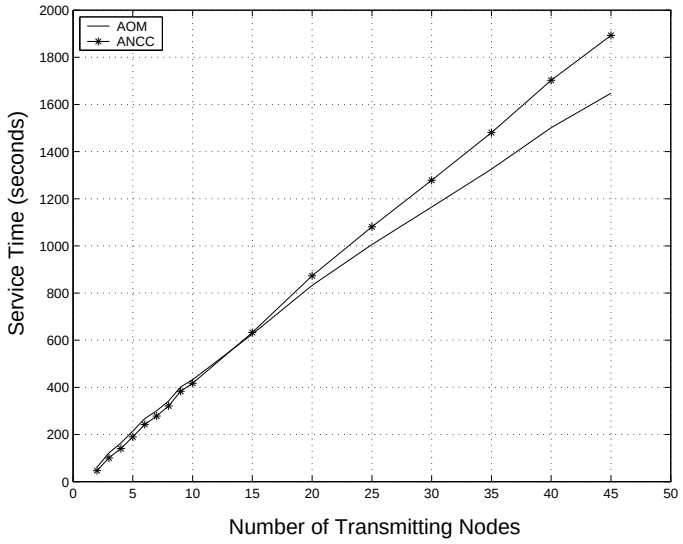


Figure 4.14: Performance of AOM and ANCC for minimizing the maximum service time metric.

$M/2$. In terms of complexity, OM requires M correlation receivers. In comparison, convolutional codes require only two correlation receivers for coded binary OM, and one decoder. The complexity of decoding higher OM order can be alleviated by the use of the Fast Walsh Transform method [9, 114], which requires only $M\log_2 M$ real additions and subtractions.

CHAPTER 5

PERFORMANCE OF WIRELESS CDMA NETWORKS UNDER OPTIMAL LINK-LAYER ADAPTATION

5.1 Introduction

Link-layer adaptive transmission techniques have recently emerged as powerful tools for improving the utilization of the limited wireless spectrum. In these techniques, one or more tunable parameters are used for link adaptation. These parameters are dynamically varied depending on the channel conditions, interference levels, rate and BER requirements, and energy constraints [40, 83]. In the single-link case, the user can adapt the transmission power, the coding, and/or the modulation scheme to meet the transmission rate and BER requirements [43]. For a given fading condition (i.e., channel gain), the maximum instantaneous rate in this case is upper-bounded by Shannon's capacity $C = W \log_2(1 + \text{SNR})$, where W is the Fourier bandwidth of the channel and SNR is the instantaneous received signal-to-noise ratio [40].

For multiuser systems, CDMA has been one of the leading technologies of the last decade [101]. Traditionally, CDMA networks have been designed to carry voice traffic only, for which rate adaptation is not desirable due to the delay-intolerant nature of voice [57]. Thus, transmission power control has been the single most important adaptation parameter [39], and has been thoroughly studied (see [110] and the references therein). For such a fixed-rate design, the "capacity" is naturally defined in terms of the maximum number of users that can be supported at a given SNR. This notion of capacity has been well-established in [39], and continues to serve as a benchmark for fixed-rate systems.

Traffic in forthcoming wireless networks is expected to be a mixture of real-time traffic (e.g., voice) and data traffic (e.g., email, messaging, web browsing, file transfer, etc.). The delay-tolerant nature of data traffic along with the increasing demand

for spectrally-efficient systems has motivated *rate* adaptation. Several rate-control techniques have been proposed based on multiple CDMA codes [106], parallel combinatory spread spectrum [134], multiple chip rates [127], adaptive modulation and coding (AMC) [15, 50], adaptive orthogonal modulation [82], and variable processing gain (VPG) techniques [29, 57, 66, 71, 119]. In particular, the authors in [57] have demonstrated the significant throughput advantage of adapting both the transmission power and rate in VPG CDMA networks employing a conventional matched filter (MF) receiver. The authors in [82] considered an additional dimension of the adaptation space by also adapting the orthogonal modulation order.

Given the growing interest in adaptation techniques, there is a need to answer a fundamental question: For given channel conditions, users requirements (e.g., minimum rate, maximum BER), and system constraints (e.g., maximum power), *what is the maximum instantaneous achievable performance of a CDMA network that employs a MF receiver and that operates under "optimal" link-layer adaptation?* In this context, "optimal adaptation" refers to any combination of transmission powers, user rates, modulation orders, and coding rates, *optimized for given instantaneous channel conditions*. The main contribution of this chapter is to answer the above question for two important throughput-related performance metrics: minimizing the maximum service time and maximizing the sum of users rates. As explained later, the two metrics differ in the time scale at which rate adaption is carried out and in the involved hardware complexity. The importance of establishing performance bounds is akin to the significance of Shannon's capacity in the single-link case and to the maximum number of users in [39] for fixed-rate systems; such bounds serve as benchmarks against which adaptive multirate systems subject to QoS requirements can be compared.

Another contribution of our work is that our analysis is applicable to both point-to-point (PTP) and multipoint-to-point (MultiPTP) networks. PTP networks is the more general communication paradigm. It can represent a completely distributed mobile ad hoc network, or a microcellular network in which mobile-base station pairs compete for the same frequency spectrum. In MultiPTP networks, multiple nodes

transmit to one node, as in the case of a cluster-based ad hoc or sensor network [88] or in the case of the uplink of a single cell in a CDMA-based cellular network (e.g., IS-95 [92]). Note that most previous capacity results considered only MultiPTP networks.

To compute the "optimal" network performance, we pursue an algorithmic approach whereby we jointly optimize users powers and rates assuming that each link individually operates at the Shannon capacity. We optimize the two throughput-related objective functions (i.e., maximum service time and sum of rates) subject to constraints on the maximum transmission power, the minimum transmission rates, and the BER. Our optimization framework can capture the wide variations in QoS requirements that characterize future multimedia networks. The elegance of the method stems not only from its ability to find the global optimum solution, but also from the computational efficiency with which this solution can be obtained.

We obtain the optimum solution to the problem of minimizing the maximum service time in both PTP and MultiPTP networks by first formulating the problem as a GGP [18] and then transforming this GGP into a GP. This GP can itself be transformed into a nonlinear convex program for which the global optima that can be efficiently found [19]. In the case of MultiPTP networks, we derive a simple expression for computing the optimal powers and rates that minimize the maximum service time. Our results make it possible to determine the feasibility of a set of rates and BER requirements under given constraints, thus allowing for the use of admission control policies.

For the maximization of the sum of rates, the optimization problem exhibits several local maxima. So there are no computationally efficient algorithms to *exactly* solve this problem. To address this issue, we approximate the sum-of-rates objective function by another function that lends itself to global optimization.

It is important to emphasize that the maximum sum of rates studied in this chapter is *not* the same as the information-theoretic capacity of the multiuser channel, which has been addressed in several previous papers. Specifically, the information capacity regions for fading channels studied in [48, 76, 77, 113] can be achieved

only with a (complex) optimum receiver structure that uses successive decoding. Furthermore, these capacity results were derived subject to *average* power constraints, with averaging being performed over the fading statistics. Our work differs in three key aspects. First, we focus on CDMA systems that employ a simple *MF receiver*[1]. Second, instead of averaging over some fading statistics, we optimize the network performance subject to given (instantaneous) channel gains. Our results are thus independent of the statistical characteristics of the fading process. Third, in our formulation we impose a *peak* rather than a mean constraint on the transmission powers (in practice, a peak-power constraint is always present, but an average-power constraint may or may not be present). Such a constraint was considered in [59] but only for multiaccess broadcast channels. The information-theoretic result of most relevance to our work is the sum-rate capacity of a Gaussian multiaccess channel (C_{Gaussian}) [28], defined as the maximum sum of rates taken over all rate vectors in the capacity region [40]. However, this C_{Gaussian} can be achieved only with a (complex) optimum receiver structure and not an MF receiver. Relative to C_{Gaussian}, our results provide *tighter upper-bounds* on the performance of CDMA systems with MF receivers, which have been the focus of many papers (e.g., [29, 57, 62, 66, 71, 82, 86, 105, 119, 129]) and are also part of 3G wireless standards. In addition to studying the performance with respect to the traditional throughput metric, we also study it with respect to the overall *service time*, which is a novel metric in this context.

Some previous work focused on finding the maximum number of users for a set of classes of users (e.g., [89]). The results in [26] apply only to the downlink of a cellular network. The authors in [96, 62, 25] studied the rate maximization for MQAM modulation wireless systems. The authors in [120] considered the max-min rate allocation problem for Aloha systems. Significant body of work was devoted to deriving the Erlang capacity (e.g., [117, 36]), defined as the average load that can be supported at a given blocking probability. These efforts do not consider adaptivity

[1]MF receivers, popular for their low complexity and cost, treat other users interference as Gaussian noise.

in all dimensions, but rather focus on power control or power/rate control[2]. Our results provide bounds on any adaptive system, including adaptivity of the coding and modulation scheme/order.

The rest of the chapter is organized as follows. In the next section, we review the capacity results for the single-link case, which provides a starting point for the analysis of the multiuser case. In Section 5.3, we present the objective functions, formulate the optimization problems, and present their solutions. Performance results are presented in Section 5.4.

5.2 Single-Link Case

In this section, we briefly analyze CDMA networks from the perspective of a single user. We then tie the analysis with Shannon's capacity and use this as a groundwork for the *multiuser* case, which is treated in subsequent sections.

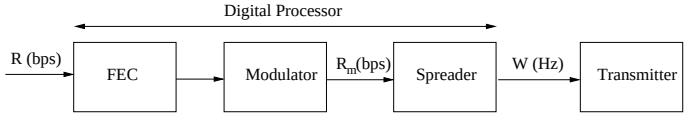


Figure 5.1: Simplified block diagram of the transmitter circuit.

The transmitter circuit of the system under study is shown in Figure 5.1. It consists of digital forward error correction (FEC), modulation, direct-sequence spreading, and analog amplification and transmission [39]. Let I be the set of active links in the network, $P_t^{(i)}$ be the transmission power of link i, and h_{ji} be the channel gain between the receiver of link i and the transmitter of link j. Consider packet reception for link i. The SNR at receiver i is:

$$\mathrm{SNR}^{(i)} = \frac{h_{ii}P_t^{(i)}}{\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\mathrm{thermal}}} \tag{5.1}$$

[2]They assume that the required E_b/N_0 (defined shortly) is fixed. This ratio can vary when the coding or modulation schemes/rates are adaptive.

where P_{thermal} is the thermal noise. For reliable communication in CDMA networks, a more relevant metric than $\text{SNR}^{(i)}$ is the effective bit energy-to-noise spectral density ratio at the detector, denoted by $\mu^{(i)}$ and given by [39]:

$$\mu^{(i)} \stackrel{\text{def}}{=} \frac{E_b}{N_0} = \frac{W}{R_i} \frac{h_{ii}P_t^{(i)}}{\sum\limits_{j \in I-\{i\}} h_{ji}P_t^{(j)} + P_{\text{thermal}}} \tag{5.2}$$

where R_i is the data rate of i's intended signal. In the case of fixed FEC scheme/rate and fixed modulation scheme/order, the required $\mu^{(i)}$ for a certain BER, denoted by μ_{req}, is constant. Thus, the required SNR for operating at that BER with data rate R_i is:

$$\text{SNR}_{\text{req}}^{(i)} = \mu_{\text{req}} \frac{R_i}{W}. \tag{5.3}$$

Both (5.2) and (5.3), which hold for any CDMA system, do not explicitly indicate the effects of FEC and modulation on the required SNR. These effects appear indirectly through the value of μ_{req}. For example, the stronger the FEC code (i.e., the lower the code rate[3]), the smaller are the values of μ_{req} and SNR_{req}, i.e., the lower is the required transmission power. This analysis is inline with Viterbi's findings [116], where he showed that the *jamming margin* is actually increased by coding; the idea is that with FEC coding, μ_{req} can be reduced, and so more interference is allowed for the same rate, i.e., $\text{SNR}_{\text{req}}^{(i)}$ in (5.3) becomes smaller.

The question of how much reduction in $\text{SNR}_{\text{req}}^{(i)}$ (through μ_{req}) can be achieved with FEC/modulation was answered by Shannon. Specifically, using Shannon's capacity equation for the single link, the *minimum* required SNR is given by[4]:

$$\left[\text{SNR}_{\text{req}}^{(i)}\right]_{\min} = 2^{R_i/W} - 1. \tag{5.4}$$

Note that in addition to providing the "average" capacity for a band-limited AWGN channel, Shannon's equation also gives the "instantaneous" capacity under the given

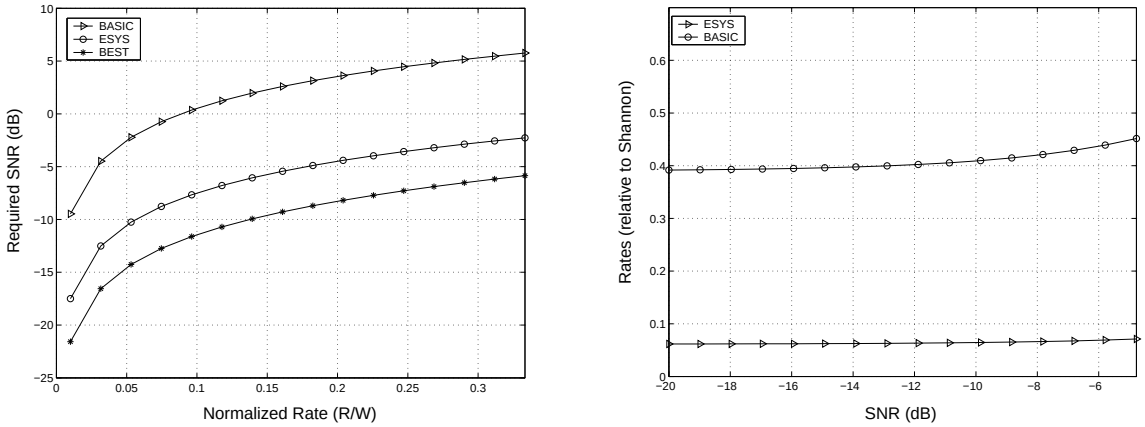
[3]The lower the code rate, the higher is the FEC-coded bit rate. However, for spread spectrum systems, this has no impact on system bandwidth since the signal will be spread using a high-rate CDMA code.

[4]This bound assumes that the probability distribution of the multiple access interference is Gaussian, which is a reasonable assumption when strong FEC codes are employed [39].

channel conditions [40]. So given a deterministic channel gain as an input, the maximum instantaneous achievable rate is provided by Shannon's equation. This result serves as a basis for the multiuser case treated in the next section.

Using the relationships between $\text{SNR}_{\text{req}}^{(i)}$ and R_i, we are now in a position to compare different FEC/modulation designs and contrast their performances against Shannon's bound reflected in (5.4). At this point, we are still dealing with the single-link case. For a given BER, the controllable parameters are the data rate, transmission power (equivalently, the SNR), and μ_{req}. To understand the interactions among these three parameters, we study the relationship between the transmission rate and the SNR for different values of μ_{req}. We consider two systems with extreme FEC capabilities: (i) a *basic system* (BASIC) that uses BPSK modulation but no FEC, and (ii) an *enhanced system* (ESYS) that uses BPSK and a relatively strong FEC code; a rate-1/3 convolutional code with a constraint-length $K = 41$ and sequential decoding with hard decisions (see [92] for details of this code). At a BER of 10^{-6}, the required E_b/N_0 for ESYS is only 2.5 dB, compared to 10.53 dB for the BASIC scheme. Note that since the FEC-coded bit rate must not exceed W, the maximum allowable data rate is $W/3$ (because the code rate in ESYS is 1/3).

Figure 5.2 demonstrates the relationship between the data rate and SNR_{req} for BASIC, ESYS, and Shannon's (referred to as BEST) at a BER of 10^{-6}. Part (a) of the figure depicts SNR_{req} versus the normalized rate R/W. It is clear that to achieve a certain data rate, BEST requires less SNR than ESYS, which itself requires a significantly less SNR than BASIC. Such an improvement essentially reflects a *power gain*. Alternatively, one may reflect the difference between these schemes in terms of the maximum achievable rate for a given SNR. Part (b) shows the achievable rates of BASIC and ESYS relative to BEST. It is shown that the relative rates of BASIC and ESYS increase slightly as the SNR increases, but it is also shown that even with a strong FEC (i.e., the ESYS case), the achievable rate does not exceed half of the channel capacity.



(a) Required SNR versus data rate.

(b) Relative rates of ESYS and BASIC.

Figure 5.2: Shannon capacity and the performance of ESYS and BASIC for a single link.

5.3 Optimal Performance in the Multiuser Case

For a network of users, increasing one user's power increases that user's SNR, and consequently allows this user to increase his rate. However, this comes at the expense of lowering the SNRs of other users, whose data rates must now be reduced to combat the added interference. This interaction among network users makes the determination of the optimal powers and rates that maximize the sum of rates or minimize the maximum service time less obvious. The goal of this section is to obtain such powers and rates. Consider first the following proposition.

Proposition 6 *The optimal* network *performance in terms of either maximizing the sum of users rates or minimizing the maximum service time can be achieved only if each link* individually *operates at the Shannon capacity in (5.4).*

The proof of this proposition is trivial, and is omitted for brevity. The idea is that the MAI is minimized when each user individually operates at its Shannon capacity. The smaller the MAI, the higher are the achievable rates, and so the lower is the maximum service time.

This proposition says that if the relationship between the rate and the required SNR of each link in the network is according to (5.4), then the powers and rates that

optimize that network's performance, if can be realized, would produce a bound on the performance of any kind of adaptation in that network. Adaptation here refers to power control, rate control, FEC scheme/rate, modulation scheme/order, or any combination of these parameters. Our goal therefore is to formulate optimization problems under the assumption that each link operates according to (5.4). We now look at how to find the optimum network performance with respect to the two metrics.

5.3.1 Objective Function I: Minimizing the Maximum Service Time

Let L_i be the load (in bits) to be transmitted over link i within a given time interval. The service time for link i, denoted by S_i, is L_i/R_i. A scheme that minimizes the maximum service time $S_{\max} \stackrel{\text{def}}{=} \max\{S_i,\ i \in I\}$ has the attractiveness of being easy to integrate in current wireless network standards. For example, the access point (AP) of an IEEE 802.11 WLAN or the Piconet controller of an IEEE 802.15.3 WPAN can utilize its polling medium access mechanism to measure the channel gains between the AP and each mobile node, and to probe nodes about their loads. Using channel gains and load values, the AP can compute the optimum powers and rates that minimize $S_{\max}$. A scheme that minimizes $S_{\max}$ does not require a user to receive any feedback from the AP while this user is transmitting, i.e., only one transceiver is required at a node. Furthermore, rate adaptation is carried out on a per-packet basis (i.e., the whole packet is transmitted at one rate), which is easy to integrate in current wireless networks standards [83].

To minimize the maximum service time, there are generally three constraints that must be satisfied. First, the transmission power at node i must not exceed a certain maximum power $P_{\max}^{(i)}$, which is set by the hardware. Second, the SNR of link i must be above the one specified in (5.4). The third constraint is that the transmission rate of node i (R_i) must be above a certain minimum rate $R_{\min}^{(i)}$, where $R_{\min}^{(i)} \geq 0$. This minimum-rate constraint is typically required for real-time traffic (e.g., voice over IP). For example, it would be required by the UMTS *conversational traffic class* that transports real-time voice or video over a WCDMA radio interface [52].

In this case, a group of active (voice) users share the spectrum using DS-CDMA. The powers and rates of these users can be periodically adapted on a per-frame basis (every 10 milliseconds). Note that our formulation accommodates the special case of $R_{\min}^{(i)} = 0$, which could result in dynamically changing the number of active users whenever the optimization is carried out (e.g., users that experience extremely bad channels may be forced to set their rates to zero).

Given the channel gains (h_{ij}) and the loads $L_i \; \forall i \in I$, the goal is to find the transmission powers and rates (i.e., $P_t^{(i)}$ and R_i, $\forall i \in I$) that minimize $S_{\max}$. Formally, this problem is stated as follows:

$$\begin{cases} \underset{\{R_i, P_t^{(i)},\ i \in I\}}{\text{minimize}} \left\{ \max_{i \in I} \frac{L_i}{R_i} \right\} & \\ \text{subject to:} & \\ \frac{h_{ii} P_t^{(i)}}{\sum\limits_{j \in I - \{i\}} h_{ji} P_t^{(j)} + P_{\text{thermal}}} \geq 2^{R_i/W} - 1, & \forall i \in I \\ P_t^{(i)} \leq P_{\max}^{(i)}, & \forall i \in I \\ R_i \geq R_{\min}^{(i)}, & \forall i \in I \end{cases} \tag{5.5}$$

Note that this formulation is applicable to both PTP and MultiPTP networks. We now proceed to solve this problem. Consider the following proposition.

Proposition 7 *The optimization problem in (5.5) can be modeled as a GGP, which can be transformed into a GP. This GP can be transformed into a nonlinear convex program.*

Proof: An optimization problem is a GGP if the objective function and the constraints are generalized posynomials[5]. The function in the first constraint in (5.5) is actually not a posynomial because of the exponential and the negative terms. However, its right-hand side (RHS) can be approximated as:

$$2^{\frac{R_i}{W}} - 1 = e^{\frac{R_i}{W'}} - 1 \approx \frac{\frac{R_i}{W'}}{1!} + \frac{\left(\frac{R_i}{W'}\right)^2}{2!} + \ldots + \frac{\left(\frac{R_i}{W'}\right)^m}{m!} \stackrel{\text{def}}{=} f\left(\frac{R_i}{W'}\right) \tag{5.6}$$

[5]See [18] for the definition of generalized posynomials.

where $W' = W/\ln 2$ and m is the approximation degree (an $m = 6$ was found to be accurate enough for this study). With simple algebraic manipulations, (5.5) can be expressed as:

$$\begin{cases} \underset{\{R_i, P_t^{(i)},\ i\in I\}}{\text{minimize}} \left\{ \max_{i\in I}\{L_i R_i^{-1}\} \right\} \\ \text{subject to:} \\ \left[\sum_{j\in I-\{i\}} h_{ji} P_t^{(j)} + P_{\text{thermal}} \right] \left[h_{ii} P_t^{(i)} \right]^{-1} f\left(\frac{R_i}{W'}\right) \le 1 \\ P_t^{(i)}/P_{\max}^{(i)} \le 1 \\ R_i^{-1} R_{\min}^{(i)} \le 1 \end{cases} \tag{5.7}$$

where the constraints in (5.7) are to be satisfied for all $i \in I$. Since $f(R_i/W')$ is a posynomial, the formulation in (5.7) is a GGP. However, in its current form, this GGP cannot be solved optimally and efficiently. Therefore, we make two transformations. The first one transforms the above GGP into a GP. To this end, we introduce a new auxiliary variable t such that $t \ge \frac{L_i}{R_i}$, $\forall i \in I$. With the introduction of t, (5.7) becomes:

$$\begin{cases} \underset{\{t, R_i, P_t^{(i)},\ i\in I\}}{\text{minimize}} \quad t \\ \text{subject to:} \\ L_i R_i^{-1} t^{-1} \le 1 \\ \left[\sum_{j\in I-\{i\}} h_{ji} P_t^{(j)} + P_{\text{thermal}} \right] \left[h_{ii} P_t^{(i)} \right]^{-1} f\left(\frac{R_i}{W'}\right) \le 1 \\ P_t^{(i)}/P_{\max}^{(i)} \le 1 \\ R_i^{-1} R_{\min}^{(i)} \le 1. \end{cases} \tag{5.8}$$

It is obvious that (5.7) and (5.8) are equivalent forms, meaning that the powers and rates that minimize t also minimize the objective function in (5.7). Formulation

(5.8) is an example of a GP, which can be easily transformed into a nonlinear convex program using a logarithmic change of variables [19]. Formally, let $z \stackrel{\text{def}}{=} \ln t$, $x_i \stackrel{\text{def}}{=} \ln P_t^{(i)}$, and $y_i \stackrel{\text{def}}{=} \ln R_i \ \forall i \in I$ (so that $t = e^z$, $P_t^{(i)} = e^{x_i}$, and $R_i = e^{y_i}$). Instead of minimizing the objective function t, we now minimize $\ln t$. Also, each constraint of the form $g \leq 1$ is changed to $\ln g < 0$. This results in the following (equivalent) optimization problem:

$$\begin{cases} \underset{\{z, x_i, y_i,\ i \in I\}}{\text{minimize}} \quad z \\ \text{subject to:} \\ \ln\left(L_i e^{-y_i} e^{-z}\right) \leq 0 \\ \ln\left(\left[\sum\limits_{j \in I - \{i\}} h_{ji} e^{x_j} + P_{\text{thermal}}\right] h_{ii}^{-1} e^{-x_i} f\left(\frac{e^{y_i}}{W'}\right)\right) \leq 0 \\ \ln\left(e^{x_i} / P_{\max}^{(i)}\right) \leq 0 \\ \ln\left(e^{-y_i} R_{\min}^{(i)}\right) \leq 0. \end{cases} \tag{5.9}$$

At first, the above formulation may look more complicated than (5.8). However, unlike (5.8), (5.9) is a *convex* optimization problem that can be solved efficiently (see [19] for more details). Once (5.9) is solved for x_i and y_i, $\forall i \in I$, the optimal power and rate allocation is simply given by $P_t^{(i)} = e^{x_i}$ and $R_i = e^{y_i}\ \forall i \in I$.

Proposition 7 applies to both PTP and MultiPTP networks. In the case of MultiPTP networks, the problem can be further simplified to allow for faster computation of the optimal solution. The following proposition enables the subsequent derivation of this solution.

Proposition 8 *The powers and rates that optimize (5.5) are such that the first constraint is satisfied with equality.*

Proof: See Appendix E.

In MultiPTP networks, the receiver is common to all transmitters, and so the channel gains h_{ji} and h_{ii} can be simply written as h_j and h_i, respectively. Hence, utilizing

Proposition 8, the optimal power and rate allocation in the case of MultiPTP networks must satisfy the following set of linear equations:

$$\frac{h_i P_t^{(i)}}{\sum\limits_{j \in I-\{i\}} h_j P_t^{(j)} + P_{\text{thermal}}} = f\left(\frac{R_i}{W'}\right), \quad \forall i \in I. \tag{5.10}$$

Using the same derivation methodology as in [105], (5.10) can be reduced to:

$$\sum_{j \in I} \frac{1}{\left(\frac{1}{f\left(\frac{R_j}{W'}\right)} + 1\right)} = 1 - \frac{P_{\text{thermal}}}{P_t^{(i)} h_i \left(\frac{1}{f\left(\frac{R_i}{W'}\right)} + 1\right)}, \forall i \in I. \tag{5.11}$$

By imposing the constraint $P_t^{(i)} < P_{\max}^{(i)}$ and noting that (5.11) is valid $\forall i \in I$, we get the following inequality:

$$\sum_{j \in I} \frac{1}{\left(\frac{1}{f\left(\frac{R_j}{W'}\right)} + 1\right)} \leq 1 - \frac{P_{\text{thermal}}}{\min\limits_{i \in I} \left[P_{\max}^{(i)} h_i \left(\frac{1}{f\left(\frac{R_i}{W'}\right)} + 1\right)\right]}. \tag{5.12}$$

This equation determines the feasibility of a set of rates, BER requirements, and maximum power constraints. We will use it to derive the optimal solution for (5.5). First, consider the following proposition:

Proposition 9 *The powers and rates that optimize (5.5) are such that* $\frac{L_i}{R_i} = \frac{L_j}{R_j}$ $\forall\ i, j \in I$.

Proof: *See Appendix F.*

This proposition says that, at the optimal solution to (5.5), all users have the same service time (S). Hence, $R_i = L_i/S\ \forall i \in I$. Accordingly, (5.12) can be written as:

$$\sum_{j \in I} \frac{1}{\left(\frac{1}{f\left(\frac{L_j/S}{W'}\right)} + 1\right)} \leq 1 - \frac{P_{\text{thermal}}}{\min\limits_{i \in I} \left[P_{\max}^{(i)} h_i \left(\frac{1}{f\left(\frac{L_i/S}{W'}\right)} + 1\right)\right]}. \tag{5.13}$$

The only unknown in this equation is S, and so it can be easily solved for the minimum S. Note that a unique solution always exists, since the left-hand side (LHS) of (5.13) is 0 at $S = \infty$, and it increases as S decreases, while the RHS is 1 at $S = \infty$, and it decreases as S decreases.

5.3.2 Objective Function II: Maximizing the Sum of Users Rates

The goal of this objective function is to maximize network throughput, subject to constraints on the BER, the maximum transmission power, and the minimum transmission rates. This function, which has been the focus of much previous research, requires fast rate adaptation; for the network to operate at the optimal point, whenever a user completes the transmission of a packet, all other transmitters must update their rates in the midst of transmitting their packets. This means that users must use intra-packet rate adaptation (i.e., different portions of the same packet must be transmitted at different rates). Furthermore, maximizing the sum of rates requires users to be able to receive feedback about their new rates while transmitting, which may necessitate the use of a multiple-transceiver architecture.

Note that the minimum-rate constraint, which has been overlooked in most previous studies, is crucial for multimedia networks. Without this constraint, some users may never be allowed to transmit, particularly if they experience a "bad" channel relative to other users (i.e., their channel gains are relatively small).

To maximize the sum of rates when nodes individually operate at the Shannon capacity, the power/rate optimization problem can be formulated as follows:

$$\begin{cases} \underset{\{R_i, P_t^{(i)},\ i\in I\}}{\text{maximize}} \quad \sum_{i\in I} R_i \\ \text{subject to:} \\ \dfrac{h_{ii}P_t^{(i)}}{\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\text{thermal}}} \geq 2^{R_i/W} - 1, \quad \forall i \in I \\ P_t^{(i)} \leq P_{\max}^{(i)}, \quad \forall i \in I \\ R_i \geq R_{\min}^{(i)}, \quad \forall i \in I. \end{cases} \tag{5.14}$$

Unfortunately, in its current form, this objective function cannot be transformed into the minimization of a posynomial, as was done in the previous section. So it is not possible to formulate this problem as a GGP, a GP, or a nonlinear convex program. In fact, this form of the problem exhibits an unknown number of local maxima, and there are no efficient algorithms to solve it optimally. Therefore, we follow a

different approach that will produce an approximate solution to the problem. First, consider the following result.

Proposition 10 *The $P_t^{(i)}$'s and R_i's that optimize (5.14) are such that the first constraint is satisfied with equality.*

Proof: The proof is along the lines of the one for Proposition 8, and is omitted for brevity.

Using this proposition, the first constraint in (5.14) can be integrated into the objective function, thus reducing the optimization variables to include only the transmission powers. Thus, (5.14) is reduced to:

$$\begin{cases} \underset{\{P_t^{(i)},\ i\in I\}}{\text{maximize}} & \displaystyle\sum_{i\in I} W\ \log_2\ \left(\frac{h_{ii}P_t^{(i)}}{\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)}+P_{\text{thermal}}}+1\right) \\ \text{subject to:} & \\ P_t^{(i)} \le P_{\max}^{(i)} & \\ \lambda_i\ \left[\displaystyle\sum_{j\in I-\{i\}} h_{ji}P_t^{(j)}+P_{\text{thermal}}\right]\left[h_{ii}P_t^{(i)}\right]^{-1} \le 1 & \end{cases} \tag{5.15}$$

where $\lambda_i \stackrel{\text{def}}{=} 2^{R_{\min}^{(i)}/W}-1$ and the constraints are to be satisfied for all $i\in I$. Noting that $\sum_i \log_2 x_i = \log_2 \prod_i x_i$, that the maximization of x is equivalent to the minimization of $-x$, and that the set of values that minimize $\alpha \log_2 \prod_i f(x_i)$ are the same set of values that minimize $\prod_i f(x_i)$, where $\alpha>0$, (5.15) can be put into the following *equivalent* form:

$$\begin{cases} \underset{\{P_t^{(i)},\ i\in I\}}{\text{minimize}} & \displaystyle\prod_{i\in I}\left(\frac{h_{ii}P_t^{(i)}}{\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)}+P_{\text{thermal}}}+1\right)^{-1} \\ \text{subject to:} & \\ P_t^{(i)} \le P_{\max}^{(i)} & \\ \lambda_i\ \left[\displaystyle\sum_{j\in I-\{i\}} h_{ji}P_t^{(j)}+P_{\text{thermal}}\right]\left[h_{ii}P_t^{(i)}\right]^{-1} \le 1. & \end{cases} \tag{5.16}$$

Note that (5.14) and (5.16) are equivalent. Similar to (5.14), (5.16) cannot be formulated as a GGP or a GP. To avoid the possibility of converging to a highly suboptimal solution, we consider a slightly different objective function from the one in (5.16). For the modified objective function, we are able to formulate the problem as a GGP. The modified objective function is:

$$\underset{\{P_t^{(i)},\ i\in I\}}{\text{minimize}} \quad \prod_{i\in I}\left(\frac{h_{ii}P_t^{(i)}}{\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\text{thermal}}}\right)^{-1} \tag{5.17}$$

where the constraints are the same as in (5.16). Let $\mathbf{P}_t^o$ denote the transmission power vector that minimizes (5.17) subject to the associated constraints. While we do not claim that $\mathbf{P}_t^o$ is optimal for (5.16), due to the functional similarities between (5.17) and the objective function in (5.16), $\mathbf{P}_t^o$ provides a sub-optimal solution to (5.16). It is noteworthy to mention that the same type of approximation has been used in many other papers (e.g., [96],[62], and [25]), but for different setups. For example, the work in [62] studied the rate maximization for wireless systems that use MQAM modulation. Our work, on the other hand, focuses on the maximum sum-of-rates that can be achieved irrespective of what the modulation is. The modified (approximating) optimization problem is given by:

$$\begin{cases} \underset{\{P_t^{(i)},\ i\in I\}}{\text{minimize}} \quad \prod\limits_{i\in I}\left(\left[h_{ii}P_t^{(i)}\right]^{-1}\left[\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\text{thermal}}\right]\right) \\ \text{subject to:} \\ P_t^{(i)}/P_{\max}^{(i)} \le 1 \\ \lambda_i \left[\sum\limits_{j\in I-\{i\}} h_{ji}P_t^{(j)} + P_{\text{thermal}}\right]\left[h_{ii}P_t^{(i)}\right]^{-1} \le 1. \end{cases} \tag{5.18}$$

The formulation in (5.18) is a GGP, and can be solved optimally and efficiently after transforming it into a convex problem using a logarithmic change of variables,

similar to the one described in the previous section. In Section 5.4, we show the approximate maximum achievable network throughput, and contrast that with what two practical adaptation VPG systems can achieve.

5.4 Numerical Results

In this section, we use the analysis of the previous sections to compute the best network performance (BEST) obtained under optimal adaptation of system parameters. We contrast BEST with the performance of two VPG schemes that are based on the modulation and coding of BASIC and ESYS (described in Section 5.2). For both schemes, the transmission power and data rate are adapted, but the coding and modulation schemes/rates are kept fixed (no FEC coding is used in BASIC). The idea of comparing with these two VPG systems is to study how well existing systems perform when compared to the benchmark derived in this chapter. Without loss of generality, we focus on MultiPTP networks in this section. Note that the analysis in this chapter applies to both PTP and MultiPTP networks. However, the exact optimal solution of VPG is only known for the MultiPTP case [57]. Our results are based on numerical experiments conducted using MATLAB. Our performance metrics include the service time (S)[6], the sum of users rates, and the average energy consumption per bit (E_b), defined as $\frac{\sum_{i=1}^{N} P_i}{\sum_{i=1}^{N} R_i}$. Note that E_b is a more significant measure than the average transmission power. In fact, it is misleading to compare the average transmission power of two systems that transmit at different data rates, as the cost of transmitting a certain number of bits depends on both the transmission power and rate. To simulate the channel gains, we use a two-ray propagation model with a path loss factor of 4. The total bandwidth of the system is $W = 1$ MHz, and $P_{\text{thermal}} = -169$ dBm. For simplicity, we assume that nodes have the same $P_{\max} = 20$ dBm and the same $R_{\min}$.

We consider N transmitting nodes that are randomly placed within a square area

[6] At the optimal solution for the first objective function, all users have the *same* service time S (see Proposition 9).

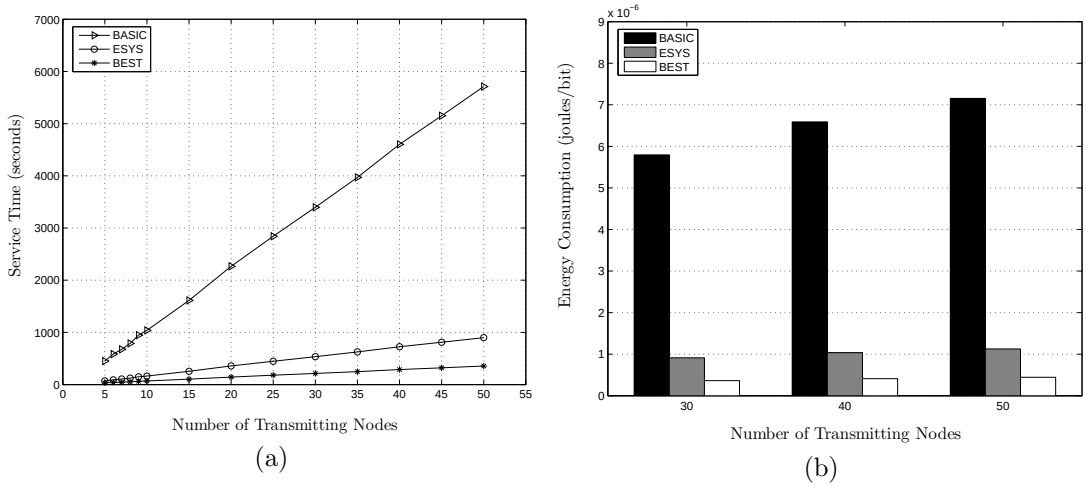


Figure 5.3: Performance of BEST, ESYS, and BASIC based on the minimization of the maximum service time in MultiPTP networks.

of length 200 meters. The common receiver is placed at the center of the square. Given the locations of the N transmitters, the channel attenuation matrix H is computed. Whenever the solution set is empty for the generated H (i.e., $R_{\min}$ cannot be satisfied for at least one user), a new set of transmitters is randomly generated. The results are obtained based on the average of 100 independent realizations of the matrix H.

For the first objective function, the workload at each transmitter is selected randomly between 1 and 20 Mbits. Since ESYS uses a coding rate of 1/3, $R_{\max} = W/3$. For this system, we set $R_{\min} = R_{\max}/100$. Figure 5.3 depicts the performance under the first objective function. Part (a) of the figure depicts the service time S versus N. It is shown that for all values of N, ESYS achieves a considerably lower S than BASIC. But the minimum S that can be achieved (i.e., under BEST) is still much lower than S under ESYS. For example, when $N = 50$, S under ESYS is about 2.5 times the achievable minimum service time. It is also shown that for all schemes, S increases with N. This is expected since as N increases, the MAI also increases, and users are forced to transmit at lower rates, which increases their service times.

Part (b) of Figure 5.3 depicts E_b versus N. It shows that although ESYS reduces

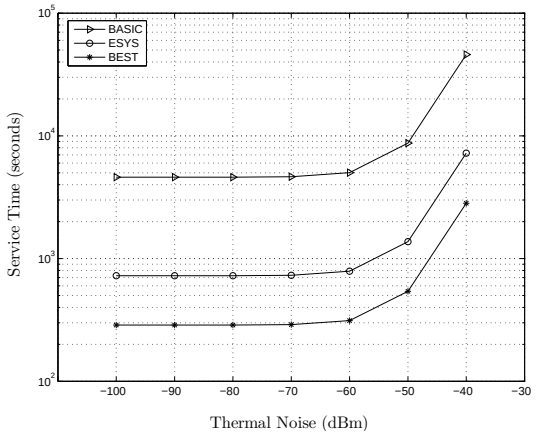


Figure 5.4: Service time S in MultiPTP networks as a function of P_{thermal}.

the energy consumption by more than 84% relative to BASIC for $N = 50$, its consumption is about 2.5 times the minimum that can be achieved.

Next, we study the impact of increasing P_{thermal}. Figure 5.4 shows S vs. P_{thermal} for $N = 40$. The workload is generated as in Figure 5.3. For all values of P_{thermal}, it can be observed that even with a strong FEC code (i.e., ESYS), VPG systems stay at about 2.5 times the minimum achievable service time. For all schemes, however, S starts to increase exponentially when P_{thermal} exceeds -60 dBm. The reason is that at this value, P_{thermal} becomes comparable to the maximum received powers for bad-channel users. Hence, the SNR of the users deteriorates significantly, causing a fast drop in their rates and a corresponding dramatic increase in S.

In the case of the second objective function (i.e., maximizing the sum of users rates), we let $R_{\max} = W/30$ providing a 10 dB processing gain relative to FEC-coded bits at maximum rate. As before, $R_{\min} = R_{\max}/100$. Part (a) of Figure 5.5 depicts the throughput performance versus N. It is shown that ESYS saturates at about half the number of users at which BEST saturates. It can also be seen that at $N = 50$, the capacity of the channel is about 2.45 times the capacity of ESYS, which itself is about 5.18 times the capacity of BASIC.

An interesting observation can be made here. The difference in the achievable rates between ESYS and BEST for the throughput objective function is about

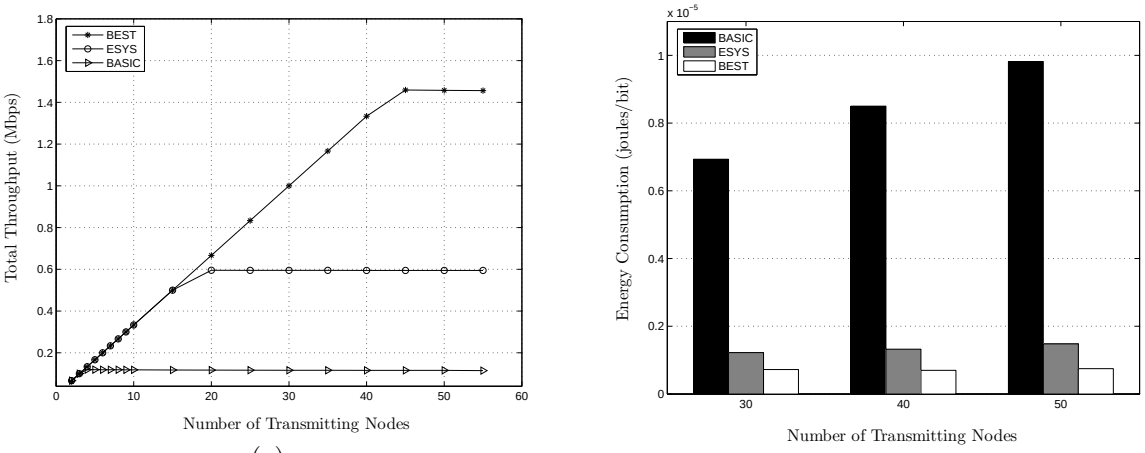


Figure 5.5: Performance versus N under the throughput maximization criterion for MultiPTP networks.

3.9 dB. This difference is very close to the one in the case of minimizing the maximum service (≈ 4 dB).

Part (b) of Figure 5.5 depicts the energy consumption of BASIC, ESYS, and BEST as a function of N. As in the case of minimizing the service time, although ESYS reduces the energy consumption by more than 85% relative to BASIC for $N = 50$, its consumption is about twice the minimum that can be achieved.

Finally,we study the impact of increasing P_{thermal} on the throughput. Figure 5.6 shows the throughput as a function of P_{thermal} for $N = 40$. It is observed that the performance of BASIC and ESYS is relatively constant with respect to BEST up to $P_{\text{thermal}} = -60$. After that, P_{thermal} becomes significant, causing a fast drop in users rates. Note that for $P_{\text{thermal}} = -40$ dBm, the throughput point for BASIC is not reported as there does not exist a solution that satisfies the $R_{\min}$ requirements of all users.

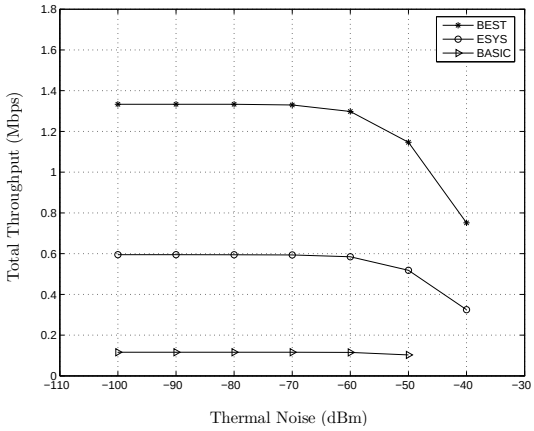


Figure 5.6: Impact of P_{thermal} on the total throughput under the second objective function.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

In Chapter 2, we proposed a CDMA-based power controlled MAC protocol for wireless ad hoc networks. This protocol, called CA-CDMA, accounts for the multiple access interference, thereby solving the near-far problem that undermines the throughput performance in MANETs. CA-CDMA uses channel-gain information obtained from overheard RTS and CTS packets over an out-of-band control channel to dynamically bound the transmission power of mobile terminals in the vicinity of a receiver. It adjusts the required transmission power for data packets to allow for interference-limited simultaneous transmissions to take place in the neighborhood of a receiving terminal. We compared the performance of our protocol with that of the IEEE 802.11 scheme. Our simulation results showed that CA-CDMA can improve the network throughput by up to 280% and, at the same time, achieve 50% reduction in the energy consumed to successfully deliver a packet from the source to the destination. To the best of our knowledge, CA-CDMA is the first protocol to provide a solution to the near-far problem in CDMA ad hoc systems at the protocol level.

In Chapter 3, we proposed a power controlled MAC protocol for MANETs, known as POWMAC. Similar to the 802.11 scheme, POWMAC is based on a single-transceiver circuitry, and it operates over a single channel for data and control packets. POWMAC adjusts the transmission powers of data packets to allow for some interference margin at the receiver. Multiple interference-limited transmissions in the vicinity of a receiver are allowed to overlap in time, provided that their MAI effects do not lead to collisions at nearby receivers. We compared the performance of POWMAC with that of the IEEE 802.11 scheme. Our simulation results showed

that POWMAC can improve the network throughput by up to 50% in random-grid topologies and much more than that in clustered topologies. Furthermore, POWMAC can achieve some reduction in the energy consumed to successfully deliver a packet from the source to the destination. To the best of our knowledge, POWMAC is the first single-channel protocol that utilizes TPC to increase network throughput while preserving the collision avoidance property of the 802.11 scheme.

In Chapter 4, we investigated the potential performance gains of using adaptive orthogonal modulation (AOM) in multirate CDMA networks. We showed that, relative to a variable processing gain (VPG) system that uses fixed orthogonal modulation (OM) order, AOM can significantly increase the network throughput while simultaneously reducing energy consumption. We studied the problem of optimal joint rate/power control for AOM-based systems under two objective functions: minimizing the maximum service time and maximizing the sum of users rates. For the first objective function, we showed that the optimization problem can be formulated as a GGP, which can be transformed into a nonlinear convex program, and be solved optimally and efficiently. In the case of the second objective function, we obtained a lower bound on the achievable gain of AOM over fixed-modulation schemes. Unlike previous work on adaptive transmission, which have focused mainly on cellular networks, ours is applicable to both PTP and MultiPTP networks. In PTP networks, our results show that, when compared with fixed OM order VPG schemes, AOM can achieve more than 50% improvement in the service time and, simultaneously, more than 40% reduction in energy consumption. In MultiPTP networks, we derived a simple algorithm for finding the optimal powers and rates for VPG, and explained the intuition behind using that algorithm as a heuristic for AOM. Our results show that the achievable throughout gain can be up to 30% compared to VPG. Furthermore, AOM achieves more than 45% reduction in the service time relative to VPG.

In Chapter 5, We investigated the achievable network performance under optimal link-layer adaptation. The derived performance bounds serve as benchmarks against which adaptive CDMA systems can be compared, and are applicable to

both point-to-point and multipoint-to-point networks. By assuming that each user individually operates at the Shannon capacity, we studied the problem of optimal joint rate/power control under two objective functions: minimizing the maximum service time and maximizing the sum of users rates. For the first objective function, we showed that the optimization problem can be modeled as a GGP, which can be transformed into a nonlinear convex program and solved optimally and efficiently. We also proved that the at the optimal solution, all users have the same service time, and used that to derive a compact formula for the MultiPTP case. In the case of maximizing the sum of rates, we defined an approximating objective function for which we were able to compute the optimal solution. The results show that relative to a multiuser VPG system that uses rate 1/3 convolutional code with constraint-length $K = 41$, the best achievable *network* performance is about 4 dB higher. This difference is quite close to the one between the Shannon capacity and the performance of that convolutional code in the *single-link* case.

6.2 Future Work

Our future work will focus on other capacity optimizations such as the use of directional and/or multiple-input multiple-output (MIMO) in CDMA-based protocols. While transmitting, a directional antenna concentrates the power in a certain direction with less interference to other directions. Furthermore, while receiving, a directional antenna has a greater sensitivity for electromagnetic radiation in a certain direction, thus it is less sensitive to interference from other directions. These preferred directions of transmission and reception reduce MAI, and thus, larger number of users can be accommodated and the system capacity is increased. We will also explore the potential of integrating CDMA into the design of the POWMAC protocol. POWMAC enjoys a simple receiver design and extending its use to CDMA seems to be natural.

Adaptive orthogonal modulation (AOM) is still a newly explored area of research. Several challenges remain to be addressed, including finding the optimal solution for

maximizing the sum of rates for AOM in MultiPTP networks, the optimal algorithm for maximizing the sum of rates for VPG and AOM in PTP networks, and closed-form approximations to the optimal solutions. In addition to solving for these theoretical limits, our future work will focus on how to integrate these algorithms within current wireless networks protocols. Furthermore, in our analysis, we let the modulation order M to take real positive value. However, in reality, M is restricted to a finite set. Our future work will focus on studying the impact of restricting M to a finite set of values.

The work presented in this dissertation provided insights for improving the spatial reuse and bandwidth efficacy of the limited wireless spectrum. However, as we discussed above, adaptive transmission in wireless networks remain a relatively new research area. While the "digital communication" research community has significantly improved the single-link performance (less than 1 dB away from Shannon's capacity), the "wireless networking" community still needs maybe decades of research to improve the overall system performance; this dissertation provides few forward steps in that direction.

APPENDIX

A. Geometric Programming

Let $x_1, \ldots, x_n$ be n variables in $\mathbb{R}^+$, and let $\mathbf{x} \stackrel{\text{def}}{=} (x_1, \ldots, x_n)$. A function f is called a *posynomial* in $\mathbf{x}$ if it can be written in the form $f(x_1, \ldots, x_n) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \ldots x_n^{a_{nk}}$, where $c_k \geq 0$ and $a_{ik} \in \mathbb{R}$. If $K = 1$, then f is called a *monomial* function. A GP is an optimization problem of the form [19]:

$$\begin{cases} \underset{\{x_i,\ i \in I\}}{\text{minimize}} \quad f_0(\mathbf{x}) & \\ \text{subject to:} & \\ f_i(\mathbf{x}) \leq 1, & i = 1, \ldots, v \\ g_i(\mathbf{x}) = 1, & i = 1, \ldots, u \\ x_i \geq 0, & i = 1, \ldots, n \end{cases} \tag{6.1}$$

where $f_0, \ldots, f_v$ are posynomial functions and $g_1, \ldots, g_u$ are monomial functions. A function f is a *generalized posynomial* if it can be formed by the addition, multiplication, positive (fractional) power, or maximum of posynomials [18]. A GGP is an optimization problem of the form (6.1), where $f_0, \ldots, f_v$ are generalized posynomial functions and $g_1, \ldots, g_u$ are monomial functions.

B. Proof of Proposition 2

This proof is by contradiction. Denote the assigned powers and rates by the vectors $\mathbf{P}_t$, and $\mathbf{R}$, where $\mathbf{P}_t \stackrel{\text{def}}{=} (P_t^{(1)}, \ldots, P_t^{(N)})$ and $\mathbf{R} \stackrel{\text{def}}{=} (R_1, \ldots, R_N)$. Let $(\mathbf{P}_t^o, \mathbf{R}^o)$ be the optimal power and rate allocation that optimize (4.7), i.e., $\max_{i \in I} \left\{ \frac{L_i}{R_i^o} \right\} \leq \max_{i \in I} \left\{ \frac{L_i}{R_i} \right\}$ for any feasible $(\mathbf{P}_t, \mathbf{R})$. Given $(\mathbf{P}_t^o, \mathbf{R}^o)$, suppose that one of the equalities, e.g., the mth link, in the first constraint in (4.7) is not satisfied, i.e.,

$$\frac{h_{mm} P_t^{o(m)}}{\sum_{j \in I - \{m\}} h_{jm} P_t^{o(j)} + P_{\text{thermal}}} > \frac{f(R_m^o)}{W}. \tag{6.2}$$

The LHS of the first constraint in (4.7) is a strictly increasing function of $P_t^{(i)}$ and is a strictly decreasing function of $P_t^{(j)}$ for $j \neq i$, while the RHS is a strictly increasing

function of R_i. Hence, there must be some power decrement $-\triangle P < 0$ for link m and some rate increment $\triangle R > 0$ for all the links, that makes the allocation $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, where $\mathbf{P}_t^{o'} = (P_t^{o(1)}, \ldots, P_t^{o(m-1)}, P_t^{o(m)} - \triangle P, P_t^{o(m+1)}, \ldots, P_t^{o(N)})$ and $\mathbf{R}^{o'} = (R_1^o + \triangle R, \ldots, R_N^o + \triangle R)$, a feasible solution to (4.7). That is, the following inequalities are still satisfied under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$:

$$\frac{h_{mm}(P_t^{o(m)} - \triangle P)}{\sum_{j \in I - \{m\}} h_{jm} P_t^{o(j)} + P_{\text{thermal}}} \geq \frac{f(R_m^o + \triangle R)}{W}, \tag{6.3}$$

$$\frac{h_{ii} P_t^{o(i)}}{\sum_{j \in I - \{i\} - \{m\}} h_{ji} P_t^{o(j)} + h_{mi}(P_t^{o(m)} - \triangle P) + P_{\text{thermal}}} \geq \frac{f(R_i^o + \triangle R)}{W} \quad \forall i \in I - \{m\}. \tag{6.4}$$

Under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, we have $\max_{i \in I} \left\{ \frac{L_i}{R_i^o + \triangle R} \right\} < \max_{i \in I} \left\{ \frac{L_i}{R_i^o} \right\}$. This is a contradiction to the optimality assumption that $\max_{i \in I} \left\{ \frac{L_i}{R_i^o} \right\} \leq \max_{i \in I} \left\{ \frac{L_i}{R_i} \right\}$ for any feasible $(\mathbf{P}_t, \mathbf{R})$. Therefore, the assumption that there is a link that does not satisfy the equality of the first constraint in (4.7) can not be true.

C. Proof of Proposition 3

This proof is by contradiction. Denote the optimal powers and rates by the vectors $\mathbf{P}_t^o$, and $\mathbf{R}^o$, respectively, i.e., $\max_{i \in I} \left\{ \frac{L_i}{R_i^o} \right\} \leq \max_{i \in I} \left\{ \frac{L_i}{R_i} \right\}$ for any feasible $(\mathbf{P}_t, \mathbf{R})$. Suppose $\max_{i \in I} \left\{ \frac{L_i}{R_i^o} \right\} > \min_{i \in I} \left\{ \frac{L_i}{R_i^o} \right\}$. Assume that the transmission time of the mth link is the minimum among all users, i.e., $\frac{L_m}{R_m^o} = \min_{i \in I} \left\{ \frac{L_i}{R_i^o} \right\}$. The LHS of the first constraint in (4.7) is a strictly increasing function of $P_t^{(i)}$ and is a strictly decreasing function of $P_t^{(j)}$ for $j \neq i$, while the RHS is a strictly increasing function of R_i. Hence, there must be some power decrement $-\triangle P < 0$ and some small rate decrement $-\triangle R_1 < 0$ for link m, and some small rate increment $\triangle R_2 > 0$ for all the other users, that makes the allocation $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, where $\mathbf{P}_t^{o'} = (P_t^{o(1)}, \ldots, P_t^{o(m-1)}, P_t^{o(m)} - \triangle P, P_t^{o(m+1)}, \ldots, P_t^{o(N)})$ and $\mathbf{R}^{o'} = (R_1^o + \triangle R_2, \ldots, R_{m-1}^o + \triangle R_2, R_m^o - \triangle R_1, R_{m+1}^o + \triangle R_2, \ldots, R_N^o + \triangle R_2)$, a

feasible solution to (4.7). That is, under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, the following inequalities are still satisfied:

$$\frac{h_{mm}(P_t^{o(m)} - \triangle P)}{\sum_{j\in I-\{m\}} h_{jm}P_t^{o(j)} + P_{\text{thermal}}} \geq \frac{f(R_m^o - \triangle R_1)}{W}, \tag{6.5}$$

$$\frac{h_{ii}P_t^{o(i)}}{\sum_{j\in I-\{i\}-\{m\}} h_{ji}P_t^{o(j)} + h_{mi}(P_t^{o(m)} - \triangle P) + P_{\text{thermal}}} \geq \frac{f(R_i^o + \triangle R_2)}{W} \quad \forall i \in I - \{m\}. \tag{6.6}$$

The small rate variations in $\mathbf{R}^{o'}$ is in the sense that $\frac{L_m}{R_m^o - \triangle R_1} \leq \max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o + \triangle R_2}\right\}$. Under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, we have $\max_{i\in I}\left\{\frac{L_i}{R_i^o}\right\} = \max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o}\right\} > \max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o + \triangle R_2}\right\} = \max\left\{\max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o + \triangle R_2}\right\}, \frac{L_m}{R_m^o - \triangle R_1}\right\}$. This is a contradiction to the optimality assumption that $\max_{i\in I}\left\{\frac{L_i}{R_i^o}\right\} \leq \max_{i\in I}\left\{\frac{L_i}{R_i}\right\}$ for any feasible $(\mathbf{P}_t, \mathbf{R})$. Therefore, it must be that $\max_{i\in I}\left\{\frac{L_i}{R_i^o}\right\} = \min_{i\in I}\left\{\frac{L_i}{R_i^o}\right\}$.

D. Proof of Proposition 5

Consider two networks A and B. Network A has v_1 elements in I_{v_1}, v_2 (where $v_2 > 1$) elements in I_{v_2}, one element in U, and v_3 elements in I_{v_3}, where I_{v_1}, I_{v_2}, U, and I_{v_3} are as defined in Section 4.3.2. Such allocation of powers and rates adheres to the optimal structure proved in [57]. Network B has v_1 elements in I_{v_1}, only one element in the set $I_{v_2}^{'} \stackrel{\text{def}}{=} \{I_{v_2} \bigcup U\}$, and v_3 elements in I_{v_3}. Our goal is to show that network B always has a higher throughput than A.

Consider network B. Utilizing (4.17), and assuming that P_{thermal} is small compared to the MAI [1], the throughput of this network is:

$$T_B = v_1 R_{\max} + v_3 R_{\min} + \frac{W}{\mu_{\text{req}}} \frac{h_u P_u}{v_1 P_{r_B}^{v_1} + v_3 P_{r_B}^{v_3}}, \tag{6.7}$$

where h_u and P_u are the channel gain and transmission power of the single element of $I_{v_2}^{'}$, $P_{r_B}^{v_1}$ is the received power of elements in I_{v_1}, $P_{r_B}^{v_3}$ and is the received power of the elements in I_{v_3}.

[1] This assumption is quite reasonable in CDMA networks [132].

For network A, the transmission powers are different, and so is the MAI. Therefore, users in I_{v_1} must increase or decrease their powers so as to attain their rates ($R_{\max}$) and to fulfill their BER constraints. The new received powers $P_{r_A}^{v_1}$ must be such that the SNR of users in I_{v_1} in network A is the same as their SNR in network B. Similarly, users in I_{v_3} must increase or decrease their powers to $P_{r_A}^{v_3}$ to maintain the same SNR. For the SNR of users in I_{v_1} to stay the same, the following must hold:

$$\frac{P_{r_A}^{v_1}}{(v_1-1)P_{r_A}^{v_1}+v_3P_{r_A}^{v_3}+\sum\limits_{j\in I'_{v_2}}h_jP_j}= \frac{P_{r_B}^{v_1}}{(v_1-1)P_{r_B}^{v_1}+v_3P_{r_B}^{v_3}+h_uP_{\mathrm{u}}}, \tag{6.8}$$

It is quite easy to verify that if $P_{r_A}^{v_1}=\alpha P_{r_B}^{v_1}$ and $P_{r_A}^{v_3}=\alpha P_{r_B}^{v_3}$, where $\alpha=\sum_{j\in I'_{v_2}}h_jP_j/h_uP_{\mathrm{u}}$, then (6.8) will be satisfied. It can also be shown that the same value of α results in equal SNR of users in I_{v_3} for networks A and B. Having decided the values of $P_{r_A}^{v_1}$ and $P_{r_A}^{v_3}$, we are now able to express the throughput of network A as:

$$T_A = v_1R_{\max}+v_3R_{\min}+ \frac{W}{\mu_{\mathrm{req}}}\sum_{i\in I'_{v_2}}\frac{h_iP_i}{v_1P_{r_A}^{v_1}+v_3P_{r_A}^{v_3}+\sum\limits_{j\in I'_{v_2}-\{i\}}h_iP_j} \tag{6.9}$$

The first two terms in (6.9) and (6.7) are equal, so to determine whether T_A is bigger or lesser than T_B, we only need to consider the last term in each equation. To this end, we divide the last term in (6.9) by the one in (6.7), and use $P_{r_A}^{v_1}=\alpha P_{r_B}^{v_1}$ and $P_{r_A}^{v_3}=\alpha P_{r_B}^{v_3}$; thus, after some manipulation, we obtain:

$$\sum_{i\in I'_{v_2}}\frac{\left[v_1P_{r_B}^{v_1}+v_3P_{r_B}^{v_3}\right]h_iP_i}{\left[v_1P_{r_B}^{v_1}+v_3P_{r_B}^{v_3}+h_uP_{\mathrm{u}}\frac{\sum_{j\in I'_{v_2}-\{i\}}h_iP_j}{\sum_{j\in I'_{v_2}}h_jP_j}\right]\sum_{j\in I'_{v_2}}h_jP_j} < \sum_{i\in I'_{v_2}}\frac{h_iP_i}{\sum_{j\in I'_{v_2}}h_jP_j}=\frac{\sum_{i\in I'_{v_2}}h_iP_i}{\sum_{j\in I'_{v_2}}h_jP_j}=1$$

Thus, $T_A<T_B$. So far, we have shown that there is only one user that is operating at power P_u that is higher than the minimum power required to achieve $R_{\min}$, but we have not shown *which* user

is that. It is not difficult to see that T_B in (6.7) is an increasing function of the received power (i.e., $h_u P_u$). Hence, the best channel user in $I - \{I_{v_1}\}$ must be chosen to operate at P_u, so the order of that user is $v_1 + 1$.

E. Proof of Proposition 8

The proof is by contradiction. Denote the assigned power and rate vectors by $\mathbf{P}_t$ and $\mathbf{R}$, where $\mathbf{P}_t \stackrel{\text{def}}{=} (P_t^{(1)}, \ldots, P_t^{(N)})$ and $\mathbf{R} \stackrel{\text{def}}{=} (R_1, \ldots, R_N)$. Let $(\mathbf{P}_t^o, \mathbf{R}^o)$ be the optimal power and rate vectors that optimize (4.7). Given $(\mathbf{P}_t^o, \mathbf{R}^o)$, suppose that one of the equalities in the first constraint in (4.7) is not satisfied, i.e.,

$$\frac{h_{mm} P_t^{o(m)}}{\sum_{j \in I - \{m\}} h_{jm} P_t^{o(j)} + P_{\text{thermal}}} > f\left(\frac{R_m^o}{W'}\right) \tag{6.10}$$

for some $m \in I$. The LHS of the first constraint in (4.7) is strictly increasing in $P_t^{(i)}$ and strictly decreasing in $P_t^{(j)}$ for $j \neq i$, while the RHS is strictly increasing in R_i. Hence, there must be some power decrement $-\triangle P < 0$ for link m and some rate increment $\triangle R > 0$ for all the links that make the allocation $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$ a feasible solution to (4.7), where $\mathbf{P}_t^{o'} = (P_t^{o(1)}, \ldots, P_t^{o(m-1)}, P_t^{o(m)} - \triangle P, P_t^{o(m+1)}, \ldots, P_t^{o(N)})$ and $\mathbf{R}^{o'} = (R_1^o + \triangle R, \ldots, R_N^o + \triangle R)$. That is, the following inequalities are still satisfied under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$:

$$\frac{h_{mm}(P_t^{o(m)} - \triangle P)}{\sum_{j \in I - \{m\}} h_{jm} P_t^{o(j)} + P_{\text{thermal}}} \geq f\left(\frac{R_m^o + \triangle R}{W'}\right) \tag{6.11}$$

$$\frac{h_{ii} P_t^{o(i)}}{\sum_{j \in I - \{i\} - \{m\}} h_{ji} P_t^{o(j)} + h_{mi}(P_t^{o(m)} - \triangle P) + P_{\text{thermal}}} \geq f\left(\frac{R_i^o + \triangle R}{W'}\right) \quad \forall i \in I - \{m\} \tag{6.12}$$

Under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, we have $\max_{i \in I} \left\{\frac{L_i}{R_i^o + \triangle R}\right\} < \max_{i \in I} \left\{\frac{L_i}{R_i^o}\right\}$. This is a contradiction to the optimality assumption that $\max_{i \in I} \left\{\frac{L_i}{R_i^o}\right\} \leq \max_{i \in I} \left\{\frac{L_i}{R_i}\right\}$ for any feasible $(\mathbf{P}_t, \mathbf{R})$. Therefore, the assumption that there is a link that does not satisfy the equality of the first constraint in (4.7) can not be true.

F. Proof of Proposition 9

The proof is by contradiction. Let $\mathbf{P}_t^o$ and $\mathbf{R}^o$ denote the optimal power and rate vectors, respectively. Suppose $\max_{i\in I}\left\{\frac{L_i}{R_i^o}\right\} > \min_{i\in I}\left\{\frac{L_i}{R_i^o}\right\}$. Assume that the transmission time of the mth link is the minimum among all users, i.e., $\frac{L_m}{R_m^o} = \min_{i\in I}\left\{\frac{L_i}{R_i^o}\right\}$. The LHS of the first constraint in (4.7) is a strictly increasing function of $P_t^{(i)}$ and is a strictly decreasing function of $P_t^{(j)}$ for $j \neq i$, while the RHS is a strictly increasing function of R_i. Hence, there must be some power decrement $-\triangle P < 0$, some rate decrement $-\triangle R_1 < 0$ for link m, and some rate increment $\triangle R_2 > 0$ for all the other users that make the allocation $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$ a feasible solution to (4.7), where $\mathbf{P}_t^{o'} = (P_t^{o(1)}, \ldots, P_t^{o(m-1)}, P_t^{o(m)} - \triangle P, P_t^{o(m+1)}, \ldots, P_t^{o(N)})$ and $\mathbf{R}^{o'} = (R_1^o + \triangle R_2, \ldots, R_{m-1}^o + \triangle R_2, R_m^o - \triangle R_1, R_{m+1}^o + \triangle R_2, \ldots, R_N^o + \triangle R_2)$. That is, under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, the following inequalities are still satisfied:

$$\frac{h_{mm}(P_t^{o(m)} - \triangle P)}{\sum_{j\in I-\{m\}} h_{jm} P_t^{o(j)} + P_{\text{thermal}}} \geq f\left(\frac{R_m^o - \triangle R_1}{W'}\right) \tag{6.13}$$

$$\frac{h_{ii} P_t^{o(i)}}{\sum_{j\in I-\{i\}-\{m\}} h_{ji} P_t^{o(j)} + h_{mi}(P_t^{o(m)} - \triangle P) + P_{\text{thermal}}} \geq f\left(\frac{R_i^o + \triangle R_2}{W'}\right) \quad \forall i \in I-\{m\}.$$

The small rate variations in $\mathbf{R}^{o'}$ is in the sense that $\frac{L_m}{R_m^o - \triangle R_1} \leq \max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o + \triangle R_2}\right\}$. Under $(\mathbf{P}_t^{o'}, \mathbf{R}^{o'})$, we have $\max_{i\in I}\left\{\frac{L_i}{R_i^o}\right\} = \max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o}\right\} > \max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o + \triangle R_2}\right\} = \max\left\{\max_{i\in I-\{m\}}\left\{\frac{L_i}{R_i^o + \triangle R_2}\right\}, \frac{L_m}{R_m^o - \triangle R_1}\right\}$. This is a contradiction to the optimality assumption that $\max_{i\in I}\left\{\frac{L_i}{R_i^o}\right\} \leq \max_{i\in I}\left\{\frac{L_i}{R_i}\right\}$ for any feasible $(\mathbf{P}_t, \mathbf{R})$. Therefore, it must be that $\max_{i\in I}\left\{\frac{L_i}{R_i^o}\right\} = \min_{i\in I}\left\{\frac{L_i}{R_i^o}\right\}$.

E. Notation Guide

The following tables summarize the notations used in Chapter 2 through 5.

$P_j^{(i)}$	Reception power of j's signal at node i
$\mu^{(i)}$	Bit energy-to-noise spectral density ratio at node i
W	Processing gain
μ^*	Target energy-to-noise spectral density ratio
P_{thermal}	Thermal noise power
$P_{\text{MAI}}^{(i)}$	Total multiple access interference at receiver i
$\xi^{(i)}$	Noise rise
ξ_{max}	Maximum planned noise rise
n	Path loss exponent
$P_{\text{CA-CDMA}}$	Transmission power in CA-CDMA
$P_{802.11}$	Transmission power in 802.11
$R_{\text{CA-CDMA}}$	Bit rate in CA-CDMA
$R_{802.11}$	Bit rate in 802.11
k	Constant related to the antenna parameters
$P_{\text{noise}}^{(i)}$	Additional noise power that each of the neighbors can add to terminal i without impacting i's current reception
P_{max}	Maximum transmission power
$P_{\text{map}}^{(j)}$	Maximum allowable power level that terminal j can use that will not disturb any ongoing reception in j's neighborhood
G_{ji}	Channel gain between nodes j and i
$P_{\text{min}}^{(ji)}$	Minimum transmission power required for node i to correctly decode the data packet if transmitted by node j
$P_{\text{MAI-current}}^{(i)}$)	Current MAI from all ongoing (interfering) transmissions at node i
$P_{\text{allowed}}^{(ji)}$	Power that terminal j is allowed to use to send to i according to the link budget
$P_{\text{MAI-future}}^{(i)}$	Interference power tolerance that i can endure from *future* unintended transmitters
$K^{(i)}$	Number of terminals in the vicinity of i that are to share $P_{\text{MAI-future}}^{(i)}$
α	Factor to account for interference received from nodes outside the transmission range
M	Number of nodes
λ	Packet generation rate per node
p	Probability of communication with a different cluster

Table 6.1: Table of notation for Chapter 2.

P_j	Transmission power of node j
μ^*	Target energy-to-noise spectral density ratio
P_{thermal}	Thermal noise power
$\xi^{(i)}$	Noise rise
$P_{\text{MAI}}^{(i)}$	Total multiple access interference at receiver i
$\xi_{\max}$	Maximum planned noise rise
n	Path loss exponent
$\mu^{(i)}$	Bit energy-to-noise spectral density ratio at node i
W	Processing gain
$P_{\min}^{(ji)}$	Minimum transmission power required for node i to correctly decode the data packet if transmitted by node j
P_{POWMAC}	Transmission power in CA-CDMA
$P_{802.11}$	Transmission power in 802.11
k	Constant related to the antenna parameters
$\pi_i(u)$	Maximum TP that terminal i can use without disturbing u's reception
$P_{\text{MAP}}^{(j)}$	Maximum allowable power level that terminal j can use that will not disturb any ongoing reception in j's neighborhood
G_{ji}	Channel gain between nodes j and i
$\Psi(i)$	Set of terminals in i's vicinity whose receptions overlap with i's transmission
$P_{\text{MAI-add}}^{(i)}$	Maximum additional interference power that node i can endure from future unintended transmitters so that the SINR at i does not drop below μ^*
$P_{\text{MTI}}^{(i)}$	Maximum tolerable interference that a *single* future interferer can add to terminal i
$P_{\max}$	Maximum transmission power
$N_{AW}^{(i)}$	Number of terminals in the vicinity of i that are to share $P_{\text{MAI-add}}^{(i)}$
ϱ	Factor to account for interference received from nodes outside the transmission range
$\tau_{Ack}^{(ji)}$	Delay before sending the ACK
α	Predetermined System parameter
β	Predetermined System parameter
p_r	Loss probability experienced by terminal r
x_r	Access probability of terminal r
f_m	Maximum Doppler shift
T_c	Channel coherence time
L_c	Total length (in bits) of the IEEE 802.11 RTS plus CTS packets
L_d	Average data packet length
R_d	Data packets transmission rate
R_c	Control packets transmission rate
T_{POWMAC}	Duration of time it takes to send N data packets according to POWMAC
$T_{802.11}$	Duration of time it takes to send N data packets according to 802.11
M	Number of nodes

Table 6.2: Table of notation for Chapter 3.

$P_t^{(i)}$	Transmission power of node i
h_{ij}	Channel gain between nodes j and i
P_{thermal}	Thermal noise power
$P_{\max}$	Maximum transmission power
$R^{(i)}$	Data rate at node i
W	Fourier bandwidth of the channel
$\mu^{(i)}$	Bit energy-to-noise spectral density ratio at node i
μ_{req}	Target energy-to-noise spectral density ratio
E_b/N_0	Bit-energy-to-noise spectral density ratio
M	Orthogonal modulation order
k	$\log_2 M$
R_m	Modulated bit rate
R_c	Coded bit rate
Z	Maximum allowed modulated bit rate
P_b	Probability of bit error
L_i	Load (in bits) to be transmitted over link i
I	Set of active links in the network
S_i	Service time of node i
N	Number of transmitting nodes
$R_{\min}$	Minimum transmission rate
$R_{\max}$	Maximum transmission rate

Table 6.3: Table of notation for Chapters 4 and 5.

REFERENCES

[1] 3GPP TR 25.848 V4.0.0, Physical layer aspects of UTRA high speed downlink packet access (release 4), 2001.

[2] International Standard ISO/IEC 8802-11; ANSI/IEEE Std 802.11, 1999 Edn. Part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications.

[3] Mesquite Software Incorporation, http://www.mesquite.com.

[4] The Cisco Aironet 350 Series of wireless LAN, http://www.cisco.com/warp/public/cc/pd/witc/ao350ap.

[5] The CMU Monarch Project., The CMU monarch projects wireless and mobility extensions to NS, http://www.monarch.cs.cmu.edu/cmuns. html.

[6] The WiMedia Alliance, http://www.wimedia.org.

[7] A. Acharya, A. Misra, and S. Bansal. MACA-P : a MAC for concurrent transmissions in multi-hop wireless networks. In *Proceedings of the First IEEE PerCom 2003 Conference*, pages 505–508, Mar. 2003.

[8] S. Agarwal, R. H. Katz, S. V. Krishnamurthy, and S. k. Dao. Distributed power control in ad-hoc wireless networks. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 2, pages 59–66, Oct. 2001.

[9] N. U. Ahmed and K. R. Rao. *Orthogonal Transforms for Digital Signal Processing*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1975.

[10] S. M. Alamouti and S. Kallel. Adaptive trellis-coded multiple-phase-shift keying for rayleigh fading channels. *IEEE Transactions on Communications*, 42(6):2305–2314, 1994.

[11] M. Alouini and A. Goldsmith. Adaptive modulation over nakagami fading channels. *Journal on Wireless Personal Communications, Kluwer Academic Publishers*, 13:119–143, 2000.

[12] C. Andren. Short PN Sequences for Direct Sequence Spread Spectrum Radios, Harris Semiconductor, http://www.sss-mag.com/pdf/shortpn.pdf.

[13] K. Aretz, M. Haardt, W. Konhauser, and W. Mohr. The future of wireless communications beyond the third generation. *Computer Networks*, 37:83–92, 2001.

[14] N. Bambos. Toward power-sensitive networks architecture in wireless communications: Concepts, issues, and design aspects. In *IEEE Personal Communications Magazine*, volume 5, pages 50–59, June 1998.

[15] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, and A. Viterbi. CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users. *IEEE Communications Magazine*, 38(7):70–77, July 2000.

[16] A. A. Bertossi and M. A. Bonuccelli. Code assignment for hidden terminal interference avoidance in multihop packet radio networks. *IEEE Transactions on Communications*, 3(4):441–449, Aug. 1995.

[17] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. MACAW: A media access protocol for wireless LANs. In *Proceedings of the ACM SIGCOMM Conference*, volume 24, pages 212–225, 1994.

[18] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. Technical report, Stanford University, Dept. of Electrical Engineering, Sept. 2004. http://www.stanford.edu/∼boyd/gp_tutorial.html.

[19] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[20] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proceedings of the ACM MobiCom Conference*, pages 85–97, 1998.

[21] R. Caceres and L. Iftode. Improving the performance of reliable transport protocols in mobile computing environment. *IEEE Journal on Selected Areas in Communications*, 13:850–857, 1995.

[22] J. K. Cavers. Variable-rate transmission for rayleigh fading channels. *IEEE Transactions on Communications*, 20(1):15–22, 1972.

[23] D. Chase. Code combining–a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets. *IEEE Transactions on Communications*, 33(5):385 – 393, May 1985.

[24] M.-S. Chen and R. Boorstyn. Throughput analysis of code division multiple access (CDMA) multihop packet radio networks in the presence of noise. In *Proceedings of the IEEE INFOCOM Conference*, pages 310–316, 1985.

[25] M. Chiang and J. Bell. Balancing supply and demand of bandwidth in wireless cellular networks: utility maximization over powers and rates. In *Proceedings of the IEEE INFOCOM Conference*, volume 4, pages 2800–2811, 2004.

[26] W. Choi and J. Y. Kim. Forward-link capacity of a DS/CDMA system with mixed multirate sources. *IEEE Transactions on Vehicular Technology*, 50(3):737–749, 2001.

[27] I. Cidon and M. Sidi. Distributed assignment algorithms for multihop packet radio networks. *IEEE Transactions on Computers*, 38(10):1353–1361, Oct. 1989.

[28] T. M. Cover and J. A. Thomas. *Eelements of information theory*. John Wiley, 1991.

[29] R. L. Cruz and A. V. Santhanam. Optimal link scheduling and power control in CDMA multihop wireless networks. In *Proceedings of the IEEE GLOBECOM Conference*, volume 1, pages 52–56, 2002.

[30] R. L. Cruz and A. V. Santhanam. Optimal routing, link scheduling and power control in multihop wireless networks. In *Proceedings of the IEEE INFOCOM Conference*, volume 1, pages 702–711, 2003.

[31] J. Deng and Z. Haas. Dual busy tone multiple access (DBTMA): A new medium access control for packet radio networks. In *Proceedings of the IEEE ICUPC Conference*, pages 973–977, 1998.

[32] O. deSouza, M.-S. Chen, and R. R. Boorstyn. A comparison of the performance of protocols in pakcet radio networks. In *Proceedings of the IEEE MILCOM Conference*, volume 2, pages 455–460, Oct 85.

[33] T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad-hoc networks. In *Proceedings of the IEEE INFOCOM Conference*, pages 976–984, 2002.

[34] T. A. ElBatt, S. V. Krishnamurthy, D. Connors, and S. Dao. Power management for throughput enhancement in wireless ad-hoc networks. In *Proceedings of the IEEE ICC Conference*, pages 1506–1513, 2000.

[35] A. Ephremides and T. V.Truong. Scheduling broadcasts in multihop radio networks. *IEEE Transactions on Communications*, 38(4):456–460, Apr. 1990.

[36] H. Fattah and C. Leung. Erlang capacity of a CDMA cellular system with multirate sources. In *Proceedings of the IEEE Vehicular Tech. Conference*, volume 5, pages 3483–3487, 2003.

[37] P. K. Frenger, P. Orten, T. Ottosson, and A. B. Svensson. Rate-compatible convolutional codes for multirate ds-cdma systems. *IEEE Transactions on Communications*, 47(6), July 1999.

[38] J. Garcia-Luna-Aceves and J. Raju. Distributed assignment of codes for multihop packet-radio networks. In *Proceedings of the IEEE MILCOM Conference*, volume 1, pages 450–454, 1997.

[39] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatley III. On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology*, 40(2):303–312, May 1991.

[40] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[41] A. Goldsmith and S. Chua. Adaptive coded modulation for fading channels. *IEEE Transactions on Communications*, 46(5):595–602, 1998.

[42] A. Goldsmith and P. Varaiya. Increasing spectral efficiency through power control. In *Proceedings of the IEEE ICC Conference*, pages 600 – 604, 1993.

[43] A. J. Goldsmith and S. Chua. Variable-rate variable-power MQAM for fading channels. *IEEE Transactions on Communications*, 45(10):1218–1230, Oct. 1997.

[44] J. Gomez, A. T. Campbell, M. Naghshineh, and C. Bisdikian. PARO: supporting dynamic power controlled routing in wireless ad hoc networks. *ACM/Kluwer Journal on Wireless Networks*, 9(5):443–460, 2003.

[45] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, Mar. 2000.

[46] D. Haccoun and Z. Gherbi. On the application of very low rate error control coding to CDMA. In *IEEE Canadian Conference on Electrical and Computer Engineering*, volume 2, pages 466 – 469, 1997.

[47] J. Hagenauer. Rate-compatible punctured convolutional codes (RCPC codes) and their applications. *IEEE Transactions on Communications*, 36(4):389 – 400, Apr. 1988.

[48] S. V. Hanly and D. N. C. Tse. Multiaccess fading channels-Part II: Delay-limited capacities. *IEEE Transactions on Information Theory*, 44(7):2816 – 2831, Nov. 1998.

[49] J. F. Hayes. Adaptive feedback communications. *IEEE Transactions on Communications*, 16:29–34, Feb 1968.

[50] K. Higuchi, A. Fujiwara, and M. Sawahashi. Multipath interference canceller for high-speed packet transmission with adaptive modulation and coding scheme in W-CDMA forward link. *IEEE Journal on Selected Areas in Communications*, 20(2):419–43, Feb. 2002.

[51] G. Holland, N. Vaidya, and P. Bahl. A rate-adaptive mac protocol for multihop wireless networks. In *Proceedings of the ACM MobiCom Conference*, pages 236–251, 2001.

[52] H. Holma and A. Toskala, editors. *WCDMA for UMTS*. Wiley, John & Sons, Incorporated, 2002.

[53] L. Hu. Distributed code assignments for CDMA packet radio networks. *IEEE/ACM Transactions on Networking*, 1:668–677, Dec. 1993.

[54] K.-W. Hung and T.-S. Yum. The coded tone sense protocol for multihop spread-spectrum packet radio networks. In *Proceedings of the IEEE GLOBECOM Conference*, volume 2, pages 712–716, 1989.

[55] C.-L. I and K. K. Sabnani. Variable spreading gain cdma with adaptive control for true packet switching wireless network. In *Proceedings of the IEEE ICC Conference*, volume 2, pages 725–730, 1995.

[56] C.-D. Iskander and P. T. Mathiopoulos. Performance of multicode DS/CDMA with noncoherent M-ary orthogonal modulation in multipath fading channels. *IEEE Transactions on Wireless Communications*, 3(1):209–223, Jan. 2004.

[57] S. A. Jafar and A. Goldsmith. Adaptive multirate CDMA for uplink throughput maximization. *IEEE Transactions on Wireless Communications*, 2(2):218–228, Mar. 2003.

[58] L. M. Jalloul and J. M. Holtzman. Performance analysis of DS/CDMA with noncoherent m-ary orthogonal modulation in multipath fading channels. *IEEE Journal on Selected Areas in Communications*, 12:862–870, 1994.

[59] N. Jindal and A. Goldsmith. Capacity and optimal power allocation for fading broadcast channels with minimum rates. *IEEE Transactions on Information Theory*, 49(11):2895–2909, Nov. 2003.

[60] M. Joa-Ng and I.-T. Lu. Spread spectrum medium access protocol with collision avoidance in mobile ad-hoc wireless network. In *Proceedings of the IEEE INFOCOM Conference*, volume 2, pages 776–783, 1999.

[61] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark. Scenario-based performance analysis of routing protocols for mobile ad-hoc

networks. In *Proceedings of the ACM MobiCom Conference*, pages 195–206, 1999.

[62] D. Julian, M. Chiang, D. O'Neill, and S. Boyd. QoS and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks. In *Proceedings of the IEEE INFOCOM Conference*, volume 2, pages 477–486, 2002.

[63] E.-S. Jung and N. H. Vaidya. A power control MAC protocol for ad hoc networks. In *Proceedings of the ACM MobiCom Conference*, pages 36–47, 2002.

[64] Y. Kamio, S. Sampei, H. Sasaoka, and N. Morinaga. Performance of modulation-level-controlled adaptive-modulation under limited transmission delay time for land mobile communications. In *Proceedings of the IEEE Vehicular Tech. Conference*, volume 1, pages 221–225, 1995.

[65] S. Kandukuri and N. Bambos. MDMA in wireless packet networks. In *Proceedings of the IEEE INFOCOM Conference*, 2001.

[66] S. Kandukuri and S. Boyd. Simultaneous rate and power control in multirate multimedia CDMA systems. In *IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications*, volume 2, pages 570–574, 2000.

[67] P. Karn. MACA - a new channel access method for packet radio. In *Proceedings of the 9th ARRL Computer Networking Conference*, pages 134–140, 1990.

[68] V. Kawadia and P. R. Kumar. Power control and clustering in ad hoc networks. In *Proceedings of the IEEE INFOCOM Conference*, pages 459–469, 2003.

[69] R. Kerr. CDMA digital cellular: An ASIC overview. In *Appl. Microwave and Wireless*, pages 30–41, Fall 1993.

[70] K. I. Kim. On the error probability of a DS/SSMA system with a noncoherent M-ary orthogonal modulation. In *Proceedings of the IEEE Vehicular Tech. Conference*, volume 1, pages 482–485, 1992.

[71] S. W. Kim and Y. H. Lee. Combined rate and power adaptation in DS/CDMA communications over Nakagami fading channels. *IEEE Transactions on Communications*, 48(1):162–168, Jan. 2000.

[72] T. J. Kwon and M. Gerla. Clustering with power control. In *Proceedings of the IEEE MILCOM Conference*, pages 1424–1428, 1999.

[73] S. W. Lee and D. H. Cho. Distributed reservation CDMA for wireless lan. In *Proceedings of the IEEE GLOBECOM Conference*, volume 1, pages 360–364, 1995.

[74] S. Lefrancois and D. Haccoun. Very low rate quasi-optimal convolutional codes for CDMA. In *IEEE Canadian Conference on Electrical and Computer Engineering*, volume 1, pages 210 – 213, Sept. 1994.

[75] J. Li, C. Blake, D. S. Couto, H. I. Lee, and R. Morris. Capacity of ad hoc wireless networks. In *Proceedings of the ACM MobiCom Conference*, pages 61–69, 2001.

[76] L. Lifang and A. Goldsmith. Capacity and optimal resource allocation for fading broadcast channels-Part I. Ergodic capacity. *IEEE Transactions on Information Theory*, 47(3):1083–1102, Mar. 2001.

[77] L. Lifang and A. Goldsmith. Capacity and optimal resource allocation for fading broadcast channels-Part II. Outage capacity. *IEEE Transactions on Information Theory*, 47(3):1103–1127, Mar. 2001.

[78] Z. Lin and A. Svensson. New rate-compatible repetition convolutional codes. *IEEE Transactions on Information Theory*, 46(7):2651 – 2659, Nov. 2000.

[79] H. Matsuoka, S. Sampei, N. Morinaga, and Y. Kamio. Adaptive modulation system with variable coding rate concatenated code for high quality multi-media communication systems. In *Proceedings of the IEEE Vehicular Tech. Conference*, volume 1, pages 487–491, 1996.

[80] J. Monks, V. Bharghavan, and W.-M. Hwu. A power controlled multiple access protocol for wireless packet networks. In *Proceedings of the IEEE INFOCOM Conference*, pages 219–228, 2001.

[81] A. Muqattash and M. Krunz. Power controlled dual channel (PCDC) medium access protocol for wireless ad hoc networks. In *Proceedings of the IEEE INFOCOM Conference*, pages 470–480, 2003.

[82] A. Muqattash, M. Krunz, and T. Shu. Performance enhancement of adaptive orthogonal modulation in wireless CDMA systems (extended version). Technical report, University of Arizona, Dept. of Electrical & Computer Engineering, Jan. 2005. http://www.ece.arizona.edu/∼alaa/aom_tr05.pdf.

[83] S. Nanda, K. Balachandran, and S. Kumar. Adaptation techniques in wireless packet data services. *IEEE Communications Magazine*, 38(1):54–64, Jan. 2000.

[84] T. Nandagopal, T.-E. Kim, X. Gao, and V. Bharghavan. Achieving mac layer fairness in wireless packet networks. In *Proceedings of the ACM MobiCom Conference*, pages 87–98, 2000.

[85] S. Narayanaswamy, V. Kawadia, R. S. Sreenivas, and P. R. Kumar. Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the COMPOW protocol. In *Proceedings of the European Wireless Conference*, pages 156–162, 2002.

[86] S.-J. Oh and K. M. Wasserman. Optimality of greedy power control and variable spreading gain in multi-class CDMA mobile networks. In *Proceedings of the ACM MobiCom Conference*, pages 102–112, 1999.

[87] T. Ojanperä and R. Prasad. *Wideband CDMA For Third Generation Mobile Communications*. Artech House, Incorporated, 1998.

[88] J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen. Topology control for wireless sensor networks. In *Proceedings of the ACM MobiCom Conference*, pages 286–299, 2003.

[89] V. K. Paulrajan, J. A. Roberts, and D. L. Machamer. Capacity of a CDMA cellular system with variable user data rate. In *Proceedings of the IEEE GLOBECOM Conference*, volume 3, pages 1458–1462, 1996.

[90] D. R. Pauluzzi and N. C. Beaulieu. A comparison of SNR estimation techniques for the AWGN channel. *IEEE Transactions on Communications*, 48:Pages:1681–1691, 2000.

[91] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. Theory of spread spectrum communications-a tutorial. *IEEE Transactions on Communications*, 30:855–884, May 1982.

[92] J. G. Proakis. *Digital Communications*. McGraw-Hill, Inc., 2001.

[93] M. B. Pursley. Performance evaluation for phase-coded spread-spectrum multiple-access commuincation-part 1: System analysis. *IEEE Transactions on Communications*, 25(8):795–799, Aug. 1977.

[94] M. B. Pursley, H. B. Russell, and J. S. Wysocarski. Energy-efficient transmission and routing protocols for wireless multiple-hop networks and spread spectrum radios. In *Proceedings of the EUROCOMM Conference*, pages 1–5, 2000.

[95] D. Qiao, S. Choi, A. Jain, and K. G. Shin. Miser: an optimal low-energy transmission strategy for IEEE 802.11a/h. In *Proceedings of the 9th Annual*

International Conference on Mobile Computing and Networking, pages 161–175, 2003.

[96] X. Qiu and K. Chawla. On the performance of adaptive modulation in cellular systems. *IEEE Transactions on Communications*, 47:884–895, 1999.

[97] D. C. R. Jain, W. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. Technical report DEC-TR-301, Ohio State University, Dept. of Computer Science and Engineering, Sept. 1984. http://www.cis.ohio-state.edu/∼jain/papers/fairness.htm.

[98] R. k. Morrow, Jr. and J. S. Lehnert. Packet throughput in slotted ALOHA DS/SSMA radio systems with random signature sequences. *IEEE Transactions on Communications*, 40(7):1223–1230, July 1992.

[99] R. Ramanathan and R. Rosales-Hain. Topology control of multihop wireless networks using transmit power adjustment. In *Proceedings of the IEEE INFOCOM Conference*, pages 404–413, 2000.

[100] T. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, 2002.

[101] T. S. Rappaport, A. Annamalai, R. M. Buehrer, and W. H. Tranter. Wireless communications: past events and a future perspective. *IEEE Communications Magazine*, 40(5):148–161, May 2002.

[102] T. S. Rappaport and L. B. Milstein. Effects of radio propagation path loss on DS-CDMA cellular frequency reuse efficiency for the reverse channel. *IEEE Transactions on Vehicular Technology*, 41:231–242, Aug. 1992.

[103] V. Rodoplu and T. Meng. Minimum energy mobile wireless networks. *IEEE Journal on Selected Areas in Communications*, 17(8):1333–1344, 1999.

[104] V. Rodoplu and T. Meng. Position based CDMA with multiuser detection (P-CDMA/MUD) for wireless ad hoc networks. In *IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications*, volume 1, pages 336–340, 2000.

[105] A. Sampath, P. S. Kumar, and J. M. Holtzman. Power control and resource management for a multimedia CDMA wireless system. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 1, pages 21–25, 1995.

[106] H. D. Schotten, H. Elders-Boll, and A. Busboom. Adaptive multi-rate multi-code CDMA systems. In *Proceedings of the IEEE Vehicular Tech. Conference*, volume 2, pages 782–785, 1998.

[107] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt. *Spread Spectrum Communications Handbook*. McGraw-Hill, Inc., 1994.

[108] E. Sousa and J. A. Silvester. Spreading code protocols for distributed spread-spectrum packet radio networks. *IEEE Transactions on Communications*, 36(3):272–281, Mar. 1988.

[109] E. Sousa and J. A. Silvester. Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network. *IEEE Journal on Selected Areas in Communications*, 8(5):762–771, June 1990.

[110] S. Tantaratana and K. M. Ahmed, editors. *Wireless Applications of Spread Spectrum Systems: Selected Readings*. IEEE Press, 1998.

[111] 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*, 23(12):1417–1433, 1975.

[112] D. N. C. Tse and S. V. Hanly. Multiaccess fading channels: I. Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Transactions on Information Theory*, 44(7):2796–2815, Nov. 1998.

[113] D. N. C. Tse and S. V. Hanly. Multiaccess fading channels-Part I: Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Transactions on Information Theory*, 44(7):2796 – 2815, Nov. 1998.

[114] A. Viterbi. *CDMA Principles of Spread Spectrum Communication*. Addison Wesley, 1995.

[115] A. J. Viterbi. Very low rate convolution codes for maximum theoretical performance of spread-spectrum multiple-access channels. *IEEE Journal on Selected Areas in Communications*, 8(4):641–649, 1990.

[116] A. J. Viterbi. Spread spectrum communications: myths and realities. *IEEE Communications Magazine*, 40(34-41), 2002.

[117] A. M. Viterbi and A. J. Viterbi. Erlang capacity of a power controlled CDMA system. *IEEE Journal on Selected Areas in Communications*, 11(6):892–900, 1993.

[118] B. Vucetic. An adaptive coding scheme for time-varying channels. *IEEE Transactions on Communications*, 39(5):653–663, 1991.

[119] M. Wang and R. Kohno. A wireless multimedia DS-CDMA network based on adaptive transmission rate/power control. In *Proceedings of the International Zurich Seminar on Broadband Communications, Accessing, Transmission, Networking*, pages 45–50, 1998.

[120] X. Wang and K. Kar. Distributed algorithms for max-min fair rate allocation in aloha networks. In *Proceedings of Annual Allerton Conference, Urbana-Champaign, USA*, Oct. 2004.

[121] C. Ware, J. Chicharo, and T. Wysocki. Simulation of capture behaviour in IEEE 802.11 radio modems. In *Proceedings of the IEEE Vehicular Tech. Conference*, volume 3, pages 1393–1397, Fall 2001.

[122] R. Wattenhofer, L. Li, P. Bahl, and Y.-M. Wang. Distributed topology control for power efficient operation in multihop wireless ad hoc networks. In *Proceedings of the IEEE INFOCOM Conference*, pages 1388–1397, 2001.

[123] W. T. Webb and R. Steele. Variable rate QAM for mobile radio. *IEEE Transactions on Communications*, 43(7):2223–2230, 1995.

[124] J. E. Wieselthier, A. Ephremides, and J. A. B. Tarr. A distributed reservation-based CDMA protocol that does not require feedback information. *IEEE Transactions on Communications*, 36(8):913–923, Augut 1988.

[125] S.-L. Wu, C.-Y. Lin, Y.-C. Tseng, and J.-P. Sheu. A new multi-channel mac protocol with on-demand channel assignment for multi-hop mobile ad hoc networks. In *Proceedings of the International Symposium on Parallel Architectures, Algorithms and Networks*, pages 232–237, 2000.

[126] S.-L. Wu, Y.-C. Tseng, and J.-P. Sheu. Intelligent medium access for mobile ad hoc networks with busy tones and power control. *IEEE Journal on Selected Areas in Communications*, 18(9):1647–1657, 2000.

[127] T. H. Wu and E. Geraniotis. CDMA with multiple chip rates for multi-media communications". In *Proceedings of the Conference on Information Sciences and Systems*, pages 992–997, 1994.

[128] K. Xu, M. Gerla, and S. Bae. How effective is the IEEE 802.11 RTS/CTS handshake in ad hoc networks? In *Proceedings of the IEEE GLOBECOM Conference*, volume 1, pages 72–76, Nov. 2002.

[129] L. Xu, X. Shen, and J. W. Mark. Performance analysis of adaptive rate and power control for data service in DS-CDMA systems. In *Proceedings of the IEEE GLOBECOM Conference*, volume 1, pages 627–631, 2001.

[130] C.-H. Yeh. ROAD: a variable-radius MAC protocol for ad hoc wireless networks. In *Proceedings of the IEEE Vehicular Tech. Conference*, volume 1, pages 399–403, 2002.

[131] T. B. Zahariadis. Migration toward 4G wireless communications. *IEEE Wireless Communications*, 11(3):6–7, 2004.

[132] J. Zander. Distributed cochannel interference control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41:305–311, 1992.

[133] H. Zhou, C.-H. Yeh, and H. Mouftah. A solution for multiple access in variable-radius mobile ad hoc networks. In *Proceedings of the IEEE International Conference on Communications, Circuits and Systems*, volume 1, pages 150–154, 2002.

[134] J. Zhu and G. Marubayashi. Properties and application of parallel combinatory SS communication system. In *Proceedings of the IEEE International Symposium on Spread Spectrum Techniques and Applications*, pages 227–230, 1992.