

DISSERTATION

RESOURCE MANAGEMENT IN  
QOS-AWARE WIRELESS CELLULAR NETWORKS

Submitted by

Zhi Zhang

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2011

Doctoral Committee:

Advisor: Edwin K. P. Chong

Mahmood R. Azimi-Sadjadi

Peter M. Young

William S. Duff

Copyright by Zhi Zhang 2011

All Rights Reserved

## ABSTRACT

### RESOURCE MANAGEMENT IN QOS-AWARE WIRELESS CELLULAR NETWORKS

Emerging broadband wireless networks that support high speed packet data with heterogeneous quality of service (QoS) requirements demand more flexible and efficient use of the scarce spectral resource. Opportunistic scheduling exploits the time-varying, location-dependent channel conditions to achieve multiuser diversity. In this work, we study two types of resource allocation problems in QoS-aware wireless cellular networks. First, we develop a rigorous framework to study opportunistic scheduling in multiuser OFDM systems. We derive optimal opportunistic scheduling policies under three common QoS/fairness constraints for multiuser OFDM systems—temporal fairness, utilitarian fairness, and minimum-performance guarantees. To implement these optimal policies efficiently, we provide a modified Hungarian algorithm and a simple suboptimal algorithm. We then propose a generalized opportunistic scheduling framework that incorporates multiple mixed QoS/fairness constraints, including providing both lower and upper bound constraints.

Next, taking input queues and channel memory into consideration, we reformulate the transmission scheduling problem as a new class of Markov decision processes (MDPs) with fairness constraints. We investigate the throughput maximization and the delay minimization problems in this context. We study two categories of fairness constraints, namely temporal fairness and utilitarian fairness. We consider two

criteria: infinite horizon expected total discounted reward and expected average reward. We derive and prove explicit dynamic programming equations for the above constrained MDPs, and characterize optimal scheduling policies based on those equations. An attractive feature of our proposed schemes is that they can easily be extended to fit different objective functions and other fairness measures. Although we only focus on uplink scheduling, the scheme is equally applicable to the downlink case. Furthermore, we develop an efficient approximation method—temporal fair rollout—to reduce the computational cost.

## ACKNOWLEDGEMENTS

First and foremost, I like to express my sincerest gratitude to my advisor and mentor, Professor Edwin K. P. Chong. I am blessed to have this opportunity to work with him, learn from him in the past years. He gave me copious amounts of insightful guidance, constant encouragement, constructive criticism, and enormous patience throughout all these years. Specially, his logical thinking and high standards to scientific work influenced me greatly, which will benefit me for my whole life. His enthusiasm and dedication to his students are truly inspiring; it is my very privilege and pride to have been one of them.

I would also like to thank Professors Mahmood R. Azimi-Sadjadi, Peter M. Young, and William S. Duff for being on my advisory committee and for all their helps along the way. I am also indebted to Professors Peter J. Brockwell, Anura P. Jayasumana, Thomas Lee, H. J. Siegel, Louis Scharf, and ECE department head Anthony A. Maciejewski for their instruction and support all along.

Next, I am very grateful to the many friends and colleagues I have had at Colorado State. Many thanks to Dr. Ying He, Dr. Yun Li, Dr. Hua Li, Dr. Patricia R. Barbosa, Dr. Zhifei Fan, Dr. Michael Buehner, Professor Ali Pezeshki, Professor Hongya Ge, Professor Jeffrey D. Herdtner, Professor Peter Schreier, Sudhir Moola, Ramin Zahedi, Yang Zhang, Sowmya Lolla, Lucas W. Krakow, Sathish Chittibabu for many illuminating conversations and everlasting friendship. I would also like to say “thank you” to many other friends, especially Jun Lei, Xiaofan Cao, Hua Shao, Xin He, Yong Wang, Shanshan Li, Lifang Xu, Yan Liu, Jun Pan, Jia Yue, Sisi Zhang, Tao Yuan, and our basketball team; my years at Fort Collins would not have been

the same without them.

Finally, I want to thank my wife, Dr. Yu Yang, for her love, understanding, patience, and encouragement, to my boy “DuDu” Feifan for being my another accomplishment. I could not have accomplished this without them. Very special thanks to my parents and my brother for their constant and unconditional love and support.

*To my parents, my wife, and my son*

# TABLE OF CONTENTS

<b>ABSTRACT</b> . . . . .	<b>ii</b>
<b>ACKNOWLEDGEMENTS</b> . . . . .	<b>iv</b>
<b>DEDICATION</b> . . . . .	<b>vi</b>
<b>LIST OF TABLES</b> . . . . .	<b>x</b>
<b>LIST OF FIGURES</b> . . . . .	<b>xi</b>
<b>1 INTRODUCTION</b> . . . . .	<b>1</b>
1.1 Overview . . . . .	1
1.2 Organization of This Dissertation . . . . .	5
<b>2 BACKGROUND AND RELATED WORK</b> . . . . .	<b>7</b>
2.1 Overview of Wireless Cellular Communication Networks . . . . .	7
2.1.1 Evolution of Wireless Cellular Communication Systems . . . . .	8
2.1.2 Key Techniques of Wireless Cellular Communication Systems . . . . .	10
2.2 Mobile Radio Propagation Model . . . . .	18
2.3 Transmission Scheduling in Wireless Networks . . . . .	22
2.3.1 Scheduling in Wireline Networks . . . . .	22
2.3.2 Challenges for Transmission Scheduling in Wireless Networks . . . . .	23
2.3.3 Fairness . . . . .	25
2.3.4 Scheduling Schemes in Wireless Networks . . . . .	27
2.4 Cross Layer Design . . . . .	32
2.5 Markov Decision Processes . . . . .	35
2.5.1 Markov Decision Process Formulation . . . . .	35
2.5.2 Approximate Dynamic Programming . . . . .	37
2.5.3 Constrained Markov Decision Processes . . . . .	39



2.6	Related Work . . . . .	40
<b>3</b>	<b>OPPORTUNISTIC SCHEDULING FOR OFDM SYSTEMS WITH FAIRNESS CONSTRAINTS . . . . .</b>	<b>46</b>
3.1	Motivation . . . . .	47
3.2	System Model . . . . .	49
3.3	Opportunistic Scheduling under Various Fairness Constraints . . . . .	52
3.3.1	Temporal Fairness Scheduling . . . . .	53
3.3.2	Utilitarian Fairness Scheduling . . . . .	55
3.3.3	Minimum-Performance Guarantee Scheduling . . . . .	58
3.4	Implementation Issues . . . . .	60
3.4.1	Control Parameter Estimation . . . . .	60
3.4.2	Optimal User Subset Search Methods . . . . .	61
3.5	Simulation Results . . . . .	64
3.5.1	Performance Gain . . . . .	65
3.5.2	Fairness . . . . .	68
3.6	Conclusions . . . . .	71
<b>4</b>	<b>OPPORTUNISTIC SCHEDULING WITH MIXED CONSTRAINTS . . . . .</b>	<b>72</b>
4.1	Motivation . . . . .	73
4.2	System Model . . . . .	75
4.3	Scheduling with Maximum and Minimum Bounds . . . . .	76
4.3.1	Performance Guarantee Scheduling . . . . .	76
4.3.2	Temporal Fairness Scheduling . . . . .	79
4.3.3	Utilitarian Fairness Scheduling . . . . .	81
4.4	Scheduling with Mixed Constraints . . . . .	83
4.4.1	Minimum Temporal Fairness and Performance Guarantee . . . . .	83
4.4.2	Minimum Temporal and Utilitarian Fairness . . . . .	85
4.5	A Generalized Framework for Opportunistic Scheduling . . . . .	88
4.6	Conclusions . . . . .	89

<b>5</b>	<b>STOCHASTIC DYNAMIC PROGRAMMING FOR OPPORTUNISTIC SCHEDULING</b>	<b>91</b>
5.1	Introduction	92
5.2	System Model And Problem Formulation	94
5.2.1	System Model	94
5.2.2	MDP Problem Formulation	95
5.3	Temporal Fairness scheduling	98
5.3.1	Expected Total Discounted Reward Criterion	98
5.3.2	Expected Average Reward Criterion	105
5.4	Utilitarian Fairness Scheduling	109
5.4.1	Expected Total Discounted Reward Criterion	109
5.4.2	Expected Average Reward Criterion	112
5.5	Temporal Fair Rollout Algorithm	115
5.5.1	Rollout Algorithm	115
5.5.2	Temporal Fair Rollout Algorithm	116
5.6	Stochastic Approximation for Parameter Estimation	117
5.7	Numerical Results	119
5.7.1	Simulation Setup	119
5.7.2	Channel Model	121
5.7.3	Simulation Results	123
5.8	Conclusions	124
<b>6</b>	<b>CONTRIBUTIONS AND FUTURE WORK</b>	<b>130</b>
6.1	Summary of Contributions	130
6.2	Future Work	132
	<b>REFERENCES</b>	<b>134</b>

## LIST OF TABLES

3.1	Modified Hungarian algorithm . . . . .	63
5.1	Channel state transition probabilities . . . . .	123

## LIST OF FIGURES

2.1	Received signal strength as a function of time for vehicle velocity at 10 mph. . . . .	21
3.1	Downlink scheduling over multiuser OFDM system . . . . .	49
3.2	System throughput gain in the temporal fairness scheduling. . . . .	65
3.3	System throughput gain in the utilitarian fairness scheduling. . . . .	66
3.4	System throughput gain in the minimum-performance guarantee scheduling. . . . .	67
3.5	Portion of resource shared by users in the temporal fairness scheduling. . . . .	68
3.6	User average performance in the utilitarian fairness scheduling. . . . .	69
3.7	User average performance in the minimum-performance guarantee scheduling. . . . .	70
5.1	Uplink queueing model of a single-cell TDM system. . . . .	94
5.2	Block diagram of the scheduling policy with online parameter estimation. . . . .	117
5.3	State transition for Rayleigh fading channel model. . . . .	122
5.4	Time fraction allocation for throughput maximization problem. . . . .	126
5.5	Average system throughput vs. normalized arrival rate. . . . .	127
5.6	Time fraction allocation for delay minimization problem. . . . .	128
5.7	Average system queue length vs. normalized arrival rate. . . . .	129

# CHAPTER 1

## INTRODUCTION

### 1.1 Overview

Unprecedented advances in wireless technology have been playing active roles in re-defining our modern lifestyle. With the application of the third-generation wireless technology, cellular systems now support interactive multimedia and high speed data services. Besides making phone calls, we are now able to use cell phones to access the Internet, conduct monetary transactions, send text messages, watch streaming videos, etc.

According to a recent International Telecommunication Union (ITU) report [1], till 2009 around six in ten people across the world use cell phones, compared to just under 15 percent of the global population used cell phones in 2002. The size of the annual wireless communication business in the United States has grown from \$100 million to more than \$50 billion in less than 15 years [2]. Meanwhile, WiFi access is now widely supported in schools, hotels, airports, coffee shops, and many other public areas. The number of WiFi hotspots in US has grown from only 15k in 2002 to over 100k in 2005. Cities like San Francisco and Philadelphia are creating citywide WiFi networks.

With the advancement of wireless technologies, wireless networking has become *ubiquitous* owing to the great demand of pervasive mobile applications. Some fundamental challenges exist for the next generation wireless network design, provisioning

of heterogeneous type of services, etc. One of the fundamental characteristics of wireless networks is the *time-varying* and *location-dependent* channel conditions due to multipath fading. A wireless channel can change rapidly and can be seriously affected by the radio propagation parameters and interferences, thus the topology and link characteristics are dynamically varying in wireless networks. The performance of a wireless network is mainly restrained by the interferences and the time-varying nature of wireless channels. The co-channel interference (CCI) is caused by users sharing the same channel due to the multiple access in wireless networks. Due to the effects such as multipath fading, shadowing, path loss, propagation delay, and noise level, the signal-to-interference-noise-ratio (SINR) at a receiver output can fluctuate in the order of tens of dBs. So how to overcome these difficulties and improve the system performance has always been the central issue in wireless community.

Dynamic resource allocation is a general strategy to control the interferences and enhance the performance of wireless networks. The basic idea behind dynamic resource allocation is to utilize the channel more efficiently by sharing the spectrum and reducing interference through optimizing parameters such as the transmission power, symbol transmission rate, modulation scheme, coding scheme, bandwidth, or combinations of these parameters. Moreover, the network performance can be further improved by introducing more diversity, such as multiuser, time, frequency, and space diversity. In addition, cross layer approach for resource allocation can provide advantages such as low overhead, more efficiency, and direct end-to-end QoS provision.

Moreover, there are other constraints such as fairness, heterogenous QoS provisioning, and practical implementation constraints. Since each user pays the same for his service, it is desirable to have fair resource allocation scheme. In order to provide fair services to all users, we need to define the new fairness concepts. In this dissertation, we consider three kinds of fairness—*temporal fairness*, *utilitarian fairness*, and

*minimum-performance guarantees*. For various applications, the QoS requirements can be very different. For example, voice payload is very sensitive for delay, data payload requires low bit error rate (BER), and video payload has burst transmission. Also, there are many practical constraints for wireless system implementation such as maximal transmitted power, minimal throughput, computation capability, implementation cost, etc. So how to optimally allocate the resources under these constraints has become an important wireless research topic.

Good scheduling schemes in wireless networks should *opportunistically* seek to exploit the time-varying channel conditions to improve spectrum efficiency thereby achieving *multiuser diversity* gain. However, the potential to transmit at higher data rates opportunistically also introduces an important tradeoff between wireless resource efficiency and level of satisfaction among individual users (fairness). For example, allowing only users close to the base station to transmit at high transmission rate may result in very high throughput, but sacrifice the transmission of other users. Such a scheme cannot satisfy the increasing demand for QoS provisioning in broadband wireless networks.

Orthogonal frequency division multiplexing access (OFDMA) is a popular multiple access and signaling scheme for wireless broadband networks. Adaptive modulation techniques in OFDMA provide the potential to vary the number of transmitted bits for a sub-channel, according to instantaneous sub-channel quality, while maintaining an acceptable BER. Resource allocation for OFDMA networks has three major tasks: sub-channel assignment, throughput allocation, and power control.

To enhance the system performance, we explore the multi-dimension diversity. By using throughput control in MAC layer, we can apply multiuser diversity and time diversity to allocate resources efficiently to different users over time according to their channel conditions. By using OFDM technique, we can apply frequency diversity to fully utilize the limited bandwidth. All these diversity can be combined together to

combat the detrimental effects such as time varying channels, CCI, or heterogeneous QoS requirements, etc.

Practical radio channels are commonly modeled as multipath Rayleigh fading channels, which are *correlated* random processes. However, much of the prior work on scheduling, is based on the relatively simple *memoryless* channel models. Finite-state Markov channel (FSMC) models have been found to be accurate in modeling such channels with memory. When channel memory is taken into consideration, the existing work on memoryless channels does not apply directly. Also, much of the previous work focused on “*elastic*” traffic, and assumed that the system has *infinite backlogged* data queues, which is not always an appropriate assumption. This assumption makes it impossible to consider the data arrival queues and further evaluate the system delay performance. The widely studied Markov decision processes (MDPs) and the associated dynamic programming methodology provide us with a general tool for posing and analyzing such sequential decision making problems with an underlying Markov process.

In this dissertation, we will study several resource allocation problems in QoS-aware wireless cellular networks. We first develop a rigorous framework to study opportunistic scheduling in multiuser OFDM systems. We derive optimal opportunistic scheduling policies under three QoS/fairness constraints for multiuser OFDM systems. We then propose a generalized opportunistic scheduling framework that incorporates multiple mixed QoS/fairness constraints, including providing both lower and upper bound constraints. Taking input queues and channel memory into consideration, we reformulate the transmission scheduling problem as a new class of Markov decision processes with fairness constraints. As an example, we investigate the throughput maximization and the delay minimization objectives in this context. Furthermore, we develop an efficient approximation method—temporal fair rollout—to reduce the computational cost in implementation.



## 1.2 Organization of This Dissertation

The organization of this dissertation is as follows:

In Chapter 2, we give the basic background knowledge. We overview the wireless cellular communication systems from the history to the key techniques. The mobile radio channel model is briefly discussed. Then we introduce the transmission scheduling in wireless networks. we address the major challenges for wireless scheduling, and explain the notion of fairness. Then, we briefly review the existing wireless transmission scheduling schemes. We also discuss the existing cross-layer design approaches in this chapter. A critical issue of dynamic resource allocation is the cross-layer optimization over time-varying, heterogeneous environments. We outline the theory of stochastic control that we will use in Chapter 5. In the end, we survey the related work to this dissertation.

In Chapter 3, we investigate the application of opportunistic scheduling in multiuser OFDM systems. we develop a rigorous framework to study opportunistic scheduling in multiuser OFDM systems. We derive optimal opportunistic scheduling policies under three QoS/fairness constraints for multiuser OFDM systems—temporal fairness, utilitarian fairness, and minimum-performance guarantees. To implement these optimal policies, we propose a modified Hungarian algorithm and a heuristic suboptimal algorithm, and compare them with non-opportunistic schemes via extensive numerical experiments.

In Chapter 4, we consider the problem of downlink transmission scheduling with general constraints. We start with considering the scheduling problems with both minimum and maximum constraints. We derive the corresponding opportunistic scheduling policies for the three long-term QoS/fairness constraints. Then we deal with scheduling problems with multiple type mixed QoS/fairness constraints. Finally, we develop a unified framework for generalized opportunistic scheduling problems which accommodates all the aforementioned scheduling schemes.

In Chapter 5, we consider the problem of fair scheduling of queued data transmissions in wireless heterogeneous networks. We deal with both the throughput maximization problem and the delay minimization problem. Taking fairness constraints and the data arrival queues into consideration, we formulate the transmission scheduling problem as a Markov decision process with fairness constraints. We study two categories of fairness constraints, namely temporal fairness and utilitarian fairness. We consider two criteria: infinite horizon expected total discounted reward and expected average reward. Applying the dynamic programming approach, we derive and prove explicit optimality equations for the above constrained MDPs, and give corresponding optimal fair scheduling policies based on those equations. A practical stochastic-approximation-type algorithm is applied to calculate the control parameters online in the policies. Furthermore, we develop a novel approximation method—temporal fair rollout—to achieve a tractable computation. Numerical results show that the proposed scheme achieves significant performance improvement for both throughput maximization and delay minimization problems compared with other existing schemes.

Chapter 6 summarizes the major contributions of this dissertation and outlines proposals for future work.

## CHAPTER 2

### BACKGROUND AND RELATED WORK

In this chapter, we provide an outline of the background knowledge which we utilize to develop our results, and review the related work. The organization of this chapter is as follows: in Section 2.1, we overview the nowadays wireless cellular communication networks from the history to the key techniques. Section 2.2 briefly discuss the mobile radio channel model. In Section 2.3, we present an introduction of the transmission scheduling in wireless networks. We first address the challenges for wireless scheduling, and introduce the notion of fairness. Then, we briefly review the existing wireless transmission scheduling schemes, including opportunistic scheduling. Section 2.4 explains the cross-layer design approaches. For enhancing the end-to-end quality of links, different layers of communication protocol should be coordinated together by cross-layer design. Section 2.5 briefly outlines the theory of stochastic control that we will utilize in Chapter 5. Finally, Section 2.6 surveys the related work in the field.

#### 2.1 Overview of Wireless Cellular Communication Networks

Over the past two decades, wireless communications have witnessed an explosive growth, and have become pervasive much sooner than anyone could have imagined [3]. Wireless networks are expected to be the dominant and ubiquitous telecommunication tools in the next few decades. The widespread success of cellular and WLAN systems

prompts the development of advanced wireless systems to provide other information services beyond voice, such as telecommuting, video conferencing, interactive media, etc., at *anyplace, anywhere, anytime*. To satisfy growing demands of heterogeneous applications, the future wireless networks are characterized by *broadband, high data rate capabilities, integration of services, flexibility, and scalability*. Many technical challenges yet remain to achieve these requirements because of the adverse natures of wireless channels.

### **2.1.1 Evolution of Wireless Cellular Communication Systems**

The cellular era just started thirty years ago with the operation of the first generation (1G) analog cellular radio systems in 1980s and each year their subscribers increased at a very fast rate. The representing 1G standard systems include Advanced Mobile Phone Service (AMPS) [4] in the United States, Nordic Mobile Telephones (NMT) in Europe, Total Access Communication Systems (TACS) in the United Kingdom, and Nippon Telephone and Telegraph (NTT) in Japan. All the 1G systems used analog frequency modulation (FM) for speech, and frequency shift keying (FSK) for signaling, and the access technique used was frequency division multiple access (FDMA).

In the 1990's, the second generation (2G) wireless cellular systems began to be introduced. Various 2G systems have been deployed around the world. Leading the pack are Global System for Mobile communications (GSM) deployed in Europe and Asia, IS-54/136 and IS-95 standards in North America, and Personal Digital Cellular (PDC) systems in Japan.

GSM was the first universal digital cellular system with modern network features extended to each mobile user. First deployed in Europe, now it is the most widely adopted standard for cellular radio systems throughout the world. The 2G systems

provide digital speech and short message services (SMS) with higher spectrum efficiency. Most 2G standards use time division multiple access (TDMA) as the access technique, except for IS-95, which is based on code division multiple access (CDMA).

2.5G is a transition step between 2G and 3G cellular wireless technologies. 2.5G provides some of the benefits of 3G (e.g. it is packet-switched) and can use some of the existing 2G infrastructure in GSM and CDMA networks. There have been several deployments of 2.5G across the world. In the USA, the 2.5G extension to CDMA systems are known as 1xEV-DO and 1xEV-DV. General packet radio service (GPRS) and Enhanced Data rate for GSM Evolution (EDGE) have been used by major GSM operators.

The third generation (3G) standard is currently being pushed as the new global standard for cellular communications. 3G networks enable network operators to offer users a wider range of more advanced services while achieving greater network capacity through improved spectral efficiency. Services include wide-area wireless voice telephony, video calls, and broadband wireless data, all in a mobile environment.

The first commercial 3G network was launched by NTT DoCoMo in Japan in 2001. Till December 2007, 190 3G networks were operating in 40 countries were operating in 71 countries, according to the Global mobile Suppliers Association (GSA). 3G networks are based on the ITU family of standards under the International mobile telecommunications 2000 (IMT-2000). IMT-2000 family includes three major radio interfaces: W-CDMA, CDMA2000, and TD-CDMA/TD-SCDMA.

The 4G (also known as Beyond 3G) [5] wireless systems represent the next complete evolution in wireless communications. It will be a complete replacement for current networks and be able to provide a comprehensive and secure all-IP solution where voice, data, and streamed multimedia can be given to users on an “anytime, anywhere” basis, and at much higher data rates than previous generations. The proposed principal technologies for 4G networks include OFDM, smart antenna, Turbo

codes, IPv6, and etc. Technologies considered to be early 4G include Flash-OFDM, the 802.16e (mobile WiMAX), and HC-SDMA.

### 2.1.2 Key Techniques of Wireless Cellular Communication Systems

In this subsection, we go over some key techniques of a wireless cellular system. We concentrate on the topics like modulation, channel coding, diversity, power control, admission control, multiple access, OFDM and OFDMA, and cellular concept.

#### Modulation

Modulation is the process of encoding information to form a message source in a manner suitable for transmission. It generally involves translating a base band message signal (called the source) to a bandpass signal at frequency that is much higher than the baseband frequency. The bandpass signal is called the modulated signal and the baseband message signal is called modulating signal. Modulation may be done by varying the amplitude, phase, or frequency of a high frequency carrier in accordance with the amplitude of the message signal. Demodulation is the process of extracting the baseband message from the carrier so that it may be processed and interpreted by the intended receiver (called the sink) [3].

For digital modulation techniques, the performance of a modulation scheme is often measured in terms of *bandwidth efficiency* and *power efficiency*. Bandwidth efficiency describes the ability of a modulation scheme to accommodate data within a limited bandwidth. The popular bandwidth efficient modulations are M-ary PAM, M-ary FSK, M-ary PSK, M-ary QAM, CPM, and MSK. Power efficiency is the ability of a modulation technique to transmit digital message with limited power. The popular power efficient modulations are M-ary orthogonal modulation and M-ary bi-orthogonal modulation. In addition to the efficiencies, other factors, such as robustness to nonlinear amplifier, performance in fading condition, and cost of transceiver, etc. also influence the choice of digital modulation.

Adaptive modulation is a promising technique to increase the data rate that can be reliably transmitted over fading channels. For this reason, some forms of adaptive modulation are being proposed or implemented in the next generation wireless systems. The basic premise of adaptive modulation is a real-time balancing of the link budget in flat fading through adaptive variation of the transmitted power level, symbol transmission rate, constellation size, BER, or any combination of these parameters. Thus, without wasting power or sacrificing BER, adaptive modulation provides a higher average link spectral efficiency (bps/Hz) by taking advantage of fading through adaptation [6].

### **Channel error control coding**

Channel error control coding arose from the seminal contribution in communication theory made by Shannon [7] that establishes fundamental limits on reliable communication, and presents the challenge of finding specific families of codes that achieve the capacity limit. Channel coding adds redundancy in the transmitted message so that if instantaneous errors occur in the received signal, the receiver can detect the errors or the data still can be recovered. The channel encoder is located between the source encoder where user's digital message sequence is produced and the modulator where the signal is modulated for transmission in the wireless channel.

There are two main types of channel codes, namely block codes, convolutional codes. The commonly used block codes are Hamming codes, Hadamard codes, Golay codes, cyclic codes, BCH codes, Reed-Solomon Codes, etc. Convolutional codes are the most widely used channel codes in practical communication systems. The main decoding strategy for convolutional codes is based on the Viterbi algorithm.

In the last decade, significant work has been done on the newly found Turbo coding/decoding [8]. Turbo coding can potentially achieve performance that is close to the Shannon capacity limits at the expense of complexity. The current 3G cellular systems employ turbo codes for relatively long (e.g., larger than 300 bits) block length

messages.

### **Diversity**

Diversity is a powerful communication technique that provides significant wireless link improvement with little added cost. It is a method for improving the reliability of a message signal by utilizing two or more communication channels with different characteristics. Diversity plays an important role in combating fading and co-channel interference and avoiding error bursts. It is based on the fact that individual channels experience different levels of fading and interference. Multiple versions of the same signal may be transmitted and/or received and combined in the receiver. Diversity techniques may exploit the multipath propagation, resulting in a diversity gain. A simple example can explain the diversity concept: If one radio path undergoes a deep fade, another independent path may have a strong signal, so the transmitted signal can still be correctly received.

The popular diversity methods are listed as follows:

- Frequency diversity

Frequency diversity is implemented by transmitting information on several frequency channels or spread over a wide spectrum to combat *frequency-selective fading*. A good example is OFDM modulation. OFDM modulation exploits frequency diversity by providing simultaneous modulation signals with error control coding across a large bandwidth, such that if a particular frequency undergoes a fade, the composite signal from all frequencies will still be demodulated.

- Space diversity

Space diversity, also known as antenna diversity, is an important diversity technique, where the signals received from spatially separated antennas have essentially uncorrelated envelopes for antenna separations of one half wavelength or



more. Space diversity reception methods can be classified into four categories: selection diversity, feedback diversity, maximal ratio combining, and equal gain diversity.

- Time diversity

Time diversity repeatedly transmits information at time spacings that exceed the coherence time of the radio channel, so that multiple repetitions of the signal will be received with independent fading conditions, thereby providing diversity. Before it is transmitted, a redundant forward error correction code is inserted and the message is spread in time by means of interleaving to resist burst errors. Rake receiver for CDMA systems is an application of time diversity by exploring the redundancy in the received signals over multipath channels.

- Space-time diversity

Multiple-input-multiple-output (MIMO) systems employing multiple transmit and receive antennas will inarguably play a significant role in the development of future broadband wireless communications. By taking diversity of the larger number of propagation paths between the transmit and receive antennas, the detrimental effects of channel fading can be significantly reduced. It has been shown that MIMO systems offer a large potential capacity increase compared to single antenna systems. To exploit this diversity, a considerable number of MIMO modulation and coding methods, also known as space-time coding (STC), have been proposed.

- Multiuser diversity

In multiuser communications, different users have different channel conditions because they are located in different locations and experience different fadings. By adaptively assigning resources such as time slots, frequency subchannels, we can take advantage of this channel diversity, which is called *multiuser diversity*.

Multiuser diversity stems from channel diversity including independent path loss and fading of users.

### **Power control**

In wireless systems, the received power represents signal strength to the desired receiver, but also interference to all other users. Power control is intended to provide each user an acceptable connection by eliminating unnecessary interference. The elegant work of Yates [9] abstracts the important properties of various power control algorithms and presents a unified treatment of power control. While power control is widely implemented in CDMA systems, such as IS-95, it has also been shown to increase the call carrying capacity for channelized systems, such as TDMA/FDMA systems [10]. Furthermore, beyond the conventional concept of power control as a means to eliminating the “near-far” effect, power control is also an effective resource management mechanism. It plays an important role in interference management, channel-quality/service-quality provisioning, and capacity management [10–12].

### **Admission control**

Empirical studies have shown that a typical user is far more irritated when an ongoing call is dropped than a call blocked from the very beginning. Hence, the purpose of admission control is to admit as many users as possible to maximize the revenue of the system while maintaining a certain level of quality of service (QoS) for ongoing connections. A new call is admitted if and only if its QoS constraints can be satisfied without jeopardizing the QoS constraints of existing calls in the network. An admission control decision is made using a traffic descriptor that specifies traffic characteristics and QoS requirements. Admission control is closely coupled with other resource allocation schemes, such as dynamic channel allocation, power control, and mobility prediction, etc. Furthermore, admission control becomes more challenging in the content of supporting multimedia services with different and multi-faceted QoS requirements in a wireless environment.

## Multiple access

Multiple access is used to allow many mobile users to share a common medium for communications. The sharing of spectrum is required to achieve high capacity by simultaneously allocating the available bandwidth to multiple users.

Frequency division multiple access (FDMA), time division multiple access (TDMA), and code division multiple access (CDMA) are the three major access techniques used to share simultaneously the limited bandwidth of radio spectrum in a wireless communication system. These multiple access techniques have been widely used in current wireless communication systems such as GSM, IS-95, cdma2000, and DECT.

FDMA assigns individual frequency bands or channels to individual subscribers. These channels are assigned on demand to users who request service. During the period of the call, no other user can share the same frequency band. TDMA systems divide the radio spectrum into time slots, and in each time slot only one subscriber is allowed to either transmit or receive. TDMA systems transmit data in a buffer-and-burst method, thus the transmission for any subscriber is noncontinuous. CDMA systems allow many subscribers to simultaneously access a given frequency allocation. User separation at the receiver is possible because each subscriber spreads the modulated waveform over a wide bandwidth using unique spreading codes. There are two basic types of CDMA, direct-sequence CDMA (DS-SS) and multi-carrier CDMA (MC-SS). Analog cellular systems (1G) use FDMA. TDMA and CDMA techniques are implemented in both 2G and 3G digital cellular systems.

By using the antenna signal processing technique, space division multiple access (SDMA) separates users' signals in different direction of arrivals(DOA). With SDMA, multiple users with different DOA are able to communicate at the same time using the same channel. In addition, the antenna can collect transmission powers from multipath components, combine them in an optimal manner, suppress interferences

from other users, and improve the received SINR. Consequently, less power is required.

In random access protocols, the channels are utilized by users attempting to access a single channel in an uncoordinated manner. Consequently, the transmissions are due to collisions by multiple users. Many packet radio (PR) access techniques are developed to handle the collisions. PR is very easy to implement, but has low spectral efficiency and may have delays. Some of the available PR access techniques are Aloha, carrier sense multiple access (CSMA), carrier sense multiple access with collision detection (CSMA/CD), data sense multiple access (DAMA), and packet reservation multiple access (PRMA) [13].

### **OFDM and OFDMA**

To provide high-data rate service, wideband transmission is necessary. In a wide-band single-carrier system, we face the problems of frequency-selective-fading and inter-symbol-interference (ISI). Furthermore, to make high-rate-data service affordable, a higher spectrum efficiency has to be achieved.

Frequency division multiplexing (FDM) is a technology that transmits multiple signals simultaneously over a single transmission path, such as a cable or wireless system. Each signal travels within its own unique frequency range (carrier), which is modulated by the data message(text, voice, video, etc.).

OFDM divides the data stream into multiple substreams to be transmitted over different *orthogonal* subchannels centered at different subcarrier frequencies. The number of substreams is chosen to make the symbol time on each substream much greater than the delay spread of the channel or, equivalently, to make the substream bandwidth less than the channel coherence bandwidth. This insures that the substreams will not experience significant ISI.

OFDM is a promising transmission technique [14,15] to combat ISI over multipath fading channels and provide efficient frequency utilization. This technique shows a great promise for high-speed wireless/ wireline data communications. A properly

coded and interleaved OFDM system is reported to exceed the performance of many other existing systems.

The discrete implementation of OFDM is sometimes called multi-carrier or discrete multi-tone modulation (DMT). It has been widely used in many applications, including Digital Audio Broadcasting (DAB) in Europe, high-speed digital subscriber lines (HDSL), asymmetric digital subscriber lines (ADSL), wireless LANs (IEEE 802.11a, 802.16), and ultra wideband (UWB) systems [13]. It is also a promising modulation schemes of choice proposed for many future cellular networks such as 4G and cognitive radio systems.

OFDMA [16] is an OFDM based multi-access technique, which has been proposed as the wireless access and signaling scheme in several next generation wireless standards. In OFDMA, the available spectrum is divided into multiple orthogonal narrowband subchannels and information symbols are transmitted in parallel over these low-rate subchannels. This method results in reduced ISI and multipath delay spread, thus improvement in capacity and attainable data rates. The rationale is that the fading on each individual subchannel is independent from user to user, so that adaptive resource allocation gives each their “best” subchannels and adapts optimally to these channels.

### **Cellular concept**

The cellular concept offers very high capacity in the limited available spectrum by applying many low power transmitters, which provide coverage to a small portion of the service area. In a cellular system, a large coverage area is broken into many small geographic areas called *cells*. Each cell is assigned with a small proportion of the total channels, and the adjacent cells are assigned with different groups of channels. The same group of channels can be reused in the cells that are enough far away so that the transmitted powers are attenuated enough and the interferences between cells are minimized. The cellular wireless networks provide a method to use limited

spectrums to serve a large number of users by reusing the channels throughout the coverage region [17].

Channels are assigned to different cells to efficiently utilize the spectrum by fixed or dynamic policies. In a fixed channel assignment, each cell is allocated a certain set of channels and each cell handles its own channel allocation independently, which is simple for implementation and fits a network with spatially uniform traffic density. In a dynamic channel assignment, the network will allocate a channel to a cell at call setup. The minimum allowable distance between co-channel cells and traffic density is considered in order to minimize the probability of blocking.

Handoff occurs when a mobile leaves the coverage area of a cell and enters the coverage area of another cell. In channelized wireless system, different radio channels will be assigned during a handoff, which is called *hard* handoff. In CDMA system such as IS-95, the assigned channel to user is not changed, but a different base station is selected for communication. This kind of handoff is called *soft* handoff.

## 2.2 Mobile Radio Propagation Model

The mobile radio channel places fundamental limitations on the performance of wireless communication systems. The three basic propagation mechanisms which impact propagation in a mobile communication system are *reflection*, *diffraction*, and *scattering* [3]. Reflection occurs when a propagating electro-magnetic wave impinges upon an object that has very large dimensions compared to the wavelength of the propagating wave. Diffraction occurs when the radio path between the transmitter and receiver is obstructed by a surface that has sharp irregularities (edges). Scattering occurs when a medium through which the wave travels consists of objects with dimensions that are small compared to the wavelength, and where the number of obstacles per unit volume is large. Scattered waves are produced by rough surfaces, small objects, or by other irregularities in the channel.

Therefore, modeling the radio channel has been one of the most difficult parts of mobile radio system design. The statistical models are applied based on measurements. In this section, we briefly describe the three major propagation models that reflect the impact of these three basic propagation mechanisms.

### **Power-law propagation**

Path loss is caused by propagation loss, where the signal is attenuated due to the distance between the transmitter and the receiver. Both theoretical and measurement-based propagation models indicate that the average received signal power decreases logarithmically with distance, whether in outdoor or indoor environments. The average large-scale path loss for an arbitrary transmitter-receiver separation is expressed as a function of distance by using a path loss exponent. For example, in the famous Lee's model [18], the path loss  $l_p$  in (dB) is

$$l_p = K + 10\alpha \log_{10}(d) - \alpha_0,$$

where  $d$  is the distance between the transmitter and receiver,  $\alpha$  is the path loss factor,  $\alpha_0$  is a correction factor used to account for different base station and mobile station (MS) antenna heights, transmit powers, and antenna gains, and  $K$  is a constant, which has different values in different environments.

### **Log-normal shadowing**

In addition to path loss, the average received signal power may be affected by shadowing from large obstacles, such as trees, buildings, or mountains. An explanation for log-normal shadowing is as follows. Consider the received signal to be the result of the transmitted signal passing through or reflecting off some random number of objects such as buildings, hills, and trees. The individual processes each attenuate the signal to some degree and the final received value is thus the product of many transmission efficiency factors. Therefore, the logarithm of the received signal equals the sum of a large number of factors, each expressed in decibels (dBs). As the number of factors becomes large, the central limit theorem shows that the distribution of

the sum can be modeled as a Gaussian distribution under fairly general assumptions. The shadowing term  $s(k)$  (dB) is modeled as a zero-mean stationary Gaussian process with autocorrelation function given by

$$E(s(k)s(k+m)) = \sigma_o^2 \xi_d^{vT/D},$$

where  $\xi_d$  is the correlation between two points separated by a spatial distance  $D$  (meters), and  $v$  is velocity of the mobile user.

### Fading

In wireless channel, reflections from small scatterers generate multiple replicas of the transmitted signal with different delay, phase, and amplitudes at the receiver. The constructive or destructive combination of these multipath signals causes signal strength fluctuation or fading. A typical time response for a multipath fading channel is shown in Figure 2.1.

If the delay spread of the received signal is significantly smaller than the symbol interval, fading only causes amplitude fluctuations. When there is no specular component in the received signal, fading can be modeled by a Rayleigh distribution:

$$p(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{\sigma^2}\right).$$

When there are scattering components as well as a dominant path, the received signal amplitude has a Ricean distribution:

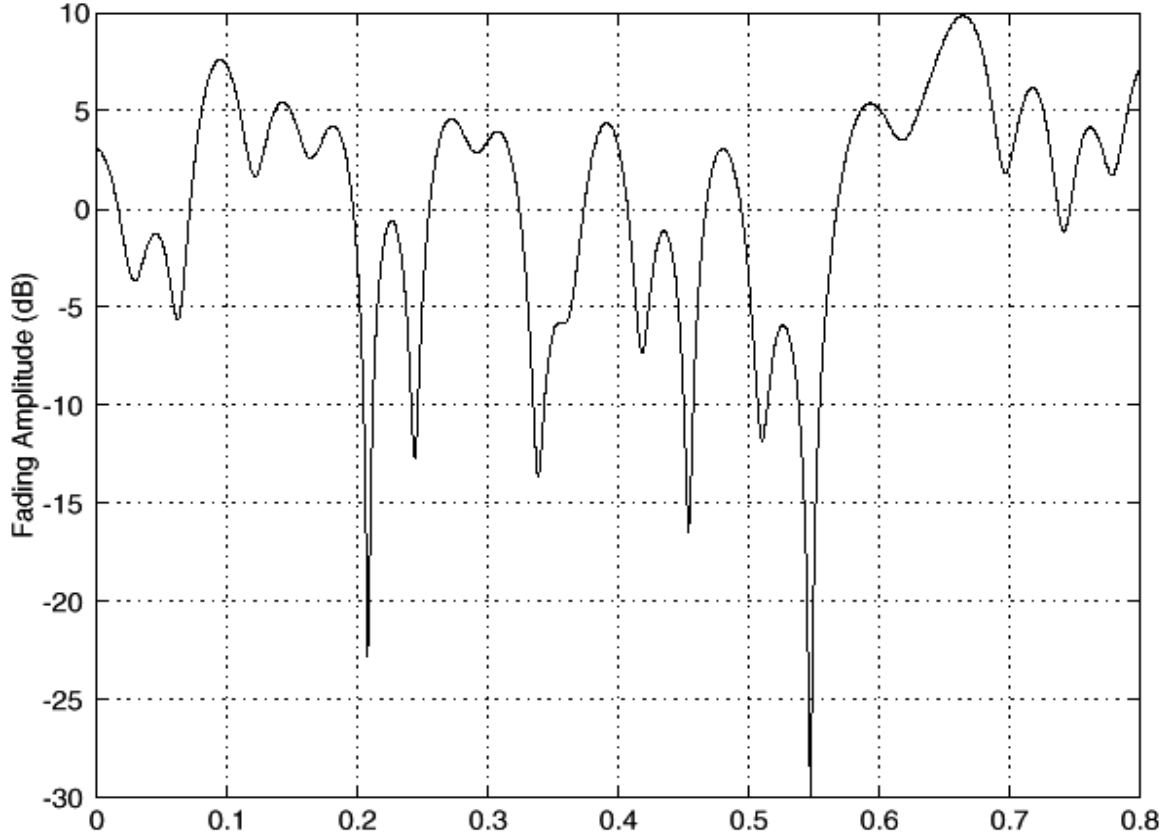
$$p(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2+A^2}{2\sigma^2}} I_0\left(\frac{Ar}{\sigma^2}\right) & A \geq 0, r \geq 0 \\ 0 & r \leq 0, \end{cases}$$

where  $I_0$  is the Bessel function of first kind and zero-order, and  $A$  denotes the peak amplitude of the dominant signal.

If the difference in time of arrival from different paths is larger than a fraction of symbol interval, in addition to fluctuations in amplitude, fading will cause frequency selective distortion as well. Received signal due to multipath signals is given by

$$r(t) = A \sum_{l=1}^L \sqrt{\alpha_l} u(t - \tau_l) e^{j(-2\pi f \tau_l)} + n(t),$$





**Figure 2.1:** Received signal strength as a function of time for vehicle velocity at 10 mph.

where  $n(t)$  is the thermal noise, and  $\tau_l$  is the delay associated with the  $l^{\text{th}}$  path.

In the meanwhile, random movement of scatters or mobile will cause doppler spread. If the mobile or scatterers are moving with speed  $v$ , the doppler shift is given by  $f_d = \frac{v}{\lambda}$ . If the doppler spread is larger than a fraction of signal bandwidth, fading can bring variation in channel response or time-selective fading. The received signal with delay and doppler shift is

$$r(t) = A \sum_{l=1}^L \sqrt{\alpha_l} u(t - \tau_l) e^{j(2\pi f_d \cos \phi_l t - 2\pi f \tau_l)} + n(t),$$

where  $\phi_l$  is the angle between the path direction and the velocity vector.

In summary, radio propagation can be roughly characterized by three nearly independent phenomena: path-loss variation, slow log-normal shadowing, and fast multi-path fading. Path losses vary with the movement of mobile stations. Slow log-normal

shadowing and fast multipath fading are time-varying with different timescales. Also the interference a user received due to other transmissions is time-varying. Furthermore, background noise is also constantly varying. All these contribute to the *time-varying* characteristics of a radio channel and motivate the need for scheduling technologies.

## 2.3 Transmission Scheduling in Wireless Networks

It is always not easy to achieve QoS in wireless networks due to several properties: *unpredictable radio link properties* as discussed above, *node mobility*, *limited energy*, and *interference from transmitters and receivers*. Therefore, transmission scheduling technologies play an important role in meeting QoS requirements for wireless networks. In this section, first, we present the resource allocation schemes and scheduling policies for wireline networks. Second, we address the characteristics and challenges for transmission scheduling in wireless networks. Third, we explain the important notion of *fairness* in network engineering and review some fairness concepts. Then, we briefly review the existing transmission scheduling schemes for wireless networks, including opportunistic scheduling, which is one of the major motivations for our research in the dissertation.

### 2.3.1 Scheduling in Wireline Networks

In wireline networks, resource allocation schemes and scheduling policies play important roles in providing service performance guarantees, such as throughput, delay, delay-jitters, fairness, and loss rate [19]. There are basically two types of scheduling disciplines: *work-conserving* and *non-work-conserving*. A work-conserving server is never idle when there is a packet to be sent. A non-work-conserving server will delay a packet until it is eligible, even when the server is idle.

Examples of work-conserving scheduling disciplines are: Delay Earliest-Due-Date (Delay-EDD), Virtual Clock, Fair Queuing (FQ), Weighted Fair Queuing (WFQ), and

Worst-case Fair Weighted Fair Queueing (WF<sup>2</sup>Q). However, with work-conserving disciplines, the traffic pattern is distorted inside the network due to fluctuations in the network load. For services that require guaranteed performance, the more important performance index is the end-to-end delay bound rather than the average delay. This is the major motivation for non-work-conserving scheduling policies. Several non-work-conserving disciplines have been proposed for packet switching networks [19]: Jitter Earliest-Due-Date (Jitter-EDD), Stop-and-Go, Hierarchical Round Robin (HRR), and Rate-Controlled Static Priority (RCSP). In addition to the challenge of providing service performance guarantees, scheduling disciplines must be simple and scalable to be implemented in real networks due to the size of wireline networks.

### 2.3.2 Challenges for Transmission Scheduling in Wireless Networks

Different assignments of the wireless resource can affect the system performance dramatically, hence, resource allocation and scheduling policies are critical in wireless networks. However, an important point to note is that the resource allocation and scheduling schemes from the wireline domain do not carry over *directly* to the wireless domain thanks to the wireless channels' unique characteristics [20],

- Channel conditions are time-varying.
- Network performance depends on channel conditions and transmission techniques.
- If the same resource is given to different users, the resultant network performance (e.g., throughput) could be different from user to user.

In wireless networks, the channel conditions of mobile users are time-varying. As discussed in Section 2.2, radio propagation can be roughly characterized by three phenomena. Path losses vary with the movement of mobile stations. Slow log-normal

shadowing and fast multipath fading are time-varying with different time-scales. Furthermore, a user receives interference from other transmissions, which is time-varying; and background noise is also constantly varying. Hence, mobile users perceive time-varying channel conditions, and these variations for each user may be independent of one another. SINR (signal to interference plus noise ratio) is a common measure of channel conditions. Apart from SINR, BER and FER (frame error rate) are also used as measure of channel conditions.

Since channel conditions are time-varying, users experience time-varying service quality and quantity. For voice users, better channel conditions may result in better voice quality. For packet data service, better channel conditions (or higher SINR) can be used to provide higher data rates using rate adaption techniques. Research had shown that cellular spectral efficiency can be increased by a factor of two or more if users with better channel conditions are served at higher data rates [6]. Procedures to exploit this are already in place for all the major cellular standards: adaptive modulation and coding schemes are implemented in TDMA, and variable spreading and coding are implemented in CDMA. In general, a user is served with better quality and/or a higher data rate when the channel condition is better. Hence, good scheduling schemes should be able to exploit the variability of channel conditions to achieve higher utilization of wireless resources.

The performance (e.g., throughput) of a user depends on the channel condition it experiences, hence, we will expect different performance when the same resource (e.g., time-slots) is assigned to different users. For example, consider a cell with two users. Suppose that user 1 has a good channel, e.g., it is close to the base station. User 2 is at the edge of the cell, where the path-loss is significant and the user experiences larger interference from adjacent cells. If the same amount of resource (power, time-slots, etc.) is assigned, it is likely that the throughput of user 1 will be much larger than that of user 2.

### 2.3.3 Fairness

The notion of *fairness* in network engineering is totally different from any respected definition of fairness from philosophy or the social sciences. The issue of fairness has been an important component in the design of optimal network flow control since it has been shown that there exist situations where a given scheme might optimize network throughput while denying access to a particular (or a set of) user(s). Loosely speaking, fairness can be thought of as a situation in which no individual class or user is denied access to the network or overly penalized. The Nash equilibrium or competitive equilibrium can be shown to be a point where no user is denied access to the network and in particular if the performance objectives are the same then it corresponds to equal throughput for each class [21]. However, fairness is difficult to quantify in the absence of a proper framework. Fairness criteria may have different implications in wireline and wireless networks. Next we will review several important fairness concepts.

#### **Max-min fairness**

The objective of max-min fairness is to maximize the minimum performance of each user can obtain under the practical constraints. Max-min fairness basically relies on the following principle: In the domain of feasible resource allocation, one user's (user 1) performance cannot be increased without decreasing some other user's (user 2) performance such that user 1's performance is better than user 2's. The compactness and convexity of the feasible region imply that such a max-min solution exists and is unique. However, the max-min fairness criterion gives an absolute priority to the user with bad conditions, which in turn will reduce the system performance.

#### **Proportional fairness**

A fairness criterion which favors the users with bad conditions less emphatically, is proportionally fair [22].

**Definition 2.1** A feasible resource allocation vector  $x_s$  for user  $s$  is *proportionally fair*, if and only if for any other feasible resource allocation vector  $x'_s$ , the sum of relative change is not positive, i.e.,

$$\sum_s \frac{x'_s - x_s}{x_s} \leq 0.$$

The physical meaning of proportional fairness is that an increase in the allocation of network resources for one user must be compensated by corresponding decreases in the allocations of one or more other users. Proportional fairness presents a tradeoff between the overall throughput and each user's throughput.

### **Utilitarian fairness**

Utilitarian fairness means that each user gets a certain share of the overall system capacity. Two extreme cases of utilitarian fairness is the system throughput maximization and max-min throughput fairness. The only objective of the system throughput maximization is to maximize the overall system throughput “*greedily*” regardless the performance of each individual user. In other words, each user is guaranteed zero percent of the system throughput. In this case, a small number of users with very good channel conditions may consume all the resource and starve other users. The objective of the max-min throughput fairness is to maximize the minimum throughput of all users. Let  $N$  be the number of users in the system. Each user is guaranteed  $1/N$  portion of the system throughput. This objective is “strictly” fair. However, when there exist users with very poor channel conditions, to achieve max-min throughput fairness will cause the significant system performance penalty [20].

### **Temporal resource-sharing fairness**

Temporal fairness means that each user is guaranteed a certain portion of the resource, i.e., time-slots. Note that temporal resource-sharing fairness is different from the utilitarian fairness. In wireline networks, when a certain amount of resource is assigned to a user, it is equivalent to granting the user a certain amount of throughput/performance value. However, the situation is different in wireless networks, where

the amount of resource and the performance value are not directly related (though closely correlated). By limiting the resource of each individual user, a user is guaranteed a certain throughput (based on its channel conditions). Resource consumed by a user can be directly connected with the price the user should pay. Premium users will obtain better services in a stochastic sense.

### **Minimum-performance guarantees**

From a user's viewpoint, a direct QoS is defined in terms of minimum-performance guarantees. In this case, each user is guaranteed a minimum-performance requirement. This type of QoS constraint is desirable for users, but difficult for the system where feasibility is a major concern.

### **2.3.4 Scheduling Schemes in Wireless Networks**

Various scheduling schemes and associated performance problems have been widely studied in wireline networks [23, 24]. However, as mentioned above, scheduling schemes from the wireline domain do not directly carry over to wireless systems because wireless channels have unique characteristics not found in wireline channels.

Transmission scheduling for wireless networks has recently attracted a lot of attention. First, scheduling policies of wireline networks are extended to wireless networks, by taking into account the bursts of errors in wireless channels. To elaborate, a wireless channel can be modeled by a two-state Markov chain also called the Gilbert-Elliot model [25, 26]. In this model, the channel at any time is assumed to be either in "good" state or "bad" state. Using such a channel model, various wireless fair scheduling policies have been proposed [27–29]. These efforts provide various degrees of performance guarantees, including short-term and long-term fairness, as well as short-term and long-term throughput bounds. A good survey of these algorithms can be found in [30]. The common limitation of these works is that channels are modeled as either "good" or "bad," which is too simple to characterize realistic wireless

channels, especially for data services.

The IS-856 system has been developed at Qualcomm to provide a versatile wireless Internet solution [31]. This system is also known as High Data Rate (HDR) [32]. The first fundamental design choice of HDR is to separate the services by including two interoperable modes: that is, 1x mode for voice and low-rate data and 1xEV mode for high-rate data services. In 1xEV mode, a single user is served at any instant (e.g, time-multiplexed CDMA); therefore avoiding power sharing and allocating the entire access point (e.g., base station) power to the user being served. The IS-856 systems use the proportional fairness scheduler. An access point always transmit at full power achieving very high peak rates for users that are in a good coverage area. An access terminal, on a slot-by-slot basis (1.67 ms), measures the pilot strength, and continuously requests an appropriate data rate based on the channel conditions.

In [33–35], the authors study scheduling algorithms for the transmission of data to multiple users. Both delay and channel conditions are taken into account. Roughly speaking, the algorithm can be described as:

$$\operatorname{argmax}_i \rho_i W_i R_i,$$

where  $W_i$  is the head-of-the-line packet delay for user queue  $i$ ,  $R_i$  is the channel capacity, and  $\rho_i$  is some constant. The proposed scheduler achieves throughput optimality, defined as follows [33]: a scheduling algorithm is *throughput optimal* if it is able to keep all queues stable if this is at all feasible to do with any scheduling algorithm. Furthermore, the authors of [34] prove the following result: to maximize the system throughput with minimum-throughput requirements, there exists some constant  $c_i$  such that one should choose a user with the maximum value of  $c_i R_i$ . In these papers, however, there is no discussion on how to obtain the values of  $c_i$  or how feasibility plays a role. Furthermore, in [36, 37], the authors study an exponential rule:

$$\operatorname{argmax}_i \rho_i R_i \exp\left(\frac{a_i W_i}{\beta + \overline{W}^\eta}\right),$$



where  $W_i$  is the queue length (or waiting time), and  $\overline{W}$  is the average queue length (or waiting time) over users. Using the exponential rule, when all queues are filled to similar capacity, the channel condition plays a significant part. On the other hand, if one queue is much longer than others, then the queue length becomes dominant and the longer queue gets a higher chance to transmit. Hence, this algorithm balances the tradeoff between queue length and throughput. The exponential rule is also *throughput optimal*.

The authors of [38] investigate a scheduling algorithm to maximize the minimum weighted throughput of users. The objective function is given by

$$\text{maximize} \quad \min_i \lim_{N \rightarrow \infty} E \left( \sum_{t=1}^N \mathbf{1}_{\{i\}} R_i(t) \right),$$

where  $R_i(t)$  is the rate of user  $i$  at time  $t$ ,  $\mathbf{1}_{\{i\}} = 1$  if time-slot  $t$  is assigned to user  $i$ , and zero otherwise. The optimal solution is

$$\operatorname{argmax}_i c_i R_i(t),$$

where  $c_i$  can be interpreted as the shadow price or reward, whose value depends on the distributions of  $R_i$ . The authors also propose an adaptive algorithm to determine the parameters  $c_i$ , and study the transient behavior.

In [39], the authors study scheduling problems for real-time traffic with fixed deadlines. Scheduling in a time-slotted system is considered; the capacity of the channel is time-varying; and the BS can estimate the channel of the current time-slot. The users achieve different QoS based on the unit prices that they are willing to pay. The objective of the base station is to maximize the revenue of the base station. The scheduling is preemptive and the base station obtains a partial revenue if a request is served partially. The unit price of a request is a non-increasing function of the time. The offline optimal scheduling scheme is shown to be NP-complete if only one user can be assigned to a time-slot. The authors then propose a greedy algorithm that chooses the request with the largest revenue in the current time-slot

to serve. The authors show that the greedy algorithm is  $1/2$  competitive against the offline optimal algorithm. Further, they show that no deterministic online algorithm can achieve a competitive ratio higher than  $1/2$ . (This does not mean that the greedy algorithm will always do better than other deterministic online algorithms.) Then the authors extend the work to various scenarios such as multi-carrier case, the case where a single slot can be shared by several users, and the case where the price is a non-increasing function of the total data that has been served to this request.

Downlink scheduling in CDMA systems for data transmission is studied in [40]. The work considers a performance metric called “stretch”, which is defined as the delay experienced by a packet normalized by its minimum achievable delay. The stretch can be considered as normalized delay. A near optimal, offline, polynomial time algorithm is proposed to minimize the maximum stretch under the assumption of continuous rates, and various online algorithms for continuous/discrete-rates are studied.

In [41], the authors investigate scheduling algorithms for uplink scheduling in CDMA. They assume that the system operates in TD/CDMA mode, with time-slotted scheduling of transmissions, assisted by periodic feedback of channel and/or congestion information through control channels. One of their observations is that it is advantageous on the uplink to schedule “strong” users one-at-a-time, and “weak” users in larger groups. This contrasts with the downlink where one-at-a-time transmission for all users has been shown to be the preferred mode.

In [42], the authors study transmission schemes for time-varying wireless channels with partial state information. A finite-state Markov chain is used to model the channel, and channel information is only available at the end of the time-slot if the transmission occurs during the time-slot. It is assumed the channel transmission matrix is unknown. The objective is to minimize a discounted infinite-horizon cost function, which can be used to indicate the balance between power cost and

throughput. An example of the cost function is:

$$C(g, s) = \begin{cases} c_0 & s = 0 \\ c_1 s + c_2 P_e(g, s) & s > 0, \end{cases}$$

where  $g$  is the state,  $s$  is the transmission power, and  $P_e$  is the error probability. The resulting optimal solution is a threshold back-off scheme: suppose a packet transmission occurs during the last time-slot and the channel state is known. If the current minimum cost is greater than  $c_0$  (no transmission cost, penalizes a scheme for placing too much emphasis on energy efficiency), then the system keeps silent for a certain number of time-slots, and then resumes transmission. The optimal transmission power is the power that minimizes the current cost function. The paper studies the effects of channel memory with partial state information, while the result may depend on the accuracy of the POMDP (Partially Observable Markov Decision Process) channel models and the estimation of transmission matrix.

Opportunistic scheduling exploits the channel fluctuations of users. Hence, the larger the channel fluctuation, the higher the scheduling gain. Thus a natural question to ask is what we should do in environments with little scattering and/or slow fading. In [43], the authors use multiple transmission antennas to “induce” channel fluctuations, and thus exploit multiuser diversity. Consider a static channel (static in the time-scale of interest) and  $N$  multiple transmission antennas. Let  $h_{ni}(t)$  be the channel gain from antenna  $n$  to user  $i$  at time  $t$ . At time  $t$ ,  $\mathbf{x}(t)$  is multiplied by  $\sqrt{a_n(t)}e^{j\theta_n(t)}$  and transmitted at antenna  $n$ ,  $i = 1; \dots; N$ , where  $\sum_{n=1}^N a_n(t) = 1$  to preserve the total transmission power. Here,  $a_n(t)$  and  $\theta_n(t)$  are random variables used to “induce” channel fluctuation. Each user feeds back the overall SINR of its “induced” channel to the base station. The base station selects the user with a largest peak value of SINR to transmit according to a certain scheduling rule. When there are a large number of users, the base station can always find a user with its peak SINR to transmit. Hence, the system performance is asymptotically as good as a

solution with an optimal beam-forming configuration, while using only the overall SINR as feedback. Note that the optimal beam-forming configuration is:

$$\begin{cases} a_n = \frac{|h_{ni}|^2}{\sum_{n=1}^N |h_{ni}|^2} & n = 1, \dots, N \\ \theta_n = -\arg(h_{ni}) & n = 1, \dots, N, \end{cases}$$

which requires individual channel information (amplitude and phase) from each antenna.

However, the potential to transmit at higher data rates opportunistically also introduces an important tradeoff between wireless resource efficiency and level of satisfaction among individual users (fairness). For example, allowing only users close to the base station to transmit at high transmission rate may result in very high throughput, but sacrifice the transmission of other users. Such a scheme cannot satisfy the increasing demand for QoS provisioning in broadband wireless networks.

To solve this problem, Liu et al. described a framework for opportunistic scheduling to exploit the multiuser diversity while at the same time satisfying three long-term QoS/fairness constraints—temporal fairness, utilitarian fairness, and minimum-performance guarantees [20, 44–46].

Opportunistic scheduling exploits the variation of channel conditions, and thus provides an additional degree of freedom in the time domain. Moreover, it can be coupled with other resource management mechanisms to further increase network performance. In the literature, opportunistic scheduling is also referred as multiuser diversity [43]. Occasionally, these two terms may have slightly different meanings. An example is the case where there is only one user in the system and the objective is to minimize transmission power while maintaining a certain data rate.

## 2.4 Cross Layer Design

It is well known that the success of today’s Internet has been based on independent and transparent protocol design in different layers, a traditional network design

approach that defines a stack of protocol layers. According to the Open System Interconnection (OSI) reference model, a communication system can be divided into seven layers from top to bottom: *Application, Presentation, Session, Transport, Network, Data Link, and Physical Layers*. Using the services provided by the lower layer, each protocol layer deals with a specific task and provides transparent service to the layer above it. Such an architecture allows the flexibility to modify or change the techniques in a protocol layer without significant impact on overall system design. However, this strict layering architecture may not be efficient for wireless networks when heterogeneous traffic is served over a wireless channel with limited and time varying capacity and high BER [47]. Efficiently utilizing the scarce radio resources with QoS provisioning requires a cross-layer joint design and optimization approach. Better performance can be obtained from information exchanges across protocol layers [48]. In this section, we briefly discuss three kinds of cross layer designs that are of great research interests recently.

#### **MAC layer and PHY layer**

The media access control (MAC) layer is a sublayer of data link layer. It provides addressing and channel access control mechanisms that make it possible for several terminals or network nodes to communicate within a network.

The physical (PHY) layer deals with signals, and provides a service to communicate bits. The concern of PHY layer is to transmit raw bits over a communication channel. The design issues largely deal with mechanical, electrical, and timing interfaces, and the physical transmission medium, which lies below the physical layer.

Because wireless channels are shared for different users, one user's transmission power is the interference for other users. Moreover in order to fully utilize the multiuser diversity, different users' rates should be controlled in such a way to optimize the overall system performance. So how to consider the resource allocation such as

rate adaptation between MAC layer and PHY layer is essential for wireless communication design. Most of our research works are focused on this type of cross layer design.

### **Application layer, MAC layer, and PHY layer**

The application layer contains a variety of protocols that are commonly needed by users. The most popular application payloads for wireless networks are voice, video, and data. For voice payload, the concern is subjective perception which is affected by the transmission delay and source encoder rate. So MAC layer and PHY layer controls are important means to guarantee the recovered voice packet qualities. For video transmission, the transmission is very bursty because of different frames and different video contents. The variable rate transmission over the lower layers can substantially improve the system performances. For data transmission, the reliability reception of data streams is the most important design issue. Therefore, powerful channel coding or ARQ is necessary for this type of application.

### **Network layer and PHY layer**

The network layer determines how packets are transferred from source to destination, which is called *routing*. A wireless ad hoc network consists of a collection of wireless nodes without a fixed infrastructure. Each node in the network serves as a router that forwards packets for other nodes. Each flow from the source to the destination traverses multiple hops of wireless links. Compared with wireline networks where flows contend only at the router with other simultaneous flows through the same router, the unique characteristics of multi-hop wireless networks show that, data stream flows also compete for shared channel bandwidth if they are within the transmission ranges of each other. This presents the problem of designing an appropriate topology aware resource allocation algorithm, so that contending multi-hop flows share the scarce channel capacity, while the total system performance is optimized [49].

## 2.5 Markov Decision Processes

Problems of sequential decision making under uncertainty are common in manufacturing, computer and communication systems. Many such decision-making problems can be formulated as Markov Decision Processes (MDPs). In these systems, an underlying Markov chain specifies the behavior of the system. However, the chain has the property that its transition law depends on the action chosen by the decision maker as well as the state of the system at each time step. MDPs originated in the study of stochastic optimal control and operations research [50] and have remained a key problem in those areas ever since. More recently, MDP models have gained recognition in such diverse fields as ecology, economics, artificial intelligence (AI), and communication engineering.

### 2.5.1 Markov Decision Process Formulation

A Markov Decision Process is a framework containing states, actions, rewards (or costs), transition probabilities and the decision horizon for the problem of optimizing a stochastic discrete-time dynamic system. The dynamic system equation is

$$x_{t+1} = f_t(x_t, u_t, w_t), \quad t = 0, 1, \dots, T,$$

where  $t$  indexes a time epoch;  $x_t$  is the state of the system;  $u_t$  is the action to be chosen at time  $t$ ;  $w_t$  is a random disturbance which is characterized by a conditional probability distribution  $P(\cdot|x_t, u_t)$ ; and  $T$  is the decision horizon. We denote the set of possible system states by  $S$  and the set of allowable actions in state  $i \in S$  by  $U(i)$ . Usually, we can assume  $S$ ,  $U(i)$ , and  $P(\cdot|x_t, u_t)$  do not vary with  $t$ . We further assume that the sets  $S$  and  $U(i)$  are finite sets, where  $S$  consists of  $n$  states denoted by  $0, 1, \dots, n - 1$ .

If, at some time  $t$ , the system is in state  $x_t = i$  and action  $u_t = u$  is applied, we incur a stage reward  $g(x_t, u_t) = g(i, u)$ , and the system moves to state  $x_{t+1} = j$  with

probability  $p_{ij}(u) = P(x_{t+1} = j | x_t = i, u_t = u)$ .  $p_{ij}(u)$  must be given a priori or may be calculated from the system equation and the known probability distribution of the random disturbance.  $g(i, u)$  is assumed bounded.

Consider the infinite horizon expected total discounted reward problem, where there is a discount factor less than one. Given an initial state  $x_0$ , we want to find a policy  $\pi = \{\mu_0, \mu_1, \dots\}$ , where  $\mu_t : S \rightarrow U$ ,  $\mu_t(i) \in U(i)$ , for all  $i$  and  $t$ , that maximizes the expected total discounted reward function

$$J_\pi(i) = \lim_{T \rightarrow \infty} E_\pi \left[ \sum_{t=0}^{T-1} \alpha^t g(x_t, \mu_t(x_t)) \middle| X_0 = i \right],$$

where  $E_\pi$  represents expectation given that a policy  $\pi$  is employed,  $\alpha$  is the discount factor with  $0 < \alpha < 1$ . A stationary policy is an admissible policy of the form  $\pi = \{\mu, \mu, \dots\}$ .

The methodology for solving MDPs is dynamic programming, based on Bellman's "*Principle of Optimality*": "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" [3]. This principle is often expressed by a system of equations called *Bellman's optimality equations*.

Therefore, under certain assumptions, the optimal expected reward-to-go function satisfies the Bellman's optimality equations

$$J_{\pi^*}(i) = \max_{u \in U(i)} \left[ g(i, u) + \alpha \sum_{j=0}^{n-1} p_{ij}(u) J_{\pi^*}(j) \right],$$

and in fact there is a unique solution of this equation. The optimal policy is implicitly specified by the above Bellman's optimality equation through the expression

$$\pi^*(i) = \operatorname{argmax}_{u \in U(i)} \left[ g(i, u) + \alpha \sum_{j=0}^{n-1} p_{ij}(u) J_{\pi^*}(j) \right].$$

The above two equations can be used to determine the optimal policy and its expected reward.



Two basic dynamic programming methodologies for solving infinite horizon MDPs are *policy iteration* and *value iteration*. Policy iteration includes a sequence of policy evaluation and policy improvement at each iteration. For problems with a total reward criterion, each policy evaluation corresponds to calculating the expected long-term reward (reward-to-go) from each state by solving linear equations with the same number of equations as the number of states; for problems with an average reward criterion, the evaluated rewards are the average rewards and differential rewards, instead of reward-to-go. Each policy improvement step involves choosing an action for each state, where the action is “greedy” with respect to the evaluated rewards. Value iteration calculates successively the optimal reward-to-go for total reward problems, or the optimal average reward and differential rewards, by turning the optimality equations into update rules; the process continues until the difference between two sequential values of the evaluated rewards is within some error bound.

### 2.5.2 Approximate Dynamic Programming

Both policy iteration and value iteration work well when an MDP model has a small or moderate size, measured by the number of states and number of actions. In many systems, however, MDP models are very large. When the number of states is very large, there would be heavy storage and computational burdens due to the large number of reward-to-go functions and the large size of the transition probability matrix in the MDP model. As the number of states increases linearly, the computational requirement increases exponentially, which leads to the so-called “*curse of dimensionality*”. When there are a large number of actions available in each state, the greedy search algorithm may lead to another form of “curse of dimensionality”.

This has motivated a broad class of approximation methods that involve more tractable computation, but yield suboptimal policies, which we refer to as *approximate*

*dynamic programming* (ADP) methods [51]. It is based on replacing the reward-to-go function  $J_\pi$  in the right-hand side of the Bellman’s optimality equation by an approximation  $\tilde{J}_\pi$ . There are two categories of approaches for selecting or calculating the functions  $\tilde{J}_\pi$ :

**Explicit reward-to-go approximation**

Here  $\tilde{J}_\pi$  is computed *offline* in one of a number of ways. One approach is to solve (optimally) a related simpler problem, obtained for example by state aggregation or by some form of enforced decomposition. The functions  $\tilde{J}_\pi$  are derived from the optimal reward-to-go functions of the simpler problem. Another approach is by introducing a parametric approximation architecture, such as a neural network or a weighted sum of basis functions or features. The idea here is to approximate the optimal reward-to-go  $J_\pi$  with a function of a given parametric form  $\tilde{J}_\pi(i) = \hat{J}_\pi(i, r_i)$ , where  $r_i$  is a parameter vector. This vector is tuned by some form of heuristic method (as for example in computer chess) or some systematic method (for example, of the type provided by the neuro-dynamic programming and reinforcement learning methodologies, such as temporal difference and Q-learning methods [51]).

**Implicit reward-to-go approximation**

Here the values of  $\tilde{J}_\pi$  are computed *online* as needed, via some computation of future rewards, starting from these states (optimal or suboptimal/heuristic, with or without a rolling horizon). We will focus on three popular schemes.

**Rollout**, where the reward-to-go of a suboptimal/heuristic policy (called the *base policy*) is used as  $\tilde{J}_\pi$ . This reward is computed as necessary in whatever from is most convenient, including by online simulation. The suboptimal policy obtained by rollout is *a one-step lookahead policy, with the optimal value function approximated by the value function of a known base policy  $\pi$* . The salient feature of the rollout algorithm is its *reward-improvement property*: the rollout policy is no worse than the performance of the base policy. In many cases, the rollout

policy is substantially better than the base policy [52, 53].

**Open-loop feedback control (OLFC)**, where an optimal open-loop computation is used, starting from the state  $i$  (in the case of perfect state information) or the conditional probability distribution of the state (in the case of imperfect state information).

**Model predictive control (MPC)**, where an optimal control computation is used in conjunction with a rolling horizon. This computation is deterministic, possibly based on a simplification of the original problem via certainty equivalence.

### 2.5.3 Constrained Markov Decision Processes

The MDPs discussed above consider only a single objective (criteria). On the other hand, there might exist several *possibly conflicting* objectives requiring a strategy that mediates between them. This is a common situation in communication networks, project management, robot control, and etc. Instead of introducing a single utility that is to be maximized (or cost to be minimized) that would be some function (e.g. some weighted sum) of the different objectives, we consider a situation where one type of reward is to be maximized while keeping the other types of rewards above some given bounds. Therefore, our control problem can be viewed as a constrained optimization problem over a given class of policies [54–57].

A constrained Markov Decision Process (CMDP) is similar to a Markov Decision Process, with the difference that the policies are now these that verify additional constraints. That is, to determine the policy  $\pi$  that:

$$\max_{\pi} J_{\pi}(i) \quad \text{subject to} \quad D_{\pi}^k(i) \geq V_k, \quad k = 1, \dots, K,$$

where  $V_k$ ,  $k = 1, \dots, K$  are some given constants;  $J_{\pi}$  and  $D_{\pi}^k$ ,  $k = 1, \dots, K$  are some reward criteria related to the policy  $\pi$ .

Several methods have been explored to solve the above CMDPs. The first one, based on on a Linear Program (LP), was introduced by Derman [54]. It is based on the fact that a CMDP can be shown to be equivalent to a LP, whose decision variables correspond to the occupation measure. There is a one-to-one correspondence between the optimal solutions of the LP and the optimal policies of the CMDP. This method is quite efficient in terms of complexity of computations, and in the amount of decision variables, and hence memory requirements for calculating the value of the CMDPs. A second method was introduced by Beutler and Ross [58] for the case of a single constraint, and is based on a Lagrangian approach. It allowed them to characterize the structure of optimal policies for the constrained problem, but it does not provide explicit computational tools. A third method, also based on an LP, was introduced by Altman and Schwartz [59] and further studied by Ross [60]. It is based on some mixing (by time-sharing) of stationary deterministic policies (these are policies that depend only on the current state and do not require randomization). A similar LP approach was later introduced by Feinberg [57] for finite MDPs (finite state and action spaces), where the mixing is done in a way that equivalent to having an initial randomization between stationary deterministic policies. These approaches require in general a huge number of decision variables. However, there are special applications where this LP can have an extremely efficient solution.

## 2.6 Related Work

Next generation wireless networks, which support high-speed packet data while providing heterogeneous QoS guarantees, require flexible and efficient radio resource scheduling schemes. Transmission scheduling for wireless networks has attracted a lot of recent attention [27–30, 61]. In contrast to wireline networks, one of the fundamental characteristics of wireless networks is the time-varying and location-dependent channel conditions due to multipath fading.

From an information-theoretic viewpoint, Knopp and Humblet showed that the system capacity is maximized by exploiting inherent *multiuser diversity* gain in the wireless channel [62]. The basic idea is to schedule a single user with the best instantaneous channel condition to transmit at any one time. Technology based on this idea has already been implemented in the current 3G systems: HDR [32] and high-speed downlink packet access (HSDPA) [63].

Good scheduling schemes in wireless networks should *opportunistically* seek to exploit the time-varying channel conditions to improve spectrum efficiency thereby achieving *multiuser diversity* gain. However, the potential to transmit at higher data rates opportunistically also introduces an important tradeoff between wireless resource efficiency and level of satisfaction among individual users (fairness). For example, allowing only users close to the base station to transmit at high transmission rate may result in very high throughput, but sacrifice the transmission of other users. Such a scheme cannot satisfy the increasing demand for QoS provisioning in broadband wireless networks.

To solve this problem, Liu et al. described a framework for opportunistic scheduling to exploit the multiuser diversity while at the same time satisfying three long-term QoS/fairness constraints—temporal fairness, utilitarian fairness, and minimum-performance guarantees [44–46]. In that work, only a single user can transmit at each scheduling time. The authors of [62] show that this is optimal for single-channel systems such as TDMA. However, the same is not the case for multiple-channel systems.

Opportunistic scheduling exploits the channel fluctuations of users. In [43], the authors use multiple “dumb” antennas to “induce” channel fluctuations, and thus exploit multiuser diversity in a slow fading environment. The authors of [64] show that with multiple antennas, transmitting to a carefully chosen subset of users has superior performance.

The authors of [27, 28] extend the scheduling policies for wireline networks to

wireless networks to provide short-term and long-term fairness bounds. However, they model a channel as being either “good” or “bad,” which may be too simple in some situations. In [34, 35, 65], the authors study wireless scheduling algorithms when both delay and channel conditions are taken into account. Scheduling with short-term fairness constraints is also discussed in [44, 66]. In [32, 67], the authors present a scheduling scheme for the Qualcomm IS-856 (also known as HDR: High Data Rate) system. Their scheduling scheme exploits time-varying channel conditions while satisfying a certain fairness constraint known as *proportional fairness* [22]. For a detailed survey of wireless scheduling techniques, see Section 2.3.4.

OFDM is a popular multiaccess scheme widely used in DVB, wireless LANs (e.g., 802.16, ETSI HIPERLAN/2), and ultra wideband (UWB) systems [13]. It is also a promising modulation scheme of choice proposed for many future cellular networks such as cognitive radio systems [14, 15]. To enhance spectrum efficiency, the spectrum pooling system allows a license owner to share underutilized licensed spectrum with a secondary wireless system during its idle times [15]. A preferred transmission mode of the secondary system is OFDM due to its inherent flexibility. In [68], the authors discuss the desired properties in designing physical layers of cognitive radio systems and claim that the modulation scheme based on OFDM is a natural approach that satisfies the desired properties.

The resource management problem in OFDM systems has attracted a lot of research interest [69, 70]. In [69], the authors propose an algorithm to minimize the total transmission power with minimum-rate constraints for users. Specifically, the algorithm allocates a set of subcarriers to each user and then determines the number of bits and transmission power on each subcarrier. In [70], the authors study the problem of dynamic subcarrier and power allocation with the objective to maximize the minimum of the users’ data rates subject to a total transmission power constraint. All these studies show that dynamic resource allocation (in terms of bit, subcarrier, and

power) schemes can achieve significant performance gains over traditional static allocations (such as TDMA-OFDM and FDMA-OFDM). However, none of the schemes described above exploit multiuser diversity. For delay-insensitive data service, we can expect to reap long-term performance benefits by exploiting multiuser diversity.

Recently, there has been significant interest in opportunistic scheduling and fairness issues for multiple-channel systems [71–75]. In [73], the authors consider a total-throughput maximization problem with deterministic and probabilistic constraints for multiple-channel systems. In [75], the authors consider opportunistic fair scheduling in downlink TDMA systems employing multiple transmit antennas and beamforming.

In [76], the authors introduce cross-layer optimization for OFDM wireless networks. The interaction between the physical (PHY) layer and media access control (MAC) layer is exploited to balance the efficiency and fairness of wireless resource allocation. The authors consider proportional and max-min fairness.

Although there has been considerable recent efforts on proportional fairness scheduling for multiple-channel systems [43, 77, 78], to the best of our knowledge there is currently no work considering multiuser OFDM systems with the three QoS fairness constraints we mentioned above. Therefore, in Chapter 3, we propose an opportunistic scheduling framework for multiuser OFDM systems under three long-term QoS/fairness constraints—temporal fairness, utilitarian fairness, and minimum-performance guarantees. We build on Liu’s work by going from the single-channel to the multiple-channel case. We show how the system performance can be optimized by serving multiple users simultaneously over the different subcarriers.

Practical radio channels are commonly modeled as multipath Rayleigh fading channels, which are *correlated* random processes. However, much of the prior work on scheduling, including [44–46], is based on the relatively simple *memoryless* channel models. Finite-state Markov channel (FSMC) models have been found to be accurate in modeling such channels with memory [79]. When channel memory is taken into

consideration, the existing work on memoryless channels does not apply directly. Also, much of the previous work focused on “elastic” traffic [22], and assumed that the system has *infinite backlogged* data queues, which is not always an appropriate assumption. This assumption makes it impossible to consider the data arrival queues and further evaluate the system delay performance.

The widely studied MDPs and the associated dynamic programming methodology provide us with a general framework for posing and analyzing problems of sequential decision making under uncertainty [80–82]. Constrained Markov decision processes have been studied mostly via linear programming and Lagrangian methods [54–57]. However, the present dynamic programming approach does not directly treat long-term fairness constraints—we show later how such constraints represent the users’ fairness guarantees.

In [60], the authors considered a Markov decision problem to maximize the long-run average reward subject to multiple long-run average cost constraints. A linear program produces the optimal policy with limited randomization. In [83], the authors considered controlled Markov models with total discounted expected losses. They used a dynamic programming approach to find optimal admissible strategies, where *admissibility* means meeting a set of given constraint inequalities. The solution in their approach is a function of a probability distribution and the admissible expected loss, and randomization is allowed. In [84], the authors derived the dynamic programming equations for discounted cost minimization problems subject to a single identically structured constraint. The equations allowed them to characterize the contraction-type structure of optimal policies, but the authors did not provide exact solution schemes and explicit computational tools for their approximation.

On the other hand, because of the notorious “curse of dimensionality,” even exact solution schemes, such as value iteration and policy iteration, often cannot be applied directly in practice to solve MDP problems. This has motivated a broad



class of approximation methods that involve more tractable computation, but yield suboptimal policies, which we refer to as *approximate dynamic programming* (ADP) methods [51]. For example, Bertsekas and Castanon proposed an approximate approach called *rollout* for deterministic and stochastic problems that are usually computationally intractable [52, 53].

## CHAPTER 3

# OPPORTUNISTIC SCHEDULING FOR OFDM SYSTEMS WITH FAIRNESS CONSTRAINTS

In this chapter, we consider the problem of downlink scheduling for OFDM systems. *Opportunistic scheduling* exploits the time-varying, location-dependent channel conditions to achieve multiuser diversity. Previous work in this area has focused on single-channel systems. Multiuser OFDM allows multiple users to transmit simultaneously over multiple channels. Here, we develop a rigorous framework to study opportunistic scheduling in multiuser OFDM systems. We derive optimal opportunistic scheduling policies under three QoS/fairness constraints for multiuser OFDM systems—temporal fairness, utilitarian fairness, and minimum-performance guarantees. Our scheduler decides not only *which time-slot*, but also *which subcarrier* to allocate to each user. Implementing these optimal policies involves solving a maximal bipartite matching problem at each scheduling time. To solve this problem efficiently, we apply a modified Hungarian algorithm and a relatively simple suboptimal algorithm. At last, we compare our schemes with non-opportunistic schemes via numerical experiments.

This chapter is organized as follows. In Section 3.1, we give the motivation for our work. The system model is described in Section 3.2. In Section 3.3, we derive opportunistic scheduling policies under various fairness constraints, and prove their

optimality. In Section 3.4, we address some implementation issues, including control parameter estimation and the assignment problem that arises in implementing these policies. An optimal algorithm and an efficient suboptimal algorithm are proposed here. In Section 3.5, we present the numerical results to illustrate the performance of our policies. Finally, concluding remarks are given in Section 3.6.

## 3.1 Motivation

Emerging broadband wireless networks which support high-speed packet data with different quality of service (QoS) demand more flexible and efficient use of the scarce spectral resource. In contrast to wireline networks, one of the fundamental characteristics of wireless networks is the time-varying and location-dependent channel conditions due to multipath fading. From an information-theoretic viewpoint, Knopp and Humblet showed that the system capacity can be maximized by exploiting inherent *multiuser diversity* in the wireless channel [62]. The basic idea is to schedule a single user with the best instantaneous channel condition to transmit at any one time. The technology has already been implemented in the current 3G systems, i.e., 1xEV-DO [31] and high-speed downlink packet access (HSDPA) [63]. The idea has also recently been adopted in cognitive radio systems which are novel intelligent wireless communication systems providing highly reliable and efficient communications by exploiting unused radio spectrum [85, 86].

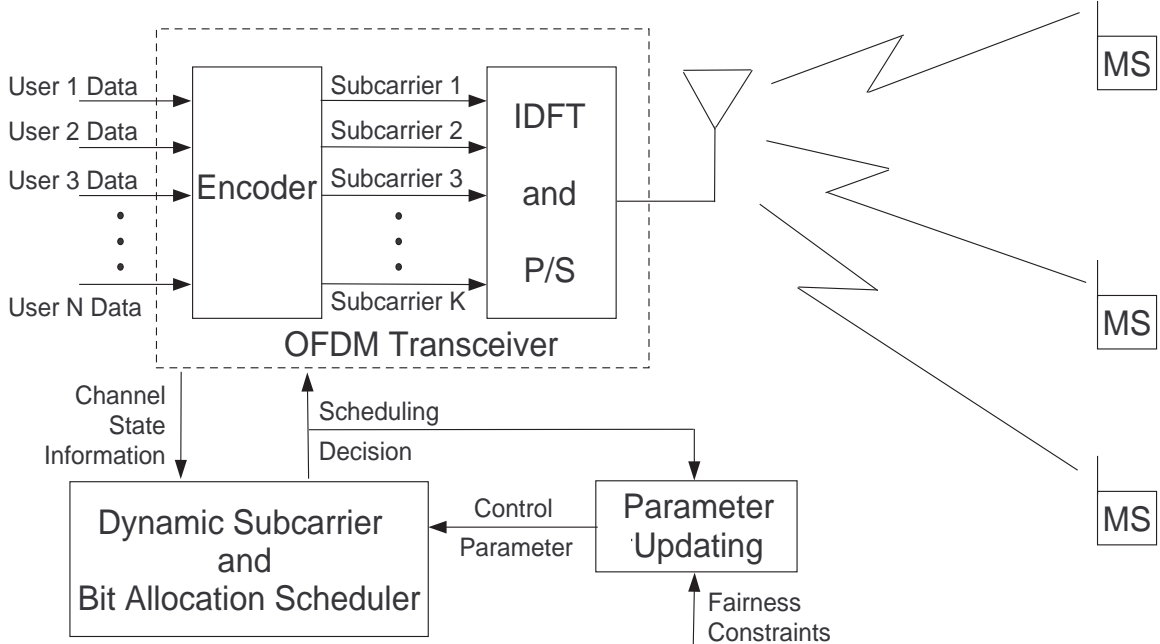
OFDM is a popular multiaccess scheme widely used in DVB, wireless LANs (e.g., 802.16, ETSI HIPERLAN/2), and UWB systems [13]. It is also a promising modulation scheme of choice proposed for many future cellular networks such as cognitive radio systems [14, 15]. OFDM divides the total bandwidth into many narrow-band orthogonal subcarriers, which are transmitted in parallel, to combat frequency-selective fading and achieve higher spectral utilization. OFDMA, a multiuser version

of OFDM, allows multiple users to transmit simultaneously on the different subcarriers [16].

Good scheduling schemes in wireless networks should *opportunistically* seek to exploit the time-varying channel conditions to improve spectrum efficiency thereby achieving *multiuser diversity* gain. However, the potential to transmit at higher data rates opportunistically also introduces an important tradeoff between wireless resource efficiency and level of satisfaction among individual users (fairness). For example, allowing only users close to the base station to transmit at high transmission rate may result in very high throughput, but sacrifice the transmission of other users. Such a scheme cannot satisfy the increasing demand for QoS provisioning in broadband wireless networks.

To solve this problem, Liu et al. described a framework for opportunistic scheduling to exploit the multiuser diversity while at the same time satisfying three long-term QoS/fairness constraints [44–46]. In that work, only a single user can transmit at each scheduling time. The authors of [62] show that this is optimal for single-channel systems such as TDMA. However, the same is not the case for multiple-channel systems.

In this chapter, we propose an opportunistic scheduling framework for multiuser OFDM systems. We build on Liu’s work by going from the single-channel to the multiple-channel case. We show how the system performance can be optimized by serving multiple users simultaneously over the different subcarriers. We focus on the downlink of an OFDM system. We derive our opportunistic scheduling policies under three long-term QoS/fairness constraints—temporal fairness, utilitarian fairness, and minimum-performance guarantees, which are similar in form to those of [46], but adapted to the setting of multiuser OFDM systems. We also state optimality conditions under each of these constraints. In particular, our scheduler decides not only *which time-slot*, but also *which subcarrier* to allocate to each user under the given QoS/fairness constraints. A stochastic approximation algorithm is used to calculate



**Figure 3.1:** Downlink scheduling over multiuser OFDM system

the control parameters online in the policies. To search over the optimal user subsets efficiently, we apply a modified bipartite matching algorithm. We also develop an efficient, low-complexity suboptimal algorithm—our experimental results illustrate that this algorithm achieves near-optimal performance.

## 3.2 System Model

In this section, we describe the system model, assumptions, notation, and formulation of the scheduling problem.

The architecture of a downlink data scheduler for a single-cell multiuser OFDM system is depicted in Figure 3.1. There is a base station (transmitter) with a single antenna communicating with  $N$  mobile users (receivers). Each user has different channel conditions over different subcarriers. By inserting pilot symbols in the downlink, the users can effectively estimate the channels. Every user should report its channel-state information over every subcarrier to the base station. All the channel-state information is sent to the subcarrier and bit allocation scheduler in the base station through

feedback channels from all mobile users. The scheduling decision made by the scheduler is conveyed to the OFDM transmitter. The transmitter then assigns different transmission rates to scheduled users on corresponding subcarriers. The scheduler makes decisions once every time-slot based on the channel-state information and the control parameters for fairness guarantees.

We assume that the base station knows the perfect channel-state information for each user over each subcarrier. The channel conditions for different users are usually independently varying in a multiuser system. Owing to frequency-selective fading, one user may experience deep fading in some subcarriers, but relatively good in other subcarriers. By dynamically assigning users to favorable subcarriers, the overall performance of the network can be increased from the multiuser diversity. In practice, requiring “perfect” channel-state information results in significant feedback overhead burden, which might be difficult to implement. We can view our current work as providing fundamental performance bounds on what is achievable with channel feedback.

The OFDM signaling is time-slotted. The length of a time-slot is fixed and the channel does not vary significantly during a time-slot. The length of a time-slot in the scheduling policy can be different from an actual time-slot in the physical layer. It depends on how fast the channel conditions vary and how fast we want to track such changes.

We assume that there is always data for each user to receive, i.e., the system has *infinite backlogged* data queues. We also assume that the transmission power is uniformly allocated to all subcarriers. In principle, performance can be improved further by allocating a different power level to each subcarrier. In a system with a large number of users, this improvement could be *marginal* because of statistical effects [43].

In this work, we will focus on scenarios with large numbers of users, or heavy-traffic

systems, where the number of users is greater than the number of available OFDM subcarriers. These scenarios can be regarded as an extreme situation for OFDM. But it is important to determine the impact of a large number of users, such as in [43]. Our goal is *to maximize the system performance by exploiting the time-varying and frequency-varying channel conditions while maintaining certain QoS/fairness constraints.*

Let  $i = 1, \dots, N$  be the index of users, and  $k = 1, \dots, K$  be the index of subcarriers. Let  $\omega_{i,k}^t$  be the instantaneous performance value that would be experienced by user  $i$  if it were scheduled to transmit over subcarrier  $k$  at time-slot  $t$ . The  $\omega_{i,k}^t$  comprise a  $N \times K$  matrix, denoted as  $\omega^t$ . Usually, the better the channel condition of user  $i$  over subcarrier  $k$ , the larger the value of  $\omega_{i,k}^t$ . Throughput (in term of data rate bits/sec) is the most straightforward form of a time-varying and channel-condition-dependent performance measure. For convenience, the reader can think of throughput as the performance measure in this chapter. However, our formulation applies in general.

Let  $\bar{A} = (A_1, A_2, \dots, A_K)$  represent a *scheduling action*, which is a vector of the indices of the users scheduled over all  $K$  subcarriers. The decision rule  $\pi^t(\cdot)$ , which is a function of  $\omega^t$ , specifies which action be chosen, i.e.,  $\pi^t(\omega^t) = \bar{A}^t = (A_1^t, A_2^t, \dots, A_K^t)$ , where the value of  $A_k^t$  is the index of the user scheduled over subcarrier  $k$  at time  $t$ . We call  $\pi(\cdot) = \{\pi^1(\cdot), \pi^2(\cdot), \dots, \pi^t(\cdot), \dots\} \in \Pi$  a *policy*, where  $\Pi$  is the set of all scheduling policies. Note that a policy may involve a *time-varying* rule for deciding scheduling actions. We are only interested in so-called *feasible* policies, those that satisfy specific QoS/fairness requirements (described in the next section).

Let  $U_i^T(\pi)$  be the average throughput of user  $i$  up to time  $T$ , and  $R_i^T(\pi)$  the

average resource consumption of user  $i$  up to time  $T$ , i.e.,

$$U_i^T(\pi) = \frac{1}{T} \sum_{t=1}^T \sum_{k=1}^K \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}}, \quad i = 1, \dots, N,$$

$$R_i^T(\pi) = \frac{1}{T} \sum_{t=1}^T \sum_{k=1}^K \mathbf{1}_{\{A_k^t=i\}}, \quad i = 1, \dots, N,$$

where  $\mathbf{1}_A$  is the indicator function of the event  $A$ , i.e.,  $\mathbf{1}_A$  takes value 1 if  $A$  occurs, and is 0 otherwise.

Let  $U^T(\pi) = \sum_{i=1}^N U_i^T(\pi)$ ; i.e.,  $U^T(\pi)$  is the average overall throughput up to time  $T$ . Then we define

$$U(\pi) = \limsup_{T \rightarrow \infty} U^T(\pi),$$

which can be considered as the asymptotic best-case system performance of policy  $\pi$ .

Using the above notation, our goal can be formally stated as: *find a feasible policy  $\pi$  that maximizes the system performance  $U(\pi)$  while maintaining certain QoS/fairness constraints*. In the following section, we derive optimal policies for three categories of scheduling problems, each with a unique QoS/fairness requirement.

### 3.3 Opportunistic Scheduling under Various Fairness Constraints

Good scheduling schemes should be able to exploit the time-varying channel conditions of users to achieve higher utilization of wireless resources, while at the same time guarantee some level of fairness among users. Fairness is central to scheduling problems in wireless systems. Without a good fairness criterion, the system performance can be trivially optimized, but might prevent some users from accessing the network resource. In this section, we will study scheduling problems under three fairness criteria for multiuser OFDM systems—temporal fairness, utilitarian fairness, and minimum-performance guarantees. These categories of fairness are adopted from [46] and are extended to multiuser OFDM systems. It turns out that the form of the optimal policies here bear a resemblance to those of [46].



### 3.3.1 Temporal Fairness Scheduling

A natural fairness criterion is to give each user a certain long-term fraction of time, because time is the basic resource shared among users. The problem of multiuser OFDM scheduling with temporal fairness can be expressed as:

$$\begin{aligned} & \max_{\pi \in \Pi} U(\pi) \\ & \text{subject to } \liminf_{T \rightarrow \infty} R_i^T(\pi) \geq r_i, \quad i = 1, \dots, N, \end{aligned} \quad (3.1)$$

where  $r_i$  denotes the minimum time-fraction that should be assigned to user  $i$ , with  $r_i \geq 0$  and  $\sum_{i=1}^N r_i \leq 1$ . Recall that  $R_i^T(\pi)$  is the average resource consumption of user  $i$  up to time  $T$ . The  $r_i$ s are predetermined and serve as the prespecified fairness constraints. The value of  $r_i$  denotes the minimum fraction of time that user  $i$  should transmit over all the subcarriers in the long run, which is usually determined by the user's class, the price paid by the user, etc.

Define the policy  $\pi^*$  as follows:

$$\pi^*(\omega^t) = \operatorname{argmax}_{\bar{A}^t} \left\{ \sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t=i\}} \right\}, \quad (3.2)$$

where the control parameters  $v_i^*$  are chosen such that:

1.  $v_i^* \geq 0, \forall i$ ;
2.  $\liminf_{T \rightarrow \infty} R_i^T(\pi^*) \geq r_i, \forall i$ ;
3. If  $\liminf_{T \rightarrow \infty} R_i^T(\pi^*) > r_i$ , then  $v_i^* = 0, \forall i$ .

Similar to [44], we can think of  $\bar{v}^* = (v_1^*, \dots, v_N^*)$  in (3.2) as an “offset” or “threshold” to satisfy the temporal fairness constraints. Under this constraint, the scheduling policy schedules the “relatively-best” subset of users to transmit. The subset of users selected by action  $\bar{A}^t$  is “relatively-best” if  $\sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t=i\}}$  is maximum over all actions. If  $v_i^* > 0$ , then user  $i$  is an “unfortunate” user, i.e., the channel conditions it experiences over all subcarriers are relative poor. (For example, it is far

from the base station.) Hence, it has to take advantage of other users (e.g., users with  $v_i^* = 0$ ) to satisfy its fairness requirement. But to maximize the overall system performance, we can only give the “unfortunate” users their minimum time-fraction requirements; hence condition 3.

The policy  $\pi^*$  defined in (3.2), which represents our opportunistic scheduling policy, is optimal in the following sense.

**Theorem 3.1** *If  $\lim_{T \rightarrow \infty} R_i^T(\pi^*)$  exists for all  $i$  for  $\pi^*$ , then the policy  $\pi^*$  is an optimal solution to the problem defined in (3.1), i.e., it maximizes the average OFDM system performance under the temporal fairness constraints.*

*Proof:* Let  $\pi$  be a policy satisfying the temporal fairness constraints, and  $v_i^*$  satisfies conditions 1–3. Hence, we have

$$\begin{aligned}
U(\pi) &\leq U(\pi) + \sum_{i=1}^N v_i^* \left( \liminf_{T \rightarrow \infty} R_i^T(\pi) - r_i \right) \\
&= \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{k=1}^K \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}} + \sum_{i=1}^N v_i^* \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{k=1}^K \mathbf{1}_{\{A_k^t=i\}} - \sum_{i=1}^N v_i^* r_i \\
&\leq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{k=1}^K \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}} \\
&\quad + \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{k=1}^K v_i^* \mathbf{1}_{\{A_k^t=i\}} - \sum_{i=1}^N v_i^* r_i \tag{3.3}
\end{aligned}$$

$$\leq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t=i\}} - \sum_{i=1}^N v_i^* r_i. \tag{3.4}$$

By the definition of  $\pi^*$ , we have

$$\sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t=i\}} \leq \sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{(A_k^t)^*=i\}}.$$

Thus,

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t=i\}} \leq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{(A_k^t)^*=i\}}.$$

Therefore,

$$\begin{aligned}
U(\pi) &\leq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^N \sum_{k=1}^K (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{(A_k^t)^* = i\}} - \sum_{i=1}^N v_i^* r_i \\
&\leq U(\pi^*) + \limsup_{T \rightarrow \infty} \sum_{i=1}^N v_i^* R_i^T(\pi^*) - \sum_{i=1}^N v_i^* r_i \tag{3.5}
\end{aligned}$$

$$\leq U(\pi^*) + \sum_{i=1}^N v_i^* \left( \limsup_{T \rightarrow \infty} R_i^T(\pi^*) - r_i \right). \tag{3.6}$$

Since  $\lim_{T \rightarrow \infty} R_i^T(\pi^*)$  exists,  $\limsup_{T \rightarrow \infty} R_i^T(\pi^*) = \liminf_{T \rightarrow \infty} R_i^T(\pi^*)$ . Thus,

$$\begin{aligned}
U(\pi) &\leq U(\pi^*) + \sum_{i=1}^N v_i^* \left( \liminf_{T \rightarrow \infty} R_i^T(\pi^*) - r_i \right) \\
&= U(\pi^*), \tag{3.7}
\end{aligned}$$

where the second part of (3.7) equals zero because of condition 3 on  $v_i^*$ . ■

Inequalities (3.3), (3.4), (3.5), and (3.6) follow from the following properties of  $\limsup$  and  $\liminf$  [87]: If  $\{x_n\}$  and  $\{y_n\}$  are real sequences, we have

$$\begin{aligned}
\liminf_{n \rightarrow \infty} x_n + \liminf_{n \rightarrow \infty} y_n &\leq \liminf_{n \rightarrow \infty} (x_n + y_n) \leq \limsup_{n \rightarrow \infty} x_n + \liminf_{n \rightarrow \infty} y_n \\
&\leq \limsup_{n \rightarrow \infty} (x_n + y_n) \leq \limsup_{n \rightarrow \infty} x_n + \limsup_{n \rightarrow \infty} y_n.
\end{aligned}$$

It is possible that the optimal policy is confronted with a tie between two or more users. When ties occur in the argmax in the policy, they can be broken arbitrarily.

### 3.3.2 Utilitarian Fairness Scheduling

In the last section, we studied the opportunistic scheduling problem for multiuser OFDM with temporal fairness constraints. In wireline networks, when a certain amount of resource is assigned to a user, it is equivalent to granting the user a certain amount of throughput. However, the situation is different in wireless networks, where the performance value and the amount of resource are not directly related. Therefore, a potential problem in wireless network is that the temporal fairness scheme has no way of explicitly ensuring that each user receives a certain guaranteed fair amount

of utility. Hence, in this section we will describe an alternative scheduling problem that would ensure that all users get at least a certain fraction of the overall system performance.

The problem of multiuser OFDM scheduling with utilitarian fairness can be expressed as:

$$\begin{aligned} & \max_{\pi \in \Pi} U(\pi) & (3.8) \\ & \text{subject to } \liminf_{T \rightarrow \infty} U_i^T(\pi) \geq a_i U(\pi), \quad i = 1, \dots, N, \end{aligned}$$

where  $a_i$  denotes the minimum fraction of the overall average throughput required by user  $i$ , with  $a_i \geq 0$  and  $\sum_{i=1}^N a_i \leq 1$ . Recall that  $U_i^T(\pi)$  is the average throughput of user  $i$  up to time  $T$  using policy  $\pi$ , and  $U(\pi)$  is the average overall throughput. The  $a_i$ s are predetermined fairness constraints here. This constraint requires long-term fairness in terms of performance value (throughput) instead of resource consumption (time) as in Section 3.3.1.

We define the policy  $\pi^*$  as follows:

$$\pi^*(\omega^t) = \operatorname{argmax}_{\bar{A}^t} \left\{ \sum_{i=1}^N \sum_{k=1}^K (\kappa + \gamma_i^*) \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}} \right\}, \quad (3.9)$$

where  $\kappa = 1 - \sum_{i=1}^N a_i \gamma_i^*$ , and the control parameters  $\gamma_i^*$  are chosen such that:

1.  $\gamma_i^* \geq 0, \forall i$ ;
2.  $\liminf_{T \rightarrow \infty} U_i^T(\pi^*) \geq a_i U(\pi^*), \forall i$ ;
3. If  $\liminf_{T \rightarrow \infty} U_i^T(\pi^*) > a_i U(\pi^*)$ , then  $\gamma_i^* = 0, \forall i$ .

Analogous to  $\vec{v}^*$  in the last section,  $\vec{\gamma}^* = (\gamma_1^*, \dots, \gamma_N^*)$  in (3.9) can be considered as a “scaling” to satisfy the utilitarian fairness constraints. The scheduling policy always schedules the “relatively-best” subset of users to transmit. Here the subset of users selected by action  $\bar{A}^t$  is “relatively-best” if  $\sum_{i=1}^N \sum_{k=1}^K (\kappa + \gamma_i^*) \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}}$  is

maximum over all actions. If  $\gamma_i^* > 0$ , then user  $i$  is an “unfortunate” user, and its average performance value equals its minimum requirement.

The policy  $\pi^*$  defined in (3.9), which represents our opportunistic scheduling policy, is optimal in the following sense.

**Theorem 3.2** *If  $\lim_{T \rightarrow \infty} U_i^T(\pi^*)$  exists for all  $i$  for  $\pi^*$  defined in (3.9), then the policy  $\pi^*$  is an optimal solution to the problem defined in (3.8), i.e., it maximizes the average OFDM system performance under the utilitarian fairness constraints.*

*Proof:* Let  $\pi$  be a policy satisfying the utilitarian fairness constraints, and  $\gamma_i^*$  satisfies conditions 1–3. Hence, we have

$$\begin{aligned}
U(\pi) &\leq U(\pi) + \sum_{i=1}^N \gamma_i^* \left( \liminf_{T \rightarrow \infty} U_i^T(\pi) - a_i U(\pi) \right) \\
&= \limsup_{T \rightarrow \infty} \sum_{i=1}^N \kappa U_i^T(\pi) + \sum_{i=1}^N \gamma_i^* \liminf_{T \rightarrow \infty} U_i^T(\pi) \\
&\leq \limsup_{T \rightarrow \infty} \sum_{i=1}^N \kappa U_i^T(\pi) + \liminf_{T \rightarrow \infty} \sum_{i=1}^N \gamma_i^* U_i^T(\pi) \\
&\leq \limsup_{T \rightarrow \infty} \sum_{i=1}^N (\kappa + \gamma_i^*) U_i^T(\pi),
\end{aligned}$$

where  $\kappa = 1 - \sum_{i=1}^N a_i \gamma_i^*$ . By the definition of  $\pi^*$ , we get

$$\sum_{i=1}^N \sum_{k=1}^K (\kappa + \gamma_i^*) \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}} \leq \sum_{i=1}^N \sum_{k=1}^K (\kappa + \gamma_i^*) \omega_{i,k}^t \mathbf{1}_{\{(A_k^t)^*=i\}}.$$

Thus,

$$\sum_{i=1}^N (\kappa + \gamma_i^*) U_i^T(\pi) \leq \sum_{i=1}^N (\kappa + \gamma_i^*) U_i^T(\pi^*).$$

Therefore,

$$\begin{aligned}
U(\pi) &\leq \limsup_{T \rightarrow \infty} \sum_{i=1}^N (\kappa + \gamma_i^*) U_i^T(\pi^*) \\
&\leq U(\pi^*) + \sum_{i=1}^N \gamma_i^* \left( \liminf_{T \rightarrow \infty} U_i^T(\pi^*) - a_i U(\pi^*) \right) \\
&= U(\pi^*),
\end{aligned} \tag{3.10}$$

where the second part of (3.10) equals zero because of condition 3 on  $\gamma_i^*$ . Similar to the proof of Theorem 1, the properties of lim sup and lim inf are applied here. ■

### 3.3.3 Minimum-Performance Guarantee Scheduling

So far, we have discussed two optimal multiuser OFDM scheduling policies that provide users with different fairness guarantees. However, while they satisfy a relative measure of performance (e.g., fairness), they do not consider any absolute measures such as data rate. This motivates the study of a category of scheduling problems with minimum-performance guarantees [33, 45].

The problem to maximize the OFDM system performance while satisfying each user's minimum performance requirement can be stated as:

$$\begin{aligned} & \max_{\pi \in \Pi} U(\pi) & (3.11) \\ & \text{subject to } \liminf_{T \rightarrow \infty} U_i^T(\pi) \geq C_i, \quad i = 1, \dots, N, \end{aligned}$$

where  $\vec{C} = (C_1, C_2, \dots, C_N)$  is a feasible predetermined *minimum-performance requirement vector*. Feasible here means that there exists some policy that solves (3.11).

The QoS constraints here offer users a more direct service guarantee. For example, a user requires a minimum data rate guarantee, then the performance measure here can be data rate. Every user is guaranteed a minimum data rate, which may be more appealing from the user viewpoint. However, it can be quite difficult in practice to appeal because of the difficulty to determine if a requirement vector is feasible.

Suppose  $\vec{C} = (C_1, C_2, \dots, C_N)$  is feasible. We define the policy  $\pi^*$  for the problem in (3.11) as follows:

$$\pi^*(\omega^t) = \operatorname{argmax}_{\bar{A}^t} \left\{ \sum_{i=1}^N \sum_{k=1}^K \beta_i^* \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}} \right\}, \quad (3.12)$$

where the control parameters  $\beta_i^*$  are chosen such that:

1.  $\beta_i^* \geq 1, \forall i$ ;

2.  $\liminf_{T \rightarrow \infty} U_i^T(\pi) \geq C_i, \forall i$ ;
3. If  $\liminf_{T \rightarrow \infty} U_i^T(\pi) > C_i$ , then  $\beta_i^* = 1, \forall i$ .

Note that the parameter  $\vec{\beta}^* = (\beta_1^*, \dots, \beta_N^*)$  “scales” the performance values of users, and the scheduling policy always schedules the “relatively-best” subset of users to transmit. Here the subset of users selected by action  $\bar{A}^t$  is “relatively-best” if  $\sum_{i=1}^N \sum_{k=1}^K \beta_i^* \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}}$  is maximum over all actions. If  $\beta_i^* > 1$ , then user  $i$  is an “unfortunate” user, and it is granted only its minimum-performance requirement.

The policy  $\pi^*$  defined in (3.12), which represents our opportunistic scheduling policy, is optimal in the following sense.

**Theorem 3.3** *If  $\lim_{T \rightarrow \infty} U_i^T(\pi^*)$  exists for all  $i$  for the  $\pi^*$  defined in (3.12), then the policy  $\pi^*$  is an optimal solution to the problem defined in (3.11), i.e., it maximizes the average OFDM system performance under the minimum-performance guarantee constraints.*

*Proof:* Let  $\pi$  be a policy satisfying the minimum-performance guarantee constraints, and  $\beta_i^*$  satisfies conditions 1–3. Hence, we have

$$\begin{aligned}
U(\pi) &\leq U(\pi) + \sum_{i=1}^N (\beta_i^* - 1) \left( \liminf_{T \rightarrow \infty} U_i^T(\pi) - C_i \right) \\
&= \limsup_{T \rightarrow \infty} \sum_{i=1}^N U_i^T(\pi) + \sum_{i=1}^N (\beta_i^* - 1) \liminf_{T \rightarrow \infty} U_i^T(\pi) - \sum_{i=1}^N (\beta_i^* - 1) C_i \\
&\leq \limsup_{T \rightarrow \infty} \sum_{i=1}^N U_i^T(\pi) + \liminf_{T \rightarrow \infty} \sum_{i=1}^N (\beta_i^* - 1) U_i^T(\pi) - \sum_{i=1}^N (\beta_i^* - 1) C_i \\
&\leq \limsup_{T \rightarrow \infty} \sum_{i=1}^N \beta_i^* U_i^T(\pi) - \sum_{i=1}^N (\beta_i^* - 1) C_i.
\end{aligned}$$

By the definition of  $\pi^*$ , we get

$$\sum_{i=1}^N \sum_{k=1}^K \beta_i^* \omega_{i,k}^t \mathbf{1}_{\{A_k^t=i\}} \leq \sum_{i=1}^N \sum_{k=1}^K \beta_i^* \omega_{i,k}^t \mathbf{1}_{\{(A_k^t)^*=i\}}.$$

Thus,

$$\sum_{i=1}^N \beta_i^* U_i^T(\pi) \leq \sum_{i=1}^N \beta_i^* U_i^T(\pi^*).$$

Therefore,

$$\begin{aligned} U(\pi) &\leq \limsup_{T \rightarrow \infty} \sum_{i=1}^N \beta_i^* U_i^T(\pi^*) - \sum_{i=1}^N (\beta_i^* - 1) C_i \\ &\leq U(\pi^*) + \sum_{i=1}^N (\beta_i^* - 1) \left( \liminf_{T \rightarrow \infty} U_i^T(\pi^*) - C_i \right) \\ &= U(\pi^*), \end{aligned} \tag{3.13}$$

where the second part of (3.13) equals zero because of condition 3 on  $\beta_i^*$ . Similar to the proof of Theorem 1, the properties of  $\limsup$  and  $\liminf$  are applied here. ■

## 3.4 Implementation Issues

In this section, several implementation issues including parameter estimation and efficient policy search methods will be considered. An optimal algorithm and a low-complexity suboptimal algorithm are developed here for policy search.

### 3.4.1 Control Parameter Estimation

The opportunistic scheduling policies described in Section 3.3 involve some control parameters to be estimated online:  $\bar{v}^*$  in temporal fairness,  $\bar{\gamma}^*$  in utilitarian fairness, and  $\bar{\beta}^*$  in the minimum-performance guarantee policy. Those parameters are determined by the distribution of performance value matrix  $\{\omega^t\}$  and the predetermined constraints. In practice, the distribution is unknown, and hence we need to estimate the control parameters.

In [46], Liu et al. give a *stochastic approximation* technique to estimate such parameters. The basic idea is to find the root of a unknown continuous function  $f(x)$ . We approach the root by adapting the weighted observation error. For example, for user  $i$  in temporal fairness scheduling, the base station updates the parameter  $\bar{v}^{\#+1}$



using a stochastic approximation algorithm

$$v_i^{t+1} = v_i^t - \epsilon^t \left( \sum_{k=1}^K \mathbf{1}_{\{A_k^t=i\}} - r_i \right),$$

where, e.g., the step size  $\epsilon^t = 1/t$ . The initial estimate  $\vec{v}^1$  can be set to  $\vec{0}$  or some value based on the history information.

Using standard methods, it can be shown that  $\vec{v}^t$  converges to  $\vec{v}^*$  with probability one [88]. The computation burden above is  $O(N)$  per time slot, where  $N$  is the number of users, which suggests that the algorithm is easy to implement online. For our OFDM scheduling schemes, we have found that this stochastic approximation algorithm also works well. For the detailed procedure, we refer the reader to [46].

### 3.4.2 Optimal User Subset Search Methods

In our optimal OFDM policies (for example, in the temporal fairness policy), all the “relative performance values” ( $\omega_{i,k}^t + v_i^*$ ), denoted  $c_{ik}$  for convenience, comprise an  $N \times K$  matrix  $[c_{ik}]$ . Therefore, the operator  $\operatorname{argmax}_{\bar{A}^t}$  is to find an action  $\bar{A}^t$  that indicates which  $K$  elements in  $[c_{ik}]$  have the maximal sum over all  $K$  selected elements. This operator is obviously different from the  $\operatorname{argmax}_i$  in [46], which simply returns the index of the largest element from a vector.

It is straightforward to compute the  $\operatorname{argmax}$  if no hard physical limitations are considered. The operator can simply select the largest  $K$  elements. However, a common physical constraint is that in any time-slot, the scheduler cannot assign two users to the same subcarrier, or two subcarriers to the same user. Mathematically, at any time-slot  $t$ , for any two subcarriers  $j$  and  $k$ ,  $j \neq k \Leftrightarrow A_j^t \neq A_k^t$ . When this physical constraint is considered, the computation of the  $\operatorname{argmax}$  in the optimal policy is nontrivial. A brute-force approach is exhaustively searching over the  $\binom{N}{K}$  possible assignments, which obviously has very high computational complexity. Since this optimal user subset search operation should be performed online at each slot, we need to use more efficient algorithms.

It turns out that the problem of computing the argmax can be posed as an integer linear program (ILP) [89]:

$$\begin{aligned}
& \text{maximize } \sum_{i=1}^N \sum_{k=1}^K c_{ik} x_{ik} \\
& \text{subject to } \sum_{i=1}^N x_{ik} = 1, \quad k = 1, \dots, K, \\
& \quad \quad \quad \sum_{k=1}^K x_{ik} \leq 1, \quad i = 1, \dots, N, \\
& \quad \quad \quad x_{ik} \in \{0, 1\}, \quad c_{ik} \geq 0, \quad N \geq K,
\end{aligned}$$

where the decision variables  $x_{ik}$  indicate which elements to choose, and the weights  $c_{ik}$  are relative performance values defined above. This problem is called the *maximal weighted bipartite matching problem* in graph theory, or the *assignment problem* in combinatorial optimization [90].

It is interesting to see that the argmax operator in optimal multiuser OFDM scheduling problem can be interpreted as a graph problem  $(U, S, E, w)$ , where  $U$  represents the set of all users,  $S$  represents the set of all subcarriers, and  $E$  represents the set of all the feasible choices for specific users to select specific subcarriers. Each choice in  $E$  is weighted by a function  $w(E)$ . The problem is to find a matching  $M \in E$  for  $U$  and  $S$  that maximizes the sum of the weights over all edges in  $M$ .

The Hungarian algorithm is one of many algorithms that have been devised to solve the assignment problem in polynomial time ( $O(N^3)$  when  $N=K$ ) [91]. We modify the Hungarian algorithm to solve our general *unbalanced* ( $N \geq K$ ) problem here by introducing a number of *slack* variables to convert the ILP problem into standard form. Note that the standard form ILP with the slack variables is algebraically equivalent to the original problem [92]. It is proved in [91] that the Hungarian algorithm can always find the maximum assignment, i.e., it is an optimal solution to this problem.

The following is our modified Hungarian algorithm:

---

---

**Table 3.1:** Modified Hungarian algorithm

---

---

- Input:** An  $N \times K$  nonnegative matrix  $[c_{ik}]$ .
- Step 1:** Initialization:
- Append  $(N-K)$  all-zero columns to the matrix.
  - In each row, subtract the smallest entry from every entry in that row. In each column, subtract the smallest entry from every entry in that column.
- Step 2:** Cover all zeros with the minimum number of (horizontal and/or vertical) lines. If the minimum number  $=N$ , goto Step 4.
- Step 3:** Subtract the smallest uncovered entry from every uncovered entry; add it to every intersection of lines. Goto Step 2.
- Step 4:** Make the assignment at zeros. If any row or column has only one 0, make that assignment. Cross out the corresponding row and column, and move to the next assignment.
- 
- 

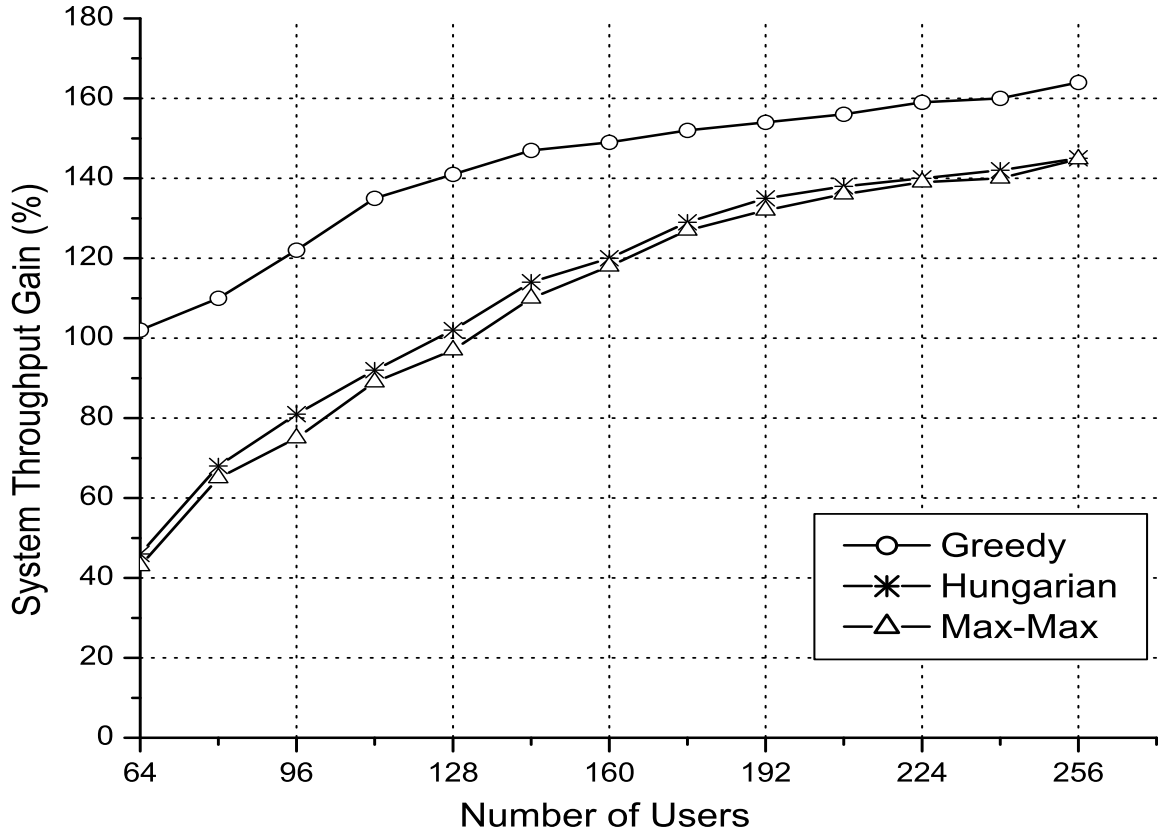
Ideally, the OFDM scheduler should repeat the above procedure at every scheduling slot. However, this still poses a heavy computational burden on the base station. Hence suboptimal algorithms with lower complexity are of interest for practical implementation.

We develop a suboptimal algorithm called “*max-max*” to perform the above argmax operation with much lower complexity. This algorithm is a variation of the “*min-min*” method for task mapping in heterogeneous computing [93]. The basic idea is this: first find the overall maximal element in the matrix  $[c_{ik}]$ , then assign the corresponding subcarrier to the corresponding user. Next, remove the newly-assigned user-subcarrier pair from the selection table. In other words, the corresponding row and column are removed from the matrix. Continue to repeat the above procedure on the reduced matrix until all subcarriers are assigned. In the simulations in the next section, the suboptimal scheme shows near-optimal performance with a lower complexity.

## 3.5 Simulation Results

In this section, we present numerical results to illustrate the performance of the various OFDM scheduling schemes developed in this chapter. For the purpose of comparison, we also simulate two special scheduling policies. *Round-robin* [94] is a non-opportunistic scheduling policy that schedules users over all subcarriers in a predetermined order. It is simple but lacks flexibility. The round-robin policy can serve as a performance benchmark to measure how much gain results from using our opportunistic scheduling policies. The other policy for comparison is a greedy scheduling scheme that always selects the user with the maximum performance to transmit for each subcarrier at each time-slot. The greedy policy will in general violate the QoS/fairness constraints, but provide an upper-bound on the system performance. It is used here to expose the tradeoff between the QoS constraints for individual users and the overall system throughput. The more relaxed the fairness constraints, the higher the overall achievable throughput, therefore the closer to we will get to the performance of the greedy scheme.

In our simulation, we consider the downlink of a heavy-traffic single-cell OFDM system with fixed 64 subcarriers. There is one base station serving all the users in the cell. Each user suffers from multipath Rayleigh fading with the bad-urban (BU) scenario of the COST 259 channel model [18, 95], and we assume a path-loss exponent of four. Every user is assumed to be stationary or slowly moving so that the maximum Doppler shift is 20 Hz. The performance values used by different users usually is a nondecreasing function of their SINR, can be in various forms, such as linear functions, step functions, or S-shape functions. For simplicity, here we take all the performance values as linear functions of users' SINR (in dB). We assume that the physical limitation on scheduling discussed in Section 3.2 applies: at each time-slot, no two users can be scheduled on the same subcarrier and each user is scheduled exactly one subcarrier.

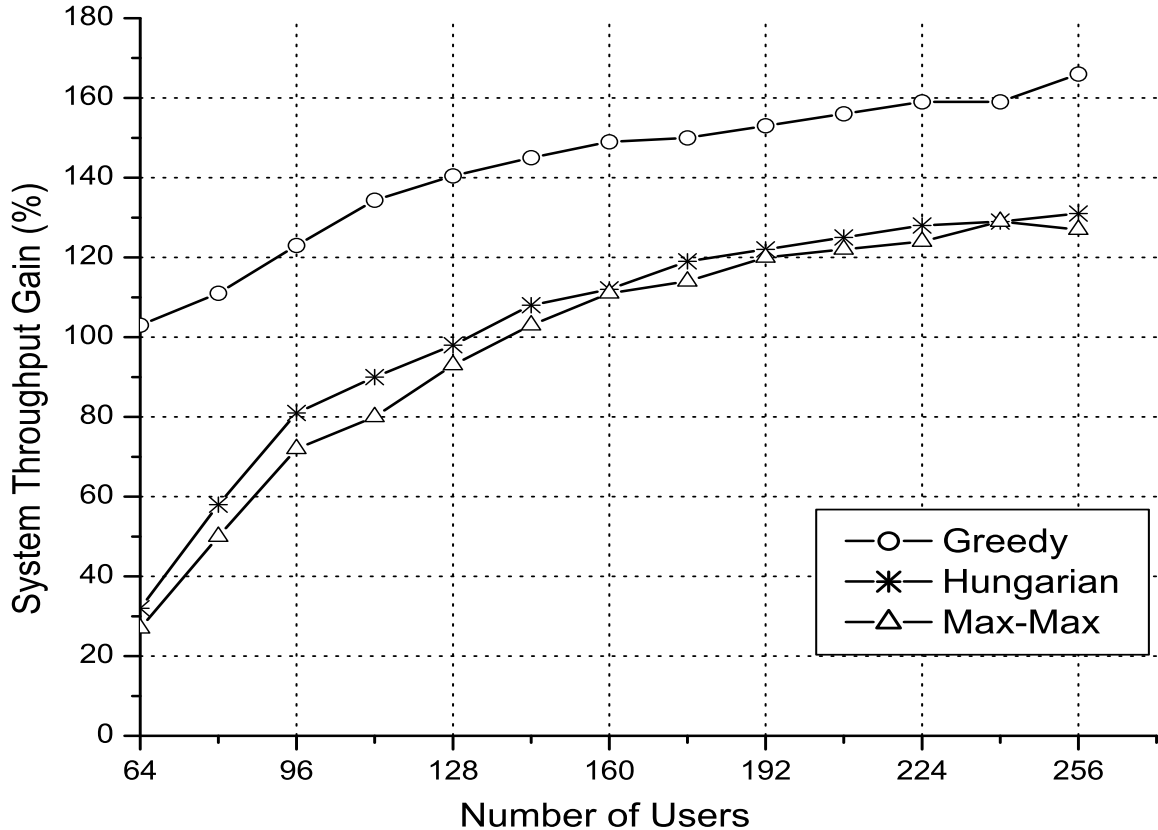


**Figure 3.2:** System throughput gain in the temporal fairness scheduling.

### 3.5.1 Performance Gain

First we assume the locations of all users are distributed uniformly in the cell and examine the impact of the number of users on the average system throughput. We use the round-robin policy as the baseline, and define the system throughput gain as  $(U_S - U_R)/U_R$ , where  $U_S$  and  $U_R$  denote the average system throughput of a given scheduling policy and the round-robin policy, respectively.

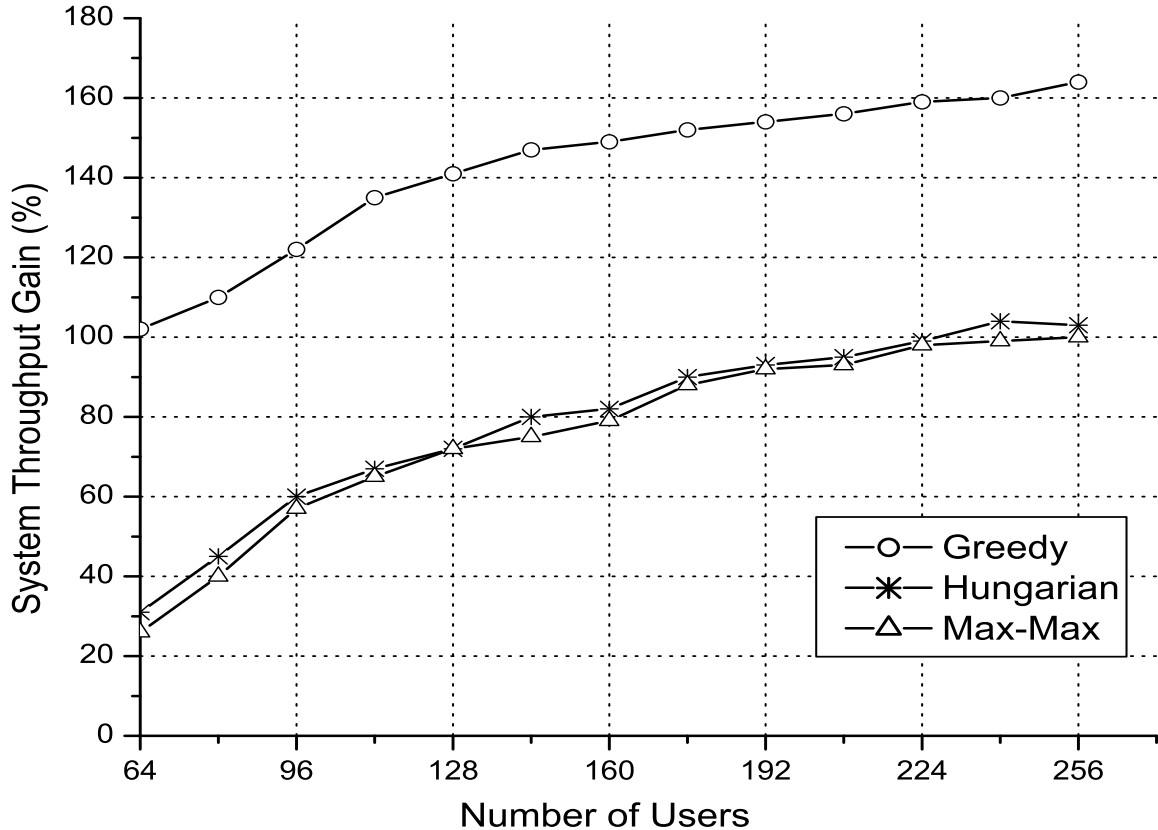
Figure 3.2 shows the system throughput gain relative to round-robin from the different policies in the temporal fairness scheduling simulations. For the purpose of simulation, we assume the time-fraction assignment is done using *fair sharing*, i.e., the total resources are evenly divided among the users. Therefore, if there are  $N$  users in the cell, we set  $r_i = 1/N$  for all users. From Figure 3.2, it is evident that the system throughput gain increases with the number of users. This is reflective of the



**Figure 3.3:** System throughput gain in the utilitarian fairness scheduling.

*multiuser diversity* gain. For 64 users, our optimal policy (Hungarian) achieves about 46% overall throughput gain, while the greedy policy has an improvement of 101%. This is not surprising since the greedy policy achieves the highest overall performance at the cost of unfairness among the users. The suboptimal policy (max-max) shows surprisingly near-optimal performance. Its performance gap with the optimal policy is less than 1–2%, and even smaller when we increase the number of users.

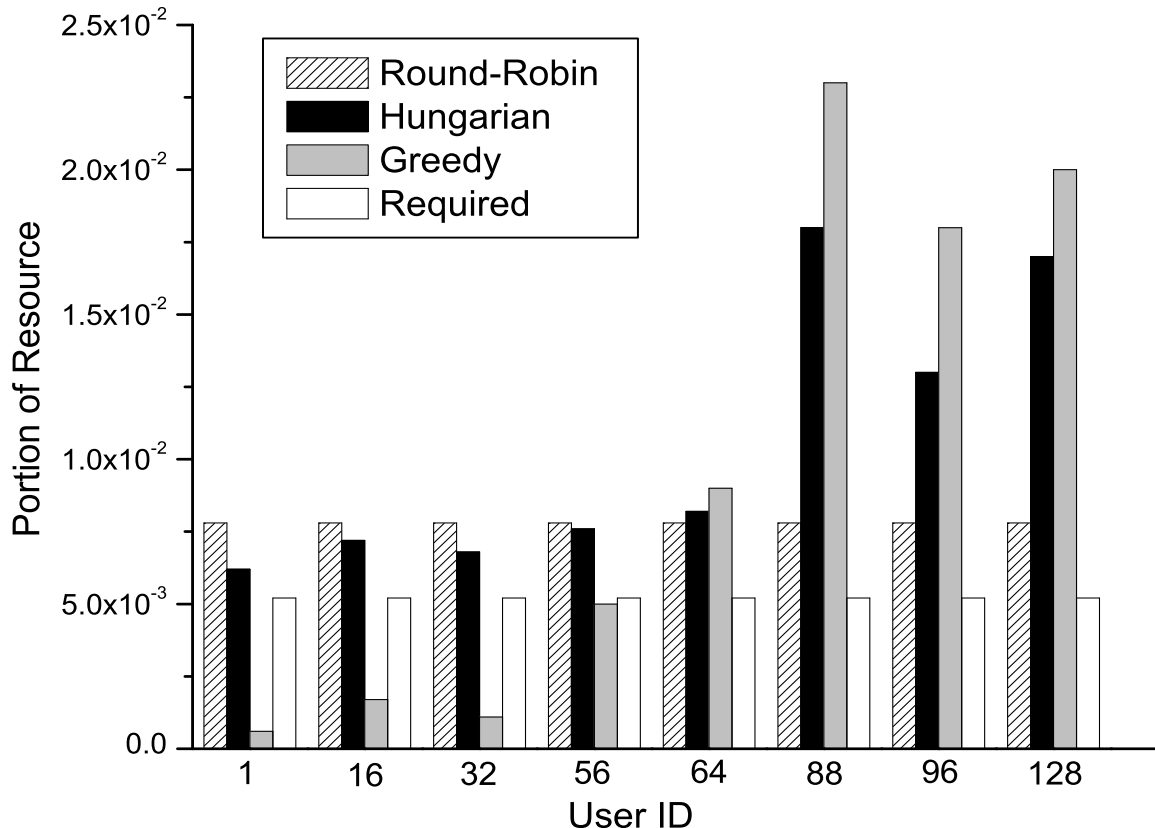
Figure 3.3 shows the system throughput gain relative to round-robin from the different policies in the utilitarian fairness scheduling simulations. We also assume *fair sharing* in the throughput-fraction assignment. This means we set  $a_i = 1/N$  for all users in a  $N$ -user system. As expected, the increasing trend similar to Figure 3.3 can be also seen here. For 64 users, our optimal policy (Hungarian) achieves about 32% overall throughput gain, while the greedy policy has an improvement of 102%.



**Figure 3.4:** System throughput gain in the minimum-performance guarantee scheduling.

The suboptimal policy (max-max) also improves the system performance by 27%.

Next, we investigate the performance of the opportunistic scheduling schemes with minimum-performance guarantees. First we run the simulation for 1,000,000 time-slots using the round-robin policy, where the resource (time) is equally distributed among all users. Then we compute an average performance value and use it as the minimum-performance requirement for each user. It is easy to see that this minimum-performance requirement vector is feasible. Figure 3.4 shows the system throughput gain relative to round-robin from the different policies in the minimum-performance guarantee scheduling simulations. For 64 users, our optimal policy (Hungarian) achieves about 31% overall throughput gain, while the greedy policy (which violates the minimum-performance requirements) has an improvement of about 100%. The suboptimal policy (max-max) also performs well with 24% overall gain.

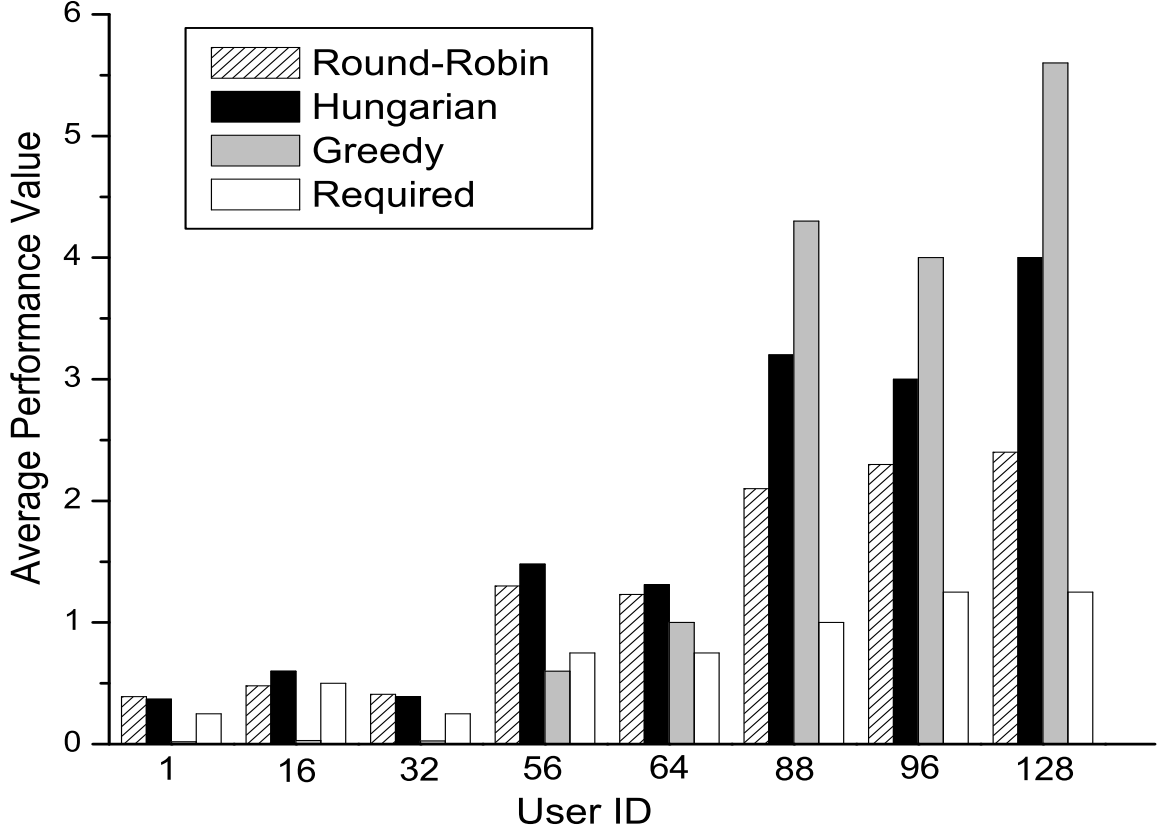


**Figure 3.5:** Portion of resource shared by users in the temporal fairness scheduling.

### 3.5.2 Fairness

Using the temporal fairness scheduling scenario as an example, we study the fairness among the users by applying the different policies. We use the same single-cell system with 64 subcarriers; and there are 128 users in the system. The users are divided into three “distance” groups. Users 1–48 belong to the “far” group, users 49–80 belong to the “middle” group, and users 81–128 belong to the “near” group. Obviously a user in the “near” group has a much higher probability to get a strong SINR than a user in the “far” group. We set all users to have the same minimum time-fraction requirement. Specifically, each user has a resource (time) requirement  $r_i = 2/(3N)$  for a  $N$ -user system, where  $\sum_i r_i = 2/3 < 1$ . Therefore, the system has the freedom to assign the remaining  $1/3$  portion of the resource to some “better” users (beyond their minimum requirements) to further improve the system performance.

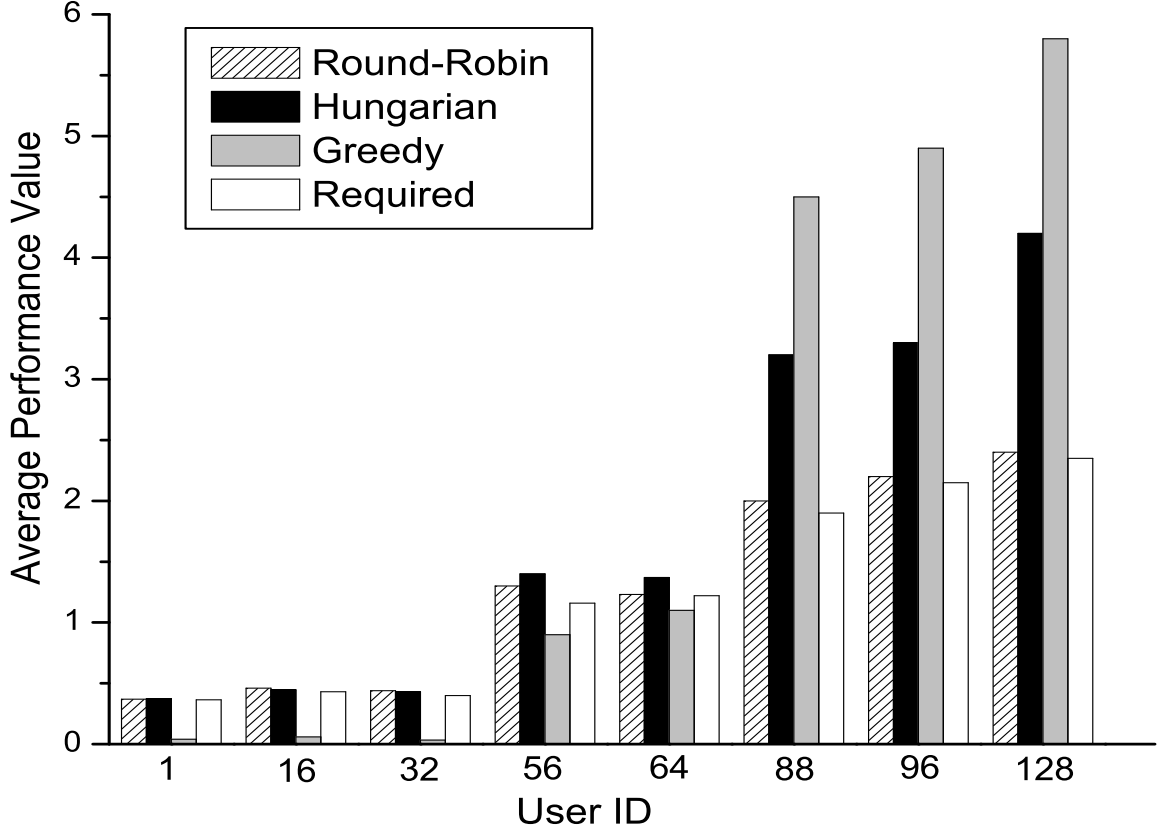




**Figure 3.6:** User average performance in the utilitarian fairness scheduling.

Figure 3.5 indicates the amount of resource consumed by selected users in the temporal fairness scheduling simulations. The first bar represents that of round-robin, where the resource is equally shared by all users. The second bar represents our optimal policy (Hungarian). The third bar is the greedy policy. The rightmost bar shows the minimum requirements of user. The second bar is higher than the fourth bar for all the users, which indicates that our temporal fairness optimal scheduling policy meets the minimum time-fraction requirements for all users. In the greedy policy, users 1, 16, and 32 get very little resource (far below the minimum requirement line) while users 88, 96, and 128 have very large shares. As expected, the greedy algorithm is heavily *biased* though it achieves the highest overall performance.

In the following, we simply check the fairness among the users with utilitarian fairness and minimum-performance guarantee scheduling. We use the same cellular



**Figure 3.7:** User average performance in the minimum-performance guarantee scheduling.

system and user group settings as temporal fairness.

In Figure 3.6, we show the average performance values of selected users in the utilitarian fairness scheduling simulations. The preset performance requirements of the selected users 1, 16, 32, 56, 64, 88, 96, and 128 are  $[0.001, 0.002, 0.001, 0.003, 0.003, 0.004, 0.005, 0.005]$ . The values represent the minimum fraction of overall average performance for individual users.

In Figure 3.7, we show the average performance values of selected users in the minimum-performance guarantee scheduling simulations. Similar to the previous section, we first run a round-robin simulation, then use the obtained average performance as minimum-performance requirement for each user. From the figure, we see that our optimal scheduling policy (Hungarian) meets all the requirements and outperforms round-robin policy everywhere.

In summary, the simulation results show that using our OFDM opportunistic scheduling policies, the system can achieve significant performance gains over the non-opportunistic round-robin policy, while satisfying the various QoS/fairness requirements. Also the low-complexity suboptimal policy shows near-optimal performance in every scenario.

## 3.6 Conclusions

Opportunistic transmission scheduling is a promising technology to improve spectrum efficiency by exploiting time-varying channel conditions. We investigated the application of opportunistic scheduling in multiuser OFDM systems, which dynamically allocates resource in both temporal and spectral domains. Optimal scheduling policies were presented and proved to be optimal under the temporal fairness, utilitarian fairness, and minimum-performance QoS constraints. We developed optimal and suboptimal algorithms to implement these optimal policies efficiently. The simulation showed that the schemes achieve improvements of about 30%–140% in network efficiency compared with a scheduling scheme that does not take into account channel conditions.

# CHAPTER 4

## OPPORTUNISTIC SCHEDULING WITH MIXED CONSTRAINTS

Opportunistic scheduling exploits the time-varying, location-dependent channel conditions to achieve multiuser diversity. In previous work, different type QoS constraints are only treated individually as different scheduling problems. In this chapter, we consider the problem of downlink transmission scheduling with more general constraints than in the last chapter.

First, instead of only considering the lower bounds, we consider the scheduling problems with both lower and upper bounds constraints here. We derive the corresponding opportunistic scheduling policies for three long-term QoS/fairness constraints—temporal fairness, utilitarian fairness, and minimum-performance guarantees. Then we deal with scheduling problems with multiple type mixed QoS/fairness constraints. As examples, we derive our opportunistic scheduling policies with a mixture of three long-term QoS/fairness constraints and prove their optimality. At last, we propose a generalized opportunistic scheduling framework to include all these scheduling schemes and more. We show that the structure of the optimal opportunistic scheduling policy can be carried over to the problem with general constraints. The proposed framework can be viewed as a theoretical generalization of the work by Liu et al. [44–46].

Although we focus on the downlink of a wireless network, our scheduling schemes can also be applied to uplink. However, the uplink may experience synchronization difficulties due to different distances between users and the base station when the duration of a time-slot is short. In general, downlink transmission is more important for data traffic due to the highly asymmetric nature of the data service.

This chapter is organized as follows. In Section 4.1, we give the motivation for this work. The system model is briefly explained in Section 4.2. In Section 4.3, we derive the corresponding opportunistic scheduling policies with both maximum and minimum constraints, and prove their optimality. In Section 4.4, we discuss the opportunistic scheduling problems with multiple mixed type QoS/fairness constraints. In Section 4.5, we proposed a generalized opportunistic scheduling framework. Finally, concluding remarks are given in Section 4.6.

## 4.1 Motivation

The unique *time-varying* characteristics of wireless networks requires that the scheduler should *opportunistically* seek to exploit channel conditions to achieve higher network performance. On the other hand, the potential to transmit at higher data rates opportunistically also introduces an important tradeoff between wireless resource efficiency and level of satisfaction among different users (fairness).

To address this problem, as we introduced in Chapter 2, Liu et al. described a framework for opportunistic scheduling to exploit the multiuser diversity while at the same time satisfying different QoS constraints [44–46]. The framework enables us to investigate different categories of scheduling problems involving two fairness requirements (*temporal fairness* and *utilitarian fairness*) and a minimum-performance requirement. The optimal scheduling solutions for these scheduling problems turn out to be index policies, and a stochastic-approximation-based algorithm can be used to efficiently estimate the key parameters of the scheduling schemes online.

In the work of [44–46], the different type QoS constraints are only treated individually as different scheduling problems. It is interesting to consider scheduling problems with multiple mixed QoS/fairness constraints. For example, a user might ask for both minimum temporal fraction and minimum performance guarantees. Future broadband wireless networks, which will fully support multimedia communications such as high-speed data access and video conferencing, demand more flexible and efficient scheduling schemes. Therefore, such scheduling problems with multiple different QoS constraints for a single user become more important and are definitely of practical interest.

The work of [44–46] also only consider the users' *minimum* (lower bound) constraints—namely, *minimum temporal fairness*, *minimum utilitarian fairness*, and *minimum-performance guarantees* constraints. But a user might be constrained by both *maximum* and *minimum* requirements of wireless resource. It is of practical interest to investigate such scheduling problems with both minimum (lower bound) and maximum (upper bound) constraints.

Providing a minimum guarantee on resource or performance is natural and arguably the simplest QoS guarantee. Multiple reasons why we feel it is important to provide minimum rate constraints are:

- Some bandwidth-sensitive applications such as VoIP and streaming video need a minimum rate in order to perform well.
- Even for static TCP-based applications such as web browsing if the bandwidth is too small then we typically get a large queue buildup which can lead to TCP timeouts and poor performance [96].
- Providing a minimum rate guarantee can help to smooth out the effects of a variable wireless channel.

- By having different minimum constraints for different users, we can ensure that high-paying premium customers receive better service than regular customers.

It seems undesirable to place maximum constraints (upper bounds) on individual users, because these bounds may limit the achievable system performance. However, the following are several reasons for imposing such constraints.

- If a user has only paid for a cheap data service, the operator might wish to cap their data rate in order to give them an incentive to upgrade to a more expensive premium service.
- Maximal constraints are useful for implementing multi-tiered services.
- Maximal constraints can decrease the subscribers' QoS sensibility to the number of subscribers in the network. For example, when only one subscriber is active, then all the system resource is available to that subscriber. This data rate will decrease as more and more subscribers become active. Thus, there will be considerable variance in QoS. By imposing maximal constraints on data rate can help to decrease this variance.

Note that if the system operator does not wish to have maximum constraints, this is easily accomplished by setting the upper bound to infinity (or some suitably large value).

## 4.2 System Model

We consider the downlink of a time-slotted system where time is the resource to be shared among all users. We focus on the scheduling problem for a single channel. Such a system model includes TDMA systems as well as time-slotted CDMA systems (e.g., HDR).

Recall that in Chapter 3, we used a stochastic model to capture the *time-varying* performance of each user. For simplicity, here we assume that  $\{U_i^t\}$ , the stochastic process associated with user  $i$ , is *stationary* and *ergodic*. Specifically, we use the notation  $\vec{U} = (U_1, \dots, U_N)$ , where  $U_i$  is a random variable representing the performance value of user  $i$  at a generic time-slot, and  $N$  is the number of users. At a generic time-slot, if a policy  $\pi$  schedules user  $i = \pi(\vec{U}) \in \{1, \dots, N\}$  to transmit, then the system receives a “reward” of  $U_i$ . Note that  $E(U(\pi))$  is the average system performance value associated with policy  $\pi$ , and it is the sum of all users’ average performance values, i.e.,  $E(U(\pi)) = \sum_{i=1}^N E\left(U_i \mathbf{1}_{\{\pi(\vec{U})=i\}}\right)$ . The objective is to find a policy  $\pi$  that maximizes the average system performance value  $E(U(\pi))$ , while satisfying specific QoS/fairness constraints.

## 4.3 Scheduling with Maximum and Minimum Bounds

In this section, we consider scheduling problems with both maximum and minimum constraints. We will still focus on the three fairness criteria—temporal fairness, utilitarian fairness, and minimum-performance guarantees.

### 4.3.1 Performance Guarantee Scheduling

In this subsection, we consider a system where each user is subject to certain maximum and minimum performance (data-rate) constraints. More precisely, the problem to maximize the system performance while satisfying each user’s maximum and minimum performance requirements can be stated as:

$$\begin{aligned} & \max_{\pi \in \Pi} E(U(\pi)) & (4.1) \\ \text{subject to } & E\left(U_i \mathbf{1}_{\{\pi(\vec{U})=i\}}\right) \geq C_i, & i = 1, 2, \dots, N, \\ & E\left(U_i \mathbf{1}_{\{\pi(\vec{U})=i\}}\right) \leq D_i, & i = 1, 2, \dots, N, \end{aligned}$$



where  $\vec{C} = (C_1, C_2, \dots, C_N)$  is a feasible predetermined *minimum-performance requirement vector* and  $\vec{D} = (D_1, D_2, \dots, D_N)$  is a feasible predetermined *maximum-performance requirement vector*; and  $\forall i, D_i \geq C_i \geq 0$ .

Define the policy  $\pi^*$  as follows:

$$\pi^*(\vec{U}) = \operatorname{argmax}_i ((\theta_i - \mu_i)U_i), \quad (4.2)$$

where the control parameters  $\theta_i$  and  $\mu_i$  are chosen such that:

1.  $\theta_i \geq 1, \mu_i \geq 0, \forall i$ ;
2.  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) \geq C_i, \forall i$ ;
3. If  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) > C_i$ , then  $\theta_i = 1, \forall i$ ; and
4.  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) \leq D_i, \forall i$ ;
5. If  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) < D_i$ , then  $\mu_i = 0, \forall i$ .

The policy  $\pi^*$  defined in (4.2), which represents our opportunistic scheduling policy, is optimal in the following sense.

**Theorem 4.1** *The policy  $\pi^*$  defined in (4.2) is an optimal solution to the problem defined in (4.1), i.e., it maximizes the system performance while satisfying the maximum and minimum performance requirements for individual users.*

*Proof:* Let  $\pi$  be a policy satisfying the maximum and minimum performance

requirements, and  $\theta_i$  and  $\mu_i$  satisfy conditions 1–5. Hence, we have

$$\begin{aligned}
E(U(\pi)) &\leq E(U(\pi)) + \sum_{i=1}^N (\theta_i - 1) \left( E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - C_i \right) \\
&\quad + \sum_{i=1}^N \mu_i \left( D_i - E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \right) \\
&= \sum_{i=1}^N \theta_i E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - \sum_{i=1}^N \mu_i E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \\
&\quad - \sum_{i=1}^N (\theta_i - 1) C_i + \sum_{i=1}^N \mu_i D_i \\
&= E \left( \sum_{i=1}^N (\theta_i - \mu_i) U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - \sum_{i=1}^N (\theta_i - 1) C_i + \sum_{i=1}^N \mu_i D_i
\end{aligned}$$

By the definition of  $\pi^*$ , we have

$$\sum_{i=1}^N (\theta_i - \mu_i) U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \leq \sum_{i=1}^N (\theta_i - \mu_i) U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}$$

Thus,

$$E \left( \sum_{i=1}^N (\theta_i - \mu_i) U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \leq E \left( \sum_{i=1}^N (\theta_i - \mu_i) U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right)$$

Hence,

$$\begin{aligned}
E(U(\pi)) &\leq E \left( \sum_{i=1}^N (\theta_i - \mu_i) U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - \sum_{i=1}^N (\theta_i - 1) C_i + \sum_{i=1}^N \mu_i D_i \\
&= E(U(\pi^*)) + \sum_{i=1}^N (\theta_i - 1) \left( E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - C_i \right) \\
&\quad + \sum_{i=1}^N \mu_i \left( D_i - E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) \right) \\
&= E(U(\pi^*)).
\end{aligned} \tag{4.3}$$

where the second and third parts of (4.3) equal zero because of conditions 3 and 5 on  $\theta_i$  and  $\mu_i$ . ■

Note that the parameters  $\theta_i$  and  $\mu_i$  “scale” the performance values of users, and the scheduling policy always schedules the “relatively best” subset of users to transmit.

If  $\theta_i > 1$ , then user  $i$  is an “unfortunate” user. The setting of parameter  $\theta_i$  will scale the performance causing more frequent allocation of slots to user  $i$  to improve its data rate. If user  $i$  satisfies the lower performance bounds, then the corresponding  $\theta_i$  will be 1, i.e., no scaling is needed in this case. Similarly,  $\mu_i$  compensates any violations of the upper bound. For users having scheduled data rate less than upper bound, the corresponding  $\mu_i$  values will be equal to 0. On the other hand, if the wireless channel of user  $i$  is such that the scheduled data rate will be higher than  $D_i$ ,  $\mu_i$  will be greater than 0. This value will scale down the corresponding performance value of the user; therefore, the user will be scheduled less often. The value of  $\theta_i$  and  $\mu_i$  can be efficiently estimated online via a stochastic-approximation-based algorithm provided in [46].

### 4.3.2 Temporal Fairness Scheduling

A natural fairness criterion is to give each action a certain portion of time because time is the basic resource shared among users. The scheduling problem with maximum and minimum temporal fairness bounds can be expressed as:

$$\begin{aligned} & \max_{\pi \in \Pi} E(U(\pi)) && (4.4) \\ & \text{subject to } P\{\pi(\vec{U}) = i\} \geq r_i, && i = 1, 2, \dots, N, \\ & P\{\pi(\vec{U}) = i\} \leq s_i, && i = 1, 2, \dots, N, \end{aligned}$$

where  $r_i$  denotes the minimum time-fraction that should be assigned to user  $i$ , with  $r_i \geq 0$  and  $\sum_{i=1}^N r_i \leq 1$ ;  $s_i$  denotes the maximum time-fraction that could be assigned to user  $i$ , with  $s_i \geq 0$  and  $\sum_{i=1}^N s_i \leq 1$ ; and  $\forall i, s_i > r_i$ .

Define the policy  $\pi^*$  as follows:

$$\pi^*(\vec{U}) = \underset{i}{\operatorname{argmax}} (U_i + \alpha_i - \beta_i), \quad (4.5)$$

where the control parameters  $\alpha_i$  and  $\beta_i$  are chosen such that:

1.  $\alpha_i \geq 0, \beta_i \geq 0, \forall i$ ;
2.  $P\{\pi^*(\vec{U}) = i\} \geq r_i, \forall i$ ;
3. If  $P\{\pi^*(\vec{U}) = i\} > r_i$ , then  $\alpha_i = 0, \forall i$ ; and
4.  $P\{\pi^*(\vec{U}) = i\} \leq s_i, \forall i$ ;
5. If  $P\{\pi^*(\vec{U}) = i\} < s_i$ , then  $\beta_i = 0, \forall i$ .

**Theorem 4.2** *The policy  $\pi^*$  defined in (4.5) is an optimal solution to the problem defined in (4.4), i.e., it maximizes the system performance, while satisfying maximum and minimum temporal fairness constraints for individual users.*

*Proof:* Let  $\pi$  be a policy satisfying the temporal fairness constraints, and  $\alpha_i$  and  $\beta_i$  satisfy conditions 1–5. Hence, we have

$$\begin{aligned}
E(U(\pi)) &\leq E(U(\pi)) + \sum_{i=1}^N \alpha_i \left( P\{\pi(\vec{U}) = i\} - r_i \right) + \sum_{i=1}^N \beta_i \left( s_i - P\{\pi(\vec{U}) = i\} \right) \\
&= E \left( \sum_{i=1}^N (U_i + \alpha_i - \beta_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - \sum_{i=1}^N \alpha_i r_i + \sum_{i=1}^N \beta_i s_i
\end{aligned}$$

By the definition of  $\pi^*$ , we have

$$\sum_{i=1}^N (U_i + \alpha_i - \beta_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \leq \sum_{i=1}^N (U_i + \alpha_i - \beta_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}}$$

Thus,

$$E \left( \sum_{i=1}^N (U_i + \alpha_i - \beta_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \leq E \left( \sum_{i=1}^N (U_i + \alpha_i - \beta_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right)$$

Hence,

$$\begin{aligned}
E(U(\pi)) &\leq E\left(\sum_{i=1}^N (U_i + \alpha_i - \beta_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right) - \sum_{i=1}^N \alpha_i r_i + \sum_{i=1}^N \beta_i s_i \\
&= E(U(\pi^*)) + \sum_{i=1}^N \alpha_i \left(P\{\pi^*(\vec{U}) = i\} - r_i\right) \\
&\quad + \sum_{i=1}^N \beta_i \left(s_i - P\{\pi^*(\vec{U}) = i\}\right) \\
&= E(U(\pi^*)).
\end{aligned} \tag{4.6}$$

where the second and third parts of (4.6) equal zero because of conditions 3 and 5 on  $\alpha_i$  and  $\beta_i$ . ■

### 4.3.3 Utilitarian Fairness Scheduling

In the last section, we studied the opportunistic scheduling problem with maximum and minimum temporal fairness bounds. In this section, we will describe an alternative scheduling problem that would ensure that all users get at least a certain fraction of the overall system performance. The opportunistic scheduling problem with maximum and minimum utilitarian fairness bounds can be expressed as:

$$\begin{aligned}
&\max_{\pi \in \Pi} E(U(\pi)) && (4.7) \\
&\text{subject to } E\left(U_i \mathbf{1}_{\{\pi(\vec{U})=i\}}\right) \geq a_i E(U(\pi)), && i = 1, 2, \dots, N, \\
&E\left(U_i \mathbf{1}_{\{\pi(\vec{U})=i\}}\right) \leq b_i E(U(\pi)), && i = 1, 2, \dots, N,
\end{aligned}$$

where  $a_i$  denotes the minimum fraction of the overall performance required by user  $i$ , with  $a_i \geq 0$  and  $\sum_{i=1}^N a_i \leq 1$ ;  $b_i$  denotes the maximum fraction of the overall performance available for user  $i$ , with  $b_i \geq 0$  and  $\sum_{i=1}^N b_i \leq 1$ ; and  $\forall i, b_i > a_i$ .

Define the policy  $\pi^*$  as follows:

$$\pi^*(\vec{U}) = \underset{i}{\operatorname{argmax}} ((\kappa + \gamma_i - \eta_i)U_i), \tag{4.8}$$

where  $\kappa = 1 - \sum_{i=1}^N a_i \gamma_i + \sum_{i=1}^N b_i \eta_i$ , and the control parameters  $\gamma_i$  and  $\eta_i$  are chosen such that:

1.  $\gamma_i \geq 0, \eta_i \geq 0, \forall i$ ;
2.  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) \geq a_i E(U(\pi)), \forall i$ ;
3. If  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) > a_i E(U(\pi))$ , then  $\gamma_i = 0, \forall i$ ; and
4.  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) \leq b_i E(U(\pi)), \forall i$ ;
5. If  $E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) < b_i E(U(\pi))$ , then  $\eta_i = 0, \forall i$ .

**Theorem 4.3** *The policy  $\pi^*$  defined in (4.8) is an optimal solution to the problem defined in (4.7), i.e., it maximizes the system performance, while satisfying maximum and minimum utilitarian fairness constraints for individual users.*

*Proof:* Let  $\pi$  be a policy satisfying the utilitarian fairness constraints, and  $\gamma_i$  and  $\eta_i$  satisfy conditions 1–5. Hence, we have

$$\begin{aligned}
E(U(\pi)) &\leq E(U(\pi)) + \sum_{i=1}^N \gamma_i \left( E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - a_i E(U(\pi)) \right) \\
&\quad + \sum_{i=1}^N \eta_i \left( b_i E(U(\pi)) - E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \right) \\
&= \left( 1 - \sum_{i=1}^N a_i \gamma_i + \sum_{i=1}^N b_i \eta_i \right) E(U(\pi)) + \sum_{i=1}^N \gamma_i E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \\
&\quad - \sum_{i=1}^N \eta_i E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \\
&= E \left( \sum_{i=1}^N (\kappa + \gamma_i - \eta_i) U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right)
\end{aligned}$$

By the definition of  $\pi^*$ , we have

$$\sum_{i=1}^N (\kappa + \gamma_i - \eta_i) U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \leq \sum_{i=1}^N (\kappa + \gamma_i - \eta_i) U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}$$

Thus,

$$E \left( \sum_{i=1}^N (\kappa + \gamma_i - \eta_i) U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \leq E \left( \sum_{i=1}^N (\kappa + \gamma_i - \eta_i) U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right)$$

Hence,

$$\begin{aligned}
E(U(\pi)) &\leq E\left(\sum_{i=1}^N (\kappa + \gamma_i - \eta_i) U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right) \\
&= E(U(\pi^*)) + \sum_{i=1}^N \gamma_i \left(E\left(U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right) - a_i E(U(\pi^*))\right) \\
&\quad + \sum_{i=1}^N \eta_i \left(b_i E(U(\pi^*)) - E\left(U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right)\right) \\
&= E(U(\pi^*)).
\end{aligned} \tag{4.9}$$

where the second and third parts of (4.9) equal zero because of conditions 3 and 5 on  $\gamma_i$  and  $\eta_i$ . ■

## 4.4 Scheduling with Mixed Constraints

In the last section, we studied scheduling problems with both maximum and minimum constraints. Therein, both maximum and minimum constraints are of the same type, i.e., it is either the data-rate constraint, or else. Furthermore, scheduling problems with multiple mixed QoS/fairness constraints is definitely of practical interests. For example, a user might ask for both minimum temporal fraction and minimum performance guarantees. In this section, we discuss the opportunistic scheduling problems with multiple mixed type QoS/fairness constraints. We still focus on the three fairness criteria—temporal fairness, utilitarian fairness, and minimum-performance guarantees.

### 4.4.1 Minimum Temporal Fairness and Performance Guarantee

Here we consider a system where each user requires both minimum temporal fraction and minimum performance guarantees. More precisely, we are interested to solve the

following constrained optimization problem:

$$\begin{aligned}
& \max_{\pi \in \Pi} E(U(\pi)) & (4.10) \\
& \text{subject to } P\{\pi(\vec{U}) = i\} \geq r_i, \quad i = 1, 2, \dots, N, \\
& E\left(U_i \mathbf{1}_{\{\pi(\vec{U})=i\}}\right) \geq C_i, \quad i = 1, 2, \dots, N,
\end{aligned}$$

where  $r_i$  denotes the minimum time-fraction that should be assigned to user  $i$ , with  $r_i \geq 0$  and  $\sum_{i=1}^N r_i \leq 1$ ; and  $\vec{C} = (C_1, C_2, \dots, C_N)$  is a feasible predetermined *minimum-performance requirement vector*.

Define the policy  $\pi^*$  as follows:

$$\pi^*(\vec{U}) = \underset{i}{\operatorname{argmax}} (\theta_i U_i + \alpha_i), \quad (4.11)$$

where the control parameters  $\alpha_i$  and  $\theta_i$  are chosen such that:

1.  $\alpha_i \geq 0, \theta_i \geq 1, \forall i$ ;
2.  $P\{\pi^*(\vec{U}) = i\} \geq r_i, \forall i$ ;
3. If  $P\{\pi^*(\vec{U}) = i\} > r_i$ , then  $\alpha_i = 0, \forall i$ ; and
4.  $E\left(U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right) \geq C_i, \forall i$ ;
5. If  $E\left(U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right) > C_i$ , then  $\theta_i = 1, \forall i$ .

**Theorem 4.4** *The policy  $\pi^*$  defined in (4.11) is an optimal solution to the problem defined in (4.10), i.e., it maximizes the system performance while satisfying the minimum time fraction and performance requirements for individual users.*

*Proof:* Let  $\pi$  be a policy satisfying the minimum time fraction and performance requirements for individual users, and  $\alpha_i$  and  $\theta_i$  satisfy conditions 1–5. Hence, we have



$$\begin{aligned}
E(U(\pi)) &\leq E(U(\pi)) + \sum_{i=1}^N \alpha_i \left( P\{\pi(\vec{U}) = i\} - r_i \right) \\
&\quad + \sum_{i=1}^N (\theta_i - 1) \left( E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - C_i \right) \\
&= \sum_{i=1}^N \theta_i E \left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) + \sum_{i=1}^N \alpha_i P\{\pi(\vec{U}) = i\} \\
&\quad - \sum_{i=1}^N \alpha_i r_i - \sum_{i=1}^N (\theta_i - 1) C_i \\
&= E \left( \sum_{i=1}^N (\theta_i U_i + \alpha_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - \sum_{i=1}^N \alpha_i r_i - \sum_{i=1}^N (\theta_i - 1) C_i
\end{aligned}$$

By the definition of  $\pi^*$ , we have

$$\sum_{i=1}^N (\theta_i U_i + \alpha_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \leq \sum_{i=1}^N (\theta_i U_i + \alpha_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}}$$

Thus,

$$E \left( \sum_{i=1}^N (\theta_i U_i + \alpha_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \leq E \left( \sum_{i=1}^N (\theta_i U_i + \alpha_i) U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right)$$

Hence,

$$\begin{aligned}
E(U(\pi)) &\leq E \left( \sum_{i=1}^N (\theta_i U_i + \alpha_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - \sum_{i=1}^N \alpha_i r_i - \sum_{i=1}^N (\theta_i - 1) C_i \\
&= E(U(\pi^*)) + \sum_{i=1}^N \alpha_i \left( P\{\pi^*(\vec{U}) = i\} - r_i \right) \\
&\quad + \sum_{i=1}^N (\theta_i - 1) \left( E \left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - C_i \right) \\
&= E(U(\pi^*)).
\end{aligned} \tag{4.12}$$

where the second and third parts of (4.12) equal zero because of conditions 3 and 5 on  $\alpha_i$  and  $\theta_i$ . ■

#### 4.4.2 Minimum Temporal and Utilitarian Fairness

Here we consider a system where each user requires both a minimum time fraction and a minimum fraction of the system performance. More precisely, the constrained

optimization problem can be expressed as:

$$\begin{aligned} & \max_{\pi \in \Pi} E(U(\pi)) & (4.13) \\ & \text{subject to } P\{\pi(\vec{U}) = i\} \geq r_i, \quad i = 1, 2, \dots, N, \\ & E\left(U_i \mathbf{1}_{\{\pi(\vec{U})=i\}}\right) \geq a_i E(U(\pi)), \quad i = 1, 2, \dots, N, \end{aligned}$$

where  $r_i$  and  $a_i$  are identical as those defined in Sections 4.3.2 and 4.3.3.

Define the policy  $\pi^*$  as follows:

$$\pi^*(\vec{U}) = \underset{i}{\operatorname{argmax}} ((\kappa + \gamma_i)U_i + \alpha_i), \quad (4.14)$$

where  $\kappa = 1 - \sum_{i=1}^N a_i \gamma_i$  and the control parameters  $\alpha_i$  and  $\gamma_i$  are chosen such that:

1.  $\alpha_i \geq 0, \gamma_i \geq 0, \forall i$ ;
2.  $P\{\pi^*(\vec{U}) = i\} \geq r_i, \forall i$ ;
3. If  $P\{\pi^*(\vec{U}) = i\} > r_i$ , then  $\alpha_i = 0, \forall i$ ; and
4.  $E\left(U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right) \geq a_i E(U(\pi)), \forall i$ ;
5. If  $E\left(U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}}\right) > a_i E(U(\pi))$ , then  $\gamma_i = 0, \forall i$ .

**Theorem 4.5** *The policy  $\pi^*$  defined in (4.14) is an optimal solution to the problem defined in (4.13), i.e., it maximizes the system performance, while satisfying the minimum temporal and utilitarian fairness constraints for individual users.*

*Proof:* Let  $\pi$  be a policy satisfying the minimum temporal and utilitarian

fairness constraints, and  $\alpha_i$  and  $\gamma_i$  satisfy conditions 1–5. Hence, we have

$$\begin{aligned}
E(U(\pi)) &\leq E(U(\pi)) + \sum_{i=1}^N \alpha_i \left( P\{\pi(\vec{U}) = i\} - r_i \right) \\
&\quad + \sum_{i=1}^N \gamma_i \left( E\left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - a_i E(U(\pi)) \right) \\
&= \left( 1 - \sum_{i=1}^N a_i \gamma_i \right) E(U(\pi)) + \sum_{i=1}^N \gamma_i E\left( U_i \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \\
&\quad + \sum_{i=1}^N \alpha_i P\{\pi(\vec{U}) = i\} - \sum_{i=1}^N \alpha_i r_i \\
&= E\left( \sum_{i=1}^N ((\kappa + \gamma_i)U_i + \alpha_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - \sum_{i=1}^N \alpha_i r_i
\end{aligned}$$

By the definition of  $\pi^*$ , we have

$$\sum_{i=1}^N ((\kappa + \gamma_i)U_i + \alpha_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \leq \sum_{i=1}^N ((\kappa + \gamma_i)U_i + \alpha_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}}$$

Thus,

$$E\left( \sum_{i=1}^N ((\kappa + \gamma_i)U_i + \alpha_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) \leq E\left( \sum_{i=1}^N ((\kappa + \gamma_i)U_i + \alpha_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right)$$

Hence,

$$\begin{aligned}
E(U(\pi)) &\leq E\left( \sum_{i=1}^N ((\kappa + \gamma_i)U_i + \alpha_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - \sum_{i=1}^N \alpha_i r_i \\
&= E(U(\pi^*)) + \sum_{i=1}^N \alpha_i \left( P\{\pi^*(\vec{U}) = i\} - r_i \right) \\
&\quad + \sum_{i=1}^N \gamma_i \left( E\left( U_i \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - a_i E(U(\pi^*)) \right) \tag{4.15} \\
&= E(U(\pi^*)).
\end{aligned}$$

where the second and third parts of (4.15) equal zero because of conditions 3 and 5 on  $\alpha_i$  and  $\gamma_i$ . ■

Similarly, we expect to derive the optimal opportunistic scheduling schemes for different combinations of QoS/fairness constraints. In general, there can be more than two different types of constraints (including maximum and minimum).

## 4.5 A Generalized Framework for Opportunistic Scheduling

We notice that the scheduling problems presented above, though with different fairness constraints, share some kind of general form. Therefore, we are interested to find out a generalized opportunistic scheduling framework to model and solve this category of scheduling problems with fairness constraints.

For generalization purpose, we start by defining a new set of notation.

- $f_i(x)$  denotes the utility function associated with user  $i$ . Suppose that  $f_i(x)$  is a monotonically increasing function of  $x$ .
- $h_i^j$  denotes the  $j^{\text{th}}$  constraint function associated with user  $i$ , and  $g_i^k$  denotes the  $k^{\text{th}}$  constraint function associated with user  $i$ . We assume that the  $h_i^j$  and  $g_i^k$  are convex functions in their arguments.
- $H_i^j$  and  $G_i^k$  denotes the  $j^{\text{th}}$  and  $k^{\text{th}}$  predetermined constraint requirement associated with user  $i$  respectively.

The generalized QoS constrained scheduling problem can be formulated as a constrained optimization problem as follows:

$$\begin{aligned}
 & \max_{\pi \in \Pi} E \sum_{i=1}^N f_i(U_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \\
 & \text{subject to } E \left( h_i^j(U_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - H_i^j \geq 0, \quad i = 1, 2, \dots, N, j = 1, 2, \dots, J, \\
 & \quad E \left( g_i^k(U_i) \mathbf{1}_{\{\pi(\vec{U})=i\}} \right) - G_i^k \leq 0, \quad i = 1, 2, \dots, N, k = 1, 2, \dots, K.
 \end{aligned} \tag{4.16}$$

We define the policy  $\pi^*$  as follows:

$$\pi^*(\vec{U}) = \underset{i}{\operatorname{argmax}} \left\{ f_i(U_i) + \sum_{j=1}^J \lambda_i^j h_i^j(U_i) - \sum_{k=1}^K \rho_i^k g_i^k(U_i) \right\}, \tag{4.17}$$

where the control parameters  $\lambda_i^j$  and  $\rho_i^k$  are chosen such that:

1.  $\lambda_i^j \geq 0, \rho_i^k \geq 0, \forall i, \forall j, \forall k$ ;
2.  $E \left( h_i^j(U_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - H_i^j \geq 0, \forall i, \forall j$ ;
3. If  $E \left( h_i^j(U_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - H_i^j > 0$ , then  $\lambda_i^j = 0, \forall i, \forall j$ ;
4.  $E \left( g_i^k(U_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - G_i^k \geq 0, \forall i, \forall k$ ;
5. If  $E \left( g_i^k(U_i) \mathbf{1}_{\{\pi^*(\vec{U})=i\}} \right) - G_i^k > 0$ , then  $\rho_i^k = 0, \forall i, \forall k$ .

**Theorem 4.6** *The policy  $\pi^*$  defined in (4.17), if one exists, is an optimal solution to the problem defined in (4.16), i.e., it maximizes the system performance, while satisfying the general fairness constraints for individual users.*

*Proof:* By formulating the constraints as the Lagrangian multipliers, the proof follows the similar steps as the proofs in the previous sections. The details are omitted here. ■

The proposed framework is a generalization and abstraction of Liu et al.'s work. The framework can accommodate various fairness constraints, not just limited to the temporal fairness and minimum-performance guarantees. Our proposed scheduling problems with mixed type constraints and maximum-minimum bounds also fits in very well.

The generalized optimal framework for opportunistic scheduling provides us an efficient tool to link successful optimization control and economy models to the engineering problems, especially in designing and analyzing the scheduling problems with the heterogeneous users' QoS/fairness constraints over wireless networks.

## 4.6 Conclusions

Opportunistic transmission scheduling is a promising technology to improve spectrum efficiency by exploiting time-varying channel conditions. In this chapter, we first consider the scheduling problems with both minimum and maximum constraints.

We derive the corresponding opportunistic scheduling policies for the three long-term QoS/fairness constraints—temporal fairness, utilitarian fairness, and minimum-performance guarantees. Then we deal with scheduling problems with multiple type mixed QoS/fairness constraints. Finally, we develop a generalized opportunistic scheduling framework to accommodate those scheduling schemes. We show that the structure of the optimal opportunistic scheduling policy is carried over to the problem with general constraints. The generalized optimal framework for opportunistic scheduling provides us an efficient tool to design and analyze the scheduling problems with the heterogeneous users' QoS/fairness constraints over wireless networks.

## CHAPTER 5

# STOCHASTIC DYNAMIC PROGRAMMING FOR OPPORTUNISTIC SCHEDULING

In this chapter, we consider the problem of fair scheduling of queued data transmissions in wireless networks. We deal with both the throughput maximization problem and the delay minimization problem. Taking fairness constraints and the data arrival queues into consideration, we formulate the transmission scheduling problem as a Markov decision process (MDP) with fairness constraints. We consider two criteria: infinite horizon expected total discounted reward and expected average reward. Applying the dynamic programming approach, we derive and prove explicit optimality equations for the above constrained MDPs, and give corresponding optimal fair scheduling policies based on those equations. A practical stochastic-approximation-type algorithm is applied to calculate the control parameters online in the policies. Furthermore, we develop a novel approximation method—temporal fair rollout—to achieve a tractable computation. Numerical results show that the proposed scheme achieves significant performance improvement for both throughput maximization and delay minimization problems compared with other existing schemes.

This chapter is organized as follows. In Section 5.1, we give the introduction for our research. In Section 5.2, we describe our system model and MDP formulation. In Section 5.3, we derive the dynamic programming equation and the optimal scheduling

policy for the temporal fair constrained problem with both expected total discounted reward and expected average reward criteria. In Section 5.4, we derive the dynamic programming equations and the optimal scheduling policies for the utilitarian fair constrained problem with both expected total discounted reward and expected average reward criteria. In Section 5.5, we propose an efficient approximation algorithm—temporal fair rollout. We discuss stochastic approximation method for parameter estimation in Section 5.6. In Section 5.7, we present and analyze the channel model and the simulation results. Finally, concluding remarks are given in Section 5.8.

## 5.1 Introduction

Next generation wireless networks, which support high-speed packet data while providing heterogeneous QoS guarantees, require flexible and efficient radio resource scheduling schemes. One of the fundamental characteristics of wireless networks is the *time-varying* and *location-dependent* channel conditions due to multipath fading. Efficient exploitation of such channel variation has attracted significant research interest in the past decade [32, 34, 62].

From an information-theoretic viewpoint, Knopp and Humblet showed that the system capacity is maximized by exploiting inherent *multiuser diversity* gain in the wireless channel. The basic idea is to schedule a single user with the best instantaneous channel condition to transmit at any one time. Technology based on this idea has already been implemented in the current 3G systems: High Data Rate (HDR) and high-speed downlink packet access (HSDPA).

Good scheduling schemes in wireless networks should *opportunistically* seek to exploit the time-varying channel conditions to improve spectrum efficiency thereby achieving *multiuser diversity* gain. In this context, it is also important to consider the tradeoff between wireless resource efficiency and level of satisfaction among individual users (fairness).



Fairness criteria are critical to the scheduling problem in wireless networks. For example, allowing only users close to the base station to transmit at high transmission rate may result in very high throughput, but sacrifice the transmission of other users. Liu et al. developed a unified *opportunistic scheduling* framework for multimedia communication in a cellular system, while providing three long-term QoS/fairness guarantees—temporal fairness, utilitarian fairness, and minimum-performance guarantees.

In this chapter, we consider an opportunistic fair scheduling problem for the uplink of a single-cell Time Division Multiplexing (TDM) system. We provide a novel formulation of the scheduling problem with a Markov channel model as an MDP with explicit fairness constraints. We deal with both the *throughput maximization* problem and the *delay minimization* problem. We consider two criteria: *infinite horizon expected total discounted reward* and *expected average reward*. In either case, we characterize the corresponding optimal MDP-based fair scheduling scheme. We focus on two categories of fairness constraints, namely *temporal fairness* and *utilitarian fairness*. Owing to the particular characteristics of the constraints, we are able to derive and prove explicit dynamic programming equations for MDPs with fairness constraints. Based on these optimality equations, we obtain the exact corresponding optimal scheduling policies. A practical stochastic approximation algorithm is applied to calculate the control parameters online in the policies. Furthermore, based on the rollout algorithm, we develop a novel approximation method—temporal fair rollout—to achieve a tractable computation.

Our work addresses heterogeneity of networks in three dimensions. First, there is heterogeneity in the channel conditions, owing to factors such as path loss, shadowing, and fading. Second is heterogeneity in the utility of the channel, which depends on factors such as the heterogeneity of the end-user devices and their capability (e.g., transmission power, battery capacity, signal-processing hardware, and application

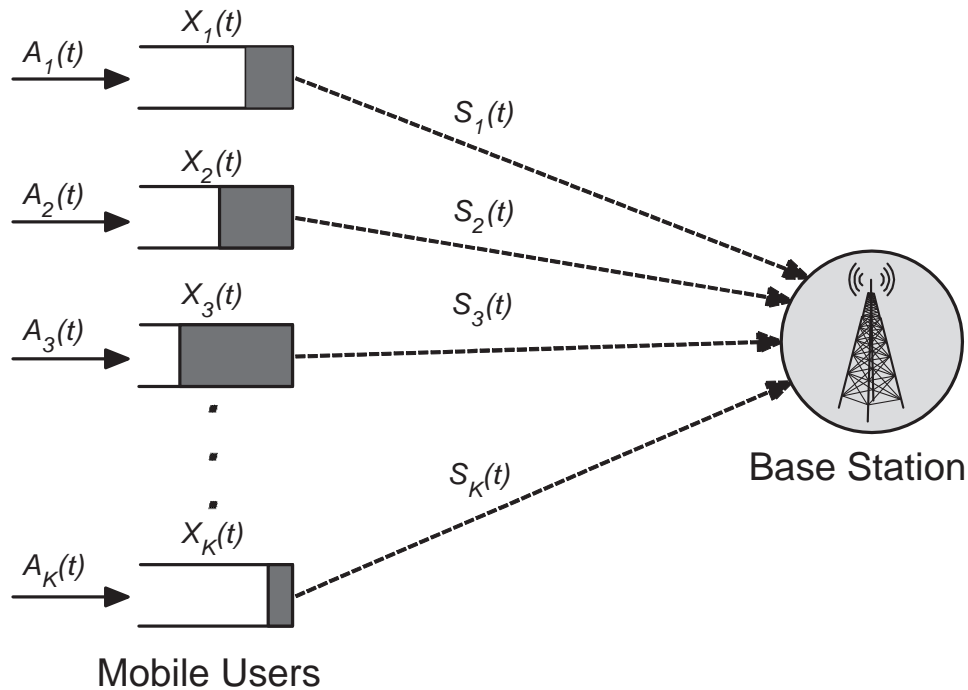
software). Third, there is heterogeneity in the end-user QoS and fairness requirements.

Our proposed scheme can easily be extended to different objective functions and other fairness measures. Although we only focus on uplink scheduling, the scheme is equally applicable to the downlink case.

## 5.2 System Model And Problem Formulation

### 5.2.1 System Model

Fig. 5.1 depicts an uplink data queueing model for a single-cell TDM system. We assume that there is a base station receiving data from  $K$  mobile users. A scheduler, located at the base station, decides at the start of each scheduling interval which (single) user to serve. We call a decision rule for scheduling which user to transmit at each interval a *scheduling policy*.



**Figure 5.1:** Uplink queueing model of a single-cell TDM system.

The wireless channel for each user differs depending on the location, the surrounding environment, and the mobility. Here we assume that the base station knows the channel state information (CSI) of all users perfectly. Each user has its own packet queue for transmission with unlimited queue capacity. We assume that packets arrive in each queue randomly (according to some distribution) and independently. The length of a scheduling interval (*time slot*) is fixed, and the channel does not vary significantly during a time slot. We also assume that all users have the same fixed packet size.

In practice, before each scheduling interval, all users need to report their current CSI to the base station. So, the perfect CSI assumption here potentially involves significant feedback signaling cost [97]. This issue has motivated the recent research interest in opportunistic scheduling with partial or reduced feedback [98–100].

Let  $t = 0, 1, \dots$  be the index of time slots, and  $k = 1, \dots, K$  the index of users. For user  $k$  at time slot  $t$ , we use  $X_k(t)$ ,  $S_k(t)$ , and  $A_k(t)$  to denote the queue length, the channel state, and the exogenous packet arrivals respectively (all in terms of number of packets). The channel state here is measured by the maximum number of packets each user can transmit to the base station at each time slot.

Let  $\pi_t$  be the user scheduled at time slot  $t$  given a scheduling policy  $\pi$ . Using this notation, the queue length evolution is given by, for all  $k \in 1, \dots, K$ ,

$$X_k(t+1) = X_k(t) + A_k(t) - \min(X_k(t), S_k(t)) \mathbf{1}_{\{\pi_t=k\}},$$

where  $\mathbf{1}_{\{\cdot\}}$  is the indicator function.

### 5.2.2 MDP Problem Formulation

Recall that in Section 2.5, a discrete-time, finite-state Markov decision process (MDP) is specified by a tuple  $(S, A, P(\cdot|\cdot, \cdot), r(\cdot, \cdot))$ . The *state space*  $S$  and the *action space*  $A$  are finite sets. At time slot  $t$ , if the system is in state  $s \in S$  and action  $a \in A$  is chosen, then the following happens:

1. a reward  $r(s, a)$  is earned immediately;
2. the process moves to state  $s' \in S$  with *transition probability*  $P(s'|s, a)$ , where  $P(s'|s, a) \geq 0$  and  $\sum_{s'} P(s'|s, a) = 1$  for all  $s$  and  $a$ .

The goal is to determine a *policy*, a decision rule for action selection at each time, to optimize a given performance criterion. This optimization involves balancing between immediate reward and future rewards: a high reward now may lead the process into a bad situation later.

We formulate our scheduling problem as a MDP as follows:

- **State:** The state space  $S$  is the set of all vectors  $s \in \mathbb{R}^{2K}$  of the form

$$s = (x_1, x_2, \dots, x_K, s_1, s_2, \dots, s_K),$$

where  $x_k$  and  $s_k$  are the queue length and the channel state of user  $k$  during a generic time slot. The state of the system at time slot  $t$  is

$$X_t = (X_1(t), X_2(t), \dots, X_K(t), S_1(t), S_2(t), \dots, S_K(t)).$$

- **Action:** The action at each time is to choose one of  $K$  users for transmission; thus an action here corresponds to a user. The action space  $A$  is thus

$$A = \{1, 2, \dots, K\}.$$

- **Transition probability function:** Since a state consists of all queue lengths and channel state, the transition probability function is determined by the queue length evolution formula and the dynamics of the channels.
- **Reward:** We consider the following two problems: the *throughput maximization* problem and the *delay minimization* problem.

The throughput maximization problem involves maximizing the system throughput with the fairness constraints (described below). The throughput

in a time slot is defined as the actual number of packets transmitted between a user and the base station in the time slot. The corresponding reward function is given by

$$r(X_t, \pi_t) = \sum_{k=1}^K \mathbf{1}_{\{\pi_t=k\}} \min(X_k(t), S_k(t)). \quad (5.1)$$

Note that the throughput for user  $k$  is the minimum of the queue length  $X_k(t)$  and the available channel transmission packets  $S_k(t)$  because at most  $X_k(t)$  packets can be transmitted at time slot  $t$ .

The delay minimization problem is to minimize the sum of the user queue lengths with the fairness constraints. The corresponding reward function is given by

$$r(X_t, \pi_t) = - \sum_{k=1}^K X_k(t). \quad (5.2)$$

(The negative sign accounts for minimization.)

Each of these reward functions leads to an overall objective function to be maximized, defined roughly as the long-term cumulative reward; these are defined precisely in the next two sections.

- **Policy:** In this chapter, the space of policies under consideration is restricted to stationary policies. A stationary policy is a mapping  $\pi : S \rightarrow A$  from the state space  $S$  to the action space  $A$ ; i.e., the stationary policy  $\pi$  selects action  $\pi(s)$  when the process is in state  $s$ . Let  $\Pi$  be the set of all stationary policies.

A natural fairness criterion is to give each user a certain long-term fraction of time, because time is the basic resource shared among users. This is called the *temporal fairness* [46], which is closely related to generalized processor sharing (GPS) in wire-line networks [23]. An alternative fairness criterion, called *utilitarian fairness*, would ensure that all users get at least a certain fraction of the overall system performance.

Based on the MDP model described above, our goal can be formally stated as: *find a policy  $\pi$  that maximizes the specified objective function  $J_\pi$  while satisfying the*

*corresponding fairness constraints.*

In this chapter, we consider infinite-horizon models. In the following sections, we will discuss the above scheduling problems with two types of objective functions: the expected total discounted reward and the expected average reward criteria.

## 5.3 Temporal Fairness scheduling

### 5.3.1 Expected Total Discounted Reward Criterion

Discounting arises naturally in applications in which we account for the time value of the rewards, such as in economic problems. The discount factor  $\alpha$  measures the present value of one unit of currency received in the future. The meaning of  $\alpha < 1$  is that future rewards matter to us less than the same reward received at the present time [80, 81].

In this subsection, we study the *infinite horizon expected total discounted reward* MDP problem with the expected discounted temporal fairness constraints. We derive and prove an explicit dynamic programming equation for the constrained MDP, and give an optimal scheduling policy based on that equation.

#### Problem formulation

For any policy  $\pi$ , we define the expected discounted reward objective function as

$$J_\pi(s) = \lim_{T \rightarrow \infty} E_\pi \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t) \middle| X_0 = s \right], \quad s \in S,$$

where  $E_\pi$  represents expectation given that a policy  $\pi$  is employed,  $\alpha$  is the discount factor with  $0 < \alpha < 1$ , and  $X_0$  is the initial state. Since  $r(X_t, \pi_t)$  is the immediate reward received at time  $t$ , it follows that  $J_\pi(s)$  represents the expected total discounted reward received when the policy  $\pi$  is employed and the initial state is  $s$ . A policy  $\pi^*$  is said to be  $\alpha$ -optimal if

$$J_{\pi^*}(s) = \max_{\pi \in \Pi} J_\pi(s), \quad \forall s \in S. \quad (5.3)$$

Hence, a policy is  $\alpha$ -optimal if its expected discounted reward is maximal for every initial state.

The expected discounted temporal fairness constraint is

$$\lim_{T \rightarrow \infty} E_\pi \left[ \sum_{t=0}^{T-1} \alpha^t \mathbf{1}_{\{\pi_t=a\}} \middle| X_0 = s \right] \geq C(a), \quad \forall a \in A, \quad (5.4)$$

where  $C(a)$  denotes the minimum discounted time-fraction in which action (user)  $a$  should be chosen, with  $0 \leq C(a) \leq 1$  and  $\sum_{a \in A} C(a) \leq 1$ .

Therefore, our goal can be stated as: *find an  $\alpha$ -optimal policy  $\pi^*$  subject to the expected discounted temporal fairness constraint.*

### Optimal scheduling policy

Theoretically, the above constrained optimization problem can be solved directly by linear programming or Lagrangian methods [55–57]. Practically, those methods are computationally formidable even for problems with moderate state spaces. Moreover, they cannot be used if the state-transition distribution is not available explicitly. Dynamic programming can be used to solve such problems online iteratively. Here we will derive and prove an explicit dynamic programming equation for the above constrained MDP. Then we characterize an optimal solution to find (5.3) subject to (5.4).

Given a function  $u : A \rightarrow \mathbb{R}$ , for any policy  $\pi$ , we define

$$V_\pi(s) = \lim_{T \rightarrow \infty} E_\pi \left[ \sum_{t=0}^{T-1} \alpha^t [r(X_t, \pi_t) + u(\pi_t)] \middle| X_0 = s \right], \quad s \in S,$$

and let

$$V_\alpha(s) = \sup_{\pi \in \Pi} V_\pi(s), \quad s \in S.$$

**Lemma 5.1** *Given a function  $u : A \rightarrow \mathbb{R}$ ,  $V_\alpha$  satisfies the optimality equation*

$$V_\alpha(s) = \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\}, \quad s \in S. \quad (5.5)$$

*Proof:* Let  $\pi$  be an arbitrary policy, and suppose that  $\pi$  chooses action  $a$  at time slot 0 with probability  $P_a$ ,  $a \in A$ . Then,

$$V_\pi(s) = \sum_{a \in A} P_a \left[ r(s, a) + u(a) + \sum_{s' \in S} P(s'|s, a) W_\pi(s') \right],$$

where  $W_\pi(s')$  represents the expected discounted weighted reward with the weight  $u(\pi_t)$  incurred from time slot 1 onwards, given that  $\pi$  is employed and the state at time 1 is  $s'$ . However, it follows that

$$W_\pi(s') \leq \alpha V_\alpha(s')$$

and hence that

$$\begin{aligned} V_\pi(s) &\leq \sum_{a \in A} P_a \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\} \\ &\leq \sum_{a \in A} P_a \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\} \\ &= \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\}. \end{aligned} \quad (5.6)$$

Since  $\pi$  is arbitrary, (5.6) implies that

$$V_\alpha(s) \leq \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\}. \quad (5.7)$$

To go the other way, let  $a_0$  be such that

$$r(s, a_0) + u(a_0) + \alpha \sum_{s' \in S} P(s'|s, a_0) V_\alpha(s') = \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\} \quad (5.8)$$

and let  $\pi$  be the policy that chooses  $a_0$  at time 0; and, if the next state is  $s'$ , views the process as originating in state  $s'$ ; and follows a policy  $\pi_{s'}$ , which is such that  $V_{\pi_{s'}}(s') \geq V_\alpha(s') - \varepsilon$ ,  $s' \in S$ . Hence,

$$\begin{aligned} V_\pi(s) &= r(s, a_0) + u(a_0) + \alpha \sum_{s' \in S} P(s'|s, a_0) V_{\pi_{s'}}(s') \\ &\geq r(s, a_0) + u(a_0) + \alpha \sum_{s' \in S} P(s'|s, a_0) V_\alpha(s') - \alpha \varepsilon \end{aligned}$$



which, since  $V_\alpha(s) \geq V_\pi(s)$ , implies that

$$V_\alpha(s) \geq r(s, a_0) + u(a_0) + \alpha \sum_{s' \in S} P(s'|s, a_0) V_\alpha(s') - \alpha \varepsilon.$$

Hence, from (5.8), we have

$$V_\alpha(s) \geq \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\} - \alpha \varepsilon. \quad (5.9)$$

Since  $\pi_{s'}$  could be arbitrary, then  $\varepsilon$  is arbitrary, from (5.7) and (5.9), we have

$$V_\alpha(s) = \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\}, \quad s \in S.$$

■

Now let  $B(S)$  be the Banach space of real-valued bounded functions on the state space  $S$ . Note that since rewards are bounded,  $V_\pi \in B(S)$  for any policy  $\pi$ . For any stationary policy  $\pi$  we define the mapping

$$T_\pi : B(S) \rightarrow B(S)$$

in the following manner:

$$(T_\pi v)(s) = r(s, \pi(s)) + u(\pi(s)) + \alpha \sum_{s' \in S} P(s'|s, \pi(s)) v(s').$$

We can interpret  $T_\pi v$  at  $s$  as representing the expected weighted reward if we use policy  $\pi$  but terminate it after one period and receive a final reward  $\alpha v(s')$  when the final state is  $s'$ .

The following lemma and theorem characterize the optimal policy for our temporal fair constrained MDP and the corresponding optimal discounted reward.

**Lemma 5.2** *Let  $u : A \rightarrow \mathbb{R}$  satisfy  $u(a) \geq 0$  for all  $a \in A$ . Let  $\pi^*$  be a stationary policy that when the process is in state  $s$ , selects an action maximizing the right-hand side of (5.5):*

$$\pi^*(s) = \operatorname{argmax}_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, a) V_\alpha(s') \right\}. \quad (5.10)$$

Then

$$V_{\pi^*}(s) = V_\alpha(s), \quad \forall s \in S.$$

*Proof:* By applying the mapping  $T_{\pi^*}$  to  $V_\alpha$ , we obtain

$$\begin{aligned} (T_{\pi^*}V_\alpha)(s) &= r(s, \pi^*(s)) + u(\pi^*(s)) + \alpha \sum_{s' \in S} P(s'|s, \pi^*(s))V_\alpha(s') \\ &= \max_{a \in A} \left\{ r(s, a) + u(a) + \alpha \sum_{s' \in S} P(s'|s, \pi^*(s))V_\alpha(s') \right\} \\ &= V_\alpha(s), \end{aligned}$$

where the last equation follows from Lemma 5.1. Hence,

$$T_{\pi^*}V_\alpha = V_\alpha,$$

which implies that

$$T_{\pi^*}^2V_\alpha = T_{\pi^*}(T_{\pi^*}V_\alpha) = T_{\pi^*}V_\alpha = V_\alpha$$

and by induction we have,

$$T_{\pi^*}^nV_\alpha = V_\alpha, \quad \forall n.$$

Letting  $n \rightarrow \infty$  and using Banach fixed-point theorem yields the result,

$$V_{\pi^*}(s) = V_\alpha(s), \quad \forall s \in S.$$

■

We now show that, under certain assumptions, the policy  $\pi^*$  in Lemma 5.2 is an  $\alpha$ -optimal policy for the discounted temporal fair constrained MDP.

**Theorem 5.1** *Suppose there exists a function  $u : A \rightarrow \mathbb{R}$  such that:*

1.  $\forall a \in A, u(a) \geq 0;$
2.  $\forall a \in A, \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t \mathbf{1}_{\{\pi_t^*=a\}} | X_0 = s \right] \geq C(a);$
3.  $\forall a \in A, \text{ if } \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t \mathbf{1}_{\{\pi_t^*=a\}} | X_0 = s \right] > C(a), \text{ then } u(a) = 0.$

Then  $\pi^*$  defined in (5.10) is an  $\alpha$ -optimal policy as defined by (5.3) subject to (5.4).

The corresponding optimal discounted reward is

$$J_{\pi^*}(s) = V_{\pi^*}(s) - \sum_{a \in A} u(a)C(a), \quad \forall s \in S. \quad (5.11)$$

*Proof:* Let  $\pi$  be a policy satisfying the expected discounted temporal fairness constraint. And suppose there exists  $u : A \rightarrow \mathbb{R}$  satisfying conditions 1-3. Then,

$$\begin{aligned} J_{\pi}(s) &= \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t) \middle| X_0 = s \right] \\ &\leq \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t) \middle| X_0 = s \right] \\ &\quad + \sum_{a \in A} u(a) \left( \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t \mathbf{1}_{\{\pi_t = a\}} \middle| X_0 = s \right] - C(a) \right) \\ &= \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t) \middle| X_0 = s \right] \\ &\quad + \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t u(\pi_t) \middle| X_0 = s \right] - \sum_{a \in A} u(a)C(a) \\ &= \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t [r(X_t, \pi_t) + u(\pi_t)] \middle| X_0 = s \right] - \sum_{a \in A} u(a)C(a) \\ &= V_{\pi}(s) - \sum_{a \in A} u(a)C(a). \end{aligned}$$

Since  $V_\pi(s) \leq V_\alpha(s) = V_{\pi^*}(s)$  from Lemma 5.2, we have

$$J_\pi(s) \leq V_{\pi^*}(s) - \sum_{a \in A} u(a)C(a) \quad (5.12)$$

$$\begin{aligned} &= \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t [r(X_t, \pi_t^*) + u(\pi_t^*)] \middle| X_0 = s \right] - \sum_{a \in A} u(a)C(a) \\ &= \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t^*) \middle| X_0 = s \right] \\ &+ \sum_{a \in A} u(a) \left( \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t \mathbf{1}_{\{\pi_t^* = a\}} \middle| X_0 = s \right] - C(a) \right) \quad (5.13) \\ &= \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t^*) \middle| X_0 = s \right] \\ &= J_{\pi^*}(s), \end{aligned}$$

where the second part of (5.13) equals zero because of condition 3 on  $u$ . From (5.12), we get the corresponding optimal discounted reward is

$$J_{\pi^*}(s) = V_{\pi^*}(s) - \sum_{a \in A} u(a)C(a), \quad \forall s \in S.$$

■

Lemma 5.2 and Theorem 5.1 provide an optimal scheduling policy for the discounted temporal fair constrained MDP. The  $\alpha$ -optimal scheduling policy  $\pi^*$  is given by (5.10), and the corresponding optimal discounted reward is given by (5.11).

We can think of the parameter  $u(a)$  in Theorem 5.1 as an “offset” or “threshold” for each user (action) to satisfy the fairness constraint, analogous to the result of [46]. Under this constraint, the scheduling policy schedules the “relatively best” user to transmit. It is straightforward to see that by setting  $u(a) = 0$  for all  $a \in A$ , the optimal policy reduces to an optimal policy for a standard (unconstrained) MDP. However, that policy could be unfair to certain users. If  $u(a) > 0$ , then user  $a$  is an “unfortunate” user, i.e., the channel condition it experiences is relatively poor. Hence, it has to take advantage of other users (e.g., users with  $u(a) = 0$ ) to satisfy its

fairness requirement. But to maximize the overall system performance, we can only give the “unfortunate” users their minimum resource requirements, hence condition 3 for  $u(a)$ .

### 5.3.2 Expected Average Reward Criterion

In the previous subsection, we posed the temporal fairness scheduling problem as an expected discounted reward MDP with constraints. In this subsection, we consider optimization problems with average reward criteria. Such problems are common in economic, computer, and communication systems. Some examples are inventory control problems and computer communication networks, where decisions are made based on throughput rate or average time a job or packet remains in the system [80].

In this subsection, we study the problem as an *infinite horizon average reward* MDP with expected average temporal fairness constraints. Analogous to the results in the last subsection, we derive and prove an explicit dynamic programming equation for the constrained MDP, and give an optimal scheduling policy based on that equation.

#### Problem formulation

For any policy  $\pi$ , we define the expected average reward objective function as

$$J_\pi(s) = \lim_{T \rightarrow \infty} E_\pi \left[ \frac{1}{T} \sum_{t=0}^{T-1} r(X_t, \pi_t) \middle| X_0 = s \right], \quad s \in S, \quad (5.14)$$

where  $E_\pi$  represents conditional expectation given that the policy  $\pi$  is employed. Since  $r(X_t, \pi_t)$  is the immediate reward received at time  $t$ , it follows that  $J_\pi(s)$  represents the expected average reward received per stage when the policy  $\pi$  is employed and the initial state is  $s$ . If the limit in (5.14) does not exist, then we agree to use limsup instead of lim. We say that the policy  $\pi^*$  is *average-reward-optimal* if

$$J_{\pi^*}(s) = \max_{\pi \in \Pi} J_\pi(s), \quad \forall s \in S. \quad (5.15)$$

The expected average temporal fairness constraint is defined as

$$\lim_{T \rightarrow \infty} E_\pi \left[ \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}_{\{\pi_t = a\}} \middle| X_0 = s \right] \geq C(a), \quad \forall a \in A, \quad (5.16)$$

where  $C(a)$  denotes the minimum relative frequency at which action  $a$  should be taken, with  $C(a) \geq 0$  and  $\sum_{a \in A} C(a) \leq 1$ .

Therefore, our goal can be stated as: *find an average-reward-optimal policy  $\pi^*$  subject to the expected average temporal fairness constraint.*

### Optimal scheduling policy

Here we derive and prove an explicit dynamic programming equation for the above constrained MDP, and give an optimal scheduling policy based on that equation.

**Theorem 5.2** *Suppose the system is unichain.<sup>1</sup> Suppose we have a bounded function  $h : S \rightarrow \mathbb{R}$ , a function  $u : A \rightarrow \mathbb{R}$ , a constant  $g$ , and a stationary policy  $\pi^*$  such that for  $s \in S$ ,*

1.  $\forall a \in A, u(a) \geq 0$ ;
2.  $\forall a \in A, \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}_{\{\pi_t^* = a\}} | X_0 = s \right] \geq C(a)$ ;
3.  $\forall a \in A, \text{if } \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}_{\{\pi_t^* = a\}} | X_0 = s \right] > C(a), \text{ then } u(a) = 0$ ;

4.

$$g + h(s) = \max_{a \in A} \left\{ r(s, a) + u(a) + \sum_{s' \in S} P(s' | s, a) h(s') \right\}; \quad (5.17)$$

5.  $\pi^*$  is a policy which, for each  $s$ , prescribes an action which maximizes the right-hand side of (5.17):

$$\pi^*(s) = \operatorname{argmax}_{a \in A} \left\{ r(s, a) + u(a) + \sum_{s' \in S} P(s' | s, a) h(s') \right\}. \quad (5.18)$$

Then  $\pi^*$  is an average-reward-optimal policy as defined by (5.15) subject to (5.16).

The corresponding optimal average reward is

$$J_{\pi^*}(s) = g - \sum_{a \in A} u(a) C(a), \quad \forall s \in S. \quad (5.19)$$

---

<sup>1</sup>An MDP is unichain if the transition matrix corresponding to every deterministic stationary policy consists of one single recurrent class plus a possibly empty set of transient state [80].

*Proof:* Let  $\pi$  be a policy satisfying the expected average temporal fairness constraint; and let  $H_t = (X_0, \pi_0, \dots, X_{t-1}, \pi_{t-1}, X_t, \pi_t)$  denote the history of the process up to time  $t$ . First, we have

$$E_\pi \left\{ \sum_{t=1}^T [h(X_t) - E_\pi(h(X_t)|H_{t-1})] \right\} = 0,$$

since

$$\begin{aligned} E_\pi \left\{ \sum_{t=1}^T [h(X_t) - E_\pi(h(X_t)|H_{t-1})] \right\} &= \sum_{t=1}^T E_\pi [h(X_t) - E_\pi(h(X_t)|H_{t-1})] \\ &= \sum_{t=1}^T \{E_\pi[h(X_t)] - E_\pi[E_\pi(h(X_t)|H_{t-1})]\} \\ &= \sum_{t=1}^T \{E_\pi[h(X_t)] - E_\pi[h(X_t)]\} = 0. \end{aligned}$$

Also,

$$\begin{aligned} E_\pi[h(X_t)|H_{t-1}] &= \sum_{s' \in S} h(s')P(s'|X_{t-1}, \pi_{t-1}) \\ &= r(X_{t-1}, \pi_{t-1}) + u(\pi_{t-1}) + \sum_{s' \in S} h(s')P(s'|X_{t-1}, \pi_{t-1}) \\ &\quad - r(X_{t-1}, \pi_{t-1}) - u(\pi_{t-1}) \\ &\leq \max_{a \in A} \left\{ r(X_{t-1}, a) + u(a) + \sum_{s' \in S} P(s'|X_{t-1}, a)h(s') \right\} \\ &\quad - r(X_{t-1}, \pi_{t-1}) - u(\pi_{t-1}) \\ &= g + h(X_{t-1}) - r(X_{t-1}, \pi_{t-1}) - u(\pi_{t-1}) \end{aligned}$$

with equality for  $\pi^*$ , since  $\pi^*$  is defined to take the maximizing action. Hence,

$$\begin{aligned} 0 &\geq E_\pi \left\{ \sum_{t=1}^T [h(X_t) - g - h(X_{t-1}) + r(X_{t-1}, \pi_{t-1}) + u(\pi_{t-1})] \right\} \\ \Leftrightarrow g &\geq E_\pi \frac{h(X_T)}{T} - E_\pi \frac{h(X_0)}{T} + E_\pi \frac{1}{T} \sum_{t=1}^T r(X_{t-1}, \pi_{t-1}) + E_\pi \frac{1}{T} \sum_{t=1}^T u(\pi_{t-1}). \end{aligned}$$

Letting  $T \rightarrow \infty$  and using the fact that  $h$  is bounded, we have that

$$\begin{aligned}
g &\geq J_\pi(X_0) + \lim_{T \rightarrow \infty} E_\pi \frac{1}{T} \sum_{t=0}^{T-1} u(\pi_t) \\
\Leftrightarrow g - \sum_{a \in A} u(a)C(a) &\geq J_\pi(X_0) + \lim_{T \rightarrow \infty} E_\pi \frac{1}{T} \sum_{t=0}^{T-1} u(\pi_t) - \sum_{a \in A} u(a)C(a) \\
&= J_\pi(s) + \lim_{T \rightarrow \infty} E_\pi \left[ \frac{1}{T} \sum_{t=0}^{T-1} \sum_{a \in A} u(a) \mathbf{1}_{\{\pi_t=a\}} \middle| X_0 = s \right] \\
&\quad - \sum_{a \in A} u(a)C(a) \\
&= J_\pi(s) + \sum_{a \in A} u(a) \left( \lim_{T \rightarrow \infty} E_\pi \left[ \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}_{\{\pi_t=a\}} \middle| X_0 = s \right] - C(a) \right). \tag{5.20}
\end{aligned}$$

Since we know that  $u \geq 0$ , and that the policy  $\pi$  satisfies the temporal fairness constraints, the second part of (5.20) is greater than or equal to zero. We get

$$g - \sum_{a \in A} u(a)C(a) \geq J_\pi(s).$$

With policy  $\pi^*$ , we have

$$\begin{aligned}
g - \sum_{a \in A} u(a)C(a) &= J_{\pi^*}(s) + \sum_{a \in A} u(a) \left( \lim_{T \rightarrow \infty} E_\pi \left[ \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}_{\{\pi_t^*=a\}} \right] - C(a) \right) \tag{5.21} \\
&= J_{\pi^*}(s),
\end{aligned}$$

where the second part of (5.21) equals to zero because of condition 3 on  $u(a)$ . Hence, the desired result is proven.  $\blacksquare$

The *average-reward-optimal* policy  $\pi^*$  is given by (5.18), and the corresponding optimal average reward is given by (5.19).

Analogous to  $u(a)$  in the last section,  $u(a)$  in (5.17) can be considered as an “offset” for each user to satisfy the average temporal fairness constraint. If we relax the fairness constraint by letting  $C(a) = 0$  for all  $a \in A$ , the optimal policy would reduce to an optimal policy for a standard (unconstrained) average reward MDP, as expected.



## 5.4 Utilitarian Fairness Scheduling

In the previous section, we studied the scheduling problem with temporal fairness constraints. In wireline networks, when a certain amount of resource is assigned to a user, it is equivalent to granting the user a certain amount of throughput. However, the situation is different in wireless networks, where the performance value and the amount of resource are not directly related. Therefore, a potential problem in wireless network is that the temporal fairness scheme has no way of explicitly ensuring that each user receives a certain guaranteed fair amount of utility (e.g., data rate). Hence, in this section we will describe an alternative fair scheduling problem that would ensure that all users get at least a certain fraction of the overall system performance, called *utilitarian fairness scheduling*.

We consider both the infinite horizon expected total discounted and average reward criteria here. In either case, we characterize the corresponding optimal MDP-based fair scheduling scheme.

### 5.4.1 Expected Total Discounted Reward Criterion

#### Problem formulation

As defined in Section 5.3.1, for any policy  $\pi$ , the expected discounted reward objective function is

$$J_\pi(s) = \lim_{T \rightarrow \infty} E_\pi \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t) \middle| X_0 = s \right], \quad s \in S.$$

The expected discounted utilitarian fairness constraint is

$$\lim_{T \rightarrow \infty} E_\pi \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t) \mathbf{1}_{\{\pi_t = a\}} \middle| X_0 = s \right] \geq D(a) J_\pi(s), \quad \forall a \in A, \quad (5.22)$$

where  $D(a)$  denotes the minimum discounted fraction of overall system performance in which action (user)  $a$  should be chosen, with  $0 \leq D(a) \leq 1$  and  $\sum_{a \in A} D(a) \leq 1$ .

Therefore, our goal can be stated as: *find an  $\alpha$ -optimal policy  $\pi^*$  subject to the expected discounted utilitarian fairness constraint.*

### Optimal scheduling policy

Given a function  $\omega : A \rightarrow \mathbb{R}$ , for any policy  $\pi$ , we define

$$U_\pi(s) = \lim_{T \rightarrow \infty} E_\pi \left[ \sum_{t=0}^{T-1} \alpha^t [(\kappa + \omega(\pi_t))r(X_t, \pi_t)] \middle| X_0 = s \right], \quad s \in S,$$

where  $\kappa = 1 - \sum_{\pi_t \in A} D(\pi_t)\omega(\pi_t)$  and let

$$U_\alpha(s) = \sup_{\pi \in \Pi} U_\pi(s), \quad s \in S.$$

**Lemma 5.3** *Given a function  $\omega : A \rightarrow \mathbb{R}$ ,  $U_\alpha$  satisfies the optimality equation*

$$U_\alpha(s) = \max_{a \in A} \left\{ (\kappa + \omega(a))r(s, a) + \alpha \sum_{s' \in S} P(s'|s, a)U_\alpha(s') \right\}, \quad s \in S. \quad (5.23)$$

*Proof:* The proof is similar to that of Lemma 5.1 in Section 5.3.1. The details are omitted for the sake of space. ■

The following lemma and theorem characterize the optimal policy for our utilitarian fair constrained MDP and the corresponding optimal discounted reward.

**Lemma 5.4** *Let  $\omega : A \rightarrow \mathbb{R}$  satisfy  $\omega(a) \geq 0$  for all  $a \in A$ . Let  $\pi^*$  be a stationary policy that when the process is in state  $s$ , selects an action maximizing the right-hand side of (5.23):*

$$\pi^*(s) = \operatorname{argmax}_{a \in A} \left\{ (\kappa + \omega(a))r(s, a) + \alpha \sum_{s' \in S} P(s'|s, a)U_\alpha(s') \right\}. \quad (5.24)$$

*Then*

$$U_{\pi^*}(s) = U_\alpha(s), \quad \forall s \in S.$$

*Proof:* The proof is similar to that of Lemma 5.2 in Section 5.3.1. ■

We now show that, under certain assumptions, the policy  $\pi^*$  in Lemma 5.4 is an  $\alpha$ -optimal policy for the discounted utilitarian fair constrained MDP.

**Theorem 5.3** *Suppose there exists a function  $\omega : A \rightarrow \mathbb{R}$  such that:*

1.  $\forall a \in A, \omega(a) \geq 0$ ;

2.  $\forall a \in A, \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t^*) \mathbf{1}_{\{\pi_t^*=a\}} \middle| X_0 = s \right] \geq D(a)J_{\pi^*}(s);$
3.  $\forall a \in A, \text{ if } \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t^*) \mathbf{1}_{\{\pi_t^*=a\}} \middle| X_0 = s \right] > D(a)J_{\pi^*}(s), \text{ then } \omega(a) = 0.$

Then  $\pi^*$  defined in (5.24) is an  $\alpha$ -optimal policy as defined by (5.3) subject to (5.22).

The corresponding optimal discounted reward is

$$J_{\pi^*}(s) = U_{\pi^*}(s), \quad \forall s \in S. \quad (5.25)$$

*Proof:* Let  $\pi$  be a policy satisfying the expected discounted utilitarian fairness constraint. And suppose there exists  $\omega : A \rightarrow \mathbb{R}$  satisfying conditions 1-3. Then,

$$\begin{aligned} J_{\pi}(s) &\leq J_{\pi}(s) + \sum_{a \in A} \omega(a) \left( \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t) \mathbf{1}_{\{\pi_t=a\}} \middle| X_0 = s \right] - D(a)J_{\pi}(s) \right) \\ &= J_{\pi}(s) + \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t \omega(\pi_t) r(X_t, \pi_t) \middle| X_0 = s \right] - \sum_{a \in A} \omega(a) D(a) J_{\pi}(s) \\ &= \lim_{T \rightarrow \infty} E_{\pi} \left[ \sum_{t=0}^{T-1} \alpha^t [(\kappa + \omega(\pi_t)) r(X_t, \pi_t)] \middle| X_0 = s \right] \\ &= U_{\pi}(s), \end{aligned}$$

where  $\kappa = 1 - \sum_{\pi_t \in A} D(\pi_t) \omega(\pi_t)$ . Since  $U_{\pi}(s) \leq U_{\alpha}(s) = U_{\pi^*}(s)$  from Lemma 5.4, we have

$$J_{\pi}(s) \leq U_{\pi^*}(s) \quad (5.26)$$

$$\begin{aligned} &= \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t^*) \middle| X_0 = s \right] \\ &\quad + \sum_{a \in A} \omega(a) \left( \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \sum_{t=0}^{T-1} \alpha^t r(X_t, \pi_t^*) \mathbf{1}_{\{\pi_t^*=a\}} \middle| X_0 = s \right] - D(a)J_{\pi^*}(s) \right) \end{aligned} \quad (5.27)$$

$$= J_{\pi^*}(s),$$

where the second part of (5.27) equals zero because of condition 3 on  $\omega$ . From (5.26), we get the corresponding optimal discounted reward is

$$J_{\pi^*}(s) = U_{\pi^*}(s), \quad \forall s \in S.$$

■

## 5.4.2 Expected Average Reward Criterion

### Problem formulation

As defined in Section 5.3.2, for any policy  $\pi$ , the expected average reward objective function is

$$J_{\pi}(s) = \lim_{T \rightarrow \infty} E_{\pi} \left[ \frac{1}{T} \sum_{t=0}^{T-1} r(X_t, \pi_t) \middle| X_0 = s \right], \quad s \in S.$$

The expected average utilitarian fairness constraint is

$$\lim_{T \rightarrow \infty} E_{\pi} \left[ \frac{1}{T} \sum_{t=0}^{T-1} r(X_t, \pi_t) \mathbf{1}_{\{\pi_t=a\}} \middle| X_0 = s \right] \geq D(a) J_{\pi}(s), \quad \forall a \in A, \quad (5.28)$$

where  $D(a)$  denotes the minimum fraction of overall system performance in which action (user)  $a$  should be chosen, with  $0 \leq D(a) \leq 1$  and  $\sum_{a \in A} D(a) \leq 1$ .

Therefore, our goal can be stated as: *find an average-reward-optimal policy  $\pi^*$  subject to the expected average utilitarian fairness constraint.*

### Optimal scheduling policy

The following theorem characterize the optimal policy for our average utilitarian fair constrained MDP.

**Theorem 5.4** *Suppose the system is unichain. Suppose we have a bounded function  $h : S \rightarrow \mathbb{R}$ , a function  $\omega : A \rightarrow \mathbb{R}$ , a constant  $g$ , and a stationary policy  $\pi^*$  such that for  $s \in S$ ,*

1.  $\forall a \in A, \omega(a) \geq 0$ ;

2.  $\forall a \in A, \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \frac{1}{T} \sum_{t=0}^{T-1} r(X_t, \pi_t^*) \mathbf{1}_{\{\pi_t^*=a\}} \middle| X_0 = s \right] \geq D(a) J_{\pi^*}(s)$ ;

3.  $\forall a \in A$ , if  $\lim_{T \rightarrow \infty} E_{\pi^*} \left[ \frac{1}{T} \sum_{t=0}^{T-1} r(X_t, \pi_t^*) \mathbf{1}_{\{\pi_t^* = a\}} \mid X_0 = s \right] > D(a) J_{\pi^*}(s)$ , then  $\omega(a) = 0$ ;

4.

$$g + h(s) = \max_{a \in A} \left\{ (\kappa + \omega(a)) r(s, a) + \sum_{s' \in S} P(s' | s, a) h(s') \right\}, \quad (5.29)$$

where  $\kappa = 1 - \sum_{a \in A} D(a) \omega(a)$ ;

5.  $\pi^*$  is a policy which, for each  $s$ , prescribes an action which maximizes the right-hand side of (5.29):

$$\pi^*(s) = \operatorname{argmax}_{a \in A} \left\{ (\kappa + \omega(a)) r(s, a) + \sum_{s' \in S} P(s' | s, a) h(s') \right\}. \quad (5.30)$$

Then  $\pi^*$  is an average-reward-optimal policy as defined by (5.15) subject to (5.28).

The corresponding optimal average reward is

$$J_{\pi^*}(s) = g, \quad \forall s \in S. \quad (5.31)$$

*Proof:* Let  $\pi$  be a policy satisfying the expected average utilitarian fairness constraint; and let  $H_t = (X_0, \pi_0, \dots, X_{t-1}, \pi_{t-1}, X_t, \pi_t)$  denote the history of the process up to time  $t$ . First, we have

$$E_{\pi} \left\{ \sum_{t=1}^T [h(X_t) - E_{\pi}(h(X_t) | H_{t-1})] \right\} = 0.$$

Also,

$$\begin{aligned}
E_\pi[h(X_t)|H_{t-1}] &= \sum_{s' \in S} h(s')P(s'|X_{t-1}, \pi_{t-1}) \\
&= (\kappa + \omega(\pi_{t-1}))r(X_{t-1}, \pi_{t-1}) + \sum_{s' \in S} h(s')P(s'|X_{t-1}, \pi_{t-1}) \\
&\quad - (\kappa + \omega(\pi_{t-1}))r(X_{t-1}, \pi_{t-1}) \\
&\leq \max_{a \in A} \left\{ (\kappa + \omega(a))r(X_{t-1}, a) + \sum_{s' \in S} P(s'|X_{t-1}, a)h(s') \right\} \\
&\quad - (\kappa + \omega(\pi_{t-1}))r(X_{t-1}, \pi_{t-1}) \\
&= g + h(X_{t-1}) - (\kappa + \omega(\pi_{t-1}))r(X_{t-1}, \pi_{t-1})
\end{aligned}$$

with equality for  $\pi^*$ , since  $\pi^*$  is defined to take the maximizing action. Hence,

$$\begin{aligned}
0 &\geq E_\pi \left\{ \sum_{t=1}^T [h(X_t) - g - h(X_{t-1}) + (\kappa + \omega(\pi_{t-1}))r(X_{t-1}, \pi_{t-1})] \right\} \\
&\Leftrightarrow g \geq E_\pi \frac{h(X_T)}{T} - E_\pi \frac{h(X_0)}{T} + E_\pi \frac{1}{T} \sum_{t=1}^T (\kappa + \omega(\pi_{t-1}))r(X_{t-1}, \pi_{t-1}) \\
&\Leftrightarrow g \geq E_\pi \frac{h(X_T)}{T} - E_\pi \frac{h(X_0)}{T} + E_\pi \frac{1}{T} \sum_{t=1}^T r(X_{t-1}, \pi_{t-1}) \\
&\quad + E_\pi \frac{1}{T} \sum_{t=1}^T \left( \omega(\pi_{t-1}) - \sum_{a \in A} D(a)\omega(a) \right) r(X_{t-1}, \pi_{t-1}).
\end{aligned}$$

Letting  $T \rightarrow \infty$  and using the fact that  $h$  is bounded, we have that

$$\begin{aligned}
g &\geq J_\pi(X_0) + \lim_{T \rightarrow \infty} E_\pi \frac{1}{T} \sum_{t=0}^{T-1} \left( \omega(\pi_{t-1}) - \sum_{a \in A} D(a)\omega(a) \right) r(X_{t-1}, \pi_{t-1}) \\
&\Leftrightarrow g \geq J_\pi(X_0) + \sum_{a \in A} \omega(a) \left( \lim_{T \rightarrow \infty} E_\pi \left[ \frac{1}{T} \sum_{t=0}^{T-1} r(X_t, \pi_t) \mathbf{1}_{\{\pi_t=a\}} \middle| X_0 = s \right] - D(a)J_\pi(s) \right).
\end{aligned} \tag{5.32}$$

Since we know that  $\omega \geq 0$ , and that the policy  $\pi$  satisfies the utilitarian fairness constraints, the second part of (5.32) is greater than or equal to zero. We get

$$g \geq J_\pi(s).$$

With policy  $\pi^*$ , we have

$$\begin{aligned}
 g &= J_{\pi^*}(s) + \sum_{a \in A} u(a) \left( \lim_{T \rightarrow \infty} E_{\pi^*} \left[ \frac{1}{T} \sum_{t=0}^{T-1} r(X_t, \pi_t^*) \mathbf{1}_{\{\pi_t^*=a\}} \middle| X_0 = s \right] - D(a) J_{\pi^*}(s) \right). \\
 &= J_{\pi^*}(s),
 \end{aligned} \tag{5.33}$$

where the second part of (5.33) equals to zero because of condition 3 on  $\omega(a)$ . Hence, the desired result is proven.  $\blacksquare$

## 5.5 Temporal Fair Rollout Algorithm

### 5.5.1 Rollout Algorithm

In the previous sections, we derived optimal policies for the expected total discounted reward and the expected average reward criteria MDP problems with the corresponding temporal fairness and utilitarian fairness constraints. Note that the optimal policies may be obtained in principle by maximizing the right-hand side of (5.5), (5.17), (5.23), or (5.29), but this requires the calculation of the optimal value function in the right-hand side, which for many problems is overwhelming.

The rollout algorithm yields *a one-step lookahead policy, with the optimal value function approximated by the value function of a known base policy  $\pi$* . The base policy  $\pi$  is typically heuristic and suboptimal, which is calculated either analytically or by simulation. The policy thus obtained is called the *rollout policy* based on  $\pi$ .

The salient feature of the rollout algorithm is its *reward-improvement property*: the rollout policy is no worse than the performance of the base policy. In many cases, the rollout policy is substantially better than the base policy. The rollout algorithm can also be viewed as the *policy improvement* step of the *policy iteration* method, which is a primary method for solving infinite horizon MDP problems [81].

### 5.5.2 Temporal Fair Rollout Algorithm

We will extend the idea of rollout to our temporal fair constrained MDPs to propose an efficient approximation method—*temporal fair rollout*. (A similar treatment applies to the utilitarian fairness case, but is omitted here for the sake of brevity.) We use the expected total discounted reward MDP as an example here (a similar approach applies to the expected average reward case).

Suppose that there exists a function  $u : A \rightarrow \mathbb{R}$  satisfying the conditions in Theorem 5.1. Then we have

$$V_{\pi^*}(s) = \max_a \{r(s, a) + u(a) + \alpha E [V_{\pi^*}(s')|s, a]\}, \quad \forall s \in S,$$

where  $E[\cdot|s, a]$  is the conditional expectation given the state  $s$  and action  $a$ . Moreover, an optimal policy is given by

$$\pi^*(s) = \operatorname{argmax}_{a \in A} \{r(s, a) + u(a) + \alpha E [V_{\pi^*}(s')|s, a]\}, \quad s \in S.$$

Applying Theorem 5.1, we can rewrite the optimal policy as:

$$\pi^*(s) = \operatorname{argmax}_{a \in A} \left\{ r(s, a) + u(a) + \alpha E [J_{\pi^*}(s')|s, a] + \alpha \sum_{a \in A} u(a)C(a) \right\}, \quad s \in S.$$

Removing the last constant term, we get

$$\pi^*(s) = \operatorname{argmax}_{a \in A} \{r(s, a) + u(a) + \alpha E [J_{\pi^*}(s')|s, a]\}, \quad s \in S.$$

Instead of calculating the optimal value function directly, we approximate it with the value function of a base policy that also satisfies the discounted temporal fairness requirements. Let  $\pi^b$  be a base policy and  $J_{\pi^b}$  the value function of the policy. Then the temporal fair rollout policy is

$$\pi^{\text{tfr}}(s) = \operatorname{argmax}_{a \in A} \{r(s, a) + u(a) + \alpha E [J_{\pi^b}(s')|s, a]\}, \quad s \in S. \quad (5.34)$$

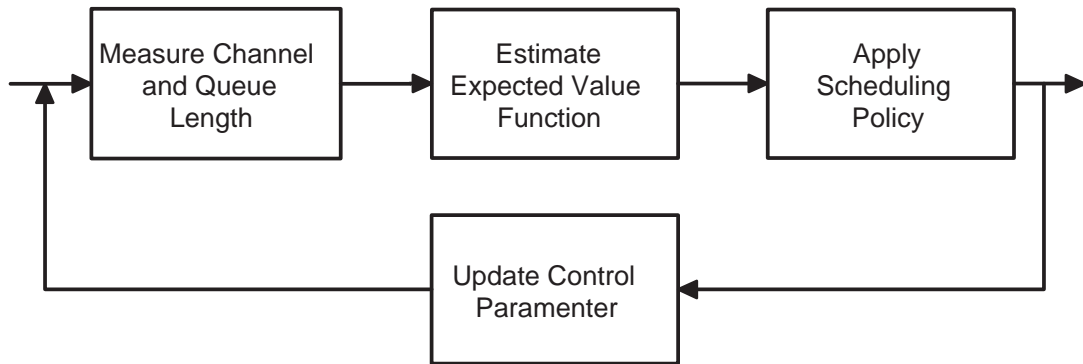
The expected value of the base policy is obtained by Monte Carlo simulation. The selection of the base policy is problem specific. In our experiments, we use the temporal fair opportunistic scheduling policy of [46] as the base policy. We will show by



simulation that the temporal fair rollout policy in fact performs better than the base policy.

## 5.6 Stochastic Approximation for Parameter Estimation

The temporal fair rollout policy (5.34) described in the previous section involves some control parameters  $u(a)$  that need to be estimated. Fig. 5.2 shows a block diagram of a general iterative procedure to estimate these control parameters online. We use a practical *stochastic approximation* technique, similar to the one in [46], to estimate such parameters.



**Figure 5.2:** Block diagram of the scheduling policy with online parameter estimation

We first briefly explain the idea of the stochastic approximation algorithm used here. Suppose we wish to find a zero root of an unknown continuous function  $f(\cdot)$ . If we can evaluate the value of  $f(x)$  at any  $x$ , then we can use the iterative algorithm

$$x^{t+1} = x^t - \beta^t f(x^t),$$

which will converge to a point  $x^*$  such that  $f(x^*) = 0$  as long as the step size  $\beta^t$  is appropriately chosen. Suppose that we cannot have the exact value of  $f(x^t)$  at  $x^t$ ; instead, we only have a noisy observation  $g^t$  of  $f(x^t)$ , i.e.,  $g^t = f(x^t) + e^t$  where  $e^t$  is the observation error (noise) and  $E[e^t] = 0$ . Then the iterative stochastic approximation

algorithm

$$x^{t+1} = x^t - \beta^t g^t,$$

converges to  $x^*$  with probability 1 under appropriate conditions on  $\beta^t$  and  $f$ . We refer readers to [88, 101] for a systematic and rigorous study of stochastic approximation algorithms.

We now use a stochastic approximation algorithm to estimate  $\vec{u}$  (the vector of  $u(a)$ ). Note that we can write  $\vec{u}$  as a root of the equation  $f(\vec{u}) = 0$ , where the  $k$ th component of  $f(\vec{u})$  is given by

$$f_k(\vec{u}) = \lim_{T \rightarrow \infty} E_{\pi^{\text{tfr}}} \left[ \sum_{t=0}^{T-1} \alpha^t \mathbf{1}_{\{\pi_t^{\text{tfr}}=k\}} \middle| X_0 = s \right] - C(k), \quad \forall k \in A.$$

Next, we use a stochastic approximation algorithm to generate a sequence of iterates  $\vec{u}^1, \vec{u}^2, \dots$  that represent estimates of  $\vec{u}$ . Each  $\vec{u}^t$  defines a policy given by

$$\pi^{\text{tfr},t}(s) = \operatorname{argmax}_{a \in A} \{ r(s, a) + u^t(a) + \alpha E [J_{\pi^b}(s') | s, a] \}, \quad \forall s \in S.$$

To construct the stochastic approximation algorithm, we need an estimate  $g^t$  of  $f(\vec{u}^t)$ . Although we cannot obtain  $f(\vec{u}^t)$  directly, we have a noisy observation of its components:

$$g_k^t = \alpha^t \mathbf{1}_{\{\pi^{\text{tfr},t}(s)=k\}} - C(k), \quad \forall k \in A.$$

Hence, we can get a stochastic approximation algorithm of the form

$$u^{t+1}(k) = u^t(k) - \beta^t (\alpha^t \mathbf{1}_{\{\pi^{\text{tfr},t}(s)=k\}} - C(k)),$$

where the step size  $\beta^t$  is appropriately chosen; for example,  $\beta^t = 1/t$ . The initial estimate  $\vec{u}^1$  can be set to  $\vec{0}$  or some value based on the history information. The computation burden above is  $O(K)$  per time slot, where  $K$  is the number of users, which suggests that the algorithm is easy to implement online. Simulations show that with the above stochastic approximation algorithm,  $\vec{u}^t$  converges to  $\vec{u}$  relatively quickly.

## 5.7 Numerical Results

In this section, we present numerical results to illustrate the performance of the proposed temporal fair rollout algorithm. We first describe our simulation setup of a cellular system, as well as the channel model. We then show the simulation results for each scheduling policy using the model.

### 5.7.1 Simulation Setup

We consider the uplink of a single-cell system with 10 mobile users and 1 base station in our simulation. The preset temporal fairness requirements for users 1–10 are  $1/11$ ,  $1/11$ ,  $1/13$ ,  $1/13$ ,  $1/13$ ,  $1/13$ ,  $1/11$ ,  $1/11$ ,  $1/13$ ,  $1/13$ . Note that the temporal fairness requirements are *nonuniform* and the summation of these is less than 1. This gives the system the freedom to assign the remaining fraction of the resource to some “better” users to further improve the system performance.

We assume that the packet arrivals at each queue are *independent Poisson processes*. For simplicity, we assume that we know the maximum arrival rate for each user. We denote the arrival rate and the maximum arrival rate for user  $k$  by  $\lambda_k$  and  $\lambda_k^{\max}$  respectively. We define the normalized-arrival-rate for user  $k$  as  $\lambda_k/\lambda_k^{\max}$ .

We divide the 10 users into five groups based on their heterogeneous arrival rates and mean channel conditions. Specifically, users 1 and 2 have low arrival rates and low mean channel conditions. Users 3 and 4 have high arrival rates and high mean channel conditions. Users 5 and 6 have high arrival rates and moderate mean channel conditions. Users 7 and 8 have low arrival rates and high mean channel conditions. Finally, users 9 and 10 have moderate arrival rates and moderate channel conditions. With this range of heterogeneous user environments, we can study how the arrival rates and channel conditions affect the users’ performances under different scheduling schemes.

For the purpose of comparison, we evaluate six related scheduling policies including temporal fair rollout:

1. Round-robin: A well-known non-opportunistic scheduling policy that schedules users in a predetermined order. At time slot  $t$ , the user  $(t \bmod K + 1)$  is chosen.
2. Greedy: The natural greedy policy always selects the user with maximum possible throughput to transmit at any time. At time slot  $t$ , we choose user according to the following index policy:

$$\operatorname{argmax}_{a \in A} \{ \min(X_a(t), S_a(t)) \}.$$

3. Rollout: At time slot  $t$ , the user chosen is

$$\operatorname{argmax}_{a \in A} \{ r(X_t, a) + \alpha E [J_{\pi^b}(X_{t+1}) | X_t, a] \},$$

where the base policy  $\pi^b$  is the above greedy policy.

4. Opportunistic scheduling-1: In [46], Liu et al. proposed an opportunistic scheduling scheme with temporal fairness constraints for memoryless channels. The policy is

$$\operatorname{argmax}_{a \in A} \{ S_a(t) + v_1(a) \},$$

where  $v_1(a)$  is estimated online via stochastic approximation.

5. Opportunistic scheduling-2: This policy is a variation of the above opportunistic scheduling-1 policy with the consideration of the queue lengths. The policy is

$$\operatorname{argmax}_{a \in A} \{ \min(X_a(t), S_a(t)) + v_2(a) \},$$

where  $v_2(a)$  is also estimated online via stochastic approximation.

6. Temporal fair rollout: We select the above opportunistic scheduling-2 policy as our base policy  $\pi^b$ . It not only satisfies the discounted temporal fairness

constraints, but also takes the queue lengths into account. At time slot  $t$ , the user chosen is

$$\operatorname{argmax}_{a \in A} \{r(X_t, a) + u(a) + \alpha E [J_{\pi^b}(X_{t+1}) | X_t, a]\}.$$

The primary motivation of this paper is to improve wireless resource efficiency by exploiting time-varying channel conditions while also satisfying certain QoS constraints among users. However, it turns out that policies (1)–(3) above violate the temporal fairness constraints (see Figs. 5.4 and 5.6), which means they are *infeasible* for our problem. The reason we include them in our comparison is that either they are very simple and widely used, or they can serve as a performance benchmark/bound.

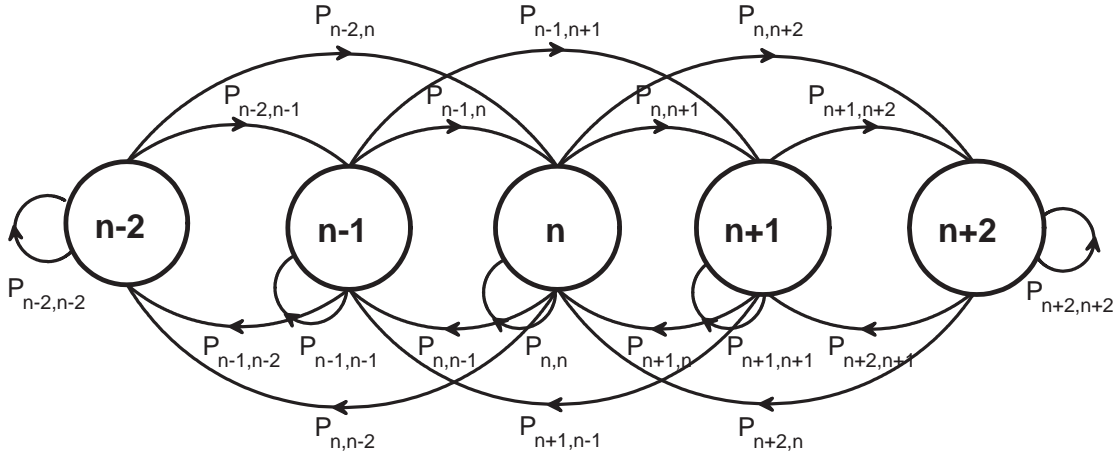
In our evaluation, our focus is not on the effect of the discount factor (which was introduced primarily for analytical tractability). Therefore, in our simulation, we treat  $\alpha$  as a number very close to 1, and replace all normalized discounted sums by finite-horizon averages. For example, in the throughput maximization problem, we calculate the throughput as a time average (without discounting). Similarly, in the delay minimization problem, the delay is calculated as the time average of the queue length. The constraints are also calculated as time averages.

### 5.7.2 Channel Model

The digital cellular radio transmission environment usually consists of a large number of scatterers that result in multiple propagation paths. Associated with each path is a propagation delay and an attenuation factor depending on the obstacles in the path that reflect electromagnetic waves. Multipath fading results in a correlated random process, i.e., a random process with memory. This kind of channel is known as the multipath Rayleigh fading channel.

Finite-State Markov Channel (FSMC) models have been found to be accurate in modeling such channels with memory [79]. The base station uses the pilot channels to estimate the SNRs at the receiver. The SNR is used as a measure of channel

condition here. The study of the FSMC emerges from a two-state Markov channel known as the Gilbert-Elliott channel [25,26]. However in some cases, modeling a radio communication channel as a two-state Gilbert-Elliott channel is not adequate when the channel quality varies dramatically. We need more than two states to capture the channel quality and take advantage of rate adaptation techniques used in cellular networks.



**Figure 5.3:** State transition for Rayleigh fading channel model.

In our simulation, we use an 8-state Markov channel model described in [102] to capture the channel conditions. Fig. 5.3 shows the state transition for the Rayleigh fading channel model. We partition the range of possible SNR values into eight equal intervals where each interval corresponds to a state in the Markov chain. We denote the set of states by  $N = \{0, 1, 2, 3, 4, 5, 6, 7\}$ , where state 0 corresponds to an SNR range of 0 db to 5 db, state 1 to 5 db to 10 db, and so on. The time interval between channel measurements for this model is 1 ms, also called the time granularity of the model. For convenience, we assume that the length of time slot is also 1 ms, so that the granularity of the channel model and the scheduling intervals are consistent. The channel state transition probabilities are given in the Table 5.1 [102].

**Table 5.1:** Channel state transition probabilities

$n \backslash n'$	$n - 2$	$n - 1$	$n$	$n + 1$	$n + 2$
0	-	-	0.677567	0.319746	0.002687
1	-	0.109712	0.676353	0.212175	0.001760
2	0.000739	0.137678	0.671957	0.188237	0.001389
3	0.000885	0.149863	0.670962	0.176897	0.001393
4	0.001099	0.157779	0.671564	0.168380	0
5	0.001102	0.164497	0.670785	0.162652	0.000964
6	0.001252	0.169662	0.670743	0.158343	-
7	0.000248	0.041396	0.958356	-	-

### 5.7.3 Simulation Results

Figs. 5.4 and 5.5 show the performance of the six policies (described in Section 5.7.1) for the throughput maximization problem where we use (5.1) as the reward function. Fig. 5.4 indicates the long-term time fraction allocations of all 10 users under the various scheduling policies for the problem. We plot the 95% confidence intervals for each user. For each user, the rightmost bar shows the minimum time fraction requirement. The remaining six bars represent the time fraction allocated to this user in the six policies evaluated here. We see that only the opportunistic scheduling-1, opportunistic scheduling-2, and temporal fair rollout policies satisfy the minimum temporal fairness requirements for all users. Therefore, these three policies are feasible solutions for our constrained problem.

Fig. 5.5 evaluates the scheduling policies by examining the impact of the arrival traffic on the average system throughput (packets/time slot). We take the normalized-arrival-rate for each user to be the same for every simulation, varying from 0.1 to 1.0. We also plot the 95% confidence intervals for each step. From Fig. 5.5, we see that our *temporal fair rollout* policy outperforms (higher means better) all the other policies except the rollout policy with the greedy base policy. This is not surprising since the latter policy achieves the best overall performance at the cost of unfairness among the users (and is thus not a feasible solution to the problem).

Figs. 5.6 and 5.7 show the performance of the six policies for the delay minimization problem where we use (5.2) as the reward function. Similar to Fig. 5.4, Fig. 5.6 indicates the time-fraction allocations of 10 users in the various scheduling policies for the problem. Also we can see that only the opportunistic scheduling-1, opportunistic scheduling-2, and our *temporal fair rollout* policies are feasible solutions.

Fig. 5.7 evaluates the scheduling policies by examining the impact of the arrival traffic on the average system queue length (packets/time slot). It is evident that the average system queue length increases significantly with increasing arrival traffic. Similar to Fig. 5.5, we also see that our *temporal fair rollout* policy outperforms (lower means better) all the other policies except the rollout policy with the greedy base policy (which, again, is not a feasible solution).

In summary, the simulation results show that our *temporal fair rollout* policy performs significantly better than other policies, including the two opportunistic scheduling policies that also satisfy the temporal fairness requirements, especially for the delay minimization case.

## 5.8 Conclusions

In this chapter, we formulated the opportunistic fair scheduling problem as an MDP with explicit fairness constraints. We derived the dynamic programming optimality equations for MDPs with temporal fairness and utilitarian fairness constraints with two criteria: *infinite horizon expected discounted reward* and *expected average reward*. Based on the optimality equations, we obtained the corresponding optimal scheduling policies for the two criteria. We applied the methods on two common scheduling objectives: throughput maximization and delay minimization problems. Our approach can naturally be extended to fit different objective functions and many other fairness measures. To compute the optimal policies efficiently, we developed



a practically viable approximation algorithm called *temporal fair rollout*. Simulations showed that the algorithm achieve significant performance gains over the other existing opportunistic and non-opportunistic scheduling schemes.

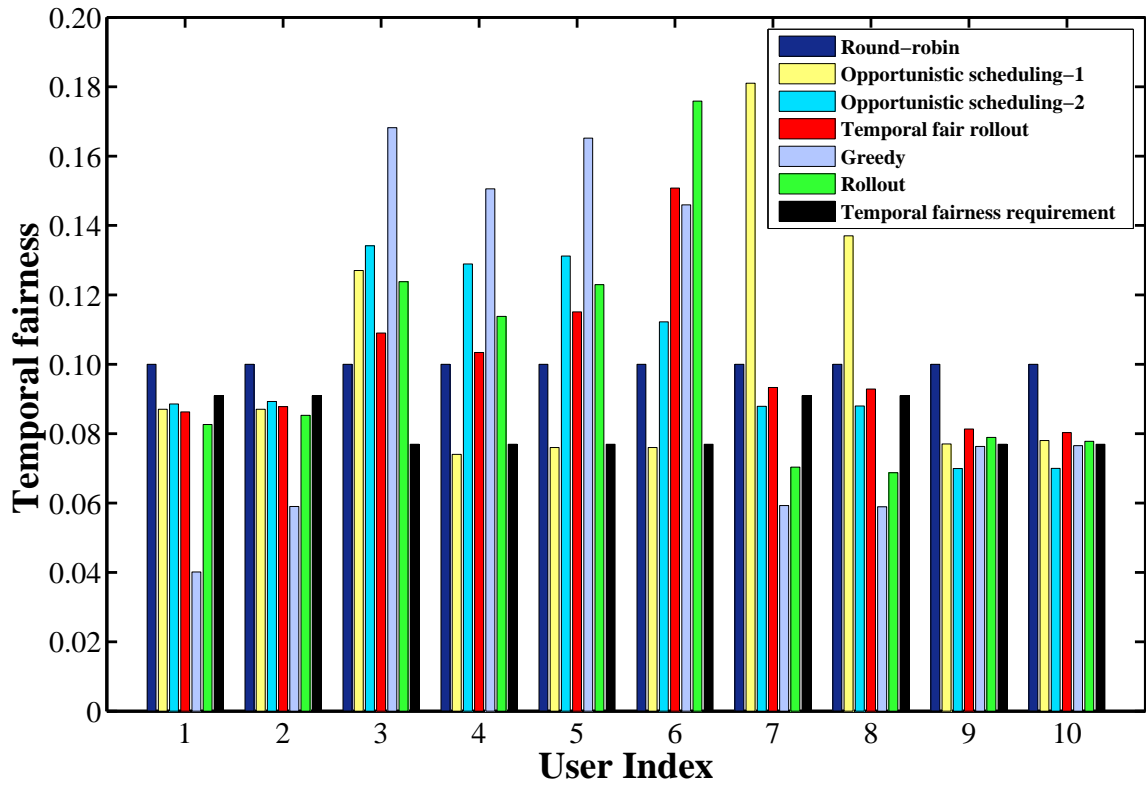


Figure 5.4: Time fraction allocation for throughput maximization problem.

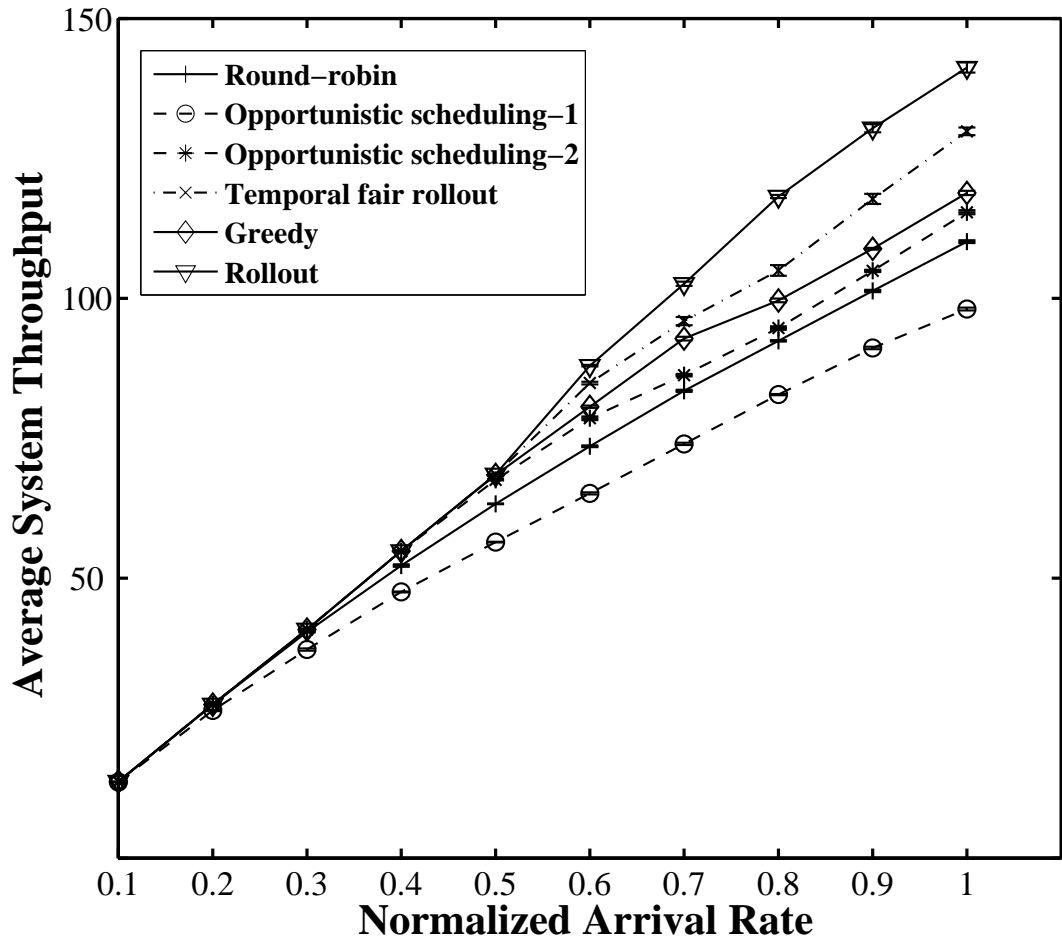


Figure 5.5: Average system throughput vs. normalized arrival rate.

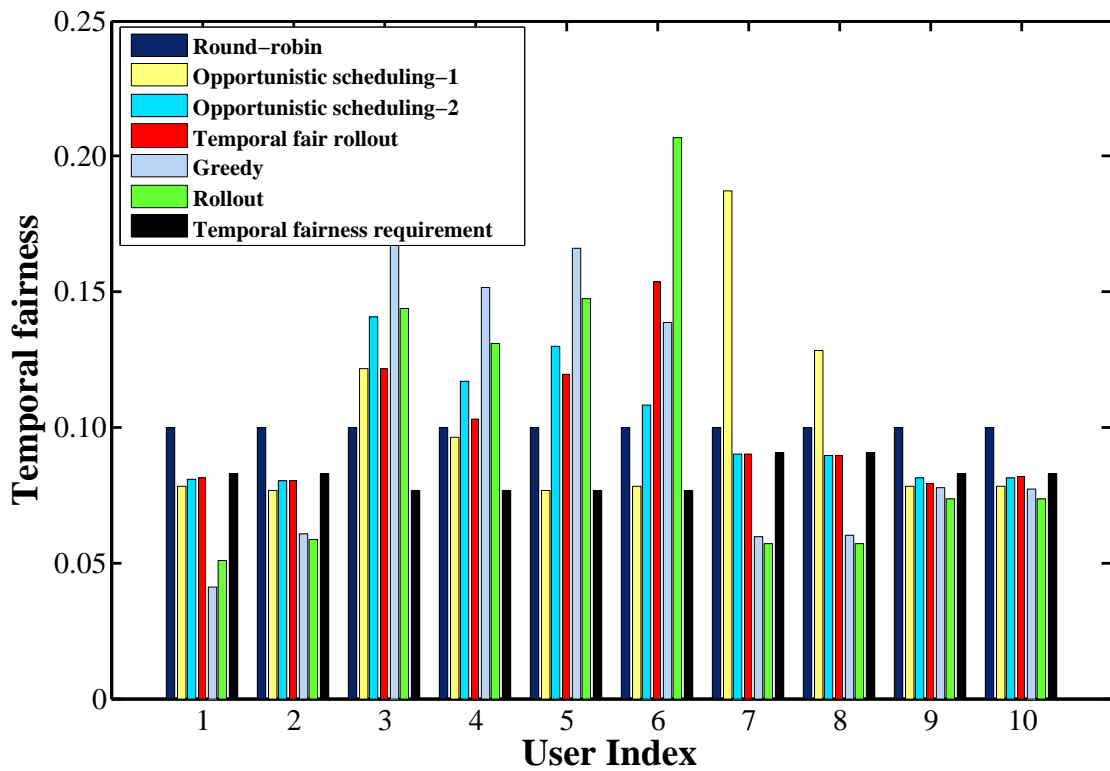


Figure 5.6: Time fraction allocation for delay minimization problem.

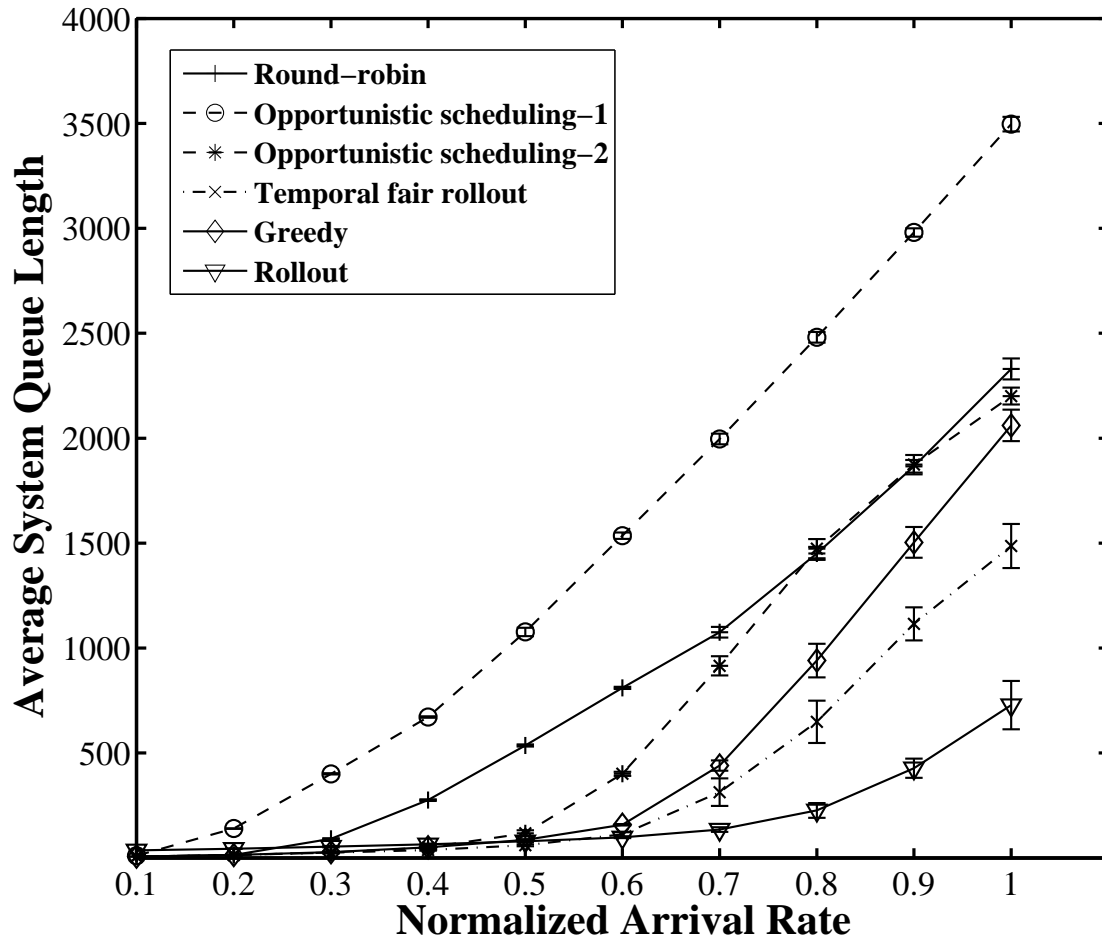


Figure 5.7: Average system queue length vs. normalized arrival rate.

## CHAPTER 6

### CONTRIBUTIONS AND FUTURE WORK

This chapter briefly summarizes the major contributions of this dissertation and outlines proposals for future work.

#### 6.1 Summary of Contributions

To meet the increasing demand for wireless services, especially affordable wireless internet services, wireless spectrum efficiency is becoming increasingly important. In wireless networks, users experience unreliable, location-dependent, and time-varying channel conditions. So dynamic resource allocation for wireless networks has become an important research topic. In this dissertation, we study several resource allocation problems in QoS-aware wireless cellular networks.

The main contributions of this dissertation are as follows:

First, we develop a rigorous framework for *opportunistically* scheduling user transmissions to exploit the time-varying channel conditions in multiuser OFDM systems, which dynamically allocates resource in both temporal and spectral domains. The objective is to maximize the OFDM system performance while satisfying various QoS requirements. Our framework enables us to investigate three categories of scheduling problems involving two fairness requirements (temporal fairness and utilitarian fairness) and a minimum-performance requirement. We provide optimal scheduling

solutions, discuss the advantages and disadvantages of the various scheduling formulations. Our scheduler decides not only *which time-slot*, but also *which subcarrier* to allocate to each user. To implementing these optimal policies involves solving a maximal bipartite matching problem at each scheduling time. To solve this problem efficiently, we propose a modified Hungarian algorithm and a simple suboptimal algorithm. Numerical results demonstrate that our schemes achieve significant improvement in system performance compared with non-opportunistic schemes.

Second, we generalize the work by Liu et al. in two ways. Beginning with the scheduling problems with both minimum and maximum constraints, we derive the corresponding optimal opportunistic scheduling policies for the three long-term QoS/fairness constraints. Then we deal with scheduling problems with multiple type mixed QoS/fairness constraints. We also provide optimal scheduling solutions. Finally, we present a generalized opportunistic scheduling framework to accommodate those scheduling schemes. We show that the structure of the optimal opportunistic scheduling policy is carried over to the problem with general constraints. The generalized optimal framework for opportunistic scheduling provides us an efficient tool to design and analyze the scheduling problems with the heterogeneous users' QoS/fairness constraints over wireless networks.

Third, taking input queues and channel memory into consideration, we reformulate the above transmission scheduling problem as a Markov decision process (MDP) with fairness constraints. We investigate the throughput maximization and the delay minimization problems in this context. We study two categories of fairness constraints, namely temporal fairness and utilitarian fairness. We consider two criteria: infinite horizon expected total discounted reward and expected average reward. We derive and prove explicit dynamic programming equations for the above constrained MDPs, and characterize optimal scheduling policies based on those equations. An attractive feature of our proposed schemes is that they can easily be extended to fit different

objective functions and other fairness measures. Although we only focus on uplink scheduling, the scheme is equally applicable to the downlink case. Furthermore, we develop an efficient approximation method—temporal fair rollout—to reduce the computational cost. Numerical results show that the proposed scheme achieves significant performance improvement for both throughput maximization and delay minimization problems compared with other existing schemes.

## 6.2 Future Work

Resource allocation and scheduling schemes are important in wireless networks, especially to provide high speed packet data and seamless service. There are many unanswered questions and problems yet to be solved in this area.

Opportunistic transmission scheduling is a promising technology to improve spectrum efficiency by exploiting time-varying channel conditions. In order to bring such benefit to future wireless networks, we can extend the opportunistic scheduling idea to efficiently support *multicast traffic*, *partial channel information*, and *ad hoc network*. We can extend the scheduling algorithms to efficiently support non-additive utility functions, which arise naturally in multicast applications. To avoid low multicast throughput caused by serving users with poor channel conditions, we need develop a comprehensive framework to tradeoff multicast throughput and the number of serviced users. Owing to the general difficulty in obtaining precise global channel states, it is necessary to extend the scheduling framework to support partial channel information. We model partial channel information in terms of instantaneous channel state distributions. The objective is to derive scheduling policy based on channel distributions rather than precise channel states. For ad hoc network environments, we assume local schedulers, each managing a subset of wireless terminals. We propose to achieve such objective without extensive information exchange among local schedulers.



We have developed a framework of opportunistic scheduling in multiuser OFDM systems. Our research demonstrated the significant improvement brought by opportunistic scheduling on the effective system capacity of multi-channel systems. Furthermore, we plan to investigate the significant feedback overhead involved in assuming perfect channel-state information feedback in OFDM systems, especially in fast fading channels. Scenarios with relatively small numbers of users in the system will be of practical interest to be explored. That means two or more subcarriers could be available for each user. The effects of finite-length data arrival queues or explicit delay requirement for certain users also will be studied. The application of multiple-channel opportunistic scheduling for MAC layer QoS control in the popular cognitive radio systems will be considered.

## REFERENCES

- [1] ITU, “Measuring the information society: The ict development index,” white paper, International Telecommunication Union, 2009.
- [2] CNN, “Your wireless future,” white paper, CNNMoney.com, May 2006.
- [3] T. S. Rappaport, *Wireless Communications: Principles & Practice*. Prentice Hall, 2nd ed., 2002.
- [4] W. R. Young, “Amps: Introduction, background, and objectives,” *Bell System Technical Journal*, vol. 58, pp. 1–14, January 1979.
- [5] Y. K. Kim and R. Prasad, *4G Roadmap and Emerging Communication Technologies*. Artech House, 2006.
- [6] S. Nanda, K. Balachandran, and S. Kumar, “Adaptation techniques in wireless packet data services,” *IEEE Communications Magazine*, pp. 54–64, January 2000.
- [7] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July, October 1948.
- [8] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near shannon limit error-correcting coding and decoding: Turbo codes,” *Proceedings of the International Conference on Communications*, pp. 1064–1070, 1993.
- [9] R. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 3, pp. 1341–1347, 1995.
- [10] J. Zander, “Performance of optimum transmitter power control in cellular radio systems,” *IEEE Transactions on Vehicular Technology*, vol. 41, pp. 51–62, February 1992.
- [11] D. Goodman and N. Mandayam, “Power control for wireless data,” *IEEE Personal Communications Magazine*, vol. 7, pp. 48–54, April 2000.
- [12] C. Saraydar, N. Mandayam, and D. Goodman, “Efficient power control via pricing in wireless data networks,” *IEEE Transactions in Communications*, vol. 50, pp. 291–303, February 2002.

- [13] R. Prasad, ed., *Universal Wireless Personal Communications*. Boston, MA, USA: Artech House, 1998.
- [14] B. Bing, ed., *Wireless Local Area Networks: The New Wireless Revolution*. New York, USA: John Wiley and Sons, Inc., 2002.
- [15] T. Weiss and F. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *Communications Magazine, IEEE*, vol. 42, no. 3, pp. S8–14, Mar 2004.
- [16] J. Chuang and N. Sollenberger, "Beyond 3g: wideband wireless data access based on ofdm and dynamic packet assignment," *Communications Magazine, IEEE*, vol. 38, no. 7, pp. 78–87, Jul 2000.
- [17] W. C. Jakes, *Microwave Mobile Communications*. New York: John Wiley, 1974.
- [18] G. Stüber, *Principles of Mobile Communication*. Norwell, MA, USA: Kluwer Academic Publishers, 2nd ed., 2001.
- [19] H. Zhang, "Service disciplines for guaranteed performance service in packet switching networks," *Proceedings of IEEE*, vol. 83, pp. 1374–96, October 1995.
- [20] X. Liu, *Opportunistic Scheduling In Wireless Communication Networks*. PhD thesis, Purdue University, December 2002.
- [21] R. M. Mazumdar and C. L.G. Douligieris, "Fairness in network optimal flow control: Optimality of product forms," *IEEE Transactions On Communications*, vol. 39, pp. 775–782, 1991.
- [22] F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.
- [23] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Trans. on Networking*, vol. 1, no. 3, pp. 344–357, Jun. 1993.
- [24] A. Demers, S. Keshav, and S. Shenker, "Analysis and simulation of a fair queueing algorithm," *Proceedings of ACM SIGCOMM'89*, pp. 3–12, 1989.
- [25] E. Gilbert, "Capacity of a burst-noise channel," *Bell Systems Technology Journal*, vol. 39, pp. 1253–1265, Sept. 1960.
- [26] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Systems Technology Journal*, vol. 42, pp. 1977–1997, Sept. 1963.
- [27] S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," *Networking, IEEE/ACM Transactions on*, vol. 7, no. 4, pp. 473–489, Aug 1999.

- [28] T. Ng, I. Stoica, and H. Zhang, "Packet fair queueing algorithms for wireless networks with location-dependent errors," *INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, pp. 1103–1111, 29 Mar-2 Apr 1998.
- [29] S. Shakkottai and R. Srikant, "Scheduling real-time traffic with deadlines over a wireless channel," in *Workshop on Wireless and Mobile Multimedia*, pp. 35–42, 1999.
- [30] Y. Cao and V. Li, "Scheduling algorithms in broadband wireless networks," *Proceedings of the IEEE*, vol. 89, no. 1, pp. 76–87, 2001.
- [31] Qualcomm, "1xEV: 1x EVolution, IS-856 TIA/EIA standard," white paper Rev. 7.2, Qualcomm, USA, 2001.
- [32] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi, "Cdma/hdr: a bandwidth efficient high speed wireless data service for nomadic users," *Communications Magazine, IEEE*, vol. 38, no. 7, pp. 70–77, Jul 2000.
- [33] M. Andrews, S. Borst, F. Dominique, P. Jelenkovic, K. Kumaran, K. Ramakrishnan, and P. Whiting, "Dynamic bandwidth allocation algorithms for high-speed data wireless networks," tech. rep., Bell Labs, 2000.
- [34] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *Communications Magazine, IEEE*, vol. 39, no. 2, pp. 150–154, Feb 2001.
- [35] S. Shakkottai and A. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in hdr," tech. rep., Bell Laboratories, 2000.
- [36] S. Shakkottai and A. Stolyar, "Scheduling for multiple flows sharing a timevarying channel: The exponential rule," *Bell Laboratories Technical Report*, December 2000.
- [37] S. Shakkottai and A. Stolyar, "Scheduling of a shared a time-varying channel: The exponential rule stability," *INFORMS Applied Probability Conference*, July 2001.
- [38] S. C. Borst and P. Whiting, "Dynamic rate control algorithms for CDMA throughput optimization," in *INFOCOM*, pp. 976–985, 2001.
- [39] M. Agarwal and A. Puri, "Base station scheduling of requests with fixed deadlines," *Proceedings of IEEE INFOCOM 2002*, 2002.
- [40] N. Joshi, S. Kadaba, S. Patel, and G. Sundaram, "Downlink scheduling in cdma data networks," *In Proceedings of ACM Mobicom 2000*, 2000.
- [41] K. Kumaran and L. Qian, "Uplink scheduling in CDMA packet-data systems," *in Proceedings of IEEE INFOCOM 2003*, 2003.

- [42] D. Zhang and K. Wasserman, “Transmission schemes for time-varying wireless channels with partial state observations,” *In Proceedings of IEEE INFOCOM 2002*, 2002.
- [43] P. Viswanath, D. Tse, and R. Laroia, “Opportunistic beamforming using dumb antennas,” *Information Theory, IEEE Transactions on*, vol. 48, no. 6, pp. 1277–1294, Jun 2002.
- [44] X. Liu, E. Chong, and N. Shroff, “Opportunistic transmission scheduling with resource-sharing constraints in wireless networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 19, no. 10, pp. 2053–2064, Oct 2001.
- [45] X. Liu, E. Chong, and N. Shroff, “Transmission scheduling for efficient wireless resource utilization with minimum-performance guarantees,” *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, vol. 2, pp. 824–828, 2001.
- [46] X. Liu, E. Chong, and N. Shroff, “A framework for opportunistic scheduling in wireless networks,” *Computer Networks*, vol. 41, no. 4, pp. 451–474, Mar. 2003.
- [47] H. Jiang, W. Zhuang, and X. Shen, “Cross-layer design for resource allocation in 3g wireless networks and beyond,” *IEEE Communications Magazine*, vol. 43, pp. 120–126, 2005.
- [48] H. Zhu, *An Optimization Theoretical Framework for Resource Allocation over Wireless Networks*. PhD thesis, University of Maryland, College Park, 2003.
- [49] M. Cardei, *Resource Efficient Approaches in Wireless Networking*. PhD thesis, University of Minnesota, 2003.
- [50] R. E. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.
- [51] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Belmont, MA: Athena Scientific, 1996.
- [52] D. P. Bertsekas, J. N. Tsitsiklis, and C. Wu, “Rollout algorithms for combinatorial optimization,” *Journal of Heuristics*, vol. 3, pp. 245–262, 1997.
- [53] D. P. Bertsekas and D. A. Castanon, “Rollout algorithms for stochastic scheduling problems,” *Journal of Heuristics*, vol. 5, no. 1, pp. 89–108, 1999.
- [54] C. Derman, *Finite State Markovian Decision Processes*. New York, NY: Academic Press, 1970.
- [55] E. Altman, *Constrained Markov Decision Processes*. Chapman and Hall/CRC, 1998.
- [56] A. B. Piunovskiy, *Optimal Control of Random Sequences in Problems with Constraints*. Dordrecht: Kluwer Academic Publishers, 1997.

- [57] E. A. Feinberg and A. Shwartz, eds., *Handbook of Markov Decision Processes: Methods and Applications*. Boston, MA: Kluwer Academic Publishers, 2002.
- [58] F. J. Beutler and K. W. Ross, “Optimal policies for controlled markov chains with a constraint,” *Journal of Mathematical Analysis and Application*, vol. 112, pp. 236–252, 1985.
- [59] A. S. Eitan Altman, “Time-sharing policies for controlled markov chains,” *OPERATIONS RESEARCH*, vol. 41, no. 6, pp. 1116–1124, 1993.
- [60] K. W. Ross, “Randomized and past-dependent policies for markov decision processes with multiple constraints,” *Operations Research*, vol. 37, no. 3, pp. 474–477, 1989.
- [61] M. Andrews, “A survey of scheduling theory in wireless data networks,” in *Proc. 2005 IMA summer workshop on wireless communications*, 2005.
- [62] R. Knopp and P. Humblet, “Information capacity and power control in single-cell multiuser communications,” *Communications, 1995. ICC '95 Seattle, 'Gateway to Globalization', 1995 IEEE International Conference on*, vol. 1, pp. 331–335, 18-22 Jun 1995.
- [63] S. Parkvall, E. Dahlman, P. Frenger, P. Beming, and M. Persson, “The high speed packet data evolution of wcdma,” in *Proc. IEEE VTC Spring 2001*, vol. 3, pp. 2287–2291, 2001.
- [64] H. Viswanathan and K. Kumaran, “Rate scheduling in multiple antenna downlink wireless systems,” *Communications, IEEE Transactions on*, vol. 53, no. 4, pp. 645–655, April 2005.
- [65] P. Liu, R. Berry, and M. Honig, “Delay-sensitive packet scheduling in wireless networks,” *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, vol. 3, pp. 1627–1632, 16-20 March 2003.
- [66] S. Kulkarni and C. Rosenberg, “Opportunistic scheduling policies for wireless systems with short term fairness constraints,” *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 1, pp. 533–537, 1-5 Dec. 2003.
- [67] A. Jalali, R. Padovani, and R. Pankaj, “Data throughput of cdma-hdr a high efficiency-high data rate personal communication wireless system,” *Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo. 2000 IEEE 51st*, vol. 3, pp. 1854–1858, 2000.
- [68] D. Čabrić and R. Brodersen, “Physical layer design issues unique to cognitive radio systems,” *Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. IEEE 16th International Symposium on*, vol. 2, pp. 759–763, 11-14 Sept. 2005.

- [69] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, “Multiuser ofdm with adaptive subcarrier, bit, and power allocation,” *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 10, pp. 1747–1758, Oct 1999.
- [70] W. Rhee and J. Cioffi, “Increase in capacity of multiuser ofdm system using dynamic subchannel allocation,” *Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo. 2000 IEEE 51st*, vol. 2, pp. 1085–1089, 2000.
- [71] G. Song and Y. Li, “Adaptive resource allocation based on utility optimization in ofdm,” *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 2, pp. 586–590, 1-5 Dec. 2003.
- [72] J. Zhu, B. Bing, Y. G. Li, and J. Xu, “An adaptive subchannel allocation algorithm for ofdm-based wireless home networks,” *Consumer Communications and Networking Conference, 2004. CCNC 2004. First IEEE*, pp. 352–356, 5-8 Jan. 2004.
- [73] Y. Liu and E. Knightly, “Opportunistic fair scheduling over multiple wireless channels,” *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, vol. 2, pp. 1106–1115, 30 March-3 April 2003.
- [74] H. Seo and B. G. Lee, “Proportional-fair power allocation with cdf-based scheduling for fair and efficient multiuser ofdm systems,” *Wireless Communications, IEEE Transactions on*, vol. 5, no. 5, pp. 978–983, May 2006.
- [75] C. Li and X. Wang, “Adaptive opportunistic fair scheduling over multiuser spatial channels,” *Communications, IEEE Transactions on*, vol. 53, no. 10, pp. 1708–1717, Oct. 2005.
- [76] G. Song and Y. Li, “Cross-layer optimization for ofdm wireless networks part i: Theoretical framework & part ii: Algorithm development,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 614–634, Mar. 2005.
- [77] M. Hu, J. Zhang, and J. Sadowsky, “Traffic aided opportunistic scheduling for wireless networks: Algorithms and performance bounds,” *Computer Networks Journal*, vol. 46, no. 4, pp. 505–518, Nov. 2004.
- [78] Z. Shen, J. Andrews, and B. Evans, “Optimal power allocation in multiuser ofdm systems,” *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 1, pp. 337–341, 1-5 Dec. 2003.
- [79] H. S. Wang and N. Moayeri, “Finite-state markov channel—a useful model for radio communication channels,” *Vehicular Technology, IEEE Transactions on*, vol. 44, no. 1, pp. 163–171, Feb 1995.
- [80] M. L. Puterman, *Markov Decision Processes*. New York, NY: John Wiley and Sons, Inc., 1994.

- [81] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 2nd ed., 2001.
- [82] S. M. Ross, *Applied Probability Models with Optimization Applications*. New York, NY: Dover Publications, Inc., 1970.
- [83] A. B. Piunovskiy and X. Mao, “Constrained markovian decision processes: the dynamic programming approach,” *Oper. Res. Lett.*, vol. 27, pp. 119–126, 2000.
- [84] R. C. Chen and G. L. Blankenship, “Dynamic programming equations for discounted constrained stochastic control,” *IEEE Trans. Automat. Contr.*, vol. 49, no. 5, pp. 699–709, May, 2004.
- [85] I. Mitola, J. and J. Maguire, G.Q., “Cognitive radio: making software radios more personal,” *Personal Communications, IEEE [see also IEEE Wireless Communications]*, vol. 6, no. 4, pp. 13–18, Aug 1999.
- [86] C. Peng, H. Zheng, and B. Zhao, “Utilization and fairness in spectrum assignment for opportunistic spectrum access,” *ACM Mobile Networks and Applications (MONET)*, vol. 11, no. 4, pp. 555–576, Aug. 2006.
- [87] H. L. Royden, *Real Analysis*. New York, USA,: The Macmillan Company, 2nd ed., 1968.
- [88] H. J. Kushner and G. G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*. New York: Springer-Verlag, 2nd ed., 2003.
- [89] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. New Jersey, USA: Prentice-Hall, Inc., 1982.
- [90] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network flows: Theory, Algorithms and Applications*. New Jersey, USA,: Prentice-Hall, Inc., 1993.
- [91] H. W. Kuhn, “The hungarian method for the assignment problem,” *Naval Research Logistics Quarterly* 2, pp. 83–97, 1955.
- [92] E. Chong and S. H. Żak, *An Introduction to Optimization*. John Wiley and Sons, Inc., 2nd ed., 2001.
- [93] T. D. Braun, H. J. Siegel, N. Beck, L. L. Bölöni, M. Maheswaran, A. I. Reuther, J. P. Robertson, M. D. Theys, B. Yao, D. Hensgen, and R. F. Freund, “A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems,” *Journal of Parallel and Distributed Computing*, vol. 61, pp. 810–837, 2001.
- [94] E. L. Hahne and R. G. Gallager, “Round robin scheduling for fair flow control in data communication networks,” *Proc. IEEE International Conference on Communications 1986*, pp. 103–107, June 1986.



- [95] L. M. Correia, *Wireless Flexible Personalized Communications, COST 259: European Co-operation in Mobile Radio Resource*. New York, USA: John Wiley and Sons, Inc., 2001.
- [96] R. Chakravorty, S. Katti, I. Pratt, and J. Crowcroft, “Flow aggregation for enhanced tcp over wide area wireless,” *Proc. IEEE INFOCOM*, 2003.
- [97] D. Gesbert and M. Slim-Alouini, “How much feedback is multi-user diversity really worth?,” in *Proc. IEEE International Conference on Communications*, pp. 234–238, 2004.
- [98] F. Floren, O. Edfors, and B. A. Molin, “The effect of feedback quantization on the throughput of a multiuser diversity scheme,” in *Proc. IEEE Global Telecommunication Conference (GLOBECOM03)*, pp. 497–501, December 2003.
- [99] P. Svedman, S. Wilson, J. Cimini, L.J., and B. Ottersten, “A simplified opportunistic feedback and scheduling scheme for ofdm,” in *Proc. IEEE Vehicular Technology Conference 2004*, 2004.
- [100] Y. Al-Harthi, A. Tewfik, and M. S. Alouini, “Multiuser diversity with quantized feedback,” *IEEE Transactions on Wireless Communication*, vol. 6, no. 1, pp. 330–337, 2007.
- [101] H. F. Chen, *Stochastic Approximation and Its Applications*. Dordrecht: Kluwer Academic Publishers, 2002.
- [102] F. Swarts and H. Ferreira, “Markov characterization of channels with soft decision outputs,” *Communications, IEEE Transactions on*, vol. 41, no. 5, pp. 678–682, May 1993.