MULTI-LAYER OPTIMIZATION IN WIRELESS AD HOC NETWORKS

BY XIANGFANG LI

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Electrical and Computer Engineering

Written under the direction of

Prof. Zoran Gajic

and approved by

____

____

____

____

New Brunswick, New Jersey

October, 2007

ABSTRACT OF THE DISSERTATION

Multi-Layer Optimization in Wireless Ad Hoc Networks

by Xiangfang Li

Dissertation Director: Prof. Zoran Gajic

The growing prevalence of wireless ad hoc networks calls for an innovative design to support Quality of Service (QoS) while maintaining high energy efficiency and bandwidth efficiency. In this dissertation, a multi-layer optimization approach is developed in view of benefits and necessities of sharing information among different layers. Specifically, given the traffic demands with QoS requirements, how to jointly design physical, MAC and network layers to optimize the network performance is considered in this dissertation.

Firstly, a joint power control and maximally disjoint multipath routing scheme is proposed for QoS provisioning of end-to-end traffic with minimum rate constraint. A framework of power control with QoS constraints is introduced and both centralized and distributed solutions are derived. It is demonstrated by simulations that the proposed scheme provides high energy efficiency and the prolonged network lifetime, as well as robustness when augmented with a dynamic traffic monitoring and switching mechanism.

In order to fulfill the QoS requirement at the link layer, TD/CDMA has been chosen as the MAC scheme due to its support for a high network throughput in a multihop environment. The multi-link versions of proportional fair and throughput optimal scheduling algorithms are proposed for multihop wireless ad hoc networks. In addition, a generic token counter mechanism is employed to satisfy the minimum and maximum rate requirements. Approximative algorithms are suggested to reduce the computational complexity. In networks that lack centralized control, distributed scheduling algorithms are derived and fully distributed implementations are provided. Simulation results demonstrate the effectiveness of the proposed schemes.

In order to further improve bandwidth efficiency, cognitive radio is considered for more efficient spatial and temporal spectrum sharing. Specifically, we consider the scenario where a cognitive radio ad hoc network is formed by low power personal/portable devices operating simultaneously in the same frequency band along with a legacy system. A power control problem is formulated to maximize the energy efficiency of the ad hoc network, as well as to guarantee QoS for both legacy network users and ad hoc network users. The results show that cognitive radio greatly improves bandwidth efficiency of wireless ad hoc networks.

Acknowledgements

First, I would like to express my deep gratitude to my advisor Prof. Zoran Gajic for his patience and encouragement during my study at Rutgers University. I would also like to thank Prof. David Daut, Prof. Ivan Marsic from CAIP, and Dr. Muhidin Lelic from United Techologies Research Center for their valuable advice and helpful suggestions.

It has been my pleasure to work and study at the Department of Electrical and Computer Engineering (ECE) and the Wireless Information Network Laborartory (WINLAB) at Rutgers University. The academic environment has provided me great opportunities to perform my research and build my knowledge base. I sincerely acknowledge the help and asisitance offered by the faculty and the staff of the ECE Department and WINLAB.

This research work was supported in part by the National Science Foundation under grant ANIR-0106857, I would like to thank the National Science Foundation for supporting this research.

I also would like to mention the support received from my family, without them, this work would not have been accomplished. My lovely daughter Iris was born during this period, her sweet smiles gave me strength to keep on working. My mother's help permitted me to have the time necessary to do this research. Finally and most importantly, my husband's encouragement and help was a constant source of inspiration.

Dedication

To my daughter, Iris.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABR	Associativity-Based Routing
ABR	Avaliable Bit Rate
AODV	Ad hoc On-demand Distance Vector
AODV-BR	AOAV with Backup Routes
AOMDV	Ad hoc On-demand Multi-path Distance Vector
ARQ	Automatic Repeat Request
ATM	Asynchronous Transfer Mode
AWGN	Additive white Gaussian Noise
BER	Bit Error Rate
BESMR	Balanced Energy Split Multi-Path Routing
BS	Base Station
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
CDF	Cumulative Distribution Function
CR	Cognitive Radio
CGSR	Clusterhead Gateway Switch Routing
CR	Cognitive Radio
CPC	Common Power Control
CSMA	Carrier Sense Multiple Access
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
CSOPC	Constrained Second-Order Power Control
CTS	Clear to Send
DCPC	Distributed Constrained Power Control
DCF	Distributed Coordination Function
DERR	Distributed Elastic Round Robin
DFS	Distributed Fair Scheduling
DPC	Distributed Power Control
DS-CDMA	Direct Sequence Code Division Multiple Access
DSDV	Destination-Sequenced Distance Vector Routing
DSR	Dynamic Source Routing
EVDO	Evolution Data Optimized
FAMA	Floor Acquistion Multiple Access
FCC	Federal Communication Commission
FDMA	Frequency Division Multiple Access
FDD	Frequency Division Duplexing
FER	Frame Error Rate
GPS	Global Positioning System
IPC	Individual Power Control

IETF	The Internet Engineering Task Force
IP	Internet Protocol
LLC	Logical Link Control
MAC	Medium Access Control
MACA	Multiple Access Collision Avoidance
MACA-BI	Multiple Access Collision Avoidance By Invitation
MACAW	MACA for Wireless
MANET	Mobile Ad Hoc Networks
MGMR	Multi-link Gradient Algorithm with Minimum and Maximum Rate
MIMO	Multiple-Input Multiple-Output
MMT	Multi-link Maximum Throughput
MMTMR	Multi-link Maximum Throughput with Minimum and Maximum Rate
MPF	Multi-link Proportional Fair
MPFMR	Multi-link Proportional Fair with Minimum and Maximum Rate
MPSMR	Minimum Power Split Multi-Path Routing
MQR	Multi-link Throughput Optimal
MQRMR	Multi-link Throughput Optimal with Minimum and Maximum Rate
MS	Mobiel Station
OLSR	Optimized Link State Routing
PCF	Point Coordination Function
PDA	Personal Digital Assistant
PF	Proportional Fair
QoS	Quality of Service
RF	Radio Frequency
RTS	Request to Send
RRM	Radio Resource Management
RREQ	Route Request
RREP	Route Reply
SIR	Signal to Interference Ratio
SINR	Signal to Interference plus Noise Ratio
SMR	Split Multipath Routing
SOR	Successive Overrelaxation Iteration
TBRPF	Topology Broadcast based on Reverse-Path Forwarding
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TDD	Time Division Duplexing
TORA	Temporally Ordered Routing Algorithm
UAV	Unmanned Airborne Vehicles
VBR	Variable Bit Rate
WRP	Wireless Routing Protocol
ZRP	Zone Routing Protocol

Chapter 1

Introduction

In this chapter, a brief overview of wireless ad hoc networks is given. Several design challenges in power control, Medium Access Control (MAC), routing, and Quality of Service (QoS) support are listed. The importance of jointly designing several layers is pointed out and a multi-layer optimization approach is proposed to address the design challenges. The cognitive radio techonology is also considered. The overall outline of the dissertation is given at the end of this chapter.

1.1 Background

There has been a great interest in wireless ad hoc networks recently since they have tremendous military and commercial potential. Wireless ad hoc networks are defined as the category of wireless networks that utilize multi-hop radio relaying and are capable of operating without the support of any fixed infrastructure (hence, they are also called infrastructureless networks) [1]. The term "ad hoc" implies that it is a network established for a special, often extemporaneous service customized to applications. Hence, the typical wireless ad hoc network is set up for a limited period of time, and the protocols are tuned to the particular application. The application may be mobile and the environment may change dynamically. If nodes are mobile, the network is termed as Mobile Ad hoc NETworks (MANET). Consequently, the ad hoc protocols must self-configure to adjust to environment, traffic and mission changes. Spurred by the growing interest in ad hoc networking, a number of standard activities and commercial standards evolved in the mid to late 1990s. Within the IETF [2], the MANET working group was born, and sought to standardize routing protocols for wireless ad hoc networks. The 802.11 [3] subcommittee standardized a medium access protocol that was based on collision avoidance. HIPERLAN and Bluetooth are some other standards that addressed and benefited ad hoc networking.

1.1.1 Applications of Wireless Ad Hoc Networks

Wireless ad hoc networks, due to their quick and economically less demanding deployment, possess goals that are very different from mobile telephony and Internet access. They find applications in several areas. The first category of applications is to set up communications for specialized, customized, extemporaneous applications in areas where there is no pre-existing infrastructure, e.g., battlefield, jungle explorations, or the forestry or lumber industry. The second category of possible applications is to set up communications where the infrastructure has failed, e.g., earthquake rescue. The third category of applications is motivated by lack of convenient, low cost infrastructure, e.g., sensors scattered throughout a city for biological detection, an infrastructureless network of notebook computers in a conference or campus setting.

1.1.2 Characteristics of Wireless Ad Hoc Networks

A wireless ad hoc network is a collection of possibly mobile nodes that wish to communicate, but have no fixed infrastructure available, and have no pre-determined organization of available links. Individual nodes are responsible for dynamically discovering which other nodes they can directly communicate with. A key assumption is that not all nodes can directly communicate with each other, so they are required to relay packets on behalf of other nodes in order to deliver data across the network. A significant feature of wireless ad hoc networks is that rapid changes in connectivity and link characteristics are introduced due to node mobility and power control practices. In the following, we review main characteristics of wireless ad hoc networks in detail.

- Mobility

Mobility of nodes is not a mandatory requirement for wireless ad hoc networks. For example, the nodes deployed for periodic monitoring of soil properties are not required to be mobile. However, in many cases, the nodes in wireless ad hoc networks can be rapidly repositioned and/or move. The rapid deployment in areas with no infrastructure often implies that the users must explore an area and perhaps form teams and coordinate among themselves to create a taskforce. We can have individual random mobility, group mobility, motion along preplanned routes, etc. The mobility model can have major impact on the selection of a routing scheme and can thus influence performance.

- Scalability

Scalability in wireless ad hoc networks can be broadly defined as whether the network is able to provide an acceptable level of service to packets even in the presence of a large number of nodes in the network. In some applications, for example, large environmental sensor fabrics, battlefield deployment, the wireless ad hoc network can grow to several thousand nodes. For wireless "infrastructure" networks (e.g. cellular network), scalability is simply handled by a hierarchical construction. The limited mobility of infrastructure networks can also be easily handled using mobile IP or local handoff techniques. In contrast, because of the more extensive mobility and the lack of fixed reference, pure wireless ad hoc networks do not tolerate mobile IP or a fixed hierarchy structure. Thus, mobility, jointly with large scale is one of the most critical challenges in wireless ad hoc networks design.

- Multihopping

A multihop network is a network where the path from source to destination traverses several other nodes. Wireless ad hoc networks consist of autonomous nodes that collaborate in order to transport information. Usually, these nodes act as end systems and routers at the same time. Because of most wireless ad hoc networks' mobility and scalability nature, they often exhibit multiple hops for obstacle negotiation, spectrum reuse, and energy conservation.

- Energy conservation

Most ad hoc nodes (e.g., laptops, sensors, PDAs) have limited power supply and no capability to generate their own power. Energy efficient design is critical for longevity of the mission. Physical- and network-level power conservation to extend battery life is an important design consideration. Recently, many wireless data link layer standards define power-saving sleep mode, to allow a node to temporarily turn off certain components, when it is not actively engaged in communicating.

- Self-organization

Wireless ad hoc networks must autonomously determine their own configuration parameters including: addressing, routing, clustering, position identification, power control, etc. Some of the popular ad hoc network applications require some special nodes, e.g. unmanned robotic components, to coordinate their motion and dynamically distribute in a geographic area to provide coverage of disconnected islands. For example, Unmanned Airborne Vehicles (UAV) can cooperate in maintaining a large ground ad hoc network interconnected in spite of physical obstacles, propagation channel irregularities and enemy jamming.

- Security

The challenges of wireless security are well known. Wireless ad hoc networks, however, are even more vulnerable to attacks than their infrastructure counterparts. The reasons: 1) open wireless medium; 2) capture of unattended roaming nodes and impersonation; 3) decentralized coordination protocols vulnerable to attack (e.g., contention based MAC); 4) lack of centralized certificate authority for key exchange, etc. Both active and passive attacks are possible. An *active* attacker tends to disrupt operations, for example, an imposter posing as a legitimate node intercepts control and data packets; reintroduces bogus control packets; damages the routing tables beyond repair; unleashes denial of service attacks, etc. Due to the complexity of the ad hoc network protocols, active attacks are far more difficult to detect than in infrastructure networks. However, the active attacker could be eventually discovered and physically disabled, the *passive* attacker is never discovered by the network. Like a "bug", it could be placed in a sensor field or at a street corner. It monitors data and control traffic patterns, the information is relayed back to the enemy headquarters via special communications channels with

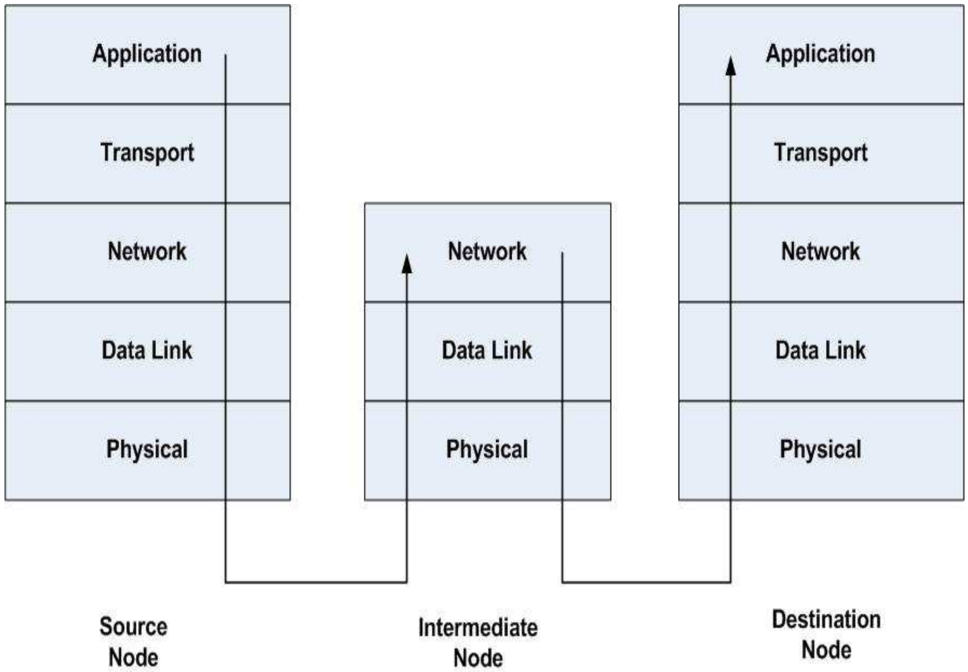

Figure 1.1: A five-layer protocol model.

low energy and low probability of detection. Defense against passive attacks requires powerful novel encryption techniques coupled with careful network protocol design.

1.2 Design Challenges of Wireless Ad Hoc Networks

Based on the characteristics of wireless ad hoc networks summarized in the previous section, the associated design challenges will be discussed in this section. The notion of protocol layering provides a conceptual basis for understanding how a complex set of protocols works together with the hardware to provide a powerful communicaiton system. The five layer reference model as shown in Figure 1.1 is considered in this work.

The purpose of the *physical layer* is to efficiently transmit and receive data bits with as few errors as possible. It handles modulation, error coding, transmisson, and reception. Because energy conservation is critical in MANET, we will focus on power control at the physical layer.

The *data link layer* groups data bits into frames, and handles frame errors and controls the flow of frames. Basically, it is needed to coordinate the transfer of information in any of the several ways, such as one-to-one (unicast), one-to-many (multicast), one-to-all (broadcast), or many-to-one (multi-access) within the radio range as determined by the physical layer. It is the responsibility of the data link layer to perform error correction for anomalies occurring in the physical layer. The data link layer is commonly defined as having two sublayers, the logical link control (LLC) sublayer and the medium access control (MAC) sublayer. The LLC is responsible for realizing a point-to-point link between endpoints and can provide error detection and control functions. The MAC sublayer allows multiple nodes to share wireless media. Since the MAC sublayer has a direct bearing on how reliably and efficiently data can be transmitted between two nodes along the routing path in the network, it affects the Quality of Service (QoS) of the network. Naturally, how to design MAC to improve the network utility is one of the focuses of this dissertation.

When nodes are connected over multiple hops, *network layer* protocol is needed. The primary duty of the network layer is to determine how to route information from the source to the destination. The characterisitcs, such as mobility, lack of infrastructure, and battery-operated, throw lots of challenges in routing protocols design. Hence, routing design will be addressed in this work.

Reliable end-to-end communications is typically provided by a *transport layer* protocol. For instance, the Transmission Control Protocol (TCP) provides a reliable connection-oriented service. The error characteristics of a wireless link can significantly influence the performance of TCP. TCP responds particularly poorly to packet loss because it is designed to treat packet loss as an indication of congestion. This is a reasonable assumption in typical wired networks, but it is often not valid for wireless links. How to adjust TCP protocol to make it performs well in wireless ad hoc networks is a big challenge. Sitting atop these layers is the *application layer*, which runs application processes like electronic mail, web services, or provides the interface presented to the user.

Ultimately, the performance of wireless ad hoc networks depends on the performance of its links, its point-to-point protocols, and its end-to-end protocols and on the interaction among these protocols. Hence, this work adopts a multi-layer optimization approach to improve the performance and efficiency of wireless ad hoc networks. In the following section, several major design challenges in wireless ad hoc networks are summarized, and current approaches in the literature are reviewed.

1.2.1 Power Control

Power control is a fundamental issue in wireless networks because it reduces nodes' power consumption and it increases the number of successful simultaneous transmissions by decreasing multi-user interference. Highly mobile wireless ad hoc networks suffer from limitations of bandwidth and mobility of nodes and users. In addition, the power used by the radio to transmit user information often exceeds the needed power and therefore tends to be wasteful and can cause interference in the coverage areas. The power control schemes proposed in the literature for wireless ad hoc networks can be categorized into two different types, common power control (CPC) where all nodes use the same transmission power; and individual power control (IPC) where nodes decide their transmission power individually. The connectivity of the network is also closely associated with the power control scheme applied.

- Common Power Control

In the case of common power control, the minimum common transmission power is defined as the minimum transmission power that keeps the whole network fully connected, thus the transmission power between one and the other is employed for all nodes. An example of a wireless ad hoc network is shown in Figure 1.2. When CPC is used, the transmission power between node 2 and node 4 (example) will be employed by all other nodes [4]. The algorithm may works well if nodes are distributed homogeneously in space, but even a single outlying node could cause every node to use a high transmission power level. Hence, when the spacial distribution of nodes is not homogeneous, which is mostly the case, it is obviously not optimal to use a common power level throughout the network. In summary, the CPC has the following deficiencies:

1. The nodes use excessive power for transmitting to other nodes which have varying coverage distances.

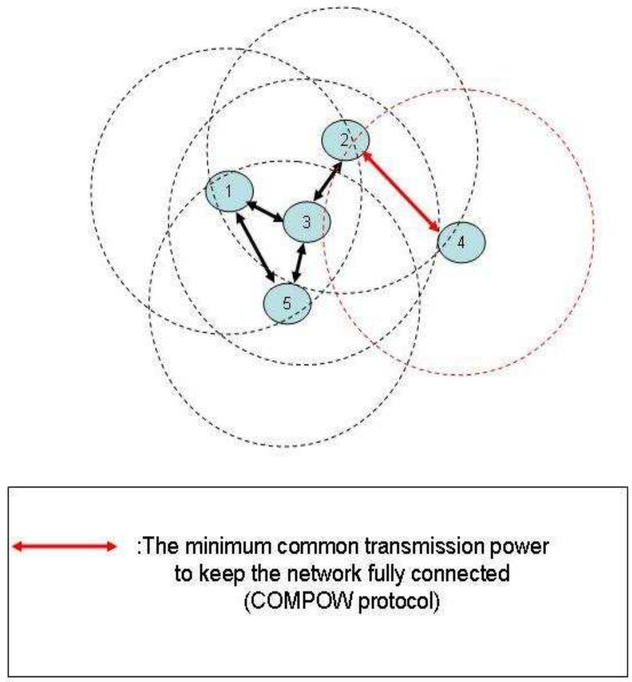

Figure 1.2: Common power control vs. individual power control

2. The use of excessive power increases the interference in communications across the network.

3. The update rate for computation of Common Power increases due to mobility of nodes, as paths get changed.

- Individual Power Control

The individual power control scheme has the advantage of fully exploiting frequency reuse (or in other words, space diversity) to maximize the network throughput or maximize the power efficiency or both [5] [6] [7]. In addition, it will adapt to channel changes and mobility. The study by Gomez and Campbell [8] suggests that using individual variable-range transmission power control improves the traffic carrying capacity by *a factor of 2* than using common transmission power control. Indeed, CPC may be viewed as a special case of IPC. Hence, IPC is chosen as the power control method for wireless ad hoc networks in this research work.

1.2.2 Medium Access Control

The popular Carrier Sense Multiple Access (CSMA) [9] MAC scheme and its variations such as CSMA with Collision Detection (CSMA/CD) developed for wired networks cannot be used directly in the wireless network because the wireless communication channel is inherently prone to errors, the topology is generally unpredictable, and wireless ad hoc networks have their own unique problems such as the hidden-terminal problem and the exposed-terminal problem, which are explained in the following.

- Hidden-terminal problem

The hidden terminal problem is inherent in wireless ad hoc networks. This problem occurs when packets originating from two or more sender nodes, which are not within the direct transmission range of each other, collide at a common receiver node. It necessitates retransmission of packets. Hence, the presence of hidden terminals can significantly reduce the throughput of a MAC protocol used in wireless ad hoc networks.

- Exposed-terminal problem

Exposed terminals, the nodes that are in the transmission range of the sender of an on-going session, are prevented from making a transmission. In order to improve the efficiency of the MAC protocol, the exposed nodes should be allowed to transmit in a controlled fashion without causing collision to the on-going data transfer.

In wireless networks, the MAC protocol must contend for access to the channel while at the same time avoiding possible collisions with neighboring nodes. How to resolve conflicts among different nodes for channel access and to ensure fair and efficient resource sharing at the same time makes the MAC protocol design a challenging task for wireless ad hoc networks.

MAC schemes developed for wireless ad hoc networks can be classified into several categories based on various criteria. For example, they can be categorized as *synchronous* or *asynchronous* in operation. They can be distinguished by who initiates a communication request, and hence categorized as *receiver-initiated* (for example MACA-BI [10]) or *sender-initiated* (for example MACA [11], MACAW [12]). They can

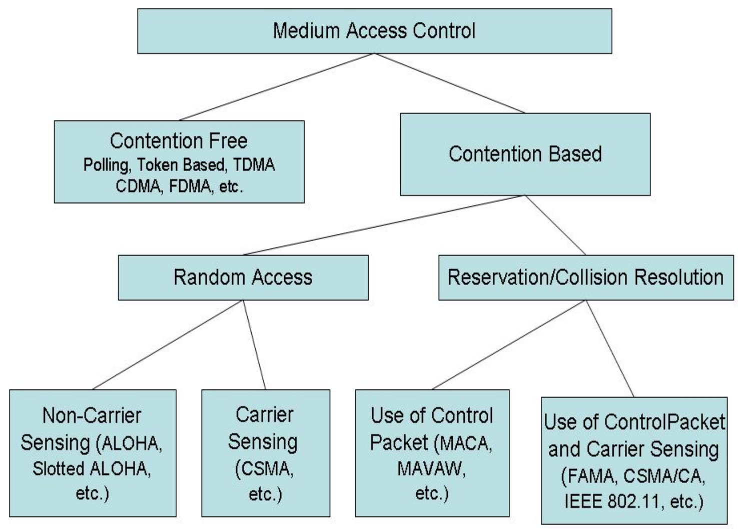

Figure 1.3: Classification of MAC schemes.

also be classified as contention-free schemes and contention-based schemes as shown in Figure 1.3. In *contention-based* MAC schemes, no central control node is needed for allocating channel resources to other nodes in the network. To transmit, each node must contend for radio resources. Collisions occur when more than one node tries to transmit at the same time. How to resolve persistent conflicts in transmissions is a major design task for contention-based schemes. On the contrary, *contention-free* MAC schemes assign dedicated channel resources to each node that wish to send packets. This works well for constant bit rate traffic. However, for bursty data traffic, channel resources will be wasted if there is no packets queued for transmission. Examples of contention-free protocols include Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), and Code Division Multiple Access (CDMA). In the following, we will briefly review contention-based schemes and contention-free schemes for wireless ad hoc networks.

Contention-Based MAC Schemes

For contention-based schemes, we can classify one step further to "*random access*" and "*dynamic reservation/collision resolution*" protocols as shown in Figure 1.3. The simplest random access based scheme is pure *ALOHA*. The basic operation of ALOHA is simple: a node may access the channel whenever it has a packet that needs to be sent. Naturally, more than one node may transfer at the same time, causing collisions. Thus, ALOHA is suitable under low system loads and it offers relatively low throughput. *Slotted ALOHA*, which is a variation of ALOHA, introduces synchronized transmission time slots similar to TDMA. Slotted ALOHA doubles the throughput as compared to the pure ALOHA.

A family of *Carrier Sense Multiple Access (CSMA)* based schemes further reduce the packet collisions and improve the throughput. The basic idea behind a CSMA protocol is: a node first senses the channel to make sure it is idle before starting to transmit. This behavior is sometimes called "listen before talking". If the channel is not busy, the node can transmit. If the channel is busy, the node will defer transmission.

Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) is commonly used in wireless local area networks, including the IEEE 802.11. CSMA/CA leverages the performance benefits of CSMA, but extends CSMA to reduce the likelihood of a collision. CSMA/CA avoids the use of collision detection, as in CSMA/CD, which is very popular in wired networks. Collision detection is not practical in a wireless network because a node's own transmission will typically obscure any transmissions at other nodes that may cause a collision at a receiver. Additionally, it is impossible to ensure that all transmitters detect a collision if one occurs at an intended receiver.

In order to solve the hidden and exposed terminal problems, many contention based (but involve some forms of *dynamic reservation/collision resolution*) schemes are proposed in the literature. Some schemes use the Request-to-Send/Clear-to-Send (RTS/CTS) control packets to prevent collisions, e.g. CSMA/CA, Multiple Access Collision Avoidance (MACA) and MACA for Wireless LANs (MACAW). The 802.11 MAC protocol supports two modes of operation, namely distributed coordination function (DCF) and point coordination function (PCF). The DCF mode provides best effort service, while the PCF mode has been designed to provide real-time traffic support in infrastructure-based wireless network configurations. The DCF mode does not use any kind of centralized control, all stations are allowed to contend for the shared medium simultaneously. CSMA/CA mechanism and random backoff scheme are used to reduce frame collisions.

Contention-Free MAC Schemes

TDMA, FDMA, and CDMA are commonly used and widely investigated collision-free medium access control protocols. They differ in how they partition physical layer resources among nodes. *TDMA* partitions physical layer channels into a set of predetermined time slots and assigns different time slots to different nodes in the network. While data tranmission from different nodes are sent at different times, they share the same frequencies in a TDMA system. *FDMA* partitions the allocated bandwidth into bands and assigns these bands to nodes in the network. In an FDMA system, data transmissions occur at different frequencies, but can occur at the same time. While TDMA and FDMA assign time slots and frequency bands to nodes, respectively, *CDMA* assigns different spreading codes to different nodes. Thus, CDMA allows simultaneous transmissions within the same frequency band, provided that the transmitters use different spreading codes.

A *hybrid TD/CDMA* medium access control scheme is adopted in this work. Each node is assigned a randomly generated spreading code. On top of that, time is split into equal sized slots where only scheduled nodes are allowed to transmit in each slot. The major advantages of a hybrid TD/CDMA schemes are greater flexibility and increased adaptability. A pure CDMA scheme assigns one or more spreading codes to a single node for the duration of its connection, while a pure TDMA scheme only allows one user to transmit during a particular time slot. Pure CDMA or pure TDMA can achieve only one degree of freedom, while a hybrid TD/CDMA scheme is more flexible since it can achieve two degrees of freedom. The flexibility can be used to adapt to different conditions, such as, for multimedia or other applications that have high data rate requirements or that require differentiated QoS. A hybrid TD/CDMA scheme may assign spreading codes to a user only at certain times, such as when the user has queued packets, thus a central controller may assign the same spreading code to different users in different time slots or it may dynamically assign different time slots to users, which assures great adaptability.

In view of low efficiency of contention-based MAC scheme, and better flexibility and increased adaptability of a hybrid TD/CDMA MAC scheme, comparing to a pure TDMA or CDMA contention-free MAC scheme, a hybrid TD/CDMA MAC scheme is adopted in this work. Specifically, a joint power control and scheduling scheme is proposed in Chapter 3 to maximize throughput while maintaining fairness among users.

1.2.3 Routing

The topology of mobile wireless ad hoc networks may change frequently and without prior notice, which makes routing in such networks a challenging task. The challenges are summarized as follows:

1. First, wireless ad hoc networks consist of autonomous nodes that collaborate in order to transport information, and those nodes are allowed to move in an uncontrolled manner. Such node mobility results in highly dynamic network with rapid topological changes that may cause frequent route failures. A good routing protocol has to adapt to the changing network topology dynamically.

2. Second, the underlying wireless channel provides much lower and more variable bandwidth than wired networks. The wireless channel working as a shared medium makes the available bandwidth per node even lower. So routing protocols should be bandwidth efficient by expending a minimal overhead for computing routes so that much of the remaining bandwidth is available for the actual data communication.

3. Third, most nodes run on batteries which have limited energy supply. In order for nodes to stay and communicate for longer periods, it is desirable that routing protocols be energy efficient as well.

Hence, routing protocols for wireless ad hoc networks must meet the conflicting goals of dynamic adaptation and low overhead to deliver good overall performance.

Routing for wireless ad hoc networks can be classified as two different approaches: *topology-based* and *position-based* routing. While in topology-based routing approaches,

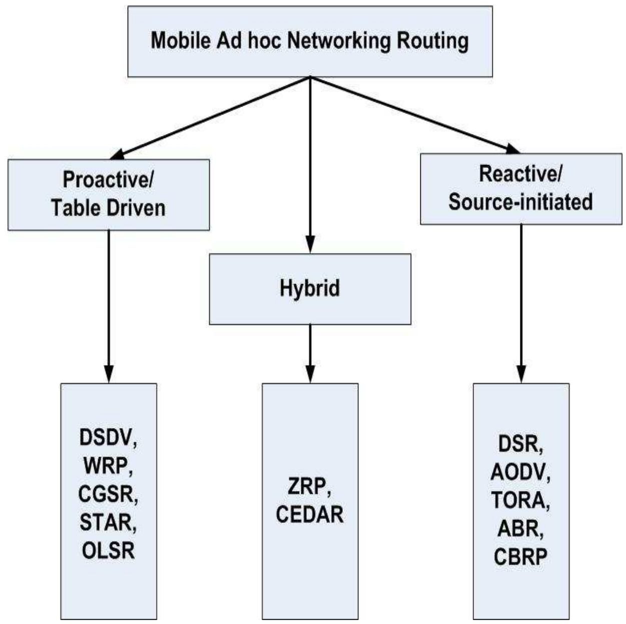

Figure 1.4: Mobile ad hoc networking routing classification.

nodes discover (partial or full) topology information by exchanging routing messages and use this information to guide future routing decisions. Position-based routing assumes that each node knows its own location by using the global positioning system (GPS) or some other indirect, localization technique. Besides, every node learns locations of its immediate neighbors by exchanging hello messages. A location service is used by the sender of a packet to determine the position of destination and to include it in the packet's destination address. The routing decision at each node is based on the destination's position contained in the packet and the position of the forwarding node's neighbors. In sum, position-based routing does not require the establishment or maintenance of routes. The nodes have neither to store routing tables nor to transmit messages to keep routing tables up to date. In this dissertation, only topology-based routing approach is considered, because location information may not be available in many situations. Based on routing information update mechanism, topology-based routing protocols can be categorized as reactive (route on-demand), proactive (routes ready-to-use) or hybrid as depicted in Figure 1.4. These categories are briefly discussed as follows:

Proactive Routing Algorithms

Proactive or table-driven routing algorithms employ classical routing strategies such as distance vector based or link state based routing. They attempt to maintain consistent up-to-date routing information from each node to every other node in the network. These protocols require each node to maintain one or more tables to store routing information, and they respond to changes in network topology by propagating updates throughout the network in order to maintain a consistent network view. The Destination-Sequenced Distance Vector routing (DSDV) [13] is a typical proactive routing protocol. DSDV routing is essentially a modification of the basic Bellman-Ford routing mechanism [14]. The modifications include the guarantee of loop-free routes and a simple route update protocol. Wireless Routing Protocol (WRP) [15], similar to DSDV, inherits the properties of the distributed Bellman-Ford algorithm. However, WRP differs from DSDV in table maintenance and in the update procedures. WRP requires each node to maintain four routing tables and the use of hello packets whenever there are no recent packet transmissions from a given node. It will lead to substantial memory requirements, and the hello packets consume bandwidth and disallow a node to enter sleep mode. WRP has an advantage over other path-finding algorithms because it avoids the problem of creating temporary routing loops that other path-finding algorithms have (through the verification of predecessor information). Optimized Link State Routing (OLSR) [16], Topology Broadcast based on Reverse-Path Forwarding (TBRPF) [17] are examples of proactive protocol using link state based routing strategy. Where as all the above routing protocols assume flat network topologies, Clusterhead Gateway Switch Routing (CGSR) assumes a hierarchical network topology. It uses DSDV as the underlying routing scheme. However, it modifies DSDV by using a hierarchical cluster-head-to-gateway routing approach to route traffic from source to destination.

In a nutshell, proactive routing protocols maintain unicast routes between all pairs of nodes even if these paths are not currently used. Therefore, when the need arises, the traffic source has a route readily available and does not have to incur any delay for route discovery. The main drawback of these approaches is that the maintenance of unused paths may occupy a significant part of the available bandwidth if the topology of the network changes frequently.

Reactive Routing Algorithms

A different approach from the proactive/table-driven routing is reactive/source-initiated on-demand routing. This type of routing creates routes only when desired by the source node. When a node requires a route to a destination, it initiates a route discovery process within the network. This process is completed once a route is found or all possible route permutations have been examined. Once a route has been established, it is maintained by a route maintenance procedure until either the destination becomes inaccessible along every path from the source or until the route is no longer desired. Ad hoc On-demand Distance Vector (AODV) [18], Dynamic Source Routing (DSR) [19], Temporally Ordered Routing Algorithm (TORA) [20], and Associativity-Based Routing (ABR) [21] are examples of existing reactive/source-initiated on-demand ad hoc routing protocols. Since reactive routing is adopted in this dissertation, DSR and AODV routing schemes are explained in detail.

- Dynamic Source Routing

As the name suggests, DSR is based on the source node that determines and specifies a route to the destination. It consists of two major phases: *route discovery* and *route maintenance*. The basic approach of DSR (and all other on-demand routing protocols) during the route discovery phase is to establish a route by flooding *Route Request* packets in the network. The destination node, on receiving a *Route Request* packet, responds by sending a *Route Reply* packet back to the source, which carries the route traversed by the *Route Request* packet received. Mobile nodes are required to maintain route caches that contain source routes of which the mobile is aware of. Entries in the route cache are continually updated as new routes are learned and obsolete routes are deleted.

When a node has a packet to send to a destination, it first consults its route cache to determine whether it already has a route to the destination. If it has an unexpired route to the destination, it will use this route to send the packet. On the other hand, if the node does not have such a route, it initiates route discovery by broadcasting a *Route Request* packet. Each intermediate node will check whether it knows a route to the destination or not. If it does not and the *RouteRequest* packet is not redundant, it will add its own address to the *route record* of the packet and forwards the packet along its outgoing links.

A *Route Reply* is generated when the route request reaches either the destination or an intermediate node which contains an unexpired route to the destination in its route cache. By the time the packet reaches either the destination or the intermediate node, it contains a route record yielding the sequence of hops taken. And the node may reverse the route in the route record to return the *Route Reply* if symmetric links are supported. If symmetric links cannot be assumed, a node generating a *Route Reply* message may need to initiate its own route discovery procedure.

Route maintenance is accomplished through the use of route error packets and acknowledgements. Acknowledgements are used to verify the correct operation of the links in a path from source to destination. *Route error* packets are generated at a node adjacent to a broken link to inform the source node. When a route error packet is received, the hop in error is removed from the node's route cache and all routes containing the hops are truncated at that point. The source node reinitiates the route discovery procedure.

The disadvantage of DSR is that the route maintenance mechanism does not locally repair a broken link. Stale route cache information could also result in inconsistencies during the route discovery phase. The route setup delay is higher than proactive routing schemes. As an on-demand routing protocol, DSR is very flexible. However, DSR introduces large routing overhead and does not scale well for use in large networks, because all the routing information has to be carried in packet headers.

- Ad hoc On-demand Distance Vector

AODV is, in essence, a combination of DSR and DSDV. AODV is an improvement on DSDV because it typically minimizes the number of required broadcasts by creating routes on a demand basis, as opposed to maintaining a complete list of routes as in the DSDV algorithm. Unlike DSR, AODV does not include route information in every data or control packet, instead, the source node and the intermediate node store the next-hop information corresponding to each flow for data packet transmission. AODV employs destination sequence numbers to identify the most recent path.

When a source node desires to send a message to a certain destination node and does not already have a valid route to that destination, it initiates a *route discovery* process by broadcasting a *Route Request (RREQ)* message to its neighbors, which then forward the request to their neighbors, and so on, until either the destinaiton or an intermediate node with a “fresh enough” route to the destination is located. Any node that receives the RREQ message updates its next hop table entry with respect to the preceding node in the path back to the source, thereby establishing a reversing path back to the initiator of the RREQ message. If a node knows an unexpired route to the destination or the node is the destination node, a *Route Reply (RREP)* message is generated and sent by unicast back to the source. Because the RREP message is forwarded along the reverse path established by the RREQ message, AODV requires symmetric links.

One unique feature in AODV is that nodes use “Hello” messages to probe their neighbors in order to validate routes. Nodes broadcast “Hello” messages in a reasonble interval. If a node does not receive a “Hello” message from a paticular neighbor for a certain period, it will delete this neighbor from its neighbor cache and mark the corresponding routes as invalid. That feature is very effective, though periodic beaconing leads to unnecessary bandwidth consumption. Another disadvantage of AODV is multiple RREP messages in reponse to a single RREQ message can lead to heavy control overhead. Unlike DSR, AODV does not include route information in every data or control packet header, which reduces overhead.

Hybrid Routing Algorithms

Proactive schemes pose a negligible delay, but use the whole network capacity to update routes and in some cases the routes determined may not be used at all. On the contrary, reactive routing protocols find and maintain only needed routes, thereby reducing the burden on the network when only a small subset of all available routes is in use at any time. This approach is attractive when the network traffic is sporadic, bursty and directed mostly toward a small subset of nodes. However, since routes are created when the need arises, data packets experience queuing delays at the source while the route is being found at session initiation and when the route is being repaired later on after a failure. Another consequence of reactive routing is that routes may become suboptimal as time progresses since a route is used until it fails with a pure reactive protocol.

In response to the above observation, it is not hard to hypothesize that a combination of proactive and reactive approaches is perhaps better than either approach. As an example, Zone Routing Protocol (ZRP) [22] combines local proactive routing and global reactive routing in order to achieve a higher level of efficiency and scalability. However, one important shortcoming of ZRP is that its design assumes a uniform traffic distribution and then optimizes the overall overhead. When the traffic is non-uniform, it may not actually be efficient.

In this dissertation, there will be more of the ad hoc routing coverage in Chapter 2, where a novel power-aware maximally disjoint routing scheme is proposed to provide QoS with high energy efficiency.

1.2.4 Quality of Service

Quality of Service is the performance level of a service offered by the network to the user. The goal of QoS provisioning is to achieve a more deterministic network behavior, so that information carried out by the network can be better delivered and network resources can be better utilized. Conventional QoS metrics include throughput, packet loss rate, end-to-end delay, and delay jitter, etc. QoS metrics for wireless ad hoc networks may include more parameters, such as power consumption and network coverage.

Power saving is important because a network of battery powered devices will not be able to provide any service if the batteries are exhausted.

Because of the challenges posed by the wireless environment, such as, 1) error prone shared radio channel, 2) lack of central coordination, 3) dynamically varying network topology, 4) limited resource availability, all protocol layers must cooperate in order to meet QoS requirements in wireless ad hoc networks. It is almost impossible to guarantee the fulfillment of QoS requirements at the physical layer at all times. Hence, adaptation mechanisms need to be implemented in higher layers to reduce the impact of unreliable physical layer on QoS as much as possible. In practice, those place most demands on the data link and network layers. To meet a given QoS requirement, the MAC sublayer needs to solve the problem of medium contention and provide adaptive scheduling and resource allocation, while the LLC sublayer needs to provide reliable communication over the link that can compensate for impairments at the physical layer. The network layer should be adaptive enough to accomodate different data traffic characteristics and QoS requirements. Much research has focused on QoS routing, which refers to the discovery and maintenance of routes that can satisfy QoS requirements under given resource constraints.

In spite of these mechanisms, QoS requirements still may not be guaranteed deterministically in a wireless ad hoc network. In other words, it is very difficult to provide *hard QoS* guarantees to user applications. Hence, almost all the approaches available in the literature provide only *soft QoS* guarantees. *Hard QoS* means QoS requirements of a connection are guaranteed to be met for the whole duration of the session. If the QoS requirements are not guaranteed for the entire session, it is termed as *soft QoS*. Besides, there are different interpretations of QoS at different communication layers. At the physical layer, QoS is synonymous to *signal to interference and noise ratio (SIR)* or to an acceptable *bit error rate (BER)*. At the MAC layer, QoS can be expressed in terms of minimum rate and maximum delay guarantees. For the multihop wireless networks, network layer QoS pertains to end-to-end provisioning of the guaranteed QoS for each session. In this research work, only soft QoS is considered, and a multi-layer optimization approach is proposed to meet the QoS requirements of users.

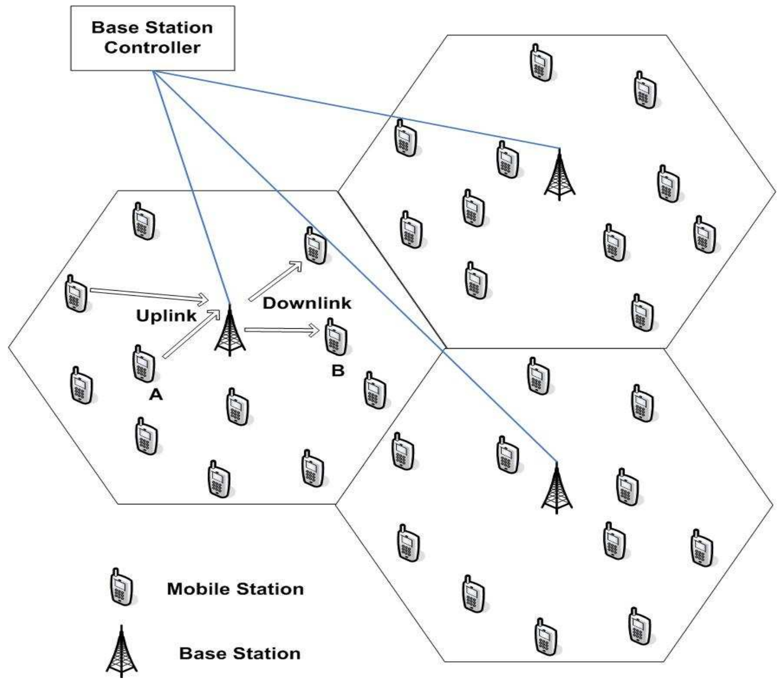

Figure 1.5: A cellular network.

1.3 Multi-Layer Optimization in Wireless Ad Hoc Networks

1.3.1 Cellular vs. Wireless Ad Hoc Networks

As mentioned earlier, wireless ad hoc networks pose many new design challenges with respect to conventional *wireless infrastructure* networks. The wireless infrastructure networks often extend, rather than replace, wired network. We use wireless cellular network as an example of wireless infrastructure networks. The topologies of cellular networks and ad hoc networks are illustrated in Figure 1.5 and Figure 1.6, respectively. The differences between them are briefly summarized in this section. Motivation for the research in this dissertation follows.

- Multi-hop vs. Single-hop

Within wireless cellular networks, wireless access to and from the wired host occurs in the last hop between base stations and mobile units that share the bandwidth of the wireless channel. Since the wireless part of cellular networks is actually one-hop network, routing is not needed. A mobile unit communicates with its nearest base station that is within its communication radius (see Figure 1.5). As it travels out of the

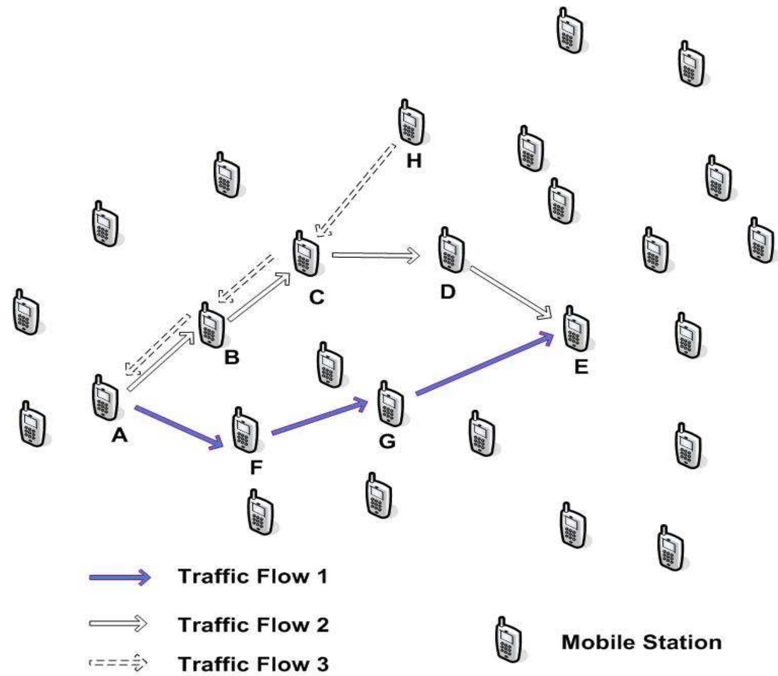

Figure 1.6: A wireless ad hoc network.

range of one base station and moves into the range of another base station, a "handoff" occurs to transit communication seamlessly. On the contrary, wireless ad hoc networks consist of autonomous nodes that collaborate in order to transport information. Usually, these nodes act as end systems and routers at the same time. The path from the source to the destination is often multihop. Hence, routing is necessary for wireless ad hoc networks. For example, two possible multi-hop traffic flows are given in Figure 1.6, if we want to communicate from mobile station A to mobile station E.

- Distributed Nature vs. Central Coordination

In cellular networks, for example, base stations act as central coordinators and allocate bandwidth to mobile terminals. However, wireless ad hoc networks usually do not have centralized coordinators. It is infeasible to assign network resources in a centralized manner where nodes keep moving continuously and no infrastructure is available. Therefore, nodes must be scheduled in a distributed fashion for gaining access to the channel. This may require exchange of control information. The MAC protocol must make sure that the additional overhead (in terms of bandwidth consumption) incurred due to this control information exchange is not very high.

- Channelization

Cellular networks use different frequency for uplink and downlink. Channelization using techniques such as TDMA, FDMA, CDMA is natural and mature. It is easy to get synchronized because it is single-hopped and the base station is the central controller.

Unlike cellular networks, there is lack of centralized control and global synchronization in wireless ad hoc networks. Channelization is not trivial any more. Since the same media are shared by multiple mobile ad hoc nodes, access to the common channel must be made in a distributed fashion, through the presence of MAC protocol. Furthermore, channelization is beyond MAC protocol, it is also related to power control and routing, which will be discussed in Chapter 2.

- Cross-Layer Issue

In cellular networks, the five-layer shown in Figure 1.1 functions are relatively clean cut and cross layer coupling is not essential. On the contrary, the five layer functions are highly coupled in wireless ad hoc networks. For example, power control (a physical layer function) will affect the network topology, hence routing (a network layer function). Thus, the optimized multi-layer design is indispensable in wireless ad hoc networks.

1.3.2 Proposed Multi-Layer Optimization Approach

The interactions among different layers have to be taken into account to improve the performance and efficiency in wireless ad hoc networks. The *physical layer* must adapt to rapid changes in link characteristics. The *data link layer* needs to minimize collisions, allow fair access, and reliably transport data over the shared wireless links in the presence of rapid changes and hidden or exposed terminals. The *network layer* needs to determine and distribute information used to calculate paths in a way that maintains efficiency when links change often and the bandwidth is at a premium. It also needs to integrate smoothly with traditional, non ad hoc aware networks and perform functions such as auto-configuration in a changing environment. The *transport layer* must be able to handle delay and packet loss statistics that are very different from wired networks. Finally, *applications* need to be designed to handle frequent disconnections and reconnections with peer applications as well as widely varying delay and packet loss characteristics.

In view of the benefits and necessities of sharing information among different layers, a multi-layer optimization approach is proposed for wireless ad hoc networks in this dissertation. Specifically, given the sessions with their source-destination pairs and QoS requirements, how to jointly design physical, MAC and network layers to optimize the network performance is considered. The *physical layer* has its key parameters, such as transmit power, modulation, coding rate, that will have a direct impact on multiple access of nodes. In this work, transmit power control will be the focus in this layer. Specifically, how to perform smart power control to extend network lifetime, optimize certain network utility and guarantee the basic QoS requirement for users is the focus of this research work. The power control problem is complex in wireless ad hoc networks [23], because the choice of power level affects many layers, for example:

1. The transmit power level determines the quality of the signal received at the receiver.

2. Power control determines the magnitude of the interference it creates for other receivers. Hence, power levels determine the performance of medium access control since the contention for the medium depends on the number of other nodes within range.

3. The transmit power level determines the range of tranmissions. Hence, power control affects network topology and thus routing.

The *MAC layer* is responsible for scheduling the transmissions and allocating the wireless channels. While concurrent transmissions lead to mutual interference, and the transmission schedule naturally affects the performance of the physical layer and may lead to transmit power adaptation in the physical layer. At the same time, the results of different schedules will eventually affect the latency and bandwidth of the routes and may change routing decisions. The *network layer* will decide the routes that carry the data packets. Different routing decisions alter the set of links to be scheduled, and thereby affect the MAC layer. In this dissertation, joint design of physical, medium access control and network layer is proposed for wireless ad hoc networks to allow interactions among these three layers. The simulation results show a significant performance gain due to the proposed multi-layer optimization design.

1.4 Cognitive Radio Technology

In the previous sections, the joint design of multiple layers to improve bandwidth efficiency, energy efficiency, and QoS assurance is reviewed. However, the most advanced systems are approaching the Shannon capacity limit, so a further increase in capacity would require the additional system bandwidth. The Federal Communication Commission (FCC) frequency chart [24] indicates multiple allocations over all frequency bands, and it seems that there is a crisis of spectrum availability at frequencies that can be economically used for wireless communications. However, actual measurements show a very low bandwidth utilization. For example, measurements taken in an urban setting reveal a typical utilization of 0.5% in the 3-4 GHz band, and the utilization drops to 0.3% in the 4-5 GHz band [25]. The discrepancy between the spectrum allocation and spectrum usage suggests that "spectrum access is a more significant problem than physical scarcity of spectrum, in large part due to legacy command-and-control regulation that limits the ability of potential spectrum users to obtain such access" [26]. In order to achieve a much better spectrum utilization and viable frequency planning, a new class of radios, termed *cognitive radio* [27], are under development to dynamically capture the unoccupied spectrum [28, 29]. Cognitive radio is able to reliably sense the spectral environment over a wide bandwidth, detect the presense/absense of legacy users (primary users) and use the spectrum only if the communication does not interfere with primary users. The FCC has recognized the promising technique and is pushing to enable it to a full realization. In this dissertation, the benefit of applying cognitive radio in wireless ad hoc networks is explored and the associated power control problem is formulated and solved.

1.5 Outline of the Dissertation

In this dissertation, a multi-layer optimization approach is proposed for wireless ad hoc networks to improve network performance. Chapter 2 is devoted to joint power control and routing in a multihop CDMA wireless ad hoc networks. Specifically, a joint power control and maximally disjoint routing scheme is proposed for routing data traffic with a minimum rate constraint while maintaining high energy efficiency and prolonged network lifetime.

Joint power control and scheduling with minimum and maximum rate constraints is discussed in Chapter 3 for multihop TD/CDMA wireless ad hoc networks, with the objective of minimizing the total transmit power subject to certain QoS guarantees. Specifically, in order to achieve a balance between throughput and fairness, two popular scheduling algorithms are discussed and extended to multi-hop wireless ad hoc networks.

Chapter 4 considers the radio resource management problem when applying a new emerging technology, cognitive radio, to wireless ad hoc networks. One of the major concerns is that the interference from the cognitive radio network should not violate QoS requirements of the legacy users. Because the co-channel interference comes from heterogeneous systems, a joint power control and admission control procedure is proposed such that the priority of the legacy users is always ensured.

Chapter 5 presents the concluding remarks and a discussion of future work.

Chapter 2

Joint Power Control and Routing in CDMA Wireless Ad Hoc Networks

In this chapter, joint power control and maximally disjoint routing is proposed for multihop CDMA wireless ad hoc networks. A framework of power control with QoS constraints in CDMA wireless ad hoc networks is introduced and the feasibility condition of the power control problem is identified. Both the centralized solution and distributed implementations are derived to calculate the transmission power given the required throughput and the set of transmitting nodes. Then a joint power control and maximally disjoint routing scheme is proposed for routing data traffic with a minimum rate constraint while maintaining high energy efficiency and the prolonged network lifetime. Furthermore, in order to provide reliable end-to-end data delivery, the proposed joint power control and maximally disjoint routing scheme is augmented by a traffic monitoring and switching mechanism to mitigate the effect of node mobility or node failure. Simulation results demonstrate the effectiveness of the proposed scheme.

2.1 Motivation

In a wireless ad hoc network architecture, the MAC protocol plays a critical role in optimizing bandwidth efficiency and resolving collisions due to the broadcast nature of wireless channels. In most standardized wireless ad hoc networks, such as in the widely deployed IEEE 802.11x networks, only one user is allowed to transmit at an instance of time. It is demonstrated in [30] that compared to the DCF (Distributed Coordination Function) mode of the IEEE 802.11x networks, CDMA-based MAC protocols achieve a significant increase in the network throughput at no additional cost in energy consumption. Hence, CDMA is employed as the MAC scheme for multihop wireless ad hoc networks (considered in this chapter), where multiple concurrent transmissions are allowed.

Power control is applied in a wireless ad hoc network to control transmission range and to keep the network fully connected [31]. Because CDMA systems are interference limited, power control also serves as a tool for interference management in CDMA wireless networks to guarantee the success of multiple concurrent transmissions. Because the transmission power of each node will decide the number of nodes in its transmission range, power control will affect the topology of a wireless ad hoc network. Thus, routing needs to be considered jointly with power control. Furthermore, using the minimum required transmission power related routing metric, energy-efficient paths can be calculated.

The instability of topology in the wireless ad hoc network, due to node mobility and changes in wireless propagation conditions, makes Quality-of-Service (QoS) support a challenging problem. Because most nodes are battery operated, energy efficiency is another important issue. In this chapter, we study joint power control and routing to address energy efficiency and QoS support in CDMA wireless ad hoc networks.

2.1.1 Design Goals

In a wireless ad hoc network, QoS support is desirable by many applications. However, as pointed out in previous research [32] [33], “hard QoS” is very difficult to support in wireless ad hoc networks because of node mobility, lack of central control and the constantly changing wireless channels. However, many applications do not require “hard QoS” and accept “soft QoS”. For example, many multimedia applications accept “soft QoS” and use rate adaptive schemes to mitigate disruptions [34]. Hence, only “soft QoS” is supported in this work.

QoS is a measure of the performance level of a service offered by the network to the user. QoS requirements include minimum data rate, maximum delay, maximum delay jitter, and maximum packet loss rate. A guarantee on the minimum data rate is arguably the simplest possible QoS guarantee. Therefore, we believe it is natural that mobile users would expect such an assurance. For example, video can become unusable if the data rate is too low. Even for static TCP-based applications such as web browsing if the data rate is too low then we typically get a large queue buildup which can lead to TCP timeouts and a poor performance. Such effects were discussed by Chakravorty et al. in [35]. Providing a minimum rate guarantee can also help to smooth out the effects of a variable wireless channel. Furthermore, by setting the minimum data rate differently for different users we can ensure service differentiation.

In order to provide QoS assurance, we propose multiple disjoint paths (minimum two) routing as opposed to single path routing. Single path routing is not reliable. The path may be broken during data transmission because of node mobility or node failure. Re-routing after detection of a broken path may incur too much extra delay in data delivery and cause loss of information. The proposed power aware maximally disjoint routing scheme to calculate two "energy-efficient maximally disjoint paths" for each data flow provides the QoS assurance for end user applications. One path acts as the primary path for sending data traffic and the other acts as a backup path and it is stored in the routing table of the sender. In case the primary path fails, the traffic will be switched to the designated backup path. The sender will monitor both paths and follow the route maintenance in standard MANET routing protocols, such as that in DSR [19].

One of the fundamental challenges in wireless ad hoc network routing is how to provide end-to-end QoS support while maintaining low energy consumption and a long network lifetime. In addition, node mobility and node failures introduce challenges for reliable data delivery. In this study, we propose a design that addresses *all* the above requirements. Both energy efficiency and QoS support (minimum rate) are considered when we jointly design power control and routing.

2.1.2 Related Works

Multipath routing techniques have been proposed for wireless ad hoc networks in many previous studies. Lee and Gerla [36] proposed AODV-BR, where alternate routes are maintained locally along the "backbone" of the primary path, and utilized when the primary path fails. Other proposals include TORA [20] and AOMDV [37]. However, disjoint paths and QoS support are not considered in the above works.

Power aware maximally disjoint routing has been considered by Srinivas and Modiano [38] [39]. It allows the data to be sent to multiple disjoint paths *simultaneously* to achieve diversity. This was not intended to handle route disruptions. It used the simplified interference model where the transmission power is proportional to the link distance only ($p_{ij} = d_{ij}^{\alpha}$ and $2 \leq \alpha \leq 4$). No required throughput is considered. Since there are major differences on how to use the obtained disjoint paths to send data between our approach and that in [38], the routing designs are completely different. Moreover, our scheme is augmented by a dynamic traffic switching mechanism to deal with node mobility or node failure.

Another related work has been done in terms of QoS provisioning [40] [41]. Iterations of power control and routing have been proposed to perform QoS provisioning for CDMA wireless ad hoc networks. The results are routes for every node pairs in the network. However, finding disjoint paths between every node pairs while achieving minimum energy may needlessly minimize energy usage over nodes that may not even be transmitting, and yields sub-optimal solutions for nodes that are transmitting. Disjoint paths are not considered in that work.

2.1.3 Outline of the Proposed Scheme

Given the current existing end-to-end traffic sessions and channel allocations across the network, the procedures of the proposed scheme are as follows.

1. Determine the power controlled connectivity graph by performing per-channel based power control. This step will find all the feasible links that are able to accommodate the coming traffic with specified QoS in terms of minimum data rate. A detailed explanation and an example are given in Section 2.2.

2. Perform Minimum-Power-Split-Multipath-Routing (MPSMR) or Balanced-Energy-Split-Multipath-Routing (BESMR) (proposed in Section 2.3) iteratively to find a primary path and the associated maximally disjoint backup path.

3. Send traffic only along the primary path and monitor both the primary path and

the backup path for available bandwidth. If the primary path is broken, switch the traffic to the backup path.

The rest of the chapter is organized as follows: Section 2.2 introduces the power control framework and the power controlled connectivity model with the minimum rate guarantee for CDMA wireless ad hoc networks. An iterative joint power control and maximally disjoint routing algorithm that may employ different energy related routing metrics is proposed in Section 2.3. The dynamic path restoration for guaranteed data delivery is proposed in Section 2.4. Performance evaluations are performed through extensive discrete-event simulations and the simulation results are given in Section 2.5. Section 2.6 contains the concluding remarks.

2.2 Power Control Framework and Power Controlled Connectivity

The topology and connectivity of a wired network are easy to determine because there exists a communication link between two nodes whenever there is a physical link between them. However, this is not the case in CDMA wireless ad hoc networks. Whether there is a communication link between two nodes or not depends on many physical layer parameters, such as transmission power, spreading gain, modulation and coding scheme, etc. As a result, we define *power controlled connectivity* in CDMA wireless ad hoc networks as follows:

Definition 1 *Given the spreading gain, modulation and coding scheme, and the desired throughput, a link between two nodes exists when the corresponding target Signal-to-Interference-Ratio (SIR) is achievable. In other words, the transmission power to achieve the target SIR is below the maximum allowable transmission power.*

We also define *power controlled connectivity graph* as

Definition 2 *The power controlled connectivity graph includes the feasible set of links (and the associated nodes) that may accommodate the traffic flow with the desired data rate* R^{tar}.

In order to obtain the power controlled connectivity graph given R^{tar}, a power control framework for CDMA wireless ad hoc networks is introduced.

2.2.1 Power Control Framework

The objective of power control is to minimize the total energy consumption, or equivalently, manage interferences intelligently to maximize the energy efficiency, and at the same time, guarantee a certain level of QoS if feasible. In this work, it is assumed that distinct channels are pre-assigned to avoid the primary conflict (a node cannot transmit and receive simultaneously [42]). It should be noted that the power control problem is formulated on a per-channel basis. In other words, only co-channel interference (the interference caused by transmitter-receiver pairs that use the same channel) need to be addressed in a multihop network. An example of channel allocation is shown for the end-to-end paths in Figure 2.1. Because a node cannot transmit and receive at the same time, transmissions of consecutive links along a path have to use different channels. For instance, in Figure 2.1 (a), two channels are allocated. Active links A to B and C to D share channel 1, active links B to C and D to E share channel 2. Moreover, at the node where multiple paths cross such as node E in Figure 2.1 (b), more channels may be necessary.

Assume that there are N_c transmitter-receiver pairs (active links) in the network using the same channel c, the power control problem can be formulated as follows
(P.1)

$$\min_{p_i} \sum_i p_i \ , \qquad i = 1, 2, \cdots, N_c. \tag{2.1}$$

subject to the constraints

$$\gamma_i \geq \gamma_i^{tar} \ , \qquad i = 1, 2, \cdots, N_c. \tag{2.2}$$

$$0 \leq p_i \leq p_i^{max} \ , \qquad i = 1, 2, \cdots, N_c. \tag{2.3}$$

where γ_i is the actual received SIR at receiver i, γ_i^{tar} is the target SIR of the i^{th} active link, p_i is the transmission power of transmitter i, p_i^{max} is the maximum power allowed for transmitter i.

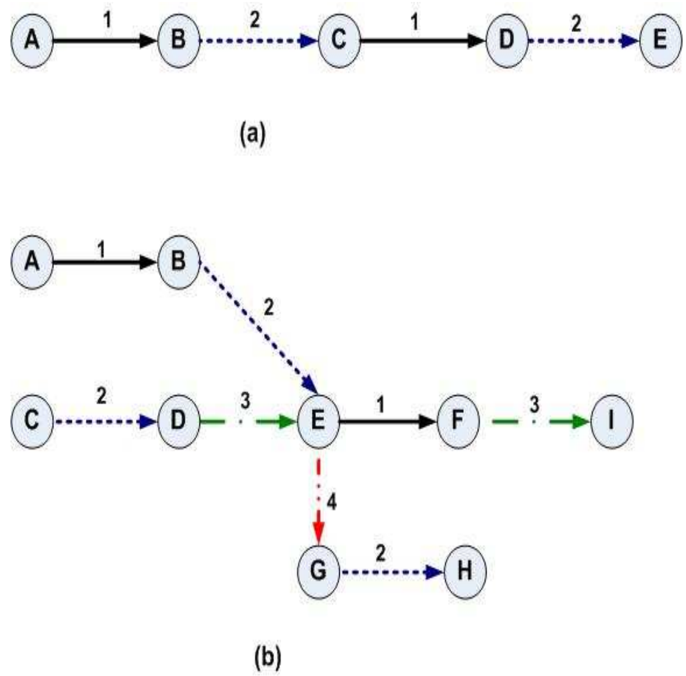

Figure 2.1: Channel allocation (indicated by numerical numbers) in multihop networks: an example.

The received SIR at receiver i is given by

$$\gamma_i = \frac{h_{ii}p_i}{\frac{1}{L}\sum_{j\neq i} h_{ij}p_j + \sigma^2} \tag{2.4}$$

where h_{ii} is the link gain from transmitter i to its designated receiver; h_{ij} is the link gain from transmitter j to receiver i; p_i and p_j are powers of transmitters i and j, respectively; and σ^2 is the background noise. The quantity L is the spreading gain for spread spectrum systems, for example, a typical value of the spreading gain $L = 64$ or 128 is used in CDMA systems. The general interference model adopted here assumes that each transmitting node in the network causes interference at any receiving node using the same channel, even if they are far apart [43]. This model is considered more realistic than the one which assumes that transmitting nodes only cause interference to their neighbors. This is because the aggregate interference from a large number of nodes may not be negligible even if the interference from each of them is small.

Given the traffic flow with the desired data rate R^{tar}, the corresponding target SIR can be expressed as

$$\gamma_i^{tar} = 2^{\frac{R_i^{tar}}{W_i}} - 1\,, \qquad i = 1, 2, \cdots, N_c \tag{2.5}$$

where W_i is the bandwidth occupied by transmission from the i^{th} transmitter to its designated receiver. $R_i^{tar} = n_i R^{tar}$, where n_i is the number of incoming and outgoing active links at the ith transmitter. Note that this formula (derived from the Shannon capacity formula) uses the achievable rate (upper bound) of the AWGN channel. However, it is justified by the fact that with the current modulation and coding technology it can be closely approximated in most practical scenarios [44].

2.2.2 Centralized Solution

The following theorem gives the *feasibility condition* of the formulated power control problem **(P.1)**

Theorem 1 *A target SIR vector* $\gamma^{\mathbf{tar}}$ *is achievable for all simultaneous transmitting-receiving pairs within the same channel as long as the feasibility condition is met, i.e., the matrix* $[I - \Gamma^{tar} Z]$ *is non-singular (thus invertible) and the inverse is element-wise positive, where matrix* Γ^{tar} *is a diagonal matrix*

$$\Gamma_{ij}^{tar} = \begin{cases} \gamma_i^{tar} & i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

and matrix Z *is the following nonnegative matrix*

$$Z_{ij} = \begin{cases} \frac{h_{ij}}{L h_{ii}} & i \neq j \\ 0 & i = j \end{cases} \tag{2.7}$$

and $\gamma_i \geq \gamma_i^{tar}$ *and* $p^{max} \geq p \geq 0, \quad i = 1, 2, \cdots, N_c.$

Proof: A target SIR vector $\gamma^{\mathbf{tar}}$ is achievable for all simultaneous transmitting-receiving pairs within the same channel if the following conditions are met [45, 46]

$$\gamma_i \geq \gamma_i^{tar} \tag{2.8}$$

$$p^{max} \geq p \geq 0 \tag{2.9}$$

where p is the vector of transmitting powers. Replacing γ_i with equation (2.4) and rewriting the above conditions in matrix form gives

$$[I - \Gamma^{tar} Z] p \geq u \tag{2.10}$$

$$p \geq 0 \tag{2.11}$$

where matrix Γ^{tar} is a diagonal matrix

$$\Gamma_{ij}^{tar} = \begin{cases} \gamma_i^{tar} & i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

and matrix Z is the following nonnegative matrix

$$Z_{ij} = \begin{cases} \frac{h_{ij}}{Lh_{ii}} & i \neq j \\ 0 & i = j \end{cases} \tag{2.13}$$

u is the vector with elements

$$u_i = \gamma_i^{tar} \sigma^2 / Lh_{ii}, \qquad i = 1, 2, ..., N \tag{2.14}$$

It is shown in [46] that if the system is feasible, the matrix $[I - \Gamma^{tar} Z]$ must be invertible and the inverse should be element-wise positive, thus prove the theorem. ■

It is also shown in [46] (Proposition 2.1) that if the system is feasible, there exists a unique (Pareto optimal) solution which minimizes the total transmitted power. A power assignment p^* is said to be *Pareto optimal* if it is feasible and any other feasible power assignment p satisfies $p \geq p^*$ componentwise [47, 48]. This solution is obtained by solving a system of linear algebraic equations

$$[I - \Gamma^{tar} Z] p^* = u \tag{2.15}$$

In the case of a CDMA network as considered in this work, since the processing gain L is a large positive number, the power control problem is usually feasible because the matrix $[I - \Gamma^{tar} Z]$ is a diagonally dominant matrix (see p. 151 Definition 6.2 in [49]). The spectral radius of $\Gamma^{tar} Z$ is less than unity (see p. 151 of [49]) in this case. And this is equivalent to the feasibility condition given in Theorem 1 [46].

Equation (2.15) provides a centralized solution to the power control problem **(P.1)**. Given the desired throughput, maximum allowable power and bandwidth of each active link i (R^{tar}, p_i^{max} and W_i), it is straightforward to calculate the optimal power vector using equation (2.16) provided that the link gain matrix is available

$$p^* = [I - \Gamma^{tar} Z]^{-1} u. \tag{2.16}$$

An $N \times N$ link gain matrix H may be formed where h_{ij} is the link gain from the j^{th} transmitter to the i^{th} receiver. Note that H is always a square matrix where the column is indexed by the transmitter and the row is indexed by the corresponding receiver.

2.2.3 Distributed Schemes

The centralized solution (equation (2.16)) needs a central controller and *global* information of all link gains. However, it is very difficult to obtain the knowledge of all link gains in an infrastructure-less wireless ad hoc network and it is usually impractical to implement a centralized solution. Also, even if the centralized scheme were to be implemented, the amount of signaling overhead would increase significantly. Therefore, a distributed implementation is suggested for realistic scenarios.

Distributed power control schemes may be derived by applying iterative algorithms to solve equation (2.16). For example, using the first-order Jacobian iterations [49], the following distributed power control scheme (also known as DCPC [50, 51] for cellular wireless systems) is obtained

$$p_i(k+1) = \min\{\frac{\gamma_i^{tar}}{\gamma_i(k)} p_i(k), \quad p_i^{max}\}\ , i = 1, 2, ..., N_c. \tag{2.17}$$

Note that each node only needs to know its own received SIR at its designated receiver to update its transmission power. This is available by feedback from the receiving node through a control channel. As a result, the algorithm is fully distributed. Convergence properties of this algorithm were studied by Yates [50]. An interference function $I(p)$, introduced in [50], is considered as standard if it satisfies three conditions: positivity, monotonicity and scalability. It was proven by Yates [50] that the standard iterative algorithm $p(k+1) = I(p(k))$ will converge to a unique equilibrium that corresponds to the minimum total transmission power. The distributed power control scheme (equation (2.17)) is a special case of the standard iterative algorithm.

Since the Jacobi iteration is a fixed-point iterative method, it usually has slow convergence speed to the sought solution. However, we select DCPC (equation (2.17)) as the power control algorithm in our proposed power aware maximally disjoint routing due to its simplicity. Other advanced algorithms with faster convergence speed can be found in [52, 53, 54, 55]. A comprehensive survey of SIR based power control for wireless networks is given by Koskie and Gajic in [56]. A review of iterative power control schemes and acceleration techniques are given in Section 3 of [56].

The complete procedure for obtaining a power controlled connectivity graph using a distributed algorithm is highlighted in Figure 2.2. The procedure will be executed for all channels. The success of concurrent transmissions within each channel is guaranteed by power control.

2.3 Power Aware Maximally Disjoint Routing

2.3.1 Multi-Path Routing vs. Single-Path Routing

In Section 1.2.3, routing algorithms for wireless ad hoc networks are briefly reviewed. The advantages and disadvantages of proactive and reactive routing schemes are summarized. Proactive routing approaches rely on an underlying routing table update mechanism that involves constant propagations of routing information. A route to every other node in the ad hoc network is always available, regardless of whether or not it is needed. This feature poses negligible delay, but incurs substantial signaling traffic and power consumption. Since both bandwidth and battery power are scarce resources in mobile ad hoc networks, this becomes a serious limitation. On the contrary, reactive routing schemes create routes on demand, which saves lots of network capacity. Hence, reactive routing schemes are desirable in wireless ad hoc networks.

The instability of the topology of mobile wireless ad hoc networks, because of node failures (due to energy loss) and link failures (due to node mobility, channel fluctuation), makes reliable data delivery a very challenging problem. To combat such problems, multi-path routing, instead of single-path routing, is a good countermeasure.

The application of multi-path routing in wireless ad hoc networks is natural, because multi-path routing helps diminishing the effect of unreliable wireless links and the constantly changing topology. Many schemes employing multipath routing have been proposed for wired networks in order to perform QoS routing [57, 58]. All of them are based on proactive routing, hence they cannot be successfully applied to wireless

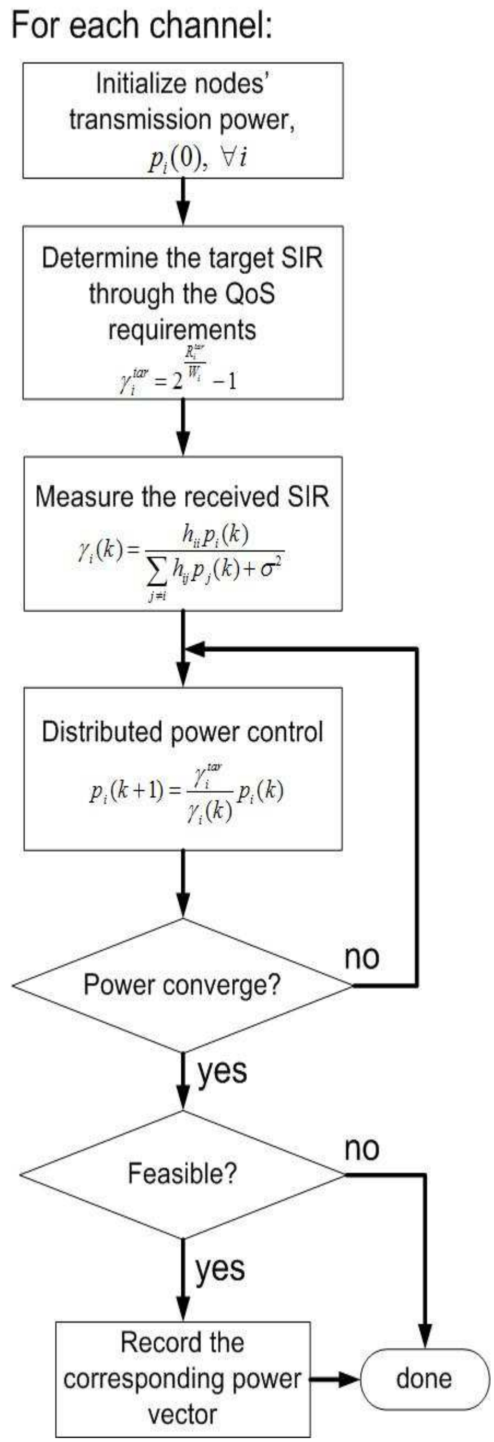

Figure 2.2: Distributed algorithm for power controlled connectivity graph.

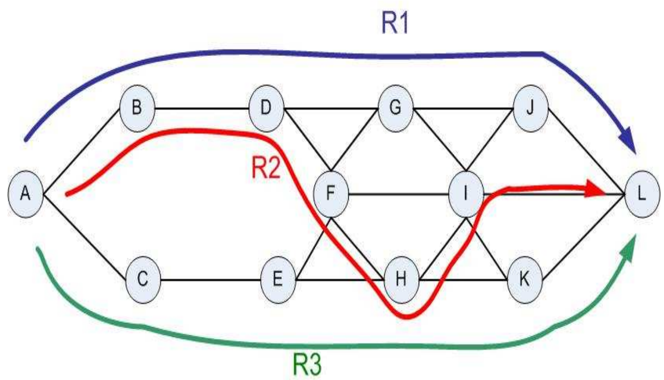

Figure 2.3: Node-disjoint vs. Link-disjoint paths.

ad hoc networks.

There are two types of disjoint paths, namely, node-disjoint paths and link-disjoint paths. Node-disjoint paths are also link-disjoint, but not vice versa. An example is illustrated in Figure 2.3. Paths R1 and R3 are node-disjoint paths (hence link-disjoint as well) since they do not share any node (except the source node A and the destination node L). On the other hand, paths R2 and R3 are link-disjoint paths because they have no common links. However, they are not node-disjoint. Since there are no common nodes along the path, node-disjoint paths are considered independent and more robust in the sense that the success or failure of one path cannot imply the success or failure of another. In this work, node-disjoint paths are preferred since they are more fault-tolerant than link-disjoint paths.

There are two ways of using the multiple paths to send data. The first approach is to send data along multiple paths *simultaneously* to achieve diversity. Examples of simultaneous transmission to achieve diversity is to either send the same data packets for redundancy [38]; or send different sub-packets using diversity coding [59, 60]. The second approach is to send data through only one path, while using the other paths as backup. Although the second approach is widely used in wired networks such as in optical networks, it has not been considered for mobile wireless ad hoc networks in the literature according to our best knowledge. The argument has been the duplicity of bandwidth and therefore for bandwidth starved wireless networks, this is a critical problem.

In our solution, we are using the second method and we are not reserving the bandwidth on the backup path. The sender keeps track of the bandwidth availability and maintain the backup path. When the primary path has failed and is not available, the backup path bandwidth is used. Therefore, for each user application, the required bandwidth is always the same and not duplicated. This solution has the following advantages:

(1). There is no complicated diversity coding scheme required. Thus, there is no excessive delay induced by waiting sub-packets from the slow path to arrive before a packet can be successfully decoded.

(2). Different traffic flows, whether they have the same source and destination or not, may share the links in their respective backup paths. This results in a much better bandwidth utilization comparing to the first approach.

(3). The packet re-ordering at the destination node during the transient phase (due to traffic shift) is much less frequent than the sub-packet re-ordering needed constantly in the first approach.

The disadvantage of the second approach is that traffic may shift back and forth if node mobility is changing much faster (orders of magnitude) than the duration of the traffic sessions. We propose a hysteresis rule for traffic shifting to mitigate this effect, as explained in detail in Section 2.4. Moreover, we should emphasize that the time constant of the mobility is on the same order or less of the duration of the traffic sessions considered in this chapter.

2.3.2 Power Aware Multipath Routing

Although the power expenditure along a route is the main concern here, other network resources such as the number of transceivers and the number of channels are also important for the success of routing, especially in the case of connection-oriented traffic [42]. In practice, routing can be done by excluding those nodes with insufficient transceivers from the topology of the network. In addition, there exist many algorithms that find efficient channel allocations, for example, see [61] and the references therein. Hence, it is assumed in this work that enough transceivers are available and proper channel allocations have been done before routing.

The routing problem is defined as follows:

(P.2)

Given the network resources at each node (such as the number of transceivers and channel allocations), find the most energy-efficient path and a maximally disjoint backup path for a given traffic flow with the required throughput in the power controlled connectivity graph.

For our approach, we consider Split Multipath Routing (SMR), introduced by Lee and Gerla [62], as the background routing scheme. SMR is an on-demand routing protocol that constructs "maximally disjoint paths". SMR is based on Dynamic Source Routing (DSR) [63] but, uses a different packet forwarding mechanism. While DSR discards duplicate routing request (RREQ), SMR allows intermediate nodes to forward certain duplicate RREQ in order to find more disjoint paths. In SMR, intermediate nodes forward the duplicate RREQ that traverse through a different incoming link than the link from which the first RREQ is received, and whose hop count is not larger than that of the first received RREQ. In SMR, a minimum power resolution is not a criterion and no desired throughput is considered. Our approach is to enhance SMR with both minimum power and balanced energy to address the minimum power resolution and to maximize the network lifetime.

Proposed Algorithm 1: Minimum Power Split Multi-Path Routing (MPSMR)

MPSMR is based on SMR. However, in MPSMR, the transmission power is used as the link metric instead of the hop count. Each RREQ has a field that records the total transmission power along a path and keeps updating the field while traversing through the network. The intermediate nodes forward the duplicate RREQ whose total power is no larger than that of the first received RREQ. The destination will choose the path with the least total transmission power and a maximally disjoint backup path.

Proposed Algorithm 2: Balanced Energy Split Multi-Path Routing (BESMR)

Sole minimization of the total consumed power per end to end delivery may drain out the power of certain nodes in the network. We argue that energy efficient routing protocols may be achieved by establishing routes that ensure that all nodes equally deplete their battery power. Therefore, instead of the transmission power, the metric p_i/E_i is proposed to balance the power efficiency and fairness among nodes. The parameters p_i and E_i are the transmission power and the remaining energy of node i, respectively. BESMR select route that minimize $\sum(p_i/E_i)$. It considers the tradeoff between the transmission power and the remaining energy of a node, thus maximizes the network's lifetime. Note that BESMR also reduces network congestions because traffic will be distributed more evenly across the network, rather than aggregated among a small set of nodes where transmission power is low.

Remark 1: In [38], a simplified interference model is used. The simplified interference model assumes that there is no interference from other transmissions, and the SIR of each link depends solely on its own received power and the background noise, i.e.,

$$\gamma_i = \frac{h_{ii}p_i}{\sigma^2} \tag{2.18}$$

where h_{ii} is the link gain from transmitter i to its designated receiver i, σ^2 is the background (receiver) noise. When the simplified interference model is applied, it is straight forward to calculate the transmission power of each link i along a path according to

$$p_i = \frac{\gamma_i \sigma^2}{h_{ii}}. \tag{2.19}$$

However, since a realistic interference model (equation (2.4)) is used in this work, an iterative algorithm is necessary to determine the transmission power and the maximally disjoint paths jointly. The procedures are listed below:

1. The transmission power of all links are initialized to the minimum power specified by the standard. An initial two maximally disjoint paths are calculated using $\sum p_i$ (for MPSMR) or $\sum(p_i/E_i)$ (for BESMR) as the routing metric. Then the

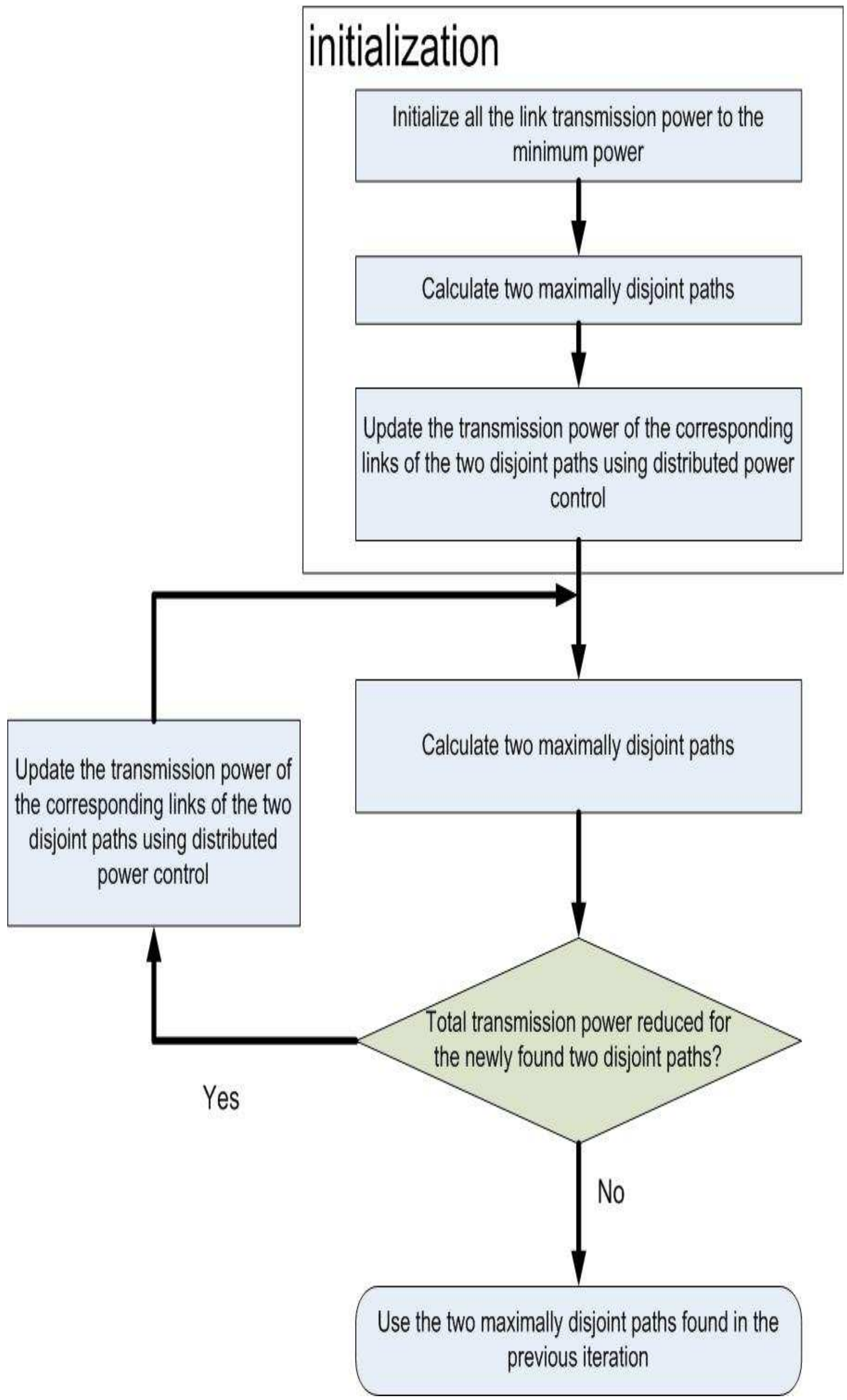

Figure 2.4: An iterative algorithm for joint power control and maximally disjoint routing.

transmission power along these two disjoint paths are updated using distributed power control (equation (4.24)) discussed in the previous section.

2. Two new maximally disjoint paths are calculated using $\sum p_i$ (for MPSMR) or $\sum(p_i/E_i)$ (for BESMR) as the routing metric.

3. If the routing metric of the two new paths are less than that of the previous two paths, then update the transmission power along these two new paths using distributed power control. Go to step 2. Otherwise, select the two disjoint paths found in the previous iteration.

The above iterative algorithm is illustrated in Figure 2.4.

Note that the proposed iterative algorithm is also valuable for call admission control. If the power control problem becomes infeasible due to a new traffic session, it will be rejected.

2.4 Dynamic Traffic Switching

The joint power control and routing scheme will be applied before each traffic session starts. In order to guarantee the required data throughput with high probability during the entire session of the traffic flow, an on-line dynamic traffic restoration scheme is indispensable to deal with node mobility or node failure. In this work, only "soft QoS" [32] is supported. In other words, there may be short transient period where QoS requirements are not guaranteed due to path break or reduced capacity. However, the QoS requirements will be ensured when the path is not broken or after the session is switched to a new path. Note that many multimedia applications accept soft QoS and use rate adaptive schemes to mitigate disruptions, for example, see [34].

There are several phases in the proposed dynamic traffic switching (restoration) scheme:

1. Initialization phase: Given the topology of a wireless ad hoc network, MPSMR or BESMR is used to find two maximally disjoint paths from the source to the destination such that the corresponding power control problem is feasible. If such

paths cannot be found, the traffic session is rejected. Otherwise, go to the next step.

2. Monitoring phase: The source node saves the two paths in its routing table and starts to send packets through the primary path. At the same time, the source also sends small amount of probe packets to monitor both paths.

3. Path switching (transient) phase: The source node monitors the throughput, delay and loss of both paths. If the throughput is below a threshold R^1_{th}, the node shifts the data traffic from the current path to the backup path. At the same time, it starts a new routing request (RREQ) using MPSMR or BESMR, and stores the newly found paths in the routing table as the new backup paths.

4. Convergence phase: If the throughput of the original path improves and increases beyond a threshold R^2_{th}, the node will shift the data traffic from the current path back to the original path.

One example of the implementation of the probe mechanism is given in [64]. The choices of the thresholds R^1_{th} and R^2_{th} depend on the traffic type (such as the compression ratio in MPEG-4) and the characteristics of the wireless ad hoc network such as node density and node mobility. Delay and loss of the path may be used to determine the traffic switching as well. The number of backup paths is another design parameter. Note that a small reduction in throughput may be compensated by adaptive modulation and coding schemes.

In order to implement the proposed scheme, a software agent for traffic monitoring and switching is installed at each node in the network. The block diagram of the agent is shown in Figure 2.5.

Usually we expect that traffic switches randomly due to the random nature of mobility pattern of mobile users, resulting topology changes, and interferences. However, it may happen that two or more traffic flows switch to paths that share the same links simultaneously. These links have a potential to be congested and the traffic flows switch simultaneously from these links. This causes instability in the traffic switching. It is resolved by enabling the source to wait a short random time before switching.

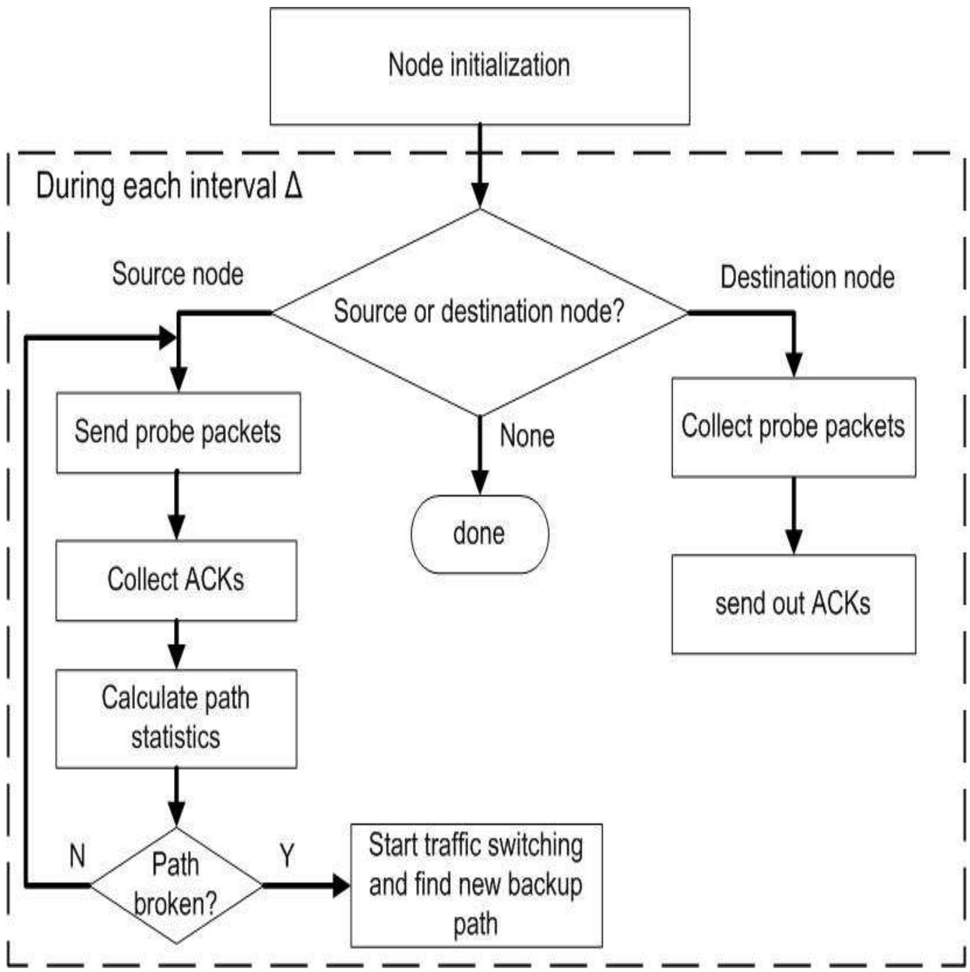

Figure 2.5: Software agent for traffic monitoring and switching.

2.5 Performance Evaluation

The performance of the proposed joint power control and maximally disjoint routing algorithms is evaluated through discrete-event simulations using OPNET. The results are compared with SMR. The dynamic traffic switching scheme is also tested.

2.5.1 Simulation Setup

In this simulation study, it is assumed that there is a fixed number ($M = 50$) of nodes located in a square area (300 meters $\times$ 300 meters). The locations of the nodes are uniformly distributed within the area. The other parameters include:

1. The required throughput, $R_i^{tar} = 250$ kbps for all the traffic sessions.

2. The bandwidth shared by all links is 1.25 MHz.

3. The link gains are assumed to be only a function of the distance, i.e., $h_{ij} = 1/d_{ij}^{\alpha}$, where $\alpha = 4$. No fading is considered here.

4. The maximum allowable transmission power p^{max} is 200 mW.

5. The background noise is $\sigma^2 = 10^{-7}$.

In addition, all the nodes are assumed to be stationary or have negligible mobility during the entire routing process such that routing and QoS provisioning will not become meaningless. However, nodes may move dramatically during traffic sessions (data forwarding).

2.5.2 Maximally Disjoint Routing with Different Interference Model

In this aspect of simulations, the source and destination are randomly chosen and the MPSMR algorithm is used to find two maximally disjoint paths with low energy expenditure. Three cases are examined with different interference models: 1) The simplified interference model (the best case); 2) The general interference model including *all* links (the worst case); 3) The general interference model including only the links within the two maximally disjoint paths. Note that the worst case corresponds to the QoS provisioning considered in [40].

In order to compare joint power control and routing schemes with different interference models, the following performance criteria are selected: 1) Average success probability (p_{succ}); 2) Energy per-bit (E_b). The first criterion (p_{succ}) focuses on the average traffic carrying capability of the network, while the second criterion (E_b) quantifies the energy efficiency of the proposed schemes.

Case	p_{succ}	E_b (in $\times 10^{-6}$ Joule/bit)	Computational Complexity
1	0.99	0.12	low
2	0.13	0.18	high
3	0.75	0.14	high

Table 2.1: Comparison of routing schemes with different interference models.

The simulation results are averaged over 100 routing attempts and are summarized in Table 2.1. It is clear that routing with the simplified interference model gives the best success probability and energy efficiency as expected. However, this model is too optimistic because it ignores the interferences. If all links (whether have data to transmit or not) are all included in the interference model, overall the worst performance is obtained due to an overly conservative situation. However, it may be useful when the network is heavily loaded. The performance of the proposed method is somewhere in between and reflects the realistic situations.

2.5.3 Comparison of SMR, MPSMR and BESMR

The performances of SMR, MPSMR and BESMR are compared in terms of energy efficiency and network lifetime. The network lifetime is defined as the time of the first node failure (running out of energy). It is assumed that all nodes have the same initial energy at the start of simulation. The source and destination of each traffic session are randomly chosen. The duration of traffic sessions are assumed to be exponentially distributed with mean equal to 1 minute. Energy efficiency is measured by the Cumulative Distribution Function (CDF) of the remaining energy at each node after the shortest lifetime of the three routing algorithms.

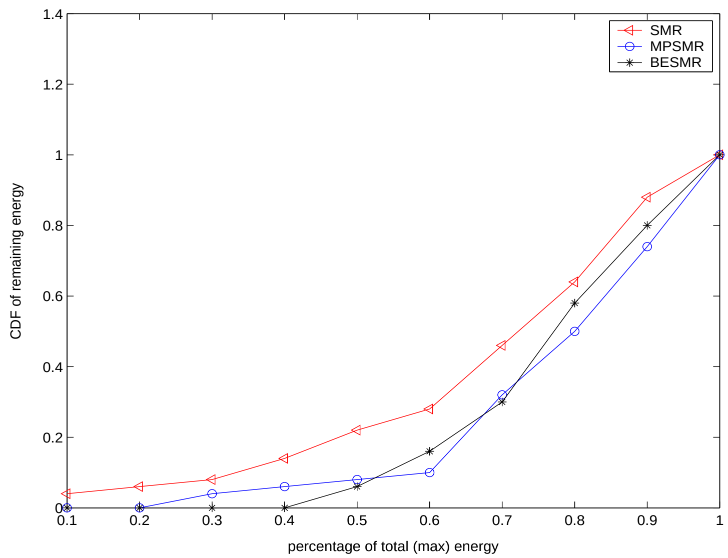

Figure 2.6: The Cumulative Distribution Function (CDF) of the remaining energy at each node.

Figure 2.6 depicts the CDF of the remaining energy at each node after the lifetime of SMR (which is the shortest among the three). It indicated that both MPSMR and BESMR have better energy efficiency than SMR (by about 15%). All nodes have more than 40% energy left using BESMR which indicates that BESMR has balanced energy usage among nodes. There are about 8% of the nodes that are heavily used (have less than 40% energy left) when MPSMR is applied.

The network lifetimes using SMR, MPSMR and BESMR are shown in Figure 2.7 for networks with 25, 50, and 100 nodes, respectively. It is clear that BESMR has the longest network lifetime because of its fairness to all nodes. A closer look at the standard deviation of the remaining energy at each node (Figure 2.8) explains that BESMR tends to balance the energy consumption among all nodes thus has the smallest standard deviation, and hence the longest network lifetime.

2.5.4 Dynamic Traffic Switching

The proposed dynamic traffic switching scheme is tested by letting a randomly selected node (other than the source and destination) on the primary path leaves the area (thus breaks the primary path) during the process of data transmission. The threshold R^1_{th}

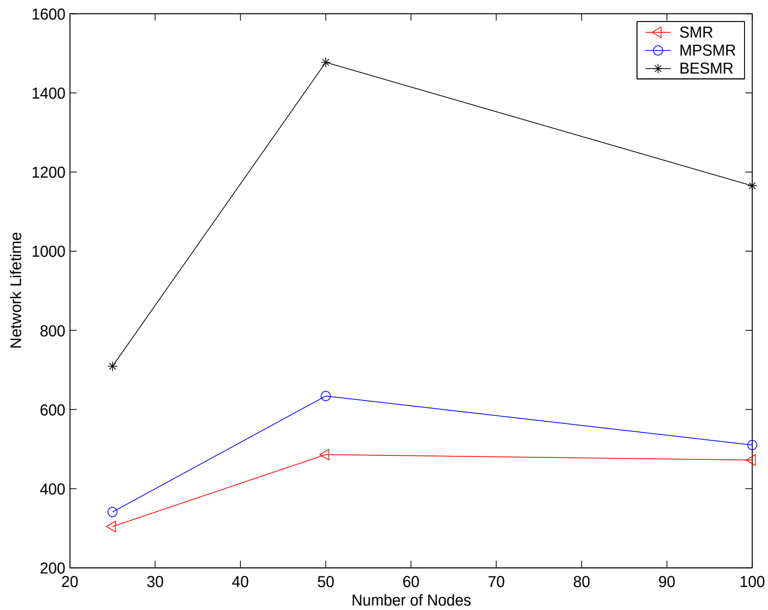

Figure 2.7: Network lifetime.

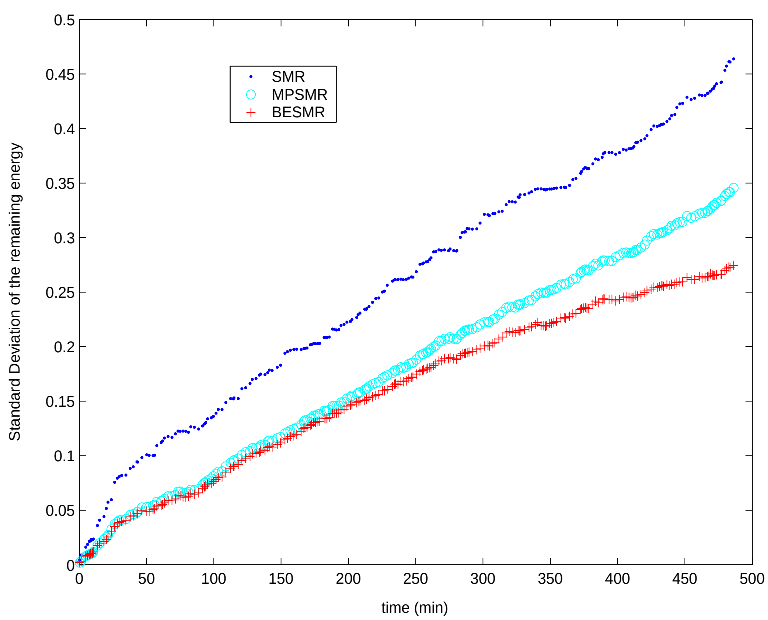

Figure 2.8: Standard deviation of the remaining energy at each node (50 nodes).

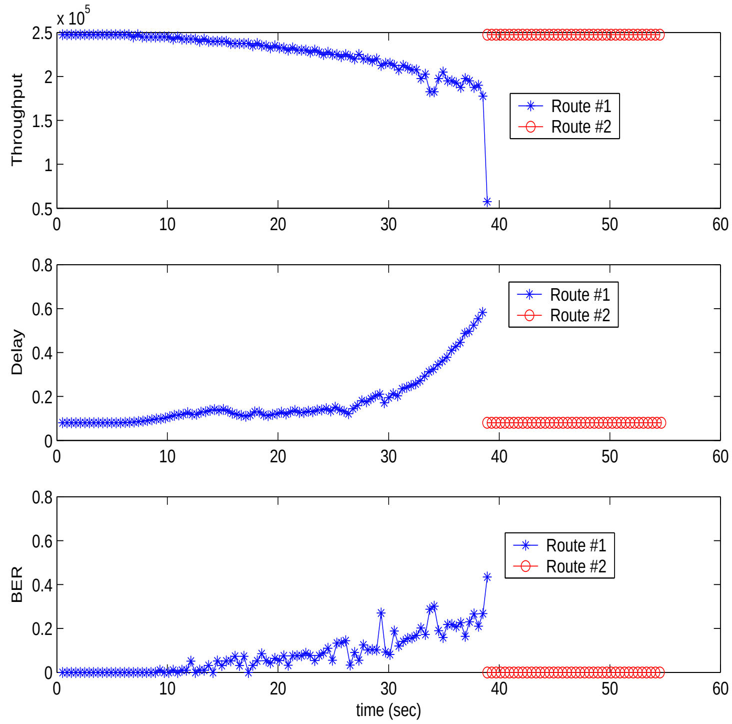

Figure 2.9: Performance index (throughput, delay and BER) during traffic switching due to node mobility.

is set to 80%.

Figure 2.9 shows the performance of the proposed traffic switching scheme when the primary path (Route #1) is broken due to node mobility. When the throughput of the primary path (Route #1) drops below 80% of the desired throughput, the traffic will be switched to the backup path (Route #2). The corresponding end-to-end delay and the bit error rate (BER) are also shown. We assume that there is only one node that moves in this simulation.

2.5.5 The Effect of Node Mobility

In this part of simulation, it is assumed that all nodes in the network are mobile and they move according to the following "random waypoint" mobility model [63]: At the beginning of each time interval, each node decides to move with probability $0 \leq q \leq 1$. If a node decides to move, it will choose a random destination and a speed vector will be sampled from a uniformly distributed random variable $v \sim [v^{min}, v^{max}]$, where v is the value of the speed. $v^{min} = 0.3$ meter/sec and $v^{max} = 0.7$ meter/sec are the lower and upper bounds of the speed, respectively.

The average number of re-routing and the average number of "effective neighbors" vs. node mobility (q) are shown in Figure 2.10. The results are averaged over 100 traffic sessions. The source and destination of each traffic session are randomly chosen. The duration of each traffic session is assumed to be exponentially distributed with the mean equal to 1 minute. Here node B is called a "effective neighbor" of node A if they are neighbors and the supported data rate between A and B is above the target data rate.

It can be observed that the number of re-routing increases with the required data rate, as expected. The number of re-routing increases with q from 0 to 0.3, however, it almost remains constant after that for the low-to-moderate required data rate. This can be explained by the average number of "effective neighbors" shown in the same figure. The average number of "effective neighbors" drops with q, however, there are still enough "effective neighbors" for the low-to-moderate required data rate. For example, there are 6 "effective neighbors" on average when $R^{tar} = 250$ kbps even when all nodes are constantly moving ($q = 1$). There are less "effective neighbors" on average for the high required data rate ($R^{tar} = 500$ kbps). The average number of neighbors drops to only 3 when all nodes are constantly moving ($q = 1$). The above simulation results are critical for network operators to set call admission control policies. Based on the estimated node mobility, traffic session duration and QoS requirements, the average number of re-routing can be estimated. Thus, the cost of supporting the traffic session with QoS can be calculated and a call admission control decision can be made accordingly.

2.5.6 Overhead and Scalability Analysis

In this aspect of the simulation, the proposed joint power control and routing plus traffic switching scheme is tested in a realistic environment. A similar setup as in Section 2.5.1 is used with the following changes:

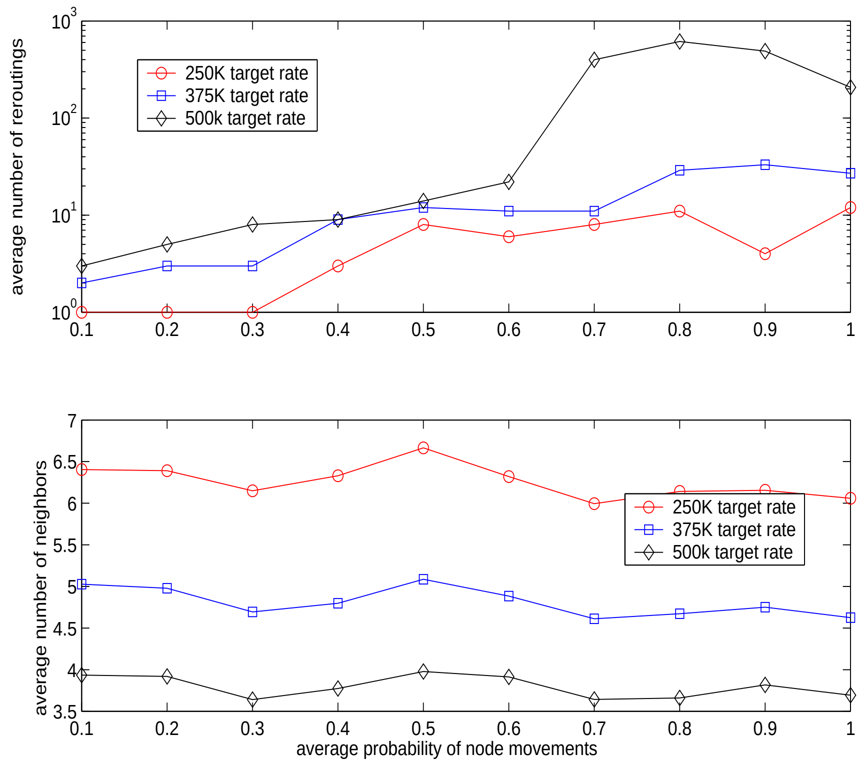

Figure 2.10: The average number of re-routing and the average number of neighbors vs. node mobility.

1. There are 80 nodes in a constrained area of 450 m $\times$ 450 m.

2. The simulation time is 10 minutes.

3. It is assumed that the link gains have the following form

$$h_{ij}(k) = d_{ij}^{-4}(k)A_{ij}(k)B_{ij}(k) \tag{2.20}$$

   where $d_{ij}(k)$ is the distance from the jth transmitter to the ith receiver at time instant k, A_{ij} is a log-normal distributed stochastic process (shadowing). B_{ij} is a fast fading factor (Rayleigh distributed).

4. It is assumed that the standard deviation of A_{ij} is 8 dB [65].

5. It is assumed that the Doppler frequency is from 8 Hz (for pedestrian mobile users) to 80 Hz (for mobile users at vehicle speed) [65].

6. All nodes in the network are constantly moving according to the "random way-point" mobility model [63], with pause time set at 10 seconds and five different velocities from 0 m/s for stationary nodes to 30 m/s for mobile users at vehicle speed.

7. Two cases with a single source/destination pair and 10 pairs are tested, respectively. All the sources are assumed to generate data packets for transmission continuously at the target rate throughout simulation. The mean packet size is 1024 bits.

The results are summarized in Table 2.2 and Table 2.3. MPSMR is chosen as the routing scheme. It is observed that there is almost no packet loss in the case of a stationary network. Routing is only needed once for each source/destination pair, and traffic switching is not required, as expected. It is also observed that the packet delivery ratio drops dramatically when all the nodes become mobile and reach vehicle speed, because the number of broken paths (thus traffic switching) increases significantly. However, it is interesting to see that 10 source/destination pairs do not yet overload the network, and the performance results (in terms of packet delivery ratio, number of traffic switching, and cost of routing) are comparable to the case of a single source/destination pair. The main reason is that data are only transmitted through one path in the proposed scheme rather than through multiple paths simultaneously, thus it avoids overloading the network. The routing overhead may be calculated as follows:

$$\eta = \frac{\#\, of\, routing\, packets \times average\, routing\, packet\, size \times \#\, of\, routing\, per\, pair}{data\, rate \times 600\, sec \times average\, \#\, of\, hops\, per\, path \times packet\, delivery\, ratio}.$$

Note that the routing overhead is about 20% in the worst case (10 source/destination pairs, 20 m/s), where the average routing packet size is 64 bits and the average number of hops per path is 5.

node velocity	packet delivery ratio		total number of traffic switching		total cost per routing (in number of routing packets)	
(m/s)	1-pair	10-pair	1-pair	10-pair	1-pair	10-pair
0	0.99	0.99	0	0	47558	55454
1	0.98	0.95	1	20	47226	71450
10	0.67	0.6	11	90	56135	90398
20	0.46	0.39	15	110	79989	83516
30	0.44	0.39	11	123	82180	68751

Table 2.2: Performance results of routing and data delivery.

The distributed power control scheme requires that the receivers provide the received SIR value (or equivalently, the link gain) to the corresponding transmitters. The power control overhead is evaluated by the number of the control packets needed for these information exchanges. It is shown in Table 2.3 that the proposed joint power control and routing scheme converges in about 5 to 6 iterations in all cases. In addition, the power control overhead does not increase too much with respect to node mobility and the number of source/destination pairs. In other words, the proposed scheme exhibits reasonable scalability in the highly mobile and high traffic load environment. Note that supporting energy efficient QoS routing in a larger network that has thousands or more nodes needs a very careful architecture design, such as a cluster based architecture and management, which is outside the scope of this work.

node velocity	number of iterations per routing		power control overhead per iteration (in number of control packets)	
(m/s)	1-pair	10-pair	1-pair	10-pair
0	6	6.9	405	405
1	5.33	5.36	642	644
10	5.19	5.99	586	598
20	5.38	5.76	569	590
30	5.88	5.64	533	561

Table 2.3: Convergence and overhead of the proposed scheme.

2.6 Conclusions

In this chapter, a joint power control and maximally disjoint routing algorithm is proposed for routing traffic between one source and destination pair with high energy efficiency. In addition, a dynamic traffic switching scheme is proposed to mitigate the effect of node mobility or node failure. Together they provide a means for reliable end-to-end data delivery with guaranteed throughput. The main contributions of this chapter are summarized as follows:

1. Joint power control and maximally disjoint multipath routing is proposed using the realistic interference model in this study rather than the simplified interference model in [38], where interference is not considered at all.

2. This study proposed the per-channel based power control, which provides a correct solution in a multihop network where only co-channel interference should be managed by power control and we do not assume that all the links are interferers to each other as has been assumed in other study [40]. Thus, our proposed design has a substantial gain (more than 300%) in terms of the capabilities of accommodating data traffic over the previous study [40] which only provided a lower bound.

3. This chapter has proposed BESMR (Balanced-Energy-Split-Multipath-Routing) as opposed to SMR in [62]. It is shown that BESMR achieves a significant gain

(more than 140%) in terms of network lifetime as compared to that of SMR technique in [62].

4. Data are only sent along the primary path rather than sending simultaneously along all the multiple paths, thus achieving high bandwidth efficiency.

5. An end-to-end traffic monitoring and switching mechanism is proposed to provide reliability against node mobility and link failures. The disturbance and delay are minimized when primary path is broken. In addition, the proposed end-to-end mechanism simplifies implementation because only the source and destination nodes are involved.

Chapter 3

Joint Power Control and Scheduling in TD/CDMA Wireless Ad Hoc Networks

In this chapter, a cluster based architecture is introduced in wireless ad hoc networks to provide centralized control within clusters, and corresponding power control and scheduling schemes are derived to maximize the network utility function and guarantee the minimum and maximum rates required by each traffic session, given routes for multiple end-to-end multihop traffic sessions. In order to achieve the high end-to-end throughput in a multihop wireless ad hoc network, TD/CDMA has been chosen as the Medium Access Control (MAC) scheme due to its support for the high network throughput in a multihop environment. The associated power control and scheduling problem is addressed to optimize the operations of TD/CDMA. Because the resulted optimal power control reveals bang-bang characteristics, i.e., scheduled nodes transmit with full power while other nodes remain silent, the joint power control and scheduling problem is reduced to a scheduling problem. The multi-link version of throughput-optimal and proportional fair scheduling algorithms for multihop wireless ad hoc networks are proposed. In addition, a generic token counter mechanism is employed to satisfy the minimum and maximum rate requirements. By ensuring a different minimum rate for different traffic sessions, service differentiation is also achieved. Approximation algorithms are suggested to reduce the computational complexity. In networks that are lacking centralized control, distributed scheduling algorithms are also derived and a fully distributed implementation is provided. Simulation results demonstrate the effectiveness of the proposed schemes.

3.1 Background and Assumptions

Wireless ad hoc networks have been the topic of extensive research recently. The interests in such networks are due to their ability to provide wireless networking capability in scenarios where no fixed wired infrastructure is available (e.g., disaster relief efforts, battlefields, etc.). The lack of fixed infrastructure introduces great design challenges. One way to reduce the difficulty is by organizing nodes into clusters and assigning management functions to certain nodes [66], such as transmission coordination. These nodes are called cluster heads. It has been shown that proper clustering in wireless ad hoc networks reduces the complexity of the link-layer and routing protocol design significantly and improves the scalability of protocols [67]. In addition, clustering increases the network capability of supporting Quality-of-Service (QoS) [68]. Clustering is also desirable because of practical reasons. For instance, in a battlefield deployment, a cluster may be naturally formed by a set of soldiers equipped with wireless communication devices with a tank serving as cluster head.

In order to resolve the issue of the low end-to-end throughput in a multihop ad hoc network, innovative medium access control protocols are indispensable. Due to their poor scalability in a multihop ad hoc network, random access protocols are not an efficient solution [69]. In [30], it is demonstrated that CDMA-based MAC protocols achieve a significant increase in the network throughput at no additional cost in energy consumption compared to 802.11x MAC protocols.

In this chapter, we restrict our interests to clustered TD/CDMA wireless ad hoc networks. It is assumed that the wireless ad hoc network is organized into clusters and each cluster has a cluster head with higher than average network resources such as power. All users/nodes within the cluster share the same frequency band and TD/CDMA is chosen as the medium access scheme. Each user/node is assigned a randomly generated orthogonal code. On top of that, time is split into equal sized slots where only scheduled users/nodes are allowed to transmit in each slot. The cluster head functions as a manager and is responsible for scheduling transmissions within a cluster. It is assumed

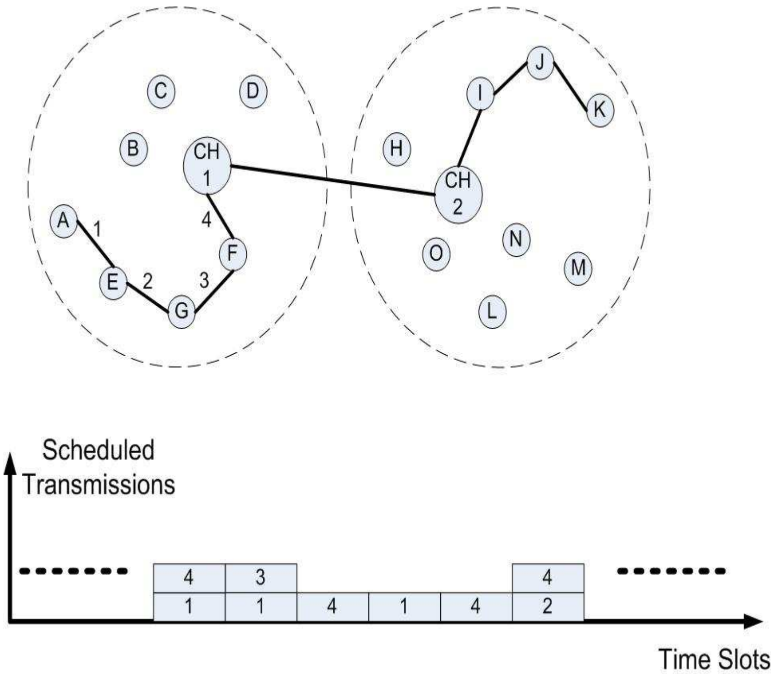

Figure 3.1: A clustered TD/CDMA wireless ad hoc network. CH: cluster head.

that communication links among cluster heads (inter-cluster communications) have sufficient bandwidth such that the bottleneck of the end-to-end traffic between nodes in different clusters resides within clusters. Hence, scheduling intra-cluster transmissions is the main concern in this chapter.

An example of a clustered TD/CDMA wireless ad hoc network is shown in Figure 3.1. There are two clusters with cluster heads $CH1$ and $CH2$, respectively. It is assumed that the intra-cluster route is given for a traffic session: $r_I = A \rightarrow E \rightarrow G \rightarrow F \rightarrow CH1$. Data traffic is forwarded in a multihop fashion. Figure 3.1 also shows a possible schedule for intra-cluster traffic transmissions.

Power control is employed in a wireless ad hoc network to control the transmission range and keep the network fully connected [31]. It is a physical layer function. However, transmission power has a direct impact on multiple access of nodes by affecting the received Signal-to-Interference-Ratio (SIR) at receivers. Hence, power control is strongly coupled with scheduling and has additional functions of reducing unnecessary interference among concurrent transmissions in TD/CDMA-based systems [70, 71, 6]. Power control and scheduling are of paramount importance of ensuring the success of multiple simultaneous transmissions and maximizing the network utility in TD/CDMA wireless ad hoc networks and it is the focus of this chapter. In this work, we are interested in traffic sessions with minimum and maximum rate constraints. The goal is to study power control and scheduling schemes that maximize certain network utility functions while providing the minimum and maximum rate of traffic sessions, given the routes and rate constraints of those sessions. Although the proposed power control and scheduling schemes focus on intra-cluster traffic transmissions, where a central controller (cluster head) is available, fully distributed versions of schemes are also developed for scenarios where no central controller is available.

The rest of the chapter is organized as follows: Section 3.2 presents an overview of works that are closely related to our problem. Section 3.3 states the wireless network model and formulates the joint power control and scheduling problem with QoS constraints. Optimal power control is given in Section 3.4. Section 3.5 gives the optimal solution to the formulated problem. Depending on the scheduling criteria, two popular schuduling algorithms are dicussed and multi-hop version of them are proposed for mobile multihop wireless ad hoc networks in Section 3.6. Section 3.7 proposes low complexity approximations, together with several algorithms that serve as lower bounds. The proposed algorithms are evaluated by extensive discrete-event simulations in Section 3.8. Distributed schemes and other various implementation issues are discussed in Section 3.9. Finally, Section 3.10 concludes the chapter.

3.2 Related Works

The power control and scheduling problem has been solved in [72] for TDMA ad hoc networks on a per frame basis and each link is assigned to a number of slots in a given frame. The authors assume that each slot has the *fixed* data rate. Using the concept of virtual links, assigning one slot to each virtual link satisfies the end-to-end session rate requirements. The joint feasibility problem is proven to be NP-complete and centralized approximation algorithms are provided. In our study, we assume a variable data rate from slot to slot due to channel fluctuations.

A centralized joint routing, scheduling and power control problem is formulated for TD/CDMA ad hoc networks and an approximation algorithm is derived in [73]. However, a simplified interference model is adopted, where no interference is assumed among different links. In [43], the authors provide long term end-to-end rate guarantees to a set of sessions at the minimum possible long term average of total transmit powers. Their main assumptions are that the system operates at significantly low SIR values and the data rate is assumed to be a linear function of SIR. Hence the transmit power is directly used as a throughput guarantee constraint. In our work, a general interference model is adopted, where each transmitting node in the network is assumed to cause interference at any receiving node, even if they are far apart. The data rate is calculated as a concave function of the SIR, which covers the entire range of SIR.

The authors in [74] proposed a joint power control and scheduling scheme based on a utility function of *instantaneous* power or *instantaneous* data rate. A degree-based greedy scheduling and an iterative power control algorithm using a penalty function approach are suggested to maximize the utility function while providing the minimum and maximum link data rates. The algorithm in [74] focused on a *snapshot* of a set of wireless links. Another work on *instantaneous* power control in wireless ad hoc networks is [6]. The authors investigate the problem of scheduling a maximum number of links in the same time slot and adapt the transmit powers to their minimum required level such that all transmissions achieve a target SIR threshold. They show that the particular system model is actually equivalent to uplink power control in TDMA cellular networks. In the case where the set of links that have buffered packets cannot be scheduled in the same time slot, these solutions do not converge and authors suggest to remove one link at a time until a feasible set of links is achieved. However, the criterion for removing the link is not precisely addressed; especially in the case of varying target SIR thresholds for each link. Also, the system model does not cover a multi-hop wireless environment. In this study, we focus on the *long-term average data rate* and the minimum average data rate requirements for traffic sessions in a routed *muti-hop* wireless ad hoc network.

A randomized policy is given to solve the multi-commodity flow problem given the long-term link capacity as weight in wireless networks [75]. Then a dynamic policy is proposed for unknown arrival and channel statistics and is proven to perform better than the randomized policy. However, no fairness among users/flows is addressed in such policies and no QoS constraint is considered in [75]. A distributed approximation is also proposed in [75] assuming that the link gains between a node and its neighbors are known. In this chapter, a family of scheduling algorithms are considered to maintain fairness and improve the performance among nodes by taking the advantage of wireless channel fluctuations. Furthermore, a token counter mechanism is introduced to maintain the minimum and maximum rate of traffic flows whenever feasible. Our proposed distributed algorithm uses a control channel to exchange link gain information [76]. In the simulation of [75], the link gains are calculated based only on distances between nodes. No fading is considered and locations of nodes are assumed to be known. In our simulation study, channel is modeled to have both shadowing and Rayleigh fading.

Power allocation and scheduling has been extensively studied for WLAN. In [77], a fully distributed algorithm for scheduling packet transmissions is proposed such that different flows are allocated bandwidth in proportion to their weights. The paper proposes a Distributed Fair Scheduling (DFS) approach obtained by modifying the Distributed Coordination Function (DCF) in IEEE 802.11 standard. A fair scheduling mechanism, distributed elastic round robin (DERR) is proposed in [78]. DERR is suitable for IEEE 802.11 wireless LANs operated in the ad hoc mode and capable of avoiding collisions through a random mapping between allowance and IFS. DERR outperforms 802.11e in terms of delay and throughput. In [79], an enhanced timer-based scheduling control algorithm is proposed to effectively manage the delay budget in IEEE 802.11e. Simulation results show that the proposed algorithm outperforms the simple scheduler control algorithm in delay and jitter in infrastructure mode. Although there are a lot of work on power allocation and scheduling for WLAN, most of them studied the infrastructure mode and focused on the random access part (DCF) in 802.11. Furthermore, to the best of our knowledge, very few papers considered an ad hoc mode and none of them considered multi-hop scenarios.

3.3 Problem Formulation: Joint Power Control and Scheduling with Minimum and Maximum Rate Constraints

In this work, it is assumed that the routes for the multiple end-to-end traffic sessions are given, and we will focus on end-to-end traffic sessions with QoS constraints. Specifically, the goal of this section is to study joint power control and scheduling schemes that maximize certain utility functions while providing the minimum and maximum rate of traffic sessions, given the routes and rate constraints of those sessions. All the links contained in the routes form the set of "active links". Each active link is uniquely identified by its transmitter and receiver. In other words, transmitter i and receiver i are the transmitter and receiver of active link i. The received SIR at the i^{th} receiver from the i^{th} transmitter (received SIR of the i^{th} active link) is defined by

$$\gamma_i = \frac{h_{ii}p_i}{\frac{1}{L}\sum_{j\neq i} h_{ij}p_j + \sigma^2} \ , \qquad i = 1, 2, \cdots, N \tag{3.1}$$

where h_{ii} is the link gain from transmitter i to its designated receiver i; h_{ij} is the link gain from transmitter j to receiver i (active link i's designated receiver); p_i and p_j are the transmission powers of transmitters i and j, respectively; σ^2 is the background (receiver) noise; and L is the spreading gain for spread spectrum systems.

In this chapter, we assume that each link has a variable rate. This rate is bounded by the feasible rate region. The link gains (channel quality) may fluctuate dramatically from one slot to another slot. In other words, the data rates of the active links are different from slot to slot during traffic sessions. A scheduling scheme should take advantage of channel fluctuations, i.e., it should be "channel-aware".

The instantaneous data rate of each active link can be evaluated by the Shannon capacity formula (for AWGN channel)

$$R_i = W_i \log_2(1 + \gamma_i) \tag{3.2}$$

where W_i is the bandwidth occupied by the transmission from the i^{th} transmitter to its designated receiver. Note that this formula gives the achievable rate (upper bound) of the AWGN channel. However, it is justified by the fact that with the current modulation and coding technology it can be closely approximated in most practical scenarios [44].

The interference model adopted here assumes that each transmitting node in the network causes interference to any receiving nodes, even if they are far apart. This model is considered as more realistic than the one which assumes that transmitting nodes only cause interference to their neighbors. This is because the aggregate interference from a large number of nodes may not be negligible even if the interference from each of them is small. The instantaneous data rate will be determined solely by the received SIR.

A guarantee on a minimum rate is arguably the simplest possible QoS guarantee. Therefore we believe it is natural that mobile users would expect such an assurance. Other reasons of ensuring a minimum rate are:

1. Some applications need a minimum rate in order to perform well. For example, streaming audio and video can become unusable if the data rate is too low.

2. Even for static TCP-based applications such as web browsing, if the data rate is too low then we typically get a large queue buildup which can lead to TCP timeouts and a poor performance. Such effects were discussed by Chakravorty et al. in [35].

3. Providing a minimum rate guarantee can help to smooth out the effects of a variable wireless channel.

4. Providing a minimum rate can allow us to ensure that a slot-based TD/CDMA service is no worse than circuit-based data systems such as wireline dialup or 3G1X wireless service.

5. By setting a minimum data rate differently for different users we can ensure service differentiation.

At first, it might seem counterintuitive to set maximum data rate constraints. However, there are some possibe reasons:

1. Maximum data rate constraints may be necessary for portable wireless devices that have limited buffering capabilities.

2. The high sensitivity of transport protocols, like TCP loss due to buffer overflow in wireless networks, makes the maximum data rate constraints necessary.

3. From a commercial market view, if a user pays only a cheap data service, the operator might wish to cap his data rate in order to give him/her an incentive to upgrade to a premium service.

We remark that if the system operator does not want to set the maximum data rate constraints then this is easily accomplished by setting ${R_i}^{max}$ to infinity. Given the routes of multiple end-to-end traffic sessions with the minimum and maximum rate constraints, our approach follows the Gradient algorithm with the Minimum/Maximum Rate constraints (GMR) developed in [80]. Let's define the long-term average rate vector $\bar{\mathbf{R}} = (\bar{R}_1, \ldots, \bar{R}_N)$ assuming that there are N active links resulting from routing, and each of the active links has the minimum rate constraint (${\bar{R}_i}^{min}$), and the maximum rate constraint (${\bar{R}_i}^{max}$). The joint power control and scheduling problem is formulated as the following optimization problem

(P.3)

$$\max_{\mathbf{R}\in\mathcal{R},\mathbf{p}\in\mathcal{P}} U(\bar{\mathbf{R}}) \tag{3.3}$$

subject to

$${\bar{R}_i}^{max} \geq \bar{R}_i \geq {\bar{R}_i}^{min} \tag{3.4}$$

where the instantaneous rate is determined by equations (3.1) and (3.2). $\mathcal{R}$ is the rate region, which is the set of long-term service rate vectors which the system is capable of providing. $\mathcal{P}$ is the set of allowable power vectors defined by

$$0 \leq p_i \leq p_i^{max} \ , \ \forall\, i \tag{3.5}$$

where p_i^{max} is the maximum allowable transmission power of transmitter i, p_i^{min} is the minimum transmission power. The utility function is of the form

$$U(\bar{\mathbf{R}}) = \sum_i U_i(\bar{R}_i) \tag{3.6}$$

where each $U_i(x)$ is an increasing concave continuously differentiable function defined for $x \geq 0$.

A node can not transmit and receive simultaneously. This primary conflict [73] is resolved by setting the link gain matrix appropriately. For example, if node i is selected to transmit in the current slot, the corresponding link gains where node i is the receiver will be set to zero.

The multi-hop nature of the problem **(P.3)** reflects the fact that the links on the same route require the same minimum rate whereas links on different routes typically have different minimum rate requirements. In other words, the order of transmissions along a route is implicitly included in the problem formulation.

3.4 Optimal Power Control

Before introducing the Multi-link Gradient algorithm with the Minimum and Maximum Rate constraints (MGMR) to solve the optimization problem **(P.3)**, we observe some useful properties of the optimal solution.

Theorem 2 *The optimal scheme has the property that each transmitting node transmits at full power, i.e.* $p_i = p_i^{max}$ *for some subset* $\mathcal{S}$ *of the nodes and with* $p_i = 0$ *for the complementary set* $\bar{\mathcal{S}}$*.*

Proof: Assume that there are N active links (transmitter-receiver pairs) in the network. Let p_{ii}^{rcv} and p_{ij}^{rcv} be the instantaneous received power at receiver i from transmitter i and j, respectively. For simplicity, we express p_{ii}^{rcv} and p_{ij}^{rcv} in units of the background noise σ^2. In order to meet the minimum rate constraints of all active links, we must have for each active link i

$$\frac{p_{ii}^{rcv}}{\frac{1}{L}\sum_{j\in\{1,2,\cdots,N\},j\neq i} p_{ij}^{rcv} + 1} \geq \gamma_i^{tar} \tag{3.7}$$

where γ_i^{tar} is the required SIR of link i. If a desired data flow rate is specified by a certain application, say, R_i^{tar}, then γ_i^{tar} can be expressed as

$$\gamma_i^{tar} \geq 2^{\frac{R_i^{tar}}{W_i}} - 1 , \qquad i = 1, 2, \cdots, N \tag{3.8}$$

The feasible SIR vectors specified in (3.7) is adapted from that in cellular wireless networks to multihop wireless networks. Given the peak received power $p_{ii}^{rcv,max} =$

$h_{ii}p_i^{max}$ and $p_{ij}^{rcv,max} = h_{ij}p_j^{max}$, we may change variables to $\theta_{ii} = \frac{p_{ii}^{rcv}}{p_{ii}^{rcv,max}}$ and $\theta_{ij} = \frac{p_{ij}^{rcv}}{p_{ij}^{rcv,max}}$ to rewrite (3.7) as

$$\frac{\theta_{ii}p_{ii}^{rcv,max}}{\frac{1}{L}\sum_{j\in\{1,2,\cdots,N\},j\neq i}\theta_{ij}p_{ij}^{rcv,max}+1} \geq \gamma_i^{tar} \tag{3.9}$$

A given SIR vector is feasible if (3.9) can be satisfied with equality with $0 \leq \theta_{ij} \leq 1$ for all i and j. We hence examine the solution subject to the set of linear equations

$$\frac{\theta_{ii}p_{ii}^{rcv,max}}{\gamma_i^{tar}} = \frac{1}{L}\sum_{j\in\{1,2,\cdots,N\},j\neq i}\theta_{ij}p_{ij}^{rcv,max}+1 \tag{3.10}$$

which can be further rewritten as

$$\theta_{ii}p_{ii}^{rcv,max}(1+\frac{L}{\gamma_i^{tar}}) = \sum_{j\in\{1,2,\cdots,N\}}\theta_{ij}p_{ij}^{rcv,max}+L \tag{3.11}$$

It can be seen by inspection that the solution is in the form $\theta_{ii}p_{ii}^{rcv,max}(1+\frac{L}{\gamma_i^{tar}}) = C$ where C is a global parameter. The value of C can be obtained by substituting the postulated solution in (3.11) to obtain $C = C\sum_j \frac{\gamma_j^{tar}}{\gamma_j^{tar}+L} + L$, which gives the final solution

$$\theta_{ii} = \frac{\gamma_i^{tar}}{(L+\gamma_i^{tar})p_{ii}^{rcv,max}}\frac{L}{[1-\sum_{j\in\{1,2,\cdots,N\}}\frac{\gamma_j^{tar}}{\gamma_j^{tar}+L}]} \tag{3.12}$$

Defining $\alpha_i = \frac{\gamma_i^{tar}}{\gamma_i^{tar}+L}$ we see that

$$\theta_{ii} = \frac{L\alpha_i/p_{ii}^{rcv,max}}{1-\sum_j \alpha_j} \tag{3.13}$$

Clearly, $0 \leq \alpha_i < 1$. Since we require that $0 \leq \theta_{ii} \leq 1$, equation (3.13) results in the following feasibility conditions to meet the required SIRs

$$\sum_j \alpha_j + \frac{L\alpha_i}{p_{ii}^{rcv,max}} \leq 1 \;\; \forall i\ . \tag{3.14}$$

Note the simple linear form of the feasible SIRs in terms of α_i.

Note that the utility function $U_i(\bar{R}_i)$ is a concave function and $\bar{R}_i$ is a linear combination of the instantaneous data rate R_i. The instantaneous data rate is again a concave function of the SIR. Since (3.14) is linear in α_i, it is more convenient to consider the R, α relationship which is now convex. The optimization problem **(P.3)** becomes

$$\max_{\mathbf{R}\in\mathcal{R},\mathbf{p}\in\mathcal{P}} \sum_i U_i(\bar{R}_i) \tag{3.15}$$

subject to

$$\sum_j \alpha_j + \frac{L\alpha_i}{p_{ii}^{rcv,max}} \leq 1 \ \ , \ \ \alpha_i > 0 \ \ \forall i \tag{3.16}$$

Equations (3.16) specify $2N$ constraints on the feasible α_i. From standard theorems on convex maximization with linear constraints, it is easy to see that the optimum occurs at *corner points* of (3.16) due to the joint-convexity of (3.15) in the α_i. Corner points of (3.16) have exactly N of $2N$ constraints binding, i.e., some subset of α_i are null, while the complementary set saturates their respective constraints in the first equation of (3.16). Combining this observation with (3.13) results in $\theta_{ii} = 1$ for the complementary set, thus proving the theorem.

■

Note that similar observations are obtained under various different contexts and assumptions [75, 43, 81, 82]. Specifically, the results reported in [43] may be viewed as a special case of the above theorem where the data rate is assumed to be a linear function of SIR instead of a more general form that is adopted in this section. Theorem 2 reveals the bang-bang characteristics of the nodes' transmission power in order to maximize the network's utility. In each time slot, selected transmitting nodes will use the maximum transmission power, while other nodes remain silent.

3.5 Multi-link Gradient Algorithm with Minimum and Maximum Rate Constraints (MGMR)

As highlighted on Theorem 2, the joint power control and scheduling problem is reduced to a scheduling problem given the bang-bang characteristics of the optimal transmission power. The Multi-link Gradient scheduling algorithm with the Minimum and Maximum Rate constraints (MGMR) will be proposed to solve the optimization problem **(P.3)**. Several types of implementation will be addressed in detail in the following section.

MGMR: *In a time slot k, select the active links according to*

$$\arg\max_{\mathbf{R}\in\mathcal{R}} \sum_i e^{a_i T_i(k)} U_i'(\bar{R}_i(k)) R_i(k) \tag{3.17}$$

where $\bar{R}_i(k)$ is the current average service rate received by link i, $T_i(k)$ is a "token counter" for link i, and $a_i > 0$ is a parameter. The values of the average rate $\bar{R}_i$ are *updated as in the Proportional Fair algorithm [83, 84]:*

$$\bar{R}_i(k+1) = (1-\beta)\bar{R}_i(k) + \beta R_i(k)$$

where $\beta > 0$ *is a small fixed parameter, and* $R_i(k)$ *is the instantaneous data rate if link* i *is actually served in slot* k *and* $R_i(k) = 0$ *otherwise. The token counter* T_i *is updated as follows:*

$$T_i(k+1) = T_i(k) + \bar{R}_i^{token} - R_i(k) \tag{3.18}$$

where $\bar{R}_i^{token} = \bar{R}_i^{\min}$ *if* $T_i(k) \geq 0$, *and* $\bar{R}_i^{token} = \bar{R}_i^{\max}$ *if* $T_i(k) < 0$. *If* $\bar{R}_i^{\max} = \infty$ *for some* i, *that means no maximum rate constraints, the token counter update rule becomes:*

$$T_i(k+1) = \max\{0, T_i(k) + \bar{R}_i^{min} - R_i(k)\} \tag{3.19}$$

If $\bar{R}_i^{\min} = 0$ *for some* i, *that means no minimum rate constraints, the rule is simplified for this* i *to:*

$$T_i(k+1) = \min\{0, T_i(k) + \bar{R}_i^{\max} - R_i(k)\} \tag{3.20}$$

Proof: We prove the optimality of the MGMR algorithm by studying the dynamics of user throughputs and token counters under the MGMR algorithm when parameters β and a_i are small. Namely, we consider the asymptotic regime such that β converges to 0, and each $a_i = \beta\alpha_i$ with some fixed $\alpha_i > 0$. We study the dynamics of *fluid sample paths* (FSP), which are possible trajectories $(r(t), \tau(t))$ of a random process which is a limit of the process $(\bar{R}(t/\beta), \beta T(t/\beta))$ as $\beta \to 0$. (Thus, $r(t)$ approximates the behavior of the vector of throughputs $\bar{R}(t)$ when β is small and we "speed-up" time by the factor $1/\beta$; $\tau(t)$ approximates the vector $T(t)$ scaled down by factor β, and with $1/\beta$ time speed-up.) The main result is a "necessary throughput convergence" condition stated in the following theorem:

Theorem 3 *Suppose FSP* (r, τ) *is such that*

$$r(t) \to \bar{R}^* \quad as\ t \to \infty$$

and $\tau(t)$ *remains uniformly bounded for all* $t \geq 0$. *Then,* $\bar{R}^*$ *is a solution to the problem* **(P.3)** *and, moreover,* $\bar{R}^* \in \mathcal{R}^{cond} \cap \mathcal{R}^* \neq \emptyset$.

The rate region $\mathcal{R}$ is a convex closed bounded polyhedron in the positive orthant. By $\mathcal{R}^*$ we denote the subset of maximal elements of $\mathcal{R}$: namely, $v \in \mathcal{R}^*$ if conditions $v \leq u$ (component wise) and $u \in \mathcal{R}$ imply $u = v$. Clearly, $\mathcal{R}^*$ is a part of the outer boundary of $\mathcal{R}$. The subset $\mathcal{R}^{cond} \subseteq \mathcal{R}$ of elements $v \in \mathcal{R}$ satisfying conditions $\bar{R}_i^{\min} \leq v_i \leq \bar{R}_i^{\max}$ for all i, is also a convex closed bounded set.

The proof of Theorem 3 is given below. Theorem 3 says that if FSP is such that the vector of throughputs $r(t)$ converges to some vector $\bar{R}^*$ as $t \to \infty$, then $\bar{R}^*$ is necessarily a solution to the problem **(P.3)**. This implies that if the user throughputs converge, then the corresponding stationary throughputs do in fact maximize the desired utility function, subject to the minimum and maximum rate constraints. ■

Suppose a stochastic matrix $\phi = (\phi_{mi}, m \in M, i = 1, \ldots, N)$ is fixed, which means that $\phi_{mi} \geq 0$ for all m and i, and $\sum_i \phi_{mi} = 1$ for every m. Consider a *Static Service Split* (SSS) scheduling rule, parameterized by the matrix ϕ. When the server is in state m, the SSS rule chooses for service queue i with probability ϕ_{mi}. (Sometimes, matrix ϕ itself is called the SSS rule.) Clearly, the vector $v = (v_1, \ldots, v_N) = v(\phi)$, where

$$v_i = \sum \pi_m \phi_{mi} \mu_i^m$$

gives the long term average service rates allocated to different flows under the SSS rule ϕ.

We define the system *rate region* to be the set $\mathcal{R}$ of all vectors $v(\phi)$ for all possible SSS rules ϕ. Thus $\mathcal{R}$ is the set of long-term service rate vectors which the system is capable of providing. The rate region $\mathcal{R}$ is a convex closed bounded polyhedron in the positive orthant. By $\mathcal{R}^*$ we denote the subset of maximal elements of $\mathcal{R}$: namely, $v \in \mathcal{R}^*$ if conditions $v \leq u$ (componentwise) and $u \in \mathcal{R}$ imply $u = v$. Clearly, $\mathcal{R}^*$ is a part of the outer ("north-east") boundary of $\mathcal{R}$.

The subset $\mathcal{R}^{cond} \subseteq \mathcal{R}$ of elements $v \in \mathcal{R}$ satisfying conditions $\bar{R}_i^{\min} \leq v_i \leq \bar{R}_i^{\max}$ for all i, is also a convex closed bounded set.

Since each function $H_i(\bar{R}_i)$ is continuous and increasing, we have the following simple fact.

Proposition 1 *If $\mathcal{R}^{cond}$ is non-empty, then at least one solution $\bar{R}^*$ of problem* **(P.3)** *exists. If $\mathcal{R}^{cond}$ contains at least one point of the set $\mathcal{R}^*$, then any solution $\bar{R}^* \in \mathcal{R}^*$. If all functions $H_i(\bar{R}_i)$ are strictly concave (for example, $H_i(\bar{R}_i) = \log(\bar{R}_i)$), a solution $\bar{R}^*$ is unique.*

We use the following notations, the sets of real numbers and non-negative real numbers are denoted by $\mathbb{R}$ and $\mathbb{R}_+$ respectively; $\mathbb{R}^N$ and $\mathbb{R}_+^N$ denote their N times products. For vectors $x, y \in \mathbb{R}^N$,

$$x \cdot y \doteq \sum_i x_i y_i \text{ is scalar product}$$

$$x \times y \doteq (x_1 y_1, \ldots, x_N y_N) \text{ is component-wise product,}$$

$$\exp(x) \doteq (\exp(x_1), \ldots, \exp(x_N))$$

The Euclidean norm $\|x\| \doteq \sqrt{x \cdot x}$ defines metric $\|x - y\|$ on $\mathbb{R}^N$. The gradient of the function H is denoted by ∇H, that is

$$\nabla H(x) = (H_1'(x_1), \ldots, H_N'(x_N))$$

For a function $\xi = (\xi(t),\ t \geq 0)$, $\theta_d \xi$ denotes its backward shift by time $d \geq 0$, namely

$$[\theta_d \xi](t) = \xi(t + d),\ t \geq 0$$

It is easy to show using *(fluid sample paths)*(FSP) properties described below in Lemmas 1 and 2, that if $\mathcal{R}^{cond} \cap \mathcal{R}^* = \emptyset$, then for any FSP the vector $\tau(t)$ cannot remain bounded and in fact $\|\tau(t)\| \to \infty$ as $t \to \infty$. Therefore, the uniform boundedness of $\tau(t)$ alone implies that $\mathcal{R}^{cond} \cap \mathcal{R}^* \neq \emptyset$.

To prove Theorem 3, we will first describe the basic FSP properties in Lemmas 1 and 2. Then we prove two special (increasingly general) cases of Theorem 3 in Lemmas 3 and 4, and conclude with the proof of Theorem 3 itself.

Lemma 1 *For any fluid sample path, all its component functions are Lipschitz continuous in $[0, \infty)$, with the Lipschitz constant upper bounded by $C + \|r(0)\|$, where $C > 0$ is a fixed constant depending only on the system parameters.*

Proof is analogous to that in [85]. ∎

Since all component functions of an FSP are Lipschitz, they are absolutely continuous, and therefore almost all points $t \geq 0$ (with respect to Lebesgue measure) are such that all component functions of an FSP have derivatives.

Lemma 2 *The family of fluid sample paths satisfies the following additional properties.*

(i) For almost all $t \geq 0$ (with respect to Lebesgue measure) we have:

$$r'(t) = v(t) - r(t) \tag{3.21}$$

where

$$v(t) \in \arg\max_{v \in \mathcal{R}} [\exp(\alpha \times \tau(t)) \times \nabla H(r(t))] \cdot v \tag{3.22}$$

and

$$\tau'(t) = \alpha(r^{token}(t) - v(t)) \tag{3.23}$$

where the components $r_i^{token}(t),\ i = 1, \ldots, N$ of vector $r^{token}(t)$ are such that

$$r_i^{token}(t) \begin{cases} = \bar{R}_i^{\min} & \text{if } \tau_i(t) > 0, \\ \in [\bar{R}_i^{\min}, \bar{R}_i^{\max}] & \text{if } \tau_i(t) = 0, \\ = \bar{R}_i^{\max} & \text{if } \tau_i(t) < 0. \end{cases} \tag{3.24}$$

(ii) "Shift property." If (r, τ) is an FSP, then for any $d \geq 0$, $(\theta_d r, \theta_d \tau)$ is also an FSP.

(iii) "Compactness." If a sequence of FSPs $(r^{(j)}, \tau^{(j)}) \to (r, \tau)$ uniformly on compact sets as $j \to \infty$, then (r, τ) is also an FSP.

The proof of properties (i)(3.21) and (iii) is completely analogous to that of the corresponding FSP properties in [85]. Property (i)(3.23) is easy to verify directly, using the definition of an FSP. The shift property (ii) (as well as compactness (iii)) is an inherent property of fluid sample paths, valid for FSPs defined in many different settings; and it is easily verified directly as well. ∎

Lemma 3 *Suppose* (r, τ) *is a stationary FSP, namely*

$$r(t) \equiv \bar{R}^* \text{ and } \tau(t) \equiv \tau^* \quad \text{for all } t \geq 0$$

Then, $\bar{R}^*$ *is a solution to the problem* **(P.3)** *and* $\bar{R}^* \in \mathcal{R}^{cond} \cap \mathcal{R}^* \neq \emptyset$.

Proof. Let vector $\eta \in \mathbb{R}^N$ be defined as $\eta \doteq \exp(\alpha \times \tau^*)$. In view of property (3.21), it follows from $r(t) \equiv \bar{R}^*$ that we have $v(t) \equiv \bar{R}^*$ as well. By (3.22), for almost all $t \geq 0$ we have

$$v(t) \in \arg\max_{v \in \mathcal{R}} [\eta \times \nabla H(r(t))] \cdot v.$$

We see (since $v(t) \equiv \bar{R}^*$ and $r(t) \equiv \bar{R}^*$) that $v = \bar{R}^*$ solves the problem

$$\max_{v \in \mathcal{R}} [\eta \times \nabla H(\bar{R}^*)] \cdot v \tag{3.25}$$

or, equivalently, the problem

$$\max_{v \in \mathcal{R}} [\nabla H(\bar{R}^*) \cdot v + \lambda^{min} \cdot v - \lambda^{max} \cdot v] \tag{3.26}$$

where the vectors $\lambda^{min}, \lambda^{max} \in \mathbb{R}^N_+$ have the following components:

$$\lambda_i^{min} = \max\{(\eta_i - 1) H_i'(\bar{R}_i^*), 0\} \geq 0,$$

$$\lambda_i^{max} = -\min\{(\eta_i - 1) H_i'(\bar{R}_i^*), 0\} \geq 0$$

Adding the constant $-\lambda^{min} \cdot \bar{R}^{min} + \lambda^{max} \cdot \bar{R}^{max}$ to the objective function in (3.26), we see that $v = \bar{R}^*$ maximizes the Lagrangian

$$\nabla H(\bar{R}^*) \cdot v + \lambda^{min} \cdot (v - \bar{R}^{min}) - \lambda^{max} \cdot (v - \bar{R}^{max})$$

for the optimization problem

$$\max_{v \in \mathcal{R}} [\nabla H(\bar{R}^*) \cdot v] \tag{3.27}$$

subject to constraints

$$v \geq \bar{R}^{min} \quad \text{and} \quad v \leq \bar{R}^{max} \tag{3.28}$$

Moreover, the complimentary slackness conditions are satisfied for the (Lagrange multipliers) λ_i^{min} and λ_i^{max}. Indeed, if for some i we have $\bar{R}_i^* > \bar{R}_i^{\min}$, then $\tau_i^* \leq 0$ (otherwise, by (3.23)-(3.24), $\tau_i(t)$ could not possibly be constant), and therefore $\eta_i \leq 1$.

This means that $\bar{R}_i^* > \bar{R}_i^{\min}$ implies $\lambda_i^{min} = 0$. Using an analogous argument, we see that $\bar{R}_i^* < \bar{R}_i^{max}$ implies $\lambda_i^{max} = 0$.

Thus, by the Kuhn-Tucker theorem (cf. [86]), $v = \bar{R}^*$ solves the problem (3.27)-(3.28), which is equivalent to the problem

$$\max_{v \in \mathcal{R}^{cond}} \nabla H(\bar{R}^*) \cdot v \tag{3.29}$$

This in turn means that point $\bar{R}^*$ is a maximal point of the set $\mathcal{R}^{cond}$ (i.e., it lies on its outer - "north-east" - boundary), and that vector $\nabla H(\bar{R}^*)$ is normal to the (convex) set $\mathcal{R}^{cond}$ at point $\bar{R}^*$. This implies that $\bar{R}^*$ is a solution to **(P.3)**.

Since $\bar{R}^*$ solves the problem (3.25), $\bar{R}^*$ is a point on the outer boundary of the entire rate region $\mathcal{R}$, i.e. $\bar{R}^* \in \mathcal{R}^*$. This implies that $\bar{R}^*$ belongs to the (non-empty) intersection of $\mathcal{R}^{cond}$ and $\mathcal{R}^*$. ∎

Lemma 4 *Suppose FSP (r, τ) is such that*

$$r(t) \equiv \bar{R}^* \quad \textit{for all } t \geq 0$$

and $\tau(t)$ remains uniformly bounded for all $t \geq 0$. Then, $\bar{R}^*$ is a solution to the problem **(P.3)** *and $\bar{R}^* \in \mathcal{R}^{cond} \cap \mathcal{R}^* \neq \emptyset$.*

Proof. As shown in the proof of Lemma 3, $v(t) \equiv \bar{R}^*$. Then, it follows from (3.23)-(3.24) that $\bar{R}_i^* \in [\bar{R}_i^{\min}, \bar{R}_i^{\max}]$ for each i - otherwise $\tau_i(t)$ could not remain bounded. Consider a function $\tau_i(\cdot)$. If $\tau_i(0) \geq 0$ and $\bar{R}_i^* = \bar{R}_i^{\min}$ then (from (3.23)-(3.24)) $\tau_i(t) \equiv \tau_i(0)$ for $t \geq 0$. If $\tau_i(0) \geq 0$ and $\bar{R}_i^* > \bar{R}_i^{\min}$ then $\tau_i(t)$ will decrease linearly at the rate $\bar{R}_i^{\min} - \bar{R}_i^*$ until it hits 0, and then will stay at 0. Similarly, if $\tau_i(0) \leq 0$, $\tau_i(t)$ either stays at $\tau_i(0)$ (in the case $\bar{R}_i^* = \bar{R}_i^{\max}$) or increases linearly until it hits 0 and then stays at 0 (in the case $\bar{R}_i^* < \bar{R}_i^{\max}$). Thus, for some fixed $d \geq 0$ and a fixed vector τ^*, we must have $\tau(t) \equiv \tau^*$ for $t \geq d$. The time shifted path $(\theta_d r, \theta_d \tau)$ is also an FSP, and, as we have shown above, it is stationary. An application of Lemma 3 completes the proof. ∎

Proof of Theorem 3. For each integer $d \geq 0$, consider the FSP $(r^{(d)}, \tau^{(d)}) \doteq (\theta_d r, \theta_d \tau)$, which is a time shifted version of (r, τ). Since all component functions of all FSPs $(r^{(d)}, \tau^{(d)})$ are uniformly Lipschitz continuous (because $\|r(t)\|$ is uniformly bounded) and the sequence of functions $r^{(d)}(\cdot)$ converges uniformly to the function identically equal to $\bar{R}^*$, we can choose a subsequence $(r^{(j)}, \tau^{(j)})$ converging (uniformly on compact sets) to a path (r°, τ°) such that $r^\circ(t) \equiv \bar{R}^*$ and $\tau^\circ(\cdot)$ being uniformly bounded. But, the path (r°, τ°) is also an FSP. Application of Lemma 4 completes the proof. ∎

3.6 Scheduling Algorithms

As highlighted in Theorem 2, the joint power control and scheduling problem is reduced to a scheduling problem given the bang-bang characteristics of the optimal transmission power. Scheduling algorithms provide mechanisms for bandwidth allocation and multiplexing at the packet level. Many QoS scheduling algorithms have been developed for wireline networks. The time-varying wireless channel makes the development of effective scheduling algorithms for wireless networks very challenging. Two facets should be considered in the scheduling problem:

1: Efficiency. Since wireless resources are scarce, it is important to efficiently use the channel by exploiting time varying channel conditions.

2: Fairness. Because the wireless channel is a shared medium over which many users compete for resources. Hence, it is important to allocate the shared resource fairly.

However, both the time varying nature of wireless channels and different channel conditions for different users, pose new requirements and important tradeoff between these two facets.

Firstly, scheduling fairness in a wireless network is more complicated than that in a wireline network. Scheduling fairness in a wireline network is usually guaranteed by dedicating a certain service rate to a flow, and the scheduling algorithm prevents different flows from interfering with each other. However, the fairness issue is different in wireless networks because of the time-dependent and user-dependent channel conditions. It may happen that a packet is scheduled for transmission on a wireless link according to a certain fairness guideline, which is independent of the link state, and the link is actually in an error state. If the packet is transmitted, it will be corrupted and the transmission resources will be wasted. In that case, deferring transmission of this packet till the link recovers from the error state is clearly a reasonable choice. The affected flow temporarily loses its share of the transmission bandwidth. To ensure fairness, the flow should be compensated for the loss later when the link recovers.

Secondly, there are tradeoffs between the efficiency and fairness. Because of the time varying channel conditions, good scheduling schemes should opportunistically exploit channel conditions to achieve a better performance, that means schedule users based on favorable channel conditions. At the same time, this leads to the fairness issue. For example, only schedule users in favorable channel conditions may result in a very good network performance, but some unfortunate users may never been served. Hence, how to balance the tradeoff between the efficiency and fairness is a design focus in this chapter.

Thirdly, there have been lots of works focused on scheduling algorithms in downlink and uplink cellular networks [53, 87]. However, the multi-hop nature of wireless ad hoc networks poses some new issues. And it has been shown in [88] that the use of multi-hop can greatly increase the network capacity. Although, there is significant research on the fairness and efficiency issues in a single-hop wireless network, research addressing multihop fairness and efficiency is still rarely found. How to design new scheduling algorithms adjusted to multi-hop wireless ad hoc networks is another design focus in this chapter.

Based on the above discussions, we can classify scheduling algorithms into *channel-state-aware* schemes and *channel-state-oblivious* schemes. We focus on two very important channel-state-aware scheduling schemes, namely, proportional fair scheduling and throughput optimal scheduling. While proportional fair scheduling achieves a good tradeoff between the efficiency and fairness, throughput-optimal scheduling considers queue stability and it is suitable for real time applications.

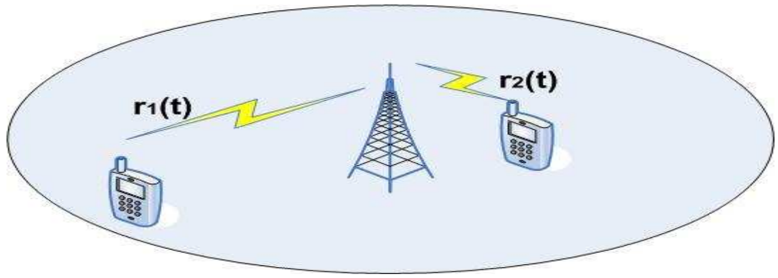

Figure 3.2: A scheduling example in 3G CDMA HDR downlink.

3.6.1 Proportional Fair Scheduling Algorithms

The "proportional fairness" is first defined by Kelly [89]. The proportional fair scheduling is proposed for wireless system in [83, 84] and further analyzed in [90, 91]. A Proportional Fair (PF) scheduling algorithm was proposed and implemented by QualComm for 3G1X EVDO (HDR) downlink. Since wireless channel capacity is time-dependent and user-dependent, in other words, the capacity (service rate) of the channel varies with time randomly and asynchronously for different users, scheduling schemes that take into account channel conditions will do a better job than channel state oblivious scheduling schemes, such as round robin.

The following example illustrates the point. We consider 3G CDMA HDR downlink (from base station to users) scenario [92], where in each time slot data can be transmitted to only one user. Each user reports to the base station its channel condition through uplink, and the base station decides which user to schedule in each time slot. Consider a simple system with two users as illustrated by Figure 3.2. Assume that channels for both users are independent, and there is unlimited amount of data to transmit for each user. User 1 can achieve the data rate of 76.8 kb/s or 153.6 kb/s with equal probabilities 0.5. User 2, which has a better channel condition on average, can achieve the data rate of 153.6 kb/s or 307.2 kb/s also with equal probabilities. Then, using the channel state oblivious *round robin* scheme to schedule users will result in users to achieve the following average rates: $R_1 = 0.5 \times (0.5 \times 76.8 + 0.5 \times 153.6) = 57.6$ kb/s and $R_2 = 0.5 \times (0.5 \times 153.6 + 0.5 \times 307.2) = 115.2$ kb/s, respectively.

Instead, if we use proportional fair scheduling proposed for HDR in [83, 84], we schedule a user with a "relatively better"channel condition (153.6 kb/s for user1 and

307.2 kb/s for user 2). In case of a tie, when a channel is relatively better or relatively worse for both users, the user to serve is chosen randomly with equal probabilities. And we can achieve these average rates: $R_1 = 0.5 \times 0.25 \times 76.8 + 0.25 \times 153.6 + 0.5 \times 0.25 \times 153.6 = 67.2$ kb/s and $R_2 = 0.5 \times 0.25 \times 153.6 + 0.25 \times 307.2 + 0.5 \times 0.25 \times 307.2 = 134.4$ kb/s, respectively. This is a 16 percent higher for each user than the round robin scheme. The above simple example shows that the proportional fair scheduling scheme takes advantage of channel state fluctuation and have a clear advantage over the channel state oblivious scheduling schemes.

We have mentioned that balancing the tradeoff between the efficiency and fairness is a very important design issue. As pointed out in [93], proportional fair scheduling maintains a balance between fairness and efficiency. We study this problem by considering two extremes, one is called "max-min" fairness, the other extreme is to maximize efficiency. Max-min fairness is a very popular scheme that is used in many wireline networking protocols, such as in the ABR mode of ATM. The intuitive notion of max-min fairness is that any user is entitled to as much performance/resource as any other user. It is a equalitarian approach by which the rate of a flow can be increased only when it is not possible to increase the rate of an already smaller flow. Max-min fairness is also used, often implicitly, in many existing wireless networks, including 802.11 networks. It turns out that the issue in a wireless network is significantly different from that in a wireline network. Due to the "solidarity" property [94] of the set of feasible rates, max-min fairness has a fundamental inefficiency problem. It is shown in [94] that the max-min fair transport rate makes all rates equal. It implies that all flows, including the most inefficient ones, have an equal rate. Given that the wireless channel condition is user dependent, max-min fairness drags all the users' performance to the worst user's performance.

The other extreme is that we maximize efficiency and ignore fairness. It is straightforward to show that we should choose the "best" users (i.e., the users with the highest achievable rate) to transmit. We propose the following algorithms as a benchmark. One is called the Multi-link Maximum Throughput (MMT) algorithm, which schedules the best users to transmit. The other is called the Multi-link Maximum Throughput with the Minimum and Maximum rate constraints (MMTMR) algorithm by incorporating a token counter mechanism inspired by the scheme developed for cellular systems [80].

MMT : *The Multi-link Maximum Throughput (MMT) algorithm corresponding to utility functions* $U(\bar{\mathbf{R}}) = \sum_i \bar{R}_i$*, and the scheduling rule is*

$$\arg\max_{\mathbf{R}\in\mathcal{R}} \sum_i R_i(k) \tag{3.30}$$

MMTMR: *The Multi-link Maximum Throughput with Minimum and Maximum Rate Constraints (MMTMR) algorithm corresponding to utility functions* $U(\bar{\mathbf{R}}) = \sum_i \bar{R}_i$*, and incorporating a token counter mechanism to meet rates requirements, the scheduling rule is*

$$\arg\max_{\mathbf{R}\in\mathcal{R}} \sum_i e^{a_i T_i(k)} R_i(k) \tag{3.31}$$

Instead of the "silly" fairness of max-min fairness, which maximizes the network efficiency ignoring user fairness at all, a better strategy is to schedule "relatively-best" users. Here, utility fairness is used as an alternative to the equalitarian approach as in max-min fairness. It corresponds to the utility metric $\sum_i U(R_i)$ where R_i is the rate of flow i, and $U()$ in an increasing concave continuously differentiable function. A special case is called proportional fairness, defined by Kelly [89], which has $U(R) = log(R)$.

It is shown by Tse [84] that proportional fairness can be achieved in a cellular network by scheduling the user which has the largest ratio of the achievable data rate at the current time slot to the average rate that it has been allocated so far. And a proportional fair algorithm provides fairness among users such that in the long run each user receives the same number of time slots of services. However, since the proportional fair scheme schedules users one-at-a-time, it needs to be modified for a multihop scenario.

In this section, we are interested in proposing and studying the multi-link version of the proportional fair algorithms for multihop wireless ad hoc networks, called Multi-link Proportional Fair (MPF). We are particularly interested in its modified version that accommodate QoS constraints required by multiple traffic sessions. MPF is modified to satisfy the minimum and maximum rate constraints using a token counter mechanism inspired by the scheme developed for cellular systems [80], thus it is named Multi-link Proportional Fair with the Minimum and Maximum Rate constraints (MPFMR).

MPFMR is a special case of the MGMR algorithm, which is based on utility function, where $U(\bar{\mathbf{R}})$ is based on proportional fair criteria:

MPFMR : *The Multi-link Proportional Fair with the Minimum and Maximum Rate constraints (MPFMR) algorithm corresponding to utility functions* $U(\bar{\mathbf{R}}) = \sum_i \log(\bar{R}_i)$*, and the scheduling rule is*

$$\arg\max_{\mathbf{R}\in\mathcal{R}} \sum_i e^{a_i T_i(k)} \frac{R_i(k)}{\bar{R}_i(k)} \tag{3.32}$$

In this study, we also considered the scheduling algorithm that solves a similar optimization problem as **(P.3)**, however, *without* the minimum and maximum rate constraints (equation (3.4)). The resulted special cases is

MPF : *The Multi-link Proportional Fair (MPF) algorithm corresponding to utility functions* $U(\bar{\mathbf{R}}) = \sum_i \log(\bar{R}_i)$*, and the scheduling rule is*

$$\arg\max_{\mathbf{R}\in\mathcal{R}} \sum_i \frac{R_i(k)}{\bar{R}_i(k)} \tag{3.33}$$

3.6.2 Throughput Optimal Scheduling Algorithms

The proportional fair scheduling algorithm discussed above is good for the best effort traffic. It achieves a long term fairness by taking advantage of channel variations and scheduling the user when it is in its "relatively better" channel state. However, this algorithm is not very efficient for real-time users. For example, suppose that the two users in the above simple example, as illustrated by Figure 3.2, running a streaming audio that requires a minimum rate of 85 kb/s. Is it possible to support this rate for both users? In this section, we will consider another type of scheduling algorithms, throughput-optimal scheduling, to achieve such goals.

Throughput optimality is defined in [92] as follows: *A scheduling algorithm is throughput optimal if it is able to keep all queues stable, if this is at all feasible to do with any scheduling algorithm.* The intuitive reasoning behind this is that if you want to achieve certain QoS requirements (suppose that users receive audio streams, a necessary QoS requirement is the delay below certain threshold, or the rate larger than the minimum rate requirement), obviously the scheduling algorithm must be able to keep all queues stable, that is, be able to handle all of the offered traffic without having the queues overflowing. Throughput-optimal scheduling algorithms are proposed in [92, 95, 75, 96], where a weighted sum of user rates is maximized for each scheduling interval. This choice has provable stability properties shown in much previous work in various contexts involving data scheduling and resource allocation. The weights may be chosen to optimize one of many possible performance measures, including average queue length, delay, or corresponding percentiles, and other similar criteria. A version of this type of algorithms that guarantees *queue stability*, i.e. boundedness of queue lengths when feasible, is specified as the rate choice that satisfies

$$\mathbf{R}^* = \arg\max_{\mathbf{R}\in\mathcal{R}} \mathbf{Q}\cdot\mathbf{R}$$

where $\mathbf{R}$, $\mathbf{Q}$ are the rate and queue vectors of the user set respectively, and $\mathcal{R}$ is the *rate region*, or the set of feasible rate vectors. The minimum/maximum instantaneous rate guarantees may be satisfied by restricting the rate region $\mathcal{R}$ appropriately. The key feature of this algorithm is that a scheduling decision depends on both current channel conditions and state of the queues, hence, it is more efficient for the real-time traffic. However, an obvious drawback of this scheme is no traffic policing is enforced [97]. If one or more sources misbehave and increase their arrival rates so that the set of arrival rates lies outside of the capacity region, then the system becomes unstable. Hence, we propose the Multi-link Throughput Optimal with the Minimum and Maximum Rate constraints (MQRMR) algorithm by incorporating a token counter mechanism to satisfy the minimum and maximum rate constraints. By setting the appropriate maximum rate constraints, user's misbehavior can be limited. And with the help of the token counter mechanism, the simple two user system example (described before) will indeed be able to achieve 85 kb/s streaming data rate. MQRMR is a special cases of the MGMR algorithm, where the utility function, $U(\bar{\mathbf{R}})$, is based on the throughput optimal criteria:

MQRMR : *The Multi-link Throughput Optimal with the Minimum and Maximum Rate constraints (MQRMR) algorithm corresponding to utility functions* $U(\bar{\mathbf{R}}) = \sum_i Q_i \bar{R}_i$*, where* Q_i *is the queue backlog at the transmitter of link* i*, and the scheduling rule is*

$$\arg\max_{\mathbf{R}\in\mathcal{R}} \sum_i e^{a_i T_i(k)} Q_i R_i(k) \ . \tag{3.34}$$

In this study, we also considered a scheduling algorithm that solves a similar optimization problem as **(P.3)**, however, *without* the minimum and maximum rate constraint (equation (3.4)). The resulted special case is

MQR : *The Multi-link Throughput Optimal (MQR) algorithm corresponding to utility functions* $U(\bar{\mathbf{R}}) = \sum_i Q_i \bar{R}_i$*, where* Q_i *is the queue backlog at the transmitter of link* i*, and the scheduling rule is*

$$\arg\max_{\mathbf{R}\in\mathcal{R}} \sum_i Q_i R_i(k) \ . \tag{3.35}$$

3.7 Low Complexity Approximations

In this part, we attempt to provide a greedy, low-complexity, approximate solution to the optimization problem **(P.3)** discussed before. The optimal solution needs to sort all possible combinations of active links. In order to run the scheduler in real-time, low complexity approximations are needed. We hence propose the following simple scheduling scheme (greedy algorithms that rank active links by their respective measure) that may be more suitable for practical implementations.

3.7.1 Greedy Algorithms

In each time slot

1. Create a list by sorting active links in a decreasing order of the measure v_i assuming no interference from other active links while computing R_i^0.

2. Add active link j, in the order starting from the top of the list, while maintaining and updating the value of $\Phi = \sum_{i \leq j} v_i$, where R_i now takes into the account interference from all added active links.

3. Stop if adding the next active link reduces Φ, and allow transmission of all added active links at their peak powers and rates as computed.

The measure v_i for different algorithms are:

1. MPFMR: $v_i = e^{a_i T_i} \frac{R_i^0}{\bar{R}_i}$,

2. MMTMR: $v_i = e^{a_i T_i} R_i^0$,

3. MQRMR: $v_i = e^{a_i T_i} Q_i R_i^0$,

4. MPF: $v_i = \frac{R_i^0}{\bar{R}_i}$,

5. MMT: $v_i = R_i^0$,

6. MQR: $v_i = Q_i R_i^0$.

We also considered several algorithms that will serve *one* active link in each time slot. These algorithms serve as the lower bound for the performance comparison.

3.7.2 One-at-a-time Algorithms

Create a list by sorting active links in decreasing order of the measure v_i assuming no interference from other active links while computing R_i^0. Serve the top on the list. The measure v_i for different algorithms are:

1. PF: $v_i = \frac{R_i^0}{\bar{R}_i}$,

2. MT: $v_i = R_i^0$,

3. PFMR: $v_i = e^{a_i T_i} \frac{R_i^0}{\bar{R}_i}$,

4. MTMR: $v_i = e^{a_i T_i} R_i^0$,

5. QRMR: $v_i = e^{a_i T_i} Q_i R_i^0$.

		Through-Optimal	Proportional Fair
Multi-Link Algorithms	without Min Rate	MQR	MPF
	with Min Rate	MQRMR	MPFMR
One-at-a-time Algorithms	without Min Rate	QR	PF
	with Min Rate	QRMR	PFMR
Implementation		Queue backlog needed	Average rate needed
Comments		Session rates maximized No fairness considered	Guarantee long term fairness

Table 3.1: Scheduling algorithms for TD/CDMA wireless ad hoc networks.

The various scheduling algorithms considered in this dissertation are summarized in Table 3.1.

3.8 Performance Evaluation

One benchmark algorithm is the optimal (centralized) MGMR algorithm given in the previous section. It gives the best possible performance. Other benchmark algorithms are one-at-a-time algorithms, which will serve as lower bounds. We will compare with these algorithms to evaluate the gains of different optimal/sub-optimal multi-link algorithms. Round Robin and fully simultaneous transmissions are considered too far from optimal and perform very poorly in most cases, and are thus ignored here.

3.8.1 Simulation Setup

In order to quantify the performance gain by applying optimal/sub-optimal scheduling algorithms, discrete-event simulations using OPNET have been performed to evaluate them in multihop TD/CDMA wireless ad hoc networks. Networks of two types of topologies and corresponding routing configurations are tested, see Figure 3.3 (linear topology) and Figure 3.4 (network with crossover traffic). It is assumed that routes are given for fixed destinations and marked with arrows in the Figures. There is one route (r_I) for destination node F in the linear network. There are three routes (r_{II}, r_{III}, and r_{IV}) for destination nodes L, J, K, respectively. The links on the routes are indexed with numerical numbers.

The routing setups represent important scenarios in multihop wireless ad hoc networks. The linear model is considered as the simplest case of relaying traffic sequentially and represents intra-cluster traffic to a fixed destination (cluster head). Figure 3.4 shows a general model where there are multiple data collection nodes such as cluster heads or data gathering gateways in wireless sensor networks.

In order to quantify the performance of different algorithms, all the nodes generate traffic such that the network is fully loaded, i.e., each node will have enough data to transmit at any time slot. It is also assumed that the traffic sources are Poisson with different inter-arrival time for different traffic sessions. The packet length is exponentially distributed with mean 1024 bits.

In this simulation study, we will use the time-averaged service rate as the criterion to compare different algorithms for fully loaded networks. Individual as well as total average rates are considered for comparison. It will quantify the traffic carrying capability of the entire network.

In order to measure the QoS-support capability for specific traffic sessions, we also define the *effective rate along a route/path* ($\bar{R}_r^{eff}$) as the minimum average rate among all the links in the path r, i.e.,

$$\bar{R}_r^{eff} = \min_{i \in r} \bar{R}_i \tag{3.36}$$

A higher effective rate of a path implies a higher QoS-support capability.

Four routes/paths are of interests here. There is route one (r_I) from node A to node F in the linear network, whereas there are three routes traversing through the network in Figure 3.4 with crossover traffic, namely, $r_{II} : A \rightarrow D \rightarrow E \rightarrow H \rightarrow I \rightarrow L$, $r_{III} : B \rightarrow E \rightarrow G \rightarrow J$ and $r_{IV} : C \rightarrow F \rightarrow H \rightarrow K$. Suppose there are traffic sessions along each route, and their respective minimum rate requirements are $\bar{R}_I^{min} = 160$ kbps, $\bar{R}_{II}^{min} = 90$ kbps, $\bar{R}_{III}^{min} = 190$ kbps and $\bar{R}_{IV}^{min} = 100$ kbps. The goal is to examine various algorithms and decide whether they could support the required minimum rate.

In the simulation we further make the following assumptions:

1. The scheduling decision is made by a central controller in every time slot. We use 1.6667 msec time slot as defined in 3G1xEV-DO (HDR) [98].

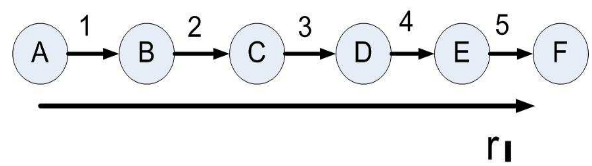

Figure 3.3: A linear TD/CDMA wireless ad hoc network.

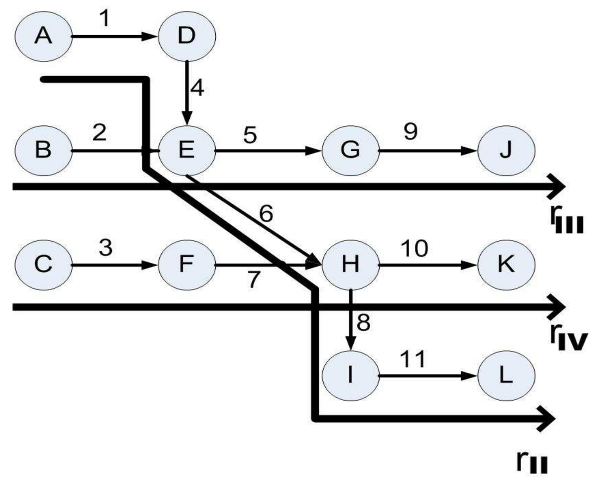

Figure 3.4: A TD/CDMA wireless ad hoc network with crossover traffic.

2. It is assumed that the link gains have the following form

$$h_{ij}(k) = d_{ij}^{-4}(k)A_{ij}(k)B_{ij}(k) \tag{3.37}$$

where $d_{ij}(k)$ is the distance from the jth transmitter to the ith receiver at time instant k, A_{ij} is a log-normal distributed stochastic process (shadowing). B_{ij} is a fast fading factor (Rayleigh distributed).

3. It is assumed that $d_{ij}(k)$ is a uniformly distributed random variable between 150 and 250 meters.

4. It is assumed that the standard deviation of A_{ij} is 8 dB [65].

5. It is assumed that the Doppler frequency is 8 Hz, corresponding to pedestrian mobile users [65].

6. It is assumed that all users share 1.25 MHz bandwidth.

7. It is assumed that the maximum allowable transmission power is $p^{max} = 200$ mW for all nodes.

Algorithms	$\bar{R}_{r_I}^{eff}$	$\bar{\mathbf{R}}$	support $\bar{R}_I^{min} = 160$kbps?
PF	95.5	314.1	No
MPF (G)	123.7	436.1	No
PFMR	155.8	204.2	No
MPFMR (G)	170.1	301.2	Yes
QR	261.9	272	Yes
MQR (G)	266.6	279.8	Yes
MQR (O)	268.2	281.2	Yes
MQRMR (G)	245.3	262.2	Yes
MMTMR (G)	186.9	474.5	Yes

Table 3.2: Effective rate and total average rate (both in kbps) in the linear network. (G):Greedy algorithm; (O):Optimal algorithm.

8. Simulation time = 40,000 slots.

In order to study the detailed behavior of each algorithm, the slot occupancy rate of each link i (η_i) is also an important quantity. It is defined as the percentage of slots assigned to link i. Note that in Multi-link algorithms, one slot may be assigned to multiple links simultaneously.

3.8.2 Linear Network

The results of the linear network are summarized in Table 3.2. We observe that the throughput-optimal family of algorithms (QR, MQR, MQRMR) have achieved better effective rates ($\bar{R}_{r_I}^{eff}$) than that of the proportional fair family of algorithms (PF, MPF, PFMR, MPFMR) for a single traffic session. In general, the throughput-optimal family of algorithms tends to balance the average rate along each traffic session/flow as long as the system is feasible because the optimization criterion (network utility function) addresses the queue backlog together with the average data rate. On the other hand, the proportional fair family of algorithms try to assign each link similar amount of slots (in the long-term) and thus will not balance the average rate along the routes. However, they tend to achieve a higher total average data rate ($\bar{\mathbf{R}}$) because they take advantage of wireless channel fluctuations and give more slots to links with a better channel quality than that of the throughput-optimal family of algorithms.

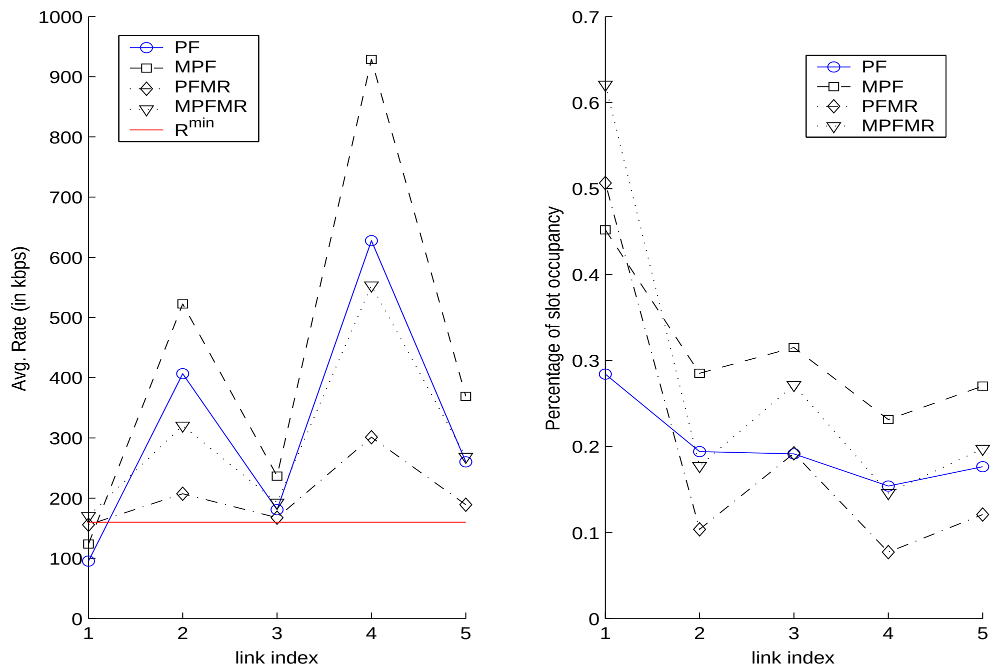

Figure 3.5: Comparison of PF-family of algorithms in a linear TD/CDMA wireless ad hoc network.

We also observe that the multi-link algorithms outperform the one-at-a-time counterparts as expected. For example, the MPF outperform PF 30% in the effective rate and 39% in the total average rate, respectively. The results also show that the greedy algorithm (for example, MQR (G)) performs very closely to the optimal algorithm (MQR (O)).

The proposed token counter mechanism helps to lift the minimum rate, and hence the effective rate. PFMR has lifted the minimum rate from PF's 95.5 kbps to 155.8 kbps, while MPFMR has lifted the minimum rate from MPF's 123.7 kbps to 170.1 kbps. Of course, this is achieved by assigning more slots to links that violate the minimum rate constraints. As a result, the links that may get higher service rates will be assigned less slots, which results in a lower total average data rate. This effect can be better observed in Figure 3.5.

In Figure 3.5, the average rate (in kbps) and the percentage of slot occupancy of all five links in the linear network are plotted when PF-family of algorithms are employed. It is clear that multi-link algorithms (MPF and MPFMR) outperform their one-at-a-time counterpart (PF and PFMR) by allowing that multiple links transmit at the same

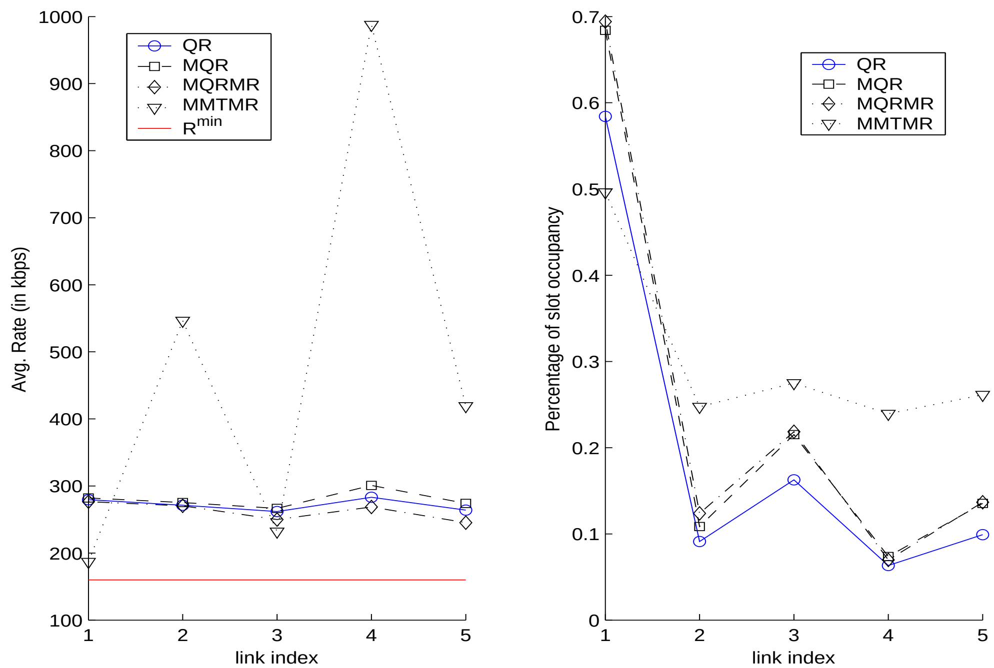

Figure 3.6: Comparison of Throughput-Optimal family of algorithms in a linear TD/CDMA wireless ad hoc network.

slot. The plot also shows that link 1 needs help to achieve the minimum rate. PFMR and MPFMR use the token counter mechanism to assign more slots to link 1 than PF and MPF, from 29% to 51% and from 45% to 62%, respectively. As a result, other links will receive less slots assignments and thus have less average rates.

Figure 3.6 shows the average rate (in kbps) and the percentage of slot occupancy of all five links in the linear network when throughput-optimal family of algorithms are employed. They tend to balance the average rate along the route as discussed before.

3.8.3 Network with Balanced Crossover Traffic

Simulations of a network with crossover traffic reveals similar observations as those obtained in the linear network. Figure 3.7 and Fig 3.8 show the average rate (in kbps) of all the links along each of the three routes of the PF-family of algorithms and the throughput-optimal family of algorithms, respectively. As long as the network load is feasible, the throughput-optimal family of algorithms provides a higher effective rate than the PF-family of algorithms. On the other hand, the PF-family of algorithms

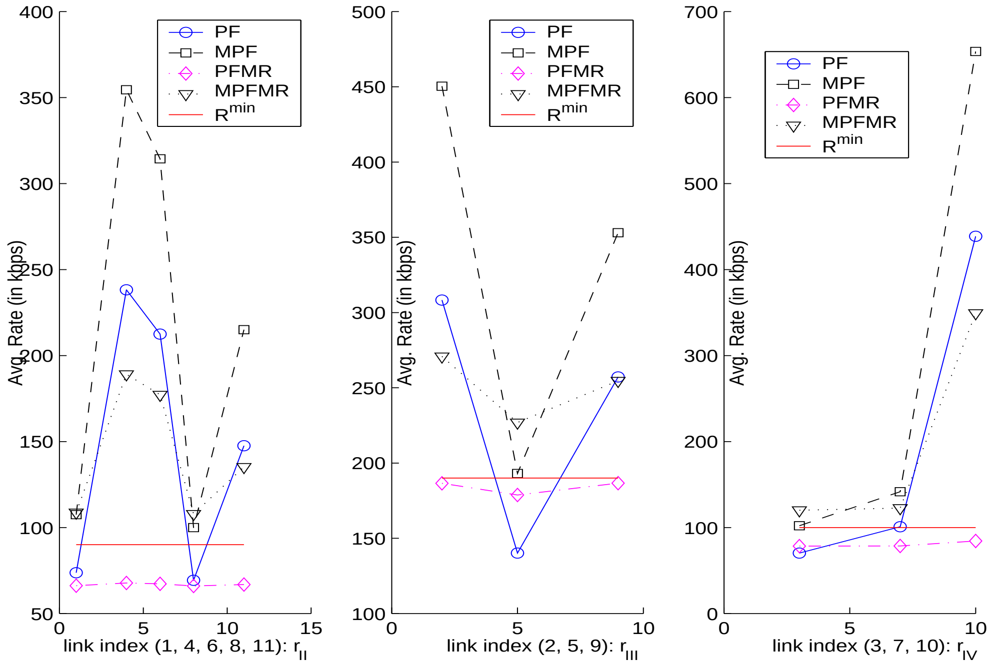

Figure 3.7: Comparison of PF-family of algorithms with balanced crossover traffic sessions.

provides higher total average rate than the throughput-optimal family of algorithms. Note that if the total average rate is the only concern, then the MMT and MMTMR algorithms should be used. However, these algorithms consider neither queue occupancy nor fairness among nodes.

In order to verify the feasibility of the network load, all the queues at all the nodes have to be bounded. A sample of the queue occupancy for all five nodes along r_{II} using algorithms MQR in the network with crossover traffic is given in Figure 3.9. All queue lengths are bounded below 10^5 bits through the entire simulation, which demonstrates the feasibility of the network load and the throughput-optimal nature of the MQR algorithm.

3.8.4 Network with Unbalanced Crossover Traffic

The above experiments show that the throughput-optimal family of algorithms outperform the PF-family of algorithms in terms of the effective rate of traffic sessions. However, it is noticeable that the throughput-optimal family of algorithms provides no

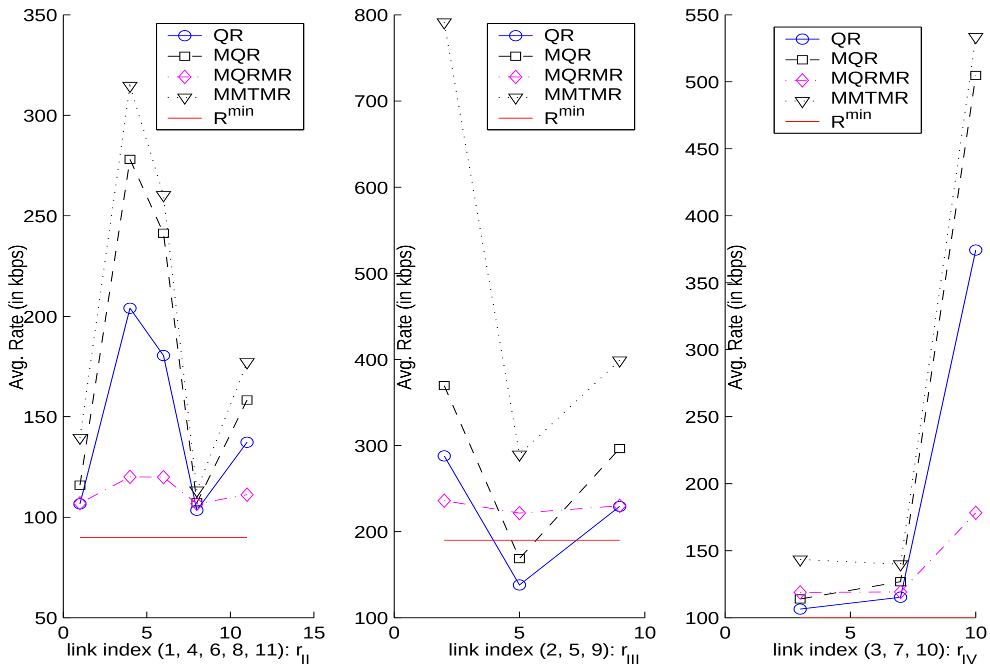

Figure 3.8: Comparison of Throughput-Optimal family of algorithms with balanced crossover traffic sessions.

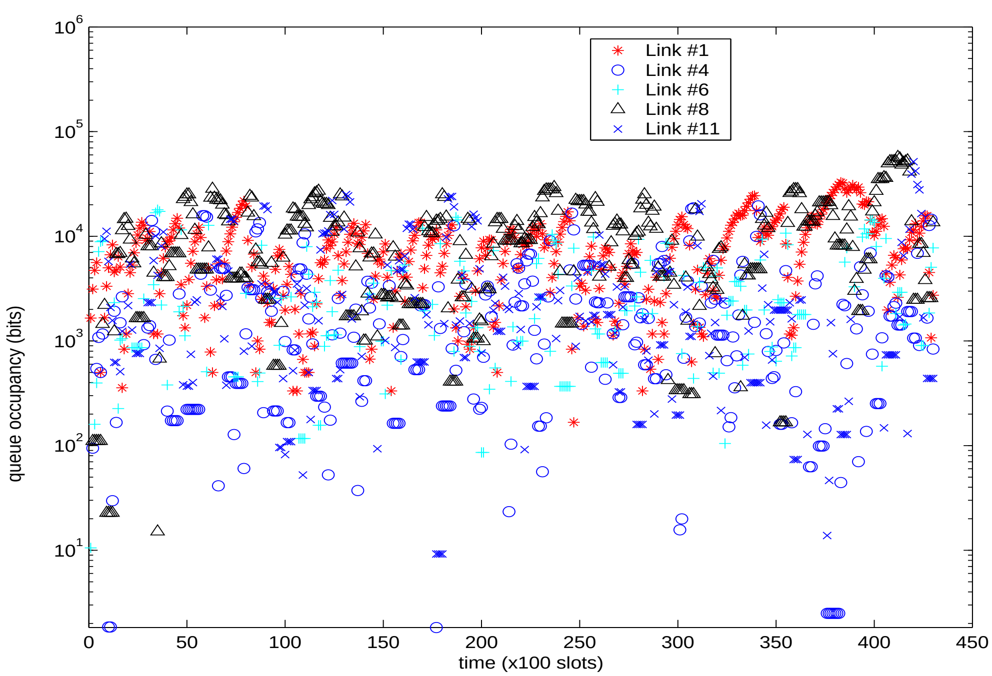

Figure 3.9: Queue occupancy of all links of r_{II} using algorithms MQR in a TD/CDMA wireless ad-hoc network with crossover traffic.

Algorithms	$\bar{R}^{eff}_{r_{II}}$	$\bar{R}^{eff}_{r_{III}}$	$\bar{R}^{eff}_{r_{IV}}$	$\bar{\mathbf{R}}$	support $\bar{R}^{min}_{II}$, $\bar{R}^{min}_{III}$, $\bar{R}^{min}_{IV}$?
PF	69.4	140.1	70.3	187.1	No
MPF (G)	101.8	191.8	101.1	271.6	Yes
PFMR	66.1	179.1	78.5	102.5	No
MPFMR (G)	108.9	226.2	122.3	188.3	Yes
QR	277.2	58.1	30.7	233.3	No
MQR (G)	371.3	66.4	44.9	256.9	No
MQRMR (G)	106.9	220.8	117.1	170.2	Yes
MMTMR (G)	150.2	281.7	135.5	303.9	Yes

Table 3.3: Effective rates of route II, III, and IV and the total average rate (all in kbps) in the network with unbalanced traffic. (G): Greedy algorithm.

fairness among the nodes, and thus may have serious unhealthy behavior when some malicious nodes take advantage of that and send large amount of data into the network.

A simple example is created to demonstrate this damaging effect. Instead of balanced traffic loads along the three routes (r_{II}, r_{III}, and r_{IV}), node A injected a lot of traffic into the network, to be exact, an order of magnitude higher than the other traffic sessions. The results are listed in Table 3.3. It is obvious that because no fairness has been considered by the throughput-optimal family of algorithms, they perform poorly with the effective rate of r_{III} and r_{IV} far below the required minimum rate. On the other hand, the PF-family of algorithms still provides the required minimum rate for all the traffic sessions and suppress the disturbance caused by the malicious node. All the multi-link PF-family of algorithms are able to support all the minimum rate requirements. However, in the throughput-optimal family of algorithms, only MQRMR is able to support all the minimum rate requirements because of the token counter mechanism. This result also indicates that the token counter mechanism indeed can help maintain the fair share of the traffic sessions specified by their minimum rate requirements.

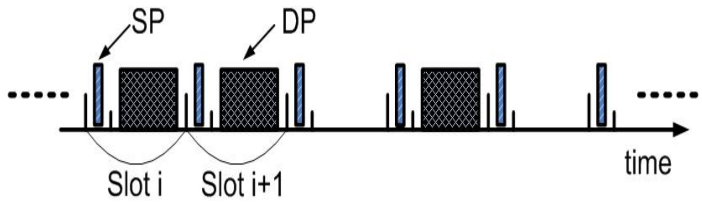

Figure 3.10: Slot format in centralized implementation. SP: scheduling packet; DP: data packet.

3.9 Centralized vs. Distributed Implementation

3.9.1 Centralized Implementation

The centralized solution needs a central controller and *global* information of all link gains. It may be implemented, for example, in a clustered wireless ad hoc network with "strong" cluster heads where centralized control is not far-fetched. In order to obtain the link gain information, each receiving node measures the received SINR.

At the beginning of each time slot, the central controller will broadcast a scheduling packet (SP) that contains the schedule for all nodes within the cluster. Each node will send an acknowledgement (ACK) that includes the measured channel gain and queue backlogs (if the throughput-optimal family of algorithms are chosen). The central controller will decode all replies and run a channel prediction algorithm to predict all channel gains for the next time slot. Then it will use the predicted channel gains and queue backlogs information to calculate the schedule for the next time slot. Note that a seperate control channel may be used for the information exchange between the central controller and each node. Alternatively, it may occupy a small percentage of each slot, as illustrated in Figure 3.10.

3.9.2 Distributed Implementation

In wireless ad hoc networks, where centralized control is not available, it may be very difficult to obtain the knowledge of all link gains, and thus it is impractical to implement a centralized solution. A distributed implementation is proposed where only local information is used to perform power control and scheduling decisions at each transmitting node individually [76]. The procedures are as follows:

1. At the beginning of each time slot, each node i in the potential transmitter set $\mathcal{S}$ select to transmit or not by flipping a coin. (This is motivated by the work of [99] and [75].)

2. Each node that decides to transmit will send a probe packet using power equal to the maximum transmission power p^{max}.

3. Each receiver detects the probe packets from all transmitting nodes nearby, and estimate the corresponding channel gain. The receiver then sends a packet including information of all the estimated link gains using power equal to the maximum transmission power p^{max}.

4. Each node i in the potential transmitter set S detects the packets from the receivers within its transmission range. From each of these receivers, node i obtains the list of all possible interfering transmitters and their link gains toward the receiver.

5. Each node i in the potential transmitter set $\mathcal{S}$ will transmit to one of the neighboring receivers where v_i (for example, $v_i = R_i/\bar{R}_i$ for MPF) is maximized.

6. Update the token counter according to equation (3.19) for algorithms using the token counter mechanism.

Note that at the start of each time slot, neighboring nodes will exchange information using a control/signaling channel. In addition, each node needs to keep a table of all the token queue lengths (for MPFMR algorithms) and the average rate for all outgoing active links.

Discrete-event simulations have been carried out to examine the performance of the proposed distributed implementation. In this simulation study, only local information is available to each node by exchanging control messages with its neighbors as described above. The overhead of the information exchange includes a one-byte (8 bits) probe packet and the reply from the receiver (which may contain multiple bytes). The exact

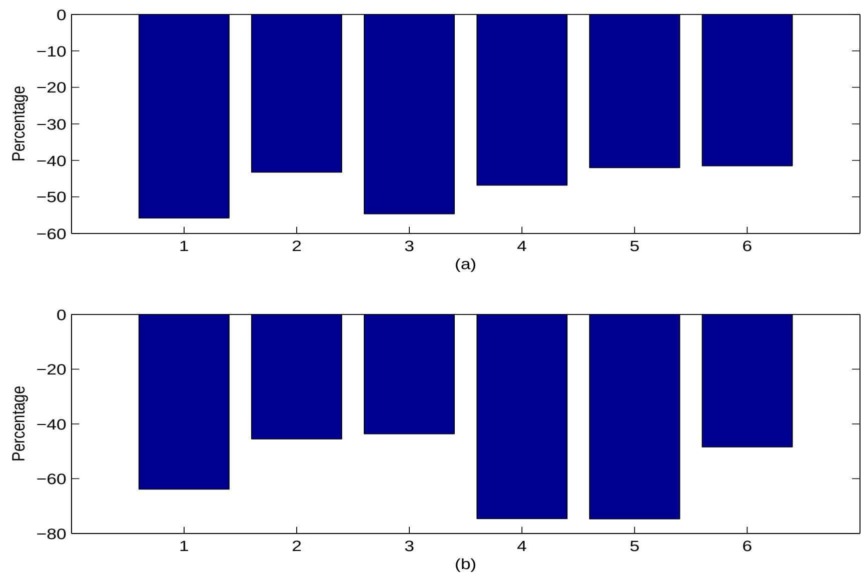

Figure 3.11: Gain/Loss of distributed algorithms over their centralized counterparts: (a). Total average rate; (b). Effective rate; 1. MPF (r_{II}), 2. MQR (r_{II}), 3. MPF (r_{III}), 4. MQR (r_{III}), 5. MPF (r_{IV}), 6. MQR (r_{IV}).

size of the reply depends on the number of probes that the receiver gets. Each link gain in the reply is counted as one byte assuming that the link gain is quantized using a 256-level quantizer. The other parameters of the simulation are the same as in Section 3.8. MPF and MQR algorithms are selected for comparison in the network with the *balanced* crossover traffic.

The percentage of rate gain/loss of distributed algorithms over their centralized counterparts is shown in Figure 3.11. The total average rate achieved by the distributed algorithms is about 40% less than their centralized counterparts because of lack of centralized control and global information. Because there is no global information about the queue backlog or the average rate, neither throughput-optimal nor fairness can be guaranteed in the distributed algorithm. The greedy nature of local decisions also results in bigger reductions (about 50%) in all effective rates, as expected.

The overhead in all cases is roughly the same 21%. This simple experiment demonstrates that the proposed distributed implementation achieves an acceptable performance (in terms of the total average rate and the effective rate comparing to the corresponding centralized algorithms) while keeps the overhead low.

3.10 Conclusions

In this chapter, the joint power control and scheduling problem for TD/CDMA wireless ad hoc networks is formulated using a utility function approach. Since the resulted optimal power control has the bang-bang characteristics, i.e., scheduled nodes transmit with full power while other nodes remain silent, the joint power control and scheduling problem is reduced to a scheduling problem. The Multi-link Gradient algorithm with the Minimum and Maximum Rate constraints (MGMR) is proposed to solve the corresponding optimization problem **(P.3)**. The main contributions of this chapter are summarized as follows:

1. The Multi-link Proportional Fair family of algorithms (MPFMR and MPF) is proposed to balance the tradeoff between efficiency and fairness.

2. While the proportional fair scheduling algorithms are good for the best effort traffic, the throughput optimal scheduling algorithms (MQRMR and MQR) are proposed to support the real-time traffic.

3. A generic token counter mechanism is proposed to satisfy the minimum and maximum rate requirements.

4. Service differentiation is achieved by ensuring a different minimum rate for different traffic sessions.

5. Greedy algorithms are proposed to achieve close to the optimal performance with much reduced computational complexity.

6. Distributed scheduling algorithms are also derived and a fully distributed implementation is provided.

7. Discrete event simulation of the scheduling algorithms using OPNET is designed and developed for validation of results. Two types of topologies, linear topology and network with crossover traffic are considered in the simulation study. The results demonstrate the effectiveness of the proposed schemes.

Chapter 4

Power Control for Cognitive Radio Ad Hoc Networks

In this chapter, the benefit of using cognitive radio in wireless ad hoc networks is explored and the associated power control problem is formulated and solved. While FCC proposes the spectrum sharing between a legacy TV system and a cognitive radio network to increase spectrum utillization, one of the major concerns is that the interference from the cognitive radio network should not violate the QoS requirements of the primary users. In this work, we consider the scenario where the cognitive radio network is formed by secondary users with low power personal/portable devices and when both systems are operating simultaneously. A power control problem is formulated for the cognitive radio network to maximize the energy efficiency of the secondary users and guarantee the QoS of both the primary users and the secondary users. The feasibility condition of the problem is derived and both centralized and distributed solutions are provided. Because the co-channel interference are from heterogeneous systems, a joint power control and admission control procedure is suggested such that the priority of the primary users is always ensured. The simulation results demonstrate the effectiveness of the proposed schemes.

4.1 Introduction

Although the U.S. government frequency allocation data [24] shows that there is fierce competition for the use of spectra, especially in the bands from 0 to 3 GHz, it is pointed out in several recent measurement reports that the assigned spectrum are highly underutilized [26, 100]. The discrepancy between spectrum allocation and spectrum use suggests that "spectrum access is a more significant problem than physical scarcity of spectrum, in large part due to legacy command-and-control regulation that limits the ability of potential spectrum users to obtain such access" [26]. In order to achieve much better spectrum utilization and viable frequency planning, Cognitive Radios (CR) are under development to dynamically capture the unoccupied spectrum [28, 101]. The Federal Communication Commission (FCC) has recognized the promising technique and is pushing to enable it to a full realization. As the first step, the FCC proposes to experiment unlicensed cognitive sharing in the TV bands (VHF and UHF bands) [27, 102, 103]. The TV bands are chosen due to the better penetration of the frequency band, "strong" received signal of the primary TV users, and TV transmitters are left on more or less continuously, and infrequently changed location or frequency [104].

Despite the advantages of using the TV bands for unlicensed cognitive spectrum sharing, there are some concerns to be solved first in order to convince FCC to finally open the TV bands. First, can secondary users (cognitive radio network) even operate without causing an excessive interference to primary users (TV users)? Second, can certain Quality-of-Service (QoS) for secondary users be provided under such constraints? So far, most of the previous works address these two issues by time sharing the spectrum between the TV system and the cognitive radio network. In such a case, there will be no co-channel interference. One of the main difficulties is to detect the presence of the TV signals accurately. Much work has been done in this area, such as [105, 106] and the references therein. In such a work, we consider a different case where the TV system and the cognitive radio network are ON simultaneously and they share the same spectrum through space separation. This case is mainly studied through MAC design, such as in [107]. Power control is only applied to address the non-intrusion to the services of the primary users [108], but not the QoS of the secondary users. We argue that the QoS of the secondary users is also very important [80]. If the capacity for the secondary users is not enough to realize their required QoS after meet the QoS constraints of the primary users, that channel might not be a good opportunity for secondary users to access.

According to the recent suggestions from the FCC [102, 103], two distinct types of unlicensed broadband devices may be used in the TV bands. One category will consist of lower power "personal/portable" unlicensed devices. The second category will consist of higher power "fixed/access" unlicensed devices that may provide wireless Internet access. This work will consider the power control problem for the first category, and we focus on the case where both the TV system and the cognitive radio network operate simultaneously. The power control problem becomes tougher than that in cellular systems or pure wireless ad hoc networks because the interference tend to be more difficult to model and control in two heterogeneous systems. In this work, we try to provide some preliminary analysis and design to address the two issues mentioned in the previous paragraph when two heterogeneous systems operate in the same channel at the same time. Specifically, a power control problem of the secondary users is formulated to maximize the energy efficiency of the secondary users and reduce the harmful interference to the primary users who have absolute priority. QoS guarantee of the secondary users is also included in the problem formulation. Feasibility conditions for the power control problem are highlighted and the corresponding joint power control and admission control procedures are provided [109, 110].

The work is organized as follows. Section 4.2 provides the model of spectrum sharing of a cognitive radio network with a TV broadcast system, and the associated power control problem is formulated. The solution of the power control problem for a single secondary transmitter is given in Section 4.3. Both centralized and distributed power control algorithms are provided for the case of multiple secondary users in Section 4.4. The effectiveness of the proposed schemes is tested through simulations in Section 4.5. Section 4.6 contains the concluding remarks.

4.2 Model and Problem Formulation

Given an existing TV station with transmission power p_{TV}, the effective receiving range is D. The effective receiving range is defined by the successful decoding of the TV signals, i.e., the received signal-to-interference-plus-noise ratio (SINR) should be above a given threshold (10 dB or higher [104], that will depend on the type of TV station) such that the received TV signal is decodable. Note that data of transmission power and effective receiving range of TV stations are publicly available, such as in [111, 112]. It is assumed that the secondary users are located in an $l \times l$ square area. The center

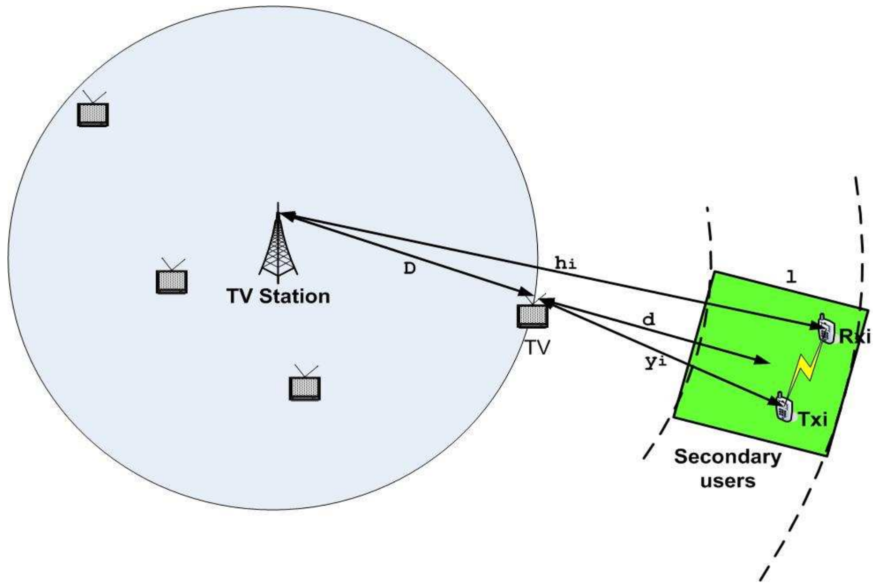

Figure 4.1: An example of spectrum sharing of a cognitive radio network with a TV broadcast system.

of the cognitive radio network is d meters away from the nearest primary receiver. The distance from the TV station to the ith secondary receiver is h_i. y_i is the distance from the i^{th} secondary transmitter to the TV receiver at the border of the TV coverage area. An example of the model is given in Figure 4.1, where only one pair of secondary users are shown. Note that although the effective receiving range of the TV station may not overlap with the transmission range of the cognitive radio network, transmissions in both systems still cause non-negligible co-channel interferences to the other system's receivers. For instance, if both systems are ON simultaneously, the transmission from the secondary users will cause interference at the primary receivers and may cause the received TV signals degraded and become unacceptable. Hence, the co-channel interference is the major barrier for the successful co-existence of the two systems.

In this chapter, we address the interference problem by considering the QoS at both the primary receivers and the secondary receivers in terms of the received SINR. Suppose there are totally N pairs of secondary users, and $p_{i,sec}$ is the transmission power of the i^{th} transmitter. Define the SINR at the m^{th} primary receiver as $\gamma_{m,TV}$, and the SINR at the i^{th} secondary receiver as $\gamma_{i,sec}$, the power control problem for energy efficiency maximization and interference suppression is formulated as follows

(P.4)

$$\min \sum_{i=1}^{N} p_{i,sec} \tag{4.1}$$

subject to

$$\gamma_{m,TV} \geq \gamma_{TV}^{tar}, \ \ \forall m \tag{4.2}$$

$$\gamma_{i,sec} \geq \gamma_{i,sec}^{tar}, \ \ i = 1, \cdots, N \tag{4.3}$$

$$p_{sec}^{min} \leq p_{i,sec} \leq p_{sec}^{max}, \ \ i = 1, \cdots, N \tag{4.4}$$

where γ_{TV}^{tar} and $\gamma_{i,sec}^{tar}$ are the target SINR for the primary receivers and the secondary receivers, respectively. p_{sec}^{min} and p_{sec}^{max} are the minimum and maximum allowable transmission power of the secondary users. These are "hard" limits including many considerations such as safety and hardware limitations, set by the standard organization or government agencies [103]. The objectives of power control in a cognitive radio network are to maximize the energy efficiency of the secondary users and suppress harmful interferences to both the primary users and the secondary users. This can be achieved by minimizing the total transmission power of the secondary users (equation (4.1)) while providing both the QoS of the primary users (equation (4.2)) and the QoS of the secondary users (equation (4.3)).

4.3 Power Control for a Single Secondary Transmitter

In this section, a simple case where there is only *one* secondary transmitter will be considered. We will first check the feasibility of the power control problem **(P.4)**. We assume that the received power is only a function of the transmitted power and path loss, i.e., fading effects (shadowing and small-scale fading) are omitted for now. We further assume that the path loss factor from the TV transmitter is α_1, and the path loss factor from the cognitive radio transmitter is α_2. Because the antenna height of the TV transmitter is usually several hundred meters higher [111] than that of the cognitive radio transmitters, it is expected that the path loss factor from the TV transmitter (α_1) will be better (smaller) than the path loss factor from the cognitive radio transmitter (α_2). The interference between the primary users and the secondary users depends on many factors such as modulation schemes and waveform design, and we assume that the orthogonality factors are f_1 and f_2, respectively.

Based on the above assumptions, the SINR of the TV receiver at the worst location of the TV coverage area is (please refer to Figure 4.1)

$$\gamma_{TV} = \frac{p_{TV}/D^{\alpha_1}}{f_2 p_{sec}/y^{\alpha_2} + \sigma^2} \tag{4.5}$$

and the SINR of the secondary receiver is

$$\gamma_{sec} = \frac{p_{sec}/r^{\alpha_2}}{f_1 p_{TV}/h^{\alpha_1} + \sigma^2} \tag{4.6}$$

where r is the distance between the secondary transmitter and the secondary receiver, σ^2 is the background noise.

In order to satisfy the two constraints on the primary and secondary SINR values, inequalities (4.2) and (4.3), we need

$$p_{sec} \leq [\frac{p_{TV}}{D^{\alpha_1}\gamma_{TV}^{tar}} - \sigma^2] y^{\alpha_2}/f_2 \ , \tag{4.7}$$

and

$$p_{sec} \geq (f_1 p_{TV}/h^{\alpha_1} + \sigma^2)\gamma_{sec}^{tar} r^{\alpha_2} \ . \tag{4.8}$$

If the power control problem is feasible, equations (4.7), (4.8), and (4.4) have to be satisfied simultaneously.

Theorem 4 *Given the transmission power of the primary transmitter (p_{TV}) and the background noise (σ^2), the target SINR values of the primary receiver and the secondary receiver (γ_{TV}^{tar} and γ_{sec}^{tar}), and the distances (D, y, h, r), the feasibility condition of the power control problem* **(P.4)** *for a single secondary transmitter is*

$$\max\{p_{sec}^{min}, \underline{p}_{sec}\} \leq p_{sec} \leq \min\{\bar{p}_{sec}, p_{sec}^{max}\} \tag{4.9}$$

where $\bar{p}_{sec} = [\frac{p_{TV}}{D^{\alpha_1}\gamma_{TV}^{tar}} - \sigma^2] y^{\alpha_2}/f_2$ *and* $\underline{p}_{sec} = (f_1 p_{TV}/h^{\alpha_1} + \sigma^2)\gamma_{sec}^{tar} r^{\alpha_2}$.

The feasibility condition given in Theorem 4 may be interpreted as follows:

Corollary 1 *Define two transmission power sets,* $S_1 = \{p_{sec}^{min} \leq p_{sec} \leq p_{sec}^{max}\}$, *and* $S_2 = \{\underline{p}_{sec} \leq p_{sec} \leq \bar{p}_{sec}\}$, *the power control problem* **(P.4)** *for a single secondary transmitter is feasible iff* $S_1 \cap S_2 \neq \emptyset$.

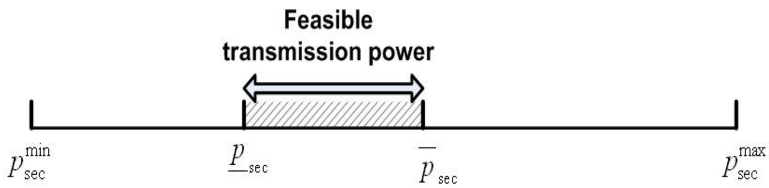

Figure 4.2: Feasible transmission power of the secondary user.

One possible case of feasible transmission power of the secondary user is shown in Figure 4.2. If the feasibility condition (inequality (4.9)) is satisfied, the optimal transmission power of the secondary user is $\max\{p_{sec}^{min}, \underline{p}_{sec}\}$. If the minimum allowable transmission power is 0, the optimal transmission power of the secondary user is $\underline{p}_{sec}$.

If the interference is dominant, i.e., if $f_2 p_{sec}/y^{\alpha_2} \gg \sigma^2$ and $f_1 p_{TV}/h^{\alpha_1} \gg \sigma^2$, which is usually the case, the sum of the SINR (in dB) of the TV receiver at the border of the TV coverage area and the SINR of the secondary receiver can be expressed as

$$\gamma_{TV}^{dB} + \gamma_{sec}^{dB} \approx \alpha_1 \frac{h}{D}(dB) + \alpha_2 \frac{y}{r}(dB) - [f_1 + f_2](dB) \ . \tag{4.10}$$

The achievable SINR of the secondary users can be estimated by substracting γ_{TV}^{tar} from the sum of the SINR.

It is observed that the sum of these two SINR values (in dB) is only a function of relative distances. One simulation result is plotted in Figure 4.3. The parameters used in simulation are given in Table 4.1 and it is assumed that $h \approx D + d$ and $y \approx d$ since $d \gg l$. It is observed that the distance between the secondary transmitter and the secondary receiver, r, has the dominant effect on the sum of the SINR values. For example, if r decreases from 300 meters ($\frac{r}{D} = 0.005$) to 60 meters ($\frac{r}{D} = 0.001$), the gain of the sum of the SINR values is about 30 dB. In addition, if r is large, say r is 480 meters ($\frac{r}{D} = 0.008$), even if the secondary user is far away from the TV coverage area (say, $\frac{d}{D} = 1$), the sum of the SINR values is still very low, about 30 dB. In other words, if the required primary SINR is 34 dB, the maximum achievable SINR for the secondary user is about -4 dB. The results suggest that only low power secondary users with short range transmissions (low power personal/portable devices [103]) are allowed when the primary users are ON. This also calls for multi-hop communications rather than single hop long range transmissions in the cognitive radio network.

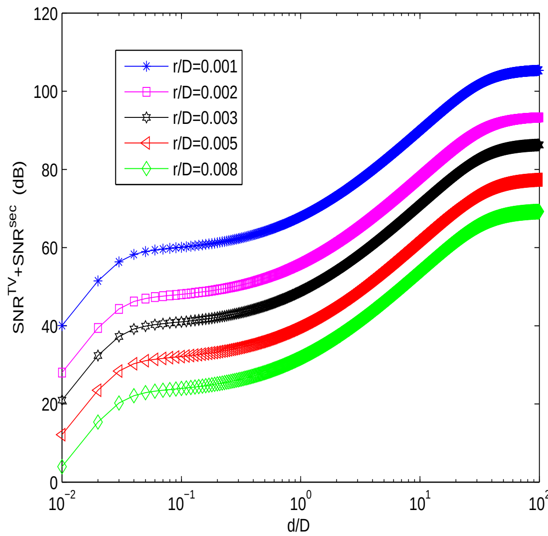

Figure 4.3: The sum SINR values (in dB) vs. $\frac{d}{D}$ and $\frac{r}{D}$, $f_1 = 1, f_2 = 1$.

We would like to point out that although the transmission powers are not explicitly included in the formula for the sum SINR, they indeed will determine the proportion of the SINR that the primary user and the secondary user will get.

4.4 Power Control for Multiple Secondary Users

In this section, we are going to provide both centralized and distributed solutions to the power control problem **(P.4)**. In order to evaluate the interference and solve the power control problem, we assume that the distances d and y_i can be estimated accurately. Indeed, geolocation devices (e.g. GPS), control signals, or spectrum sensing may be applied to detect the primary transmissions and get an accurate estimate of the distances [103].

4.4.1 Centralized Solution

The SINR of the TV receiver at the worst location of the TV coverage area is

$$\gamma_{TV} = \frac{p_{TV}/D^{\alpha_1}}{f_2 \sum p_{i,sec}/y_i^{\alpha_2} + \sigma^2} \tag{4.11}$$

The SINR of the i^{th} secondary receiver is

$$\gamma_{i,sec} = \frac{g_{ii} p_{i,sec}}{\sum_{j \neq i} g_{ij} p_{j,sec} + f_1 p_{TV} / h_i^{\alpha_1} + \sigma^2} \tag{4.12}$$

where g_{ij} is the link gain from the j^{th} secondary transmitter to the i^{th} secondary receiver.

The following theorem gives the feasibility condition of the power control problem **(P.4)**.

Theorem 5 *The power control problem* **(P.4)** *is feasible for all N simultaneous transmitting-receiving pairs of secondary users within the same channel as long as*

(1). The matrix $[I - \Gamma_{sec}^{tar} Z]$ is non-singular (thus invertible);

(2). The transmission power vector p_{sec}^* satisfies inequality (4.4) element-wise, where

$$p_{sec}^* = [I - \Gamma_{sec}^{tar} Z]^{-1} u \ , \tag{4.13}$$

matrix Γ^{tar} is a diagonal matrix

$$\Gamma_{sec\,ij}^{tar} = \begin{cases} \gamma_{i,sec}^{tar} & i = j \\ 0 & otherwise \end{cases} , \tag{4.14}$$

matrix Z is the following nonnegative matrix

$$Z_{ij} = \begin{cases} \frac{g_{ij}}{g_{ii}} & i \neq j \\ 0 & i = j \end{cases} , \tag{4.15}$$

u is the vector with elements

$$u_i = \gamma_{i,sec}^{tar} \eta_i^2 / g_{ii}, \qquad i = 1, 2, ..., N \tag{4.16}$$

and

$$\eta_i^2 = f_1 p_{TV} / h_i^{\alpha_1} + \sigma^2 \ . \tag{4.17}$$

(3). The transmission power vector p_{sec}^* also satisfies the following inequality

$$\frac{p_{TV} / D^{\alpha_1}}{f_2 \sum p_{i,sec}^* / y_i^{\alpha_2} + \sigma^2} \geq \gamma_{TV}^{tar} \ . \tag{4.18}$$

Proof: A target SINR vector $\gamma^{\mathbf{tar}}$ is achievable for all simultaneous transmitting-receiving pairs of secondary users within the same channel if the following conditions are met [45, 46]

$$\gamma_{i,sec} \geq \gamma_{i,sec}^{tar} \tag{4.19}$$

$$p \geq 0 \tag{4.20}$$

where p is the vector of transmitting powers. Define η_i^2 as in equation (4.17). Replacing $\gamma_{i,sec}$ with equation (4.12) and rewriting the above conditions in matrix form gives

$$[I - \Gamma^{tar} Z]p \geq u \tag{4.21}$$

$$p \geq 0 \tag{4.22}$$

where matrix Γ^{tar}, matrix Z and vector u are defined in equations (4.14), (4.15), and (4.16), respectively.

It is shown in [46] that if the system is feasible, the matrix $[I - \Gamma^{tar} Z]$ must be invertible and the inverse should be element-wise positive, thus prove part (1) of the theorem.

It is also shown in [46] (Proposition 2.1) that if the system is feasible, there exists a unique (Pareto optimal) solution which minimizes the transmitted power [47, 48]. This solution is obtained by solving a system of linear algebraic equations

$$[I - \Gamma^{tar} Z]p^* = u \tag{4.23}$$

In order to satisfy constraints (4.2) and (4.4) in the power control problem **(P.4)**, the transmission power vector p_{sec}^* must satisfy inequality (4.4) element-wise and inequality (4.18), thus prove the theorem. ■

The above proof highlighted the centralized solution to the problem **(P.4)**. Although it seems that the power control problem **(P.4)** is similar to that in cellular systems [51] and in wireless ad hoc networks [6], the power control problem considered here addressed the interference from *heterogeneous* systems and additional constraint (4.2) has to be satisfied and the interference between primary and secondary users has to be taken into account in the problem formulation. It also calls for the joint design of power control and admission control for the cognitive radio network such that QoS of the primary users is ensured all the time. The procedures of joint power control and admission control is summarized below.

Joint power control and admission control

1. Solve the transmission power vector p^*_{sec} using equation (4.13).

2. Check whether the transmission powers are within limit, i.e., $p^{min}_{sec} \leq p^*_{i,sec} \leq p^{max}_{sec}$, $\forall i$? If yes, goes to the next step; otherwise, the power control problem **(P.4)** is not feasible. Remove the jth secondary user that has the largest $\sum_{i=1}^{N}[Z_{ij} + Z_{ji}]$ and return to Step 1 with the reduced number of transmitters.

3. Check whether the transmission powers satisfy inequality (4.18). If yes, set the transmission power vector as p^*_{sec}; otherwise, the power control problem **(P.4)** is not feasible. Remove the secondary user that requires the largest transmission power ($p = max\{p^*_{i,sec}\} \; \forall i$) and return to Step 1 with reduced number of transmitters.

It worth pointing out that Steps 2 and 3 implement admission control for the secondary users. When the power control problem **(P.4)** is not feasible, the secondary user that caused the worst interference should be silenced. The central controller can verify the transmission power limits in a straight forward way in Step 2 after solving p^*_{sec} using equation (4.13). The worst interferer to other secondary users inside the cognitive radio network is the one that has the largest row and column sum of matrix Z. In Step 3, given that p_{TV}, γ^{tar}_{TV}, and D are publicly available data, and y_i can be estimated accurately, the central controller can verify inequality (4.18). This time the worst interferer to the primary receivers is the one that has the largest transmission power since all the secondary transmitters have more or less the same distance to the primary receivers. In a cognitive radio network with centralized management, such as in a cluster based architecture, the above procedures may be implemented.

4.4.2 Distributed Solution

The centralized solution (equation (4.13)) needs a central controller and *global* information of all link gains, and centralized power control requires extensive control signaling in the network and it is difficult to implement in practice, especially for an infrastructure-less wireless ad hoc network. Therefore, a distributed implementation which only uses local information to make a control decision is proposed for realistic scenarios.

Distributed power control schemes may be derived by applying iterative algorithms to solve equation (4.23). For example, using the first-order Jacobian iterations [49], the following distributed power control scheme is obtained

$$p_{i,sec}(k+1) = \min\{\frac{\gamma_{i,sec}^{tar}}{\gamma_{i,sec}(k)} p_i(k),\ \ p_{sec}^{max}\}\ , i = 1, 2, ..., N. \tag{4.24}$$

Note that each node only needs to know its own received SINR at its designated receiver to update its transmission power. This is available by feedback from the receiving node through a control channel. As a result, the algorithm is fully distributed. Convergence properties of this type of algorithms were studied by Yates [50, 51]. An interference function $I(p)$ is standard if it satisfies three conditions: positivity, monotonicity and scalability. It is proved by Yates [50] that the standard iterative algorithm $p(k+1) = I(p(k))$ will converge to a unique equilibrium that corresponds to the minimum use of power. The distributed power control scheme (equation (4.24)) is a special case of the standard iterative algorithm.

Since the Jacobi iteration is a fixed-point iterative method, it usually has slow convergence speed to the sought solution. However, we select equation (4.24) as the power control algorithm in cognitive radio networks due to its simplicity. Other advanced algorithms with faster convergence speed can be found in [45, 53].

The distributed power control algorithm given in equation (4.24) does not enforce the QoS requirement of the primary users represented by inequality (4.18). Thus, the secondary users who apply equation (4.24) alone may violate the QoS requirement of the primary users. In order to address this issue, we propose two possible solutions. The first solution is a direct solution, where a "genie" is placed near the primary receiver at the border of the TV coverage area. The genie will monitor the interference level and inform the secondary users (such as using a beacon signal) if the interference level is too high and the QoS requirement of the primary users will be violated. One possible implementation of the genie is a secondary user that happens to locate inside the TV coverage area. The second solution is an indirect solution. Assume that $y_i \approx y_j = d, \forall i \neq j$[1], then the inequality (4.18) may be written as

$$\sum_i p_{i,sec} \leq [p_{TV}/(D^{\alpha_1}\gamma_{TV}^{tar}) - \sigma^2]\frac{d^{\alpha_2}}{f_2} \ . \tag{4.25}$$

Suppose that all secondary users that plan to transmit will report to a manager their respective transmission powers, $p_{i,sec}$ for user i, the manager will be able to verify the QoS requirement of the primary users by checking inequality (4.25).

4.5 Simulation Results

In this section, the performance of the proposed power control algorithm is examined. It is assumed that a group of $N = 50$ transmitting-receiving pairs of secondary users using low power devices are communicating with each other in a 2000 meter $\times$2000 meter area. They share the same spectrum with a TV system, and the TV station is located $D + d$ meters away. The locations of the transmitting-receiving pairs are chosen such that $r_{ij} > 3r_{ii}$ to ensure feasibility of the power control problem, where r_{ij} is the distance from the jth transmitter to the ith receiver and $g_{ij} = 1/r_{ij}^{\alpha_2}$. The initial transmission power of the secondary users are randomly chosen between p_{sec}^{min} and p_{sec}^{max}. The rest of the simulation parameters are listed in Table 4.1.

The average achievable SINR value of the secondary users (γ_{sec}^{avg}) vs. d/D is shown in Figure 4.4. It is observed that γ_{sec}^{avg} increases monotonically with d as expected. It is also shown that the gain in γ_{sec}^{avg} decreases when d increases, because the interference between the two systems play less a role in the achievable SINR value when they are further away. When $d/D > 2$, γ_{sec}^{avg} is pretty much limited by the interference of its own system.

---

[1]This assumption is expected to be true most of the time, since typically the secondary users must reside far away enough from the TV coverage area.

Parameters	Value
p_{TV}	100 kW
γ_{TV}^{tar}	34 dB
p_{sec}^{min}	0 mW
p_{sec}^{max}	100 mW
γ_{sec}^{tar}	3 dB
σ^2	10^{-14}
D	60 km
α_1	3
α_2	4

Table 4.1: Simulation parameters

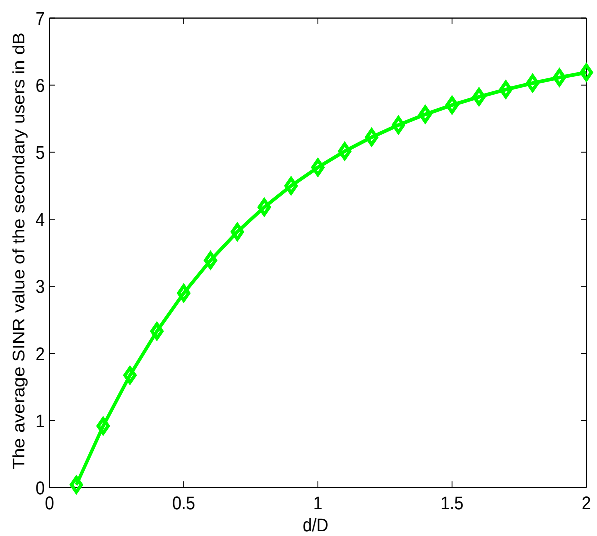

Figure 4.4: The average achievable SINR value of the secondary users (γ_{sec}^{avg}) vs. d/D.

In the following part of the simulation, $d = 0.5D$, and the distributed power control algorithm, equation (4.24), is applied. The convergence of the mean square error of the secondary user's SINR ($e_{sec}^2 = E[(\gamma_{sec} - \gamma_{sec}^{tar})^2]$) is given in Figure 4.5. It is observed that the power control algorithm converges very fast (in about 10 steps). Similarly, the convergence of the transmission power of some randomly chosen secondary users is shown in Figure 4.6.

The minimum SINR value of the primary users during the power control process of the secondary users is shown in Figure 4.7. It is confirmed that the QoS of the primary users is not violated during the power control process.

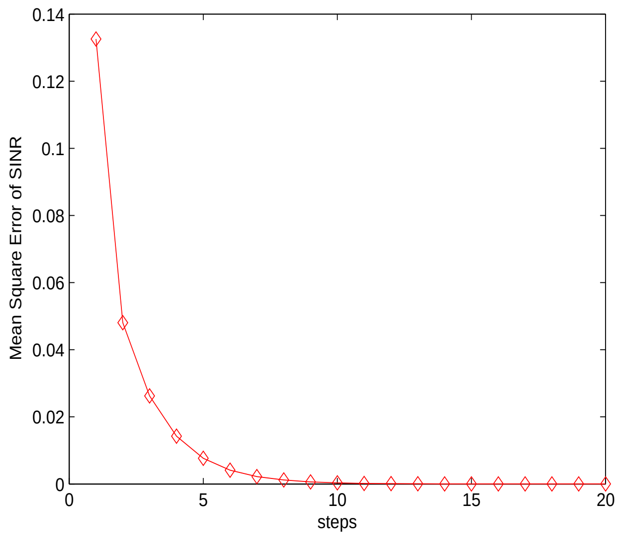

Figure 4.5: The convergence of the mean square error of the secondary user's SINR.

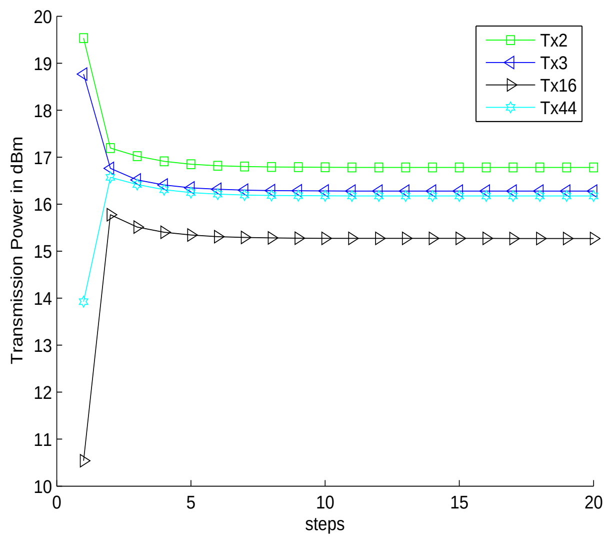

Figure 4.6: The convergence of the transmission power of the secondary users.

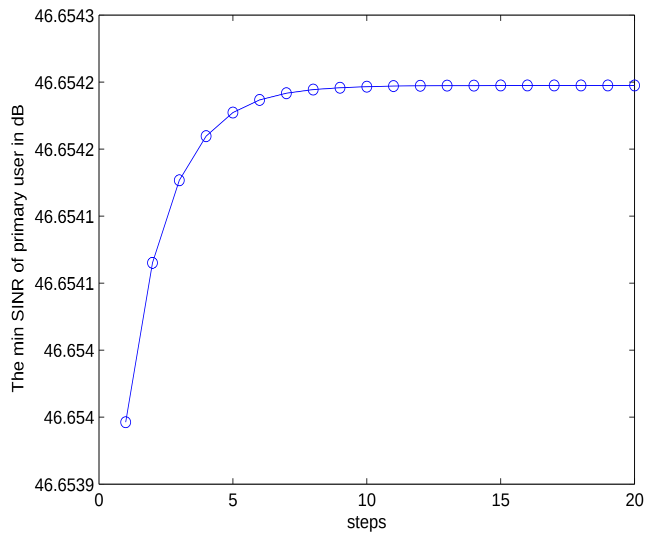

Figure 4.7: The minimum SINR value of the primary users during the power control process of the secondary users.

4.6 Conclusions

In this chapter, a power control problem is formulated for a cognitive radio network that operates simultaneously in the same frequency band with a TV system. Both centralized and distributed solutions are given to maximize the energy efficiency of the cognitive radio network and provide QoS support for both primary and secondary users. In addition, the feasibility condition is derived and the joint power control and admission control procedure is suggested such that priority of the primary users is ensured all the time. Furthermore, the proposed power control and admission control procedure may be combined with the MAC design to enhance the promise of non-intrusion to the primary system during spectrum sharing.

It worth pointing out that the results obtained in this chapter can be extended to the CDMA cognitive radio network in a straight forward manner. In the case of TDMA as the MAC scheme and only one secondary user is allowed to transmit during one time slot, the results of the single secondary transmitter case in Section 4.3 give optimal power control for one TDMA cognitive radio network. The results in Section 4.4 correspond to the power control of co-channel secondary users in multiple TDMA cognitive radio networks.

Although the TV broadcast system is chosen as an example of the primary system in this work, the proposed methods can be extended to other cases where heterogeneous systems share the same spectrum. In the current work, only one cognitive radio network is considered. The power control for multiple cognitive radio networks is one of our future research efforts.

Chapter 5

Conclusions and Future Work

In this work, the interactions among physical layer, link layer and network layer, are studied to improve the performance and efficiency of wireless ad hoc networks. Specifically, given the sessions with their source-destination pairs and QoS requirements, power control, scheduling and routing are jointly designed to improve the performance of end-to-end communications in terms of energy efficiency, bandwidth efficiency, and QoS support.

A joint power control and maximally disjoint multipath routing algorithm augmented by a traffic monitoring and switching mechanism is proposed for routing traffic between a source and destination pair with high energy efficiency and robustness. Based on realistic interference model, a framework of power control with QoS constraints in CDMA wireless ad hoc networks is introduced and both the centralized and distributed solutions are derived. Then, an iterative joint power control and maximally disjoint routing scheme is proposed for routing data traffic with the minimum rate constraint while maintaining high energy efficiency and prolonged network lifetime. Specifically, the Minimum Power Split Multi-path Routing (MPSMR) and Balanced Energy Split Multipath Routing (BESMR) algorithms are proposed. Simulations demonstrate that the proposed schemes achieve a significant gain in energy efficiency and network lifetime over SMR. Furthermore, in order to provide reliable end-to-end data delivery, the proposed scheme is augmented by a traffic monitoring and switching mechanism to mitigate the effect of node mobility or node failure. The proposed joint power control and routing is based on SMR, a multipath source routing algorithm. Extensions of the proposed scheme to other popular ad hoc routing algorithms such as Ad hoc On-demand Multi-path Distance Vector (AOMDV) [37] is of interest for future research.

The cluster based architecture is introduced in wireless ad hoc networks to provide centralized control within clusters, and the corresponding power control and scheduling schemes are derived to maximize the network utility function and guarantee the minimum and maximum rates required by each traffic session, given routes for multiple end-to-end multihop traffic sessions. In order to achieve a high end-to-end throughput in a multihop wireless ad hoc network, TD/CDMA has been chosen as the medium access control scheme due to its support for the high network throughput in a multihop environment. The associated power control and scheduling problem is addressed to optimize operations of TD/CDMA. Because the resulted optimal power control reveals bang-bang characteristics, i.e., scheduled nodes transmit with full power while other nodes remain silent, the joint power control and scheduling problem is reduced to a scheduling problem. The Multi-link Gradient algorithm with Minimum and Maximum Rate constraints (MGMR) is proposed to solve the optimization problem. Based on two popular scheduling criteria, the Multi-link Proportional Fair scheduling algorithm with Minimum and Maximum Rate constraints (MPFMR) and the Multi-link Throughput Optimal with Minimum and Maximum Rate constraints (MQRMR) algorithmis are proposed. A generic token counter mechanism is employed to satisfy the minimum and maximum rates requirements. Note that by ensuring different minimum rates for different traffic sessions, service differentiation can also be achieved. Approximative algorithms are suggested to reduce the computational complexity. In networks that lack centralized control, distributed scheduling algorithms are also derived and a fully distributed implementation is provided. Simulation results demonstrate the effectiveness of the proposed schemes. However, the overhead of the distributed scheduling algorithms is still significant, how to design more efficient distributed scheduling algorithms will be an interesting future research topic.

In order to further improve the bandwidth efficiency, cognitive radio technology is considered for more efficient spatial and temporal spectrum sharing. In this work, the power control problem is formulated for a cognitive radio ad hoc network that operates simultaneously in the same frequency band with a legacy system. Both centralized and distributed solutions are provided to maximize the energy efficiency of the cognitive radio network and provide QoS support for both legacy and ad hoc network users. In addition, the feasibility condition is derived and a joint power control and admission control procedure is suggested such that priority of legacy users is ensured at all times. It worth pointing out that although the TV broadcast system is chosen as an example of the legacy system in this dissertation, the proposed methods can be extended to other cases where heterogeneous systems share the same spectrum. In the current work, only one cognitive radio ad hoc network is considered. The power control for multiple cognitive radio networks is one of our future research efforts. Another interesting future topic would be modeling the behavior of primary and secondary networks by using Stackelberg game [113], where the primary network is the Stackelberg leader, and the secondary network is the Stackelberg follower.

References

[1] C. S. R. Murthy and B. Manoj, *Ad Hoc Wireless Networks–Architectures and Protocols*. Prentice Hall, 2004.

[2] IETF, "Mobile ad-hoc networks working group," in *[Online]. Available: http://www.ietf.org/html: Retrieved Jan. 2007*.

[3] "Ieee 802.11 wirless local area networks working group," in *[Online]. Available: http://www.ieee802.org: Retrieved Jan. 2007*.

[4] 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 European Wireless 2002. Next Generation Wireless Networks: Technologies, Protocols, Services and Application.*, Feb. 2002, pp. 156–162.

[5] T. ElBatt, S. Krishnamurthy, D. Connors, and S. Dao, "Power Management for Throughput Enhancement in Wireless Ad-Hoc Networks ," in *Proceedings of IEEE ICC*, 2000, pp. 1506–1513.

[6] T. ElBatt and A. Ephremides, "Joint Scheduling and Power Control for Wireless Ad-hoc Networks," in *Proceedings of IEEE INFOCOM*, 2002, pp. 976–984.

[7] J. Monks, J. Ebert, W. Hwu, and A. Wolisz, "Energy Saving and Capacity Improvement Potential of Power Control in Multi-hop Wireless Networks," *Computer Networks*, vol. 41, pp. 313–330, 2003.

[8] J. Gomez and A. Campbell, "A Case for Variable-Range Transmission Power Control in Wireless Multihop Networks," in *Proceedings of IEEE INFOCOM*, vol. 2, 2004, pp. 1425–1436.

[9] D. Bertsekas and R. Gallager, *Data Networks (Second Edition)*. Prentice Hall, 1991.

[10] F. Talucci and M. Gerla, "MACA-BI (MACA By Invitation), a Wireless MAC Protocol for High Speed Ad Hoc Networking," in *IEEE ICUPC'97*, vol. 2, Oct. 1997, pp. 913–917.

[11] P. Karn, "MACA - a New Channel Access Method for Packet Radio," in *Proceedings of 9th ARRL Computer Networking Conference*, 1990.

[12] V. Bharghavan, A. J. Demers, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless LAN's," in *Proceedings of SIGCOMM*, 1994, pp. 212–225.

[13] C. E. Perkins and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for mobile computers," in *Proceedings of ACM SIGCOMM*, 1994, pp. 234–244.

[14] L. R. Ford and D. R. Fulkerson, *Flows in Networks.* Princeton University Press, 1962.

[15] S. Murthy and J. J. Garcia-Luna-Aceves, "An Efficient Routing Protocol for Wireless Networks," *ACM/Baltzer Mobile Networks and Applications*, vol. 1, no. 2, pp. 183–197, 1996.

[16] T. Clausen and P. Jacquit, "Optimized Link State Routing Protocol (OLSR)," in *IETF RFC 3626, http://www.ietf.org/rfc/rfc3626.txt*, Oct 2003.

[17] R. Ogier, M. Lewis, and F. Templin, "Topology Dissemination Based on Reverse-Path Forwarding (TBRPF)," in *IETF Internet Draft, http://www.ietf.org/internet-drafts/draft-ietf-manet-tbrpf-09.txt*, June 2003.

[18] C. E. Perkins, E. Belding-Royer, and S. R. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing," in *RFC 3561, http://www.ietf.org/rfc/rfc3561.txt*, July 2003.

[19] D. Johnson, Y. Hu, and D. Maltz, "The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks," in *IETF RFC 4728, http://www1.ietf.org/mail-archive/web/ietf-announce/current/msg03407.html*, Feb. 2007.

[20] V. D. Park and M. S. Corson, "A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Network," in *Proceedings of IEEE INFOCOM*, 1997, pp. 1405–1413.

[21] C. K. Toh, "A Novel Distributed Routing Protocol to Support Ad-Hoc Mobile Computing," in *Proceedings of IEEE 15th Annual Int'l. Phoenix Conf. Comp. and Commun.*, Mar. 1996, pp. 480–486.

[22] Z. Hass and M. Pearlman, "The performance of Query Control Schemes for the Zone Routing Protocol," *IEEE/ACM Transactions on Networking*, vol. 9, no. 4, pp. 427–438, 2001.

[23] V. Kawadia and P. Kumar, "Principles and Protocols for Power Control in Wireless Ad Hoc Networks," *IEEE Journal on Selected areas in communications*, vol. 23, no. 1, pp. 76–88, Jan. 2005.

[24] "United States Frequency Allocations," in *[Online] Available: http://www.ntia.doc.gov/osmhome/allochrt.pdf.*

[25] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation Issues in Spectrum Sensing for Cognitive Radios," in *Proceedings of 38th Annual Asilomar Conference on Signals, Systems and Computers*, vol. 1, Nov. 2004, pp. 772–776.

[26] FCC, "Spectrum Policy Task Force," in *ET Docket no.02-135*, Nov. 2002.

[27] ——, "Facilitating Opportunities for Flexible, Efficient, and Reliable Spectrum Use Employing Cognitive Radio Technologies, Notice of Proposed Rule Making and Order," in *ET Docket no.03-322, [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-03-322A1.pdf*, 2003.

[28] J. Mitola and G. Q. Maguire, "Cognitive Radio: Making Software Radios More Personal," *IEEE Personal Communication.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[29] B. Fette, *Cognitive Radio Technology.* Newnes, Aug. 2006.

[30] A. Muqattash and M. Krunz, "CDMA-Based MAC Protocol for Wireless Ad Hoc Networks," in *Proceedings of MobiHoc'03*, 2003, pp. 153–164.

[31] V. Kawadia and P. R. Kumar, "Power Control and Clustering in Ad Hoc Networks," in *Proceedings of IEEE INFOCOM*, 2003, pp. 459–469.

[32] S. Chen and K. Nahrstedt, "Distributed Quality-of-Service Routing in Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, pp. 1488–1505, Aug. 1999.

[33] L. Chen and W. Heinzelman, "QoS-Aware Routing Based on Bandwidth Estimation for Mobile Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 3, pp. 561–572, Mar. 2005.

[34] A. Alwan, R. Bagrodia, N. Bamos, M. Gerla, L. Kleinrock, J. Short, and J. Villasenor, "Adaptive Mobile Multimedia Networks," *IEEE Personal Communications*, vol. 3, no. 2, pp. 34–51, Apr. 1996.

[35] R. Chakravorty, S. Katti, I. Pratt, and J.Crowcroft, "Flow Aggregation for Enhanced TCP Over Wide Area Wireless," in *Proceedings of IEEE INFOCOM*, vol. 3, 2003, pp. 1754–1764.

[36] S. Lee and M. Gerla, "AODV-BR: Backup Routing in Ad Hoc Networks," in *Proceedings of IEEE WCNC*, vol. 3, Sep 2000, pp. 1311–1316.

[37] M. K. Marina and S. R. Das, "Ad Hoc On-demand Multipath Distance Vector Routing Protocol," in *Proceedings of IEEE International Conference for Network Protocols (ICNP)*, Nov. 2001.

[38] A. Srinivas and E. Modiano, "Minimum Energy Disjoint Path Routing in Wireless Ad-Hoc Networks," in *Proceedings of MobiCom'03*, Sep. 2003, pp. 122–133.

[39] ——, "Finding Minimum Energy Disjoint Paths in Wireless Ad-Hoc Networks," *ACM Wireless Networks*, pp. 401–417, Nov. 2005.

[40] C. Comaniciu and H. Poor, "QoS Provisioning for Wireless Ad Hoc Data Networks (invited paper)," in *Proceedings of 42nd IEEE Conference on Decision and Control*, vol. 1, Dec. 2003, pp. 92–97.

[41] ——, "On Energy-Efficient Hierarchical Cross-Layer Design: Joint Power Control and Routing for Ad Hoc Networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, pp. 1–9, 2007.

[42] A. Michail and A. Ephremides, "Energy-Efficient Routing for Connection-Oriented Traffic in Wireless Ad-Hoc Networks," *Mobile Networks and Applications*, vol. 8, pp. 517–533, 2003.

[43] R. Cruz and A. Santhanam, "Optimal Routing, Link Scheduling and Power Control in Multi-hop Wireless Networks," in *Proceedings of IEEE INFOCOM*, 2003, pp. 702–711.

[44] R. Berry and E. Yeh, "Cross-Layer Wireless Resource Allocation," *IEEE Signal Processing Magazine*, pp. 59–68, Sep. 2004.

[45] G. Foschini and Z. Miljanic, "A Simple Distributed Autonomous Power Control Algorithm and Its Convergence," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, Nov. 1993.

[46] D. Mitra, "An Asynchronous Distributed Algorithm for Power Control in Cellular Radio Systems," in *Proceedings of 4th WINLAB Workshop of 3rd Generation Wireless Information Networks*, Oct. 1993, pp. 177–186.

[47] M. Xiao, N. Shroff, and E. Chong, "A Utility-Based Power-Control Scheme in Wireless Cellular Systems," *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 210–221, Apr. 2003.

[48] J. Herdtner and E. Chong, "Analysis of a Class of Distributed Asynchronous Power Control Algorithms for Cellular Wireless System," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 436–446, Mar. 2000.

[49] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.

[50] R. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1348, Sep. 1995.

[51] S. Grandhi, J. Zander, and R. Yates, "Constrained Power Control," *International Journal of Wireless Personal Communications*, vol. 1, no. 4, Apr. 1995.

[52] R. Jantti and S. Kim, "Second-Order Power Control with Asymptotically Fast Convergence," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 447–457, Mar. 2000.

[53] X. Li, "Uplink Power Control and Scheduling in CDMA Systems," in *M.S. Thesis, WINLAB, Rutgers University*, 2003.

[54] D. Lelic and Z. Gajic, "Gauss-Seidel Iterations for SIR-based Power Updates for 3G Wireless CDMA Communication Networks," *International Journal of Wireless Information Networks*, vol. 11, pp. 115–121, 2004.

[55] X. Li and Z. Gajic, "An improved SIR-based Power Control for CDMA Systems using Steffensen Iterations," in *Proceedings of IEEE CISS*, Mar. 2002, pp. 287–290.

[56] S. Koskie and Z. Gajic, “Signal-to-Interference-Based Power Control For Wireless Networks: A Survey, 1992-2005,” *Dynamics of Continuous, Discrete and Impulsive Systems*, vol. 13, pp. 187–220, 2006.

[57] I. Cidon, R. Rom, and Y. Shavitt, “Analysis of Muti-path Routing,” *IEEE/ACM Trans. Networking*, vol. 7, pp. 885–896, Dec. 1999.

[58] D. Sidhu, R. Nair, and S. Abdallah, “Finding Disjoint Paths in Networks,” in *Proceedings of ACM SIGCOMM*, 1991, pp. 43–51.

[59] A. Tsirigos and Z. Haas, “Analysis of Multipath Routing - Part I: The Effect on the Packet Delivery Ratio,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 138–146, Jan. 2004.

[60] ——, “Analysis of Multipath Routing - Part II: Mitigation of the Effects of Frequently Changing Network Topologies,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 500–511, Mar. 2004.

[61] M. Gong and S. Midkiff, “Distributed Channel Assignment Protocols: A Cross-Layer Approach,” in *Proceedings of IEEE WCNC*, vol. 4, 2005, pp. 2195–2200.

[62] S. Lee and M. Gerla, “Split Multipath Routing with Maximally Disjoint Paths in Ad Hoc Networks,” in *Proceedings of IEEE ICC*, vol. 10, May 2001, pp. 3201–3205.

[63] D. Johnson and D. A. Maltz, “Dynamic Source Routing in Ad Hoc Wireless Networks,” *Mobile Computing*, pp. 153–181, 1996.

[64] L. Qian, A. Elwalid, and I. Widjaja, “ICMP Extension for One-way Performance Metrics,” *Internet Draft*, Jul. 2000.

[65] G. Stüber, *Principles of Mobile Communication*, 2nd ed. Kluwer Academic Publishers, 2001.

[66] H. Hassanein and A. Safwat, “Virtual Base Stations for Wireless Mobile Ad Hoc Communications: an infrastructure for the infrastructure-less,” *International Journal of Commun. Syst.*, vol. 14, pp. 763–782, 2001.

[67] C. Lin and M. Gerla, “Adaptive Clustering for Mobile Wireless Networks,” *IEEE Journal on Selected Areas of Communications*, vol. 15, no. 9, pp. 1265–1275, Sep. 1997.

[68] R. Ramanathan and M. Steenstrup, “Hierarchically-Organized, Multihop Mobile Wireless Networks for Quality-of-Service Support,” *Mobile Networks and Applications*, vol. 3, no. 1, pp. 101–119, 1998.

[69] S. Xu and T. Sandawi, “Does the IEEE 802.11 MAC Protocol Work Well in Multihop Wireless Ad Hoc Networks?” *IEEE Communications Magazine*, pp. 130–137, Jun. 2001.

[70] L. Qian, X. Li, D. Vaman, and Z. Gajic, “Power Control and Proportional Fair Scheduling with Minimum Rate Constraints in Clustered Multihop TD/CDMA Wireless Ad Hoc Networks,” in *IEEE Wireless Communication and Networking Conference (WCNC)*, vol. 2, Apr. 2006, pp. 763–769.

[71] ——, “Power Control and Scheduling with Minimum Rate Constraints in Clustered Multihop TD/CDMA Wireless Ad Hoc Networks,” *Journal of Wireless Communication and Mobile Computing (WCMC)*, vol. 6, pp. 791–808, 2006.

[72] U. Kozat, I. Koutsopoulos, and L. Tassiulas, “A Framework for Cross-Layer Design of Energy-Efficient Communication with QoS Provisioning in multihop Wireless Networks,” in *Proceedings of IEEE INFOCOM*, vol. 2, 2004, pp. 1446–1456.

[73] R. Bhatia and M. Kodialam, “On Power Efficient Communication over multihop Wireless Networks: Joint Routing, Scheduling and Power Control,” in *Proceedings of IEEE INFOCOM*, vol. 2, 2004, pp. 1457–1466.

[74] J. Fang and R. Rao, “An Integrated and Distributed Scheduling and Power Control Algorithm for Maximizing Network Utility for Wireless Multihop Networks,” in *Proceedings of IEEE MILCOM*, vol. 2, Oct. 2003, pp. 1011–1017.

[75] M. Neely, E. Modiano, and C. Rohrs, “Dynamic Power Allocation and Routing for Time Varying Wireless Networks,” in *Proceedings of IEEE INFOCOM*, Apr. 2003, pp. 745–755.

[76] L. Qian, X. Li, D. Vaman, and Z. Gajic, “Joint Power Control and Proportional Fair Scheduling with Minimum Rate Constraints in Clustered Based MANET,” in *Proceedings of The 2nd International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, Dec. 2006, pp. 197–208.

[77] N. H. Vaidya, P. Bahl, and S. Gupta, “Distributed Fair Scheduling in a Wireless LAN,” in *Proceedings of ACM MobiCom, Boston, MA*, Aug. 2000, pp. 167–178.

[78] H. Ferng, C. Lee, J. Huang, and G. Chiu, “Designing a Fair Scheduling Mechanism for IEEE 802.11 Wireless LANs,” *IEEE Communications Letters*, vol. 9, pp. 301–303, Apr. 2005.

[79] Y. Fan, C. Huang, and Y. Hong, “Timer Based Scheduling Control Algorithm in WLAN for Real-Time Services,” in *Proceedings of IEEE International Symposium on Circuits and Systems*, May 2005, pp. 4533–4537.

[80] M. Andrews, L. Qian, and A. Stolyar, “Optimal Utility Based Multi-User Throughput Allocation subject to Throughput Constraints,” in *Proceedings of IEEE INFOCOM*, vol. 4, Mar. 2005, pp. 2415–2424.

[81] K. Kumaran and L. Qian, “Uplink Scheduling in CDMA Packet-Data Systems,” in *Proceedings of IEEE INFOCOM*, vol. 1, Apr. 2003, pp. 292–300.

[82] B. Radunovic and L. Boudec, “Optimal Power control, Scheduling, and Routing in UWB Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1252–1270, Sep. 2004.

[83] A.Jalali, R.Padovani, and R.Pankaj, “Data Throughput of CDMA-HDR: a High Efficiency - High Data Rate Personal Communication Wireless System,” in *Proceedings of the IEEE Semiannual Vehicular Technology Conference, VTC2000-Spring*, vol. 3, May 2000, pp. 1854–1858.

[84] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic Beamforming using Dumb Antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, 2002.

[85] A. Stolyar, "On the Asymptotic Optimality of the Gradient Scheduling Algorithm for MultiUser Throughput Allocation,," *Operations Research*, vol. 53, no. 1, pp. 12–25, Jan.-Feb. 2005.

[86] P. Gill and W. Murray, *Numerical Methods for Constrained Optimization.* Academic Press, London, 1974.

[87] X. Liu, E. Chong, and N. Shroff, "A Framework for Opportunistic Sheduling in Wirless Network," *Computer Networks*, vol. 41, pp. 451–474, 2003.

[88] S. Toumpis and A. Goldsmith, "Capacity Regions for Wireless Ad Hoc Networks," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 736–748, Jul. 2003.

[89] F. Kelly, "Charging and Rate Control for Elastic Traffic," *European Transactions on Telecommunications*, vol. 8, pp. 33–37, 1997.

[90] S. Borst, "A User-Level Performance of Channel-Aware Scheduling Algorithms in Wireless Data Networks," *IEEE/ACM Transactions on Networking*, vol. 13, pp. 636–647, Jun. 2005.

[91] H. Kushner and P. Whiting, "Asymptotic Properties of Proportional Fair Sharing Algorithms," in *Proceedings of the 40th Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2002.

[92] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R.Vijayakumar, "Providing Quality of Service over a Shared Wireless Link," *IEEE Communication Magazine*, vol. 39, no. 2, pp. 150–154, Feb. 2001.

[93] L. Jiang and S. Liew, "Proportional Fairness in Wireless LANs and Ad Hoc Networks," in *Proceedings of IEEE WCNC*, 2005, pp. 1551–1556.

[94] B. Raduovic and J. Boudec, "Rate Performance Objectives of Multi-hop Wireless Networks,," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 334–349, Oct. 2004.

[95] L. Tassiulas and A. Ephremides, "Stability Properties of Constrained Queueing Systems and Scheduling for Maximum Throughput in Multihop Radio Networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1949, Dec. 1992.

[96] N. Kahale and P. Wright, "Dynamic Global Packet Routing in Wireless Networks,," in *Proceedings of IEEE INFOCOM*, 1997, pp. 1414–1421.

[97] A. Eryilmaz and R. Srikant, "Fair Resource Allocation in Wireless Networks using Queue-length-based Scheduling and Congestion Control,," in *Proceedings of IEEE INFOCOM*, Mar 2005, pp. 1794–1803.

[98] "3GPP2 Standard," in *http://www.3gpp2.org.*

[99] M. Grossglauser and D. Tse, “Mobility Increases the Capacity of Ad-hoc Wireless Networks,” *IEEE/ACM Transactions on Networking*, vol. 10, pp. 477–486, Aug. 2002.

[100] M. McHenry, “NSF Spectrum Occupancy Measurements Project Summary,” in *http://www.sharedspectrum.com/inc/content/measurements/nsf /NSF_Project_Summary.pdf*, Aug. 2005.

[101] J. Mitola, “Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio,” in *Doctoral Dissertation, Royal Inst. Technol. (KTH), Stockholm, Sweden*, 2000.

[102] FCC, “Unlicensed Operation in The TV Broadcast Bands, ET Docket No. 04-186; Additional Spectrum for Unlicensed Devices Below 900 MHz And in The 3 GHz Band, ET Docket No. 02-380,” in *ET Docket no.04-113, [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-04-113A1.pdf*, 2004.

[103] ——, “Unlicensed Operation in the TV Broadcast Bands; Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band,” in *ET Docket no.06-156 [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-06-156A1.pdf*, 2006.

[104] M. Marcus, “Unlicensed Cognitive Sharing of TV Spectrum,” *IEEE Communication Magazine*, pp. 24–25, May 2005.

[105] A. Sahai, N. Hoven, S. Mishra, and R. Tandra, “Fundamental Tradeoffs in Robust Spectrum Sensing for Opportunistic Frequency Reuse,” in *Technical Report. [Online]. Available: http://www.eecs.berkeley.edu/sahai/Papers/CognitiveTechReport06.pdf*, 2006.

[106] X. Liu and S. Shankar, “Sensing-Based Opportunistic Channel Access,” *ACM Journal on Mobile Networks and Applications (MONET)*, vol. 11, no. 1, pp. 577–591, May 2006.

[107] Q. Zhao, L. Tong, A. Swami, and Y. Chen, “Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework,” *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 589–600, Apr. 2007.

[108] N. Hoven and A. Sahai, “Power scaling for cognitive radio,” in *Proceedings of WirelessCom 05 Symposium on Emerging Networks, Technologies and Standards*, vol. 1, Jun. 2005, pp. 250–255.

[109] L. Qian, X. Li, J. Attia, and Z. Gajic, “Power Control for Cognitive Radio Ad Hoc Networks,” in *Proceedings of 15th IEEE LAN/MAN Workshop*, Jun. 2007, pp. 7–12.

[110] ——, “Joint Power Control and Admission Control for CDMA Cognitive Radio Networks,” in *Proceedings of 15th IEEE LAN/MAN Workshop*, Jun. 2007.

[111] FCC, “KRON-TV at San Francisco,” in *[Online]. Available: http://www.fcc.gov/fcc-bin/tvq?list=0facid=65526: Retrieved Oct. 2006.*

[112] ——, "KRON-TV Service Contour Map," in *[Online]. Available: http://www.fcc.gov/fcc-bin/FMTV-service-area?x=TV2821.html: Retrieved Oct. 2006*.

[113] T. Basar and G. Olsder, *Dynamic Noncooperative Game Theory*. Academic Press, 2nd Edition, 1995.

Vita

XIANGFANG LI

Education and Employment

Sep. 1989 - Jul. 1993 B. Eng. from Beijing University of Aeronautics and Astronautics, Beijing, P.R.China.

Sep. 1993 - Apr. 1996 M.S. Eng. from Beijing University of Aeronautics and Astronautics, Beijing, P.R.China.

Apr. 1996 - Feb. 2000 Member of Technical Staff at Civil Aviation Computer Information Center of China (www.caacnet.com).

Sep. 2000 - Oct. 2003 M.S. from Rutgers University - The State University of New Jersey, Piscataway, New Jersey.

Sep. 2000 - present Graduate Student at Rutgers University - The State University of New Jersey, Piscataway, New Jersey.

Publications

1 X. Li and Z. Gajic, "An Improved SIR-based Power Control for CDMA Systems using Steffensen Iterations," *in Proceedings of Conference on Information Sciences and Systems (CISS)*, pp. 287-290, Princeton, NJ, Mar. 2002.

2 X. Li, "Uplink Power Control and Scheduling for CDMA Packet Data Systems," *in Proceedings of 40th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1308-1317, Oct. 2002.

3 X. Li and Z. Gajic, "Centralized Power Control in Coordinated CDMA Systems using Krylov Subspace Iterations," *in Proceedings of IASTED International Conference on Communications, Internet and Information Technology (CIIT)*, pp. 305-309, Nov. 2002.

4 X. Li and Z. Gajic, "Centralized Power Control in Coordinated CDMA Systems using Krylov Subspace Iterations," *WINLAB Technical Report 233*, Feb. 2003.

5 X. Li and Z. Gajic, "An Improved SIR-based Power Control for CDMA Systems using Steffensen Iterations," *WINLAB Technical Report 235*, Mar. 2003.

6 L. Qian, X. Li and Z. Gajic, "Adaptive Discrete Power Control for CDMA Systems," *in Proceedings of Conference on Information Sciences and Systems(CISS)*, pp. 1044-1049, Princeton, NJ, Mar. 2004.

7 L. Qian, N. Song, and X. Li, "Detecting and Locating Wormhole Attacks in Wireless Ad Hoc Networks through Statistical Analysis of Multi-path," *in Proceedings of IEEE Wireless Communication and Networking Conference (WCNC)*, pp. 2106-2111, New Orleans, LA, Apr. 2005.

8 L. Qian, N. Song, and X. Li, "Wormhole Attacks Detection in Wireless Ad Hoc Networks: A Statistical Analysis Approach," *in Proceedings of 1st International Workshop on Security in Systems and Networks (SSN'2005))*, Apr. 2005.

9 L. Qian, N. Song, and X. Li, "Detection of Wormhole Attacks in Multi-path Routed Wireless Ad Hoc Networks: A Statistical Analysis Approach," *Journal of Network and Computer Application*, pp. 1-23, Jul. 2005.

10 X. Li, S. Koskie, Z. Gajic, "Stefenssen Iterations for Power Updates in CDMA Wireless Networks," *in Proceedings of DCDIS 4th International Conference on Engineering Applications and Computational Algorithms*, Ontario, Canada, Jul. 2005.

11 L. Qian, X. Li, D. Vaman, and Z. Gajic, "Power Control and Proportional Fair Scheduling with Minimum Rate Constraints in Clustered Multihop TD/CDMA Wireless Ad Hoc Networks," *in Proceedings of IEEE Wireless Communication and Networking Conference (WCNC)*, pp. 763-769, Las Vegas, NV, April 3-6, 2006.

12 L. Qian, X. Li, D. Vaman, and Z. Gajic, "Joint Power Control and Maximally Disjoint Routing for Reliable Data Delivery in Multihop CDMA Wireless Ad Hoc Networks," *in Proceedings of IEEE Wireless Communication and Networking Conference (WCNC)*, pp. 300-306, Las Vegas, NV, April 3-6, 2006.

13 L. Qian, X. Li, D. Vaman, and Z. Gajic, "Power Control and Scheduling with Minimum Rate Constraints in Clustered Multihop TD/CDMA Wireless Ad Hoc Networks," *Journal of Wireless Communication and Mobile Computing (WCMC)*, vol.6, pp. 791-808, Wiley, 2006. (DOI: 10.1002/wcm.442)

14 L. Qian, N. Song, and X. Li, "Secure Anonymous Routing in Clustered Multihop Wireless Ad Hoc Networks," *in Proceedings of IEEE on Information Sciences and Systems (CISS)*, pp. 1629-1634, Princeton, NJ, Mar. 2006.

15 L. Qian, X. Li, D. Vaman, and Z. Gajic, "Joint Power Control and Proportional Fair Scheduling with Minimum Rate Constraints in Clustered Based MANET," *in Proceedings of The 2nd International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, pp. 197-208, Hong Kong, China, Dec. 2006.

16 L. Qian, X. Li, J. Attia, and Z. Gajic, "Power Control for Cognitive Radio Ad Hoc Networks," *in Proceedings of the 15th IEEE LAN/MAN Workshop*, pp. 7-12, Princeton, NJ, Jun. 2007.

17 L. Qian, X. Li, J. Attia, and Z. Gajic, "Joint Power Control and Admission Control for CDMA Cognitive Radio Networks," *in Proceedings of the 15th IEEE LAN/MAN Workshop*, Princeton, NJ, Jun. 2007.

18 L. Qian, N. Song, and X. Li, "SARC: Secure Anonymous Routing for Cluster based MANET," *Book Chapter in Wireless Communication Research Trends, Nova Scientific Publisher*, Chapter 2, 2007.