Radio Resource Management in CDMA-based Cognitive and Cooperative Networks

RADIO RESOURCE MANAGEMENT IN CDMA-BASED COGNITIVE AND COOPERATIVE NETWORKS

BY

BIN WANG, M.Sc

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2011) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE: Radio Resource Management in CDMA-based Cognitive and Cooperative Networks

AUTHOR: Bin Wang

M.Sc, (Electrical and Computer Engineering)

Southeast University, Nanjing, China

SUPERVISOR: Dr. Dongmei Zhao

This thesis is dedicated to my wife Min Shen, without her enormous love and unconditional support I would not have come even close to what I have achieved today.

Abstract

In this thesis we study radio resource management (RRM) in two types of CDMA-based wireless networks, cognitive radio networks (CRNs) and cooperative communication networks. In the networks, all simultaneous transmissions share the same spectrum and interfere with one another. Therefore, managing the transmission power is very important as it determines other aspects of the network resource allocations, such as transmission time and rate allocations. The main objective of the RRM is to efficiently utilize the available network resources for providing the mobile users with satisfactory quality of service (QoS).

In Chapters 2-4, we study the resource management in a CDMA-based ad hoc CRN with spectrum underlay, where the CRN shares the same spectrum with the uplink of a cellular CDMA network, which is the primary network. We first jointly consider the resource allocations in both the primary and the secondary networks, and study the optimum transmission power and rate allocations for supporting best effort traffic in the CRN. Fair transmission throughput is provided to the secondary links at each time slot. We then study how to support long-term best effort traffic in the CRN, and provide proportional fairness to the average throughput among links. An optimum scheduling problem is formulated and solved, and two heuristic

scheduling schemes are designed by opportunistically taking advantage of the available resources with low computation complexity. The resource allocation problem is further extended for supporting traffic with strict QoS requirements, where admission control and packet transmission scheduling are jointly considered to provide QoS for two types of traffic, streaming traffic requiring low outage probability, and non-real-time traffic with a minimum transmission rate requirement. For both types of the traffic, the average throughput of each admitted link is guaranteed by the admission control decisions. The outage performance of the streaming traffic is then guaranteed by a higher priority in the scheduling, and the instantaneous transmission power and rate for each link is obtained based on information of the measured interference and link conditions.

Chapter 5 of the thesis is devoted to study power allocations for a cellular CDMA network with cooperative relaying. Both the uplink and the downlink are considered. An optimum power distribution problem is formulated and solved for each direction, and heuristic schemes are then proposed. Different relay strategies and relay station selection criteria are considered.

Acknowledgements

First, I would like to express my sincere gratitude to Professor Dongmei Zhao, my advisor, for all her support and advice toward the accomplishment of my research work and this dissertation. She guided me to identify the problems in wireless networks, and gave me all kinds of advices and supports. I am very grateful for her suggestions to keep me on track and focused, for her insight and time sharing with me. She is my advisor not only on research, but also on morality.

I would like to thank Dr. Terence D. Todd and Dr. Ted H. Szymanski for being on my thesis committee, for their contribution of precious time in providing suggestions, comments and questions that greatly helped clarifying issues and improving quality of the thesis. To Dr. Polychronis Koutsakis, I would like to give my thanks for his advise and help on my research work when he worked with us. I would also like to thank the anonymous external examiner for reading my thesis and providing valuable comments.

It has been a great experience to work at wireless networking laboratory (WNL), where I meet and work with many great lab members. I greatly appreciate the discussion, help and collaboration from past and current WNL members.

My special thanks go to my wife, Min Shen, for her unconditional support, love and faith on me in the course of my research work. Thanks to the devotion and

happiness she brings to our family.

Last but not least, I would like to express my gratitude to my parents and relatives for their consistent love and support.

Contents

List of Tables

List of Figures

List of Abbreviations

1G	first generation
2G	second generation
3G	third generation
GSM	global system for mobile Communications
DCS	digital cellular system
PCS	personal communications service
SMS	short message service
GPRS	general packet radio service
HSCSD	high speed circuit switched data services
IMT	international mobile telecommunications
HSD	high speed data
EDGE	enhanced data rates for GSM evolution

3GPP	3G partnership project
CDMA	code division multiple access
WLAN	wireless local area network
RAT	radio access technology
UWB	ultra wide-band
DVB	digital video broadcasting
BS	base station
MS	mobile station
RRM	radio resource management
QoS	quality of service
TACS	total access communication system
BER	bit error rate
SINR	signal to interference plus noise ratio
MMPP	Markov modulated Poisson process
NE	Nash equilibrium
ILP	integer linear programming
CH	cluster head
iCAR	integrated cellular and ad hoc relaying

RS	relay station
SS	source station
DS	destination station
UCAN	unified cellular and ad-hoc network
WCDMA	wideband code division multiple access
GPS	global positioning system
OTDOA	observed time difference of arrival
DF	decode and forward
AF	amplify and forward
CSI	channel state information
CRN	cognitive radio network
DSA	dynamic spectrum access
WRAN	wireless regional area networks
CR	cognitive radio
PU	primary user
SU	secondary user
SCH	superframe control header
DOSS	dynamic open spectrum sharing

MCCG	multi-Channel contention graph
PF	proportional fairness
JPAC	joint power and admission control
AWGN	additive white gaussian noise
ERA	equal rate allocation
GP	geometric programming
SUS	single user scheduling
MUS	multi-user scheduling
ACK	acknowledgement frame
SI	scheduling interval
MUSU	multi-user scheduling utility
MUSI	multi-user scheduling interference
nrt	non-real-time
MRC	maximum ratio combining

List of Symbols

W	spreading bandwidth
γ_i^*	minimum SINR for successful decoding of link i
I_{th}	interference threshold of secondary networks with spectrum underlay
$P_{p,i}$	transmission power of the ith primary link
$P_{s,i}$	transmission power of the ith secondary link
M_p	number of primary links
M_s	number of secondary links
$R_{p,i}$	transmission rate of the ith primary link
$R_{s,i}$	transmission rate of the ith secondary link
$P_{p,\max}$	maximum primary transmission power
$P_{s,\max}$	maximum secondary transmission power

$g_{p2p,ij}$	link gain from the ith primary transmitter to the jth primary receiver
$g_{p2s,ij}$	link gain from the ith primary transmitter to the jth secondary receiver
$g_{s2p,ij}$	link gain from the ith secondary transmitter to the jth primary receiver
$g_{s2s,ij}$	link gain from the ith secondary transmitter to the jth secondary receiver
γ_p^*	minimum SINR for successful decoding of primary link
γ_s^*	minimum SINR for successful decoding of secondary link
$G_{p,i}$	processing gain of the ith primary link
η	background additive white Gaussian noise power
$\mathbf{R}_s$	vector of secondary transmission rate, $\mathbf{R}_s = (R_{s,1}, R_{s,2}, \ldots, R_{s,N_s})$
$I_{p,i}$	aggregate noise and interference that the ith primary link experiences
d_i	distance between the transmitter and the receiver of the ith link
α	path loss exponent
ΔP	primary link's transmit power difference w/o the secondary transmissions
ΔE_b	energy increase of primary link by each transmitted bit in CRN
$\mathbf{R}(t)$	vector of instantaneous secondary user rate, $\mathbf{R}(t) = [R_1(t), R_2(t), \ldots, R_N(t)]$
$\overline{\mathbf{R}}$	vector of average secondary rate, $\overline{\mathbf{R}} = [\overline{R}_1, \overline{R}_2, \ldots, \overline{R}_N]$
$U(\overline{\mathbf{R}})$	utility function of $\overline{\mathbf{R}}$
$\mathbf{r}(t)$	vector of allocated rate at time t

$\overline{\mathbf{X}}(t)$	exponentially smoothed average rate for link i up to time t
t_c	parameter to balance the past and current transmission rates
$I_{p2s,i}(t)$	interference from primary users to the ith secondary link at time t
$I_{s2s,i}(t)$	interference from other secondary links to the ith link at time t
$\hat{P}_i(t)$	maximum transmit power limited by the equipment and I_{th}
$I_{\text{th,rem}}$	remaining interference threshold
$\mathcal{T}$	set of secondary links in the same exclusive region
$Q_i(t)$	weighted contribution to interference from link i at time t
$w_i(t)$	distributed scheduling weight of link i at time t
$\mathcal{D}$	set of the nrt data connections
$\mathcal{S}$	set of the streaming connections
$R_{\text{req},i}$	average transmission throughput required by the nrt link i
$B_{\text{max},i}$	maximum buffer size of streaming connection i
$B_i(t)$	buffer size of link i at time t
$\tilde{I}_{p2s,i}$	long term average of $I_{p2s,i}(t)$
$I_{\text{th,data}}$	interference threshold for secondary nrt links
T_i	duration of one time slot
$\xi_i(t)$	counter for nrt link

$L_i(t)$	priority level of link i at time t
T_h	total amount of time for backoff
$P_{\text{MS,max}}$	maximum transmission power of MS
$P_{\text{BS,max}}$	maximum transmission power of BS
N^u	number of connections in the uplink
N^d	number of connections in the downlink
r_i^u	RS of connection i in the uplink
P_i^u	transmission power of SS i
$P_{r_i^u}$	transmission power of the RS of connection i
$\gamma_{s,i}^u$	received SINR at BS from SS i
$\gamma_{c,i}^u$	received SINR at RS from SS i
$\gamma_{r,i}^u$	received SINR at RS from RS i
P_b	total transmission power received at the BS
$\gamma_{s,i}^d$	received SINR at DS from BS i
$\gamma_{c,i}^d$	received SINR at RS from BS i
$\gamma_{r,i}^d$	received SINR at DS from RS i
C_i	set of the potential RSs of SS i

Chapter 1

Introduction

1.1 Evolution of Wireless Communication Networks

The concept of the first generation (1G) cellular systems was developed by Bell Laboratories in 1947. It was not carried out until 1979 when the technological developments such as integrated circuits, microprocessors, frequency synthesizers, etc., had made it possible. The 1G systems are analog and were designed for wireless speech services. The developments of digital signal processing methods along with the rapid development of integrated circuits and microprocessors led to the replacement of the analog 1G cellular systems by the digital second generation (2G) cellular systems [1]. The first of these, the Global System for Mobile Communications (GSM), was realized in 1992 in Europe. It operates in the 900 MHz band, and is based on FDMA/TDMA. Variants of GSM have been developed for higher frequency bands, such as the Digital Cellular System 1800 (DCS 1800) in Europe and Personal Communications Service

1900 (PCS 1900) in North America. The major improvements offered by the digital transmissions of the 2G systems over the 1G systems are better speech quality, increased capacity, global roaming, and data services like the Short Message Service (SMS) [2]. Major improvements in the data services were also the introduction of packet switched services such as the General Packet Radio Service (GPRS) [3] [4], and higher-data-rate circuit switched services such as the High Speed Circuit Switched Data services (HSCSD) [5] [6].

Although the 2G systems could already provide some basic data services, the possible data rates were relatively low and cannot satisfy the needs of future mobile services like mobile web browsing, file transfer, real-time video, digital TV, etc. The third generation (3G) cellular systems are known with the name International Mobile Telecommunications for the year 2000 (IMT-2000) [7], and are being implemented in many countries around the world. The 3G systems introduce wireless wideband packet-switched data service for wireless access to the Internet with speeds up to 2 Mb/s.

The 2G systems have been evolving with the introduction of new technology enhancements, such as GPRS and HSCSD in GSM, and High Speed Data (HSD) [8] in IS-95. A step further towards the 3G networks is the Enhanced Data rates for GSM Evolution (EDGE) technology [9], which enables three times higher data rates than those possible with the ordinary GSM/GPRS network [9]. Several standardization bodies joined their forces in 1998 in the 3G Partnership Project (3GPP) agreement [10] with the joint goal of producing globally applicable technical specifications and technical reports for a 3G mobile system.

The 3G systems from their initial designing stage are aiming at providing ubiquitous wireless networks that can support voice, multimedia and high-speed data communications. To be of interest to the customers, these new services should be provided at a low cost and with high quality. An important step for achieving these goals is the selection of the multiple access methods. Code Division Multiple Access (CDMA) has been selected as the air interface for these networks. Upon these new physical transmission technologies, flexible and efficient radio resource management schemes are applied to further enhance the efficiency of the technology.

The 3G systems greatly improve both the quality and the diversity of the communication services. This can be regarded as one major drive for the rapid development of the intelligent mobile phones, and brings a new wave transition from the personal computer end to the handset devices [11] [12]. As the counterparts, short range and high speed wireless technology has been progressed very fast in recent years to supplement those long distance cellular systems. Wireless Local Area Networks (WLANs), with the benefits of easy set-up, low cost and free of spectrum usage license, become a very popular wireless transmission method for local area services [13]. Versatile internet applications and services can be conveniently accessed at home, schools and any other hot spots deployed areas [14]. In addition, new emerging radio access technologies (RATs), like Ultra Wide-band (UWB) [15] and Digital Video Broadcasting (DVB) [16], are also commercialized and applied in different areas.

The new emerging RATs together with the traditional wireless access methods comprise a set of discrete technologies. Integration of different RATs may provide chances for innovative services with flexibility and in a cost-efficient manner. For example, users can switch to the preferable interfaces based on the required service

quality or communication costs. How to fit each of these technologies into a global integrated infrastructure, optimize the advantages of each RAT, and incorporate the future requirements of operators and user expectations, are open questions for future wireless communication networks designs [17] [18] [19].

The data-centric communication follows the similar pattern of voice communication, particularly driven by the industry push and growing user acceptance of wireless Internet. The demand for increasing the support from wireless communications [20] which requires more efficient and effective utilization of the limited radio spectrum, which has been a vital issue and is becoming even more important than developing new RATs [21]. New resource management methods should be invented to improve the efficiency of the existing wireless technologies. On the other hand, more dynamic RATs should be developed to sense the context of environment and adapt to the most suitable wireless band in order to utilize the spectrum efficiently.

Radio Resource Management (RRM)

During the network planning stage, the operators decide the right number and locations of network devices, such as base stations (BSs) and relays, to be deployed based on users' demands, transmission environment and other factors [22]. The physical layer design determines the maximum amount of available network resources. How to utilize the available radio resources in the most effective and efficient way is one of the most important objectives of RRM. The effectiveness is: to provide the users with satisfactory quality of service (QoS), and the efficiency is to increase the revenue of the network service providers.

The radio resource in traditional circuit switched wireless networks is defined by a

set of basic physical transmission parameters necessary to support a signal waveform transporting end user's information [23]. In particular, a radio resource unit in an FDMA-based network is equivalent to a certain bandwidth within a given carrier frequency. For example, in the 1G Total Access Communication System (TACS), a radio resource unit is a 25 kHz portion in the 900 MHz band. A radio resource unit in a TDMA-based network is equivalent to a pair consisting of a carrier frequency and a time slot. For example, in GSM, a radio resource unit is a time slot with duration of 0.577 ms at a 200 kHz carrier bandwidth. Given the number of the frequency channels and time slots, the resource management operation is relatively limited to assign these resource units or channels to the requesting users.

In packet switched networks, network resources are shared by different connections dynamically. This provides opportunities for more efficient resource utilization since one connection can use the resource that other connections are not using. As a result, the complexity of RRM is increased. The task of RRM becomes even more challenging when the traffic is multimedia with different QoS requirements and the users are mobile. To meet the wide-ranging QoS requirements of these applications, traffic scheduling algorithms should be employed. Scheduling algorithms are important components in the provision of guaranteed quality of service such as delay, delay jitter, packet loss rate, and throughput. A scheduling algorithm may possess some basic features [24]. First, the algorithm should utilize the channel efficiently. This implies that the scheduler should avoid assigning transmission network resource to links with poor transmission condition. Second, the algorithm should be able to provide delay bound guarantees for individual users in order to support delay-sensitive applications. Third, the algorithm should redistribute available resources fairly across

users. Fourth, the algorithm should provide guaranteed throughputs for users. It should be noted that not all these features can be achieved simultaneously in one scheduling scheme. For example, to guarantee the strict delay requirement for a packet, a user may have to transmit in poor link conditions, which consumes a large amount of network resources and reduces the resource utilization efficiency. A tradeoff is necessary between fairness and efficiency based on different system objectives.

In order to support QoS traffic, it is necessary to control the amount of traffic admitted into the system so that the system is not overloaded. This is called admission control, which usually considers the long-term resource usage (a time period much longer than a single packet transmission time) so that a connection, once admitted into the system, can be served with the required QoS throughout its lifetime. Admission control requires not only knowledge about the traffic (e.g. average rate, maximum rate, burstiness, etc), but also the network conditions (such as channel fading and interference). User mobility is another factor that affects the admission decisions, since mobility management is needed to solve handover and BS selection problems. In addition, the amount of the traffic that a network can support depends on how resources are allocated among the users, and therefore, the admission decisions should be based on specific scheduling schemes. Such complications usually make admission control a very difficult problem. It is very difficult, if not impossible, to achieve the optimum admission control strategy so that both the QoS of all admitted users is satisfied and the network resource is fully utilized at all the time. Therefore, a practical admission control scheme may either over-estimate or under-estimate the amount of traffic that can be served in the network. In the former case, it depends on the scheduling process to decline the resource usage for some users; and in the latter

case, the extra network resources may still be utilized by serving best effort traffic.

When multiple users share the same frequency bands, such as in CDMA-based networks, mutual interference may limit the system capacity or other QoS performance, such as the service rate to each user. How to manage the transmission power of each user and therefore to control the interference level in the system is important in order to efficiently utilize the spectrum resources and provide satisfactory QoS to the users. Given the bit error rate (BER) requirement, which corresponds to the signal to interference plus noise ratio (SINR) requirement for specific modulation and coding schemes, minimizing the transmission power for each user can minimize the mutual interference and maximize the total number of supportable users. On the other hand, any approach that can reduce the required SINR, for example, by adjusting the modulation and coding schemes, can reduce the required transmission power of the users and improve the system capacity.

1.2 Radio Resource Management (RRM) in CDMA-Based Wireless Networks

A CDMA-based system is typically interference-limited. In such a system users may transmit their signals simultaneously in the same frequency band. Each transmitter is assigned a dedicated spreading code, which can be reproduced at the intended receiver to regenerate the transmitted signal. The cross-correlation of different spreading codes is ideally zero, so that the desired signal can be recovered and other interfering signals can be removed at the receiver. In a practical system, the radio channel can be non-linear and the spreading codes may not be orthogonal to one another. Therefore, the

transmission from one user may cause interference to other users. The more users are in the system and the higher power they transmit, the more interference they generate to one another. The management of transmission power and mutual interference is directly related to the system capacity and QoS to the users.

Power control

Power control is originally used to solve the near-far problems in the uplink of cellular CDMA networks when homogeneous traffic (mainly voice) is supported. In the uplink, all the transmissions in the same cell share the same receiver, which is the base station (BS). When all the users transmit at the same power, their signals arrive at the BS receiver with different strength. Signals from the users far away from the BS have much poorer quality than that from the users close to the BS. This is the "near-far" effect. In order to balance the received SINRs for the signals from different users, power control is applied so that the signals from different users all arrive at the BS with the same SINR.

The function of the power management in the downlink is different, since the signal and interference from the BS arrive at a given mobile station (MS) go through the same radio channel and undergo the same attenuation. Therefore, power control is not needed to combat the near-far problem. Instead, it is used to provide more power to users located near the cell borders, where the users can suffer from high interference from the transmissions in nearby cells. In addition, the transmission power of the BS should be controlled in order to reduce the interference to nearby cells.

A great deal of the work on power control in cellular CDMA systems has focused

on how to set the transmission power so that all users in the system have acceptable SINR, or bit-energy-to-interference spectral-density ratios (E_b/I_o), which is the normalized SINR per transmitted bit. This approach is based on a fairly reasonable assumption that the BER at a receiver is a monotonically decreasing function of E_b/I_o. Let R be the information bit rate, then the energy/bit is given by $E_b = P_r/R$, where P_r is the received power for homogeneous users. The interference power spectrum density is the interference power divided by the spreading bandwidth W, i.e., $I_0 = P_r(N-1)/W$. Hence the expression for E_b/I_0 becomes

$$E_b/I_o = \frac{P_r/R}{(N-1)P_r/W} = \frac{W/R}{N-1}. \tag{1.1}$$

The quantity W/R is called the processing gain. Solving for N we can find the capacity of the cell in number of users as

$$N = 1 + \frac{W/R}{E_b/I_0} \approx \frac{W/R}{E_b/I_0} \triangleq N_w. \tag{1.2}$$

When considering multiple cells, transmissions from neighboring cells also cause interference to one another. When all the traffic is homogeneous, the interference experienced from other cells is a fixed ratio f times the interference experienced from the same cell, where $f \approx 0.4$ [25]. Then the per-cell capacity of a multi-cell system is given by

$$N_w = \frac{1}{1+f}\frac{W/R}{E_b/I_0}. \tag{1.3}$$

As the development of wireless networks, an increasing amount of traffic is heterogeneous. Each user may have a different SINR or E_b/I_0 requirement. Let R_i be the transmission rate and $P_{u,i}$ be the transmission power of user i, γ_i^* represent the

E_b/I_0 requirement of the user, $g_{u,i}$ represent the link gain from the user to the BS, and η represent the noise at the BS receiver. A more general expression about the transmission power is given by

$$\frac{W}{R_i}\frac{P_{u,i}g_{u,i}}{\sum_{j\neq i}P_{u,j}g_{u,j}+\eta} \geq \gamma_i^*, \tag{1.4}$$

for all i's. Let $\mathbf{P} = [P_{u,1}, P_{u,2}, ..., P_{u,N}]$ be the vector with $P_{u,i}$ as the ith element, and $\mathbf{1}$ be the vector of the same size as $\mathbf{P}$ with all elements equal to 1. We can rewrite the above relationship into the vector form as

$$\mathbf{W}\mathbf{P}^T \geq \eta\mathbf{1}^T, \tag{1.5}$$

where $\mathbf{W} = [\mathbf{W}_{ij}]$ is an $N \times N$ matrix defined as:

$$\mathbf{W}_{ij} = \begin{cases} \frac{Wg_{u,i}}{R_i\gamma_i^*}, & \text{when } i = j \\ -g_{u,j}, & \text{when } i \neq j. \end{cases} \tag{1.6}$$

For the downlink, the relationship of power distribution is similar except that different receivers may experience different noise power. Let $P_{d,i}$ be the transmission power of the BS to the ith MS, $g_{d,i}$ be the link gain from BS to MS i, and η_i be the noise power experienced at MS i. The downlink transmission power should satisfy the following relationship,

$$\frac{W}{R_i}\frac{P_{d,i}g_{d,i}}{\sum_{j\neq i}P_{d,j}g_{d,i}+\eta_i} \geq \gamma_i^*. \tag{1.7}$$

In terms of implementation, the BS can track the link gains between itself and

each user, then adjust the transmission power to the downlink users or inform the users to adjust their transmission power in the uplink. By controlling the transmission power, the SINR of each connection can be protected, the interference to other users can be reduced, and the system capacity can be maintained.

1.2.1 Packet transmission scheduling

The achievable performance of dynamic traffic scheduling is coupled to the time varying conditions, including both radio channel and interference conditions. For best effort services, rather than combating a poor channel by high transmission power, which generates high interference, transmissions can be postponed to when the channel conditions turn to be good. This concept has appeared in various forms for DS-CDMA. It is shown in [26] that the reduction of transmission power at deep fading helps improve the system throughput by above 200%. A truncated rate adaptation scheme is considered in [27], where data transmission is suspended when the channel gain is below a threshold and then the rate is adapted with a constant power. It has been shown that under certain conditions for a single cell case, only the user with the best channel condition should access the whole bandwidth at any time [28] in order to maximize the uplink capacity. However, when users have the minimum BER requirements and the transmission rates belong to a discrete set, the maximum capacity may be achieved by simultaneous transmissions. In [29] a combined intracell scheduling and power control is proposed, where only one user in each cell is allowed to transmit at any given time. The scheduling is made to guarantee that the average throughput is satisfied for each link. The scheduling is combined with a distributed power control algorithm to adjust cell-site powers, so that the average data rate is supported with

as low transmission power as possible.

1.2.2 Admission control

The purpose of admission control is two-fold. First, a connection request can be accepted into the system only if there is sufficient resource in the system to satisfy its required QoS. Second, all existing connections should still be able to receive their required QoS after admission of the new connection. Effectiveness and efficiency are two performance metrics to evaluate an admission control scheme. The effectiveness is to guarantee the QoS of the admitted traffic, and the efficiency is to maximize the amount of traffic admitted into the system.

While network operators are opt to accept more users for given amount of network resource in order to increase their revenue, the number of users that can be accommodated in the system is reversely proportional to the amount of resource used by each user. The simplest CAC is to limit the number of users admitted in the network. In TDMA or FDMA-based networks, each user requires several time slots or frequency channels. In a CDMA-based network, the interference-limited nature makes it difficult to achieve accurate admission control. Because of the co-channel interference, the amount of resource, such as transmission power, required by each user is dependent on the number of users in the system, their geographical locations, and physical channel conditions. Therefore, QoS of all the users sharing the same system is related. For homogeneous traffic, N_w in (1.2) gives the maximum number of users that can be supported in a CDMA-based system. When there are multiple types of traffic, such as voice and data, one threshold is set for each type of the traffic, such as in [30], and the number of admitted users for a given type of traffic cannot

exceed the respective threshold. Alternatively, for a system that can serve k types of traffic, a k-dimensional admissible region can be defined based on combinations of the number of users for each type of the traffic, such as in [31]. In [32] the cell loading (bandwidth utility) for multi-class traffic is expressed as a function of the required SINR of the users and the total number of users.

When a new call request arrives in a cell, its required bandwidth utility is calculated. Meanwhile the bandwidth utility required by existing users are recalculated. An admission control decision is then made based on the calculated bandwidth utilities.

In a circuit switched network where each user is always served with a constant rate, SINR is commonly used as the criterion for admission control, since there is a direct mapping between the received signal strength and the BER, given the coding and modulation techniques. A connection is in outage if the actual received SINR is below the minimum required SINR. Examples can be found in [33]- [34]. The CAC scheme proposed in [33] simply compares the SINR that can be provided to the new user with a threshold value and admits the new user only if the achievable SINR is larger than the threshold. In the SINR-based CAC scheme proposed in [35], different SINR thresholds are used for different traffic classes depending on the required BER. A CAC scheme is designed for the uplink in a cellular CDMA system in [36], where a residual capacity is defined based on the actual SINR and the SINR threshold. A new user is admitted if the residual capacity is greater than zero. The definition of the residual capacity is modified in [37] in order to predict the impact of admitting a new call in the target cell on its neighboring cells, including adjacent and nonadjacent cells. When the required SINR is larger, more connection requests are blocked. An upper bound of the SINR threshold is given in [34] to maintain a specified level

of the connection blocking rate. When serving variable rate traffic, predicting the long term QoS is difficult, especially for non-stationary and non-uniform traffic. The interference level and resource availability is predicted through Kalman filter in [38].

Handoff is a special issue in mobile networks. A user currently admitted in one cell may travel to a different cell later, and the user may travel through the coverage areas of multiple cells during the lifetime of its connection. An integrated SIR-based resource management scheme that encompasses CAC, handoff, power control, and dynamic channel allocations has been proposed in [33]. The station starts the handoff process received SINR is less than the predefined handoff threshold. Similarly a new call is accepted if the achievable SINR is greater than the admission threshold. Dynamic channel allocations, signal strength, base station assignment and asynchronous power control are integrated into the system model. In [39] a local interference-based CAC scheme is designed, where the handoff decision is made based on blocking probability constraints of various user classes, and a buffering scheme is proposed to cope with congestion.

Wireless networks are increasingly required to support multimedia services such as voice, video, data, and audio with various QoS profiles. Optimal admission policies are designed to maximize the network utilization while satisfying the users' requirements. In [40] [41] both the physical layer QoS and network QoS are guaranteed by cross-layer optimization for constant bit rate traffic. Optimal connection admission control schemes are proposed in [42] for CDMA-based networks to support variable bit rate traffic. The Markov-modulated Poisson process (MMPP) is used to model the packet arrival process and track the traffic variations. By introducing a small SIR outage probability, the proposed algorithms can admit more users and increase the network

utilization.

1.2.3 RRM in ad hoc wireless networks

In a CDMA-based network, simultaneous transmissions using the same bandwidth interfere with one another. A change in the power allocation on one link can induce changes in the interference conditions of all links in the surrounding area and change the performance of the ongoing flows. As the number of users increases, co-channel interference increases, which requires each transmitter to transmit at higher power in order to combat further increases the mutual interference.

Unlike the cellular networks, which provide single hop connectivity between the MSs and fixed BSs, ad-hoc networks do not have pre-configured network topologies. They do not rely on extensive and expensive installations of fixed BSs throughout the usage area, and peer nodes can directly communicate with one another or relay traffic for one another. The topology or connectivity of an ad-hoc network can change dynamically due to node mobility.

In an ad hoc network with all single hop links, if all the users behave well, an iterative power control method can achieve the desired SINR γ_i^*. Given $P_i(t)$ as the current transmission power for the transmitter of link i, and $\gamma_i(t)$ as the received SINR at the current time, the transmission power at the next time $P_i(t+1)$ is given by [43]

$$P_i(t+1) = \frac{\gamma_i^*}{\gamma_i(t)} P_i(t). \tag{1.8}$$

When there is a feasible solution, this algorithm converges [43]. This type of power control can be performed distributively by individual links. Each receiver measures its received SINR and compares it with the target SINR. It then informs the transmitter

to increase or decrease the transmission power depending on whether the actual SINR is below or above the target SINR. However, this distributed power control is difficult in multi-hop ad hoc networks, where the target SINR of each hop in a multi-hop link depends on that of other hops.

Transmission power affects the transmission range of a station. In multi-hop networks, controlling the transmission power is related to routing, which further affects the interference level in the system. By breaking a direct link between two users into multiple shorter links, the transmission power required by the sender is reduced. On the other hand, this requires more relay nodes, each contributing interference to the system. Therefore, in multi-hop networks, there is a tradeoff between the number of hops and the overall system performance [44]. The concept of controlling the transmission radius in multihop packet radio networks was first introduced in [45]. In [46], the authors developed a model for analyzing the throughput and forwarding process in a multihop network, where each mobile node may have a variable and different transmission range. The work in [47] employed transmission power as the link metric for the shortest path routing algorithms in an attempt to realize the minimum-power routing algorithm discussed in [48]. However, the congestion caused by multiuser interference was not represented in this link metric. Transmission power adjustment is used in [49] in order to control the topology of wireless ad hoc networks and facilitate the routing of the network.

Users in the ad hoc networks may transmit with non-negotiated power level and at non-negotiated time instants. The unexpected incoming interferer requires each user to have the extra transmission power reserved. A power margin is set in [50] to combat the possible increase of interference. In addition, more sophisticated power

distribution regulation is required to make the self-regulated nodes in the ad-hoc CDMA networks to cooperate and work well [51]. Some utility-based power control algorithms are proposed in [52] and [53], where the utility functions can integrate users' QoS (e.g., SINR) with the system performance. Each user is assumed to be selfish and chooses its transmission power level in order to maximize its own utility function. This is opposed to maximizing the total network utility. In order to improve the total network utility, the pricing mechanism is introduced in each user's utility function, and this discourages users from selfishly increasing their respective transmit power levels. Without a central controller, convergence is a problem for the power control schemes in ad hoc networks. Game theory is a commonly used tool to study not only users' behaviors but also the system stability [54] [55]. In [54], the utility function of different users is associated with each user in the network, and each user adapts its transmission power and waveform, and the system is guaranteed to reach a Nash equilibrium (NE). The work assumes a symmetric interference among distributed users, and the utility function does not lead to a stable NE when the mutual interference is asymmetric. A framework based on potential game theory is proposed in [55] to construct convergent power adaptation games to address the asymmetry of mutual interference.

Transmission power can affect different aspects of the network performance. Therefore, it is often considered jointly with other aspects in a network. In [56], a joint scheduling and power control strategy is proposed to limit multi-user interference to increase the single-hop throughput and reduce the power consumption. The scheme is applied for contention-based networks, and it is only constrained by the SINR requirement. In [57], scheduling is jointly considered with routing and power control. The

problem is formulated as a linear programming problem with an exponential number of constraints, and an iterative algorithm is proposed with lower computation complexity. In [58], a MAC layer scheduling algorithm is presented using integer linear programming (ILP). The licensed channels act as links to connect the secondary users with one another. Within a frame, a distributed scheduling algorithm is implemented to activate the available links without causing interference among the users. The scheduling algorithms proposed in [59] consider the interference, fading, and packet delay.

Because of the difficulty in resource and topology management in an ad-hoc network, some works try to impose an infrastructure in the network by dividing it into multiple clusters and having a cluster head (CH) in each cluster. The CH acts like a cellular base station, so that power control and resource allocations within a cluster can be managed or assisted by the CH [60] [61]. On the other hand, the ad hoc transmission mode is also used in traditional cellular communications for various performance improvement.

1.2.4 RRM in cellular networks with ad hoc relaying

A lot of work has been done to combine the advantages of the fixed infrastructure in the cellular networks and the flexibility in the ad hoc networks.

First, the relaying function of the ad hoc network can be added into a cellular network, so that the transmission power of the MSs can be reduced or the coverage area in the cellular network can he enhanced if the transmit power is unchanged [62] [63]. Furthermore, using relaying may improve the capacity of a cellular network. In [64], an integrated cellular and ad hoc relaying (iCAR) system is proposed. Fixed relay

stations (RSs) are deployed at the boundary of the cells. Traffic in the hot spot area can be forwarded to the surrounding lightly loaded cells via the RSs. In [65], a unified cellular and ad-hoc network (UCAN) multi-hop architecture is proposed. The goal of UCAN is to use multi-hop routing to improve the throughput when the quality of the signal in the downlink channel between the BS and the MS is poor. A proxy client with good signal quality relays the traffic from the MS to the BS. The studies reported on in [66] assess the potential capacity improvements in this type of the system when using a standard cellular frequency reuse plan. Relaying channels are allocated from channels that have been assigned to other cells, and therefore inter-cell interference has to be carefully considered. In [67], a mechanism called position assisted relaying (PAR) is proposed for wideband CDMA (WCDMA) cellular networks with dual-mode stations. In this scheme, a nearby station may relay transmissions for a station when that station's cellular link becomes unusable. Geo-location techniques, such as global positioning system (GPS) or observed time difference of arrival (OTDOA), are used by the BS to select a candidate RS, and the selection criteria may take into account the factors such as direction of movement. Capacity improvement using in-band ad hoc relaying in a CDMA-based system is limited as shown in [68], since the relaying traffic creates interference that limits the capacity, even with an algorithm that attempts to minimize the mutual interference. The study described in [69] proved that hotspot CDMA capacity can be substantially increased by out-of-band multi-hop traffic relaying. It has been shown that it is sometimes better to use multi-hop relaying to shrink the hotspot cell, but, in other cases, it is better to use multi-hop relaying to shed traffic from the hotspot area. The results have also shown that even a single ad hoc hop can achieve significant capacity improvement in

a cellular CDMA network.

Although relaying channel can help reduce the hop distance along a transmission route to save transmission power of individual stations, the transmission efficiency is still affected by the relay channel conditions. If the direct channel can still be retained, the gain of spatial diversity may further reduce the MS's transmission power and improve the system performance. Extensive research has been done to prove that using cooperative communications at the physical layer can greatly improve the transmission rate and reduce the transmission errors [70] [71]. Several works e.g. [72] [73], addressed the information-theoretic aspects of the cooperative communications to prove the network capacity improvement after using both the direct link and relay link.

Different practical schemes are proposed to improve the performance of existing cellular CDMA networks. In [74] an opportunistic cooperation is proposed which achieves the minimum outage probability for cooperative decode-and-forward (DF) networks by dynamically adjusting the transmission time and corresponding power of each node. However, this protocol requires perfect global channel state information (CSI), which imposes a stringent requirement on the capacity of the feedback channels. By considering a practical scenario wherein only a coarse channel feedback is available at the transmitter, [75] proposed a power control algorithm to minimize the outage probability for a wireless cooperative amplify-and-forward network. The authors in [76] studied the resource allocation problem and optimized the outage of a wireless DF network under several feedback schemes. Besides outage improvement, cooperative relaying can potentially save transmission power. This is studied in [77] for a pair of source and destination nodes and a number of RSs, in [78] [79] by combining power

minimization with routing in an ad hoc network, and in [80] [81] for cooperative broadcasting in dense wireless networks.

Both relaying and cooperative communication are aiming at signal strength enhancement. Recently, a new framework of cellular and ad hoc network, known as cognitive radio networks (CRNs) has been proposed for spectrum reuse [82] [83]. Cognitive networks can dynamically share a piece of spectrum with different priorities. Spectrum usage efficiency can be greatly improved to solve the current issue of wireless spectrum shortage and to meet the increasing demands on higher data rates and more stringent QoS requirement. The concepts of cognitive radio networks will be introduced in the next section.

1.3 Introduction to Cognitive Radio Networks

1.3.1 Cognitive radio background

Spectrum is scarce resource in wireless communications. Currently, fixed spectrum slices are licensed to each wireless service/technology. Recent studies [21] [84] have shown that 78%(97%) of spectrum is unutilized in urban (rural) areas. On the other hand, the demands for wireless communication services have increased dramatically. The unbalance between the over-crowded wireless spectrum and the low utilization efficiency on most parts of the spectrum has stimulated the efforts in engineering, economics and regulation communities in searching for better spectrum management policies. Some new ideas have been proposed to provide more flexible and efficient usage of the spectrum. The concept of dynamic spectrum access (DSA) or open spectrum is discussed in [85], which aims to dynamically manage spectrum access and

spectrum sharing by using new technology and standards, in place of the current static band allocation. The IEEE 802.22 working group on Wireless Regional Area Networks (WRAN) has been developing a new standard, focusing on constructing a consistent and point-to-multipoint WRAN that will utilize the free UHF/VHF TV bands for communication [86]. The key enabling technology of the projects mentioned above is the cognitive radio (CR), first presented in the dissertation of Joseph Mitola [82]. Cognitive radio is a paradigm for wireless communications in which either a network or a wireless node changes its transmission or reception parameters to communicate efficiently without interfering with the licensed users. This alteration of parameters is based on active monitoring of several factors in the external and internal radio environment, such as radio frequency spectrum, user behavior and network states. From learning the wireless environment, the cognitive radio terminal will tune to the under-utilized spectrum and make its own transmission without notice to the primary users (PUs).

1.3.2 Dynamic spectrum access

A cognitive radio network has to share resources with the primary network(operators with licensed spectrum), and should not destruct the service of the primary users. Dynamic spectrum access or open spectrum aims to dynamically manage spectrum access and spectrum sharing by using new technology and standards, in place of the current static band allocation [87].

There are two types of spectrum access that allows the secondary devices to use the licensed spectrum: spectrum overlay and spectrum underlay [83]. In spectrum

overlay, an unlicensed device continuously senses the licensed spectrum bands and opportunistically uses those bands that are currently not used by the licensed devices. Once the primary users use the band, the secondary users (SUs) should release the band and return it back to the primary users. This is also called opportunistic access. Spectrum underlay is based on the interference temperature model [21] proposed by the FCC, which will be discussed in more details in the following subsection. The primary and secondary links share the same spectrum and may transmit simultaneously, as long as the interference level experienced by the primary links is below a pre-negotiated threshold.

1.3.3 Interference temperature model

A model called interference temperature has been introduced by the FCC [21]. I_{temp} is a measure of the interference power within a given bandwidth [88], and it is specified in Kelvins and defined as

$$I_{\text{temp}} = P_i/\kappa B, \tag{1.9}$$

where P_i is the interference power in Watts over bandwidth B measured in Hertz, and κ is Bolzmann's constant (1.3810^{-23} J/K). The concept of interference temperature attempts to characterize interference and noise with a simple number. To be more exact, the interference temperature sets a quantity to measure the interference term caused by an interference environment.

The interference temperature limit or threshold I_{th} characterizes the maximum amount of tolerable interference for a given frequency band in a particular location where the primary receiver can operate satisfactorily. Cognitive terminals operating in licensed frequency bands have to adjust their transmission power in such a way

that they avoid raising the interference temperature over the limit at the primary user's receiver. The following must hold at all licensed receivers:

$$I_{\text{temp}} \leq I_{\text{th}}. \tag{1.10}$$

Thus, for each transmission the interference temperature is measured and suitable transmission power and bandwidth B are computed to meet QoS requirements without violating I_{th}. Since the interference is measured at the primary receiver, the secondary transmitter should know the link gain between itself and the primary receiver to adjust its transmission power. In addition, the effects of cognitive radio transmissions should be measured on all possible primary receivers. Therefore, the coexistence of the primary and secondary networks requires effective interaction between the primary receivers and the secondary transmitters [82].

1.3.4 Resource management in cognitive radio networks

Spectrum management

The licensed spectrum can only be used by secondary networks when there is no primary users on the spectrum or the spectrum is under-used. The secondary network capacity depends on the transmission activity of the active primary users. The opportunistic transmission behavior of the secondary networks imposes unique challenges for their coexistence with the primary networks and QoS provisioning of the carried services. New spectrum management functions are required for the secondary networks with critical design challenges including interference avoidance and QoS awareness. For interference avoidance, the CR networks should guarantee that the

normal operations in the primary networks are not disturbed. The secondary network should select an appropriate spectrum band to support its own QoS-aware communications in the dynamic and heterogeneous spectrum environment. To address these challenges, a set of functionalities are required for spectrum management in the CR networks. The spectrum management process consists of four major steps [89]. 1) Spectrum sensing. A secondary user can use only an unused portion of the spectrum. Therefore, a secondary user should monitor the available spectrum bands, capture their information, and then detect spectrum holes. 2) Spectrum decision. Based on the spectrum availability, a secondary user can select the best channel. This allocation not only depends on the spectrum availability, but is also determined by the scheduling within the secondary networks. 3) Spectrum sharing. Multiple secondary users attempting to access the same spectrum should coordinate with one another to prevent collisions. 4)Spectrum mobility. The CR users are regarded as visitors to the spectrum. Hence, if the specific portion of the spectrum in use is required by a primary user, the secondary communications must be switched to another vacant portion of the spectrum.

MAC protocols for cognitive radio networks

There has been some research efforts on designing MAC protocols in CRNs in both industry standardization and academic research projects. From the standardization point of view, the current IEEE 802.22 draft is the first world wide standard related to cognitive radio. Its MAC employs the superframe structure [86], where synchronized distributed sensing, fast sensing using energy detection, and fine sensing using feature detection are used. At the beginning of every superframe, the BS sends a special

preamble and superframe control header (SCH) through each TV channel (up to 3 contiguously) that can be used for communication and is guaranteed to meet the incumbent protection requirements. In addition, IEEE 802.22 is operated in the point-to-multipoint model, which is comparably easier than the cases without the control of the BS.

In the academic field, several MAC protocols for ad hoc cognitive radio networking have been proposed [90] [91] [92]. Most of them assume that there are no hardware constraints on the spectrum sensing ability and assume full spectrum sensing in a particular portion of the spectrum. The Dynamic Open Spectrum Sharing (DOSS) MAC [93] protocol allows nodes to adaptively select an arbitrary spectrum for the incipient communications subject to spectrum availability. In this protocol, after the detection of primary users' presence, three operational channels (a busy tone band, a control band, and a data band) are set up to eliminate hidden terminals, transmit control packets, and transmit data packets. This protocol is suitable for systems where each user has multiple transceivers, one transceiver for each channel. In [94], the AS-MAC protocol is proposed for a CRN to coexist with the GSM cellular system. One of the control channels in the GSM band is used as the secondary common control channel, which facilitates many spectrum sharing functionalities, such as handshakes between the transmitter and receiver, communications with a central entity, or sensing information exchange. The sensing decision under the hardware constraints of the cognitive radio is first considered in [95]. Each secondary user has limited knowledge of the channel availability, so there is no need for all the secondary users to do continual spectrum sensing. At the beginning of each slot, a secondary user with data to transmit chooses a set of channels to sense, then based on the sensing outcome, a set

of channels is chosen to be accessed. Such spectrum sensing and access decisions are made to maximize the throughput of the secondary user while limiting the interference to the primary network by fully exploiting the sensing history and the spectrum occupancy statistics.

Scheduling in cognitive radio networks

Sophisticated scheduling schemes are needed to allocate resource efficiently and fairly among the users in a CRN. Compared to the scheduling in traditional wireless networks, scheduling in a CRN is more complicated due to the opportunistic nature of the networks [96] [97] [98].

Early works focus on applying graph theory to spectrum allocation and traffic scheduling problems. In [99], optimum spectrum allocation is solved for CRNs by constructing an interference graph. However, due to high computational complexity, it only applies for an ad-hoc CRN with a limited number of fixed secondary users. In [100], the unused licensed channels are allocated opportunistically to a set of cognitive base stations so that the percentage of channel usage is maximized. The authors then formulated this problem as a graph-coloring problem and proposed several greedy heuristics for channel allocation. A similar problem to [100] is considered in [101], where a reward function is introduced that is proportional to the coverage areas of the base stations which help the collaboration among secondary users. Again, the problem was studied based on a graph-coloring formulation. Both the works in [100] and [101] are based on the binary interference model, which does not capture the aggregate interference effects when multiple transmissions simultaneously happen

on one channel. The joint spectrum allocation and scheduling in cognitive radio networks is studied in [97] using the proposed novel Multi-Channel Contention Graph (MCCG) to characterize the impact of interference. Based on the MCCG, an optimal algorithm is presented to compute the maximum throughput solutions.

Throughput optimization in cognitive radio networks is one of the major interests of existing literature. In [102], optimal and suboptimal distributed strategies are derived for the secondary users to decide which channels to access with the objective of throughput maximization under a framework of partially observable Markov decision process. In [96], the concept of time-spectrum block allocations is proposed, and a centralized allocation scheme is designed to maximize the system throughput by optimizing the time-spectrum blocks. The optimization problem is NP-complete, and then a heuristic method is used to allocate these blocks at each node.

Other objectives like maximizing spectrum utilization is considered in [103], while minimizing total transmission power is discussed in [104] to schedule the rate requirements of all links. However, the works of [103] and [104] do not consider the scenario of opportunistic spectrum access and the issue of protecting the primary users. In [98] an opportunistic scheduling policy is developed for cognitive radio networks to maximize the throughput utility of the secondary users, subject to maximum collision constraints with the primary users. The technique of Lyapunov optimization is used to design an online flow control, scheduling and resource allocation algorithm that meets the desired objectives and provides explicit performance guarantees. This method has the high computational complexity.

Auction-based schemes are proposed as another direction to solve the spectrum

allocation and traffic scheduling problems. Spectrum band auctions have been proposed in [105] [106] to allocate wireless spectrum resources, in which bidders obtain different spectrum channels free of conflict. Different pricing models are developed to address revenue and fairness, and low-complexity market-clearing algorithms are derived for prices and allocations in real-time.

Maximizing system resource usage efficiency may lead to resource allocation unfairness. The traditional max-min fairness can provide fairness but may not utilize the system resource efficiently [107]. The desired scheduling criteria can provide fairness to users and exploit the multi-user diversity gain at the same time. The concept of proportional fairness (PF) is a good scheduling objective to balance the fairness among users and the resource utilization efficiency [107] [108]. The proportional fairness is in the form of logarithmic utility function $\log(X)$, which is a specific case of Nash bargaining solution [109] [110]. Based on maintaining a balance between maximizing the system profit and providing fair QoS among the users, it tries to assign each user a data rate or a scheduling priority that is inversely proportional to its expected resource consumption. As a real system application, PF scheduling is employed at the base station to schedule the downlink flows among different users for 3G wireless data networks [111].

Admission control for cognitive radio networks

Admission control algorithms for CDMA networks have been reviewed in Section 1.2.2. Most of these algorithms are not suitable for cognitive radio networks due to the interactions between the primary and the secondary users. Some work has been done to address the admission control problems in cognitive radio networks.

In [112], a centralized user removal algorithm based on the tree-pruning technique was proposed. The proposed algorithm leads to an optimal set of the supported secondary links but with extensive computations. To deploy the scheme distributively, a game theoretic approach based on sequential play was introduced, but the approach only converges to local optimal solutions. In [113], a distributed algorithm was introduced that aims at minimizing the total transmit power by the primary and secondary links. In order to maximize the number of admitted secondary links under the QoS and interference constraints, the authors of [114] developed two centralized removal algorithms I-SMIRA and I-SMART(R), which require the knowledge of instantaneous channel gains between the system nodes to be kept at a central controller. In [115], joint admission control and rate/power allocation for secondary links was investigated using mean channel gains information. In [116], a joint power and admission control (JPAC) algorithm for CR networks was introduced. The admission control algorithm removes the SUs recursively while considering the transmission power control of the nodes, so that all remaining SUs can communicate reliably, and the QoS requirement of the PU can be satisfied.

Admission control is important for limiting the traffic load and providing desired QoS, but the complexity inherent in resource management of ad-hoc networks will make the admission control in cognitive ad-hoc network even harder. How to design an effective admission control scheme in a CRN is especially challenging if it should be implemented with low complexity.

1.4 Introduction to Cooperative Networks

1.4.1 Cooperative communications

The mobile wireless channel suffers from fading, meaning that the signal attenuation can vary significantly over different time on different channels. Traditional cellular networks use direct transmissions between the BS and MSs. When the channel condition between the BS and a given MS is poor, very high transmission power may be required from the BS (for the downlink) or from the MS (for the uplink). This not only consumes high energy but also causes high interference to other users sharing the system and reduces the system capacity. Using a relay station to relay the traffic between the BS and the MS can be an option in this situation, so that the transmission power of the MSs (for the uplink) and the BS (for the downlink) can be reduced, or equivalently the coverage area in the cellular systems can be increased. With the help of the relay stations, the interference level in the system can be reduced, which can improve the capacity of cellular networks [62]. The reason behind the exploration of user cooperative communications is the willingness of users to share power and computation with neighboring nodes and this can lead to savings of overall network resources. In particular, spatial diversity is generated by transmitting signals from different locations, thus allowing independently faded versions of the signal at the receiver. If the direct transmission of the traditional cellular system can be assisted with the relay channel, independent copies of the signal will generate diversity and can effectively combat the deleterious effects of fading which neither direct nor relay transmission can solve[117][118].

Relay channel

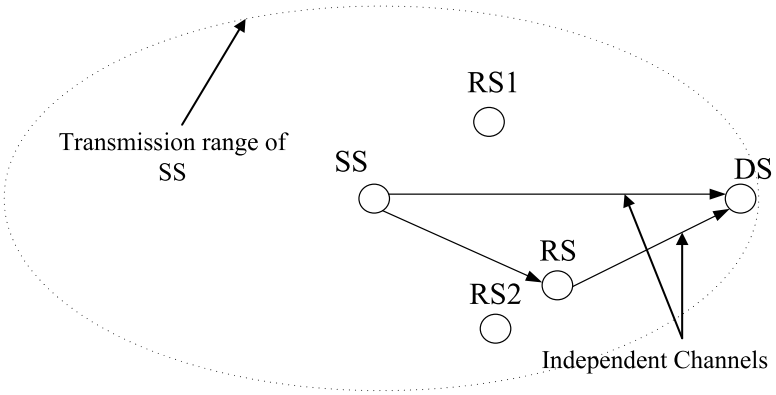

Figure 1.1: Illustration of cooperative communications

As shown in Fig 1.1, three devices are involved in the communications of a cooperative link, the source station (SS), the relay stations (RS) and destination station (DS). Due to the broadcasting nature of the wireless signal transmission, the signal sent from the SS to the DS can be received at the some stations close to the SS. If some of these stations are willing to relay the received signal copies to the DS, the same copy of signal is transmitted through several statistically independent paths, and this generates the spatial diversity.

Relay forward methods

There are different methods for the received data to be processed at the relay station before the data is sent to the destination[73][119][70]. For the Amplify and Forward (AF), the relay simply amplifies its partner's signal and forwards it to the destination. For the Decode and Forward (DF), the relay node should decodes, re-encode, and then forward the partner's signal to the destination.

AF is the traditional method for relaying, where the analogue relay acts like the

signal amplifier. The analog signal received by the relay is amplified and then sent again to the DS. For cooperative communications, the relay channel is not the only channel, it acts like the auxiliary channel to the direct channel. How much the power is amplified is determined not only by the relay channel itself but also the transmission on the direct link. The AF method has the merits of simplicity and no constraints on decoding as the supplement to the direct link. The shortcoming of AF is that the noise and interference in the received signal is amplified as well, which is the main downfall of this method.

Instead of simply amplifying the noise and interference contaminated signal, the RS can first decode the received signal then re-encode it and transmit to the DS. When decoding is successful, the noise and interference is removed. On the other hand, the decoding requires more power consumption than the front end power amplification, and the decoding requirement puts constraints on both the source to relay channel and the transmission power at the source station. Incorrectly decoded signal is meaningless for DF. Additionally, it may incur some security problems and need extra scheme to protect the privacy of relayed packets.

1.4.2 RRM challenges in cooperative communication networks

Both the characteristics of relay channel and the relay forward methods raise new questions in resource management for applying cooperation in wireless networks.

First, multiple relay stations may be available. Selecting a RS is directly related to the performance of the relay-based system. Second, after the relay stations are chosen, a suitable relay forward method should be decided. Third, power allocation

is related to the lifetime of both the mobile stations and the relay stations, so how to optimally distribute the power among mobile stations and relay stations determine the lifetime of the system. In addition, the transmission power allocation affects the interference conditions in the system, and the overall performance such as capacity and QoS.

Relaying traffic for other stations consumes the energy of the RSs. The costed power for relaying other's information will be counteracted when the user is sending its own information with diversity gain. The saved energy could be much more than the energy the user costs to relay other's information. This is the instinct motivation of users' cooperation. Incentive schemes or policies can be designed to further increase the motivation for cooperation from both the user and the network's perspective [120] [121] [122].

Related works

Extensive research has been done to demonstrate that using cooperative communications at the physical layer can greatly improve the transmission rate and reduce transmission errors [72] [119]. However, how to apply the cooperative communication into the existing wireless protocols are still facing many technical difficulties.

First, given the amount of available idle stations, schemes are needed to choose the best relay stations at a limited cost. The best relay is the one that contributes the most to the output SINR. In [117], the authors present ad hoc approaches to approximate the optimal solution with limited complexity. The authors propose two sub-optimal schemes, simple selection and sequential selection. Each destination station chooses the node from the decoding set to which it has the best channel

condition. Each user makes independent decision on the relay selection. This is sub-optimal, especially when many users pick the same relay. In [123], each terminal can act as both source and relay, and a pre-determined cooperator group is assumed. While in [124], several relay selection algorithms are proposed, which search over a set of N candidate relays and optimize the frame error rate.

After selecting the relay stations with given criteria, works need to allocate wireless resource properly for the source stations and the relay stations, including power distribution and time slot allocation. A system level simulation of the performance of a relay assisted cellular network is presented in [125]. The authors consider a hexagonal cell with the BS at one of the vertices, and five relays distributed uniformly in the cell. Two hop communications are considered. The relays repeat the data transmitted by the BS. The average throughput at each point in the cell is found through simulation. In [126], a utility maximization framework is constructed for solving the optimal relaying strategies and power allocation in cellular networks. The system model is a cellular network where users can relay for one another. The optimization problem is solved by decomposing it into multiple smaller subproblems connected hierarchically to one another. While solving one of the subproblems, the authors assume that each node uses only a finite set of modulation schemes, and hence, supports only a discrete set of rates. This enables them to do an exhaustive search to find out the best relay and the relaying strategy. [81] showed that it is possible to find the optimal transmit power and the schedule by exploiting linear networks. The authors provided a general condition under which nodes transmitting in the order of increasing distances to the source is optimal. It suggested a natural strategy for cooperative broadcast in dense wireless networks. That is to let nodes transmit according to their distance to the

source with the minimum transmission power. In [74], the authors explored the effect of partial channel state information on the performance of cooperative communications for delay limited applications. An opportunistic decode-and-forward protocol is proposed where the relay terminal is utilized depending on the overall network state with dynamic power and time allocation.

1.5 Motivation and Overview of the Thesis

Cognitive radio networking can possibly improve the efficiency of radio spectrum utilization. On the other hand, it poses significant challenges to resource allocations. Within the CRN, the amount of available resource varies and is beyond the control of the CRN. The CRN should also interact with the primary network to find available resource and allocate resources accordingly. Although extensive work has been done on channel sensing and other physical layer designs for CRNs, little work is available for QoS provisioning in CRNs. Due to the secondary nature, resource allocations in the CRNs have to be adjusted dynamically in to not violating the QoS in the primary network. In the spectrum underlay scenario, where both the primary and the secondary networks can use the channel at the same time, controlling the transmission power in the CRN is very important to control the interference to the primary network. Meanwhile, the users in the CRN should coordinate among one another so that the available resources can be utilized more efficiently, while not violating the agreement with the primary network. This work is especially challenging in ad hoc CRNs, where the peer users have limited interactions with one another. In Chapters 2-4 of this thesis, we study the resource management in an ad hoc CRN with spectrum underlay, where the CRN share the same spectrum with the uplink of a cellular CDMA

network, which is the primary network. A controller is co-located with the primary BS for measuring the interference level at the BS receiver. Based on the measured interference level, transmission power and rate for the links in the CRN are scheduled accordingly. We first jointly consider the resource allocations in both the primary and the secondary networks in Chapter 2, and study the optimum resource allocations, in terms of both transmission power and rate, for supporting best effort traffic in the CRN, subject to the transmission rate requirement and maximum power limit in the primary network. Fair transmission throughput is provided to the secondary links at each time slot. The effect of setting different interference thresholds on both networks is also studied. In Chapter 3, the focus is switched to supporting long-term best effort traffic in the CRN, and the fairness is relaxed to achieve the long term average throughput for links. In addition to formulating and solving the optimum scheduling problem, practical and heuristic schemes are designed by opportunistically taking advantage of the current available resources. The resource allocation problem is further extended in Chapter 4 for supporting traffic with strict QoS requirements, where admission control and packet transmission scheduling are jointly considered for two types of traffic, streaming traffic with both transmission rate and outage requirements, and non-real-time traffic with a minimum transmission rate requirement. For both types of the traffic, the average throughput of each admitted link is guaranteed by the admission control decisions. The outage performance of the streaming traffic is then guaranteed by a higher priority in the scheduling. The instantaneous transmission power and rate for each link is obtained based on information about the measured interference and link conditions. Fairness is achieved for the outage performance among the streaming links and throughput performance among the nrt

links.

While using cognitive radio networks can enhance the utilization of the under-utilized spectrum band, cooperative communications can improve the network performance by taking advantage of multi-path diversity through relaying. Although extensive work has been done on cooperative communication networks, not much is available for CDMA-based networks, where all transmissions can share the same spectrum. As relaying increases the number of transmitters, managing transmission power of the stations, both the source stations and the relay stations, becomes more complicated yet it is critically important for improving the system performance. In Chapter 5, power allocations are studied for both the uplink and the downlink of a cellular CDMA networks with cooperative relaying. Optimum power distribution problem is formulated and solved for each direction, and heuristic schemes are then proposed.

At last, a summary of the thesis is given and future research topics are discussed in chapter 6.

Chapter 2

Slot-by-slot Optimum Scheduling in CRNs

In a CRN with spectrum underlay, the secondary links can transmit at the same spectrum as the primary links as long as the interference that the secondary links cause to the primary links is below a pre-negotiated interference threshold. This chapter studies the transmission power and rate allocations in such a scenario. Followed by the system description in Section 2.1, an optimization problem is formulated in Section 2.2 with an objective to maximize a function of the secondary link transmission rates, subject to the SINR requirements of both the primary and secondary links and the interference constraints. The effect of setting different interference thresholds on the transmission rate of the secondary links and how the secondary transmissions affect the transmissions of the primary links is studied in Section 2.3. Numerical results in Section 2.4 show some interesting relationship among the interference threshold, transmission rate and power of the secondary links, and transmission power of the primary links.

2.1 System Description

We consider a CDMA-based cellular network as the primary network, where different frequency bands are used in the uplink and the downlink. Instead of communicating with the base station (BS) directly in the primary network, some mobile stations (MSs) near one another may form an ad hoc CRN and communicate directly with one another. For example, the MSs located in the same playground or conference center prefer to communicate directly with other MSs to save the communication cost and battery power. The secondary transmissions share the same spectrum as the uplink of the cellular network through spectrum underlay, and cause interference to the uplink transmissions in the primary network. For preliminary study, we consider only single hop transmissions in the CRN. As the BS is the receiver for all the uplink transmissions in the primary network, a controller for the CRN is co-located with the primary BS for measuring the interference level at the BS caused by the secondary transmissions. The controller can be a device independent of the primary BS. Alternatively, the primary and secondary networks can be tightly coupled in sometimes, and the primary BS can provide some information to assist the secondary network operation. In this case, the primary BS can also work as the controller for the secondary network.

Several ways can be used for the controller and the CR nodes to exchange information. If the primary and secondary networks are tightly coupled, the CRN can share the same control channels with the primary network. The primary BS can monitor the secondary-to-primary interference and centrally control the CRN as in [108] and [111]. This is not a problem if the primary network has relatively light traffic load, which is most likely the case when a secondary network is allowed and I_{th} is set

to be reasonably high, then the control channel in the primary network has low traffic load and can be well used for the secondary network as well. An alternative way to provide the common control channel in the CRN is that the CRN can lease several mini-slots in both the uplink and downlink in the primary network for transmitting control signals. The mini-slots in the uplink channel are used for the secondary devices to report the link and interference conditions to the controller, and the mini-slots in the downlink are used for the controller to broadcast information related to admission control and packet transmission scheduling to the secondary devices. In addition, the CRN can seek out-of-band control channel as done in [127], and the control channel can be in the license-free band.

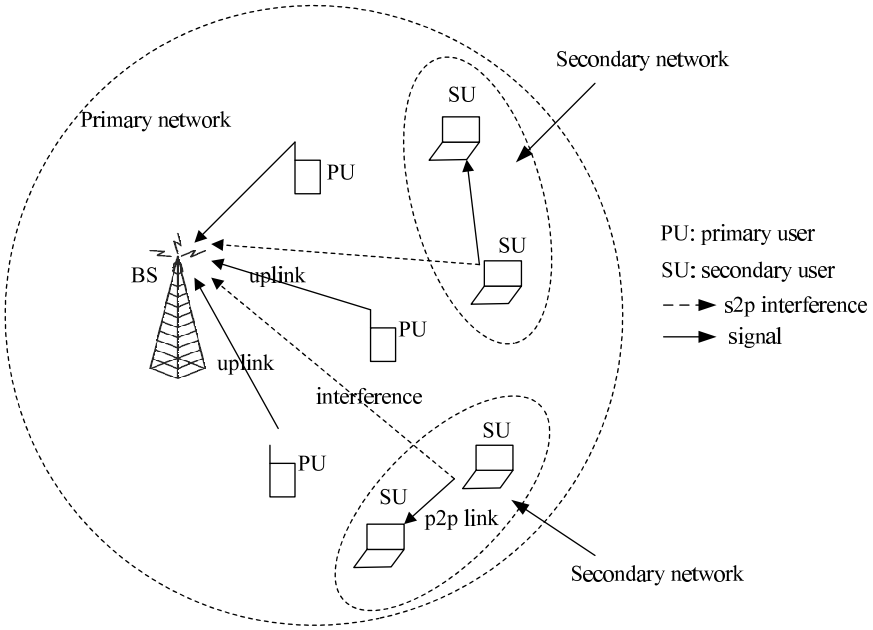

Figure 2.1: Illustration of primary and secondary networks

Since only the uplink is considered for the primary network, the transmitters are the MSs, and they share the same receiver, which is the BS. The transmission power of a primary link is denoted as $P_{p,i}$, for $i = 1, 2, \ldots, M_p$, and that of a secondary

link is denoted as $P_{s,i}$, for $i = 1, 2, \ldots, M_s$, where M_p and M_s, respectively, are the number of primary and secondary links. The transmission rate of the ith primary link is $R_{p,i}$, and that of a secondary link is $R_{s,i}$. The maximum transmission power of the transmitter is limited by $P_{p,\max}$ for a primary link and $P_{s,\max}$ for a secondary link. Let $g_{p2p,ij}$ denote the link gain from the transmitter of the ith primary link to the receiver of the jth primary link, $g_{p2s,ij}$ the link gain from the transmitter of the ith primary link to the receiver of the jth secondary link, $g_{s2p,ij}$ the link gain from the transmitter of the ith secondary link to the receiver of the jth primary link, and $g_{s2s,ij}$ the link gain from the transmitter of the ith secondary link to the receiver of the jth secondary link. Each link has a strict signal-to-interference-plus-noise ratio (SINR) requirement at the receiver, which should be above γ_p^* for the primary links and γ_s^* for the secondary links after the signal is despreaded.

There are two interference models for measuring the interference level at the primary receivers. The first is to monitor the noise and total interference from all the secondary and the primary transmitters. This measured interference is then compared with the interference threshold, and the result is used to regulate the secondary transmissions. In this way, the interference at the primary receiver caused by the secondary transmissions is not detected separately. This model does not require a priori knowledge of the RF environment, and consequently does not need to distinguish the licensed signals from the interference and noise. The second model requires that the aggregate signal strength coming from the secondary transmitters is measured at the receiver of a primary link and compared with the interference threshold. In this case, interference caused by the secondary transmitters should be separated from that caused by primary transmitters in order to calculate the interference level. Below we

formulate the power and rate allocation problem in the primary-secondary scenario based on these two interference models.

2.2 Optimum Rate and Power Allocation

We first consider the SINR requirement of each primary and secondary link. For a primary link, its required SINR can be satisfied if

$$\frac{G_{p,i}P_{p,i}g_{p2p,ii}}{\sum_{j\neq i}P_{p,j}g_{p2p,ji}+\sum_{j=1}^{N_s}P_{s,j}g_{s2p,ji}+\eta}\geq\gamma_p^*, \tag{2.1}$$

where $G_{p,i}=W/R_{p,i}$ is the processing gain of the ith primary link. In the denominator on the left-hand side of (2.1), the first term is the interference from all other primary links, the second term is the interference from all the secondary links, and η is the background additive white Gaussian noise (AWGN) power.

The SINR of the ith secondary link can be satisfied if

$$\frac{WP_{s,i}g_{s2s,ii}}{R_{s,i}\left(\sum_{j\neq i}P_{s,j}g_{s2s,ji}+\sum_{j=1}^{N_p}P_{p,j}g_{p2s,ji}+\eta\right)}\geq\gamma_s^*. \tag{2.2}$$

Between the brackets in the denominator on the left-hand side of (2.2), the first term is the interference from all other secondary links, and the second term is the interference from all the primary links.

The transmission power of the secondary links is also limited by the interference threshold

$$\sum_{j\neq i}P_{p,j}g_{p2p,ji}+\sum_{j=1}^{N_s}P_{s,j}g_{s2p,ji}+\eta\leq I_{\text{th}}, \tag{2.3}$$

for the first interference model, and

$$\sum_{j=1}^{N_s} P_{s,j} g_{s2p,ji} \leq I_{\text{th}}, \tag{2.4}$$

for the second interference model.

Based on the above conditions we formulate the following optimization problem, which maximizes a function (normally convex) of the secondary link transmission rates, subject to the SINR requirements of the primary and secondary links and the given interference threshold, where $\mathbf{R}_s = (R_{s,1}, R_{s,2}, \ldots, R_{s,N_s})$. In the optimization problem, the constraint is based on the first interference model. If the second interference model is used, the constraint (2.8) should be changed to (2.4).

$$\max f(\mathbf{R}_s) \tag{2.5}$$

$$\text{s.t. } \frac{W P_{s,i} g_{s2s,ii}}{R_{s,i}\left(\sum_{j\neq i} P_{s,j} g_{s2s,ji} + \sum_{j=1}^{N_p} P_{p,j} g_{p2s,ji} + \eta\right)} \geq \gamma_s^*, \quad i = 1, 2, \ldots, M_s \tag{2.6}$$

$$\frac{G_{p,i} P_{p,i} g_{p2p,ii}}{\sum_{j\neq i} P_{p,j} g_{p2p,ji} + \sum_{j=1}^{N_s} P_{s,j} g_{s2p,ji} + \eta} \geq \gamma_p^*, \quad i = 1, 2, \ldots, M_p \tag{2.7}$$

$$\sum_{j\neq i} P_{p,j} g_{p2p,ji} + \sum_{j=1}^{N_s} P_{s,j} g_{s2p,ji} + \eta \leq I_{\text{th}}, \quad i = 1, 2, \ldots, M_p \tag{2.8}$$

$$0 \leq P_{p,i} \leq P_{p,\max},\ i = 1, 2, \ldots, M_p, \tag{2.9}$$

$$0 \leq P_{s,i} \leq P_{s,\max},\ i = 1, 2, \ldots, M_s. \tag{2.10}$$

We consider two objective functions based on different rate allocation criteria. First, a simple equal rate allocation (ERA) is considered. That is, all secondary links transmit at the same rate. In this case, $R_{s,i} = R_s$ for all $i = 1, 2, \ldots, N_s$, and

the objective of the optimization is to maximize R_s. The ERA formulation provides perfect rate fairness among all the secondary links. However, links with poor SINR conditions transmit higher power to achieve the same rate as other links. Therefore, the link with the worst SINR condition limits the transmission rate of all links in the secondary network. The second rate allocation criterion is based on the proportional fairness (PF) for the rate allocations [107] and the objective function is $\sum_{i=1}^{N_s} \log(R_{s,i})$.

The PF is usually used as compared to the traditional max-min fairness which can provide fairness but can negatively affect the transmission throughput [107]. The desired scheduling criterion is to provide fairness and exploit the multi-user diversity gain at the same time. The concept of PF is a good scheduling objective to balance the fairness among users and the resource utilization efficiency [107] [108]. The PF is in the form of a logarithmic utility function $\log(X)$, which is a specific case of Nash bargaining solution [109] [110]. Based on maintaining a balance between system profit maximization and user throughput fairness, it tries to assign each data flow a data rate or a scheduling priority that is inversely proportional to its anticipated resource consumption.

It can be found that the above optimization problem is a geometric programming (GP) problem (Appendix A) which is non-linear and non-convex due to constraints (2.6) and (2.7). By observing the two constraints we find that the left-hand side of each of the inequalities is the division of a monomial to a posynomial and is usually named as an inverted posynomial [128]. Lower bounding an inverted posynomial is allowed in GP since it is equivalent to upper bounding a posynomial, thus the problem can be easily transformed into a standard form GP problem. By applying the convex form transformation described in Appendix A, this problem can be

converted to a convex optimization problem and solved efficiently.

2.3 Interference Threshold

A higher interference threshold allows higher transmission power from the secondary links and can potentially increase the transmission rate of the secondary links. On the other hand, as the secondary links increase their transmission power, they cause more interference to the primary links. As a result, transmission power of the primary links should also be increased. The mutual interference effect eventually reaches a balance, then neither the primary nor the secondary links can increase the transmission power. When I_{th} and $P_{s,\max}$ are both large enough, the secondary-to-primary interference is maximized when at least one MS in the primary network reaches $P_{p,\max}$.

Consider that homogeneous traffic is carried out by the primary links. Then $G_{p,i} = G_p$ for all $i = 1, 2, \ldots, M_p$. Let $I_{p,i}$ represent the aggregate noise and interference that the ith primary link experiences from all other primary links and all secondary transmitters. With perfect power control, the actual SINR for the primary link at the BS receiver input is equal to γ_p^*, and all the primary links have an equal received power, $P_{p,r}$, at the BS [129]. In this case, we have $G_p P_{p,r}/I_{p,i} = \gamma_p^*$, or $I_{p,i} = G_p P_{p,r}/\gamma_p^*$. As the transmission power is limited by $P_{p,\max}$, we have

$$P_{p,i} = \frac{P_{p,r}}{g_{p2p,ii}} \leq P_{p,\max}. \tag{2.11}$$

Then

$$I_{p,i} \leq \frac{G_p P_{p,\max} g_{p2p,ii}}{\gamma_p^*}, \tag{2.12}$$

for all $i = 1, 2, \ldots, M_p$. Let

$$I_{p,\max} = \min_i I_{p,i} = \min_i \frac{G_p P_{p,\max} g_{p2p,ii}}{\gamma_p^*}. \tag{2.13}$$

Then $I_{p,\max}$ is the upper bound of a meaningful interference threshold in the sense that if $I_{\text{th}} \leq I_{p,\max}$, the transmission power of the second links is limited by I_{th} to satisfy the SINR requirements of the primary links. When $I_{\text{th}} > I_{p,\max}$, transmission power of the secondary links may not be limited by I_{th}, since the SINR requirement at the primary network limits the transmission power of the secondary links. When $I_{\text{th}} \leq I_{p,\max}$, the SINR of the primary links can be protected (within the maximum transmission power limit of the MSs) as long as the actual secondary-to-primary interference is below the threshold.

Since $g_{p2p,ii}$'s are random, below we find the distribution of $I_{p,\max}$ and then its mean value. Given any value $y > 0$,

$$\Pr\{I_{p,\max} \leq y\} = \Pr\left\{\min_i \frac{G_p P_{p,\max} g_{p2p,ii}}{\gamma_p^*} \leq y\right\} \tag{2.14}$$

$$= \Pr\left\{\min_i g_{p2p,ii} \leq \frac{y\gamma_p^*}{G_p P_{p,\max}}\right\} \tag{2.15}$$

$$= 1 - \prod_i \Pr\left\{g_{p2p,ii} > \frac{y\gamma_p^*}{G_p P_{p,\max}}\right\} \tag{2.16}$$

$$= 1 - \left[1 - \Pr\left\{g_{p2p,ii} \leq \frac{y\gamma_p^*}{G_p P_{p,\max}}\right\}\right]^{M_p}. \tag{2.17}$$

In the above derivation we assume that all $g_{p2p,ii}$'s are independent and identically distributed. The mean of $I_{p,\max}$ can be found as

$$\mathrm{E}[I_{p,\max}] = \int_0^\infty (1 - \Pr\{I_{p,\max} \leq y\}) dy \tag{2.18}$$

When finding $\Pr\left\{g_{p2p,ii} \leq \frac{y\gamma_p^*}{G_p P_{p,\max}}\right\}$, we consider that the transmitted signals suffer from both path loss and log-normal fading. Then $g_{p2p,ii} = Ad_i^{-\alpha}e^{-\beta X}$, where A is the received power at a reference distance, d_i is the distance (normalized to the reference distance) between the transmitter and the receiver of the ith link, α is the path loss exponent, $\beta = \log 10/10$, and X is a normally distributed random variable with zero mean and standard deviation of σ. The distribution of $g_{p2p,ii}$ can be found as

$$\begin{aligned} & \Pr\left\{g_{p2p,ii} \leq \frac{y\gamma_p^*}{G_p P_{p,\max}}\right\} \\ = \; & \Pr\left\{Ad_i^{-\alpha}e^{-\beta X} \leq \frac{y\gamma_p^*}{G_p P_{p,\max}}\right\} \\ = \; & \Pr\left\{-\alpha \ln d_i - \beta X \leq \ln \frac{y\gamma_p^*}{AG_p P_{p,\max}}\right\}. \end{aligned} \tag{2.19}$$

When all MSs are uniformly distributed in a circular area of radius of D, the probability density function (pdf) of d is $f_d(z) = 2z/D^2$ for $0 \leq z \leq D$. Then

$$\begin{aligned} & \Pr\left\{g_{p2p,ii} \leq \frac{y\gamma_p^*}{G_p P_{p,\max}}\right\} \\ & = \int_0^D f_d(z)dz \int_{-\frac{\alpha}{\beta}\ln z - \frac{1}{\beta}\ln\frac{y\gamma_p^*}{AG_p P_{p,\max}}}^{\infty} \mathcal{N}_x(0,\sigma^2)dx, \end{aligned} \tag{2.20}$$

where $\mathcal{N}_x(\mu,\sigma^2)$ is the pdf of a normal distribution with mean of μ and standard deviation of σ. This integration can be done numerically.

Table 2.1: Default Simulation Parameters

Parameter	Value
Transmission rate of each primary link R_p	64kbps
SINR threshold γ_p^*, γ_s^*	5dB
Path loss exponent α	4
Standard deviation of log-normal fading σ	4dB
Maximum MS transmission power $P_{p,\max}$ and $P_{s,\max}$	0.5W
Spread spectrum bandwidth W	5MHz
Background AWGN noise power η	10^{-10}W
Reference distance d_0	10m
Radius of BS coverage D	1000m
Ad hoc transmission range of MSs	300m

2.4 Numerical Results

We consider the same system as described in Section 2.1. Default simulation parameters are listed in Table 2.1. We first show the results based on the first interference model, then compare the results based on the two interference models.

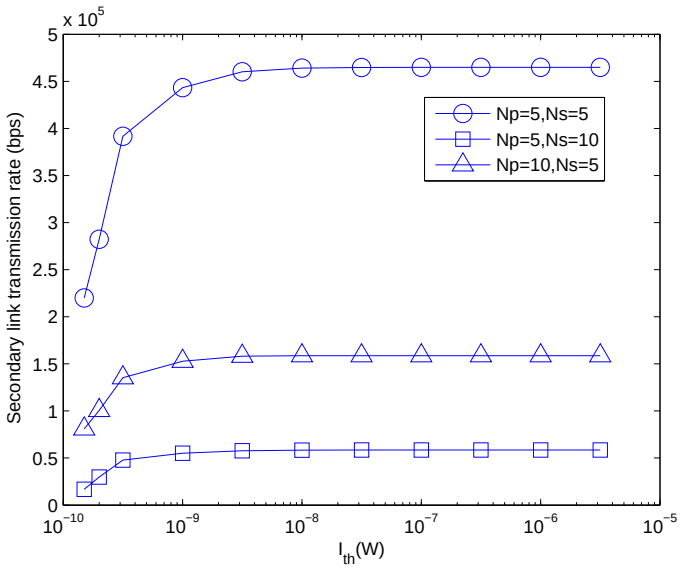

Figure 2.2: Secondary link transmission rate: ERA

Figs. 2.2 and 2.3 show that when the interference threshold is below a certain value, the secondary link transmission rate increases with the interference threshold for

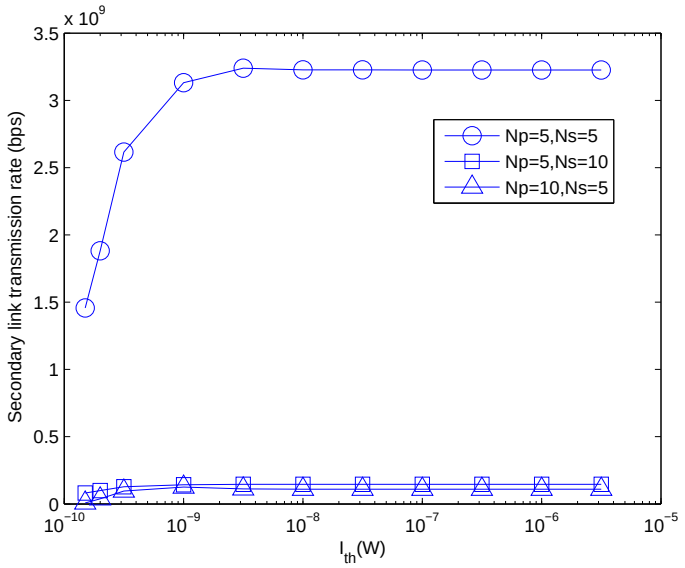

Figure 2.3: Secondary link transmission rate: PRA

both ERA and PRA. Beyond this range, further increasing the interference threshold does not affect the secondary transmission rate anymore, since the transmission of the secondary links is limited by the primary link's SINR constraint and the mutual interference between primary and secondary networks. When the secondary users increase their transmission power to achieve higher rates, the increased power generates higher interference to the primary networks. To sustain the SINR, the primary links will increase their power too. Once the maximized interference limit at the primary receiver is reached, the secondary users cannot further increase their transmission rate even if the interference threshold is not reached.

It is also seen from both Figs. 2.2 and 2.3 that increasing the number of the primary or secondary links results in lower secondary link rate due to that more links are competing for the network resources. Comparing the two figures we can find that using PRA can achieve a lot higher transmission rate for the secondary links than

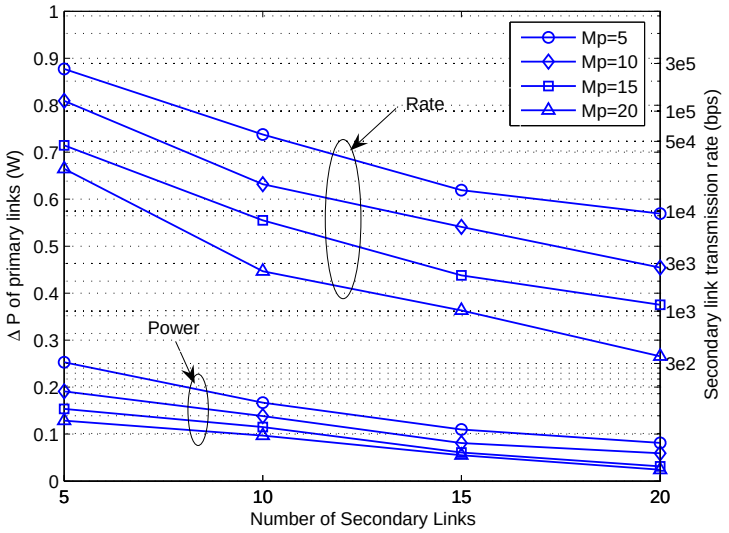

Figure 2.4: ΔP of primary links: ERA

using ERA, since the former can take better advantage of good channel conditions of individual links.

Compared to the case without the secondary links, transmission power of the primary links is increased due to extra interference from the secondary links. We define ΔP as the difference of the average transmission power of the primary links with and without the secondary transmissions. Fig. 2.4 shows that with the increase of the number of secondary links, the average primary transmission power of the primary links decreases. This is explained by the dramatic decrease of the secondary link transmission rate as the number of secondary links increases. Because of the equal rate allocation, the transmission rate of the secondary links is limited by the link with the worst link condition. In addition, mutual interference among the secondary links also reduces their transmission rate. As the transmission rate decreases, the secondary transmission power is reduced, which causes less interference to the primary links.

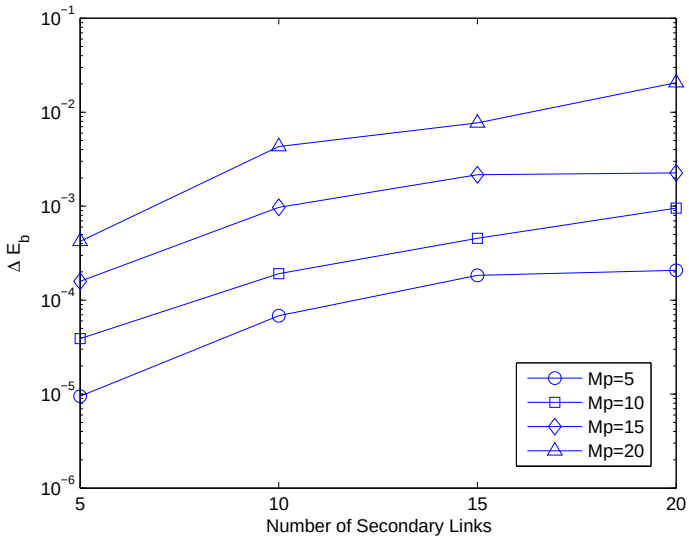

Figure 2.5: ΔE_b of primary links: ERA

Therefore, the primary link transmission power is reduced. Dividing ΔP by the total secondary link transmission rate, we have ΔE_b, which is the average amount of energy increase to the primary link caused by each transmitted bit in the secondary links, and this is shown in Fig. 2.5, which shows that it costs more energy from the primary transmitters on average to support every bit in the secondary network as the number of secondary links increases.

ERA does not allow links with better channel conditions to transmit at higher transmission rate. Figs. 2.6 and 2.7 show that both ΔP and ΔE_b of the primary links decrease with the number of secondary links for PRA. However, comparing Fig. 2.7 with Fig. 2.5 we find that ΔP in PRA is about four orders of magnitudes smaller than that in ERA. Equivalently, with the same amount of ΔP, the CRN based on PRA can achieve much higher transmission rate than that based on ERA.

Figs. 2.8-2.10 show $I_{p,\max}$ versus different system parameters. The actual average

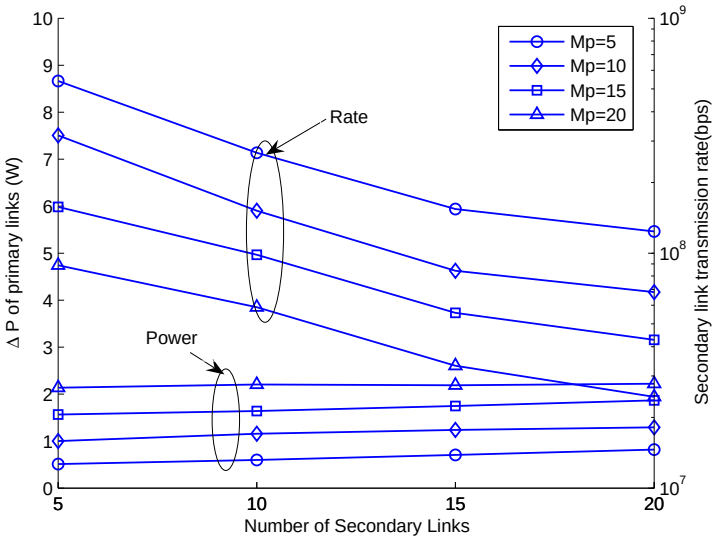

Figure 2.6: ΔP of primary links: PRA

interference at the primary link is also shown for ERA and PRA, respectively. It is seen that a primary network with a smaller cell size, larger $P_{p,\max}$, or larger fading deviation may tolerate a larger interference threshold for the same number of secondary links. This relationship can be easily seen from (2.13). Most importantly, the figures show that using PRA can cause significantly lower interference to the primary links than using ERA. With ERA, the link with the worst SINR condition may transmit much higher power than other links. When this link is close to the primary link receiver (the BS), the primary link transmitters should transmit very high power to combat the high interference. Therefore, in ERA, it is more likely that the primary link transmitters reach the maximum transmission power limit. This also explains that the interference level at the BS using ERA is almost the same as $I_{p,\max}$. In contrast, PRA does not encourage the secondary links with poor SINR to transmit as high rate as the links with good SINR, and therefore the interference level at the

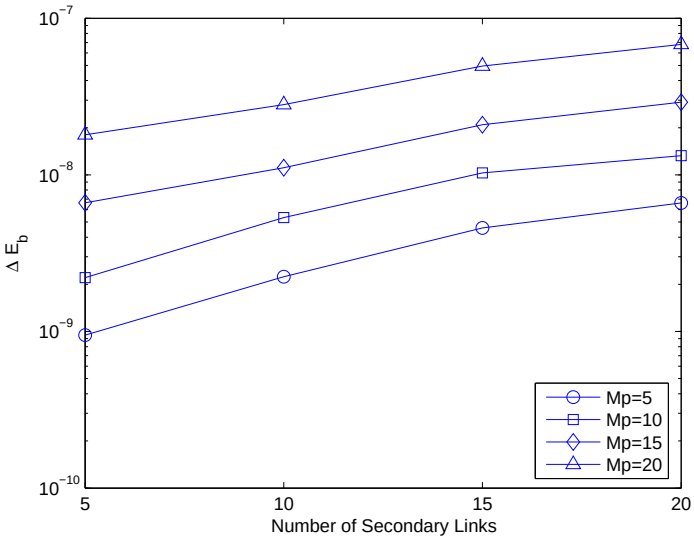

Figure 2.7: ΔE_b of primary links: PRA

primary links is limited by the maximum transmission power of both the primary and secondary transmitters.

Transmission rate of the secondary links is shown in Fig. 2.11 based on the two interference models defined in (2.3) and (2.4). The difference is more obvious when I_{th} is small and close to the noise power level. When I_{th} is relatively large, the two interference models do not make obvious difference to the secondary transmission rate. When I_{th} is small, using the second interference model achieves much higher transmission rate than using the first interference model, since in the latter case, noise and interference from the primary network can dominate the interference threshold, while in first interference model, the secondary transmissions can take advantage of all the interference allowed by I_{th}. As I_{th} increases, the secondary transmission power increases and eventually is limited by their mutual interference and SINR constraints, but not by I_{th}. At this point, the two interference models result in about the same

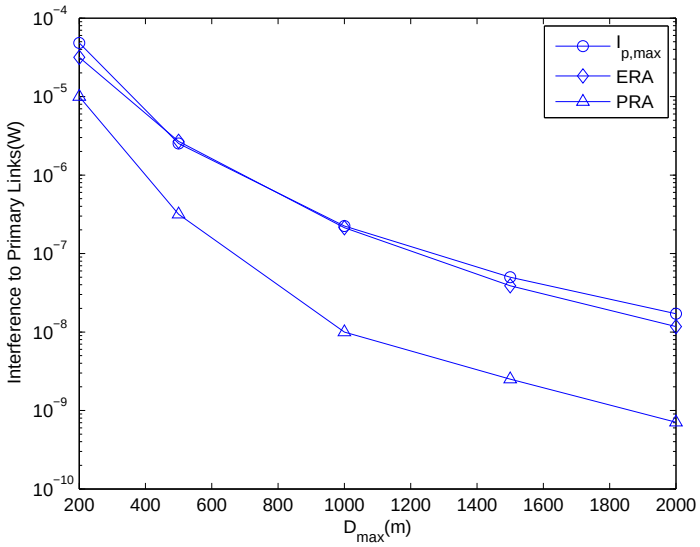

Figure 2.8: Maximum interference vs. cell size

transmission rate at the secondary network.

2.5 Summary

In this chapter, we have studied the optimum power and rate allocations in a CDMA-based cognitive wireless network with spectrum underlay. Our results indicate that there is a limit on the interference threshold beyond which the secondary link transmission rate cannot be increased by increasing the interference threshold. Furthermore, the increase of the secondary user transmission rate will consume extra power from the primary user, and the same amount of power increase from the primary users can support higher rate of the secondary links using proportional rate allocation, compared to using equal rate allocation among the secondary links.

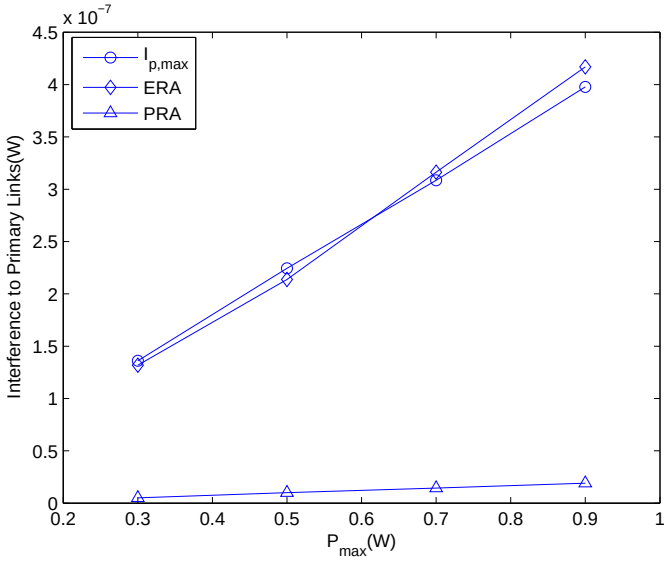

Figure 2.9: Maximum interference vs. $P_{p,\max}$

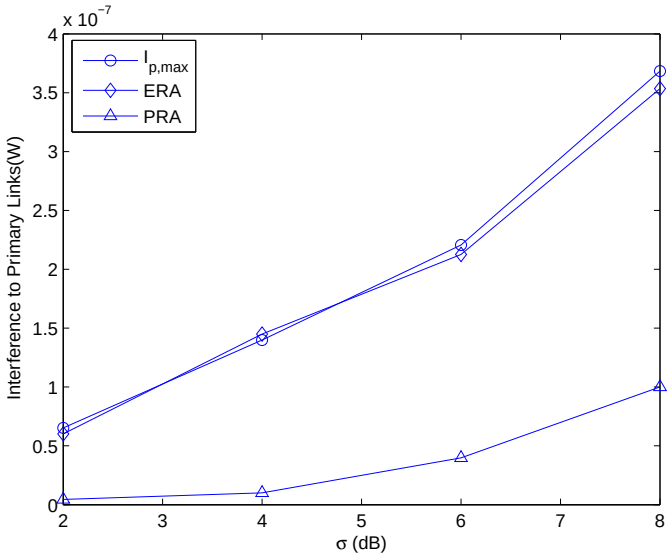

Figure 2.10: Maximum interference vs. σ

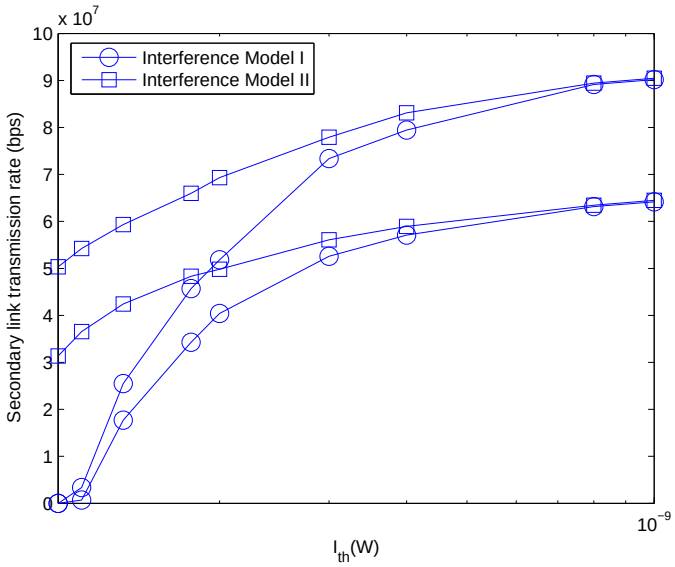

Figure 2.11: Comparison of two interference models

Chapter 3

Scheduling With Long-term Proportional Fairness in CRNs

3.1 Introduction

In the previous chapter, the transmission power and rate allocations for PRA achieve proportional fairness (PF) in every time slot. On the other hand, most traffic in ad hoc networks is elastic, and minor delay is expected and tolerable in packet transmissions. Therefore, it is acceptable that PF is achieved over multiple time slots, or even a longer time. In this way, a user having poor channel conditions can delay its transmissions so that other users can take advantage of their better channel conditions and increase their transmission throughput, provided PF is kept over a period of time. Achieving PF in an ad hoc cognitive wireless network is not trivial due to uncertainty of available radio resources and lack of a central controller. In this chapter we study long term PF scheduling in ad hoc cognitive wireless networks. We first formulate an

optimal scheduling problem with an objective to achieve PF of the long-term average transmission rates among different links, subject to the interference constraints within the cognitive network and imposed by the primary network. Implementing the optimum scheduling is unlikely in a practical system due to the difficulty to collect all required information about link and interference conditions and high complexity to solve the optimization problem. We then propose two heuristic scheduling schemes. In the first scheme, transmission priorities of the links are determined by their potential contributions to a utility function, assuming there is no co-channel interference within the network. In the second scheme, the transmission priorities are derived from both the objective utility function and interference to the primary network. We also consider using exclusive regions to limit interference among simultaneous transmissions in order to improve the system throughput without compromising the fairness. The scheduling schemes can be implemented distributively in the ad hoc cognitive wireless network with limited assistance from the primary network. The remainder of the chapter is organized as follows. In Section 3.2 we introduce some background on PF scheduling. In Section 3.3 we formulate an optimum PF scheduling problem based on the long term average rates. Two scheduling schemes are proposed in Sections 3.4 and 3.5, respectively. Numerical results are demonstrated in Section 3.6, followed by concluding remarks in Section 3.7.

3.2 PF Scheduling for Long Term Average Rates

Consider a network with N links. Channel time is divided into discrete time slots. We define the transmission rate of link i at time slot t as $R_i(t)$, and its long term average rate as $\overline{R}_i$. We further define two vectors, an instantaneous rate vector

$\mathbf{R}(t) = [R_1(t), R_2(t), \ldots, R_N(t)]$ and an average rate vector $\overline{\mathbf{R}} = [\overline{R}_1, \overline{R}_2, \ldots, \overline{R}_N]$. A general scheduling problem based on the long term average rates of links can be formulated to maximize a concave utility function[1] $U(\overline{\mathbf{R}})$. It has been proved in [130] that this objective can be achieved iteratively through slot-by-slot scheduling. At each time slot t, the iterative scheduling is to find a rate vector $\mathbf{r}(t)$ such that

$$\mathbf{r}(t) = \arg\max_{\mathbf{R}(t)} \nabla U\left(\overline{\mathbf{X}}(t)\right) \cdot \mathbf{R}(t), \tag{3.1}$$

where $\nabla U(\cdot)$ is the gradient of function $U(\cdot)$, $\overline{\mathbf{X}}(t) = [\overline{X}_1(t), \overline{X}_2(t), \ldots, \overline{X}_N(t)]$, and $\overline{X}_i(t)$ is an exponentially smoothed average rate for link i up to time t and defined as

$$\overline{X}_i(t) = (1 - 1/t_c)\overline{X}_i(t-1) + 1/t_c R_i(t-1), \tag{3.2}$$

where $t_c > 1$ is a parameter which balances the weights of the past and current transmission rates. For a special case, we have $\overline{X}_i(0) = R_i(0) = 0$ for all $i = 1, 2, \ldots, N$. In [130] it has been proved that when $t \to \infty$, both $\overline{\mathbf{X}}(t)$ and $\overline{\mathbf{R}}$ converge to the optimum solution of the average rate vector. The derivation is given below.

Utility-based objective functions can consider both the system throughput and user's fairness, and are widely accepted in designing fair scheduling strategies, among which PF scheduling is one of the most popularly used. In [107], PF is proved to be equivalent to maximizing a utility function $U(\mathbf{z}) = \sum_{i=1}^{N} \log z_i$, where z_i is the performance metric of link i.

[1] A utility or revenue function is normally increasing and concave while a cost function is usually increasing and convex in order to reflect the users' preference.

When the utility function is given by

$$U\left(\overline{\mathbf{X}}(t)\right) = \sum_{i=1}^{N} \log \overline{X}_i(t). \tag{3.3}$$

PF is achieved for the rates over a long term, and the scheduling is referred to as long term PF scheduling. Substituting (3.3) into (3.1) we can find that

$$\mathbf{r}(t) = \arg\max_{\mathbf{R}(t)} \sum_{i=1}^{N} \frac{R_i(t)}{\overline{X}_i(t)}, \tag{3.4}$$

which gives the optimum transmission rate vector at time slot t in order to achieve long term PF scheduling iteratively.

Achieving the optimum scheduling based on (3.4) normally requires a central scheduler, or multiple links should coordinate with one another — both are difficult in an ad hoc wireless network. In the scheduling scheme proposed in [107], only the transmitter with the largest $\frac{R_i(t)}{\overline{X}_i(t)}$ is allowed to transmit in every time slot. This is referred to as single user scheduling (SUS). By limiting the number of simultaneous transmitting users to one, SUS tends to waste system resources. In the remaining part of this chapter, we study multi-user scheduling (MUS) which allows multiple simultaneous transmissions in order to more efficiently utilize the available network resources while keeping long-term PF among different links. We first formulate an optimum PF scheduling problem and then propose two heuristic MUS schemes.

3.3 Problem Formulation

We consider the same system as described in Section 2.1, where the secondary transmissions opportunistically use the spectrum of the cellular network uplink through spectrum underlay, and interfere with the uplink transmissions in the cellular network. Since all the uplink transmissions in a cellular cell share the same receiver at the BS, the interference level experienced by the primary links can be measured at the BS. We use I_{th} to denote the maximum tolerable interference level that the BS can accept from the secondary transmissions and assume that the BS is able to distinguish the interference caused by the secondary transmissions from that caused by the primary transmissions. This is the second interference model described in Chapter 2, which allows us to focus our attention to the CRN since the interference threshold limits the transmissions in the CRN but not in the primary network. The BS (or the CRN controller) compares measured interference level with I_{th}, and broadcasts information to the secondary network to assist power allocation and transmission scheduling in the network.

We consider point-to-point links, i.e., single hop transmissions only in the CRN. The one-hop transmission range in the secondary network is usually much shorter than the cellular cell size. Channel time is divided into equal size time slots, each of which is for transmitting one data packet and an acknowledgement frame (ACK). The transmission power of an MS is limited by both $P_{\max}$ and I_{th}, where $P_{\max}$ is the maximum transmission power of the MS. We use transmitter i and receiver i, respectively, to represent the transmitter and receiver of link i, where $i = 1, 2, \ldots, N$, and N is the total number of links. Let $R_i(t)$ and $P_i(t)$, respectively, denote the transmission rate of link i and the transmission power of its transmitter at time slot

t. Each link requires a minimum signal-to-interference-plus-noise ratio (SINR), γ^*. Then

$$\frac{W}{R_i(t)} \frac{P_i(t) g_{s2s,ii}(t)}{I_{p2s,i}(t) + I_{s2s,i}(t) + \eta} \geq \gamma^*, \tag{3.5}$$

where W is the spread spectrum bandwidth, $g_{s2s,ii}(t)$ is the link gain from transmitter i to receiver i at time slot t, η is the power of the background additive white Gaussian noise (AWGN), $I_{p2s,i}(t)$ is the total interference at receiver i caused by all primary transmissions at time slot t, and $I_{s2s,i}(t)$ is the total interference at receiver i caused by other transmissions in the same network and given by

$$I_{s2s,i}(t) = \sum_{j=1, j\neq i}^{N} P_j(t) g_{s2s,ji}(t), \tag{3.6}$$

where $g_{s2s,ji}(t)$ is the link gain from transmitter j to receiver i.

By keeping only the equality in (3.5) we find the maximum transmission rate of link i as

$$R_i(t) = \frac{W}{\gamma^*} \frac{P_i(t) g_{s2s,ii}(t)}{I_{p2s,i}(t) + I_{s2s,i}(t) + \eta}. \tag{3.7}$$

The total interference that all secondary transmissions cause at the primary BS should not exceed the predefined threshold, i.e.,

$$\sum_{i=1}^{N} P_i(t) g_{s2p,i}(t) \leq I_{\text{th}}, \tag{3.8}$$

where $g_{s2p,i}(t)$ is the link gain from transmitter i to the BS of the primary network.

From the description in Section 3.2, the optimum PF scheduling for long term average rates can be achieved through an iterative process. At time slot t the scheduling

is to find the rate vector $\mathbf{R}(t)$ and power vector $\mathbf{P}(t) = [P_1(t), P_2(t), \ldots, P_N(t)]$ in order to maximize the summation in (3.4), subject to the constraints given in (3.6)-(3.8) as well as the MS maximum transmission power limit. We can then find the optimum transmission rate and power vectors at time t from the following optimization problem:

$$\max_{\mathbf{R}(t),\mathbf{P}(t)} \sum_{i=1}^{N} \frac{R_i(t)}{\overline{X}_i(t)} \tag{3.9}$$

$$\text{s.t.} \quad R_i(t) = \frac{W}{\gamma^*} \frac{P_i(t) g_{s2s,ii}(t)}{I_{p2s,i}(t) + I_{s2s,i}(t) + \eta}, \ i = 1, 2, \ldots, N \tag{3.10}$$

$$I_{s2s,i}(t) = \sum_{j=1, j\neq i}^{N} P_j(t) g_{s2s,ji}(t), \ i = 1, 2, \ldots, N \tag{3.11}$$

$$\textstyle\sum_{i=1}^{N} P_i(t) g_{s2p,i}(t) \leq I_{\text{th}} \tag{3.12}$$

$$0 \leq P_i(t) \leq P_{\max}, \ i = 1, 2, \ldots, N. \tag{3.13}$$

Comments:

- The iterative scheduling based on the solution to the above optimization problem is referred to as OPT-A, where "A" stands for average rate.
- If we change the objective function in (3.9) to $\max_{i=1}^{N} R_i(t)$, the solution results in the highest total transmission rate in every time slot. This is the max-rate scheduling.
- A max-min rate scheduling can also be obtained at each time slot by solving the same optimization problem but changing the objective function to $\max \min_{i=1}^{N} R_i(t)$.

The above optimization problem is a nonlinear and nonconvex problem and cannot

be efficiently solved. We rewrite the optimization problem so that it fits the format of the standard GP problem.

$$\min_{\mathbf{R}(t),\mathbf{P}(t)} \frac{1}{\sum_{i=1}^{N} \frac{R_i(t)}{\overline{X}_i(t)}} \tag{3.14}$$

$$\text{s.t.} \quad \frac{R_i(t)\gamma^*}{W} \frac{I_{p2s,i}(t) + \sum_{j=1,j\neq i}^{N} P_j(t) g_{s2s,ji}(t) + \eta}{P_i(t) g_{s2s,ii}(t)} \leq 1, \quad i = 1, 2, \ldots, N \tag{3.15}$$

$$\frac{1}{I_{\text{th}}} \sum_{i=1}^{N} P_i(t) g_{s2p,i}(t) \leq 1 \tag{3.16}$$

$$0 \leq \frac{P_i(t)}{P_{\max}} \leq 1, \ i = 1, 2, \ldots, N, \tag{3.17}$$

where the objective is to minimize the inverse of the original utility function. We change the equality in the rate constraint in (3.10) to an inequality in (3.15) without affecting the results of the problem, and replace $I_{s2s,i}(t)$ with the right hand side of (3.10). The objective function and each of the constraints are in the same format as in the standard GP problem, and the optimization problem can be solved using the method in [128].

In [128], it is demonstrated that the solution converges to the global optimal most of the cases, although it may converge to a local optimal under some conditions. In solving (3.14), we have done extensive tests using the same method and compared the results with the global optimal solution obtained using the exhaustive search method. Our results show that over 99% of the cases the iterative method converges to the global optimum solution for randomly selected initial values in the feasible region. In other cases it converges to a local optimum solution.

Solving the problem is time consuming, since it takes a large number of iterations before the method converges, especially when the number of links is large. In addition, the optimum scheduling solution cannot be easily implemented in a practical system as it requires the scheduler to know $g_{s2p,i}(t)$, $g_{s2s,ij}(t)$, and $I_{p2s,i}(t)$ for all $i, j = 1, 2, \ldots, N$. It is unlikely that all this information can be collected, even there exists a central station. Note that with a small probability the iterative method does not find the global optimum but a local optimum solution. However, we still refer to the solution as "optimum" to simplify the presentation. In the following two sections, we design two heuristic multi-user scheduling (MUS) schemes. Both the schemes are performed mainly by the MSs in the secondary network distributively, and require some limited assistance from the primary BS. The schemes can be easily implemented in a practical system.

3.4 Proposed MUSU scheme

In this section we propose a MUSU scheme, where the second "U" stands for "utility" since the scheduling decision takes into consideration the utility that each link can potentially contribute to the objective utility function.

The scheduling is performed periodically at the beginning of each scheduling interval (SI), which is defined as a period of time over which the physical channel conditions are relatively stable. We consider that a SI is much longer than a time slot so that the overhead for making the scheduling decisions has the minimum impact on the transmission throughput. At the beginning of each SI, there is no transmission in the secondary network, and each receiver measures its received $I_{p2s,i}(t) + \eta$ level, which is a sum of the interference from the primary network and the background

noise power. Receiver i then passes the value of $I_{p2s,i}(t) + \eta$ to its transmitter. This information exchange can be performed in different ways, one of which is to combine this with the estimation of link gain $g_{s2s,ii}(t)$, which is also required in making the scheduling decisions as will be seen below. The transmitter-receiver pairs either follow a pre-assigned order or are scheduled by the BS in order to exchange special channel estimation packets. The RTS/CTS exchanges can be used for this purpose. Upon receiving the packet from transmitter i, receiver i estimates $g_{s2s,ii}(t)$ based on the received signal strength and returns the value together with the value of $I_{p2s,i}(t)+\eta$ back to the transmitter. Upon receiving the message containing $g_{s2s,ii}(t)$ and $I_{p2s,i}(t) + \eta$, transmitter i estimates its maximum transmission rate, $\hat{R}_i(t)$, assuming $I_{s2s,i}(t) = 0$, as

$$\hat{R}_i(t) = \frac{W}{\gamma^*} \frac{\hat{P}_i(t) g_{s2s,ii}(t)}{I_{p2s,i}(t) + \eta}, \tag{3.18}$$

where $\hat{P}_i(t)$ is the maximum allowed transmission power of the MS limited by both the equipment maximum transmission power and the available interference threshold, i.e.,

$$\hat{P}_i(t) = \min \left\{ P_{\max}, \frac{I_{\text{th}}}{g_{s2p,i}(t)} \right\}. \tag{3.19}$$

Note that I_{th} can be considered as a network resource that is unique to a cognitive wireless network, since a higher I_{th} allows MSs to transmit with higher power. In (3.19) it is assumed that the transmission power of each MS can consume all the interference threshold, if its maximum transmission power allows. The calculation of $\hat{R}_i(t)$ also needs $g_{s2p,i}(t)$, which can be measured at the MSs in the secondary network. We assume that the link gain from the secondary transmitter to the primary BS is equal to that in the reverse direction, so that $g_{s2p,i}(t)$ can be measured at transmitter

i based on the pilot signal level from the BS. Based on the calculated $\hat{R}_i(t)$, the transmitter estimates the utility that it can contribute to the objective function in (3.9) as

$$\hat{U}_i(t) = \frac{\hat{R}_i(t)}{\overline{X}_i(t)} = \frac{W}{\gamma^* \overline{X}_i(t)} \frac{\hat{P}_i(t) g_{s2s,ii}(t)}{I_{p2s,i}(t) + \eta}. \tag{3.20}$$

Replacing $\hat{P}_i(t)$ in (3.20) using the right-hand-side of (3.19), we have

$$\hat{U}_i(t) = \frac{W g_{s2s,ii}(t)}{\gamma^* \overline{X}_i(t)(I_{p2s,i}(t) + \eta)} \min \left\{ P_{\max}, \frac{I_{\text{th}}}{g_{s2p,i}(t)} \right\}. \tag{3.21}$$

Similar to SUS, in MUSU a transmitter with higher $\hat{U}_i(t)$ is given a higher priority to transmit. However, in MUSU multiple transmitters are allowed to transmit at the same time. A lower priority transmitter can transmit if interference caused by all higher priority ones at the primary BS is below I_{th}. In order to prioritize the transmissions, each transmitter starts a backoff timer that is inversely proportional to $\hat{U}_i(t)$. Transmitters with larger $\hat{U}_i(t)$ then have their backoff timers expire earlier. During the backoff period, the transmitters listen to the BS for an updated interference threshold, $I_{\text{th,rem}}$, which specifies how much interference the BS can still accept from the secondary transmissions on top of the current interference level. $I_{\text{th,rem}}$ is equal to I_{th} at the beginning of each SI and updated after more transmitters transmit. The maximum transmission power of a transmitter is limited by both $P_{\max}$ and the current $I_{\text{th,rem}}$. Upon the time when its backoff timer expires, transmitter i finds its transmission power as

$$P_i(t) = \min \left\{ P_{\max}, \frac{I_{\text{th,rem}}}{g_{s2p,i}(t)} \right\}. \tag{3.22}$$

If $P_i(t) > 0$, the MS transmits immediately, and $I_{\text{th,rem}}$ is updated at the BS based on the updated interference level. Mathematically, $I_{\text{th,rem}}$ is updated as $I_{\text{th,rem}} =$

$I_{\text{th,rem}} - P_i(t)g_{s2p,i}(t)$. Later on, the transmitter with the next largest $\hat{U}_i(t)$ will have its backoff timer expire. The above process is repeated until either $I_{\text{th,rem}} = 0$ or all the MSs have transmitted. The transmission rates are found through adaptive rate allocations distributively based on the experienced interference at the receivers and the SINR requirements of the links. After the rate adaptation, the transmitter updates the smoothed average rate, $\overline{X}_i(t)$, at the beginning of each time slot based on (3.2), although the scheduling decisions are kept the same throughout the SI.

Let $\mathcal{T}$ represent a set of the scheduled links. The following simple pseudocode summarizes how to find the transmission rate and power vectors mathematically based on the MUSU scheduling process.

Pseudocode 1: MUSU

1: Define $\mathcal{C} = \{1, 2, \ldots, N\}$, $\mathcal{T} = \emptyset$, and $I_{\text{th,rem}} = I_{\text{th}}$.
2: Find $\hat{U}_i(t)$ for $i = 1, 2, \ldots, N$.
3: **while** $\mathcal{C}$ is not empty and $I_{\text{th,rem}} > 0$ **do**
4: Find $k = \text{argmax}_{i \in \mathcal{C}} \hat{U}_i(t)$.
5: $P_k(t) = \min\{P_{\max}, \frac{I_{\text{th,rem}}}{g_{s2p,k}(t)}\}$.
6: $\mathcal{T} = \mathcal{T} \cup \{k\}$ and $\mathcal{C} = \mathcal{C} \setminus \{k\}$.
7: $I_{\text{th,rem}} = I_{\text{th,rem}} - P_k(t)g_{s2p,k}(t)$.
8: **end while**
9: $P_i(t) = 0$ for all $i \in \mathcal{C}$.
10: Find $R_i(t)$'s from (3.10) and (3.10).

The MUSU scheduling is mainly done at the secondary network in a distributive way with limited assistance from the primary BS, which measures the amount of interference from the secondary transmissions and broadcasts the updated $I_{\text{th,rem}}$

to the secondary network. There is no need to exchange information among the secondary links. Although the transmission rates eventually are also dependent on link gains, $g_{s2s,ij}(t)$'s for all $i, j = 1, 2, \ldots, N$, these link gain values are not required in making the scheduling decisions.

Multiple MSs can transmit simultaneously using the MUSU scheduling scheme, provided the total interference to the primary network does not exceed the interference threshold. However, this does not mean that higher utility can always be achieved by allowing more MSs to transmit, since mutual interference can deteriorate the throughput. In order to increase the overall utility, the scheduling should avoid simultaneous transmissions with strong mutual interference. This can be achieved by using the concept of exclusive region as in [131] and [132]. After all higher priority transmitters have transmitted, a lower priority one can transmit only if it does not cause strong interference to any scheduled transmissions.

One practical way to implement the exclusive regions in the ad hoc network is to use short RTS/CTS messages. When the backoff timer of a transmitter expires, the transmitter sends a short RTS frame to its receiver, which responds with a short CTS frame. Other transmitters that are still in backoff overhear the messages and will not transmit in the current SI. A transmitter that does not hear any RTS/CTS frames before expiration of its backoff timer can transmit. Note that using the exclusive regions does not completely eliminate mutual interference among simultaneous transmissions. The transmission rates of the links are obtained through adaptive rate allocations. Mathematically, the rates can be found from (3.10) and (3.10).

Pseudocode 2 is a simplified code that can be used to find the rate and power vectors in the MUSU scheduling process with exclusive regions, and it is a slight

modification of Pseudocode 1.

Pseudocode 2: MUSU with exclusive regions

1: Define $\mathcal{C} = \{1, 2, \ldots, N\}$, $\mathcal{T} = \emptyset$, and $I_{\text{th,rem}} = I_{\text{th}}$.
2: Find $\hat{U}_i(t)$ for $i = 1, 2, \ldots, N$.
3: **while** $\mathcal{C}$ is not empty and $I_{\text{th,rem}} > 0$ **do**
4: Find $k = \text{argmax}_{i \in \mathcal{C}} \hat{U}_i(t)$.
5: **if** Link k is not in the exclusive region of any link in $\mathcal{T}$ **then**
6: $P_k(t) = \min\{P_{\max}, \frac{I_{\text{th,rem}}}{g_{s2p,k}(t)}\}$.
7: $\mathcal{T} = \mathcal{T} \cup \{k\}$ and $\mathcal{C} = \mathcal{C} \setminus \{k\}$.
8: $I_{\text{th,rem}} = I_{\text{th,rem}} - P_k(t) g_{s2p,k}(t)$.
9: **else**
10: $P_k(t) = 0$ and $\mathcal{C} = \mathcal{C} \setminus \{k\}$.
11: **end if**
12: **end while**
13: $P_i(t) = 0$ for all $i \in \mathcal{C}$.
14: Find $R_i(t)$'s from (3.10) and (3.10).

Note that it takes time for the primary BS to detect the interference changes after each transmission. If the time difference between the backoff timers of two transmitters is less than the maximum round-trip time between the primary BS and the secondary transmitters, the secondary transmitter whose backoff timer expires latter may not receive the most updated $I_{\text{th,rem}}$ value and therefore transmit higher power than it should. This may result in that the actual interference level at the primary BS is higher than I_{th}. As soon as the BS realizes this, it broadcasts a power down message indicating the current interference level I'. Upon receiving this message,

the transmitters that are still in backoff should not transmit in the current SI, and the transmitting stations should reduce their transmission power to $\frac{I_{\text{th}}}{I'}$ of its current transmission power.

3.5 Proposed MUSI scheme

Ideally, if mutual interference among peer links can be completely eliminated from the secondary network by using the exclusive regions, then $I_{s2s,i} = 0$ for all $i \in \mathcal{T}$. The scheduling is then to find the set $\mathcal{T}$ and the rate and power vectors so that

$$\max_{\mathbf{R}(t),\mathbf{P}(t)} \sum_{i\in\mathcal{T}} \frac{R_i(t)}{\overline{X}_i(t)} \tag{3.23}$$

$$\text{s.t.} \quad R_i(t) = \frac{W}{\gamma^*} \frac{P_i(t) g_{s2s,ii}(t)}{I_{p2s,i}(t) + \eta},\ i \in \mathcal{T} \tag{3.24}$$

$$\sum_{i\in\mathcal{T}} P_i(t) g_{s2p,i}(t) \le I_{\text{th}} \tag{3.25}$$

$$0 \le P_i(t) \le P_{\max},\ i \in \mathcal{T}. \tag{3.26}$$

Let $Q_i(t) = P_i(t) g_{s2p,i}(t)$ and replace $R_i(t)$ in (3.23) with the right-hand-side of (3.24), the above optimization problem can be rewritten as:

$$\max_{\mathbf{R}(t),\mathbf{P}(t)} \sum_{i\in\mathcal{T}} \frac{W g_{s2s,ii}(t)}{\gamma^* \overline{X}_i(t)[I_{p2s,i}(t) + \eta]} \frac{1}{g_{s2p,i}(t)} Q_i(t) \tag{3.27}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{T}} Q_i(t) \le I_{\text{th}} \tag{3.28}$$

$$0 \le Q_i(t) \le P_{\max} g_{s2p,i}(t),\ i \in \mathcal{T}. \tag{3.29}$$

From the constraint in (3.28) we can find that all $Q_i(t)$'s are equally weighted in terms of their contributions to the total secondary-to-primary interference. Meanwhile, the

contribution of each $Q_i(t)$ to the objective function is proportional to a weight $w_i(t)$ given by

$$w_i(t) = \frac{W g_{s2s,ii}(t)}{\gamma^* \overline{X}_i(t)[I_{p2s,i}(t) + \eta]} \frac{1}{g_{s2p,i}(t)}. \tag{3.30}$$

A scheduling scheme can then be designed so that a higher priority is given to the links with larger weights at the beginning of the current SI. The scheduling works in a similar way as MUSU, except that the backoff timer of each transmitter is inversely proportional to $w_i(t)$. We can find that prioritizing the transmissions based on the weight defined in (3.30) considers not only the objective function as in MUSU, but also the interference threshold, and therefore the scheduling scheme is referred to as MUSI, where "I" stands for interference threshold.

The main difference between MUSU and MUSI is the different effects of $g_{s2p,i}(t)$ on link transmission priorities. In (3.30), a smaller $g_{s2p,i}(t)$ leads to a larger $w_i(t)$ or a higher priority for link i, when other conditions are the same. However, in (3.21) $g_{s2p,i}(t)$ does not independently affect the transmission priority(or $\hat{U}_i(t)$). When $g_{s2p,i}(t)$ is small so that $P_{\max} < \frac{I_{\text{th}}}{g_{s2p,i}(t)}$, the effect of $g_{s2p,i}(t)$ on $\hat{U}_i(t)$ is overwhelmed by $P_{\max}$. As a result, a link with very low $g_{s2p,i}(t)$ may not have a chance to transmit if MUSU is used. The same link has a high priority to transmit if MUSI is used. From the optimization problem we can observe that when the values of $\frac{W g_{s2s,ii}(t)}{\gamma^* \overline{X}_i(t)(I_{p2s,i}(t)+\eta)}$ are the same for two links, the one with smaller $g_{s2p,i}(t)$ should have a higher priority to transmit as it adds little to the secondary-to-primary interference but can greatly increase the objective function. A link with smaller $g_{s2p,i}(t)$ can contribute more to the overall system utility.

The concept of exclusive region can also be used in the MUSI scheme as in the MUSU scheme for further reducing the mutual interference among peer links.

3.6 Numerical Results

Table 3.1: Default Simulation Parameters

Parameter	Value
Cellular cell radius	1000m
Total number of primary links	10
Transmission rate of each primary link	32kbps
Spread spectrum bandwidth W	5MHz
Background AWGN noise power η	10^{-10}W
Path loss exponent	4
Standard deviation of log-normal fading σ	8dB
Maximum primary user transmission power	0.5W
Maximum secondary user transmission power $P_{\max}$	0.05W
SINR threshold γ^*	5dB
Default interference threshold I_{th}	10^{-12}W
Maximum secondary user transmission range	100m
t_c	1000 time slots

We consider a generic cell of a cellular network as the primary network, where MSs are uniformly distributed in a circular coverage area of the BS. We then distribute N transmitters for the secondary network with their positions uniformly distributed in the BS coverage area. For each link, the receiver is randomly located within the circular ad hoc coverage area of the transmitter. The physical channel propagation model is a superposition of path loss and large scale lognormal shadowing. Default simulation parameters can be found from Table 3.1. Below we demonstrate performance of the proposed scheduling schemes. We use MUSU-EX and MUSI-EX, respectively, to represent the MUSU and MUSI schemes using exclusive regions, MUSU and MUSI for without exclusive region. For comparison, we also show performance of the optimum solution, SUS and max-rate scheduling. Results of max-min scheduling are not shown as its throughput is several magnitudes smaller than the other schemes.

We first demonstrate the throughput performance of the proposed scheduling schemes in Fig. 3.1. For comparison, we also show the throughput of OPT-A and

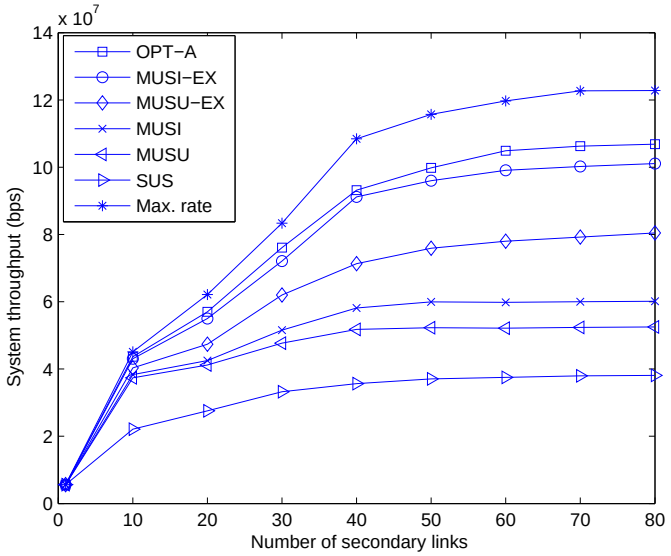

Figure 3.1: System throughput of different schemes

SUS. All the proposed schemes outperform SUS since they can take better advantage of the available resources than SUS by allowing multiple MSs to transmit simultaneously whenever possible. The MUSI scheme achieves higher throughput than MUSU, and this verifies that the prioritization metric in (3.30) is better than that in (3.21) in terms of throughput. All the proposed scheduling schemes achieve lower throughput than OPT-A, since the proposed heuristic schemes cannot take accurate mutual interference among peer links into consideration when making scheduling decisions. Among all the schemes shown in Fig. 3.1, the max-rate scheduling achieves the highest throughput. Later we will demonstrate that the high throughput is achieved at the price of significant unfairness. Comparing the scheduling schemes with and without using the exclusive regions, we find significant throughput improvement using the exclusive regions. By avoiding links with strong mutual interference to transmit simultaneously, the overall throughput can be improved greatly. This is true for both

the MUSU and MUSI schemes. This proves that when reducing the mutual interference $I_{s2s,i}$, throughput performance of the proposed scheduling schemes approaches to that of the optimum scheduling. Among all the heuristic schemes, the MUSI-EX is the one with throughput performance closest to the optimum.

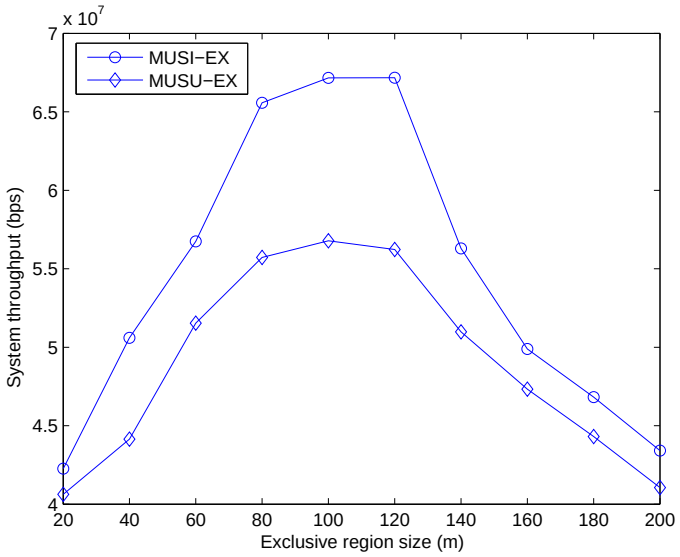

Figure 3.2: System throughput vs. exclusive region size

Although having a reasonable size of exclusive regions can eliminate simultaneous transmissions with strong interference, a very large size of exclusive region decreases the number of simultaneous transmissions and reduces the system level throughput. In Fig. 3.2 we show the effect of the exclusive region size on the system throughput, assuming the size can be perfectly controlled. We find that the system throughput increases with the exclusive region size at first, then decreases when further increasing the exclusive region size. The optimum size of the exclusive regions may depend on different system parameters, while developing an analytical model to obtain the accurate relationship can be difficult.

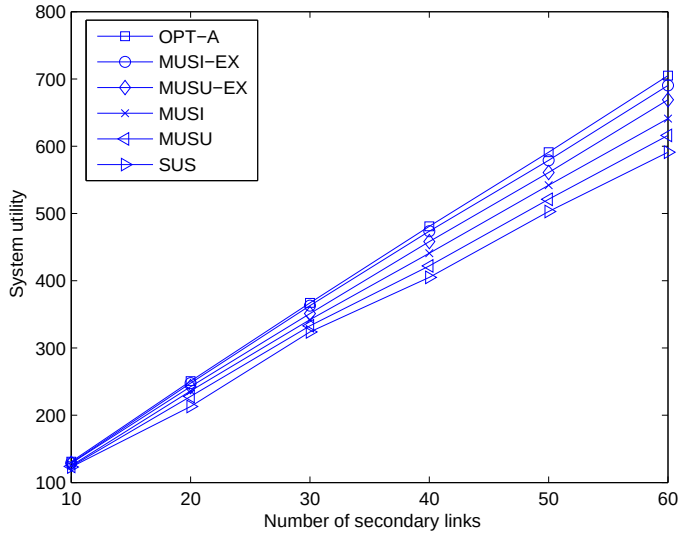

Figure 3.3: System utility $\sum_{i=1}^{N} \log \overline{R}_i$

In Fig. 3.3, we compare the system utility $\sum_{i=1}^{N} \log \overline{R}_i$ of the proposed PF scheduling schemes. For comparison we also show the utility of SUS. The optimal scheduling gives the highest utility, and SUS results in the lowest utility. For the proposed PF scheduling schemes, the ones with exclusive regions achieve higher utility. This indicates that eliminating strong mutual interference among simultaneous transmissions can improve the system level throughput and increase the overall utility function. We also find that the MUSI scheme achieves higher utility than the MUSU scheme. This verifies the intuitive comparison between MUSU and MUSI in the last paragraph of Section 3.5.

We then use the Jain's fairness index to evaluate the fairness of the proposed scheduling schemes. The fairness index is $f = \frac{\sum_{i=1}^{N}(U_i)^2}{\left(\sum_{i=1}^{N} U_i\right)^2}$ as defined in [133], where U_i's, $U_i = \frac{\overline{R}_i}{\overline{X}_i}$ for $i = 1, 2, \ldots, N$, are the performance metrics to be evaluated, and $\overline{X}_i = \lim_{t\to\infty} \overline{X}_i(t)$. In fact, U_i indicates the tradeoff between service rate and transmission

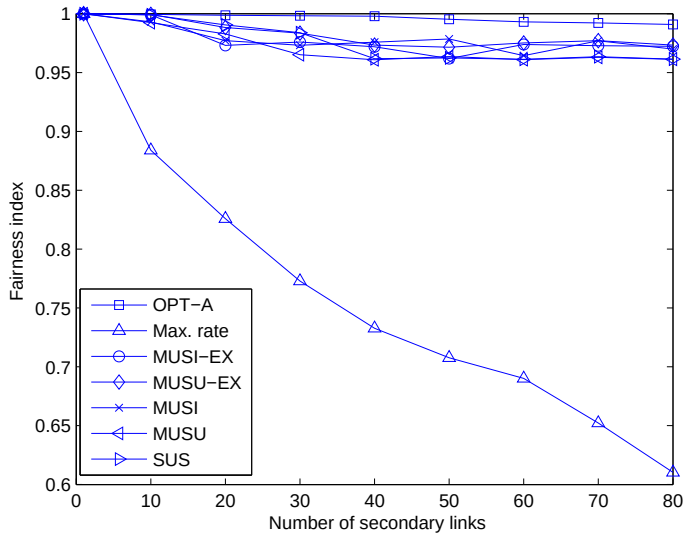

Figure 3.4: Fairness index

delay for link i. From (3.2) we can find that the instantaneous transmission rates of a link are discounted with time when calculating $\overline{X}_i(t)$. For example, consider that for the first 2 time slots, the rates for link 1 are $R_1(1) = 0$ and $R_1(2) = R_0 > 0$, and for link 2 are $R_2(1) = R_0$ and $R_2(2) = 0$. For all later time slots, $R_1(t) = R_2(t)$. When the simulation is completed, the two links have the same average rate, i.e., $\overline{R}_1 = \overline{R}_2$, but $\overline{X}_1 > \overline{X}_2$ and therefore $U_1 < U_2$. That is, the link with its traffic served later has a smaller U_i. In general, to have the same U_i values for multiple links, the links with shorter service delay should receive lower average throughput compared to the links with longer delay. For OPT-A, the fairness index is approximately equal to 1 as shown in Fig. 3.4. Theoretically, $\overline{X}_i$ should converge to $\overline{R}_i$ when t goes to infinity, and the fairness index of OPT-A should be 1. In simulations, the time required to have $\overline{X}_i$ be close to $\overline{R}_i$ increases with the total number of links. Therefore, the fairness index for OPT-A is slightly smaller than 1 when N is very large. The fairness indexes of

all the proposed scheduling schemes and the SUS scheme are very close to 1 as well, although slightly smaller than the fairness index of OPT-A. Note that the drawback of SUS is its low resource utilization but not fairness. The max-rate scheduling results in the worst fairness index, as it always gives transmission opportunity to the link with the best channel condition and achieves high transmission rate by sacrificing the transmission opportunities for links with poorer channel conditions.

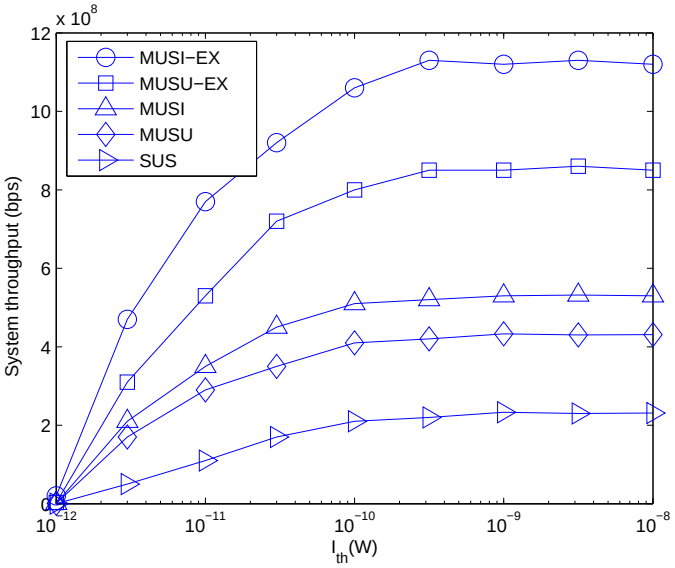

Figure 3.5: System throughput vs. interference threshold

We then look at the throughput of the scheduling schemes with different interference thresholds. From Fig. 3.5 we can find that when the interference threshold is below a certain value, the throughput increases with the interference threshold. This is true for all the schemes shown in the figure. Beyond a certain value, further increasing the interference threshold does not affect the network throughput anymore. This is due to the contradictory effects of the interference threshold on the secondary

network. On one hand, increasing the interference threshold may allow higher transmission power from the MSs. This can potentially increase the transmission rates. On the other hand, a higher interference threshold forces the primary MSs to transmit higher power and causes higher interference to the secondary transmissions, and this may reduce the secondary transmission rates. When the interference threshold is relatively low, for example, smaller than the background noise power, increasing I_{th} has little effect on the primary transmission power. Therefore, the primary to secondary interference is almost constant when I_{th} increases, and the secondary link transmission rate increases with I_{th}. When I_{th} is relative high, the high interference from the primary transmissions limits the increase of transmission power of the secondary MSs and therefore limits their throughput.

Selecting different values for t_c affects the tradeoff between system throughput and packet transmission delay. When t_c is small, a higher weight is given to the current transmission rate as shown in (3.2) and the scheduling reacts to channel condition changes quickly. In this case, fairness can be achieved in a relatively small time scale. However, the instantaneous transmission rates can change very abruptly from one SI to the next. This is shown as the cases of $t_c = 10$ and $t_c = 10^2$ in Fig. 3.6. When t_c is large, the effect of current channel condition on the scheduling is reduced, and the transmission rates change much smoother with time, for example, when $t_c = 10^3$ and $t_c = 10^4$. Meanwhile, fairness is achieved over a longer time period as t_c increases. The exact value of t_c depends on specific applications and is beyond the scope of this paper.

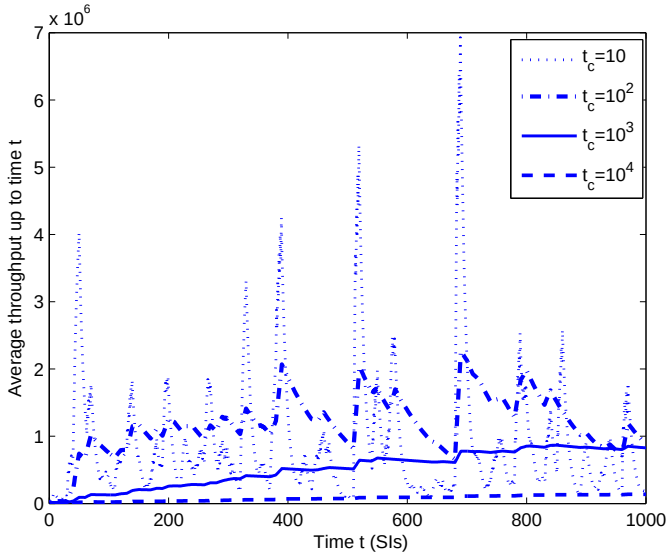

Figure 3.6: Effect of t_c on average throughput (MUSI)

3.7 Summary

In this chapter, we have proposed two proportional fair scheduling schemes for best effort traffic in a CDMA-based CRN with spectrum underlay. Based on long term average rates, both the scheduling schemes can be performed at the secondary network distributively with limited assistance from the primary network. With much reduced complexity, the proposed scheduling schemes can achieve close-to-optimum throughput while keeping good fairness.

In this work we have assumed that there is no minimum rate requirement for the secondary links. In the next chapter we will extend this work to consider that each link has a minimum average rate requirement and design admission control and scheduling schemes to guarantee this requirement.

Chapter 4

Joint Admission Control and Packet Transmission Scheduling in CRNs

4.1 Introduction

It has been demonstrated in the previous chapter that resources can be well managed in a CRN when serving best effort data traffic in order to optimize the throughput performance. In this chapter we study resource management for supporting traffic with strict quality of service requirements in a CRN with spectrum underlay. Both admission control and packet transmission scheduling are considered. When the traffic load at the primary network is relatively stationary, the capacity change in the CRN is relatively stationary. Using admission control can effectively ensure the amount of admitted traffic below the CRN system capacity over a long term. The actual QoS of the admitted traffic is achieved through packet-by-packet scheduling. Both

streaming traffic and non-real-time (nrt) traffic are considered. The streaming traffic requires a constant replay rate at the receiver, and the non-real-time data traffic requires an average transmission throughput. By opportunistically taking advantage of the random link and interference conditions during the scheduling process, very low session outage probability can be achieved for the streaming traffic, and the nrt traffic can receive the required throughput.

The remainder of this chapter is organized as follows. In Section 4.2 we describe the system model on which this work is based. Section 4.3 presents the proposed admission control scheme. The packet scheduling schemes are described in Section 4.4. Numerical results are demonstrated in Section 4.5, followed by a summary in Section 4.6.

4.2 System Description

We consider the same CRN network as described in the previous two chapters. Two types of traffic are served, non-real-time (nrt) data traffic and streaming traffic. We use $\mathcal{D}$ to represent a set of the nrt data connections and $\mathcal{S}$ to represent a set of the streaming connections. With each link carrying only one connection, $\mathcal{D}$ and $\mathcal{S}$ can also be used to represent the set of links carrying the nrt data and that of the links carrying the streaming traffic, respectively. Each nrt traffic source generates data at a random rate and the data link requires an average transmission throughput $R_{\text{req},i}$. For the streaming traffic, all data can be pre-stored in a buffer at the transmitter, or the source at the transmitter generates packets after the connection is established and stores the packets in a buffer before they are transmitted. The buffer at the transmitter is assumed to be of infinite size. After a link is established, successfully

transmitted data are first stored in a buffer at the receiver and replayed at a constant rate $R_{\text{req},i}$. The proposed work, however, can be extended for supporting variable rate streaming traffic by converting the variable rate into an equivalent constant rate based on the average rate, peak rate, tolerable packet loss rate, required delay, and other traffic characteristics. An example of this type of methods is to use the effective bandwidth [134]-[135]. The equivalent constant service rate (or effective bandwidth) is between the average rate and peak rate. It is closer to the average rate if the service can tolerate longer delay or higher packet loss rate; and it is closer to the peak rate if the service requires very stringent delay and low packet loss rate. Having larger buffer size at the receiver allows the sender to transmit at a high rate to the receiver whenever the radio resource is available provided the delay for replaying the data is tolerable.

Due to the random resource availability and the time varying channel and interference conditions, the instantaneous transmission rate of link i, $R_i(t)$, varies with time. For each nrt link, there is a buffer with an infinite size at each receiver, and therefore the instantaneous transmission rate can be as high as the channel conditions and available resources allow. For the steaming traffic, the receiver of the ith link has a buffer with a maximum size of $B_{\max,i}R_{\text{req},i}$. The receiver starts playing the data as soon as its data buffer becomes full for the first time. Replayed data is removed from the buffer. Session outage occurs when the receiver buffer does not have sufficient data to keep the required replaying rate. Intuitively, larger $B_{\max,i}$ results in fewer session outages since the session can tolerate up to $B_{\max,i}$ consecutive link outages ($R_i(t) = 0$) without a session outage. On the other hand, having larger $B_{\max,i}$ introduces longer delay to start playing the data at the receiver. The CR

nodes of each link measure the link gain, $g_{ii}(t)$, which includes both the path loss and log-normally distributed shadowing. Information of the transmission power is included in a data message, and the receiver finds the link gain based on the received signal strength. The aggregate interference that all primary transmissions cause to the ith secondary link, $I_{p2s,i}(t)$, is measured by the CR receiver at the beginning of each time slot when there is no transmission activity in the CRN. We assume that the time slots in the CRN and the primary network are perfectly synchronized, so that $I_{p2s,i}(t)$ keeps constant during one time slot. The long term average of $I_{p2s,i}(t)$ is denoted as $\tilde{I}_{p2s,i}$. The link gain between the transmitter of the ith secondary link to the primary BS, $g_{s2p,i}(t)$, is measured at the secondary transmitter based on the received signal strength of the pilot signal from the BS. The short term link gain and interference values are used for packet transmission scheduling, and the path loss and $\tilde{I}_{p2s,i}$ values are used for admission control.

4.3 Admission Control

Admission control should guarantee that all connections receive their required QoS throughout their lifetime. Therefore, it should consider the long-term channel and interference conditions, not only the conditions when the connection request is received. The interference constraint and random resource availability makes admission control in a CRN a very challenging issue. Furthermore, the admission control should consider the QoS not only for the link requesting for admission, but also for all existing links. In addition, admission control is also closely related to packet transmission scheduling, since the actual QoS of the connections is determined by how packet transmissions are scheduled. On the other hand, most practical scheduling schemes

are complicated, which generally should jointly consider not only the traffic characteristics and QoS requirements of individual connections, but also instantaneous link and interference conditions. This makes it very difficult to find the connection-level capacity based on specific scheduling schemes. Therefore, in this work we use relatively static information, such as path loss and average interference conditions, for admission control. We consider that the position information of all the stations in the CRN is available at the controller and use $\tilde{g}_{ji}$ to represent the path loss between the transmitter of link j and the receiver of link i. The path loss between the transmitter of link i and the BS, denoted as $\tilde{g}_{s2p,i}$, can also be obtained based on distance between the station and the BS. The admission control decisions based on such static information can tolerate random changes in the incoming traffic, measurement errors for the link gains and MS locations. On the other hand, the distance and average link gain information can be fairly accurate in a system where the MSs are relatively static.

Assume there are $N-1$ existing links in the CRN. When a new connection request arrives with the required rate $R_{\text{req},new}$, the request is sent to the controller, which assumes that the new request is the Nth link in the system and solves the following

optimization problem:

$$\max \tilde{R}_N \tag{4.1}$$

$$\text{s.t. } \frac{W\tilde{P}_i\tilde{g}_{ii}}{\tilde{R}_i\left(\sum_{j\neq i}^{N}\tilde{P}_j\tilde{g}_{ji}+\tilde{I}_{p2s,i}+\eta\right)} \geq \gamma^*,\ 1\leq i\leq N \tag{4.2}$$

$$\sum_{i=1}^{N}\tilde{P}_i\tilde{g}_{s2p,i} \leq I_{\text{th}},\ \text{if the new request is a streaming connection} \tag{4.3}$$

$$\sum_{i=1}^{N}\tilde{P}_i\tilde{g}_{s2p,i} \leq I_{\text{th,data}},\ \text{if the new request is a nrt connection} \tag{4.4}$$

$$\tilde{R}_i = R_{\text{req},i},\ 1\leq i\leq N-1 \tag{4.5}$$

$$0\leq \tilde{P}_i \leq P_{\max},\ 1\leq i\leq N, \tag{4.6}$$

where the unknowns are $\tilde{R}_N$ and $\tilde{P}_i$ for $i = 1,\ldots,N$, $\tilde{P}_i$ is the estimated transmission power for the transmitter of link i, and $I_{\text{th,data}} \leq I_{\text{th}}$. After solving the problem, if $\tilde{R}_N \geq R_{\text{req},new}$, the new request is accepted as the Nth existing link, and $R_{\text{req},N} = R_{\text{req},new}$. Otherwise, the user's request is rejected. Different interference thresholds are set depending on the new request is streaming or nrt traffic. When $I_{\text{th,data}} = I_{\text{th}}$, the nrt data requests have the same priority as the streaming traffic in admission control; when $I_{\text{th,data}}$ is smaller, fewer nrt data requests can be accepted, leaving more resources for potential streaming traffic requests. Outage performance for the streaming traffic is not explicitly considered in the admission control. However, the random changes of link gains and interference conditions provide the packet transmission scheduling with opportunities to achieve low outage probability. This will be detailed in Section 4.4.

When the primary users' activity is relatively stationary, it is possible for the CR stations to measure the experienced interference levels from the primary network and predict the future resource availability with high certainty. Therefore, the

QoS of admitted connections can be well guaranteed for a long term. On the other hand, if there is any abrupt change in the primary network transmission activity, QoS in the CRN may be violated and some admitted connections may have to be dropped. For example, when a new connection is generated at the primary network, especially the transmitter is close to the receiver of an existing secondary link, the primary-to-secondary interference experienced by the secondary receiver can increase significantly, both instantaneously and over a long term. The controller periodically reevaluates the resource availability at the CRN based on the updated interference levels. It checks the feasibility of all constraints in the above optimization problem. If any of the constraints cannot be satisfied, a certain criterion is used to remove one or more connections from the system until all the constraints can be satisfied for the remaining connections.

We consider three criteria for connection dropping, i) the most recently admitted connection is dropped first, ii) the connection with the highest $\tilde{I}_{p2s,i}$ is dropped first, and iii) the connection with the smallest estimated weight is dropped first. For the weight-based dropping criterion, an estimated weight is defined for each connection as follows

$$\tilde{w}_i = \frac{W}{\gamma^*} \frac{\tilde{g}_{ii}}{\tilde{g}_{s2p,i}\ \overline{X}_i\ (\tilde{I}_{p2s,i} + \eta)}. \tag{4.7}$$

This is similar to the weight defined in chapter 3 (3.30), where instantaneous link gains and interference level are used when defining the weight. Instead, the weight defined in equation (4.7) above is based on path losses and the average primary-to-secondary interference level in order to consider the long term transmission conditions. A link with smaller link gain and experiencing stronger interference may have a smaller $\tilde{w}_i$. Therefore, the intention for the weight-based dropping criterion is to

drop links with poorer transmission conditions. In designing these dropping criteria, we do not distinguish links with different type of the traffic, but this can be an option in practical systems. The exact criterion may depend on practical situations, for example, agreement between the users and the network operator, but is beyond the scope of this work.

4.4 Packet Transmission Scheduling

In the scheduling process, the first goal is to reduce the session outage for the streaming links and guarantee the average throughput for the nrt links. This will be presented in subsections 4.4.1 and 4.4.2. The scheduling should then try to utilize the available resources in a fair and efficient way. We design two scheduling methods, an optimum scheduling in Subsection 4.4.3 and a weight-based scheduling in Subsection 4.4.4. The former is an ideal method assuming global information about the link and interference conditions is available, and the latter is a heuristic scheme based on measured link and interference conditions at the CR nodes.

4.4.1 Reducing session outage probability for streaming traffic

For the streaming traffic, every effort should be made to reduce the session outage. Assuming that data can be replayed at the receiver during the packet transmission process, whether there is a session outage at time slot t depends on both the receiver buffer occupancy at the end of time slot $t-1$ and the transmission rate at time slot t. Let $B_i(t-1)$ represent the receiver buffer occupancy normalized to $R_{\text{req},i}$ for link

i at the end of time slot $t-1$, i.e., the actual buffer occupancy is $B_i(t-1)R_{\text{req},i}$. Then $B_i(t-1)$ indicates how much time the buffered data at the receiver can keep the session running without a session outage, if the link will not have a chance to transmit after time slot $t-1$. Because of the constant replay rate, the transmitter has full knowledge about $B_i(t-1)$. Define $\mathcal{S}' = \{i | B_i(t-1) < T_s, i \in \mathcal{S}\}$, where T_s is the duration of one time slot. Links in $\mathcal{S}'$ should have a higher priority to transmit at time slot t in order to avoid immediate session outage. The minimum transmission rate at time slot t to keep the session running for the entire duration of the time slot is given by

$$R_{\min,i}(t) = \frac{T_s - B_i(t-1)}{T_s} R_{\text{req},i}. \tag{4.8}$$

On the other hand, the limited buffer size at the receiver limits the maximum transmission rate. Consider that an amount of $R_{\text{req},i}T_s$ data is replayed and removed from the receiver buffer in each time slot (when there is no session outage), the maximum transmission rate for link i at time slot t is given by

$$R_{\text{buf},i}(t) = \frac{B_{\max,i} - B_i(t-1) + T_s}{T_s} R_{\text{req},i}. \tag{4.9}$$

4.4.2 Achieving average throughput for nrt traffic

For the nrt links, the objective is to maximize the resource utilization while satisfying the average throughput requirements. Meanwhile, we notice that in order to maximize the resource utilization, the average throughput can be achieved over a longer period of time for links with poorer channel and interference conditions. To have some control over the time scale over which the average throughput is achieved for the nrt traffic, we define a counter, $\xi_i(t)$, for each nrt link. The initial value of $\xi_i(t)$ is set to

zero and it is updated at the end of every time slot. It is increased by one if link i does not transmit at a time slot, and decreased by $\lfloor \frac{R_i(t)}{R_{\text{req},i}} \rfloor$ if the link transmits at rate $R_i(t)$. Define $\mathcal{D}^{'} = \{i | \xi_i(t-1) > \xi_{\max}, i \in \mathcal{D}\}$. Links in $\mathcal{D}^{'}$ have a higher priority to transmit in the next time slot. Intuitively, a smaller $\xi_{\max}$ helps to achieve the average throughput over a shorter period of time, but may force the links to transmit more frequently than necessary, and this can easily result in that some links transmit when their transmission conditions are poor.

4.4.3 Optimum scheduling

Let $\mathcal{S}^{''}$ be a set of the streaming links with $B_i(t-1) \geq T_s$, and $\mathcal{D}^{''}$ be a set of the nrt links with $\xi_i(t-1) \leq \xi_{\max}$. The objective of the scheduling for links in $\mathcal{S}^{''}$ and $\mathcal{D}^{''}$ is to improve the resource utilization, while keeping fairness among the links. Links with good transmission conditions should transmit at higher rates in order to efficiently utilize the available resources. Meanwhile, other links with poorer transmission conditions should also have chances to transmit in order to avoid very low rate over a long period of the time. The idea of packet scheduling with long-term proportional fairness (PF) can achieve this objective. It is proved that the PF scheduling is the best tradeoff between network resource utilization and user's satisfaction [107]. The long term PF scheduling further improves resource utilization compared to the slot-by-slot PF scheduling. In chapter 3 we have designed scheduling schemes to achieve long-term PF rates among links in a CRN with spectrum underlay, where the OPT-A scheme is an ideal scheme that achieves the optimum proportional fairness, and the MUSU scheme is a heuristic scheme that is based on measured link and interference conditions. The basic idea for these schemes is that links with either

good transmission conditions at the current time or low transmission rates in the past may have a higher priority to transmit. However, the schemes are specifically designed for best-effort data traffic. In this work we reconsider the scheduling schemes by jointly considering the streaming traffic and the nrt traffic. By modifying the OPT-A and MUSU schemes in chapter 3, an optimum scheduling is formulated in the remaining of this subsection, and a heuristic scheduling is designed in the next subsection.

Let $R_{\text{opt},i}(t)$ and $P_{\text{opt},i}(t)$, respectively, be the transmission rate and power for link i. The optimum scheduling scheme is formulated as follows :

$$\max \sum_{i=1}^{N} \frac{R_{\text{opt},i}(t)}{\overline{X}_i(t)} \tag{4.10}$$

$$\text{s.t.} \quad \frac{W P_{\text{opt},i}(t) g_{ii}(t)}{R_{\text{opt},i}(t) \left[\sum_{j\neq i}^{N} P_{\text{opt},j}(t) g_{ji}(t) + I_{p2s,i}(t) + \eta\right]} \geq \gamma^*, \ i = 1, 2, \ldots, N \tag{4.11}$$

$$R_{\text{opt},i}(t) > 0, \ i = 1, 2, \ldots, N, \tag{4.12}$$

$$R_{\text{opt},i}(t) \leq R_{\text{buf},i}(t), \ i \in \mathcal{S}, \tag{4.13}$$

$$R_{\text{opt},i}(t) \geq R_{\text{min},i}(t), \ i \in \mathcal{S}^{'}, \tag{4.14}$$

$$0 \leq P_{\text{opt},i}(t) \leq P_{\max}, \ i = 1, 2, \ldots, N, \tag{4.15}$$

$$\sum_{i\in\mathcal{S}'} P_{\text{opt},i}(t) g_{s2p,i}(t) \leq I_{\text{th}}, \tag{4.16}$$

$$\sum_{i\in\mathcal{D}'} P_{\text{opt},i}(t) g_{s2p,i}(t) \leq I_{\text{th}} - \sum_{j\in\mathcal{S}'} P_j g_{s2p,j}(t), \tag{4.17}$$

$$\sum_{i\in\mathcal{S}''} P_{\text{opt},i}(t) g_{s2p,i}(t) \leq I_{\text{th}} - \sum_{j\in\mathcal{S}'\cup\mathcal{D}'} P_j g_{s2p,j}(t), \tag{4.18}$$

$$\sum_{i\in\mathcal{D}''} P_{\text{opt},i}(t) g_{s2p,i}(t) \leq I_{\text{th}} - \sum_{j\in\mathcal{S}\cup\mathcal{D}'} P_j g_{s2p,j}(t), \tag{4.19}$$

where $\overline{X}_i(t)$ represents a smoothed average rate in the past and is defined as

$$\overline{X}_i(t+1) = \alpha \overline{X}_i(t) + (1-\alpha) R_{\text{opt},i}(t), \tag{4.20}$$

and $0 < \alpha < 1$. For a special case, $\overline{X}_i(t) = 0$ when $t = 0$. The constraints (4.12)-(4.14) are for the transmission rates, and the constraints (4.16)-(4.19) give different priorities of the links by setting different interference thresholds. The optimization problem is solved first using constraints (4.11)-(4.14) and (4.16) for all the links in $\mathcal{S}^{'}$. After having the results, it is then solved using constraints (4.11)-(4.14) and (4.17) for all the links in $\mathcal{S}^{''}$. After the scheduling is done for all the streaming links, it is solved using constraints (4.11)-(4.14) and (4.18) for all the links in $\mathcal{D}^{'}$, and finally solved using constraints (4.11)-(4.14) and (4.19) for all the links in $\mathcal{D}^{''}$. In this way, the streaming links in $\mathcal{S}^{'}$ have the highest priority, followed by the streaming links in $\mathcal{S}^{''}$, and the links in $\mathcal{D}^{''}$ have the lowest priority. If there is no feasible solution, link k is temporarily removed, where $k = \arg\max_{i, i \in \mathcal{S}'} B_i(t-1)$, and the optimization problem is solved again. This is repeated until a feasible solution is obtained.

The optimization problem is nonlinear and nonconvex. It can be reformulated into a general geometric programming problem as done in Section 3.3. The optimum scheduling, however, cannot be implemented in a practical system because it requires link gains and interference conditions within the CRN and between the CRN and the primary network.

4.4.4 Weight-based scheduling

In order to design a scheduling scheme that can be implemented in a practical system, we modify the MUSI scheme proposed in chapter 3 to incorporate the session outage

performance for the streaming traffic and the average throughput requirement for the nrt traffic. Define $L_i(t)$ as the priority level of link i at time slot t. A larger value of $L_i(t)$ indicates a lower transmission priority for the link. A link with $L_i(t) = 1$ has the highest priority to transmit, and a link with $L_i(t) = \infty$ cannot transmit at time slot t. Algorithm 1 is used to find the priority levels. The streaming links with $B_i(t-1) < T_s$ have the highest priority, and a higher priority is given to the link with smaller $B_i(t-1)$. After that, a higher priority is given to the nrt links with $\xi_i(t-1) > \xi_{\max}$, and a higher priority is given to the link with larger $\xi_i(t-1)$. For each of the other links, the streaming links have a higher priority than the nrt links. A weight $w_i(t)$ similar to that in chapter 3 is defined and a link with a larger weight has a higher priority to transmit. The weight for link i at time slot t is given by

$$w_i(t) = \frac{W}{\gamma^* \overline{X}_i(t) \left[I_{p2s,i}(t-1) + \eta\right]} \cdot \frac{g_{ii}(t-1)}{g_{s2p,i}(t-1)}. \tag{4.21}$$

Different from the definition in chapter 3, this weight in (4.21) is calculated based on the link and interference conditions at time slot $t-1$. Each CR transmitter calculates the weight at the end of time slot $t-1$ and passes it to the controller.

At the end of each time slot, the controller collects information from the CR nodes and calculates the transmission priority of each link for the next time slot using Algorithm 1. For the streaming links with $B_i(t-1) < T_s$ and the nrt links with $\xi_i(t-1) > \xi_{\max}$, the CR transmitters pass the respective $B_i(t-1)$ and $\xi_i(t-1)$ values to the controller at the end of the time slot $t-1$. For other links, the transmitters can pass their calculated $w_i(t)$ values to the controller, which performs Algorithm 1 and informs the CR transmitters of their priority levels through the control channel before the start of time slot t.

After performing Algorithm 1, transmission power for the links will be found. Define

$$I_{\text{th,rem}} = I_{\text{th}} - \sum_{\text{all } j, L_j(t)<L_i(t)} P_j(t) g_{s2p,j}(t) \tag{4.22}$$

as the amount of the interference that the BS can still tolerate after the transmissions of all the CR links with higher priority. The maximum transmission power of a nrt link i is then given by

$$P_i(t) = \min \left\{ P_{\max}, \frac{I_{\text{th,rem}}}{g_{s2p,i}(t)} \right\}. \tag{4.23}$$

In order to find the updated $I_{\text{th,rem}}$ for each link, the lower priority link should transmit slightly later. This is achieved by a backoff process similar to that in chapter 3. The transmitter of link i sets its backoff timer to be $L_i(t)\delta$, where δ is equal to the maximum round trip time required for a signal to propagate between an MS in the CRN and the controller. At the time when the transmitter of link i completes the backoff, the controller measures the interference level, updates $I_{\text{th,rem}}$, and broadcasts it in the CR network. Upon completing the backoff, the CR transmitter calculates its transmission power based on the current $I_{\text{th,rem}}$. Lower priority links cannot transmit if $I_{\text{th,rem}} = 0$ when they complete the backoff. Note that the total amount of time for backoff, denoted as T_h, is limited. Therefore, the number of transmitting links in a given time slot is limited to T_h/δ, and a link with $L_i(t) > T_h/\delta$ cannot transmit in the current time slot. Instead of having the MSs perform the backoff and calculate the transmission power, the CR transmitters can also pass values of $g_{s2p,i}(t-1)$ to the controller, which calculates the transmission power for each CR transmitter and passes the values back to the transmitters at the beginning of the time slot t. In either way, the controller plays an important role because the transmission power at the CRN is

limited by the interference level that is measured at the controller. Instead of using the backoff process to select the links that are allowed to transmit, an optimization problem can be formulated and find the optimum T_h/δ links. However, this requires the controller to collect information about the mutual interference among the links, which is practically difficult and therefore not considered here.

In addition, the concept of exclusive region is adopted to prevent strong mutual interference among the peer links. A link i in $\mathcal{S}''$ or $\mathcal{D}''$ cannot transmit if it is within the exclusive region of any link with a higher priority. The actual transmission power for each link is limited by the maximum transmission power of the transmitter node.

For a streaming link, the transmission power is also dependent on $R_{\text{buf},i}(t)$, which is the maximum rate limited by the receiver buffer space. The required transmission power for achieving $R_{\text{buf},i}(t)$ is estimated as

$$P_{\text{buf},i}(t) = \frac{\gamma^* R_{\text{buf},i}(t)[I_{p2s,i}(t) + \eta]}{W g_{ii}(t)}, \tag{4.24}$$

where mutual interference from peer links is ignored due to the use of exclusive regions which prevent strong mutual interference among the peer links. By taking into consideration the interference threshold and the maximum transmission power of the node, the actual transmission power is given by

$$P_i(t) = \min\left\{P_{\max}, \frac{I_{\text{th,rem}}}{g_{s2p,i}(t)}, P_{\text{buf},i}(t)\right\}. \tag{4.25}$$

Note that mutual interference cannot be completely eliminated by using the exclusive regions. Given the transmission power of each node, the actual transmission

Algorithm 1 Calculating priority levels

Define $\mathcal{A}_s^{'} = \mathcal{S}^{'}$, $\mathcal{A}_d^{'} = \mathcal{D}^{'}$, $\mathcal{A}_s^{''} = \mathcal{S}^{''}$, and $\mathcal{A}_d^{''} = \mathcal{D}^{''}$. Initialize $C = 0$, and $L_i(t) = \infty$ for all links.
while $\mathcal{A}_s^{'}$ is not empty **do**
  $i = \arg\min_{k \in \mathcal{A}_s^{'}} B_k(t-1)$
  $C = C + 1$, $L_i(t) = C$, and $\mathcal{A}_s = \mathcal{A}_s^{'} \setminus \{i\}$.
end while
while $\mathcal{A}_d^{'}$ is not empty **do**
  $i = \arg\max_{k \in \mathcal{A}_d^{'}} \xi_k(t-1)$
  $C = C + 1$, $L_i(t) = C$, and $\mathcal{A}_d^{'} = \mathcal{A}_d^{'} \setminus \{i\}$.
end while
while $\mathcal{A}_s^{''}$ is not empty **do**
  $i = \arg\max_{k \in \mathcal{A}_s^{''}} w_k(t)$
  $C = C + 1$, $L_i(t) = C$, and $\mathcal{A}_s^{''} = \mathcal{A}_s^{''} \setminus \{i\}$.
end while
while $\mathcal{A}_d^{''}$ is not empty **do**
  $i = \arg\max_{k \in \mathcal{A}_d^{''}} w_k(t)$
  $C = C + 1$, $L_i(t) = C$, and $\mathcal{A}_d^{''} = \mathcal{A}_d^{''} \setminus \{i\}$.
end while

rate is obtained through rate adaptation so as to meet the SINR requirements. Mathematically, the transmission rate for link i can be found from (4.26) as

$$R_i(t) \leq \frac{W P_i(t) g_{ii}(t)}{\gamma^* \left[\sum_{j=1, j\neq i}^{N} P_j(t) g_{ji}(t) + I_{p2s,i}(t) + \eta\right]}, \tag{4.26}$$

where the equality holds if the actual SINR is equal to γ^*.

The scheduling decisions strongly depend on the interference levels at the primary BS. In a practical system, the measured interference levels may not be accurate. As a result, the actual interference at the BS caused by the secondary transmissions, denoted by I_{me}, may exceed I_{th}. When this occurs, the controller notifies all the transmitting MSs to reduce their transmission power to $P_i(t) \times \frac{I_{\text{th}}}{I_{\text{me}}}$. Furthermore,

the allowed interference threshold can be reduced to βI_{th}, where β is between 0 and 1, in order to accommodate potential measurement errors. When β is smaller, more measurement errors can be tolerated in the scheduling, given the same probability that (4.16) is violated; or with the same measurement errors, (4.16) is violated with a smaller probability.

4.5 Numerical Results

We consider a generic cell of a cellular network as the primary network. The initial locations of the MSs in the primary network are uniformly distributed in a circular coverage area of the BS. For the CRN, the initial locations of the transmitters are uniformly distributed in the BS coverage area, and the receivers are randomly located within the circular ad hoc coverage area of the respective transmitters. The moving speed of the MSs is 5m/s and the direction is uniformly distributed between 0 and 2π. Whenever an MS moves to the boundary of the BS coverage, it changes to another direction randomly. The link gains are generated using the two-dimensional correlated shadowing model in [136], which considers both spatial and temporal correlations of the shadowing factors. The temporal correlation reflects the correlation of link gains over time due to moving of the MSs. The spatial correlation considers the fixed location of obstacles and finds the correlation in close distance. The autocorrelation function is determined by two factors: the resolution distance (which is 1m) gives the accuracy of the correlation, and the decorrelation length (which is 20m) is a constant recommended in [136]. Each $\tilde{I}_{p2s,i}$ is an average of the measured $I_{p2s,i}(t)$ values over one second interval.

In the primary network, transmissions from MS i to the BS in the uplink require a constant rate, $R_{p,i}$, and a minimum required SINR, γ_p^*, at the BS receiver. Let $P_{p,i}$ represent the transmission power of MS i in the primary network, and $g_{p,i}(t)$ the link gain between the MS and the BS. Assuming perfect power control is achieved, the values of $P_{p,i}$ for all i can be jointly solved from the following relationship, $\frac{WP_{p,i}(t)g_{p,i}(t)}{R_{p,i}\left[\sum_{j=1,j\neq i}^{M_p} P_{p,j}(t)g_{p,j}(t)+I_{\text{th}}+\eta\right]} \geq \gamma_p^*$, for $i = 1, 2, \ldots, M_p$, where M_p is the total number of primary links, and the equality is used so that each MS transmits at the minimum power.

For the CRN, two independent Poisson processes are used to generate the connection requests for the streaming traffic and nrt data traffic, respectively. The new connection is randomly (with equal chance) assigned to one of the transmitter-receiver pairs that do not have active connections at the time. The duration of each admitted connection follows an exponential distribution. For each admitted nrt connection, data packets each with a fixed size are generated using another Poisson process. Default parameters are listed in Table 4.1.

Fig. 4.1 shows the connection blocking rates of both types of the traffic. When generating Fig. 4.1, we fix the connection arrival rate for the streaming traffic and vary that for the nrt traffic. As the nrt connection arrival rate increases, the overall traffic load admitted into the network increases, and therefore the blocking rates for both types of the traffic may increase. Although the blocking rate for the nrt traffic keeps increasing, that for the streaming traffic is upper bounded. This is due to the use of $I_{\text{th,data}}$, which limits the amount of admitted nrt traffic. Therefore, when the arrival rate for the nrt connections is high, the high blocking rate for the nrt traffic guarantees that a certain amount of resource is always available for the

Table 4.1: Default Simulation Parameters

Parameter	Value
Cell radius	1000m
Path loss exponent	4
Standard deviation of log-normal fading	8dB
Spread spectrum bandwidth W	5MHz
SINR threshold γ^* (primary and secondary)	5dB
Background AWGN noise power η	10^{-10}W
Number of primary links	10
Transmission rate of primary link	32kbps
Maximum transmission power of primary MS	0.5W
Maximum transmission power of secondary MS $P_{\max}$	0.05W
Exclusive region size (radius)	100m
Interference threshold I_{th}	10^{-12}W
Secondary user transmission range	100m
Streaming connection arrival rate	100 requests/sec
Nrt data connection arrival rate	100 requests/sec
Average streaming connection duration	60 sec
Average nrt connection duration	120 sec
Data packet arrival rate	2.5 packets/sec
Time slot duration T_s	5ms
Overhead time per time slot T_h	0.5ms

streaming traffic. Furthermore, the blocking rate for the nrt traffic is much smaller when $I_{\text{th,data}} = 70\%$ than that when $I_{\text{th,data}} = 30\%$ since more system resource is available for the nrt traffic in the former case. Correspondingly, the blocking rate for the streaming traffic is higher when $I_{\text{th,data}} = 70\%$ than that when $I_{\text{th,data}} = 30\%$. Fig. 4.2 further shows that setting different values for $I_{\text{th,data}}$ can change the relative amounts of the streaming and nrt traffic admitted into the system. When $I_{\text{th,data}} = 0$, all the system resource is used for the streaming traffic. As $I_{\text{th,data}}$ increases, more nrt data connections and fewer streaming connections can be admitted into the system. Below we consider performance of the admitted streaming and nrt traffic separately.

Figs. 4.3 and 4.4 show the session outage performance for the streaming traffic. One session outage is recorded if for a given time slot, the receiver does not have

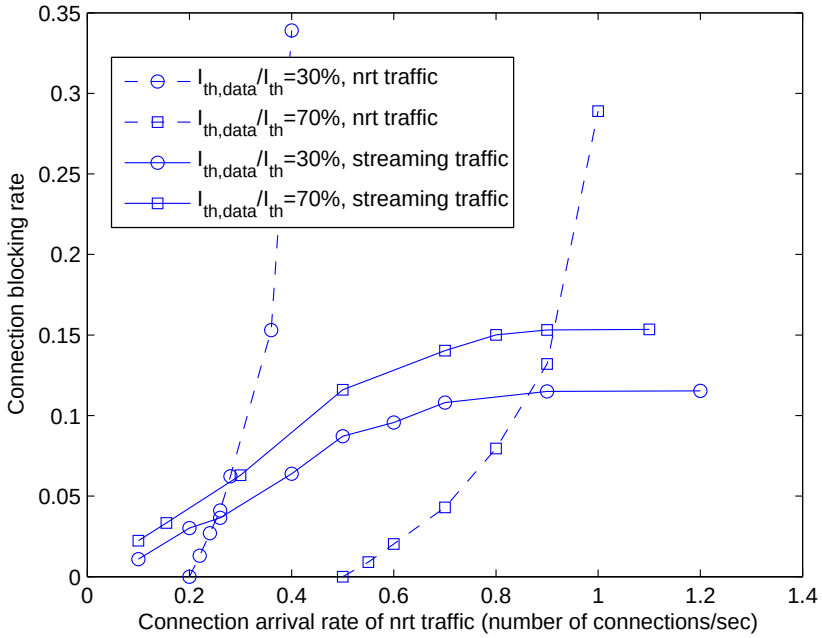

Figure 4.1: Connection blocking rate

enough data in the buffer to keep the replay rate at $R_{\text{req},i}$. Fig. 4.3 shows that both the optimum scheduling and the weight-based scheduling achieve very low session outage probability. This indicates that the admission control scheme and the scheduling together can effectively protect QoS of admitted traffic. By using only static link gains and average interference levels, the admission control ignores the random channel fading and interference changes in predicting the resource availability, and this gives chance for the scheduling scheme to take advantages of the random channel fading and interference changes so to achieve low session outage. The outage performance using the weight-based scheduling is very close to that using the optimum scheduling. This proves the effectiveness of the weight-based scheduling in determining the link transmission priority based on link and interference conditions.

In Fig. 4.3, we also plotted outage performance of two other schemes: buffer-based

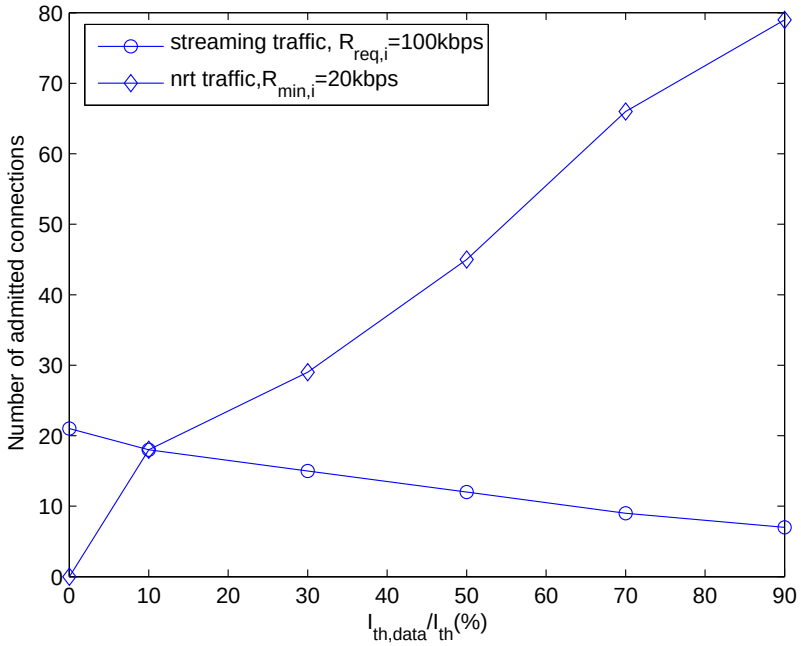

Figure 4.2: Effect of $I_{\text{th,data}}$ on amount of admitted traffic

and rate-based scheduling. In the buffer-based scheduling, links with smaller $B_i(t)$ values are given a higher priority to transmit; and in the rate-based scheduling, links with higher $R_{\text{opt},i}(t)$ are given a higher priority to transmit. Both the buffer-based and rate-based schemes result in about one magnitude higher outage probability than the proposed scheduling schemes. The rate-based scheme does not consider the buffer occupancy, which is directly related to the session outage, and therefore results in the highest outage probability. The buffer-based scheduling does not consider link transmission conditions. Links with small receiver buffer occupancy but poor transmission conditions may have to transmit at high power which leaves small $I_{\text{th,rem}}$ for other links and causes high interference to other links.

Fig. 4.4 shows fairness of the scheduling schemes in terms of outage performance, where we use Jain's fairness index [137]. The buffer-based scheme has the best fairness

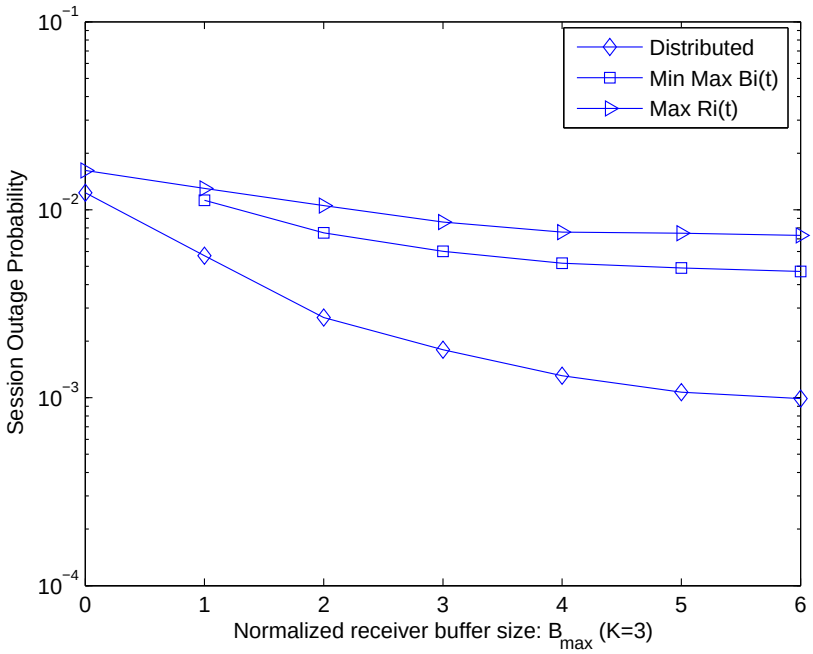

Figure 4.3: Session outage probability

among all the schemes since it always balances the buffer occupancy, which is directly related to the outage performance. Fairness of the weight-based scheme is very close to that of the buffer-based and the optimum scheduling since buffer occupancy is the first aspect considered when determining transmission priority of the streaming links. The rate-based scheme has the worst fairness because it does not consider buffer occupancy when determining the transmission priority.

Fig. 4.5 shows performance of the nrt traffic when using the weight-based scheduling, where $R_{avr,i}$ in the vertical axis is the cumulative average throughput of an admitted nrt connection over the time interval indicated in the horizontal axis. It is shown that when the time interval is sufficiently long (about 1500 time slots with the given parameters), all connections can receive their required average throughput. Within a short period of time, using smaller $\xi_{\max}$ may result in smoother rate changes

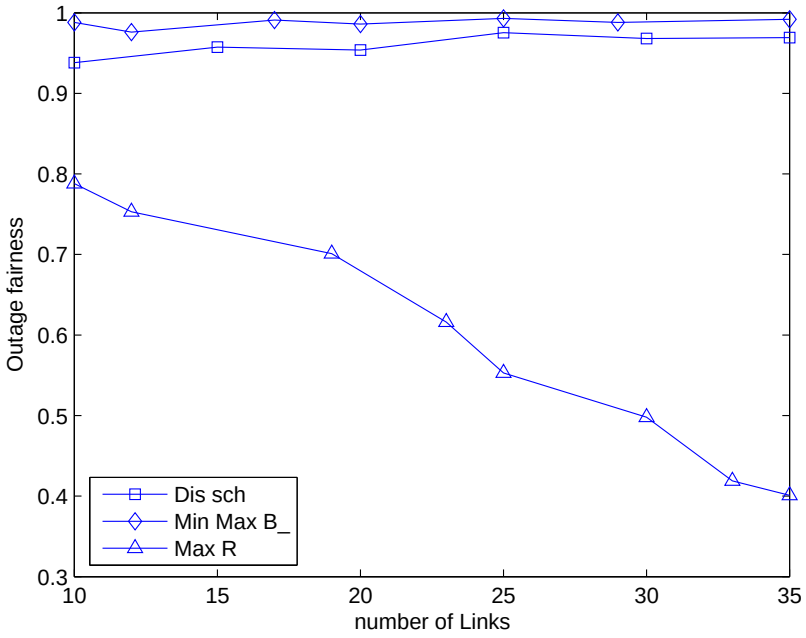

Figure 4.4: Fairness of session outage

since it forces the nrt connections to transmit more frequently. On the other hand, setting a larger value for $\xi_{\max}$ may benefit the system resource utilization.

Fig. 4.6 compares the packet transmission delay and throughput performance of the admitted nrt traffic. The figure shows that the performance of the optimum scheduling and the weight-based scheduling is quite close, in terms of both throughput and delay. This shows the effectiveness of the weight-based scheduling. Fig. 4.6 also shows that as $\xi_{\max}$ increases, the throughput increases, while data transmission delay increases too. This shows the tradeoff between the resource utilization and the latency.

Fig. 4.7 further shows the fairness of average packet transmission delay among the nrt links, where we can see that very good fairness can be achieved among different links.

Accurate interference measurements are assumed in generating the above results.

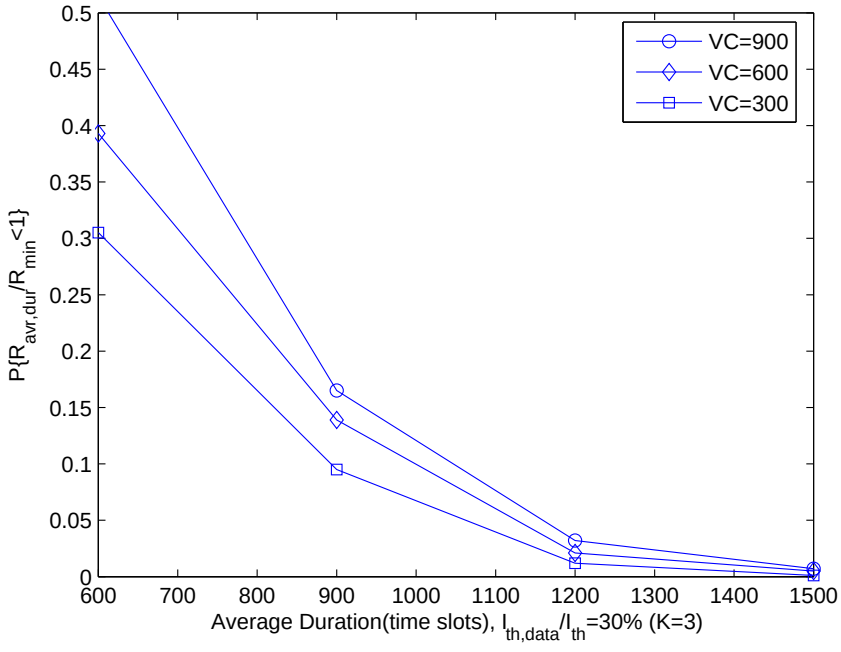

Figure 4.5: Cumulative average data throughput

In practice, the measured interference at the BS may not be accurate. As a result, the actual interference at the BS caused by the secondary transmissions, denoted by I_{me}, may exceed I_{th}. Whenever $I_{\text{me}} > I_{\text{th}}$, the controller notifies all the transmitting MSs to reduce their transmission power to $\frac{I_{\text{th}}}{I_{\text{me}}}$ of their current transmission power, in order to protect the transmissions in the primary network. To ensure the QoS of the traffic admitted into the CRN, the allowed interference threshold during admission control can be reduced to βI_{th}, where β is between 0 and 1. That is, I_{th} is replaced with βI_{th} during admission control, although the acceptable interference threshold packet transmissions is I_{th}. We assume that the measured interference from the secondary transmissions at the controller has an error which is log-normally distributed with mean of zero and standard deviation of δ. Figs. 4.8 and 4.9 show the effect of estimation error on the performance of the streaming traffic and the nrt traffic,

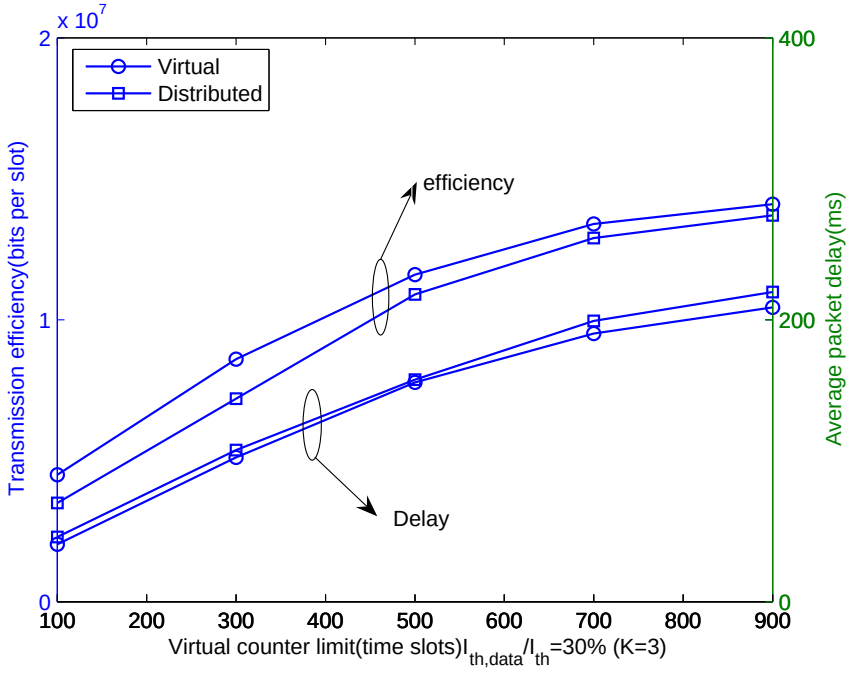

Figure 4.6: Throughput and delay performance of nrt traffic

respectively, based on the above approaches. Fig. 4.8 indicate that larger estimation errors can negatively affect the performance of the streaming traffic. On the other hand, this depends on the tolerable outage probability. If a slightly higher outage probability, such as 1% for the streaming traffic, the estimation error has very little effect on the performance of the streaming traffic. Reducing β during the admission control can reduce the outage probability of admitted streaming traffic at a price of a fewer admitted connections. Similar observations can be obtained from the nrt traffic from Fig. 4.9, although performance degradation is larger, compared to the streaming traffic, because of the lower priority of the nrt traffic. Mathematically analyzing the relationship among β, δ, and the required outage performance can be difficult, but the value of β can be adjusted dynamically based on the measured outage performance.

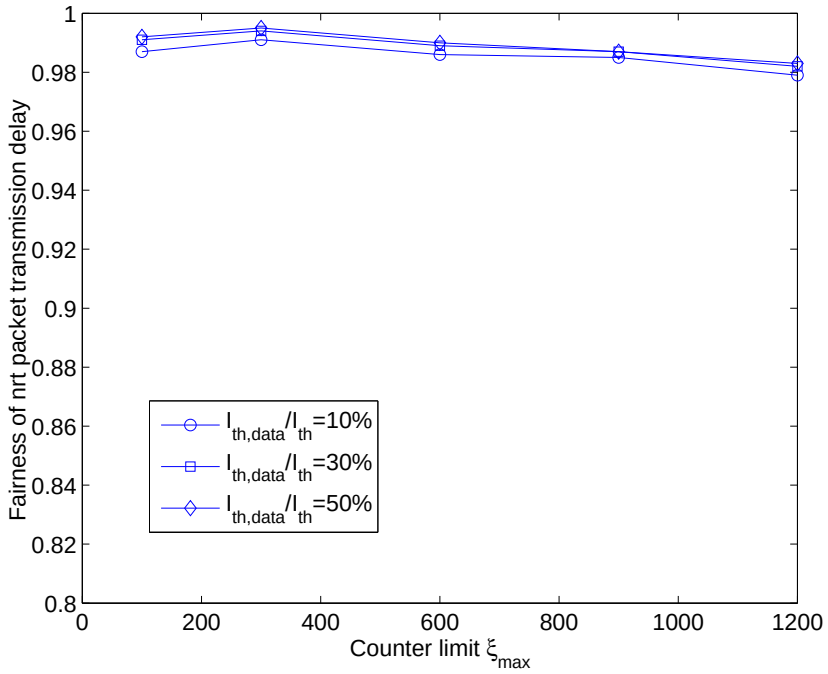

Figure 4.7: Fairness of transmission delay for admitted nrt connections

In generating all previous results, we considered a fixed number of primary connections. Next we consider a more dynamic scenario in the primary network, where each MS can have two states. When the MS generates a new connection, its state changes from "off" to "on". When an existing connection ends, the MS's state changes from "on" to "off". In the simulation, we fix the probability from "on" to "off" and vary the probability from "off" to "on". Fig. 4.10 shows the connection dropping rate. We can see that it is possible to achieve very low connection dropping rate. Using the weight-based criterion achieves slightly lower connection dropping rate than the other two criteria, because this criterion takes not only interference between the two networks, but also interference within the CRN. The criterion that drops the most recent admitted connection does not consider the interference condition and therefore

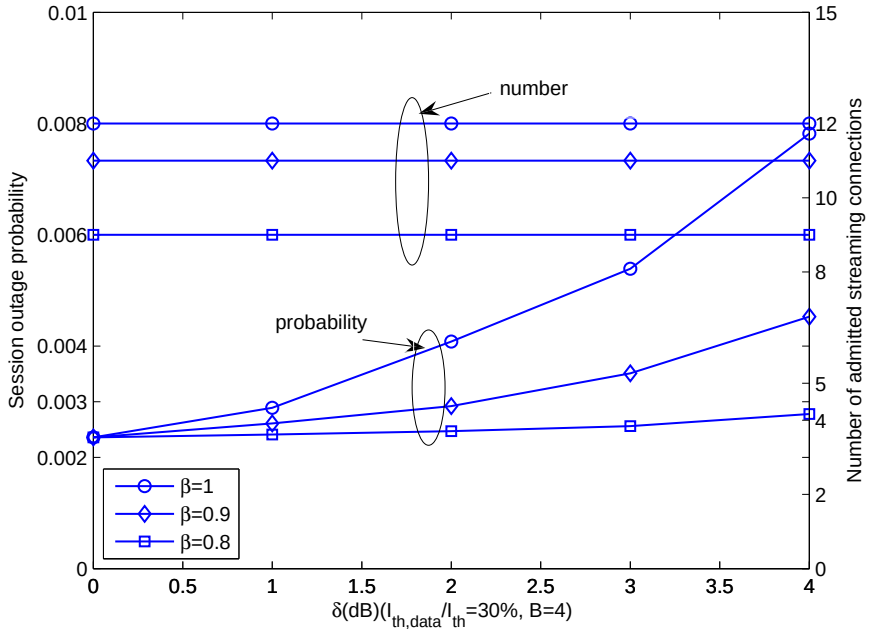

Figure 4.8: Effect of inaccurate measurement on streaming traffic performance

results in the highest dropping rate, and the criterion based on primary to secondary interference achieves drop rate performance between the other two criteria.

4.6 Conclusions

We have proposed an admission control scheme for streaming traffic and non-real-time data traffic and designed scheduling schemes for each type of the traffic. By using static channel and interference conditions, the admission control scheme effectively limits amount of the traffic admitted into the system to be below capacity. The heuristic scheduling scheme opportunistically takes advantage of the random channel and interference conditions and achieves low and fair session outage probability for the streaming traffic and high throughput and fair transmission delay for the data

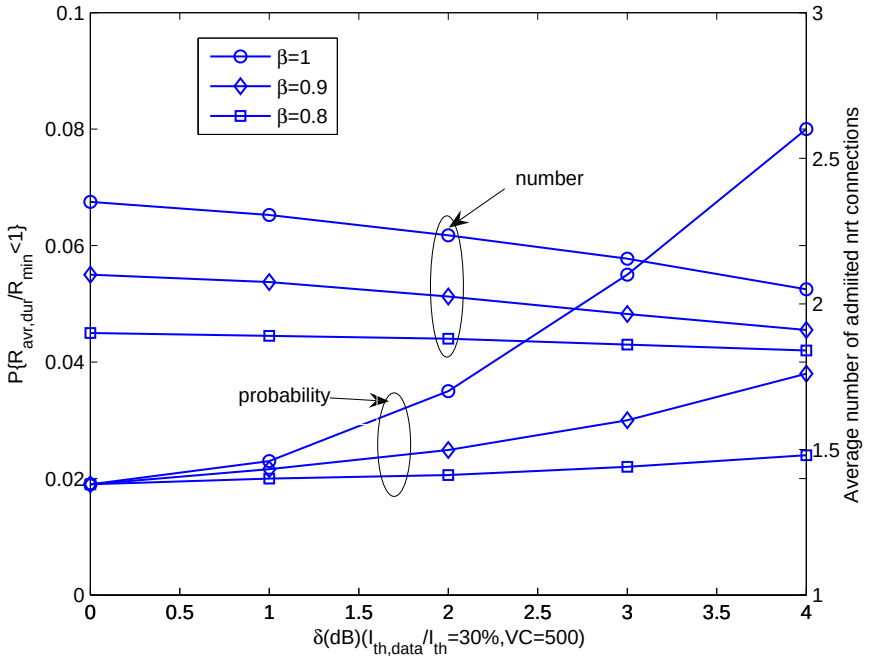

Figure 4.9: Effect of inaccurate measurement on nrt traffic performance

traffic, and its performance is close to the optimum scheduling. Extending the current work to multihop ad hoc network is underway.

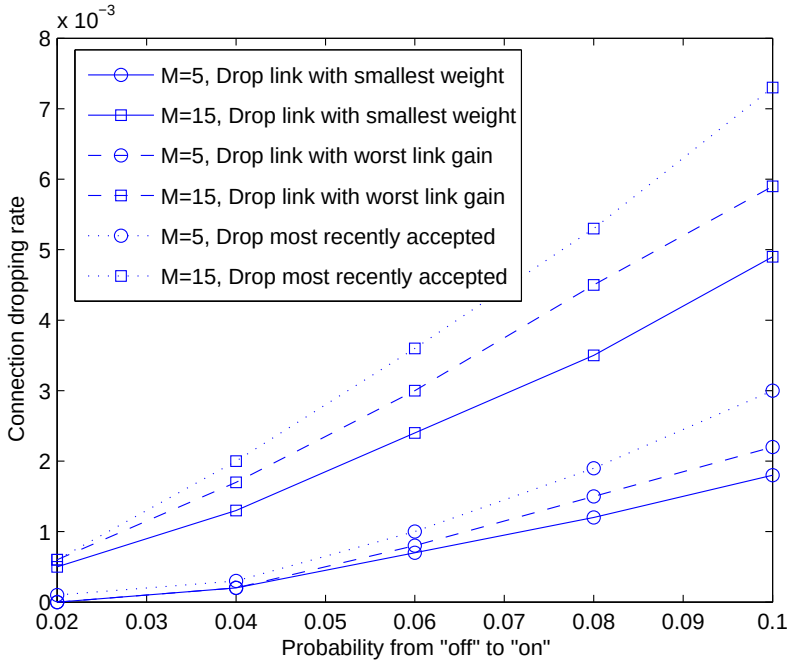

Figure 4.10: Connection dropping rate

Chapter 5

RRM in CDMA-based Cooperative Networks

5.1 Introduction

Using cooperative relaying is another approach to improving the spectrum efficiency and the transmission quality of the users. As mentioned in section 1.4.2, most of the work on cooperative communication networks available in the literature is based on orthogonal transmissions among different links, where different transmitter-receiver pairs use different frequency channels. In this chapter, we study the power distribution of a CDMA-based cellular network with cooperative relaying, where all the transmissions share the same spectrum. Two forwarding techniques, i.e., decode-and-forward (DF) and amplify-and-forward (AF), are considered. An optimization problem is formulated for the uplink and downlink, respectively. The objective is to minimize the total transmission power of the stations, subject to the average transmission rate and SINR requirement of the user traffic. A heuristic scheme is

designed for distributing transmission powers of the SSs and RSs in each direction. Transmission power and outage performance is examined for both the optimum and heuristic solutions with different RS selection criteria. We find that very significant performance improvement can be achieved using cooperative relaying. The remaining part of this paper is organized as follows. In Section 5.2 we describe the system model that this work is based on. The optimum power distribution solutions for the uplink and downlink are presented in Section 5.3 and Section 5.4, respectively. A heuristic power distribution scheme is presented in Section 5.5. RS selection criteria are described in Section 5.6. Numerical results are shown in Section 5.7. Section 5.8 concludes the paper.

5.2 System Description

We consider a cellular CDMA network, where different frequency bands are used for the uplink and downlink, so that there is no interference between the uplink and downlink transmissions. Each BS is associated with several MSs, each of which is equipped with a single radio transceiver with an omni-directional antenna. The maximum transmission power for an MS is $P_{\mathrm{MS,max}}$, and that for the BS is $P_{\mathrm{BS,max}}$. We consider cooperative transmissions in both the uplink and downlink. In the uplink, MSs carrying their own connections are the SSs, and the BS is the DS for all connections. In the downlink, the BS is the SS for all connections and the DSs are MSs. MSs that do not carry active traffic are available for cooperatively relaying packets for their peer stations.

Each BS transmits a pilot signal at a constant power. An MS measures the channel gain between itself and its associated BS according to the received pilot signal strength

and reports the channel gain to the BS. The link gains can be used by the BS to select RSs and to distribute the transmission power. We use $G_{i,j}$ to represent the link gain between station i and station j, where a station can be either the BS or an MS, and the MS can be an SS (in the uplink), RS, or DS (in the downlink). We further assume that the link gains are reciprocal, i.e., for any two stations i and j, $G_{i,j} = G_{j,i}$.

For capacity analysis, we consider that each connection generates packets at a constant rate R bits per second. All packets have the same size and bit error rate (BER) requirement. A packet is received successfully if the average BER is less than or equal to the target BER. For given transceiver design, the BER can be converted into a minimum SINR requirement, γ^*. Channel time is divided into equal length time slots, each of which is for one packet transmission. We assume that channel fading is relatively slow so that the channel condition can be considered time-invariant over a period of time which is much longer than the packet transmission time.

We consider that one RS is selected for each connection based on the RS selection criterion. The RS selection schemes are presented in Section 5.6. Since the RS cannot transmit and receive at the same time, it receives from the SS in the odd time slots and forwards to the DS in the even time slots. The DS receives from the SS in the odd time slots and from the RS in the even time slots, and combines the signals at the end of each even time slot using maximum ratio combining (MRC). As a result, the SS only transmits at the odd time slots, i.e., it is idle for half of the time. Therefore, the instantaneous data transmission rate of the source should be $2R$ in order to guarantee the average service rate of R to the end user. We consider a synchronized case as follows. In the odd time slots, all SSs transmit while both the DS and the RSs are receiving, and in the even time slots all RSs forward the received signals from the SSs

to the BS. We consider both two relay forwarding techniques: decode-and-forward (DF) and amplify-and-forward (AF).

5.3 Power Distribution in the Uplink

We consider N^u connections in the uplink. The MS carrying the ith connection in the uplink is denoted as the ith MS. We use r_i^u to represent the RS of connection i in the uplink, P_i^u to denote the transmission power of SS i, and $P_{r_i^u}$ the transmission power of the RS of connection i. Let $\gamma_{s,i}^u$ represent the received SINR of the signal at the BS receiver input for a packet from SS i, $\gamma_{c,i}^u$ the received SINR of the signal at the RS receiver input for the same packet, and $\gamma_{r,i}^u$ the received SINR at the BS receiver input for a packet forwarded from the RS. Below we formulate an optimum power distribution problem for both DF and AF with an objective to minimize the total transmission power of all MSs, including both the SSs and RSs, subject to the SINR requirement of the connections and maximum station transmission power.

We consider a typical connection in the uplink. In the odd time slots, the RS and DS receive interference from transmissions of SSs of all other connections. The received SINR for a packet transmitted from SS i at the BS after de-spreading can be found as

$$\gamma_{s,i}^u = \frac{W}{2R} \frac{P_i^u G_{b,i}}{\sum_{j=1}^{N^u} P_j^u G_{b,j} - P_i^u G_{b,i} + \eta}. \tag{5.1}$$

for $i = 1, 2, \ldots, N^u$, where W is the spread spectrum bandwidth, b is the BS that SS i is associated to, and η is the power of the background additive white Gaussian noise. Similarly, the received SINR of the same signal at the RS after de-spreading

is given by

$$\gamma_{c,i}^{u} = \frac{W}{2R} \frac{P_i^u G_{i,r_i^u}}{\sum_{j=1}^{N^u} P_j^u G_{j,r_i^u} - P_i^u G_{i,r_i^u} + \eta}. \tag{5.2}$$

For DF, $\gamma_{c,i}^{u}$ should be larger than or equal to γ^* so that the RS can successfully decode the received signal from the SS before forwarding it to the DS. When the packet is forwarded from the RS to the DS, it experiences interference from transmissions of all other RSs, and its SINR at the BS receiver after de-spreading is

$$\gamma_{r,i}^{u} = \frac{W}{2R} \frac{P_{r_i^u} G_{b,r_i^u}}{\sum_{j=1}^{N^u} P_{r_j^u} G_{b,r_j^u} - P_{r_i^u} G_{b,r_i^u} + \eta}. \tag{5.3}$$

For AF, successful decoding is not required at the RS, which simply amplifies the received analog signal together with interference and noise and forwards to the BS. Therefore, the expression of $\gamma_{r,i}^{u}$ should take the amplified interference into consideration and is given by

$$\gamma_{r,i}^{u} = \frac{W}{2R} \frac{P_{r_i^u} G_{b,r_i^u} \beta_i^u}{\sum_{j=1}^{N^u} P_{r_j^u} G_{b,r_j^u} - P_{r_i^u} G_{b,r_i^u} \beta_i^u + \eta}, \tag{5.4}$$

where $\beta_i^u = \frac{\gamma_{c,i}^u}{\gamma_{c,i}^u + 1}$.

At the end of each even time slot, the BS combines the signals from the SS and the RS for each connection, and the combined SINR after MRC is $\gamma_i^u = \gamma_{s,i}^u + \gamma_{r,i}^u$ for connection i. To ensure successful transmission, $\gamma_i^u \geq \gamma^*$.

Given the above analysis, an optimum power distribution problem can be formulated as follows:

$$\text{Problem I:} \quad \min \sum_{i=1}^{N^u} \left(P_i^u + P_{r_i^u}\right) \tag{5.5}$$

$$\text{s.t.} \quad \gamma_{s,i}^u + \gamma_{r,i}^u \geq \gamma^*, \ i = 1, 2, \ldots, N^u \tag{5.6}$$

$$\gamma_{c,i}^u \geq \gamma^*, \ i = 1, 2, \ldots, N^u, \text{ for DF only} \tag{5.7}$$

$$0 \leq P_i^u \leq P_{\text{MS,max}}, \ i = 1, 2, \ldots, N^u \tag{5.8}$$

$$0 \leq P_{r_i^u} \leq P_{\text{MS,max}}, \ i = 1, 2, \ldots, N^u \tag{5.9}$$

where $\gamma_{s,i}^u$ and $\gamma_{c,i}^u$, respectively, are given by (5.1) and (5.2), and $\gamma_{r,i}^u$ is given by (5.3) for DF and (5.4) for AF. Note that the SSs and the RSs do not transmit at the same time, therefore the objective function is equivalent to minimizing the total SS transmission power in the odd time slots and the total RS transmission power in the even time slots. If there is no solution to the optimization problem, one or more connections should be removed so that a solution can be found for the remaining connections. The removed connections are in outage.

Solving Problem I is not trivial. The constraint in (5.6) is a non-linear function of P_i^u's and $P_{r_i^u}$'s. It can be proved that Problem I is non-convex. The problem cannot be directly solved by commonly used optimization software. With some simple mathematical manipulations we can use the iterative method proposed in [128] to solve the problem. The method is summarized in Appendix. Let $A_i = P_i^u G_{b,i}$, $\bar{A}_i = \sum_{j=1,j\neq i}^{N^u} P_j^u G_{b,j}$, $B_i = P_{r_i^u} G_{b,r_i^u}$, $\bar{B}_i = \sum_{j=1,j\neq i}^{N^u} P_{r_j^u} G_{b,r_j^u}$, and $C = \frac{2R\gamma^*}{W}$. Then

we have $\gamma_{s,i}^u = \frac{W}{2R}\frac{A_i}{\bar{A}_i+\eta}$ and $\gamma_{r,i}^u = \frac{W}{2R}\frac{B_i}{\bar{B}_i+\eta}$. Then constraint (5.6) can be rewritten as

$$\frac{A_i}{\bar{A}_i+\eta} + \frac{B_i}{\bar{B}_i+\eta} \geq C, \tag{5.10}$$

which can be further rewritten as

$$\frac{C\bar{A}_i\bar{B}_i + \eta C\bar{A}_i + \eta C\bar{B}_i + \eta^2 C}{A_i\bar{B}_i + B_i\bar{A}_i + \eta A_i + \eta B_i} \leq 1. \tag{5.11}$$

In (5.11) both the numerator and the denominator on the left-hand side are posynomials of P_i^u's and $P_{r_i^u}$'s. Replacing (5.6) in Problem I with (5.11), we can use the iterative method in Appendix A to solve.

5.4 Optimum Power Distribution in the Downlink

In the downlink the BS is the SS for all connections. Let P_i^d denote the transmission power of the BS for connection i, then $P_b = \sum_{i=1}^{N^d} P_i^d$ is the total transmission power of the BS, where N^d is the total number of connections in the downlink. We use r_i^d to represent the RS for connection i in the downlink, and $P_{r_i^d}$ the transmission power of the RS. Then $\sum_{j=1}^{N^d} P_{r_j^d}$ is the total transmission power of all RSs. The objective of the power distribution is to guarantee the SINR requirement of the connections, while minimizing both the BS and the RS transmission power. Below we use $\gamma_{s,i}^d$ to represent the received SINR of the signal at the DS for a packet from the BS for connection i, $\gamma_{c,i}^d$ the received SINR at the RS for the same packet, and $\gamma_{r,i}^d$ the received SINR at the DS for the packet forwarded from the RS.

The objective of power distribution in the downlink is to minimize $P_b + \sum_{i=1}^{N^d} P_{r_i^d}^d$.

Since the BS and the RSs do not transmit at the same time, the objective function is equivalent to minimizing the total BS transmission power in the odd time slots and the total RS transmission power in the even time slots. At the odd time slots when the BS (SS) transmits, the received SINR of the signal for connection at the DS receiver after de-spreading is given by

$$\gamma_{s,i}^{d} = \frac{W}{2R} \frac{P_i^d G_{b,i}}{P_b G_{b,i} - P_i^d G_{b,i} + \eta}, \tag{5.12}$$

and the same packet also reaches the RS of connection i with the received SINR (after de-spreading) given by

$$\gamma_{c,i}^{d} = \frac{W}{2R} \frac{P_i^d G_{b,r_i^d}}{P_b G_{b,r_i^d} - P_i^d G_{b,r_i^d} + \eta}. \tag{5.13}$$

For DF, the RS should decode the received packet, re-encode, and forward to the DS. That is, $\gamma_{c,i}^{d} \geq \gamma^*$ should hold. At the DS receiver, the forwarded signal from the RS experiences interference from transmissions of all other RSs. Assuming the RS can correctly decode the signal from the BS, the SINR at the DS receiver for connection i after de-spreading is given by

$$\gamma_{r,i}^{d} = \frac{W}{2R} \frac{P_{r_i^d} G_{r_i^d,i}}{\sum_{j=1}^{N^d} P_{r_j^d} G_{r_j^d,i} - P_{r_i^d} G_{r_i^d,i} + \eta}. \tag{5.14}$$

For AF, the RS does not decode the received signal, but amplifies the received analog signal, interference, and noise and forwards to the DS. The SINR of the forwarded

signal at the DS after de-spreading is given by

$$\gamma_{r,i}^{d} = \frac{W}{2R} \frac{\beta_i^d P_{r_i^d} G_{r_i^d,i}}{\sum_{j=1}^{N^d} P_{r_j^d} G_{r_j^d,i} - \beta_i^d P_{r_i^d} G_{r_i^d,i} + \eta}, \tag{5.15}$$

where $\beta_i^d = \frac{\gamma_{c,i}^d}{\gamma_{c,i}^d+1}$.

At the end of each even time slot, the DS combines the signal from the BS in the previous time slot and the signal from the RS in the current time slot, and the combined SINR after MRC is $\gamma_i^d = \gamma_{s,i}^d + \gamma_{r,i}^d$ for connection i.

Similar to that for the uplink, the optimum power distribution problem for the downlink is then formulated as follows:

$$\text{Problem II:} \quad \min \left(P_b + \sum_{i=1}^{N^d} P_{r_i^d} \right) \tag{5.16}$$

$$\text{s.t.} \quad \gamma_{s,i}^d + \gamma_{r,i}^d \geq \gamma^*, \ i = 1, 2, \ldots, N^d \tag{5.17}$$

$$\gamma_{c,i}^d \geq \gamma^*, \ i = 1, 2, \ldots, N^d, \text{ for DF only} \tag{5.18}$$

$$0 \leq P_{r_i^d} \leq P_{\text{MS,max}}, \ i = 1, 2, \ldots, N^d \tag{5.19}$$

$$0 \leq P_b \leq P_{\text{BS,max}} \tag{5.20}$$

where $\gamma_{s,i}^d$ and $\gamma_{c,i}^d$, respectively, are given by (5.12) and (5.13), and $\gamma_{r,i}^d$ is given by (5.14) for DF and (5.15) for AF. A similar method as in Section 5.3 can be used to convert Problem II into a format that can be solved by the iterative method in Appendix. Communication outage occurs if the optimization problem does not have a feasible solution.

5.5 Heuristic Power Distribution Schemes

From the optimum problem formulation in Section 5.3 we find that for the uplink, the signals received at the RS of SS i experience interference from transmissions of all other SSs. Therefore, to solve the optimization problem, the BS should know link gains from all SSs to all RSs. This information cannot be easily obtained by the BS. The same difficulty also exists in the downlink when implementing the optimum power distribution solution since it requires link gains from all RSs to all DSs.

A practical solution is that the SS-DS link and the RS-DS link each have a fixed SINR target so that dynamic power control is possible, for example, each of the two links can have a fixed share of the total required SINR. A simple scheme is to determine the shares based on their link gains to the BS. In the uplink, we can have

$$\frac{\gamma_{s,i}^u}{\gamma_{r,i}^u} = \frac{G_{b,i}}{G_{b,r_i^u}},\ i = 1, 2, \ldots, N^u, \text{ and} \tag{5.21}$$

$$\gamma_{s,i}^u + \gamma_{r,i}^u = \gamma^*,\ i = 1, 2, \ldots, N^u. \tag{5.22}$$

Based on this policy, the target SINRs for the SS-DS link and the RS-DS link are $\frac{G_{b,i}}{G_{b,i}+G_{b,r_i^u}}\gamma^*$ and $\frac{G_{b,r_i^u}}{G_{b,i}+G_{b,r_i^u}}\gamma^*$, respectively. Assuming the dynamic power control converges and the power control is perfect, the transmission power of each SS and RS can be found from the equation set of (5.23) and (5.24) for DF and the equation set

of (5.23) and (5.25) for AF.

$$\gamma_{s,i}^{u}: \quad \frac{W}{2R}\frac{P_i^u G_{b,i}}{\sum_{j=1}^{N^u} P_j^u G_{b,j} - P_i^u G_{b,i} + \eta} = \frac{G_{b,i}\gamma^*}{G_{b,i} + G_{b,r_i^u}}, \tag{5.23}$$

$$\gamma_{r,i}^{u} \text{ for DF:} \quad \frac{W}{2R}\frac{P_{r_i^u} G_{b,r_i^u}}{\sum_{j=1}^{N^u} P_{r_j^u} G_{b,r_j^u} - P_{r_i^u} G_{b,r_i^u} + \eta} = \frac{G_{b,r_i^u}\gamma^*}{G_{b,i} + G_{b,r_i^u}}, \tag{5.24}$$

$$\gamma_{r,i}^{u} \text{ for AF:} \quad \frac{W}{2R}\frac{P_{r_i^u} G_{b,r_i^u}\beta_i^u}{\sum_{j=1}^{N^u} P_{r_j^u} G_{b,r_j^u} - P_{r_i} G_{b,r_i^u}\beta_i^u + \eta} = \frac{G_{b,r_i^u}\gamma^*}{G_{b,i} + G_{b,r_i^u}}, \tag{5.25}$$

for $i = 1, 2, \ldots, N^u$. The solution to the above equation system is feasible if all the following conditions are satisfied: i) $0 \leq P_i^u \leq P_{\text{MS,max}}$ for all $i = 1, 2, \ldots, N^u$, and ii) $0 \leq P_{r_i^u} \leq P_{\text{MS,max}}$ for all $i = 1, 2, \ldots, N^u$. Note that the power distribution does not guarantee the SINR of the SS-RS link. For DF, the solution should satisfy an additional condition, $\gamma_{c,i}^u \geq \gamma^*$, for successful decoding at the RS.

A similar heuristic power distribution solution can be designed for the downlink. That is, the total required SINR of each downlink connection is split between the BS-DS and RS-DS links based on their link gains. The SINR shares of the two links are given by

$$\frac{\gamma_{s,i}^d}{\gamma_{r,i}^d} = \frac{G_{b,i}}{G_{r_i^d,i}}, \ i = 1, 2, \ldots, N^d, \text{ and} \tag{5.26}$$

$$\gamma_{s,i}^d + \gamma_{r,i}^d = \gamma^*, \ i = 1, 2, \ldots, N^d, \tag{5.27}$$

and the transmission power of the BS and RS for each connection can be solved using the equation set in (5.28) and (5.29) for DF and (5.28) and (5.30) for AF, assuming

perfect power control.

$$\gamma_{s,i}^{d}: \quad \frac{W}{2R}\frac{P_i^d G_{b,i}}{P_b G_{b,i} - P_i^d G_{b,i} + \eta} = \frac{G_{b,i}\gamma^*}{G_{b,i} + G_{r_i^u,i}}, \tag{5.28}$$

$$\gamma_{r,i}^{d} \text{ for DF:} \quad \frac{W}{2R}\frac{P_{r_i^d} G_{r_i^d,i}}{\sum_{j=1}^{N^d} P_{r_j^d} G_{r_j^d,i} - P_{r_i^d} G_{r_i^d,i} + \eta} = \frac{G_{r_i^d,i}\gamma^*}{G_{b,i} + G_{r_i^d,i}}, \tag{5.29}$$

$$\gamma_{r,i}^{d} \text{ for AF:} \quad \frac{W}{2R}\frac{\beta_i^d P_{r_i^d} G_{r_i^d,i}}{\sum_{j=1}^{N^d} P_{r_j^d} G_{r_j^d,i} - \beta_i^d P_{r_i^d} G_{r_i^d,i} + \eta} = \frac{G_{r_i^d,i}\gamma^*}{G_{b,i} + G_{r_i^d,i}}, \tag{5.30}$$

The solution is feasible if i) $0 \leq \sum_{i=1}^{N^d} P_i^d \leq P_{\text{BS,max}}$, and ii) $0 \leq P_{r_i^d} \leq P_{\text{MS,max}}$ for $i = 1, 2, \ldots, N^d$. For DF, the solution should satisfy an additional condition, $\gamma_{c,i}^{d} \geq \gamma^*$ for $i = 1, 2, \ldots, N^d$.

5.6 RS Selection

The selection of RSs can affect the cooperative system performance. In order to find the RS that results in the best performance (e.g., in terms of total transmission power or outage probability) for a particular connection, link conditions such as SINR of potential RSs should be known. However, the link conditions depend on transmission powers of the stations, which are dependent on selected RSs of the considered connection and all other connections in the system. Therefore, the RS selection of all connections are correlated and the optimum RS selection requires information of all SS-DS links and all potential SS-RS and RS-DS links. Below we present an RS selection scheme that decouples the RS selection among different connections.

Theoretically, all idle MSs can be potential RSs for a connection. In order to save time and signaling exchange overhead for RS selection, only a small number of potential RSs are selected for each connection, and one of them is selected as the

RS forwarding traffic for the connection. In the uplink, a certain number of idle MSs closest to the SS are selected as the potential RSs for each connection. In the downlink, the same number of idle MSs closest to the DS are selected as the potential RSs for each connection. If by coincidence the same idle MS is selected by two connections as their potential RS, then the idle MS becomes potential RS for one of the connections randomly.

We first describe the SINR-based RS selection for the uplink. Initially, each SS transmits a packet assuming that cooperative relaying does not exist. The objective of the SS transmissions is to achieve the target SINR γ^* for each link, and the power control is performed in the same way as in traditional CDMA networks without relaying. During this period, potential RSs listen to the transmission from their corresponding SS and measure received SINRs. At the end of the packet transmission, the potential RSs with received SINRs satisfying a certain condition report to the BS. When using DF, the potential RSs that can correctly decode the signal from their SS report to the BS. Although correct decoding is not a requirement for AF, we set an SINR threshold, $\gamma_{th} > 0$, so that only the potential RSs that receive the SS's signal with an SINR above the threshold report to the BS. This is to avoid the case when the SINR at the RS is low and the RS amplifies most interference and noise instead of the desired signal. The effect of the threshold on the system performance will be discussed in Section 4.5. Among all the potential RSs that have reported to the BS, the one with the best RS-BS link gain is selected as the RS, which is then notified through the downlink transmissions and starts relaying for its SS.

After the initial RS selection, cooperative communications are established between the SS and the RS. While the selected RS is relaying traffic for the SS, all other

potential RSs keep listening to the SS. As the physical channel conditions change with random fading, the SINR of the current RS may drop below the threshold and should be replaced. In this case, the SS sends a message requesting for a new RS. All potential RSs then report to the BS if their received SINRs satisfying the SINR requirements, and the new RS is selected as the one with the best SINR. Since the fast fading can be well mitigated by the transceiver design, its effect on the RS selection is negligible. Therefore, the overhead for reselecting the RSs is relatively small.

The RS selection process in the downlink works in a similar way. Initially, the BS transmits a packet for each connection assuming that there is no relaying. The transmission power to each DS is to achieve the target SINR for their connections. At the end of the packet transmission, potential RSs measure received SINRs and report to the BS. This process is the same as in the uplink. Upon receiving the reports from the potential RSs, the BS selects the one with the best RS-DS link gain as the RS and notifies the RS through the downlink transmissions. The potential RSs keep listening to the SS. RSs are reselected as their SINRs are not satisfactory. If there is no potential RS satisfying the required SINR condition for certain connections, the SS transmits in both the odd and even time slots.

For comparison, we also present three simple criteria for RS selection based on link gains. Compared to the SINR-based RS selection, these link gain based RS selection criteria are less dynamic and require smaller signaling overhead. We first describe the link gain based RS selection for the uplink. In the first criterion, the potential RS with the best link gain between itself and the SS among all the potential RSs is selected as the RS. That is, $r_i^u = \text{argmax}_{j \in C_i} G_{i,j}$, where C_i is a set of the potential RSs of SS i. The RS selection criterion is referred to as "the best first hop"

(i.e., the "first hop" is the SS-RS hop). The second one is referred to as "the best second hop", which selects the RS based on the link gains from the potential RSs to the BS, and $r_i^u = \text{argmax}_{j \in C_i} G_{j,b}$ (i.e., the "second hop" is the RS-DS hop). The last criterion takes into consideration both the first and second hop link gains, and $r_i^u = \text{argmax}_{j \in C_i} \min(G_{i,j}, G_{j,b})$. This is referred to as "the best worst hop". The link gain based RS selection for the downlink is similar, except the "first hop" is the BS-RS hop, and the second hop is the RS-DS hop.

5.7 Numerical Results

We consider a generic cell in a cellular CDMA network, where the BS is located at the center of the cell. The initial locations of the MSs are uniformly distributed in the cell coverage area. MSs move randomly using a random way point model. We simulate a slow fading channel with path loss and log-normal shadowing. Default parameters are listed in Table 5.1. Based on the parameters we can find that the channel coherence time is 27ms. Therefore the channel quasi-static duration is around 40 time slots (each time slot lasts for 2/3ms). The link gain for each channel is kept constant for a period of time which follows Gaussian distribution with mean of 40 time slots and standard deviation of 4 time slots. The link gain is regenerated when this period has expired. With 10m/s moving speed on average, an MS moves less than 0.3m in 40 time slots. Therefore, we can assume that the MS is static between two channel gain updates.

When using either the optimum or heuristic power distribution, if there is no feasible solution, we apply a simple removal scheme by removing the connection with the worst SS-DS link gain. The removed connection is in outage. This process is

repeated until a feasible solution can be found for all the remaining connections. A more complicated connection removal scheme can be used and a better outage probability may be achieved, but this is at a price of increased computational overhead and longer processing delay and is not considered in this paper. Below we examine the outage probability and transmission power performance using different forwarding techniques and RS selection criteria.

Table 5.1: Default Parameters Settings

Parameter	Value
Cell size	2km
Carrier frequency	2GHz
Spread spectrum bandwidth W	5MHz
Time slot duration	2/3ms
MS velocity	0 to 20m/s
Path loss exponent	4
Log-normal fading standard deviation δ	6dB
Background noise power η	10^{-8}W
Total number of MSs	140
Number of potential RSs per connection	6
Max. transmission power of MSs, $P_{\mathrm{MS,max}}$	0.5W
Max. transmission power of BSs, $P_{\mathrm{BS,max}}$	5W
SINR threshold γ^*	7dB
AF SINR threshold for RS selection, γ_{th}	4dB
Number of uplink connections N^u	10
Number of downlink connections N^d	10
Data rate R	32kbps

5.7.1 Results for the uplink

Figs. 5.1-5.5 show the performance in the uplink. Comparing the outage performance using the three RS selection criteria based on link gains, we can find from Fig. 5.1 that the "best worst hop" one achieves the best outage performance. This observation is

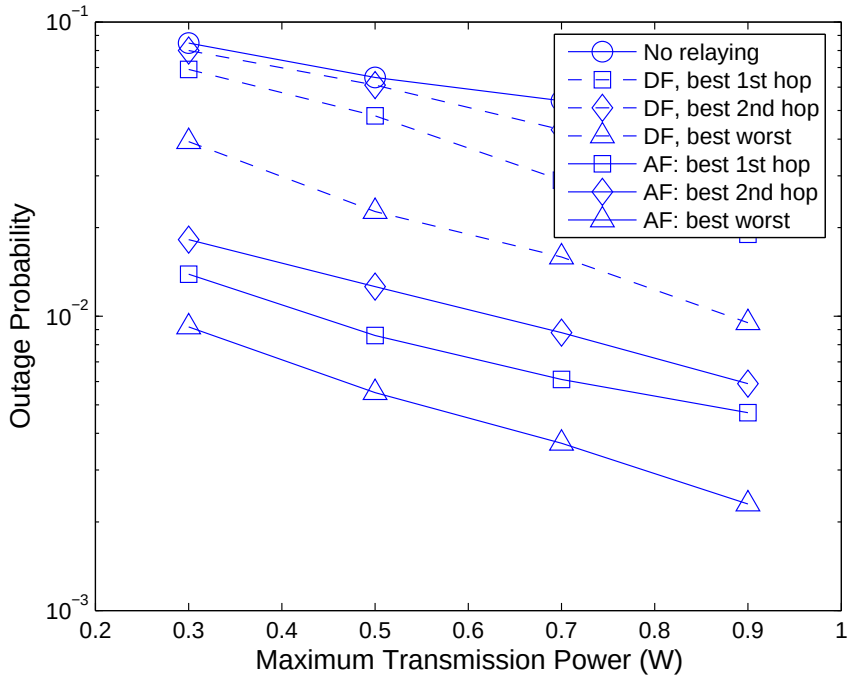

Figure 5.1: Uplink: outage performance using link gain based RS selection

consistent for both AF and DF. This is because the "best worst hop" one integrates the link conditions of both the SS-RS and RS-BS hops. On the other hand, the "best first hop" criterion can result in using an RS with extremely poor RS-BS link, and the "best second hop" one may select an RS with extremely poor SS-RS link. When this happens, very high transmission power is required from the SS or RS, causing high interference to other stations in the system. Therefore, below we only consider the "best worst hop" one for link-gain based RS selection.

It is also interesting to see from Fig. 5.1 that with the link-gain based RS selection criteria, using DF results in worse outage performance than using AF. This is due to the fact that, DF requires sufficient transmission power from the SS in order for the RS to correctly decode the received signal. The high transmission power causes

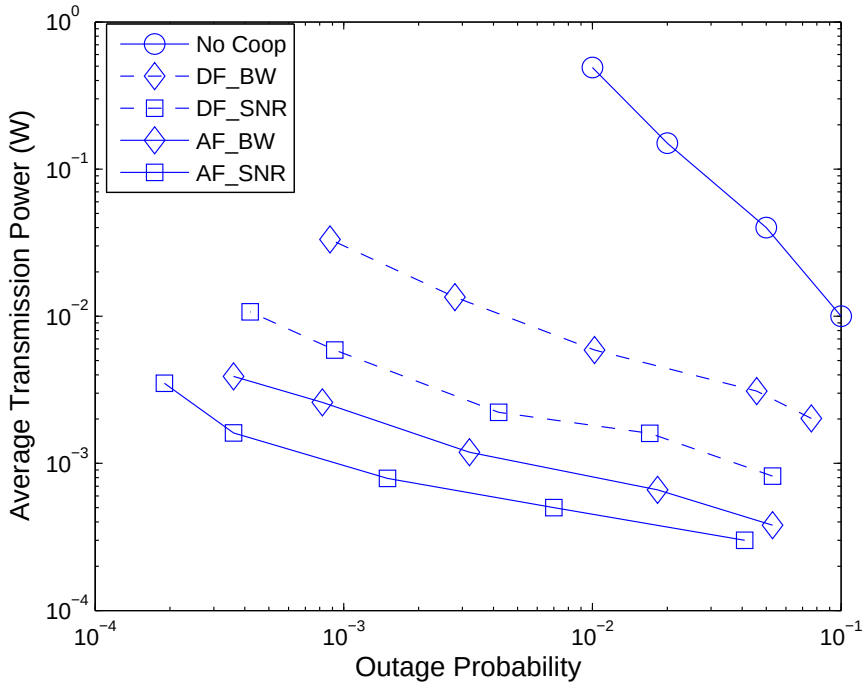

Figure 5.2: Uplink: transmission power vs. outage performance

high interference to the system and can negatively affect the overall performance. In contrast, using AF always leads to better outage performance than that without relaying. Although the RS using AF amplifies both the interference and noise together with the signal, less transmission power may be required from the SS using AF than using DF since no decoding is required in AF.

Fig. 5.2 compares the performance of cooperative relaying using the SINR-based RS selection and the "best worst hop" RS selection. It is shown that the SINR-based RS selection can achieve much better outage performance than the "best worst hop" one. It is also seen that both the "best worst hop" and the SINR-based RS selection can achieve better outage performance and significantly save transmission power, compared to without relaying.

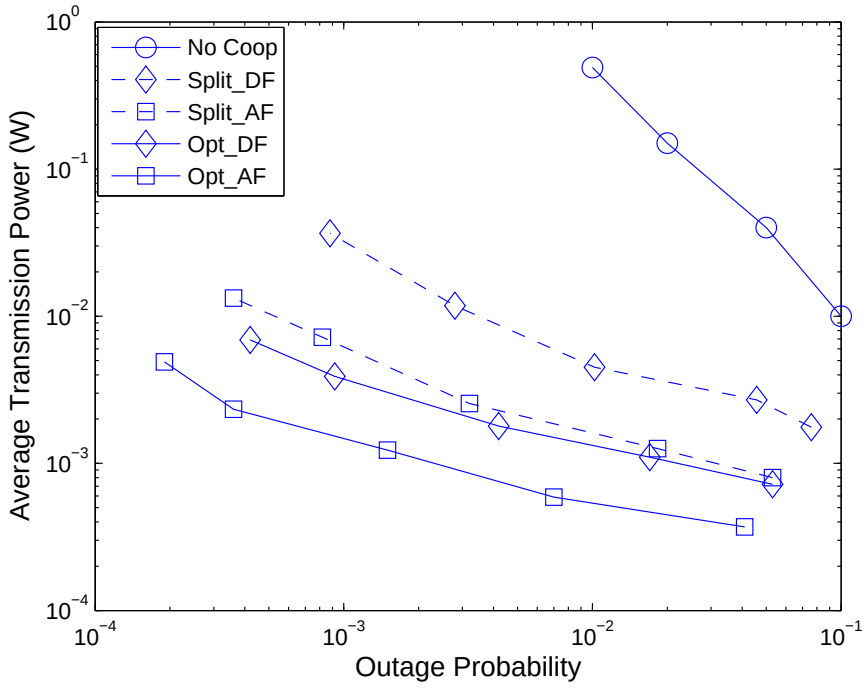

Figure 5.3: Uplink: comparison between the optimum and heuristic power allocation

All results in Figs. 5.1 and 5.2 are based on the optimum power distribution. In Fig. 5.3 we compare the performance of the optimum and heuristic power distribution using the SINR-based RS selection, where "scheme" represents the heuristic power distribution. The performance of the heuristic scheme is slightly worse than the optimum solution. Using the heuristic scheme, the link (either the SS-DS or the RS-DS link) with better link gain is allocated a larger fraction of the target SINR for each connection. However, the same link may suffer from high interference from other transmitters and require high transmission power.

Using cooperative relaying can significantly reduce the transmission power. This can be seen in both Figs. 5.2 and 5.3. Given a certain outage probability requirement, the vertical axis of each of the figures shows the total required transmission power

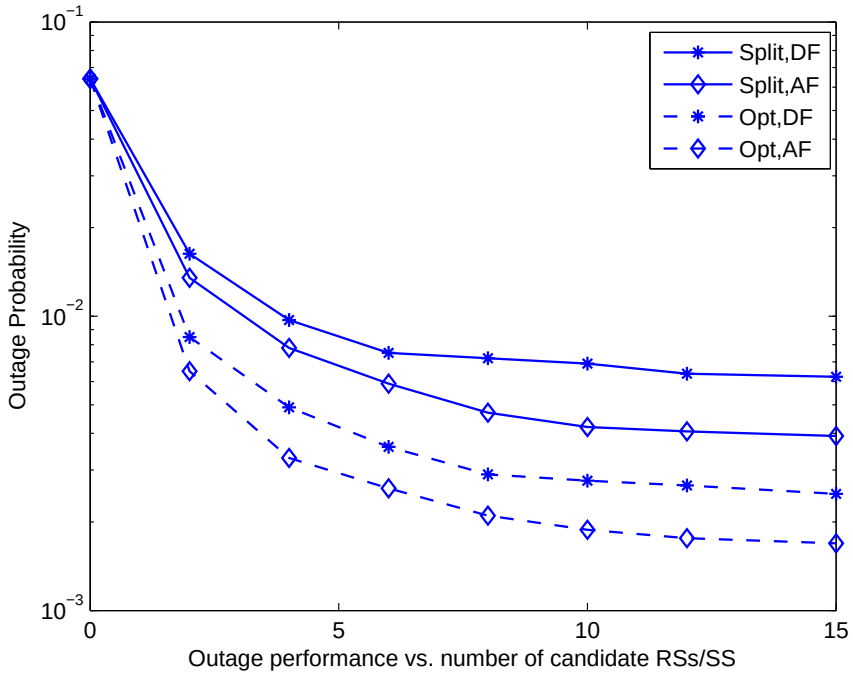

Figure 5.4: Uplink: outage performance vs. number of potential RSs per connection

from the SS and the RS for each connection. For comparison, the transmission power vs. outage curve for the original system without relaying is also shown. The transmission power is averaged over all connections. Although the number of transmitters in the system with relaying is increased, the total transmission power can be several magnitudes lower than that without relaying. As the outage probability becomes lower, the difference between required transmission power in the systems with and without cooperative relaying becomes larger. This can be a very attractive feature for encouraging mobile users to cooperatively relay traffic for one another. For a long term, every station can have its own traffic and should relay for its peers, eventually, they all benefit from the cooperative transmissions for improved communication outage probability and significantly reduced battery power consumption.

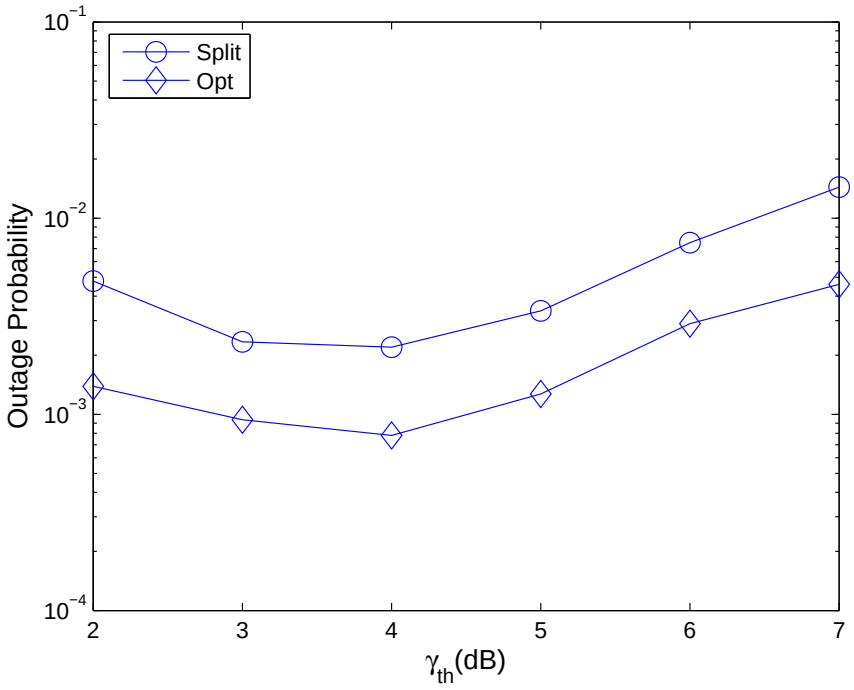

Figure 5.5: Uplink: effect of γ_{th} on outage performance using AF

Comparing Figs. 5.2 and 5.3 we find that the heuristic power distribution together with the SINR-based RS selection achieves similar performance as the optimum power distribution with the "best worst hop" RS selection. This observation is true for both AF and DF. Note that the results in Fig. 5.2 are based on optimum power distribution. That is, transmission powers of all SSs and RSs are jointly optimized. This requires a central station that knows the link gains from all SSs to all RSs, which may not be practical. For each connection, the actual SINR that each of the SS-DS and the RS-DS links contributes depends on both its own link gain and interference from other transmitters. Therefore, distributed power control cannot be used in this scenario. On the other hand, the heuristic power distribution scheme sets a clear SINR target for each of the SS-DS and RS-DS links, and the transmission power can be easily

adjusted by using distributed power control.

In Fig. 5.4 we show the outage probability by varying the number of potential RSs. For both the optimum and the heuristic power distribution, the outage performance is improved with the number of potential RSs. When the number of potential RSs per connection is relatively small, increasing the number can dramatically reduce the outage probability. However, when it exceeds a certain value, for example, 7 in the example shown, further increasing the number of potential RSs does not improve the outage probability, since the performance is then limited by the link condition of the SSs and other system parameters.

Fig. 5.5 shows the effect of γ_{th} on the outage performance in cooperative communications using AF. It is seen that for both the optimum and heuristic power distribution, there is an optimum value of γ_{th} that results in the lowest outage performance. This is due to the contradictory effects of γ_{th} in AF. When γ_{th} is very small, a selected RS may have poor link condition from the SS to itself. The poor receiving quality is then translated to poorer SINR after the RS forwards the amplified analog signal (together with interference and noise) to the BS. On the other hand, a large value of γ_{th} may prevent eligible RSs from being selected, resulting in communication outage. The optimum value of γ_{th} is between zero and γ^* in general and around 4dB in the example shown.

5.7.2 Results for the downlink

For the downlink we demonstrate transmission power of the BS (SS) and the RSs separately as normally they have different maximum transmission power limits. To save space, we do not show results of the "best first hop" and "best second hop" RS

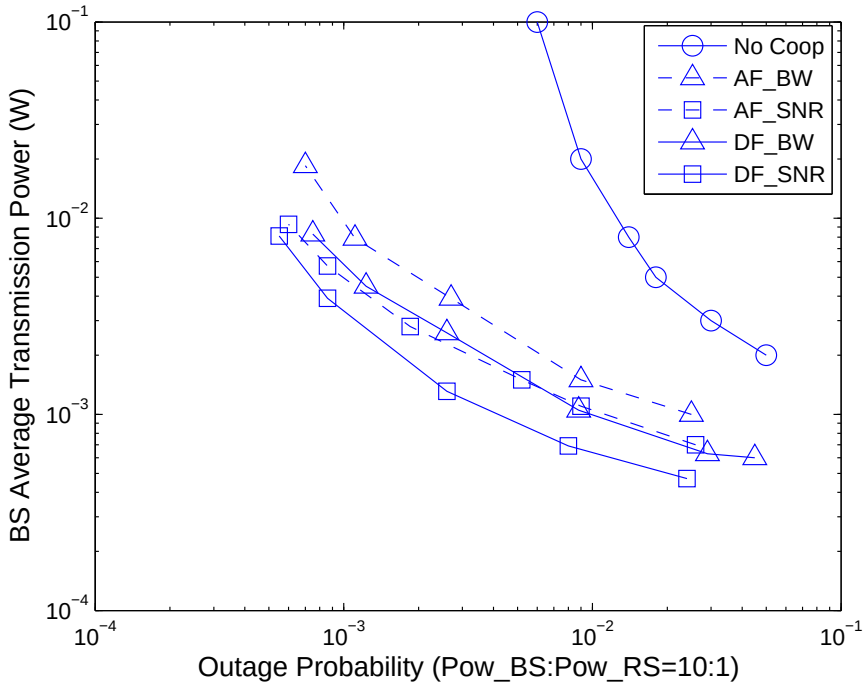

Figure 5.6: Downlink: comparison of BS transmission power

selection criteria as they perform worse than the "best worst hop" one in terms of both transmission power and outage probability.

Figs. 5.6 and 5.7 show the BS and RS transmission power, respectively, based on optimum power distribution. From the figures we find that using the SINR-based RS selection achieves better performance than the "best worst hop" one in both AF and DF. Fig. 5.6 shows that cooperative relaying can significantly save the BS transmission power. With cooperative relaying, the required BS transmission power can be several magnitudes lower when the outage probability is below 10^{-2}, and this improvement becomes higher as the outage probability is lower. Fig. 5.7 shows that the RS transmission power only increases very slightly when the required outage probability is reduced from 10^{-2} to 10^{-4}. With very low (< 10mW) RS transmission

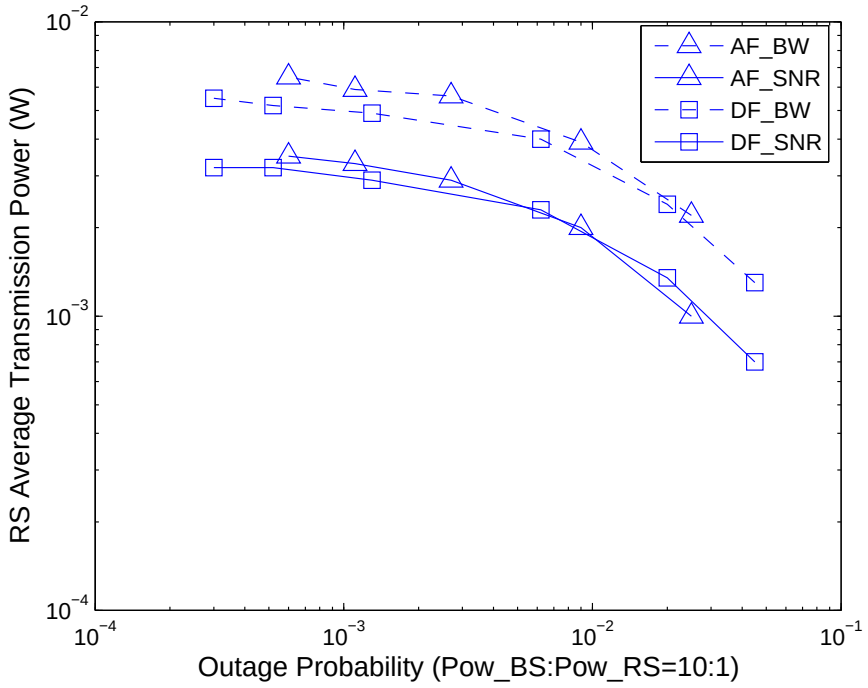

Figure 5.7: Downlink: comparison of RS transmission power

power, cooperative relaying in the downlink can significantly improve the outage probability for the same BS transmission power or reduce the BS transmission power for given outage requirement.

Figs. 5.6 and 5.7 also show that DF achieves slightly better performance than AF for the same RS selection criterion. This is because in the downlink the BS has better control over the transmission power to guarantee correct decoding at the RS as much as possible. Although similar procedure is also performed for AF, the performance is degraded due to amplified noise and interference.

Figs. 5.8 and 5.9 compare the performance of the link gain based power distribution with the optimum power distribution when using AF and the SINR-based RS

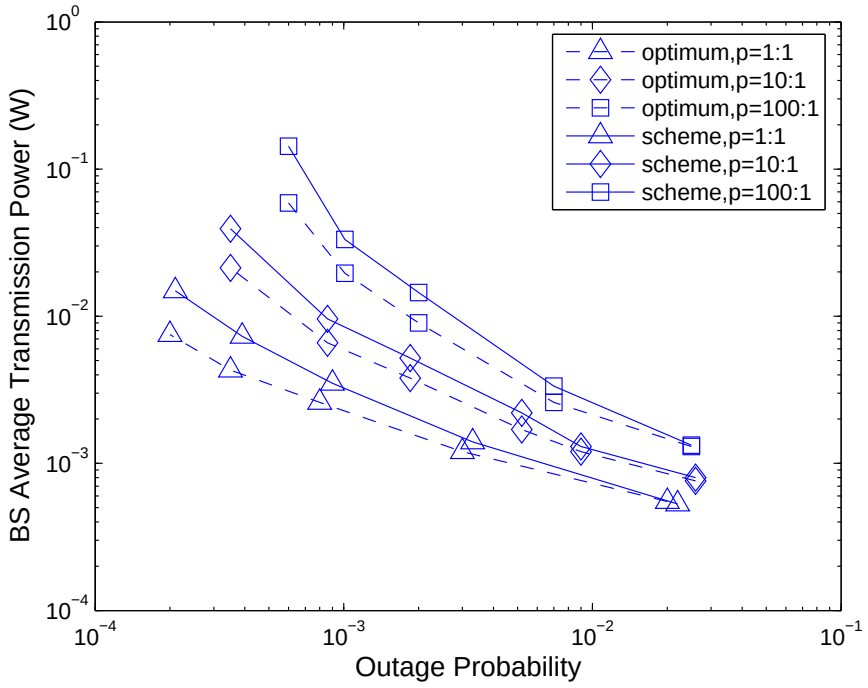

Figure 5.8: Downlink: BS transmission power using AF

selection, and similar comparison is done in Figs. 5.10 and 5.11 for DF. The performance of the heuristic scheme is slightly worse than the optimum due to its fixed allocation of the target SINR between the SS-DS and RS-DS links without considering the interference condition of the links. In the same figures we fix the maximum BS transmission power and vary the maximum RS (MS) transmission power, where $p = P_{\mathrm{BS,max}}/P_{\mathrm{MS,max}}$. It is seen that for both AF and DF, when the allowed RS (MS) transmission power increases, i.e., p is smaller, the required BS transmission power is reduced in order to achieve the same outage performance, or the outage probability can be reduced for the same BS transmission power.

We have also found that in the downlink, the heuristic power distribution together with SINR-based RS selection achieves similar performance as the optimum power

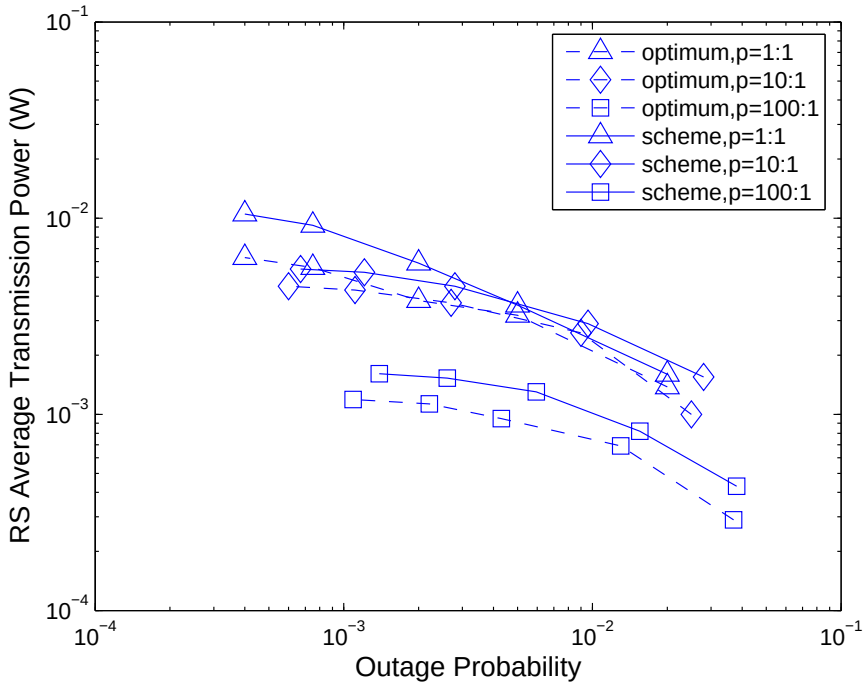

Figure 5.9: Downlink: RS transmission power using AF

distribution with the "best worst hop" RS selection, and this observation is consistent for both AF and DF.

Fig. 5.12 shows the effect of selecting different γ_{th} values in AF on the downlink outage probability, where we have similar observations as in the uplink. Fig. 5.13 shows that having more potential RSs can improve the outage performance. Comparing the figure with Fig. 5.4 we can find that the outage performance in the downlink does not improve as significantly as that in the uplink. This is due to the fact that the power allocation in the downlink can be better controlled at the BS, and therefore the requirement for better RS selection in the downlink is reduced, compared to that in the uplink.

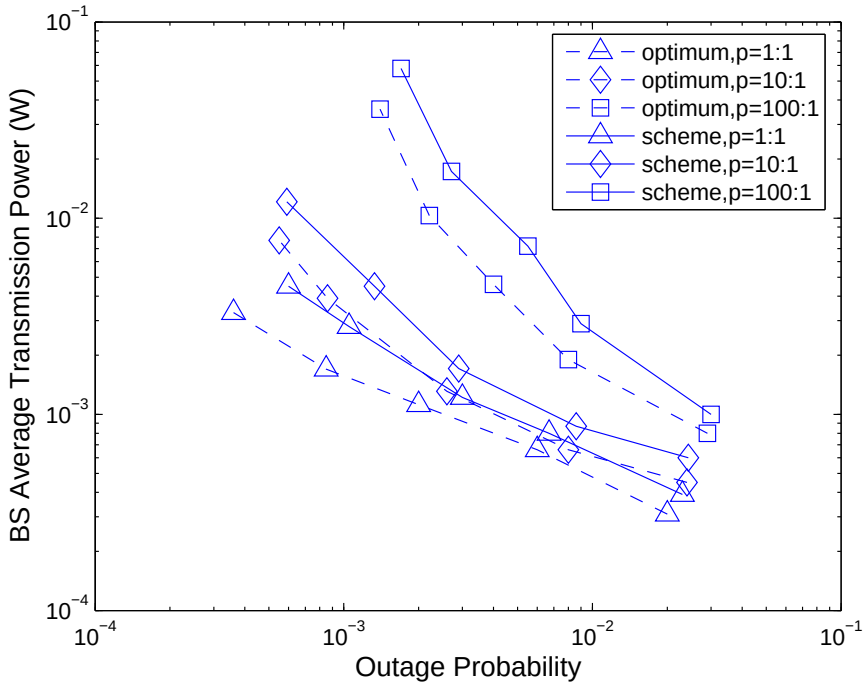

Figure 5.10: Downlink: BS transmission power using DF

5.8 Summary

We have studied power distribution of cooperative relaying in CDMA networks. Our results have shown that outage performance improvement and significant transmission power reduction is possible by using cooperative communications, provided appropriate forwarding techniques and RS selection criterion are used. The improved outage performance and reduced transmission power can be attractive features that encourage more stations to cooperate with one another, as this benefits all users in the system as well as the network providers. Our results also indicate that using AF is preferred in the uplink and DF is preferred in the downlink.

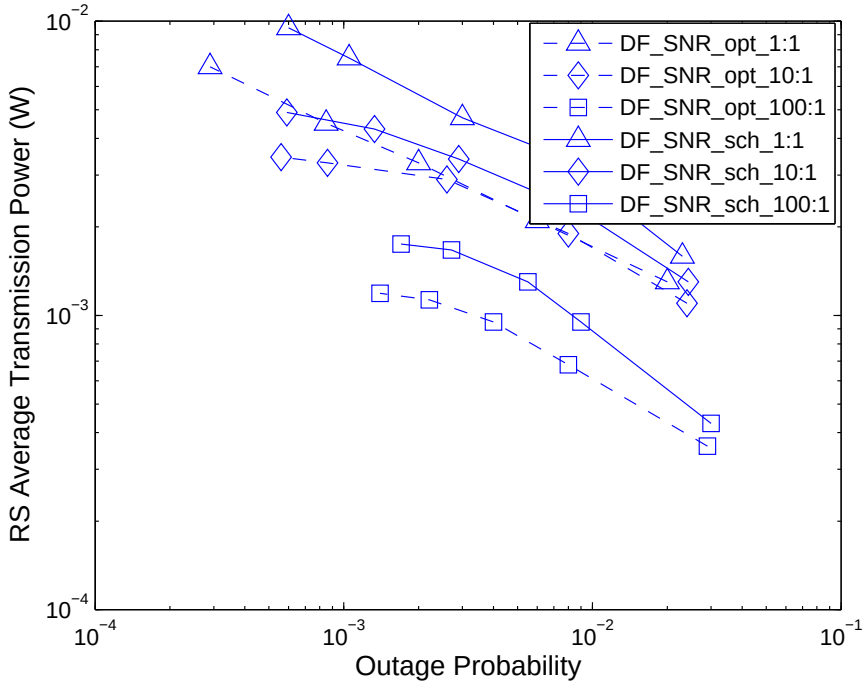

Figure 5.11: Downlink: RS transmission power using DF

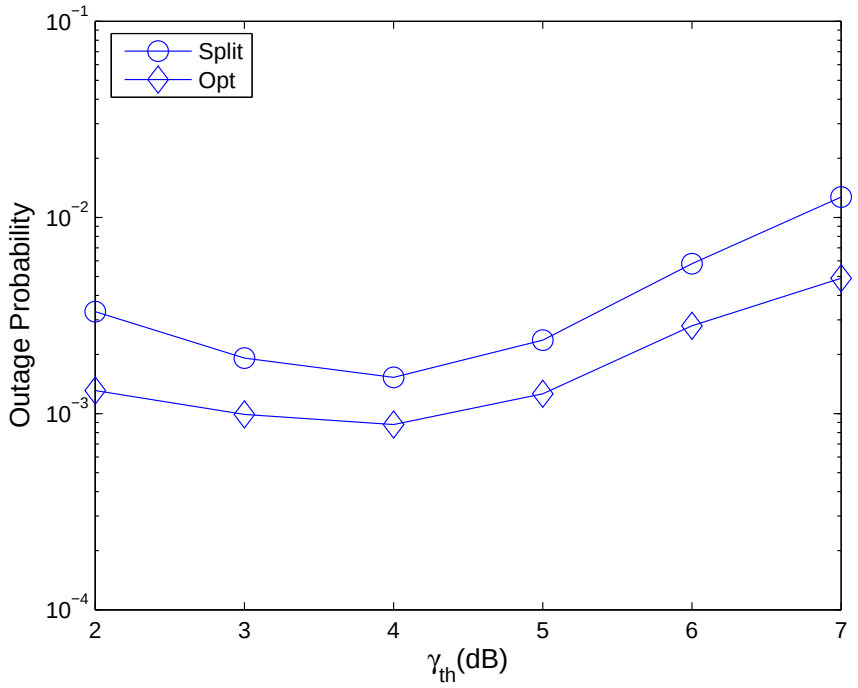

Figure 5.12: Downlink: effect of γ_{th} on outage performance using AF

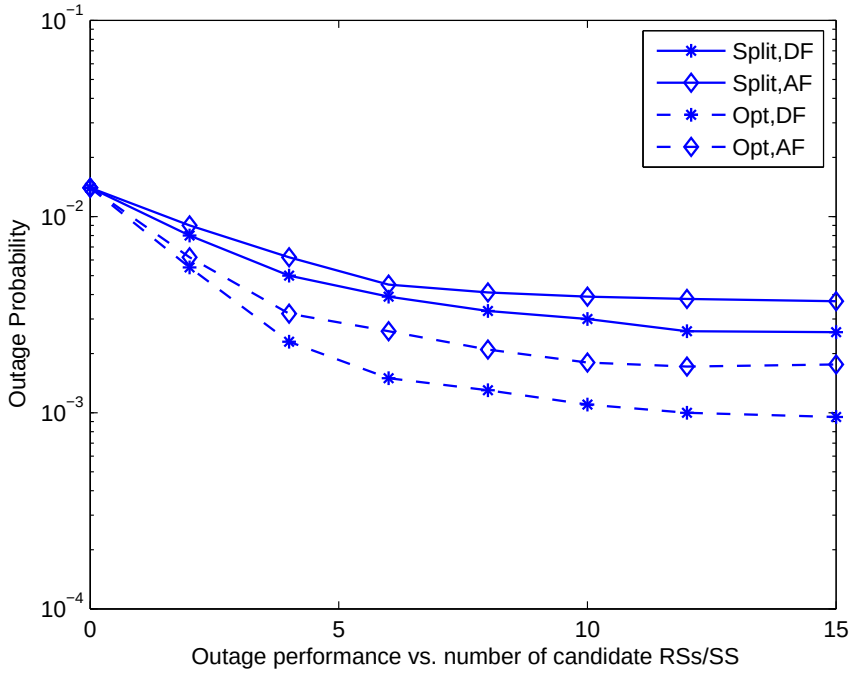

Figure 5.13: Downlink: effect of number of potential RSs/connections

Chapter 6

Conclusions and Future Works

We have studied resource allocations of two types of networks, cognitive radio networks and cooperative communication networks, in a CDMA-based scenario, where all simultaneous transmissions cause interference to one another, and transmission power management is the main factor that affects other aspects of the network performance.

We have studied resource management for supporting best effort traffic in the CRNs. Different approaches have been used in our proposed resource management schemes to managing the transmission power and interference in the CRN. Interference to the primary network is measured and fed back to the stations in the CRN, and this information is used so that the interference from the secondary transmissions to the primary network does not exceed the predefined interference threshold. Within the CRN, by preventing stations with strong mutual interference to transmit simultaneously, the available resources can be more efficiently utilized. Transmission opportunities are prioritized based on the traffic type, instantaneous transmission conditions, and the QoS requirements in terms of fairness, required throughput, or

outage performance. The work has proved that, efficient and fair resource management is possible in ad hoc CRNs for supporting both short term and long term traffic. Furthermore, it is possible to support traffic with strict quality of service requirements.

On the other hand, there are more practical and interesting topics on resource management that should be studied for CRNs. In our work, we have studied an environment with a single cell in the primary network and a single CRN. When multiple CRNs coexist in a cellular network with multiple radio cells, resource management becomes more complicated, yet more transmission opportunities can be explored for the CRNs. Resource management in multihop ad hoc CRNs should also be studied to extend the current work, which is based on single hop transmissions. Our current work is based on the assumption of stationary traffic load in the primary network. When there is significant jitter or the traffic load exhibits heavy tail distribution, the current admission control for long term throughput guarantee may not be valid, and other methods should be explored to predict the resource availability in the CRNs.

We have also studied power distribution in CDMA-based cooperative communication networks. Our work has shown that with very low transmission power from the relay stations, using cooperative relaying can significantly save the transmission power of the mobile stations or improving the outage performance. This work is based on given transmission rate requirements for the traffic. Further work can be done for supporting variable rate traffic, maximizing a certain system utility function, such the throughput, subject to providing the users with a certain fairness, supporting user mobility, etc.

Appendix A

Geometric Programming

We first introduce two definitions.

A *monomial* is defined as a function h: $\mathbf{R}^n_{++} \to \mathbf{R}$:

$$h(\mathbf{x}) = dx_1^{a^{(1)}} x_2^{a^{(2)}} \dots x_n^{a^{(n)}}, \tag{A.1}$$

where $d \geq 0$ and $a^{(j)} \in \mathbf{R}$, $j = 1, 2, \dots, n$.

Given that $d_k \geq 0$ and $a_k^{(j)} \in \mathbf{R}$ for all $k = 1, 2, \dots, K$ and $j = 1, 2, \dots, n$, a *posynomial* $f(\mathbf{x})$ is defined as

$$f(\mathbf{x}) = \sum_{k=1}^{K} d_k x_1^{a_k^{(1)}} x_2^{a_k^{(2)}} \dots x_n^{a_k^{(n)}}. \tag{A.2}$$

Consider the following nonconvex problem

$$\text{min} \quad f_0(\mathbf{x}) \tag{A.3}$$

$$\text{s.t.} \quad f_i(\mathbf{x}) \leq 1, \ i = 1, 2, \dots, m \tag{A.4}$$

where f_0 and f_i are posynomials. This type of problems in general are not convex, but can be transformed into convex problems. With a change of variables: $y_i = \log x_i$ and $b_{ik} = \log d_{ik}$, (A.3) can be converted into convex form [138]:

$$\begin{aligned} \min \quad & \tilde{f}_0(\mathbf{y}) = \log(\sum_k e^{a_{0k}^T y + b_{0k}}) \\ \text{s.t.} \quad & \tilde{f}_i(\mathbf{y}) = \log(\sum_k e^{a_{ik}^T y + b_{ik}}) \leq 0, \ i = 1, 2, \ldots, m. \end{aligned} \tag{A.5}$$

Since the functions $\tilde{f}_i$ are convex, this problem is a convex optimization problem, which can be solved globally and efficiently through the interior point primal dual method [138] with polynomial running times.

When either $f_0(\mathbf{x})$ or $f_i(\mathbf{x})$ is not posynomials but in the format of a ratio of posynomials, i.e., $f_i(\mathbf{x}) = s(\mathbf{x})/g(\mathbf{x})$, the problem is not a standard GP problem any more, but it still can be turned into GP problem by approximating the denominator of the ratio of posynomials, $g(\mathbf{x})$, with a monomial $\tilde{g}(\mathbf{x})$, leaving the numerator $s(\mathbf{x})$ unchanged. It is proved in [128] that if $g(\mathbf{x}) = \sum_i u_i(\mathbf{x})$ is a posynomial, then

$$g(\mathbf{x}) \geq \tilde{g}(\mathbf{x}) = \prod_i \left[\frac{u_i(\mathbf{x})}{\alpha_i} \right]^{\alpha_i}. \tag{A.6}$$

If, in addition, $\alpha_i = \frac{u_i(\mathbf{x}_0)}{g(\mathbf{x}_0)}$, $\forall i$, for any fixed positive $\mathbf{x}_0$, then $\tilde{g}(\mathbf{x}_0) = g(\mathbf{x}_0)$, and $\tilde{g}(\mathbf{x}_0)$ is the best local monomial approximation to $g(\mathbf{x}_0)$ near $\mathbf{x}_0$ in the sense of first order Taylor approximation. It is further proved in [128] that the approximation of a ratio of posynomials $f_i(\mathbf{x}) = s(\mathbf{x})/g(\mathbf{x})$ with $\tilde{f}_i(\mathbf{x}) = s(\mathbf{x})/\tilde{g}(\mathbf{x})$ satisfies the Karush-Kuhn-Tucker (KKT) conditions:

(1) $f_i(\mathbf{x}) \leq \tilde{f}_i(\mathbf{x})$ for all $\mathbf{x}$,

(2) $f_i(\mathbf{x}_0) = \tilde{f}_i(\mathbf{x}_0)$ where $\mathbf{x}_0$ is the optimal solution of the approximated problem in the previous iteration, and

(3) $\nabla f_i(\mathbf{x}_0) = \nabla \tilde{f}_i(\mathbf{x}_0)$.

With the above process, the denominator of $f_i(\mathbf{x})$ is approximated as a monomial, and $f_i(\mathbf{x})$ is then approximated as a posynomial. An iterative method as follows is then proposed in [128] to solve the original optimization problem:

Step 0: Choose an initial feasible point $\mathbf{x}^{(0)}$ and set $k = 1$.

Step 1: Approximate $g_i(\mathbf{x})$ with $\tilde{g}_i(\mathbf{x})$ around the previous point $\mathbf{x}^{(k-1)}$.

Step 2: Solve the approximated problem and obtain solution $\mathbf{x}^{(k)}$.

Step 3: Increase k by 1 and go back to Step 2 until the solution converges.

The convergence of this method is guaranteed by the KKT conditions in the approximation. It is demonstrated in [128] that this method can find the global optimum in over 96% cases, and a local optimum is found in other cases.

Bibliography

[1] M. Mouly and M. Pautet, *The GSM System for Mobile Communications*. Telecom Publishing, 1992.

[2] S. Redl, M. Weber, and M. Oliphant, *An Introduction to GSM*. Artech House, 1995.

[3] J. Cai and D. J. Goodman, "General packet radio service in GSM," *IEEE Communication Magazine*, vol. 35, no. 10, p. 122, 1997.

[4] R. Dettmer, "Mobilising packet data," *IEEE Review*, vol. 47, no. 4, p. 9, 2001.

[5] D. Calin and D. Zeghlache, "High speed circuit switched data over gsm: Potential traffic policies," *in Proceeding of IEEE Vehicle Technology Conference*, vol. 2, p. 1274, Ottawa, Ontario, Canada, May.1998.

[6] C. Scholefield, "Evolving GSM data services," *in Conference Record IEEE International Conference on Universal Personal Communication*, vol. 2, p. 888, San Diego, CA , USA, Oct.1997.

[7] "Third Generation Partnership Project (3GPP)."

[8] D. N. Knisely, S. Kumar, S. Laha, and S. Nanda, "Evolution of wireless data services: IS-95 to cdma2000," *IEEE Communications Magazine*, vol. 36, p. 140, 1998.

[9] A. Furuskar, S. Mazur, F. Muller, and H. Olofsson, "EDGE: Enhanced data rates for GSM and TDMA/136 evolution," *IEEE Personal Communication*, vol. 6, p. 56, 1999.

[10] P. Chaudhury, W. Mohr, and S. Onoe, "The 3GPP proposal for IMT-2000," *IEEE Communication Magazine*, vol. 37, no. 2, p. 72, 1999.

[11] C. Andersson and P. Svensson, "Mobile InternetAn industry-wide paradigm shift," *Ericsson Review*, 2005.

[12] "Mobile Multimedia Services Including Mobile Intranet and Internet Services," *European Telecommunications Standards Institute*, 1998.

[13] "Information technology Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks Specific Requirements. Part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," *IEEE Std 802.11*, 1999.

[14] J. Ala-Laurila, J. Mikkonen, and J. Rinnemaa, "Wireless LAN Access Network Architecture for Mobile Operators," *IEEE Communications Magazine*, vol. 2, no. 11, p. 82, 2001.

[15] M. Ghavami, L. B. Michael, and R. Khono, *Ultra Wideband Signals and Systems in Communications Systems.* Wiley, 2004.

[16] "Digital Video Broadcasting (DVB)," *www.dvb.org*, 2002.

[17] P. Demestichas, N. Koutsouris, and G. Koundourakis, "Management of Networks and Services in a Composite Radio Context," *IEEE Wireless Communication Magazine*, vol. 10, no. 4, p. 44, 2003.

[18] P. Demestichas and V. Stavroulaki, "Issues in Introducing Resource Brokerage Functionality in B3G, Composite Radio, Environments," *IEEE Wireless Communications Magazine*, vol. 11, no. 5, p. 32, 2004.

[19] P. Demestichas, G. Vivier, K. El-Khazen, and M. Theologou, "Evolution in Wireless Systems Management Concepts: From Composite Radio to Reconfigurability," *IEEE Communications Magazine*, vol. 42, no. 3, p. 90, 2004.

[20] Y. Kyun, Kim, Prasad, and Ramjee, *4G Roadmap and Emerging Communication Technologies*. Artech House, 2006.

[21] FCC, "Spectrum Policy Task Force Report," *ET Docket 02-135*, 2002.

[22] G.Fodor, A. Eriksson, and A. Tuoriniemi, "Providing Quality of Service in Always Best Connected Networks," *IEEE Communications Magazine*, vol. 7, no. 2, 2003.

[23] J. Zander and S.-L. Kim, *Radio Resource Management for Wireless Networks*. Artech House, 2001.

[24] R. Guerin and V. Peris, "Quality-of-service in Packet Networks: Basic Mechanisms and Directions," *Computer Networks*, vol. 31, p. 169, 1999.

[25] A. J. Viterbi, *Principles of Spread Spectrum Communication*. Addison-Wesley, 1995.

[26] B. Hashem and E. Sousa, "A combined power/rate control scheme for data transmission over a DS/CDMA system," *in Proceeding of IEEE VTC*, vol. 2, p. 1096, Sept.1998.

[27] S. W. Kim and Y. H. Lee, "Combined rate and power adaption in DS/CDMA communications over Nakagami fading channels," *IEEE Transactions on Communication*, vol. 48, no. 4, p. 162, 2000.

[28] S. Ramakrishna and J. M. Holtzman, "A scheme for throughput maximization in a dual-class CDMA system," *IEEE Jounal on Selected Areas Communication*, vol. 16, no. 7, p. 830, 1998.

[29] F. Berggren, S.-L. Kim, and J. Zander, "Joint power control and intracell scheduling of DS-CDMA nonreal time data," *IEEE Transactions on Communication*, vol. 19, no. 5, p. 1860, 2001.

[30] Y. Ma, J. Han, and K. Trivedi, "Call Admission Control for Reducing Dropped Calls in Code Division Multiple Access (CDMA) Cellular Systems," *9th Annual Joint Conference IEEE Compupter and Communication Societies*, vol. 3, p. 1481, Tel Aviv, Israel, May.2002.

[31] D. Zhao, X. Shen, and J. W. Mark, "Radio Resource Management for Cellular CDMA Systems Supporting Heterogeneous Services," *IEEE Transactions on Mobile Computers*, vol. 2, no. 6, p. 147, 2003.

[32] Y. Guo and B. Aazhang, "Call Admission Control in Multiclass Traffic CDMA Cellular System Using Multiuser Antenna Array Receiver," *Proceeding IEEE Vehicle Technology Conference*, vol. 1, p. 365, Tokyo, Japan, May.2000.

[33] C. Chuah, R. Yates, and D. Goodman, "Integrated Dynamic Radio Resource Management," *Proceeding IEEE Vehicle Technical Conference*, vol. 2, p. 584, Chicago, IL, USA, Jul.1995.

[34] D. Kim, "On Upper Bounds of SIR-based Call Admission Threshold in Power-controlled DS-CDMA Mobile Systems," *IEEE Communication Letters*, vol. 6, no. 2, p. 13, 2002.

[35] J. Chak and W. Zhuang, "Connection Admission Control for Indoor Multimedia Wireless Communications," *Proceeding IEEE Vehicle Technical Conference*, vol. 4, p. 2570, Ottawa, Ontario, Canada, May.1998.

[36] Z. Liu and M. E. Zarki, "SIR-Based Call Admission Control for DS-CDMA Cellular Systems," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 4, p. 638, 1994.

[37] M. Kim, B. C. Shin, and D.-J. Lee, "SIR-Based Call Admission Control Inter-cell Interference Prediction for DS-CDMA Systems," *IEEE Communication Letters*, vol. 4, no. 4, p. 29, 2000.

[38] Z. Dziong, M. Jia, and P. Mermelstein, "Adaptive Traffic Admission for Integrated Services in CDMA Wirelessaccess Networks," *IEEE Journal on Selected Area in Communications*, vol. 14, no. 5, p. 1737, 1996.

[39] J. Chang, J. Chung, and D. Sung, "Admission Control Schemes for Soft Handoff in DS-CDMA Cellular Systems Supporting Voice and Stream-type Data Services," *IEEE Transactions Vehicle Technology*, vol. 51, no. 6, p. 1445, 2002.

[40] C. Comaniciu and H. V. Poor, "Jointly optimal power and admission control for delay sensitive traffic in CDMA networks with LMMSE receivers," *IEEE Transactions on Signal Processing*, vol. 51, no. 6, p. 2031, 2003.

[41] S. Singh, V. Krishnamurthy, and H. V. Poor, "Integrated voice/data call admission control for wireless DS-CDMA systems with fading," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, p. 1483, 2002.

[42] F. Yu, V. Krishnamurthy, and V. C. M. Leung, "Cross-layer optimal connection admission control for variable bit rate multimedia trafiic in packet wireless CDMA networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, p. 542, 2006.

[43] S. A. Grandhi, R. Vijaya, and D. J. Goodman, "Distributed power control in cellular radio systems," *IEEE Transactions on Communications*, vol. 42, no. 5, p. 226, 1994.

[44] R. Ramanathan and R. Rosales-Hain, "Topology control of multihop wireless networks using transmit power adjustment," *IEEE INFOCOM*, vol. 2, p. 404, Tel Aviv, Israel, Mar.2000.

[45] L. Kleinrock and J. Silvester, "Optimum transmission radii packet radio networks or why six is a magic number," *Proceeding of IEEE Telecommunication Conference*, Piscataway, N.J, 1978.

[46] T. Hou and V. Li, "Transmission range control in multihop packet radio networks," *IEEE Transactions on Communication*, vol. 34, no. 3, p. 38, 1986.

[47] T. ElBatt, S. Krishnamurthy, D. Connors, and S. Dao, "Power management for throughput enhancement in wireless ad hoc networks," *Proceeding IEEE ICC*, New Orleans, LA, USA, Jun,2000.

[48] N. Bambos, "Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects," *5th International Conference on Telecommunications and Information Technology*, Krabi, May.2008.

[49] R. Ramanathan and R. Rosales-Hain, "Topology control of multihop wireless networks using transmit power adjustment," *Proceeding of IEEE INFOCOM*, Tel Aviv, Israel, Mar.2000.

[50] J. Cai, K.-H. Liu, X. Shen, J. Mark, and T. Todd, "Power allocation and scheduling for MAC layer design in UWB networks ," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 2, p. 1103, 2008.

[51] J. P. Monks, "A Study of the Energy Saving and Capacity Improvement Potential of Power Control in Multi-hop Wireless Networks," *Proceeding IEEE 26th Annual Local Computer Networks*, New York, NY, USA, Feb.2001.

[52] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interfrence compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, p. 1074, 2006.

[53] N. Feng, S. Mau, and N. B. Mandayam, "Pricing and power control for joint network-centric and user-centric radio resource management," *IEEE Transactions on Communications*, vol. 52, no. 2, p. 1547, 2004.

[54] C. W. Sung, K. W. Shum, and K. K. Leung, "Multi-objective Power Control and Signature Sequence Adaptation for Synchronous CDMA Systems a Game-theoretic Viewpoint," *Proceeding IEEE International Information Theory*, vol. 24, p. 335, Jun.2003.

[55] R. Menon, "A Game-theoretic Framework for Interference Avoidance in Ad Hoc Networks," 2005.

[56] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, p. 74, 2004.

[57] R. L. Cruz and A. V. Santhanam, "Optimal routing, link scheduling and power control in multi-hop wireless networks," *IEEE Infocom*, Mar.2003.

[58] M. Thoppian, S. Venkatesan, and R. Chandrasekaran, "MAC-layer scheduling in cognitive radio based multi-hop wireless networks," *in Proceedings of International Symposium on World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2006.

[59] K. Hamdi, W. Zhang, and K. B. Letaief, "Uplink scheduling with QoS provisioning for cognitive radio systems," *in Proceedings of Wireless Communications and Newtroking Conference*, Mar.2007.

[60] T. Kwon and M. Gerla, "Clustering with power control," *IEEE MILCOM*, vol. 2, p. 1422, Atlantic City, NJ, USA, Oct.1999.

[61] J. P. Monks, V. Bharghavan, and W. W. Hwu, "Transmission Power Control

for Multiple Access Wireless Packet Networks," *25th Annual IEEE Conference on Local Computer Networks*.

[62] A. Zadeh and B. Jabban, "Performance analysis of multihop packet CDMA cellular networks," *Proceeding of IEEE GLOBCOM*, San Antonio, TX, Nov.2001.

[63] V. Sreng, H. Yanikomeroglu, and D. Falconer, "Coverage enhancement through two-hop relaying in cellular radio systems," *IEEE Wireless Communications and Networking Conference*, vol. 01, p. 881, Aug.2002.

[64] H. Wu, C. Qiao, S. De, and O. Tonguz, "Integrated cellular and ad hoc relaying systems: iCAR," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 3, p. 2105, 2001.

[65] H. Luo, R. Ramjee, P. Sinha, L. Li, and S. Lu, "UCAN: a unified cellular and ad-hoc network architecture," *in Proceeding of Mobile Computing Networking*, 2003.

[66] V. Sreng, H. Yanikomeroglu, and D. Falconer, "Relayer Selection Strategies in Cellular Networks with Peer-to-Peer Relaying," *IEEE Vehicular Technology Conference*, Oct.2003.

[67] H. Yap, X. Yang, S. Ghaheri-Niri, and R. Tafazolli, "Position Assisted Relaying and Handover in Hybrid Ad Hoc WCDMA Cellular System," *Proceeding of 13th IEEE International Sympothm on Personal Indoor and Mobile Radio Communication*, Sept.2003.

[68] T. Harrold and A. Nix, "Performance Analysis of Intelligent Relaying in UTRA

TDD," *Proceeding of IEEE Vehicular Technology Conference*, vol. 3, p. 1374, Dec.2002.

[69] T. Todd and D. Zhao, "Cellular CDMA Capacity in Hotspots with Limited Ad Hoc Relaying," *Proceeding of 14th IEEE International Sympothm on Personal Indoor and Mobile Radio Communication*, Sept. 2003.

[70] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," *in Proceeding of IEEE Internation Symposium on Information Theory*, Cambridge, MA, USA, Aug.1998.

[71] A. Stefanov and E. Erkip, "Cooperative Coding for Wireless Networks," *IEEE Transactions on Communication*, vol. 52, no. 3, p. 1470, 2004.

[72] A. Sendonaris, E. Erkip, and B. Aazhang., "User cooperation diversity-Part I: System description," *IEEE Transactions on Communications*, vol. 51, p. 1927, 2003.

[73] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Transactions on Information Theory*, vol. 51, no. 6, p. 3037, 2005.

[74] D. Gunduz and E. Erkip, "Opportunistic cooperation by dynamic resource allocation," *IEEE Transactions on Communication*, vol. 6, no. 5, p. 1446, 2007.

[75] N. Ahmed, M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Outage minimization with limited feedback for the fading relay channel," *IEEE Transactions on Communication*, vol. 54, no. 3, p. 659, 2006.

[76] T. T. Kim, G. Caire, and M. Skoglund, "Decode-and-forward relay channels with quantized channel state feedback: an outage exponent analysis," *IEEE Transactions on Information Theory*, vol. 54, no. 5, p. 4548, 2008.

[77] J. Luo, R. S. Blum, and A. M. Haimovich, "New approaches for cooperative use of multiple antennas in ad-hoc wireless networks," *Proceeding IEEE Vehicular Technology Conference (VTC)-Fall*, vol. 4, p. 2769, Sept.2004.

[78] A. Azgin, Y. Altunbasak, and G. AlRegib, "Cooperative MAC and routing protocols for wireless ad hoc networks," *IEEE Global Telecommunication Conference*, St. Louis, MO, Dec.2004.

[79] Z. Yang, J. Liu, and A. Host-Madsen, "Cooperative routing and power allocation in ad-hoc networks," *Proceeding on IEEE Globel Telecommunications Conference*, St. Louis, MO, Dec.2005.

[80] P. Liu, Z.Tao, and S. Panwar, "A Cooperative MAC Protocol for Wireless Local Area Networks," *Proceeding IEEE ICC*, May. 2005.

[81] B. Sirkeci-Mergen and A. Scaglione, "On the power efficiency of cooperative broadcast in dense wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, p. 497, 2007.

[82] J. Mitola, "Cognitive radio: an integrated agent architecture for software defined radio," *Ph.D. Thesis, KTH Royal Inst. Technology*, Stockholm, Sweden, 2000.

[83] S. Haykin, "Cognitive radio: Brain-empowered wireless communications,"

IEEE Journal on Selected Areas in Communications, vol. 25, no. 6, p. 201, 2005.

[84] A. Petrin and P. G. Steffes, "Analysis and comparison of spectrum measurements performed in urban and rural areas to determine the total amount of spectrum usage," *Proceeding of the ISART*, 2005.

[85] R. Berger, "Open spectrum: a path to ubiquitous connectivity," *ACM Queue*, 2003.

[86] C. Cordeiro and K. Challapali, "802.22: the first worldwide wireless standard based on cognitive radios," *In First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Baltimore, MD, USA, Nov.2005.

[87] M. Akyildiz, W. Lee, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Elsevier Computer Networks*, vol. 50, p. 2127, 2006.

[88] T. Clancy, "Measuring interference temperature," *Virginia Tech Symposium on Wireless Personal Communications*, 2006.

[89] I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A Survey of Spectrum Management of Cognitive Radio Networks," *IEEE Communications Magazine*, vol. 46, no. 4, p. 40, 2008.

[90] S. Kirshamurthy, "Control Channel Based MAC-Layer Configuration, Routing and situation Awareness for Cognitive Radio Networks," *IEEE MILCOM*, May.2005.

[91] L. Le and E. Hossain, "OSA-MAC: A MAC Protocol for Opportunistic Spectrum Access in Cognitive Radio Networks," *IEEE WCNC*, Las Vegas, NV, Mar.2008.

[92] Y. R. Kondareddy and P. Agrawal, "Synchronized MAC Protocol for Multi-hop Cognitive Radio Networks," *IEEE ICC*, vol. 25, p. 36, Beijing, China, Jun.2008.

[93] L. Ma, X. Han, and C. Shen, "Dynamic open spectrum sharing MAC protocol for wireless ad hoc networks," *IEEE DySPAN*, Sept.2005.

[94] S. Sankaranarayanan, P. Papadimitratos, and A. Mishra, "A bandwidth sharing approach to improve licensed spectrum utilization," *In First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Jun. 2005.

[95] 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, no. 7, p. 576, 2007.

[96] Y. Yuan, P. Bahl, and Y. Wu, "Allocating dynamic time-spectrum blocks for cognitive radio networks," *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, May. 2007.

[97] J. Tang, S. Misra, and G. Xue, "Joint spectrum allocation and scheduling for fair spectrum sharing in cognitive radio wireless networks," *Computer Networks*, vol. 52, p. 2148, 2008.

[98] R. Urgaonkar and M. J. Neely, "Opportunistic Scheduling with Reliability

Guarantees in Cognitive Radio Networks," *Proceeding of IEEE INFOCOM*, Mar.2008.

[99] A. Hoang and Y. Liang, "Maximizing Spectrum Utilization of Cognitive Radio Networks Using Channel Allocation and Power Control," *in Vehicular Technology Conference*, Sept.2005.

[100] W. Wang and X. Liu, "List-coloring based channel allocation for open spectrum wireless networks," *in Proceedings of IEEE 62nd Vehicular Technology Conference*, Sept.2004.

[101] H. Zheng and C. Peng, "Collaboration and fairness in opportunistic spectrum access," *in Proceedings of IEEE International Conference on Communications*, vol. 5, p. 3132, Hong Kong, Jun.2005.

[102] L. T. Q. Zhao and A. Swami, "Decentralized cognitive MAC for dynamic spectrum access," *in Proceeding of IEEE DySPAN*, Sept.2005.

[103] A. Behzad and I. Rubin, "Multiple access protocol for power-controlled wireless access nets," *IEEE Transactions on Mobile Computing*, vol. 3, no. 5, p. 307, 2004.

[104] S. A. G. Kulkarni and M. Srivastava, "Subcarrier allocation and bit loading algorithms for OFDMA-based wireless networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 6, p. 652, 2005.

[105] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, "A general framework for wireless spectrum auctions," *2nd IEEE International Symposium on DySPAN*, Sept.2007.

[106] X. Zhou, S. Gandi, S. Suri, and H. Zheng, "eBay in the Sky: Strategyproof wireless spectrum auctions," *Proceedings of the 14th ACM international conference on Mobile computing and networking*, 2008.

[107] F. Kelly, A. Maulloo, and D. Tan, "Rate Control in Communication Networks: Shadow Prices, Proportional Fairness and Stability," *Journal of the Operational Research Society*, vol. 49, no. 6, p. 237, 1998.

[108] F. P. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, p. 33, 1997.

[109] J. F. Nash, "The bargaining problem," *Econometrica*, vol. 18, p. 155, 1950.

[110] L. Zhou, "The Nash bargaining theory with non-convex problems," *Econometrica*, vol. 65, p. 681, 1997.

[111] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMAHDR a high efficiency-high data rate personal communication wireless system," *IEEE Vehicular Technology Conference*, vol. 3, p. 1854, Sept.2000.

[112] Y. Xing and C. N. Mathur, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Transactions on Mobile Computing*, vol. 6, no. 6, p. 423, 2007.

[113] M. H. Islam, Y.-C. Liang, and A. T. Hoang, "Distributed power and admission control for cognitive radio networks using antenna arrays," *in Proceeding of IEEE DySPAN*, Dublin, April.2007.

[114] L. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive

radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, p. 5306, 2008.

[115] D. I. Kim, L. Le, and E. Hossain, "Joint rate and power allocation for cognitive radios in dynamic spectrum access environment," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, p. 5517, 2008.

[116] L. Qian, X. Li, J. Attia, and Z. Gajic, "Power control for cognitive radio ad hoc networks," *in Proceding of IEEE Local and Metropolitan Area Networks*, 2007.

[117] E. Beres and R. Adve, "Selection cooperation in multi-source cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, p. 118, 2008.

[118] T. M. Cover and A. E. Gamal, "Capacity Theorems for the Relay Channel," *IEEE Transactions on Information Theory*, vol. 25, no. 2, p. 572, 1979.

[119] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Wireless Communications*, vol. 42, no. 2, p. 68, 2004.

[120] F. Liu and R. Dong, "A Reputation Mechanism to Stimulate Node Cooperation in Ad Hoc Networks," *3rd International Conference on Genetic and Evolutionary Computing*, Guilin, Oct.2009.

[121] B. S. Naouel and B. Levente, "A Charging and Rewarding Scheme for Packet Forwarding In Multi-hop Cellular Networks," *4th ACM Symposium on Mobile Ad Hoc Networking and Computing*, New York, NY, US, 2003.

[122] M. Jakobsson and J.P.Hubaux, "A Micro-payment Scheme Encouraging Collaboration In Multi-hop Cellular Networks," *7th International Financial Cryptography Conference*, 2003.

[123] P. Mitran, H. Ochiai, and V. Tarokh, "Space-time diversity enhancements using collaborative communications," *IEEE Transactions on Information Theory*, vol. 51, no. 3, p. 2041, 2005.

[124] A. K. Sadek and W. Su, "Multi-node cooperative communications in wireless networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, p. 341, 2007.

[125] R. Hu, S. Sfar, and G. Charlton, "Protocols and system capacity of relay-enhanced hsdpa systems," *Proceeding of the 2008 Annual Conference on Information Sciences and Systems*, Princeton, NJ, Mar.2008.

[126] T. C.-Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allo- cations in cooperative cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 5, p. 328, 2007.

[127] I. Akyildiz, W. Lee, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 9, no. 7, p. 40, 2008.

[128] M. Chiang and C. Tan, "Power control by geometric programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, p. 2640, 2007.

[129] K. S. Gilhousen, I. M. Jacobs, and R. Padovani, "On the capacity of a cellular CDMA system," *IEEE Transactions of Vehicle Technology*, vol. 40, no. 6, p. 303, 1991.

[130] A.L.Atolyar, "On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput aloocation," *Bell Lab Report*, vol. 53, p. 12.

[131] C. Lee, "Power control for CDMA cellular radio systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, p. 77, 2006.

[132] K.-H. Liu, L. Cai, and X. Shen, "Exclusive-Region Based Scheduling Algorithms for UWB WPAN," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, p. 933, 2008.

[133] D. Chiu and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," *Computer Networks and ISDN Systems*, vol. 17, p. 1, 1989.

[134] F. Kelly, "Effective bandwidths at multi-class queues," *Queueing Systems*, vol. 9, no. 1, Oct.1991.

[135] J. Huih, "Resource allocation for broadband networks," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 12, p. 1598, Jun.1988.

[136] Y. Ohta and T. Fujii, "Fading channel measurement for static mobile terminals in outdoor NLOS environments," *in Proceeding of IEEE 69th Vehicular Technology Conference*, Barcelona, Apr.2009.

[137] R. Jain, W. Hawe, and D. Chiu, "A Quantitative measure of fairness and discrimination for resource allocation in Shared Computer Systems," *DECTR-301.*

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