4-2-2012

Resource Allocation, Scheduling and Feedback Reduction in Multiple Input Multiple Output (MIMO) Orthogonal Frequency-Division Multiplexing (OFDM) Systems

Nansong Wu
Florida International University, nwu001@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

RESOURCE ALLOCATION, SCHEDULING AND FEEDBACK REDUCTION

IN MIMO-OFDM SYSTEMS

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

by

Nansong Wu

2012

To: Dean Amir Mirmiran
       College of Engineering and Computing

This dissertation, written by Nansong Wu, and entitled Resource Allocation, Scheduling and Feedback Reduction in MIMO-OFDM Systems, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Jean H. Andrian

_____
Deng Pan

_____
Yimin Zhu

_____
Kang K. Yen, Major Professor

Date of Defense: April 2, 2012

The dissertation of Nansong Wu is approved.

_____
Dean Amir Mirmiran
College of Engineering and Computing

_____
Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2012

DEDICATION

I dedicate this dissertation to my wife and parents. Without their supports, the

completion of this work would not have been possible.

ACKNOWLEDGMENTS

ABSTRACT OF THE DISSERTATION

RESOURCE ALLOCATION, SCHEDULING AND FEEDBACK REDUCTION

IN MIMO-OFDM SYSTEM

by

Nansong Wu

Florida International University, 2012

Miami, Florida

Professor Kang K. Yen, Major Professor

The number of wireless systems, services, and users are constantly increasing and therefore the bandwidth requirements have become higher. One of the most robust modulations is Orthogonal Frequency-Division Multiplexing (OFDM). It has been considered as an attractive solution for future broadband wireless communications.

This dissertation investigates bit and power allocation, joint resource allocation, user scheduling, and limited feedback problem in multi-user OFDM systems. The following dissertation contributes to improved OFDM systems in the following manner. (1) A low complexity sub-carrier, power, and bit allocation algorithm is proposed. This algorithm has lower computational complexity and results in performance that is comparable to that of the existing algorithms. (2) Variations of the proportional fair scheduling scheme are proposed and analyzed. The proposed scheme improves system throughput and delay time, and achieves higher throughput

without sacrificing fairness which makes it a better scheme in terms of efficiency and

fairness. (3) A DCT feedback compression algorithm based on sorting is proposed.

This algorithm uses sorting to increase the correlation between feedback channel

quality information of frequency selective channels. The feedback overhead of system

is successfully reduced.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER I   INTRODUCTION

1.1 Background

The number of wireless systems, services, and users is constantly increasing and therefore the bandwidth requirements have become higher. One of the most robust modulations is MIMO-OFDM (multiple-input and multiple-output orthogonal frequency-division multiplexing). It has been considered as an attractive solution for future broadband wireless communications [1]. Three important issues in these systems are the resource allocation, user scheduling that is the allocation of the OFDM sub-carriers among the users and feedback compression [2]. By employing a combination of adaptive resource allocation, scheduling and limited feedback techniques, the throughput of a cellular communication system can be substantially increased.

OFDM is a robust modulation technique for transmitting large amounts of data over a radio wave. As shown in Figure 1.1, OFDM divides a broadband channel into narrowband sub-channels, so the signal transmitted is the sum of all sub-carriers. The orthogonal selection of sub-carriers results in a very efficient use of bandwidth, though the frequency components of the signal overlap with each other. OFDM can adapts the transmitted signal to the dynamic channel properties by using appropriate bit loading methods on the property that previously estimated.
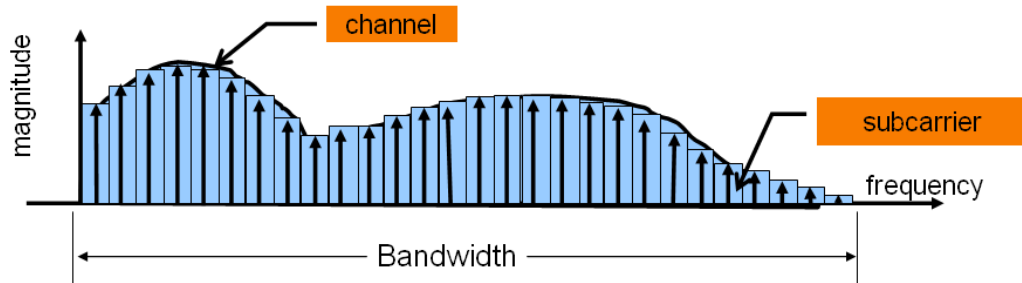
Figure 1.1   Orthogonal frequency-division multiplexing

The advantages of OFDM include: (1) outstanding spectral efficiency; (2) simple implementation based on FFT (Fast Fourier Transform); (3) multipath resistant; (4) less complex design of receiver equalization; and (5) link adaption. These advantages make OFDM an attractive solution for future broadband wireless communications.

MIMO is a technique for space diversity that uses multiple transmission and reception antennas. MIMO can be implemented in various ways either to gain diversity, to combat signal fading, or to obtain capacity gain. MIMO along with OFDM is used by several standards. MIMO-OFDM technique has been adopted by the 4th generation of wireless communication standard. It allows for transmission of different data rate requirements in multimedia service. The most advantage of MIMO-OFDM over single-carrier schemes is the ability to deal with severe channel conditions. Some wireless transmission technologies related to MIMO and OFDM become important research topics. These technologies include adaptive power allocation, multiuser scheduling, and limited feedback. They will be discussed in the following chapters.

2

Long Term Evolution (LTE) is the latest standard of mobile network technology that after GSM (Global System for Mobile Communications), EDGE (Enhanced Data rates for GSM Evolution), UMTS (Universal Mobile Telecommunications System) and HSPA (High Speed Packet Access) network technologies. LTE is designed mainly based on two existing wireless communication technologies-OFDM technology and MIMO technology. Specifically, old single-carrier transmission technology can no longer satisfy the demand of high data rate and high mobility for future mobile communication systems. The wide bandwidth and multipath delay spread of B3G (Beyond third generation) systems cause severe Inter-Symbol Interference (ISI). And design of equalizer in Code Division Multiple Access (CDMA) and other single-carrier systems becomes more difficult. For these reasons, researchers are paying attention on OFDM and multi-carrier transmission technologies [4].

Literatures [5] and [6] show that larger capacity in wireless transmission systems is achieved by adoption of MIMO technology. And under certain conditions, the increase in system capacity is proportional to the minimum number of antennas at transmission and reception ends. Because of the unique advantage in enhancing channel capacity, MIMO technology has been introduced to LTE system framework. Multi-antenna technology extends wireless communication systems to the spatial dimensions. It greatly enhances channel capacity and transmission reliability of wireless communication. Currently, two multi-antenna technologies, space-time

diversity and antenna multiplexer are broadly studied. Space-time block codes and layered space-time technologies like Vertical Bell Labs Layered Space-Time (v-BLAST) are introduced. The MIMO system studied in this dissertation is a v-BLAST system using linear pre-coding technology.

1.2 Motivation of Research

The motivation of this research is to find novel approaches to reduce feedback information and maintain high throughput and fairness in MIMO-OFDM systems.

In OFDM systems, orthogonal sub-carriers can be used for parallel data transmissions. Sub-carriers experience different channel fading rates because of frequency selective fading channel. The system allocates different transmit power and bits to each sub-carrier [6]. The allocation involves discrete optimization. Since for practical implementation, the coding rate and adaptive modulation can have only discrete values [7].

Scheduling scheme is defined as the method of user selection and how to allocate available resources to a selected user. Two crucial key performance indicators are fairness and efficiency. Traditionally, spectral efficiency is evaluated in terms of system throughput, which is sometimes unfair to those users who are far away from a base station or with bad channel conditions. On the other hand, absolute fairness may lead to low bandwidth efficiency. Therefore, an effective trade-off between efficiency and fairness is desired in the communication system.

Resource allocation needs to be done under the premise of known channel information. Channel feedback information brings performance gain to the system, however, it introduces feedback overhead. With the development of communication technology and the popularity of adaptive technology, the feedback overhead reduction has become a problem that cannot be ignored. Limited feedback algorithm is adapted to this need. How to minimize the channel feedback information overhead and ensure system performance is worth researching.

1.3 Significance of Research

In an adaptive MIMO-OFDM system, the receiver feeds back channel state information to the transmitter. Assuming perfect channel state information at the transmitter is unrealistic due to the capacity limitation on the feedback channel and its round-trip delay. It has been shown that even partial channel knowledge at the transmitter can provide significant gain in capacity compared to systems that do not take into account channel state information; however, feedback of detailed channel state information consumes valuable bandwidth on the reverse link. Consequently, there is significant interest in designing effective methods to introduce the amount of feedback of channel state information without significantly penalizing the capacity of the reverse link. Substantial gain in capacity can be achieved in multiple antenna systems when accurate channel state information is available at the transmitter.

By employing a combination of adaptive resource allocation, feedback compression, and scheduling, the throughput of the cellular communication system

can be substantially increased [8]. The throughput in a multiuser system can be maximized if a frame is transmitted exclusively to the best user, and modulation and coding can be adapted to the current channel condition. This scheme, sometimes denoted as multiuser diversity, can be viewed as a means to provide selection diversity between users and thus mitigate channel fading in wireless networks.

1.4 Research Purpose and Difficulties

Wireless communication is one of the fastest growing technologies. With the widely application of computer network and multimedia technology, the demand for high quality multimedia service and high speed data service keeps increasing. In order to meet the growing demand as well as the advanced theoretical research, new technology development for future wireless systems is already underway. As the wireless resources are limited, the keys to the future wireless communication include: multicarrier multi-antenna transmission technology that utilizes the spatial resources, and maximization of spectral utilization and power efficiency. By employing a combination of adaptive resource allocation, feedback compression, and scheduling, the throughput and fairness of the wireless communication system can be substantially increased [8].

In modern wireless networks, adaptive coding and modulation are implemented so that the mobile users and base stations can adapt their transmission rates to the quality of the wireless channel [3]. This adaptation not only increases the spectral efficiency of the wireless links between the base station and the mobile users,

but also can be exploited further by the base station through scheduling. In adaptive

MIMO-OFDM system, a feedback channel must be set up for informing the

transmitter. Feedback can be very rate-demanding especially in the multi-antenna

system, so it is necessary to apply certain compression algorithms to the feedback

channel. In every time slot, all users can have access to the N sub-carriers of OFDM.

However, each frequency slot can be dedicated to only one user. The multiuser

scheduler can select the user based on the channel quality of the users in each

frequency-slot according to feedback channel information. Then the adaptive

modulation chooses a set of suitable modulation parameters based on the feedback

information regarding the full or partial channel state information.

Since the channel SNR values are correlated in both time and frequency, data

compression algorithms exploiting such correlations can substantially reduce the

required feedback rate. There are two categories of compression algorithms, namely

lossless compression, including arithmetic coding, Lempel-Ziv-Welch (LZW) coding,

etc.; and lossy compression, including transform coding, linear predictive coding, etc.

[4]

Several recent literatures have adopted lossless algorithms in MIMO-OFDM

systems. The use of time-frequency correlation compression techniques reduces the

feedback data by 50%. However, an error in decoding propagates non-linearly [5],

and the feedback data rate is still high. Meanwhile, most of the authors assumed

max-SNR scheduling in the adaptive MIMO-OFDM system, in which the scheduling

policy is that the base station transmits exclusively to the user with the highest SNR. It is unnecessary for users with low SNR to send feedback; the probability that such a user is scheduled is very low. This can be exploited to reduce the feedback rate; however, it does not provide any fairness among the users. This scheduler always selects the user with the highest SNR and therefore the highest throughput at each frequency-slot. Without power control, users suffering from bad channel conditions may starve and will not be given a chance to transmit. This gives an unfair resource allocation among the users [2]. Since the max-SNR scheduler does not provide any fairness, other schedulers must be considered in practice. The scheduler may have a large impact on the necessary feedback rate; with fair schedulers, feedback cannot be limited to the user with the highest SNR [4]. In literature [4], several compression methods have also been studied in multicarrier MIMO-OFDM systems, which suggest a combination of a lossy compression scheme and a fair scheduler for a future feedback compressor.

In adaptive MIMO-OFDM systems, Proportional Fair Scheduling (PFS) technique tries to schedule a user whose ratio of instantaneous throughput to its own average throughput over the past window of length is the largest. In this time slot t, the scheduler selects the user with the largest value of that ratio among all users in the system. By adjusting t, the desired tradeoff between fairness and throughput can be achieved [2]. To increase computational efficiency by reducing the complexity and the computation time, a table look-up method may be used. The scheduler may

determine the continuous bandwidth allocation and/or the continuous power distribution iteratively, until a difference between proportional fair capacities corresponding to sequential calculations becomes less than a predetermined number, and/or a predetermined maximal number of iterations are reached [7].

1.5 Structure of Dissertation

The dissertation mainly investigates bit and power allocation for OFDM systems, joint resource allocation, user scheduling, and limited feedback in multiuser OFDM systems. We first propose a multi-carrier bit and power allocation algorithm with low complexity for single-user OFDM systems. A simplified joint resource allocation scheme considering user fairness in multiuser OFDM systems is proposed. Then, we study the limited feedback algorithms and channel correlation in frequency domain and time domain. A DCT compression algorithm based on sorting is proposed. Finally, we study a multiuser scheduling algorithm with low feedback rate. Considering use of limited feedback technology in MIMO systems, the dissertation also proposes a multiuser scheduling scheme based on statistical channel state information.

Chapter 2 presents a survey of studies that have been conducted relating to resource allocation, scheduling and feedback in MIMO-OFDM systems. This includes what previous researchers have accomplished using OFDM and MIMO techniques. In addition, recent trends in MIMI-OFDM research are presented.

The complete dissertation is organized as follows. Chapter 3 mainly studies the adaptive resource allocation techniques in multi-carrier OFDM systems. In a single-user OFDM system, according to user's channel fading characteristics in different sub-carriers, the system can use adaptive power and transmission bit allocation technology to improve system performance. In this chapter, we first analyze the characteristics of optimal power and bit allocation scheme in the single-user OFDM systems. According to the results we further propose a simplified adaptive resource allocation algorithm based on greedy search mechanism. Then we study the resource allocation in multiuser OFDM systems with the constraint of minimum user transmission rate, and design a low-complexity joint sub-carrier, power and bit allocation algorithm.

In Chapter 4, we present and analyze the variety of multiuser scheduling schemes. A simulation on performance is carried out and shows the significant advantages of proportional fair and the proposed variations over the conventional schemes. The proposed schemes improve system throughput and delay time. NAPF achieves higher throughput without sacrificing fairness. It is also a better proposed scheme in terms of efficiency and fairness.

In Chapter 5, we first analyze the limited feedback algorithms and channel correlation in frequency domain and time domain. After introduction of three existing lossless compression algorithms and two lossy compression algorithms, and their application in OFDM systems' feedback compression, a DCT compression algorithm

based on sorting is proposed. Appropriate data block size for this algorithm is analyzed which helps reducing the feedback overhead of the system.

Chapter 6 is the conclusions of this dissertation. Suggestions for further work are also discussed in this chapter.

CHAPTER II   LITERATURE REVIEW

Communication technology is considered one of the fastest growing high technology in the 21st century for its rapid development, and mobile communication is currently the focus of communications technology development. In recent years, with the wide application of computer network and multimedia technology in various fields, the demand for high quality multimedia services and high speed data services is increasing rapidly. To meet the growing demand for advanced theoretical study, research for future mobile communication systems is already underway. The future mobile communication system is known as Beyond 3rd Generation (B3G) system [3], [4], [5]. It strives to establish a unified all IP mobile communications network, which provides seamless access to different networks and services, and supports a higher speed data transmission service.

2.1 Resource Allocation in OFDM Systems

OFDM is one of the common schemes in multi-carrier transmission technology. It divides a broad bandwidth into multiple orthogonal narrow bands for parallel data transmission; therefore compared to common single-carrier systems it provides one more degree of freedom, frequency degree, for optimized resource allocation. In a traditional mobile wireless communication system, communication quality and reliability are usually affected by frequency selective fading channel. While in OFDM systems, each sub-carrier system can be approximately considered as a flat fading channel, taking into account that the bandwidth of sub-carrier is relative

narrow. As a result, during signal transmission, the system can intelligently allocate transmit power and adjust transmission rate for sub-carriers according to their channel quality.

Adaptive power allocation is one of the common optimization schemes of resource allocation. Water-filling algorithm, the most classic scheme out there, is to maximize system capacity [7]. In OFDM systems, adaptive power allocation technique can be extended to time domain and frequency domain. Adaptive modulation is also a common optimization scheme of resource allocation. The concept of adaptive modulation was first proposed by Cavers in 1972 [8]. In 1995, Webb et al. proposed variable-rate QAM (Quadrature Amplitude Modulation) scheme [9] in fading channel to achieve adaptive modulation. Through this approach only one modulator and demodulator are required at the transmitter and receiver respectively, which simplifies system implementation. In 1997, Goldsmith et al. theoretically studied the capacity of adaptive modulation system in fading channels. The authors proposed a power allocation method that combines the power distribution technology with the M-QAM (M-ary QAM) modulation system of adaptive transmission rate. This method achieved satisfactory results [10], [11]. On this basis, researchers begin to study the application of M-QAM adaptive modulation technique in broadband wireless communication systems. In particular, design of adaptive modulation in OFDM system attracts wide attention [12].

In order to make good use of system resources, design of the adaptive transmission scheme combined with power allocation and adaptive modulation (i.e., transmission bit allocation) become a popular research topic in OFDM multi-carrier system. Hughes-Hartogs algorithm [13], based on greedy search algorithm，was the bit and power allocation algorithm initially used in multi-carrier systems. However, increasing number of sub-carriers in multi-carrier system caused high computational complexity of the algorithm. Therefore, Chow's and Fischer's algorithms [14], [15] appeared later. Both algorithms maximized the channel capacity and obtained sub-optimal bit and power allocation. In addition, Fischer's algorithm found closed-form solution of the power and bit allocation, and reduced the complexity of the algorithm. These algorithms were originally designed for Asymmetric Digital Subscriber Line (ADSL) Multi-tone (MT) communication system. With the wide application of OFDM system in mobile wireless communications, systems require lower complex adaptive allocation algorithms. People begin further research and to design adaptive allocation algorithms for fast calculation [16], [17], [18], [19], trying to seek balance between system performance and computational complexity [20].

In multiuser OFDM systems, optimization of power and bit allocation becomes more complicated. The channel state varies from one user to another in the same sub-carrier frequency. Therefore, when multiusers access, assigning appropriate sub-carriers to different users increases spectrum resources utilization, and improves system performance [21], [22]. However, early optimization schemes were for

maximizing system capacity [23], [24], [25]. These schemes obviously ignored the different service requirements and fairness among individual users. In order to address fairness among users during resource allocation, literatures [26], [27], [28], [29] proposed a variety of criteria for fairness measurement. More rational multiuser resource allocation schemes were designed under the guidelines of these criteria. However, these schemes were of high computational complexity. As a result, designing an efficient resource allocation scheme, which takes into account specific user requirement and fairness, becomes one of the topics of this dissertation.

2.2 User Scheduling and Limited Feedback Technology in MIMO Systems

Multi-antenna technology can significantly improve the system's channel capacity [30]. This technology in recent years draws attention from researchers. Multi-antenna transmission technology used in a MIMO system makes it possible to access multiple users simultaneously and to perform parallel transmission, and therefore improves the system's transmission rate. Many literatures showed in-depth research for multiuser MIMO systems. In literature [31], [32], Jindal and Vishwanath et al. theoretically analyzed the channel capacities of multiuser MIMO broadcast channel and multiple-access channel, and gave the dual relationship between them.

In order to get the largest channel capacity in a real multiuser MIMO system, a series of multiuser access and diversity transmission method were proposed. Among them, Dirty Paper Coding (DPC) is a multiuser access scheme proved to help the

system to obtain full channel capacity [33], [34]. However, this scheme not only requires full channel information of all accessed users, but its computational complexity is also very high. Therefore, researchers began to focus on some multiuser transmission schemes of high efficiency and low complexity. One of the most common schemes is the beamforming scheme, which is based on Zero-forcing (ZF) and Block-diagonalization (BD) [35], [36], [37]. However, this scheme usually has restrictions on the number of users that can access the system in each time slot. This limitation is determined by the configured system antennas. Thus, in one cell, the base station usually needs to provide access to a small number of selected users and arranges data transmission within a single time slot. Therefore, how to select the small number of appropriate users from all active users in the system for parallel transmission, i.e., multiuser scheduling, becomes the central issue of related researches [38], [39], [40], [41].

Initially, research on scheduling problems in multiuser MIMO system is based on the assumption that the base station has full channel information of all users [38], [39], [40]. However, if the actual wireless transmission systems use Frequency-Division Duplexing (FDD), the assumption does not hold. Therefore, in order to ensure the base station obtaining the downlink channel state information, first the user side needs to perform channel estimation, and then feeds the estimated channel parameters back to the base station. These procedures require additional system feedback information. Considering the feasibility of system implementation,

the system feedback information should be reduced while ensuring the system performance. Some low-rate feedback based multiuser scheduling schemes began to receive attention. Such scheduling algorithms generally use the following methods:

- All users in one cell feed their channel state information back to the base station. However, in order to reduce the amount of feedback, each user only send partial of the full channel state information [42], [43], [44], [45]. For example, only the magnitude or direction of channel information is sent back.

- First design a certain guideline, and then following the guideline choose some initial users in the cell. Let them feed back full channel state information. According to this information, the base station uses user scheduling algorithms to further pick out the users that have good channel conditions and are suitable for parallel transmissions. Then data transmission is arranged [46], [47], [48].

- Scheduling algorithms only use the statistical information of each user channel for user selection [49], [50], [51], [52]. We will further discuss this scheme in later chapters of the dissertation.

Compared with the previous two methods, the third one requires the least amount of system feedback information. In MIMO system, limited feedback technology is widely used to reduce the amount of feedback information. With the applications of limited feedback MIMO systems, some literatures [45], [47], [53] began to focus on the scheduling in multiuser MIMO system of limited feedback. Limited feedback mechanism is an important step to apply the multiuser MIMO

system related technologies in practice, such as Long Term Evolution (LTE) system. In the limited feedback system, the receiver and transmitter use the same codebook, and the codebook is pre-set in the system. The receiver first quantizes its instantaneous channel state information to channel vector/matrix, finds their corresponding codes in the codebook, and then feeds it back to the transmitter. In this way, system feedback can be effectively reduced by choosing an appropriate size of codebook [53]. In multiuser limited feedback systems, we should not only consider the performance of scheduling algorithm, but also the system performance loss resulting form quantization error of feedback channel. Therefore, design of an improved scheduling algorithm for a limited feedback system is also one of the topics of this dissertation.

After discussion of the multiuser scheduling in limited feedback MIMO systems, the following two chapters of this dissertation focus on optimization of bandwidth allocation in limited feedback systems. In MIMO systems that use hybrid retransmission, literatures [54], [55], [56], [57] suggested designing different precoding for retransmitted users to improve system performance. But there is little literature on how to design limited feedback scheme for retransmitted users. Therefore, this dissertation study this issue in later chapters. We also note that in existing studies on limited feedback MIMO systems, most of the literatures [45], [53], [58], [59] considered all users as homogeneous users, and therefore the number of bits fed back by the each user is the same. However, in the real system, different geographic

locations of users, or different transmission distances to the base station lead to different path loss. Therefore, research on heterogeneous multiuser MIMO systems should consider different average channel SNRs of different users due to path loss. Designing and optimization of limited feedback scheme in the heterogeneous multiuser MIMO systems has also become one topic of recent research.

CHAPTER III   ADAPTIVE RESOURCE ALLOCATION IN OFDM

SYSTEMS

3.1 Introduction

In OFDM systems, orthogonal sub-carriers can be used for parallel data transmissions. Sub-carriers experience different channel fading rates because of frequency selective fading channel. The system allocates different transmit power and bits to each sub-carrier [6]. In order to optimize resource allocation, for sub-carriers in deep fading, the system can temporarily stop using them for transmission. That means no transmit power and symbols are allocated to these sub-carriers.

The bit allocation in OFDM systems involves discrete optimization. A common allocation algorithm is based on greedy search algorithm [14]. In order to obtain a more efficient allocation algorithm, Chow and Leke [15], [17], who change the criterion of bit allocation from integer to non-negative continuous real number, therefore simplify the modeling of bit allocation. They also proposed allocation algorithms of relatively low complexity. However, due to the actual number of bits allocated is from rounded optimization result, performance loss is inevitable. Meanwhile, the number of effective bits allocated is usually determined by the order of modulation method used by the system. Performance loss caused by rounding will increase accordingly with the gap between modulation stages. In order to improve the performance of the algorithm, literature [18], [19] has proposed a bit allocation algorithm based on the iterative method. Simulations show near optimal system

performance can be achieved using this method. Because this method uses a scheme similar to greedy bit allocation, its complexity is higher than that of the previous algorithms. Krongold tries to take into account both complexity of the algorithm and system performance [20]. He proposes a more efficient allocation algorithm based on Lagrange multiplier method. In literature [21], Wyglinski et al. provide a detailed analysis of the trade-off between the complexity of discrete bit allocation algorithm and the algorithm performance. Thus, the main issue for multi-carrier OFDM system is how to design an efficient joint bit power allocation algorithm of low computational complexity [60], [61]. Based on greedy allocation algorithm and its analysis, in this chapter we propose a simplified adaptive bit and power allocation algorithm. This method effectively reduces the computational complexity without scarifying system performance.

With further research, the resource allocation problem in multiuser OFDM system attracts researchers' attention. Introduction of multiuser system increases the difficulty of system design. Unlike the conventional Frequency Division Multiple Access (FDMA) system, there are multiple orthogonal sub-carriers in an OFDM system. The system can allocate sub-carriers to appropriate users according to different user channel states in sub-carriers. The proposed method puts system resource to a more efficient use, and gets both frequency diversity gain and multiuser diversity gain. Literatures [22], [23], [24], [25], [26] proposed a series of adaptive resource allocation algorithms. With the constraint of limited total transmit power,

researchers maximize the total multiuser system channel capacity, and minimize transmit power under the required transmission rate. These problems can be simplified and modeled as a typical optimization problem. Wong et al. proposed a joint sub-carrier, power and bit allocation algorithm based on Lagrange multiplier method [22]. However, the computational complexity of the algorithm is relatively high so that it is difficult to be applied to real systems. Later Wong et al. proposed a simplified allocation algorithm. This simplified algorithm first determines the number of sub-carriers and the corresponding number of transmission bits for each user [24]. In order to obtain better system performance without increasing the algorithm complexity, researchers further proposed a modified algorithm [25]. In this algorithm the sub-carrier allocation and bit and power allocation are treated as two separate processes, so the researchers successfully reduced the computational complexity [26]. Literature [23] thoroughly analyzed the features of optimal sub-carrier and power allocation schemes in multiuser OFDM systems. The results indicate that optimal allocation schemes should meet two conditions:

(1) Sub-carrier should be allocated to the user with the highest channel gain in the system;

(2) Power allocation should satisfy the classic water filling algorithm.

In the above literatures, most of the multiuser resource allocation algorithms either maximize the system capacity or minimize the total transmit power. However, in real systems, due to near-far effect of users in different cells, in long-term, these

allocation algorithms cause the system to allocate resources to users whose positions are close to the base stations, and users in good channels. Thus, the fairness among users cannot be guaranteed. To ensure user fairness and quality of service (QoS) requirements, researchers begin to focus on effective resource allocation schemes that guarantee user fairness. First, literature [27] presents a Max-Min based multiuser OFDM resource allocation algorithm. The goal is to maximize the smallest single user channel capacity. In order to better ensure user fairness, resource allocation algorithm built on proportional fairness was suggested [28]. This method maximizes system capacity; while the user accessed transmission capacity meets the given proportion. To further reduce complexity of the algorithm, Nash game theory is used in the design of a simple allocation algorithm [29]. Zhang and Letaief et al. considered an alternative equity principle [30]. Because each user in the system has a separate transmission rate requirement, they optimize the system's total transmission rate under the premise that minimum transfer rate of all users are ensured. However, the computational complexity of this algorithm is higher. In this chapter we will study the resource allocation problem under minimum transmission rate constraint, and propose an allocation algorithm of low complexity.

3.2 Joint Bit and Power Allocation Algorithm in Single-user OFDM Systems

3.2.1 Single-user OFDM System Model

OFDM systems split serial data stream into multiple slower parallel sub-data streams. The broadband frequency selective channel is divided into N parallel flat

fading channels. N is the number of sub-carriers. Assume g_i is the channel amplitude gain of the ith sub-carrier, and x_i is the transmission symbol of the ith sub-carrier. n_i is Gaussian white noise of the ith sub-carrier with average 0 and variance σ^2. As a result, the signal received by the ith sub-carrier of the system can be expressed as:

$$y_i = g_i \cdot x_i + n_i, \quad i = 1, 2, \cdots, N \tag{3.1}$$

The instantaneous SNR on ith sub-carrier can be expressed as:

$$\gamma_i = \frac{p_i \cdot g_i}{\sigma^2} \tag{3.2}$$

where p_i is the transmission order of the ith sub-carrier.

System required Bit Error Rate (BER) is defined as B_{th}. The order of modulation for ith sub-carrier is m_i. To meet the BER performance requirement, the transmit power of this sub-carrier should be at least:

$$p_i = \frac{S^{m_i} \cdot \sigma^2}{g_i^2} \tag{3.3}$$

where S^{m_i} is the smallest SNR required to get B_{th} by a system using m_i order modulation. These parameters are known constant to the system. As a result, the joint bit and power allocation algorithm in single-user OFDM system can be modeled as follows:

$$\max \left(\sum_{i=1}^{N} m_i \right) \tag{3.4}$$

where modulation order $m_i \in M$, $i = 1, 2, \cdots, N$, and M is a set of possible modulation orders of the system. The transmission order $p_i \geq S^{m_i} \sigma^2 / g_i^2$,

$i = 1, 2, \cdots, N$, and $\sum_{i=1}^{N} p_i \leq P_T$. Here P_T is the total transmit power. Because the bit allocation is a discrete value, we cannot use Lagrange multiplier method directly to solve the problem. Literature [30] proposed a mechanism based on iterative bit allocation algorithm. In the method each step only allocates one bit, and then determines which user get this bit according to the system performance gain acquired by this single bit. However, the method is complex due to the greedy iterative mechanism. In this section, we will analyze the characteristics of the optimal allocation schemes, and propose a modified joint bit and power allocation algorithm to reduce the computational complexity.

3.2.2 Resource Allocation Algorithm Based on Greedy Criterion

3.2.2.1 Analysis of Optimal Allocation Schemes

In order to get a better simplified algorithm, first we need to analyze the characteristics of existing optimal allocation schemes. Equation 3.4 shows that the optimal allocation scheme has the best system power efficiency and the original objective function is equivalent to:

$$\min \ \frac{P_T}{\sum_{i=1}^{N} m_i} \tag{3.5}$$

Thus, we consider two different sub-carriers i and j, and their orders of modulation are m_i and m_j. In order to meet the system's BER requirement, the transmit power p_i and p_j should satisfy Equation 3.3. The total transmit power can be calculated as:

$$p(i,j) = \sigma^2 \cdot \left(\frac{S^{m_i}}{g_i^2} + \frac{S^{m_j}}{g_j^2} \right) \tag{3.6}$$

We exchange the two sub-carriers' modulation. The total transmit power $p'(i,j) = \sigma^2 \cdot \left(\frac{S^{m_j}}{g_i^2} + \frac{S^{m_i}}{g_j^2} \right)$. Comparing $p'(i,j)$ with the total transmit power under the former scheme, we can get the following relationship:

$$p(i,j) - p'(i,j) = \sigma^2 \cdot \left(\frac{S^{m_i}}{g_i^2} - \frac{S^{m_j}}{g_i^2} + \frac{S^{m_j}}{g_j^2} - \frac{S^{m_i}}{g_j^2} \right) = \sigma^2 \cdot \frac{(S^{m_i} - S^{m_j})(g_j^2 - g_i^2)}{g_i^2 g_j^2} \tag{3.7}$$

Without loss of generality, assume $g_i^2 \geq g_j^2$, $m_i \geq m_j$, and $S^{m_i} \geq S^{m_j}$. Then the value of Equation 3.7 is negative. This explains less transmit power is needed when a sub-carrier of high channel gain is used for transmission. Based on this conclusion, the optimal bit allocation scheme will have the structure shown in Figure 3.1.
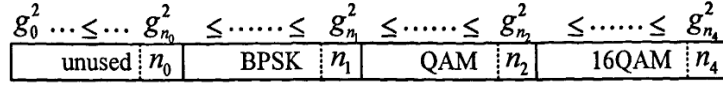
Figure 3.1   Optimal bit allocation scheme structure

In this allocation scheme, power efficiency $\Delta p_i^{m_i}$ is defined as the amount of increased transmit power to transmit one more bit on the ith sub-carrier of m_i modulation.

$$\Delta p_i^{m_i} = \frac{\sigma^2 S^{m_i}}{m_i g_i^2} \tag{3.8}$$

Assume that the alternative bit allocation scheme allocates one more bit to jth sub-carrier with m_j modulation. Then the increased transmit power of the alternative scheme is as follows:

$$\Delta p_i^{m_i} - \Delta p_j^{m_j} = \sigma^2 \cdot \left(\frac{t_{m_i}}{g_i^2} - \frac{t_{m_j}}{g_j^2} \right) \qquad (3.9)$$

Here $t_{m_i} = S^{m_i} / m_i$ is defined as a multiplication factor related to modulation m_i. When the modulation is given, the corresponding S^{m_i} is fixed. Then t_{m_i} is a constant. Equation 3.9 shows if we want the first scheme to have better power efficiency, it must meet the following condition:

$$\Delta p_i^{m_i} \leq \Delta p_j^{m_j} \Leftrightarrow \frac{g_i^2}{g_j^2} \geq \frac{t_{m_i}}{t_{m_j}} \qquad (3.10)$$

We can compare two different allocation scenarios. First, compare the required transmit power of applying different modulation orders on the same sub-carrier (let $g_i^2 = g_j^2$):

$$\Delta p_i^{m_i} \leq \Delta p_j^{m_j} \Leftrightarrow t_{m_i} \leq t_{m_j} \qquad (3.11)$$

This equation shows that bits allocated to the scheme with smaller multiplicative factor can achieve better power efficiency. Usually the higher-order modulation relates to a larger multiplier factor. Second, bits are allocated to different sub-carriers of the same modulation (let $t_{m_i} = t_{m_j}$):

$$\Delta p_i^{m_i} \leq \Delta p_j^{m_j} \Leftrightarrow g_i^2 \geq g_j^2 \qquad (3.12)$$

Equation 3.12 shows that bits allocated to the sub-carrier of higher channel gain leads to better power efficiency. Based on the above analysis, we propose a simplified greedy allocation algorithm (S-GAL) in this section. In each iteration, the algorithm starts with choosing sub-carriers with high channel gain from the sub-carriers with

same modulation, and then choose a suitable sub-carrier and allocate transmit bit to it.

In Figure 3.1, for example, in each iteration the algorithm will select a sub-carrier from n_0, n_1, n_2 and n_4, without the need for comparison of all sub-carriers.

3.2.2.2 Simplified Greedy Allocation Algorithm (S-GAL)

First we propose a simplified sub-carrier selection algorithm. Assume m_i and m_j are the modulation schemes for two sub-carriers i and j. m' is the modulation of nearest high order to modulation m. The increased transmit power required by increasing sub-carrier modulation to m'_i can be calculated as

$$\Delta p_i^I = \frac{1}{m'_i - m_i} \cdot \left(p_i^{m'_i} - p_i^{m_i} \right) = \frac{\sigma^2 (S^{m'_i} - S^{m_i})}{(m'_i - m_i) g_i^2} \tag{3.13}$$

Δp_j^I can be calculate using the same method. Comparing these two schemes, we get the following necessary and sufficient condition:

$$\Delta p_i^I \leq \Delta p_j^I \Leftrightarrow \frac{g_i^2}{g_j^2} \geq \frac{\beta(m_i, m'_i)}{\beta(m_j, m'_j)} \tag{3.14}$$

where $\beta(m, m') = (S^{m'} - S^m)/(m' - m)$. It is a constant under the modulation scheme (m, m'). These parameters can be pre-calculated and saved as a table, so comparison process in the algorithm can be simplified. Algorithm 3.1 shows the flow of simplified greedy allocation algorithm. In the algorithm, i^m represents the sub-carrier of the highest channel gain from a set of sub-carriers of m-order modulation. B represents the highest modulation order from set M of all available modulations.

The algorithm sorts all sub-carriers at the beginning. So in each iteration, the i^m needed can be obtained directly. The computational complexity of sorting is $O(N \log_2 N)$.

Algorithm 3.1 Simplified greedy allocation algorithm (S-GAL)

Step 1: In initialization all the sub-carrier is allocated with 0 bit, and sort sub-carriers according to the ascending order of g_i.

Step 2: All sub-carriers are grouped by modulation scheme. Let G^m denote the set of sub-carriers that use modulation scheme m. In initial state, $G^0 = \{1, 2, \cdots, N\}$, $G^m = \phi$ for $1 \le m \le B$, where ϕ is an empty set; B denotes the largest modulation order in set M of all available modulation schemes. Assume $P_t = 0$.

Step 3: Choose the sub-carrier with the largest channel gain from G^m, and denote it i^m.

Step 4: Choose the sub-carrier with best condition from $i^m (m = 0, 1, \cdots, B)$ according to Equation 3.14, and denote it i^*. Assume i^* belongs to G^{m*}.

Step 5: Calculate the transmit power increment $\Delta p_{i^*}^l$ and the total transmit power P_t. If $P_t + \Delta p_{i^*}^l > P_T$, the algorithm is terminated. Otherwise, go to the next step.

Step 6: Delete sub-carrier i^* from G^{m*}, and assign $P_t = P_t + \Delta p_{i^*}^l$. Return to Step 3.
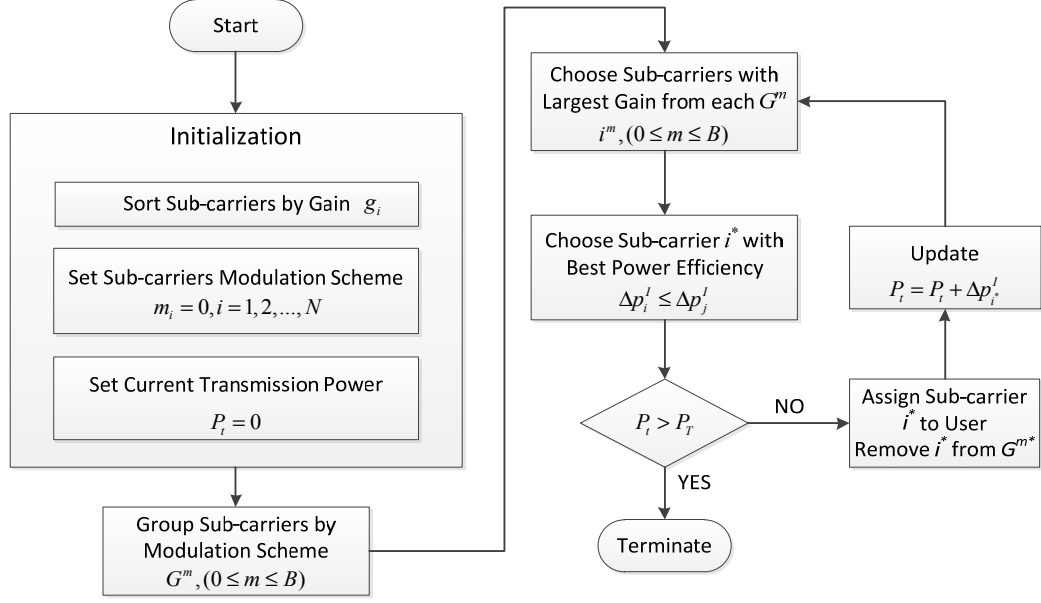
Figure 3.2   Flowchart of simplified greedy allocation algorithm (S-GAL)

3.2.2.3 Modified Greedy Allocation Algorithm (M-GAL)

Though the above simplified algorithm effectively reduces the computational complexity in iterations, the convergence rate is not improved. We further assign an appropriate initial value to speed up the convergence rate of the algorithm. The initial values in S-GAL algorithm are all zero. Differently, we assume initial power allocations on all sub-carriers are equal in M-GAL, and then we determine the initial allocation scheme. A detailed description of the modified algorithm is in Algorithm 3.2.

This algorithm not only effectively reduces the amount of computation of iterations in the original greedy algorithm, but also speeds up the convergence rate by setting an appropriate initial value. The following simulations show that its performance is almost the same as that of the original iterative algorithm.

30

Algorithm 3.2 Modified Greedy Allocation Algorithm (M-GAL)

Step 1: Sort sub-carriers according to the ascending order of g_i.

Step 2: In initial allocation for each sub-carrier: if $S^m \leq g_i^2 / \sigma^2 \leq S^{m'}$, then $m_i = m$. And calculate the least transmit power of each sub-carrier by Equation 2.3.

Step 3: Calculate current total transmit power $P_t = \sum_{i=1}^{N} p_i$.

Step 4: According to the initial allocation, group sub-carriers into sets $G^m, (0 \leq m \leq B)$, choose the sub-carrier with the largest channel gain from G^m, and denote it i^m. Gi^*

Step 5: Choose the sub-carrier i^* from set of $i^m, (0 \leq m \leq B)$ according to Equation 2.14.

Step 6: Calculate the total transmit power P_t. If $P_t > P_T$, terminated. Otherwise, next step.

Step 7: Delete sub-carrier i^* from G^{m*}, choose the sub-carrier i^{m*} with the largest channel gain from the updated G^m, and re-calculate the transmit power $P_t = P_t + \Delta p_{i^*}^l$. Return to Step 5.
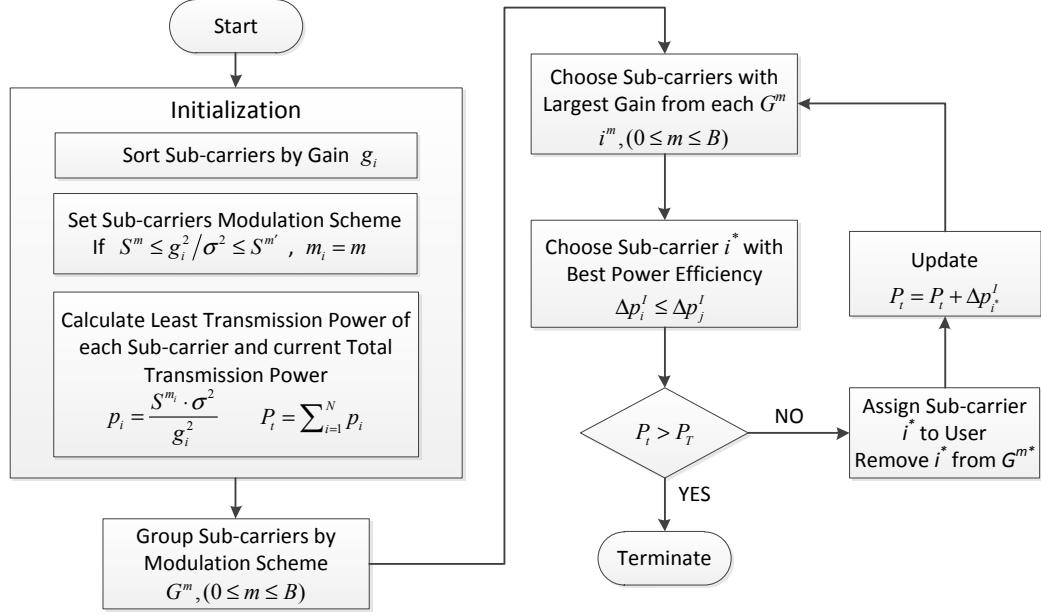
Figure 3.3   Flowchart of modified Greedy Allocation Algorithm (M-GAL)

3.2.3 Analysis of Algorithm Complexity and Simulation

In order to compare the performance of our proposed algorithm with the existing resource allocation algorithms, we consider an OFDM system with 256 sub-carriers. BER B_{th} equals to 10^{-4}. The wireless channel is six-tap multipath Rayleigh channel, and the Power Delay Profile (PDP) on each path decays exponentially. We assume the available M-QAM modulation set is $M = \{0, 1, 2, 3, 4, 5, 6\}$ where 0 represents no bit is allocated for transmission. Under this system configuration, we can use the expression of BER performance of the M-QAM modulation system given by literature [12]. Meanwhile, in order to consider the impact of encoding and decoding in a real system, we introduce a SNR correction factor Γ which is calculated from $\Gamma = -\ln(5B_{th})/1.5$. [12]
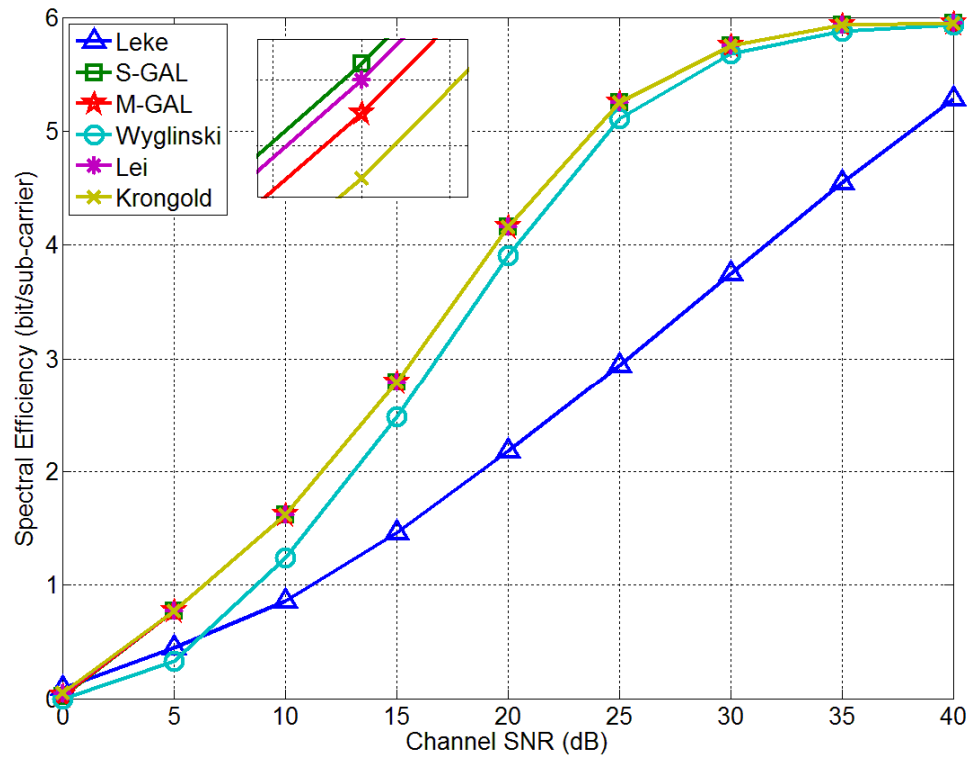
Figure 3.4   Performance comparison of resource allocation algorithms in single-user

OFDM systems

Figure 3.4 compares the throughput of the different power and bit allocation algorithms. The algorithm used by Wyglinski in this comparison assumes equal power allocation. It can be seen from the figure that the simplified algorithm S-GAL shows very close throughput to that of the existing algorithms.

The performance of the further improved M-GAL is also very close to that of S-GAL and Lei algorithms. As an early allocation algorithm, Leke's throughput performance is relatively poor.

The performance of the proposed algorithm is very close to the optimal allocation algorithms, so it is necessary to compare their computational complexity. Table 3.1 shows a quantitative comparison of the complexity of these algorithms. Variable I (I_{Leke}, I_{Lei}, I_K, I_W, I_S, I_M) in this table represents the required number of iterations of different algorithms. Figure 3.5 shows under different channel SNR, the normalized number of iterations required by different algorithms with respect to Lei's. We can use this figure to give the value of I in Table 3.1.

Table 3.1   Quantitative comparison of algorithms complexity

Algorithm	Computational Complexity
Leke's [15]	$O(N \log_2 N + (5I_{Leke} + 11)N)$
S-GAL	$O(N \log_2 N + I_S(5B + 4))$
Krongold's [18]	$O(I_K(B + 4)N)$
Wyglinski's [19]	$O(N \log_2 N + (2B + 7 + I_W(B + 2))N)$
Lei's [17]	$O(I_{Lei}(3B - 4)N)$
M-GAL	$O(N \log_2 N + 2N + I_M(5B + 4))$
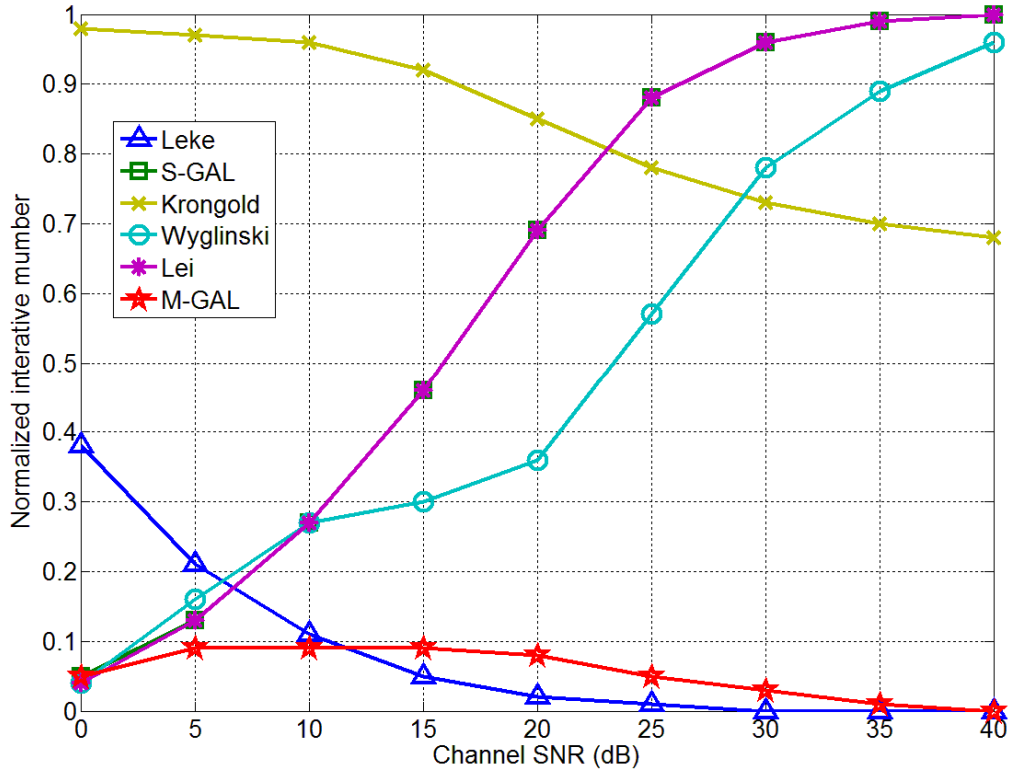
Figure 3.5   Statistic of required number of iterations of different resource allocation

algorithms in single-user OFDM systems

We further compare the complexity of the algorithms in a real scenario. Taking

N=256 and B=6 for example, the number of iterations of each algorithm is calculated

in accordance with the value of I at SNR=20dB in Figure 3.5. So the S-GAL

algorithm needs 38,290 calculations, and Lei's algorithm requires about 3.8×10^6

calculations. We can also get the results that Leke's, Krongold's and Wyglinski's

algorithms need 12,381, 25,561 and 30,939 calculations, respectively. The proposed

M-GAL algorithm needs the least number of calculations, which is 7,133 times. In

conclusion, compared with traditional algorithms, the M-GAL algorithm effectively

increases the convergence rate as a result of the appropriate initial value. It also

obtains similar throughput performance to optimal allocation algorithm as shown in

Figure 3.4 with less computation.

3.3 Joint Sub-carrier, Bit and Power Allocation Algorithm in Multiuser OFDM

System

3.3.1 Multiuser OFDM System and Resource Allocation Model

Multiuser OFDM systems have greater freedom in the design of resource

allocation compared to single-user OFDM systems. In this section, we consider an

OFDM system of K users and N sub-carriers. Similar to the previous section, we

assume that the channel information is completely known to the transmitter and

receiver of the system. At the base station end, the system allocates sub-carriers to

users with good channel conditions depending on their channel quality information,

therefore the system performance is improved, and the transmit power is reduced.

Figure 3.6 is a system diagram of the downlink of a multiuser OFDM system using

adaptive resource allocation. The base station arranges for users to use appropriate

sub-carriers for data transmission according to each user's channel quality

information on sub-carrier. The base station also allocates appropriate power and

transmission bits to these users. Meanwhile, in order to discuss multiple users and the

power and bit allocation scheme on their sub-carriers, we introduce variables $m_{k,n}$ and

$p_{k,n}$, represent transmit bit and power of *mth* sub-carrier of *kth* user, respectively. The

number of transmit bit on a single sub-carrier is usually limited by the available

system modulation, therefore the value of $m_{k,n}$ should be selected from the modulation set M.
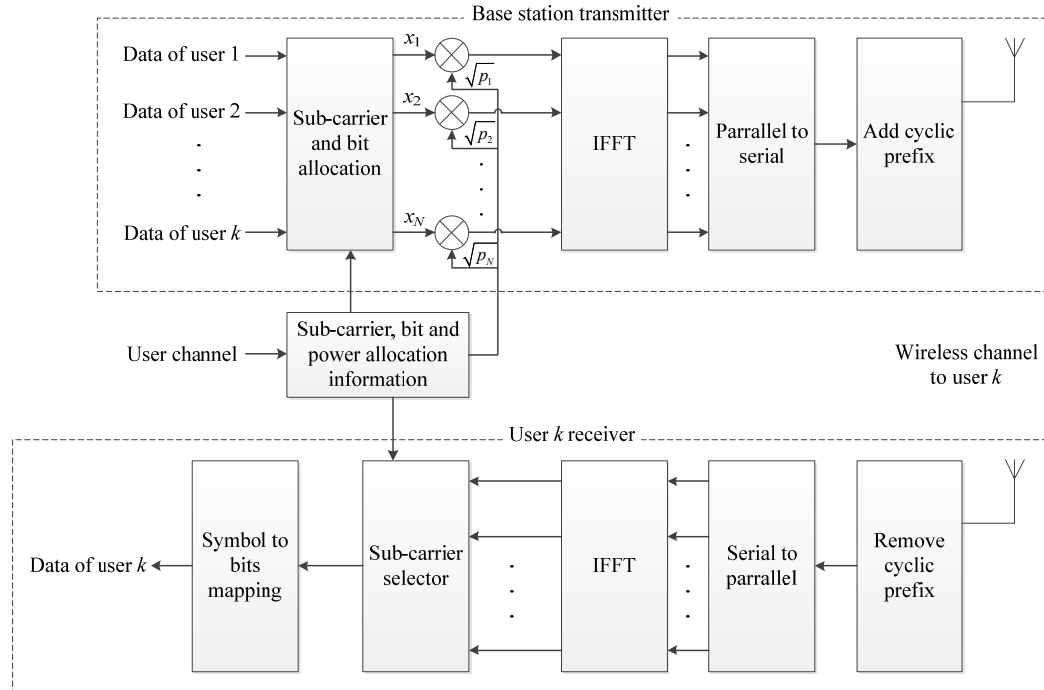
Figure 3.6   Adaptive resource allocation model for downlink of multiuser OFDM

systems

In this section, we analyze optimization of sub-carrier, power and bit allocation schemes to maximize the system throughput. It is clear that by simply maximizing the throughput of the system, the system requirements cannot be met. Thus we add a constraint to the optimization: each user's quality of service (QoS) requirement should be satisfied when maximizing the system throughput.

Each user's QoS requirement is measured by two indicators, the minimum user transmission rate requirement (minimum rate constraint) and the BER requirement of user data transmission (target BER constraint). To simplify the analysis, we assume

that all users have the same target BER constraint. The system takes into consideration of the user's received SNR during power and sub-carrier allocation, and tries to meet the user's BER requirement by ensuring a certain received SNR. SNR is a convex function of BER and $c_{k,n}$, where $c_{k,n}$ is the received information in bits per symbol on the *nth* sub-carrier of user *k*. [24] When user channel conditions are known, in order to ensure a lowest BER on the *nth* sub-carrier, the minimum required SNR should be:

$$SNR = f(BER, c_{k,n}) \tag{3.15}$$

In this system, BER constraints are the same for all users, so the constraints can be considered as system-related constants. Then equation 3.15 can be simplified to $SNR = f(c_{k,n})$.

Suppose each user channel has white Gaussian noise. In order to ensure the user's target BER, the transmit power that the system allocates for this sub-carrier should meet the following criterion:

$$p_{k,n} = \frac{f_{k,n}(c_{k,n})}{\alpha_{k,n}^2} \tag{3.16}$$

where $\alpha_{k,n}^2 = g_{k,n}^2 / \sigma^2$. g is the channel gain of the sub-carrier. σ^2 is the white Gaussian noise. Thus, resource allocation in multiuser OFDM for maximizing the system transmission bit can be modeled as an optimization problem in the following standard form:

$$\min_{\rho_{k,n},c_{k,n}} \left(-\sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \right)$$

subject to $\quad c_{k,n} \in \{0,1,2,\cdots,M\}$

$$\rho_{k,n} \in \{0,1\}$$

$$\sum_{k=1}^{K} \rho_{k,n} - 1 = 0$$

$$\sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} \frac{f(c_{k,n})}{\alpha_{k,n}^2} - P_T \leq 0$$

$$R_k - \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \leq 0$$

$for\ all \quad k \in \{0,1,2,\cdots,K\}, \quad n \in \{0,1,2,\cdots,N\}$ \hfill (3.17)

where $\{0,1,2,\cdots,M\}$ is a set of all possible values for $c_{k,n}$. M is the maximum

number of transmission bits/OFDM symbol that can be transmitted by each subcarrier.

$\rho_{k,n}$ represents whether the *nth* sub-carrier is assigned to the *kth* user for data

transmission.

Literature [23] has proved that in multiuser OFDM system, the optimal

sub-carrier allocation scheme only assigns one sub-carrier to an individual user in one

transmission time slot. Therefore, $\rho_{k,n}$ must satisfy the following relationship: if the

nth sub-carrier is assigned to the k^*th user, then $\rho_{k^*,n} = 1$; for all other $k \neq k^*$ users, the

corresponding $\rho_{k,n} = 0$. Since each sub-carrier can only be allocated to one user,

$\sum_{k=1}^{K} \rho_{k,n} - 1 = 0$. P_T represents the constraint of total system transmit power. The

constraint $\sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} \frac{f(c_{k,n})}{\alpha_{k,n}^2} - P_T \leq 0$ means the sum of allocated power must be equal

or less than the total system transmit power. R_k represents the minimum transmission

rate required by user k. $R_k - \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \leq 0$ indicates that the allocated transmission

rate of user k satisfies the minimum requirement.

3.3.2 Design of Adaptive Resource Allocation Scheme

For the resource allocation problem described in Equation 3.17, we first use

convex analysis and optimization theory, and give the conditions of the solution. The

optimization problem constraints are simplified. According to these conditions,

resource allocation schemes are analyzed for systems with (at least one user's $R_k \neq 0$)

and without (for all user k, $R_k = 0$) minimum transmission rate constraint, respectively.

3.3.2.1 Analysis of Resource Allocation Optimization Problem

In the optimization model shown in (3.17), both non-linear constraints and

integer constraints (constraints for $\rho_{k,n}$) are involved. Thus it is difficult to solve the

problem directly. So we temperately eliminates $\rho_{k,n} \in \{0,1\}$ in the following

discussion.

To simplify the problem, we first consider the system of equal power

distribution. Simplifying procedures taken in this section are mainly based on the

results given by literatures [23] and [30]: "In multiuser OFDM systems, compared

with adaptive resource allocation scheme, the system performance and throughput can

hardly be deteriorated by using equal power allocation scheme." Therefore, power

allocation of each user in each sub-carrier using equal power distribution is as

follows:

$$p_{k,n} = \frac{P_T}{N}$$

(3.18)

With Equation 3.16, we can further calculate the bits allocated to the corresponding user's sub-carrier. Therefore

$$c_{k,n} = f^{-1}(\alpha_{k,n}^2 \frac{P_T}{N})$$

(3.19)

Optimization problem in Equation 3.17 can be simplified to a optimization problem regarding $\rho_{k,n}$.

$$\min_{\rho_{k,n}} \left[-\sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \right]$$

subject to $\sum_{k=1}^{K} \rho_{k,n} - 1 = 0$

$$R_k - \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \leq 0$$

(3.20)

By using standard optimization techniques in book [17], we obtain Lagrange function of this optimization problem:

$$L = -\sum_{k=1}^{K} \sum_{n=1}^{N} \rho_{k,n} c_{k,n} + \sum_{n=1}^{N} \lambda_n \left(\sum_{k=1}^{K} \rho_{k,n} - 1 \right) + \sum_{k=1}^{K} \mu_k \left(R_k - \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \right)$$

(3.21)

where λ_n and μ_k are non-negative Lagrange multiplier constraint and minimum transmission constraint for $\rho_{k,n}$, respectively.

Taking the first derivative of L with respect to $\rho_{k,n}$, we obtain the necessary conditions for the optimal solution $\rho_{k,n}^*$. If $\rho_{k,n}^* \neq 0$, we have

$$\left. \frac{dL}{d\rho_{k,n}} \right|_{\rho_{k,n}=\rho_{k,n}^*} = -c_{k,n} - \mu_k c_{k,n} + \lambda_n$$

$$= \lambda_n - (1+\mu_k)c_{k,n} \tag{3.22}$$

If $\rho_{k,n}^* \in (0,1)$, $\left. \frac{dL}{d\rho_{k,n}} \right|_{\rho_{k,n}=\rho_{k,n}^*} = 0$

If $\rho_{k,n}^* = 1$, $\left. \frac{dL}{d\rho_{k,n}} \right|_{\rho_{k,n}=\rho_{k,n}^*} < 0 \tag{3.23}$

Now we re-consider the integer constraint $\rho_{k,n} \in \{0,1\}$ in the following discussion.

$$\rho_{k,n}^* = \begin{cases} 0 & \text{if } \lambda_n > (1+\mu_k)c_{k,n} \\ 1 & \text{if } \lambda_n < (1+\mu_k)c_{k,n} \end{cases} \tag{3.24}$$

From 3.23, we find that for each sub-carrier n, if all users $k \in \{1, 2, \cdots, K\}$ have different $(1+\mu_k)c_{k,n}$, then only the user with the largest $(1+\mu_k)c_{k,n}$ can use this nth sub-carrier. The above statement can be represented as follow.

$$\rho_{k,n}^* = 1, \, k = k^*$$

$$\rho_{k,n}^* = 0, \text{ for all } k \neq k^*$$

where $\qquad k^* = \arg \max_k (1+\mu_k)c_{k,n} \tag{3.25}$

Here *arg max* stands for the argument of the maximum. Hence, for a fixed set of Lagrange multipliers μ_k, $k \in \{1, 2, \cdots, K\}$, we can use equation 3.25 to determine k^* for each sub-carrier n. Then the $\rho_{k,n}^*$ obtained is the solution for the optimization problem.

The optimal solution of sub-carrier allocation can be obtained from Equation 3.25 when μ_k for all users are determined. Accordingly, the optimal power and bit allocation scheme can be represented as

$$p_{k,n} = \begin{cases} \dfrac{P_T}{N}, & k = k^* \\ 0, & k \neq k^* \end{cases} \tag{3.26}$$

$$c_{k,n} = \begin{cases} f^{-1}\left(\alpha_{k,n}^2 p_{k,n}\right), & k = k^* \\ 0, & k \neq k^* \end{cases} \tag{3.27}$$

We can simplify the resource allocation problem to finding the set of μ_k that satisfies the users' minimum transmission rate constraints. Let r_k denotes the allocated rate for user k.

$$r_k = \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \tag{3.28}$$

According to convex optimization, this Lagrangian multiplexer can be solved by the following equations.

$$\begin{aligned} \mu_k = 0, & \text{ if } r_k > R_k \\ \mu_k > 0, & \text{ if } r_k = R_k \end{aligned} \tag{3.29}$$

Hence, the optimal resource allocation for multiuser OFDM systems is to find the solution of μ_k under the conditions shown in Equation 3.29. In the next two sections, we will discuss the design of resource allocation schemes.

3.3.2.2 Design of Resource Allocation without Minimum Transmission Rate Constraint

First, we consider a simple situation in which the minimum transmission rate constraints for users are not required. Therefore, for all k, $R_k = 0$. Equation 3.29 can be written as $\mu_k = 0$, $k = 1, 2, \cdots, K$. In this case, Equation 3.25 is updated to

$$k^* = \arg\max_k c_{k,n} \qquad (3.30)$$

which assigns each sub-carrier to the user with the maximum bit rate. The optimal allocation scheme without minimum transmission rate constraints is described in the following algorithm.

Algorithm 3.3 Optimal Resource Allocation without Minimum Transmission Rate Constraints

    i)   Sub-carrier allocation

        for $n = 1$ to N

          for $k = 1$ to K

$$c_{k,n} = f^{-1}\left(\alpha_{k,n}^2 \frac{P_T}{N} \right)$$

          end

$$k^* = \arg\max_k c_{k,n}$$

$$\rho_{k^*,n} = 1$$

$$\rho_{k,n} = 0, \quad for \ k \neq k^*$$

end

ii) Bit and power allocation

for $n = 1$ to N

$$p_{k^*,n} = \frac{P_T}{N}$$

$$\rho_{k,n} = 0, \quad for \ k \neq k^*$$

$$c_{k^*,n} = f^{-1}\left(\alpha_{k^*,n}^2 \frac{P_T}{N}\right)$$

$$c_{k,n} = 0, \quad for \ k \neq k^*$$

end

3.3.2.3 Design of Resource Allocation Scheme with Minimum Transmission Rate Constraint

When the users minimum required transmission rates are not zero, $R_k \neq 0$. According to Equation 3.29, appropriate value of μ_k should be determined for each user to ensure that the allocated transmission rate $r_k \geq R_k$. Therefore, the allocation scheme should satisfy the following two rules.

(1) μ_k are set to zero for users whose allocated rates have exceeded their minimum required rates.

(2) for the other users whose transmission rate dose not exceed the minimum

45

rate $R_k{}^2$, proper positive values of μ_k need to be found to guarantee their minimum transmission rate constraint.

Previous analysis indicates that the allocated rate of the *kth* user may be affected not only by μ_k but also by $\mu_j (j \neq k)$ of other users. Therefore, it is difficult to get μ_k for each user from the above criterions. To solve the problem, we propose a simplified resource allocation algorithm based on iteration. First, μ_k for all users are set to zero, and the transmission rates for users are calculated. Second, the algorithm checks if the current allocated rates satisfy the minimum required rates of each user. If the minimum transmission rate of user k is not satisfied, the algorithm will increase μ_k to re-allocate sub-carriers to the user until its allocated rate meet the R_k constraints. To avoid interfere with the previous allocated users in the reallocation process, During the re-allocation for one user, the sub-carriers already allocated to the user whose minimum required rate has not been satisfied or $\mu_k > 0$ will not be reassigned to other users. The algorithm calculates μ_k for each user sequentially to satisfy the user minimum transmission rate. We notice that in some poor channel conditions, the resource allocation schemes cannot satisfy some user's minimum transmission rate constraint. In this case, the algorithm will be terminated and an outage event will be reported. We introduce outage probability to determine the performance of algorithms. The outage probability describes the probability that not all users' minimum transmission rates are satisfied using the proposed resource allocation scheme. To fight this situation, users can be sorted based on their QoS, then the transmission

requirement for users with higher QoS will be considered first. Details of the proposed sub-carrier, power and bit allocation algorithm is described in Algorithm 3.4.

Algorithm 3.4 Resource allocation algorithm with minimum transmission rate constraint

Initialization

Step 1. Find $k^* = \arg\max_{k} c_{k,n}$, $n = 1, 2, \cdots, N$. Set $\mu_k = 0$, $k = 1, 2, \cdots, K$, excluding all sub-carriers allocated to users of $r_k < R_k$.

Iterative Process

Step 2. While termination condition is not met, do Steps 3-6.

Step 3. Select user k with $r_k < R_k$. Find the solution of μ_k using Algorithm 3.5. For all $n \in S$ that $(1 + \mu_k) c_{k,n} > c_{k^*,n}$, let $r_{k^*} = r_{k^*} - c_{k^*,n}$. Then, reallocate n to user k, and remove n from $\{1, 2, \cdots, N\}$. If no solution of μ_k exists, then an outage event occurs and the algorithm stops.

Step 4. If for all $k = 1, 2, \cdots, K$ we have $\mu_k > 0$ and $r_k < R_k$, then an outage event occurs and the algorithm stops.

Step 5. If $r_k \geq R_k$, $k = 1, 2, \cdots, K$, the algorithm stops.

Step 6. Go to Step 2.

If there is an allocation scheme which satisfies the minimum transmission rate constraint of all users, the above algorithm gives a suboptimal solution. Otherwise, the algorithm will terminate and report an outage so that the convergence of the algorithm can be guaranteed. The outage probability will be discussed in the following simulation. Note that though the proposed algorithm searches for a suboptimal solution of μ_k, simulation results indicate that its performance is very close to the optimal allocation scheme. This will be further discussed in the next simulation and analysis section. Bisection search is used in Algorithm 3.4 to find the solution, and Algorithm 3.5 gives the steps of finding the solution of μ_k.

Algorithm 3.5 Solving μ_k

Step 1. Set $\mu_{low} = 0$.

Step 2. While termination condition is not met, do Steps 3-7.

Step 3. Find a sufficiently large μ_{high} that ensures $r_k \geq R_k$, $k = 1, 2, \cdots, K$.

Step 4. If such μ_{high} does not exist, then there is no solution of μ_k, terminate.

Step 5. If $\bar{R} \neq R_k$, $\mu_{new} = \dfrac{\mu_{low} + \mu_{high}}{2}$, $\bar{R} = r_k$. For all $n \in S$ that $(1 + \mu_{new}) c_{k,n} > c_{k^*,n}$, let $\bar{R} = \bar{R} + c_{k,n}$. If $\bar{R} > R_k$, $\mu_{high} = \mu_{new}$; else $\mu_{low} = \mu_{new}$.

Step 6. Return solution of $\mu_k = \mu_{new}$.

Step 7. Test termination condition. Verify if $\bar{R} = R_k$, $k = 1, 2, \cdots, K$.
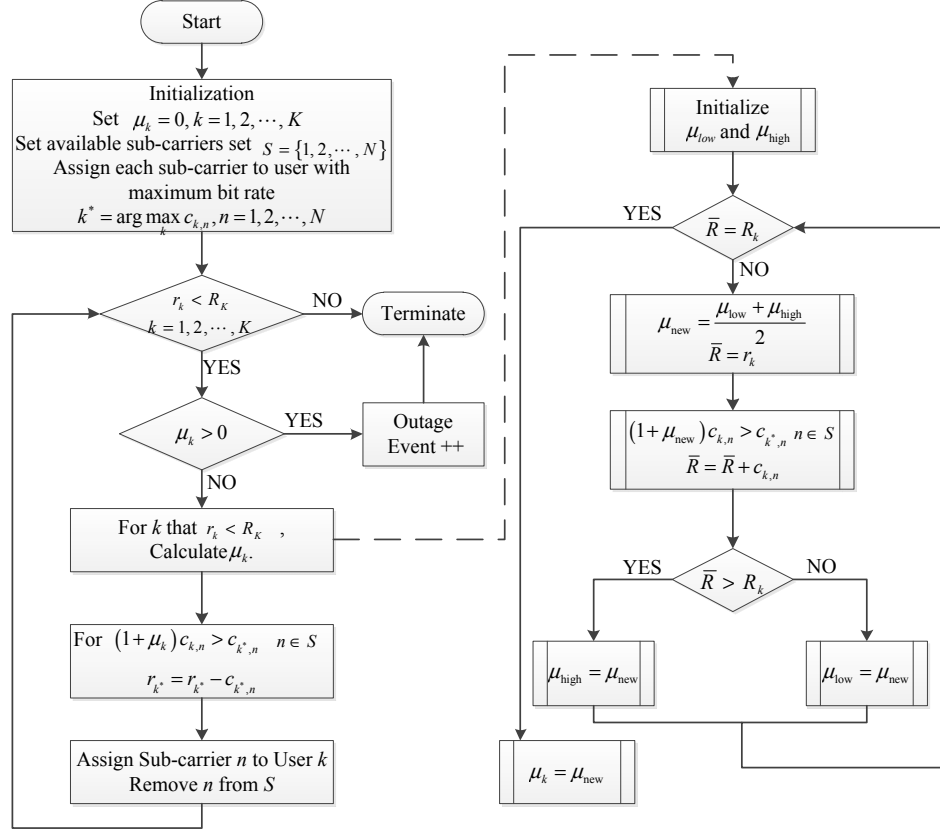
Figure 3.7   Flowchart of adaptive resource allocation algorithm with minimum

transmission rate constraint

3.3.3 Computer Simulation and Algorithm Complexity Analysis

The multiuser OFDM simulation system is set up similar to the system described in section 3.2.3, with K represents the number of users.

We first evaluate the system performance. Figure 3.8 compares the spectral efficiency of adaptive allocation Algorithm 3.5 (ASA) and the traditional fixed sub-carrier allocation (FSA) scheme. The performance is compared in scenarios of different numbers of users. We assume the system uses adaptive bit loading technique [30], and all sub-carriers are evenly allocated to users in FSA. It can be seen that

Figure 3.8   Performance comparison of fixed sub-carrier allocation scheme (FSA)

and proposed adaptive sub-carrier allocation scheme (ASA) in multiuser OFDM

systems

when there is one user ($K = 1$), performance of the proposed adaptive sub-carrier

allocation scheme is the same as that of fixed sub-carrier allocation scheme. Spectral

efficiency of the system with adaptive resource allocation scheme is improved when

the number of users increases. We observe a performance gain of about 3dB when the

number of users rises from 4 to 16. The proposed adaptive sub-carrier allocation

scheme exploits multiuser diversity and allocates the sub-carrier to the user with the

best channel condition. The gain will also increase as the number of user increases.

Figure 3.9   Spectral efficiency of fixed sub-carrier allocation scheme (FSA) and

proposed adaptive sub-carrier allocation schemes (ASA) in multiuser OFDM systems

(K=4)

Figure 3.9 and Figure 3.10 show the influence on system performance by adding the

minimum transmission rate constraint. Performance of three algorithms namely

Algorithm 3.4 (ASA-MRC), adaptive sub-carrier allocation without minimum

transmission rate constraint (ASA w/o MRC), and bit allocation algorithm under fixed

sub-carrier allocation (FSA w/o MRC) are compared. In Figure 3.9, we choose a

4-user OFDM system. The minimum transmission rate constraints of users are set to

128, 128, 64, and 64 bits per OFDM symbol. It can be seen that the adaptive
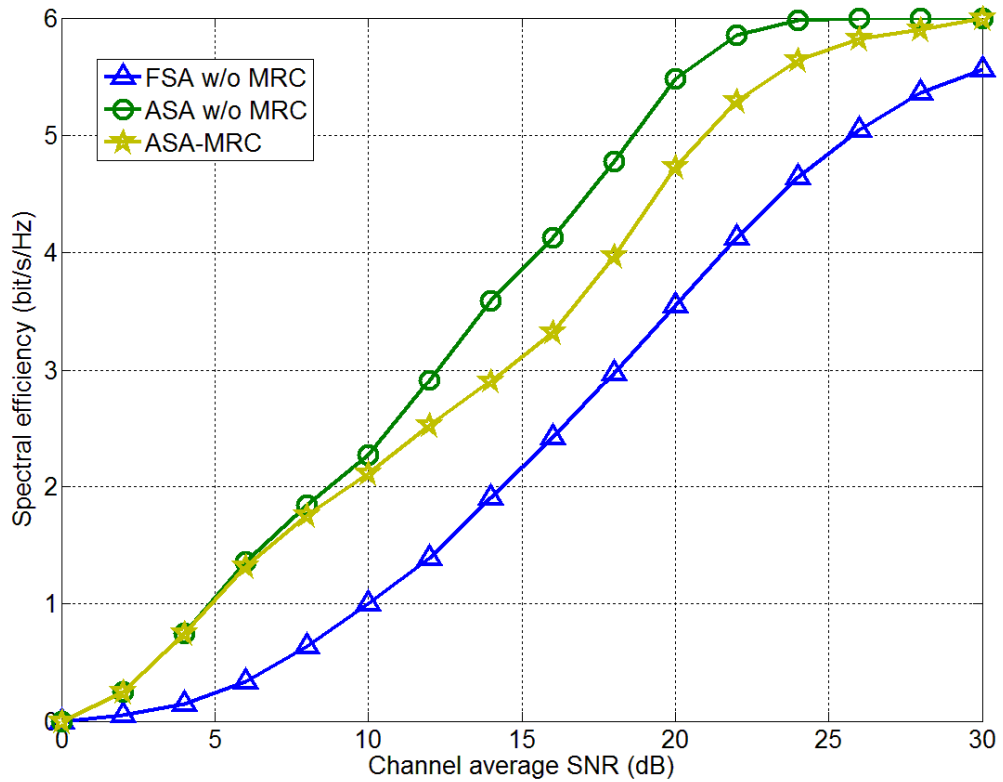
Figure 3.10 Spectral efficiency of fixed sub-carrier allocation scheme (FSA) and

proposed adaptive sub-carrier allocation schemes (ASA) in multiuser OFDM systems

(K=8)

sub-carrier allocation scheme outperform the fixed sub-carrier allocation scheme by

5dB when the spectral efficiency is above 2 bit/sub-carrier. Compared with adaptive

allocation, the added constraint causes a decrease of 1dB of throughput at mid to high

SNR area. Figure 3.10 shows similar results in an 8-user OFDM system with

minimum transmission rate constraints as 128, 128, 64, 64, 64, 32, 32, 32 bits per
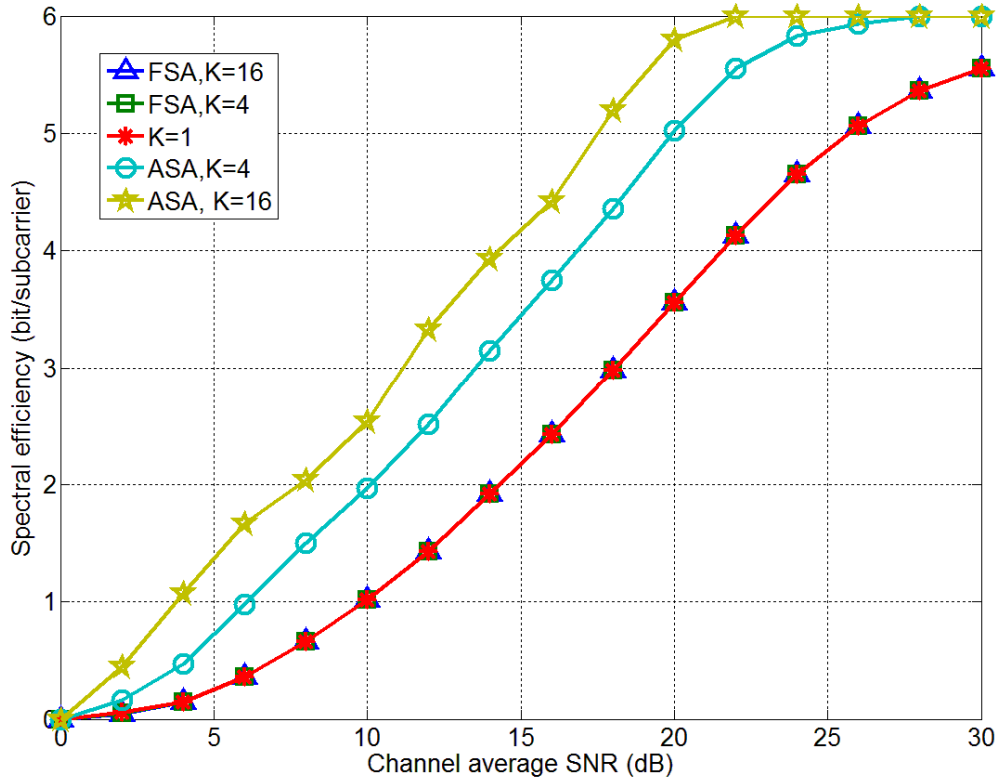
OFDM symbol.
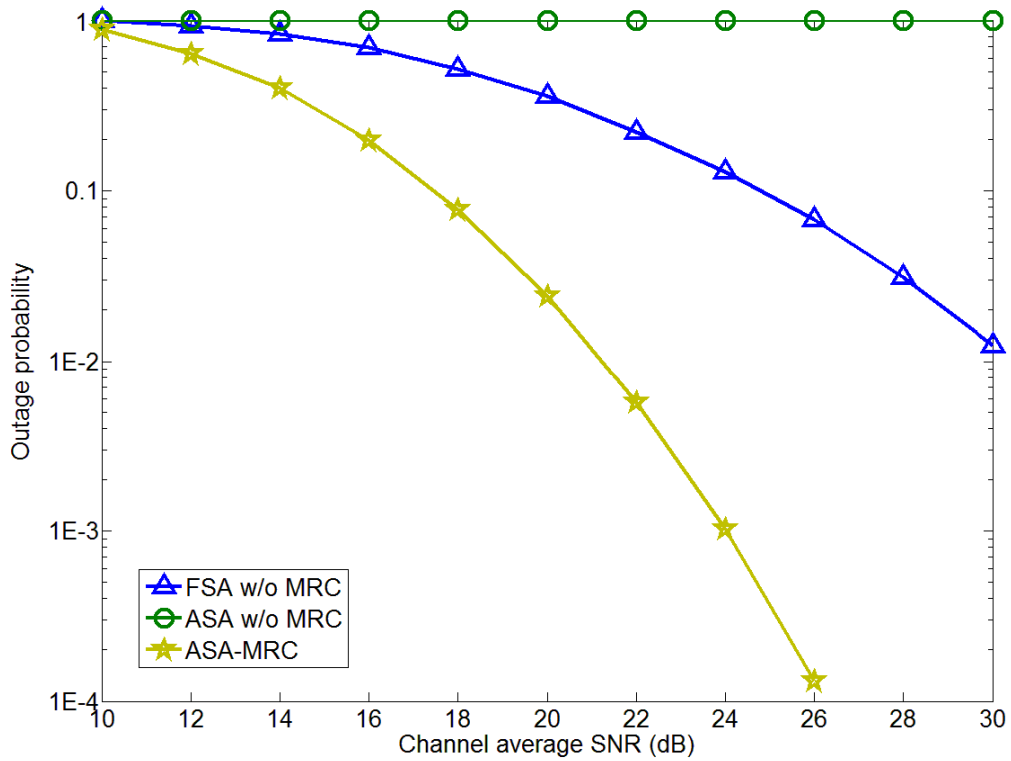
Figure 3.11 The outage probabilities of fixed sub-carrier allocation scheme (FSA)

and proposed adaptive sub-carrier allocation schemes (ASA) in multiuser OFDM

systems (K=4)

Figures 3.11 and 3.12 show the outage probability of these allocation schemes.

The proposed adaptive allocation scheme with minimum transmission rate constraint

can effectively reduce the outage probability. When the SNR is high, the minimum

transmission rate cannot be guaranteed if adaptive allocation technique is used alone.

The performance of the fixed sub-carrier allocation scheme is not the worst, but at a

heavy cost of overall throughput.

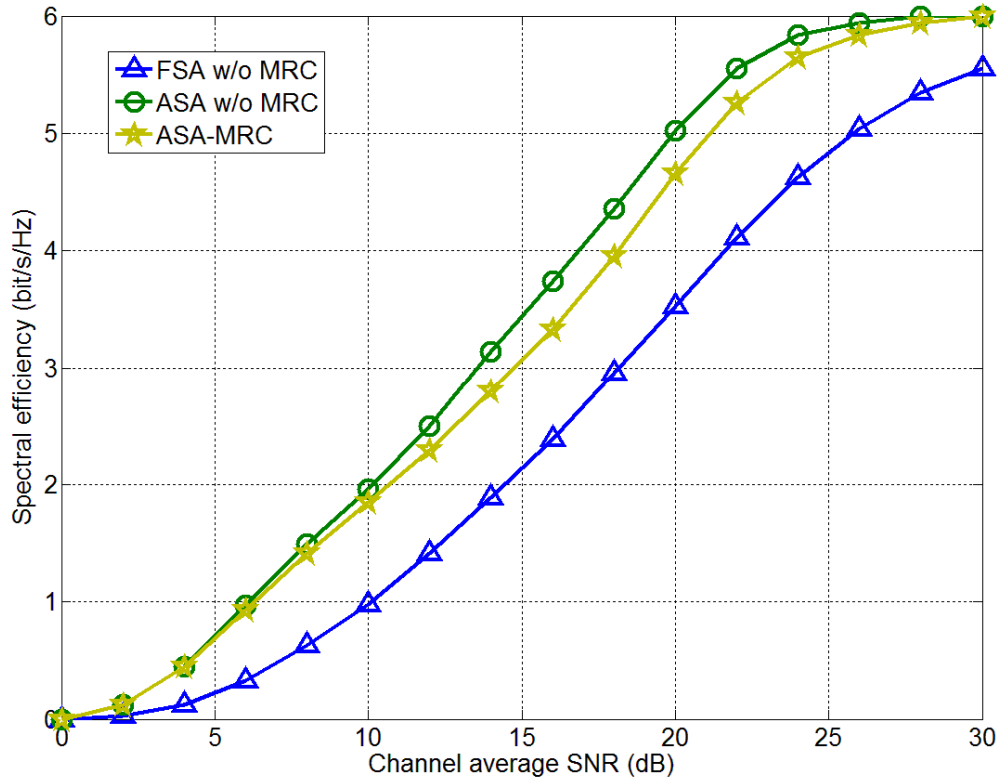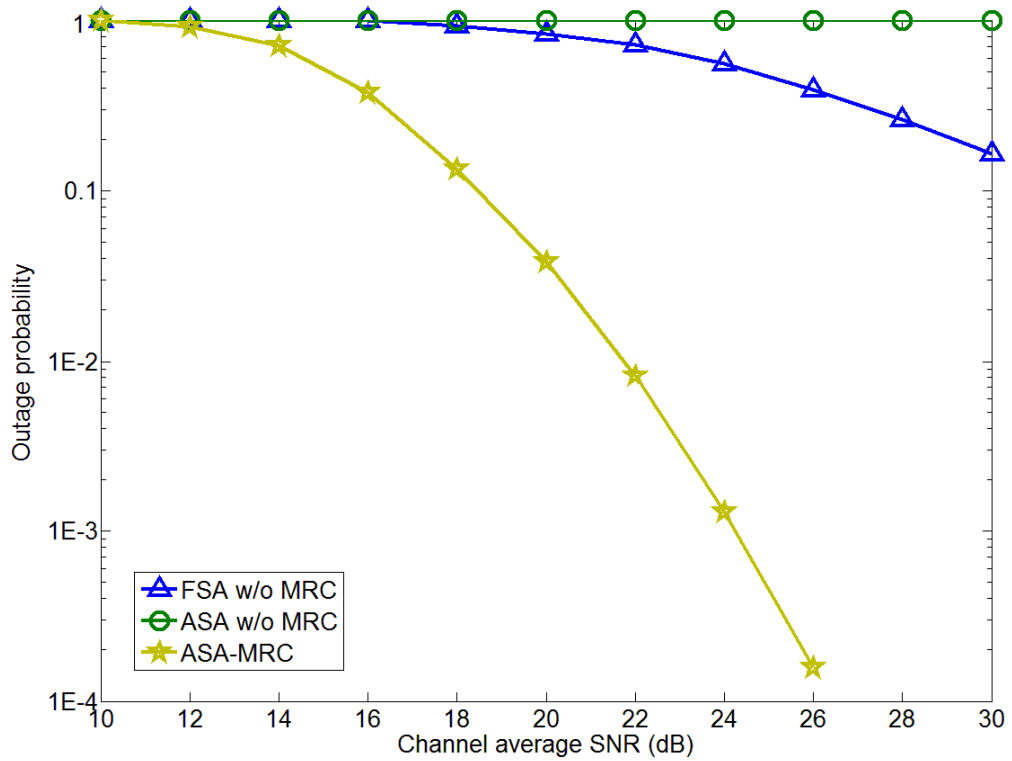Figure 3.12   The outage probabilities of fixed sub-carrier allocation scheme (FSA)

and proposed adaptive sub-carrier allocation schemes (ASA) in multiuser OFDM

systems (K=8)

In Figure 3.13 the adaptive allocation scheme with minimum transmission rate constraint is compared with the optimal results obtained by linear integer programming. The performance of the proposed algorithm is very close to linear integer programming. It is able to achieve a near optimal solution with only about 0.5% performance.
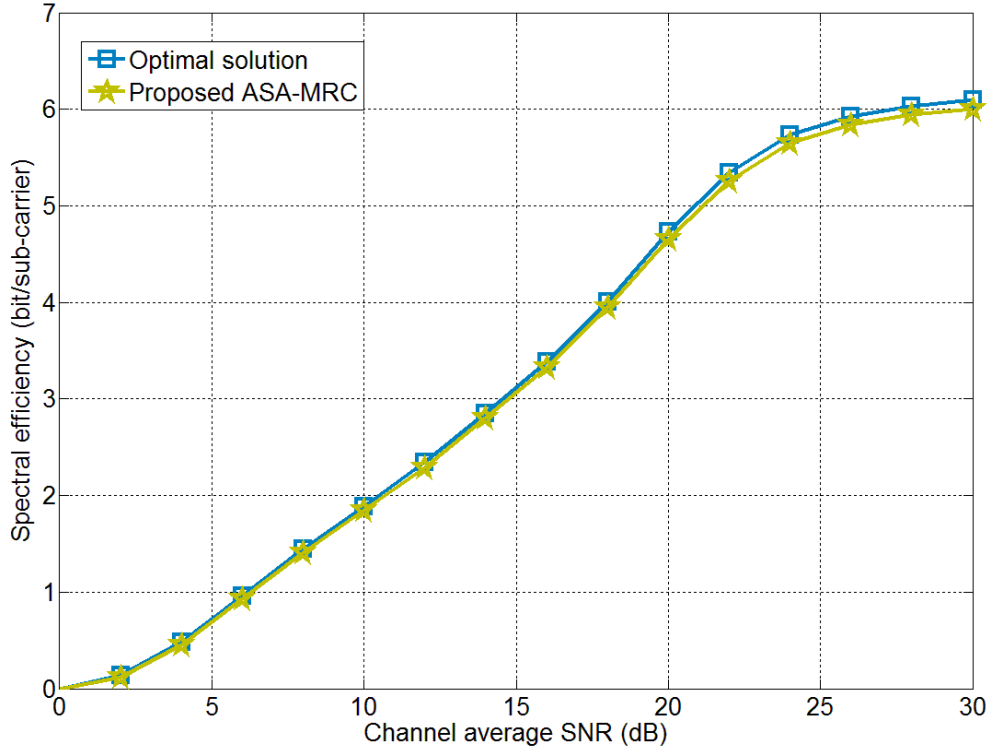
Figure 3.13 Performance comparison between proposed adaptive sub-carrier

allocation scheme and the exhaustive search

We further compare the computational complexity of the proposed algorithm,

optimal solution using linear integer programming and the algorithm proposed by

Zhang [30]. The proposed algorithm requires at most N multiplications and N

comparisons for checking the re-allocation using an updated μ_{new}. The convergence

of the bisection search needs I iterations. Since $c_{k,n}$ are integers from set M, the

solution of μ_k can be restricted to fixed values. Therefore, I is less than

$(2\log_2 L - 1)$, where L is the maximal integer in M. Since we need to solve η_k s of no

more than $(K-1)$ users, $3(K-1)N$ computations are required for re-allocations. So far,

the overall complexity of our proposed algorithm with minimum transmission rate constraints can be concluded to

$$KN + 3KN + 2N\left(2\log_2 L - 1\right)\left(K - 1\right)$$

The quantitative complexities of these algorithms are listed in Table 3.2. To get the optimal solution, the linear integer programming is the most complex of the algorithms. Zhang's algorithm is less complex. Considering $\log_2 M \leq 3$ and $K \ll N$, our proposed algorithm with minimum transmission rate constraint is more computationally efficient than Zhang's algorithm, makes it the best algorithm in this comparison.

Table 3.2   Computational complexity comparison

Algorithm	Computational Complexity
Optimal solution	$O\left(2\left(KN + K + N\right)\sqrt{2}^{KN}\left(2KN\left(K + N\right) + KN\right)\right)$
Zhang	$O\left(3KN + 2N^2\right)$
Proposed ASA-MRC	$O\left(2KN + 4KN\log_2 L\right)$

3.4 Summary

In this chapter, we study the resource allocation problem in multiuser OFDM systems, and propose an efficient resource allocation scheme. The allocation problem is designed to maximize the system spectral efficiency under the constraints of users' QoS requirements. Both BER and minimum transmission rates for users are considered. The bisection search is employed to accelerate the convergence speed of

the algorithm. Simulations and algorithm complexity comparison show that our algorithm offers a more preferable tradeoff between the performance and the complexity than existing approaches. The proposed adaptive allocation algorithm with minimum transmission rate constraints satisfies users' QoS requirements, and achieves good performance while maintaining a lower complexity.

CHAPTER IV   FAIR SCHEDULING SCHEMES IN MIMO-OFDM

SYSTEMS

4.1 MIMO-OFDM and Scheduling

Orthogonal Frequency Division Multiplexing (OFDM) is one of the most promising physical layer technologies for high data rate wireless networks. The advantages are robustness to frequency selective fading, high spectral efficiency, and low computational complexity comparing with the traditional wireless technologies. Multiple-Input Multiple-Output (MIMO) transceivers can be used in conjunction with OFDM to increase the diversity gain and the system capacity by exploiting spatial domain. MIMO-OFDM is considered a key technology in emerging high-data rate systems such as 4G, IEEE 802.16, and IEEE 802.11n [62].

Making efficient use of radio resources is a very challenging task for 3G/4G wireless communication systems as the scarcity of radio resources, diverse QoS requirements, and wireless channel conditions further complicate scheduling and radio resource management. The main purpose of employing a scheduler is to intelligently allocate radio resources to achieve high system performance in terms of efficiency and fairness. This makes the scheduling scheme a key component in optimizing the system performance across the physical (PHY) and medium access control (MAC) layers. [63] Since the scheduling scheme is defined as the method of user selection and how to allocate available resources to a selected user, there are two crucial key performance indicators, fairness and efficiency. Traditionally, spectral

efficiency is evaluated in terms of system throughput, which is sometimes unfair to those users who are far away from a base station or with bad channel conditions. On the other hand, absolute fairness may lead to low bandwidth efficiency. Therefore, an effective trade-off between efficiency and fairness is desired in the communication system.

4.2 MIMO-OFDM System Model

The model of the system under study is illustrated in Figure 4.1. We consider a multiuser MIMO-OFDM system employing N_T transmit antennas at the base station. There are K users, and each user has N_R receive antennas. We assume that there are N sub-carriers, so that the MIMO channel between the kth user and the base station on
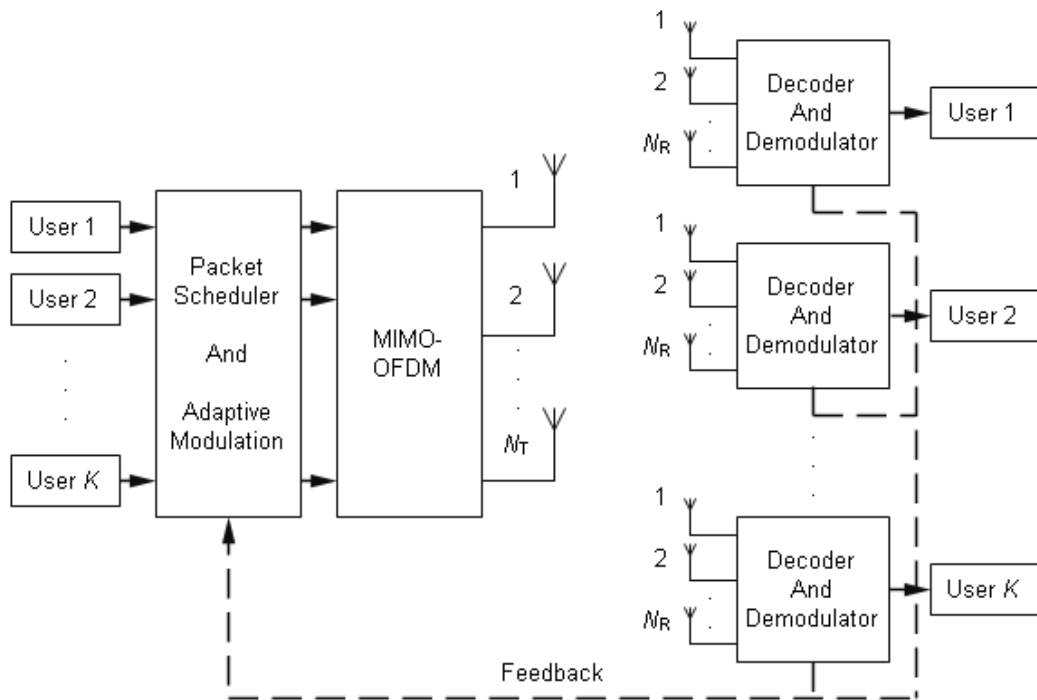
Figure 4.1   MIMO-OFDM downlink multiuser diversity with scheduling

*N*th sub-carrier can be expressed by $H[k,n]$ matrix of size $N_R \times N_T$, with elements $H_{i,j}[k,n]$ corresponding to the discrete Fourier transform (DFT) of the $h_{i,j}(k,n)$; the latter corresponds to the complex channel gain of the channel response between the *i*th transmit and *j*th receive antennas. Depending on the employed orthogonal space-time code, the number of transmit antennas N_T will be different [64].

In this system, all users have access to the *N* sub-carriers of OFDM in each time slot. However, each frequency slot is dedicated to only one selected user. The multiuser scheduler can select a user based on the channel quality of these users in each frequency slot according to the feedback channel information. Similar to the expression in [65], the received signal for the orthogonal space-frequency block coded (OSFBC) OFDM after OSFBC decoding can be written as Equation 4.1.

$$\tilde{s}[k,n] = c \parallel H[k,n] \parallel_F^2 s[k,n] + \eta[k,n] \tag{4.1}$$

where $\tilde{s}[k,n]$ is the output signal of the decoder, $s[k, n]$ is the transmitted symbol, $\parallel . \parallel_F^2$ is the squared Frobenius matrix norm as shown in Equation 4.2.

$$\parallel H[k,n] \parallel_F^2 = \sum_{i=1}^{N_T} \sum_{j=1}^{N_R} \mid H_{i,j}[k,n] \mid^2 \tag{4.2}$$

The term $\lambda[k, n]$ is a complex Gaussian noise with distribution $N\left(0, \frac{cN_0}{2} \parallel H[k,n] \parallel_F^2\right)$ per dimension. *c* is a constant which depends on the system. For the orthogonal space-time block codes, the total energy of the symbol transmitted through the N_T antennas can be normalized to N_T, therefore, we can express the instantaneous signal-to-noise ratio (SNR) per symbol at the receiver of the *k*th user as

$\gamma[k,n] = \frac{\bar{\gamma}}{N_T R_c} \| H[k,n] \|_F^2$, where $\bar{\gamma}$ is the average receive SNR per antenna, and R_c is the code rate [66].

We consider one cell with a single base station (BS) located at the cell center and K users uniformly distributed within the cell radius. In the MIMO-OFDM systems, the sub channel fading $H_{i,j}[k, n]$ is considered as a Rayleigh fading. At the beginning of each frame, the BS obtains the estimated SNR value from all users. The scheduler selects one mobile station (MS) based on the schemes introduced in the next section.   Then bit allocation is performed on the selected MS. In bit allocation, adaptive modulation and coding is applied to different tones for the selected user. That is, when the channel condition on a specific tone is good, more bits are loaded on this tone to achieve higher system throughput. The number of bits that each tone can carry is determined by the estimated SNR value as well as the SNR margin, which is required to guarantee a symbol error rate below a certain margin, and to combat any unexpected interference.   This is to meet the QoS requirements for different types of multimedia communications service.

4.3 Scheduling Schemes

In this section, first, the conventional scheduling schemes, Round Robin and Max-SNR are described. Second, the Proportional Fair (PF) [67] scheme is introduced. We modified the PF scheme to improve system throughput, delay time and user satisfaction rate. Afterwards, three variations are proposed and discussed, emphasizing the modifications made to adjust the algorithm.

4.3.1 Round Robin and Maximum SNR

Round Robin (RR) and Maximum SNR (Max-SNR) are two basic and less complex scheduling strategies in packet data services. RR services the users in a round robin fashion. All active users are identified by their IDs, and each user is allocated a number of time slots. Transmission service will not be granted to the same user before all the other users have been served. In MaxSNR, all active users are periodically ranked by their reported SNR values. At each scheduling event, the scheduler selects the user with the highest SNR value and allocates a number of time slots depending on service requirements and availability of resources. It is possible for the same user to be rescheduled for the next available resource provided that it still has the highest SNR value at the next event. It is well known that RR has the advantage of being fair in terms of giving equal slots to every user, and MaxSNR can provide high system throughput.

4.3.2 Proportional Fair

In order to improve fairness without sacrificing much in terms of throughput, both fairness and throughput should be incorporated in one scheduling scheme. Proportional Fair (PF) is one of many efforts to address this issue [68]. The scheduler selects the user with a maximal priority metric, defined as Equation 4.3.

$$m = \arg \max_{1 \le i \le K} \frac{T_i(t)}{R_i(t-1)} \tag{4.3}$$

where i is the user index, m is the selected user, K is the total number of users, $T_i(t)$

denotes the data rate potentially achievable by the channel in the present time slot, and $R_i(t-1)$ is the average data rate of this user. With this scheme, a user is selected when it has a good channel, or high $T_i(t)$, to keep system throughput high. In the meantime, users with bad channel conditions are also considered since their low average data rate $R_i(t-1)$ will increase their chance of being selected for the next scheduling event. This scheme is implemented in high data rate (HDR) networks, and has recently been studied by many authors [69]. Their research results show uneven data rates achieved by different users with different channel conditions. In other words, PF is not fair enough for users with bad channel conditions. To improve the fairness of PF, one approach is to consider what is known as the data rate control (DRC) exponent rule proposed in [70], where the user selection criteria was modified

to $m = \arg\max_{1 \le i \le K} \dfrac{T_i^q(t)}{R_i(t-1)}$ with q for a weighted parameter introduced to manage the

relationship among the data rates of users with different channel conditions. When the control parameter q is fixed, we miss two important flexibilities: the flexibility to adapt to the current wireless channel condition and the flexibility to choose a value for q that ensures fairness among all users at the same time.

The PF schemes are implemented in an 802.16e OFDM based system. The performance of PF depends on the observation window. To provide long-term fairness from user activation to termination, we define the criterion as shown in the equation below

$$I = \arg\max_{1 \le i \le K} \frac{T_{i,j}}{R_{i,j}}$$

where $T_{i,j}$ is the supportable rate at symbol j of user i, and $R_{i,j}$ is the average experienced rate until symbol j by user i, which is calculated as the total number of bits received divide by the time from its activation to the current instant, with the assumption that this user stays active all the time. In the case of user termination, it is deregistered from the BS, and the scheduler will not count it as one of its active users.

4.3.3 Adaptive Proportional Fair

We may add an adaptive exponent to the numerator of PF and name it adaptive PF. Under the assumption that all users are active all the time, the criterion becomes Equation 4.4.

$$I = \arg\max_{\substack{1 \le i \le K \\ 1 \le j \le M}} \frac{T_{i,j}^{n_{i,j}}}{R_{i,j}} \tag{4.4}$$

It is the same as PF except for adaptive component $n_{i,j}$ to equalize the value of $T_{i,j}/R_{i,j}$, and M is the total number of symbols in each frame.

Based on adaptive PF with added parameter for required data rate of user $R_{i\_r,j}$, the new criterion is described by Equation 4.5.

$$I = \arg\max_{\substack{1 \le i \le K \\ 1 \le j \le M}} \frac{T_{i,j}^{n_{i,j}} \cdot R_{i\_r,j}}{R_{i,j}} \tag{4.5}$$

4.3.4 Normalized Adaptive Proportional Fair

In the above criterions, we notice that there are two possible problems: the value of I is very large, and they cannot achieve good user satisfaction and average user rates. To solve these problems, we propose a new criterion as shown in Equation 4.6

$$I = \arg \max_{\substack{1 \leq i \leq K \\ 1 \leq j \leq M}} \left(\frac{T_{i,j}^{n_{i,j}} \cdot R_{i\_r,j}}{R_{i,j}} \right)^{n_{i,j}} \frac{T_{i,j}}{\sum_{i=1}^{N} T_{i,j}} \tag{4.6}$$

There are two changes. In the denominator, a normalization factor is added. Hence, this scheme is referred to as normalized adaptive PF. Thus, the first problem in adaptive PF is solved. The adaptive exponent is added to the ratio of the average and required rates. Aiming at improving user satisfaction rate, we let $\left(\dfrac{R_{i\_r,j}}{R_{i,j}} \right)^{n_{i,j}} = \sigma_i$ to obtain the value of $n_{i,j}$, where σ_i is the threshold defined in the user satisfaction rate. The normalized adaptive proportional fair (NAPF) is expected to give a better performance among all proposed schemes for multi-rate traffic. For comparison, we will also implement RR and MaxSNR to give the lower and upper bound of throughput.

4.4 Computer Simulation and Analysis

System simulations were used to determine the throughput performance of 802.16e OFDM based on the simulation assumptions [71]. The cell layout is Hex grid, and the cell radius is set to 1km. Number of sub-carriers is 8. There are 12 users and 1

base station. The packet size is set to 1Mbyte. The frame structure is defined as 2.5GHz of frequency. The number of OFDM carriers is 64. The modulation scheme is quaternary phase shift keying (QPSK).

Figure 4.2 is the comparisons of system throughput vs. load for different schemes. In IEEE 802.16e, the traffic load can reach 100Mbps. When the load is above 47Mbps, RR has the lowest throughput as expected, and PF improves system throughput over RR. Throughput of NAPF is even higher than MaxSNR when the system load is more than 40Mbps.
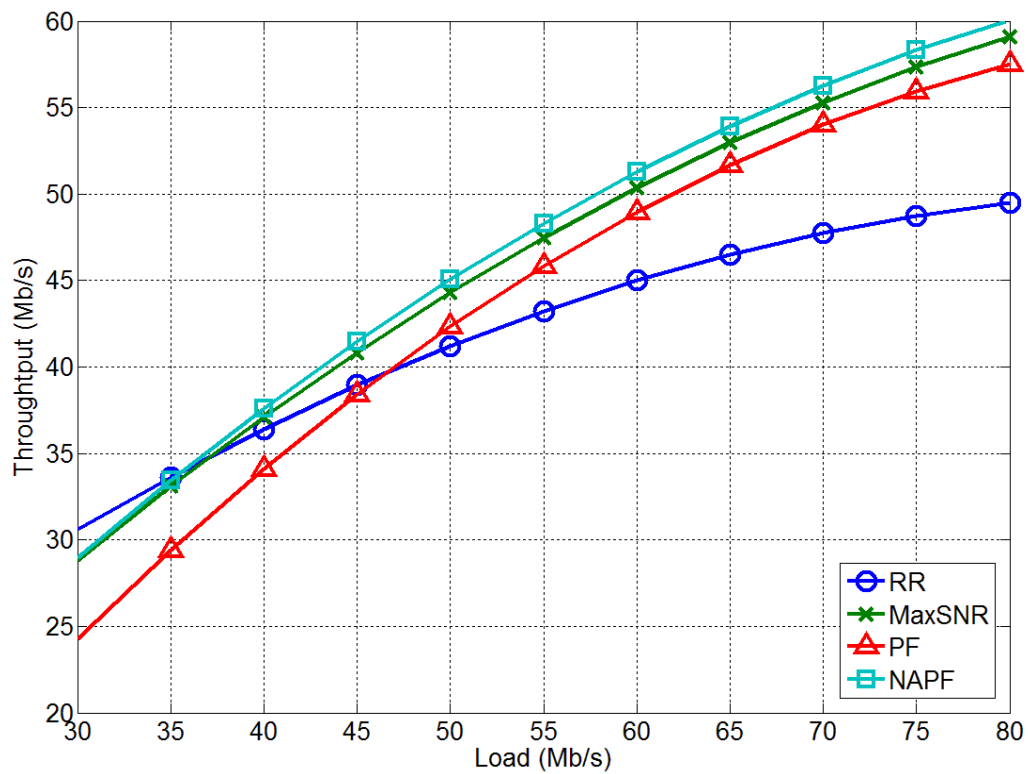
Figure 4.2   System throughput comparisons

Mean packet delay is the mean value of packet completion time, which equals to the time of instant of packet arrival until the user receives the packet. Figure 4.3 shows the mean delay comparison results. When the system throughput is more than 40Mbps, we can see that RR has the largest delay. PF scheme has less delay than RR, but its performance is still worse than MaxSNR and NAPF when the throughput exceeds 50Mbps. The delay of NAPF scheme is less or equals to MaxSNR when the system throughput is more than 35Mbps.
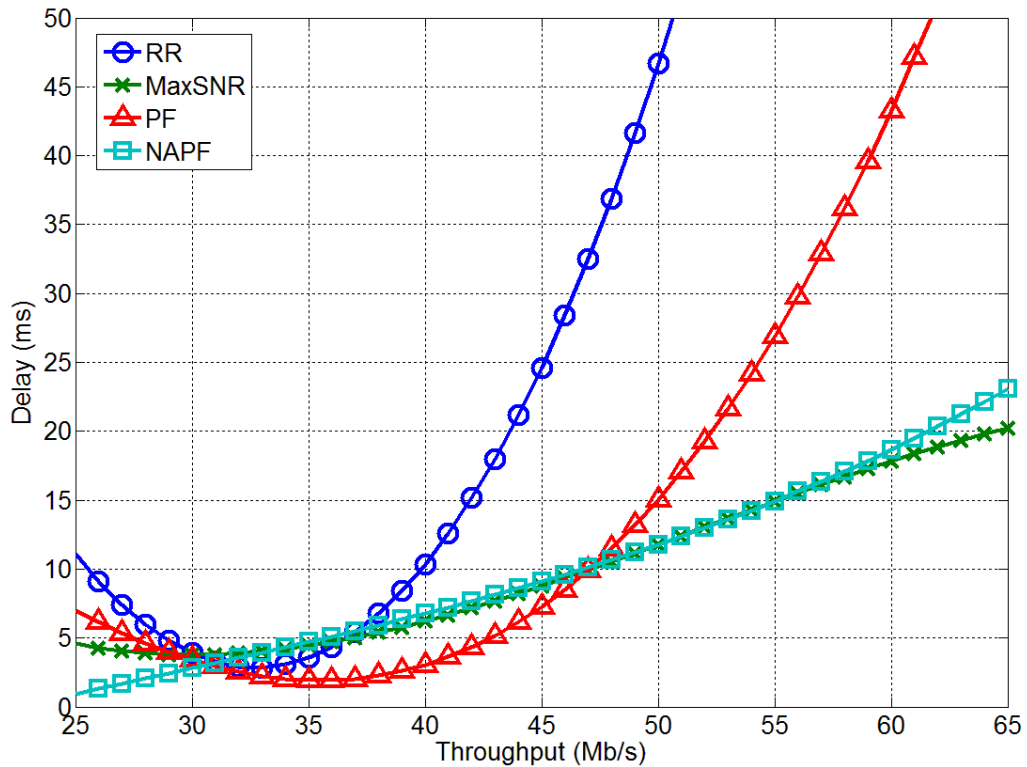
Figure 4.3　Delay comparisons

Figure 4.4 shows the fairness performance of these schemes. The user satisfaction rate is defined as the number of users with no blocking and no dropped

packets to the number of total users. When the throughput is larger than 40Mbps, both of the PF schemes perform better than RR and MaxSNR. The criterion of PF indicates that the scheduler not only favors those mobile stations with high supportable rate, but also its experienced rate. Our proposed scheme gives extra weight to users with good channel. It considers the required rate as well. Compared with PF, our proposed algorithm improves the system throughput. At the same time, it also gives enough fairness to mobile stations whose channel is not so good in the past. In other words, if a user does not receive service at previous frame due to its bad channel condition, then its precedence will increase at the next frame.
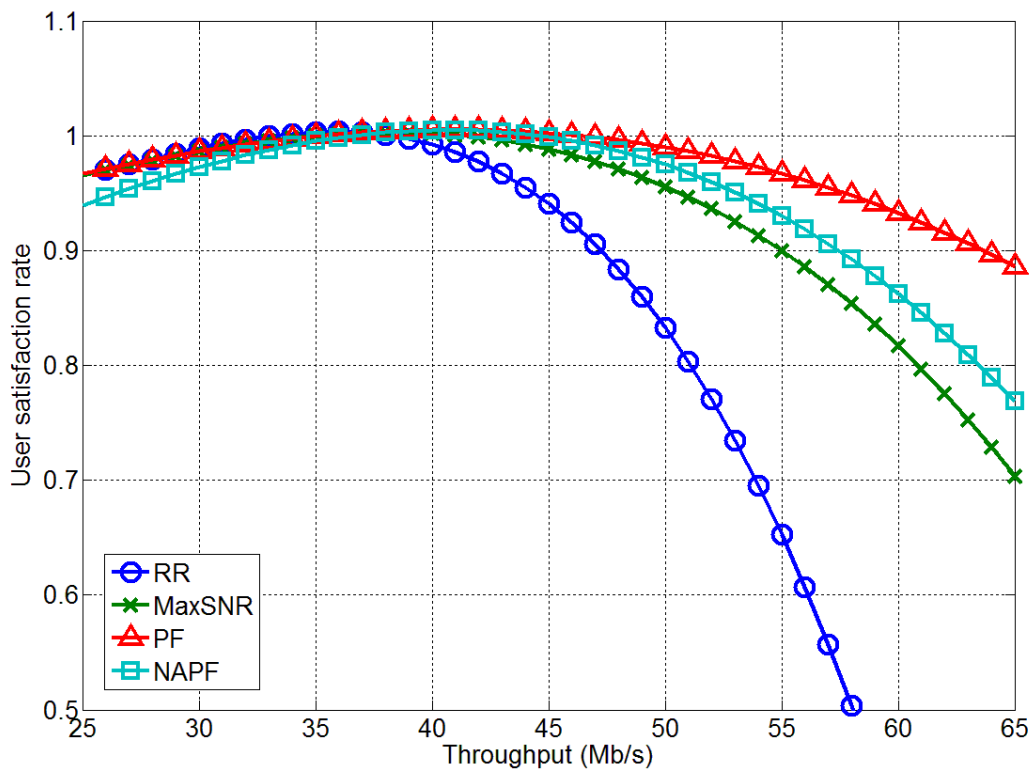
Figure 4.4   User satisfaction rate comparisons

4.5 Summary

In this chapter, we have presented and analyzed multiuser scheduling schemes. A simulation and performance evaluation is performed and shows the significant advantages of proportional fair and the proposed variations over the conventional schemes. The proposed schemes improve system throughput and delay time. The proposed normalized adaptive proportional fair scheme achieves higher throughput without sacrificing fairness. It is also a better scheme in terms of efficiency and fairness.

CHAPTER V   FEEDBACK STRATEGIES AND ALGORITHMS BASED ON
FEEDBACK COMPRESSION

5.1 Compression of Feedback

Channel SNR is relevant in time domain because of Doppler shift, and it is also relevant in frequency domain because of multipath delay spread. Data compression algorithms can utilize the relevance to reduce feedback overhead. In the following sections, we will review lossless and lossy algorithms and their applications in feedback compression.

5.1.1 Lossless Compression Algorithms

Lossless compression explores the redundancy of statistical data, and performs data compression. The original data can be recovered without causing any distortion. However, the compression rate is subject to the limit of theoretical statistical redundancy, and is usually 2:1 to 5:1. Common lossless compression algorithms include Huffman coding, Run Length Encoding (RLE), and Lempel-Ziv-Welch coding (LZW).

5.1.1.1 Huffman Coding and Compression Algorithms

Huffman coding is proposed by David A. Huffman in 1952, it is a data compression algorithm base on probability of signal. It is entirely based on the probability of character to construct the shortest average prefix. This famous coding method in data compression is known as optimal coding.

Huffman coding can be described as follow: In variable length coding, if the code length is in reverse chronological order strictly according to the probability of corresponding symbol, its average length is the smallest. Theoretical studies have shown that the compression ratio is close to the upper limit. This method have been extensively studied and applied to image, text and video compression, such as MPEG, JPEG standard.

Compression ratio of Huffman coding is usually approximately 2 to 3 for text data, 1 to 2 for executable file. Compared with the lossy compression methods such as wavelet compression, compression ratio of Huffman encoding is relatively low. Therefore, Huffman coding does not apply to files in which the majority of symbols repeat a lot. Run-length encoding (RLE) and other coding will get a higher compression ratio.

In literatures [72] and [73], the compression rate of Huffman feedback compression algorithm is not high due to the feedback values are evenly distributed. To overcome this problem, the author utilizes the correspondence in time and frequency domain of feedback information in OFDM system, and proposes two lossless compression algorithms based on Huffman algorithm: Iterative Time-Frequency Algorithm and Block Time-Frequency Algorithm.
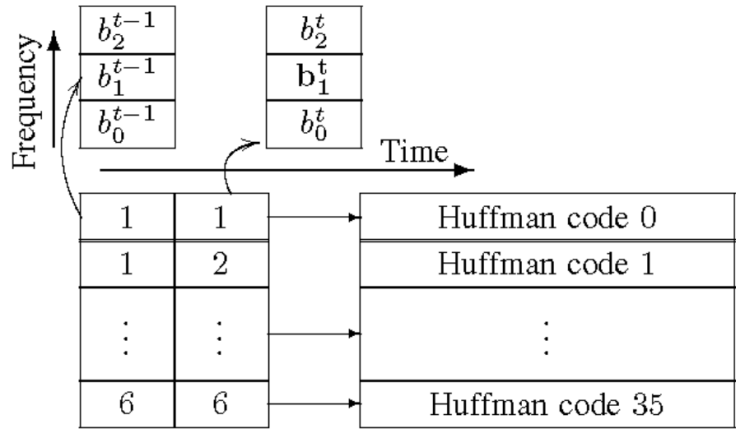
Figure 5.1   Flowchart of iterative time-frequency algorithm
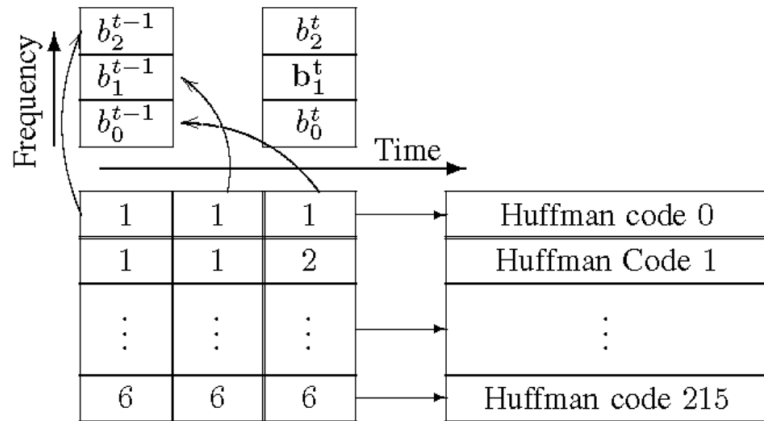
Figure 5.2   Flowchart of block time-frequency algorithm

In Figures 5.1 and 5.2, b_n^t represents the modulation order of the n-th cluster at time frame t.

For Iterative Time-Frequency Algorithm, the feedback information is the difference of modulation of current cluster from the previous cluster in time domain and the adjacent clusters in frequency domain. For Time-Frequency Algorithm, the feedback information is the difference of modulation of current cluster from the

previous cluster in time domain and the adjacent cluster in time domain of the previous cluster.

Because of channel correlation in time and frequency domains, the modulation difference of the current cluster from an adjacent cluster is usually 0, 1 and −1. Thus the feedback values are concentrated. Simulation results show that these two improved Huffman feedback compression algorithms greatly reduce the amount of feedback information.

5.1.1.2 Run Length Encoding and Feedback Compression Algorithm

Run length encoding is a compression method utilizes the space redundancy. It is a lossless compression coding method belongs to statistical encoding. The algorithm idea is simple. When a string of consecutive repeated characters appears in the data, only need the characters, length and location of the string to recover the original data. Data structure of RLE is shown in Figure 5.3.

Figure 5.3   RLE coding structure

In the data structure diagram, D is compressed character, L is the number of repeated D, S is the character not used in the data character set $\{D\}$, $S \notin \{D\}$. A string "$DD……D$" of L length in the data stream is shown in the figure. Compression can be achieved only when $L \geq 3$. Therefore encoding must first determine the value
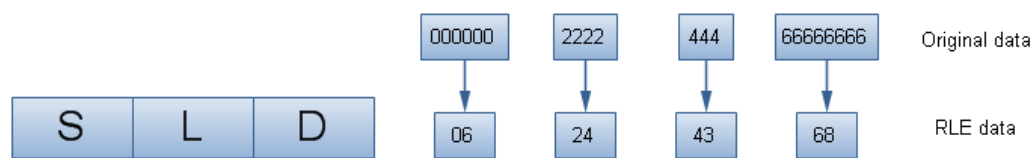
of *L*, and then decide whether to use RLE. During decoding, first check the read data

is *S* in order to determine the next data is length or character.

Traditional Best-M program does not take into account the correlation

between adjacent sub-carriers. Good performance sub-carriers may appear in a cluster,

and form a continuous high SINR sub-band. Based on this fact, literature [74]

proposes codebook construction method base on RLE.

In the literature, *v(b, R)* represents an $N \times 1$ vector. A run begins from position

b and length is *R*, i.e., for *v(b, R)*, value in position *b*, *b* + 1, … , *b* + *R*−1 equals 1, and

0 elsewhere. Proposed codebook consists of vectors of the form

$$f = S_0 v(0, N) + \sum_{i=1}^{P} \Delta_i v(b_i, R) \tag{5.1}$$

Intuitively, the feedback vector takes the form of a base-SINR S_0 added to *P*

runs each of magnitude *Δi*. Thus, a code vector is specified by (1) the quantized

base-SINR S_0, (2) the number of run-vectors *P*, (3) the magnitude *Δi*, or *Δi* start

position b_i and the length R_i of each run. By restricting the above quantities, a limited

codebook can be obtained. Simulation results show that under the proportional fair

scheduling algorithm, this RLE based scheme reduces feedback overhead, and the

performance is better than HGR scheme.

5.1.1.3 LZW Coding and Feedback Compression Algorithm

LZW algorithm utilizes the characteristic of duplicate code in data. It uses a

simple code of previously appeared string to replace the same string content

afterwards, and achieves data compression. It adaptively builds a dictionary during the compression process to reflect the correspondence between the string and code. Code dictionary is used to determine the compression code of output string.

LZW coding effectively uses the repetitive character, redundancy of character frequency and high utilization rate mode. It does not require a priori information on the statistics of input data, only needs to scan the data once. Its computation time is proportional to the length of the data. Literature [75] applies the LZW algorithm to wireless communication systems.

Literatures [76], [77], [78] introduce applications of other lossless compression algorithms in feedback compression of wireless communication systems.

5.1.2 Lossy Compression Algorithms

Lossy compression allows certain loss of information in the compression process. Although the original data cannot be completely restored, the lost part does not affect much the understanding of the original information. In return a much larger compression ratio can be achieved. Lossy compression is widely used in voice, image and video data compression. For example, the main applications of DCT is JPEG image compression standard. DWT has been applied in JPEG 2000 image compression standard. The principles of these two compression algorithms are introduced as follow.

5.1.2.1 DCT Transform and Related Feedback Compression Algorithms

DCT is closely related to Fourier transform. In Fourier series expansion, if the expanded function is a real even function, then its Fourier series only contain cosine terms. Cosine transform can be derived after discretization. This is called the discrete cosine transform.

Discrete Cosine Transform is an orthogonal transformation method proposed by N. Ahmed et al. in 1974. It is often considered to be a good solution to transform voice and image signals. In recent years the development of digital signal processing (DSP) chip, coupled with the advantages of application-specific integrated circuit (ASIC) design, make DCT establishing an important position in image coding, such as H.261, JPEG, MPEG and other international coding standard. In video compression, the most commonly used transform is DCT. DCT is considered to be a near-optimal transform which performance is close to Karhunen- Loève Transform (KLT) [79], [80].

The reason DCT can be used for compression is the energy coefficients after transform are mainly concentrated in the low-order DCT coefficients, and some of the discarded higher order coefficients will not cause a large distortion. This is because the original data is usually relevant and has a degree of redundancy.

Let us take one-dimensional DCT transform as an example to introduce the principles of DCT-based Channel Quality Information (CQI) compression:

Suppose the original CQI feedback information is $\gamma = [\gamma_1, \gamma_2, \cdots, \gamma_{N_c}]$. Apply DCT to the original information, we get

$$[DCT(\gamma)]_p = w(p)\sum_{q=1}^{N_c}\gamma(q)\cos\frac{\pi(2q-1)(p-1)}{2N_f} \tag{5.2}$$

in which $p = 1, \ldots, N_c$. $[.]_p$ represents the pth element if the vector, and

$$w(p) = \begin{cases} \dfrac{1}{\sqrt{N_f}} & p = 1 \\ \sqrt{\dfrac{2}{N_f}} & 2 \le p \le N_c \end{cases} \tag{5.3}$$

Discard coefficient which amplitude is close to zero in DCT (γ_i) to achieve compression and shorter compressed sequence. This compressed sequence is feedback to the base station through feedback link. The base station uses 0 to replace the lost high frequency coefficients due to compression.

$$y = \{[DCT(\gamma)]_p \cdot l([DCT(\gamma)]_p \ge T_c)\}_{p=1}^{N_c} \tag{5.4}$$

in which $l(x) = \begin{cases} 1 & x > 0 \\ 0 & else \end{cases}$. T_c is the compression threshold. Performing DCT on y at the user's side, we get the reconstructed CQI sequence.

Literature [81] proposes a new data compression based CQI feedback scheme. The author considers the channel correlation in frequency domain, and performs DCT on the feedback data in frequency domain. The author also considers the correlation in

time domain, and samples the feedback data in time domain. The discarded data is estimated using MMSE estimator. Simulation results show that this method can reduce 5 to 30 times the amount of feedback information without losing system throughput.

Literatures [82], [83], [84], [85], [86] also introduce current DCT feedback compression algorithms in OFDM systems.

5.1.2.2 Wavelet Transform and Related Feedback Compression Algorithms

Wavelet transform is a signal processing method. It is a development of the traditional Fourier transform. Wavelet multi-resolution analysis has good frequency characteristics. By using gradually refined time-domain step length for high-frequency, it can focus on the details of the analyzed subject. Thus it is suitable for non-stationary signal processing. Literature [87] discusses compression algorithms based on wavelet transform. By multi-resolution analysis on the input data, the original signal is decomposed to sub-band signals of different frequency, and then compressed individually.

Formula of one-dimensional orthogonal multi-resolution analysis is

$$f(t) = \sum_{k \in Z} c_{0,k} \phi_{0,k}(t) = \sum_{k \in Z} c_{J,k} \phi_{J,k}(t) + \sum_{j=1}^{J} \sum_{k \in Z} d_{j,k} \psi_{j,k}(t) \quad J \geq 1 \tag{5.5}$$

in which $c_{0,k}$ is the original data before compression $\begin{cases} c_{J,k} = \langle f, \phi_{J,k} \rangle & k \in Z \\ d_{j,k} = \langle f, \psi_{j,k} \rangle & k \in Z \end{cases}$ represents low frequency coefficients and high frequency coefficients of each level

78

after J-level wavelet decomposition. $\phi(t)$ is the wavelet scaling function. $\psi(t)$ is the wavelet function. After J-level wavelet decomposition of the data, we can get the low-frequency coefficient sequence $c_{J,k}$, and high-frequency coefficient sequence $d_{j,k}$ ($j = 1, 2, ..., J$). Where $c_{J,k}$ represents Low-frequency part of $c_{0,k}$. $d_{j,k}$ ($j = 1, 2, ..., J$) does not overlap and shows the details of the different frequency bands. If the original data $c_{0,k}$ has strong correlation, then its energy is mainly concentrated in the low frequency part, and band is narrow. The other energy distribution is in the high frequency, most of the values of $d_{j,k}$ ($j = 1, 2, ..., J$) are close to 0. Normally, discarding these high frequency coefficients close to 0 will not have significant impact on rebuilding data.

Similarly, formula of two-dimensional orthogonal multi-resolution analysis is

$$
\begin{aligned}
f(x, y) &= \sum_{k,m} c_{0,k}\phi_{0,k}(x, y) \\
&= \sum_{k,m} c_{k,m}^{J}\phi_{J,k,m} + \sum_{j=1}^{J}\sum_{k,m} d_{k,m}^{j,1}\psi_{j,k,m}^{1} + \sum_{j=1}^{J}\sum_{k,m} d_{k,m}^{j,2}\psi_{j,k,m}^{2} + \sum_{j=1}^{J}\sum_{k,m} d_{k,m}^{j,3}\psi_{j,k,m}^{3}
\end{aligned}
\tag{5.6}
$$

in which, $J \geq 1$. $\begin{cases} \psi^{1}(x, y) = \phi(x)\psi(y) \\ \psi^{2}(x, y) = \phi(x)\psi(y) \\ \psi^{3}(x, y) = \phi(x)\psi(y) \end{cases}$ are horizontal, vertical, diagonal wavelet

functions. Figure 3.4 shows an example of two-dimensional wavelet decomposition.
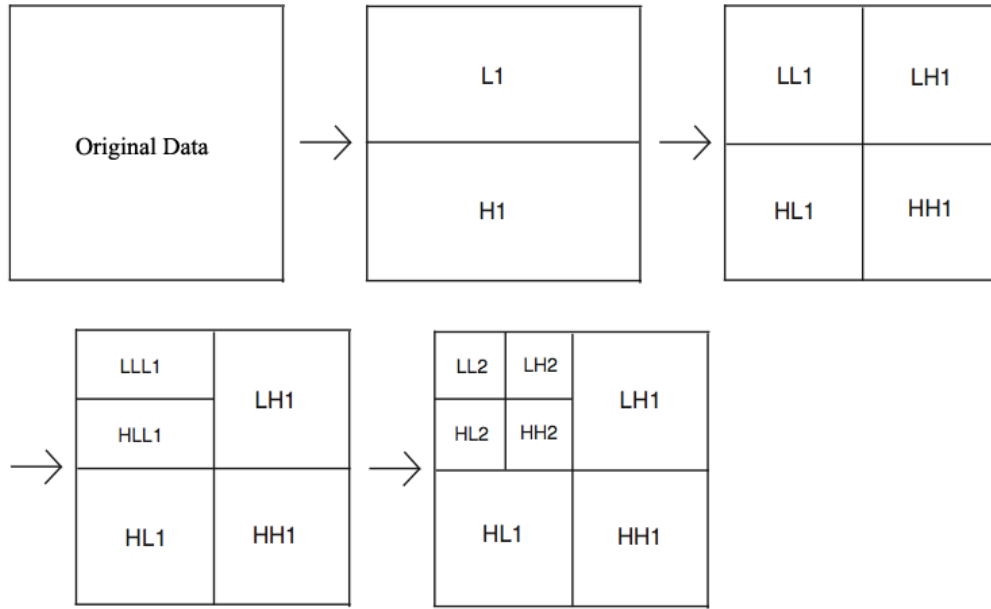
Figure 5.4   Two-dimensional wavelet decomposition

Two-dimensional wavelet transforms the original data matrix to a low-frequency part and three high-frequency parts in the horizontal, vertical and diagonal directions. LL reflects the basic characteristics of the original data, which has the majority of energy. HL, LH and HH are the horizontal, vertical and diagonal high-frequency component respectively. They represent the edge information in horizontal, vertical and diagonal direction. Decomposition can be continued in the same way for low-frequency matrix.

We take one-dimensional wavelet transform as an example to introduce the principle of DWT-based CQI compression: First the client's original CQI data is J-level wavelet decomposited in accordance with equation (5.4). Low-frequency coefficients $c_{J,k}$ containing the original data were maintained. By discarding those

high frequency coefficients $d_{j,k}$ whose amplitude is close to zero. Data is compressed. A shorter compressed sequence is obtained. This compressed sequence is fed back through the feedback link to the base station side. On the base station side, some high frequency coefficients lost due to compression will be replaced with 0. Then we get

$$\tilde{d}_{j,k} = d_{j,k} \cdot l(|d_{j,k}| \geq T_J) \quad j = 1, \cdots, J \qquad (5.7)$$

in which T_j (j = 1, 2, …, J) is the compression threshold. User perform wavelet transform on $c_{J,k}$ and $\tilde{d}_{j,k}$ ($j = 1, 2, \cdots, J$), and obtain a reconstructed CQI sequence.

In order to weigh the performance of DWT-CQI compression algorithm, we define two parameters:

Coefficient compression rate $\quad \alpha = \sum_{j=1}^{J} \sum_{k=1}^{N_j} l(|d_{j,k}| \leq T_J) \Big/ (\sum_{j=1}^{J} N_j + N_J)$

Energy reservation ratio $\quad \beta = \sum_{j=1}^{J} \sum_{k=1}^{N_j} |\tilde{d}_{j,k}|^2 + \sum_{k=1}^{N_j} |c_{J,k}|^2 \Big/ (\sum_{j=1}^{J} \sum_{k=1}^{N_j} |d_{j,k}|^2 + \sum_{k=1}^{N_j} |c_{J,k}|^2)$

in which N_j presents the number of high frequency coefficients for j level wavelet decomposition. α and β show the effect of DWT-CQI compression algorithm . The larger α is, the more feedback overhead is reduced. The larger β is, the smaller the distortion is after information reconstruction.

In literatures [88, 89], DWT feedback compression algorithm called Distributed-Haar CQI Feedback is proposed. It is an improved Best-M Haar algorithm. The author considers time domain correlation of CQI feedback, and divides all sub-bands into N_G groups. The grouping method is known to the users and base

station in advance. In each feedback time interval, only one certain group of CQI information of N_G group sub-band is fed back. The feedback method is Best-M haar compression algorithm. After N_G times of feedback, CQI information of all sub-bands are fed back. Simulation results show that when the speed of a mobile user is 3km/h and 15km/h, the distributed method used in this literature can effectively reduce the feedback overhead. Compared with the Best-M individual method and DCT partitioning method, the performance of DCT significant method is better when the TTI feedback overhead is the same.

5.2 Feedback Compression Algorithms over Frequency Selective Fading Channel

From the above analysis we can see that the compression algorithm can effectively reduce the feedback overhead in OFDM systems. Generally, lossy compression algorithms can achieve higher compression rates than lossless compression algorithms.

Multipath delay spread makes of the channel SNR correlated in the frequency domain. It provides a theoretical base for the algorithms that reduce system overhead. However, in some complex communication environment, the channels of the system may be severe frequency-selective. Correlations between adjacent sub-carriers are relatively small. In this case, traditional feedback compression algorithms are difficult to obtain good compression results.

In this section a DCT feedback compression algorithm based on sorting is proposed. Sorting is performed before the feedback compression of channel CQI. This

sorting can increase the correlation of the original data; therefore the effect of compression algorithms is enhanced.

The CQI information becomes more correlated after sorting at client's end. The overhead of feedback compression algorithm is also reduced. However, the sequence information at client's end must be transmitted to the base station to correctly recover the CQI. This also increases the feedback overhead of the system. This algorithm is practically effective only when overhead reduced from compression algorithms is greater than the overhead increased from the of sequence information. We assume that each channel coefficient and DCT compression coefficient occupy B bits. The number of CQI processed by DCT feedback compression algorithm each time is N. DCT coefficients retained after processing using original DCT feedback compression algorithm is M. DCT coefficients retained after processing using sorting DCT feedback compression algorithm is M'. We have the following equations.

Feedback compression rate of original DCT algorithm $c_1 = M/N$.

Feedback compression rate of sorting DCT algorithm

$$c_2 = \left(M'B + \lceil \log_2 N! \rceil\right)/(NB)$$

Let $B = 8$, we have $c_1 - c_2 = 8(M - M') - \lceil \log_2 N! \rceil/(8N)$. Therefore $\lim_{N \to 0}(c_1 - c_2) = -\lim_{N \to 0} \ln(N!)/(8N \ln 2)$. According to Stirling's approximation for sorting algorithms, $\ln(n!) \approx \frac{1}{2}\ln(2\pi n) + n\ln(n) - n = \frac{1}{2}\ln(2\pi) + (n + \frac{1}{2})\ln n - n$.

Then

$$\lim_{N \to 0}(c_1 - c_2) = -\lim_{N \to 0}[(N + 1/2)\ln N - N]/(8N\ln 2) = -\infty \qquad (5.8)$$

When the number of sub-carriers N increases, the codebook of the feedback sequence information is exponential growth. This is adverse to reduce the overhead for the sorting algorithm. Therefore an appropriate value of N is needed.

$$N = 4, \quad c_1 - c_2 = 8(M - M') - \lceil \log_2 4! \rceil/(8 \times 4)$$

$$= [8(M - M') - 5]/(8 \times 4) \qquad (5.9)$$

$$N = 8, \quad c_1 - c_2 = 8(M - M') - \lceil \log_2 8! \rceil/(8 \times 8)$$

$$= [8(M - M') - 16]/(8 \times 8) \qquad (5.10)$$

$$N = 16, \quad c_1 - c_2 = 8(M - M') - \lceil \log_2 16! \rceil/(8 \times 16)$$

$$= [8(M - M') - 45]/(8 \times 16) \qquad (5.11)$$

When $N = 4$, there are only 4 DCT transform coefficients. The compression rate is low. When $N = 16$, the code book of sequence information already takes up 35% of the feedback overhead. Therefore we select $N = 8$.

5.3 Simulation and Performance Analysis

We employ a multiuser OFDM system for simulation. Max-SNR scheduling is used as the multiuser diversity scheduling algorithm. The system channel bandwidth is 2MHz. The number of users is 6. The multipath number $L = 6$. Number of sub-carriers $N_c = 64$. The length of cyclic prefix (CP) is 16. Matlab is used in the simulation.

We change the power distribution and the delay distribution of 6 multipath signals of the channel, resulting in four frequency selective channels A, B, C and D which frequency selectivity reduces in order. A represents an extremely serious frequency selective fading channel, and D represents a minor frequency selective fading channel. Figure 5.5 to Figure 5.8 show comparisons of BER performance and feedback overhead of the sorting DCT algorithms and original DCT algorithms under those different frequency selective channels. In the legends of the figures, "Best" represents the system's BER performance when channel completely feedback.
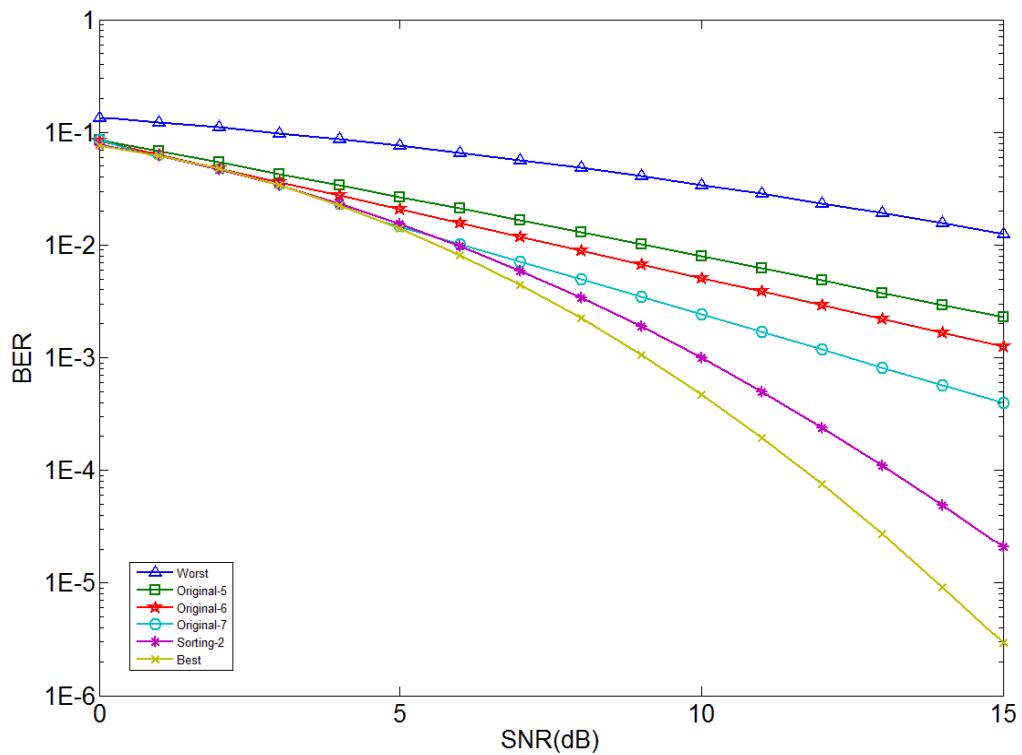
Figure 5.5   Performance of DCT feedback compression algorithms under channel condition A

"Worst" represents the system's BER performance when channel has no feedback. "Sorting" represents the system's BER performance for sorting DCT algorithm when *M* equals a certain number. "Original" represents the system's BER performance for the original DCT algorithm when *M* equals a certain number.

It can be seen from Figure 5.5 the performance of algorithm Sorting2 is better than Original5 to Original7. That is, for each group 8 CQI data, sorting DCT algorithm only needs to feedback two coefficients and its performance is better than
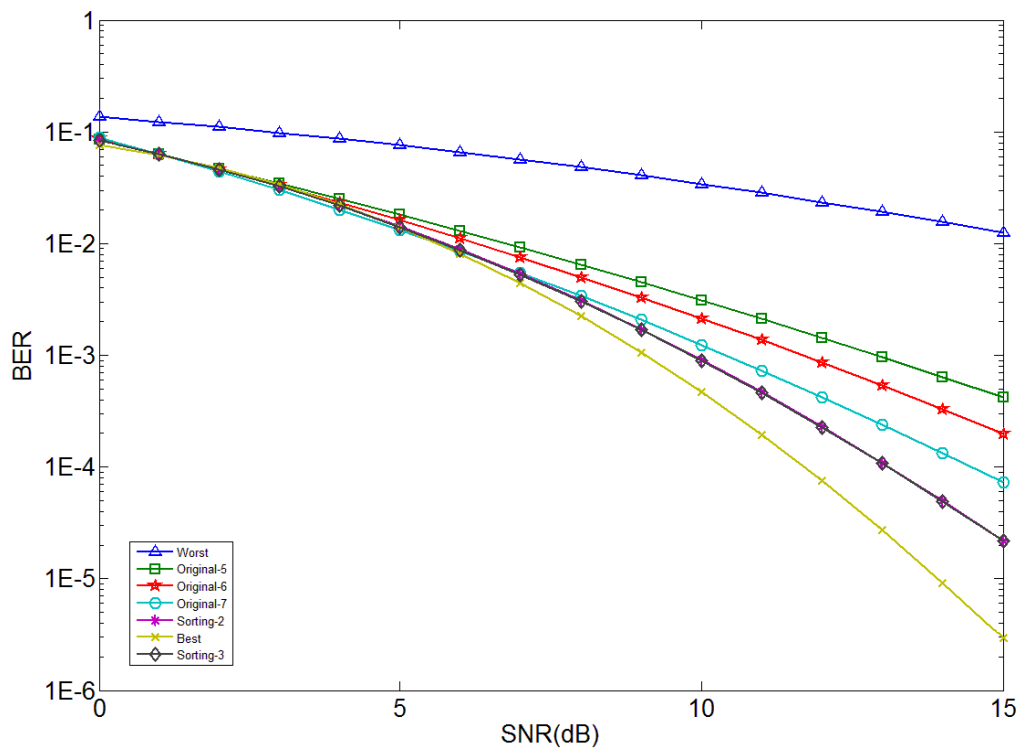
Figure 5.6   Performance of DCT feedback compression algorithms under channel condition B

the original DCT algorithm which uses 7 coefficients. The correlation in data increased a lot after sorting. Compression rate of algorithm Sorting 2 is 50% compared to compression rate of 87.5% for algorithm Original7.

Compare Figures 5.5, 5.6 and 5.7, we can see with the reduction of the channel frequency selectivity, the correlation between adjacent sub-carriers is increased. The performance of sorting DCT algorithms and original DCT algorithms are improved, and the performance gap is reduced.
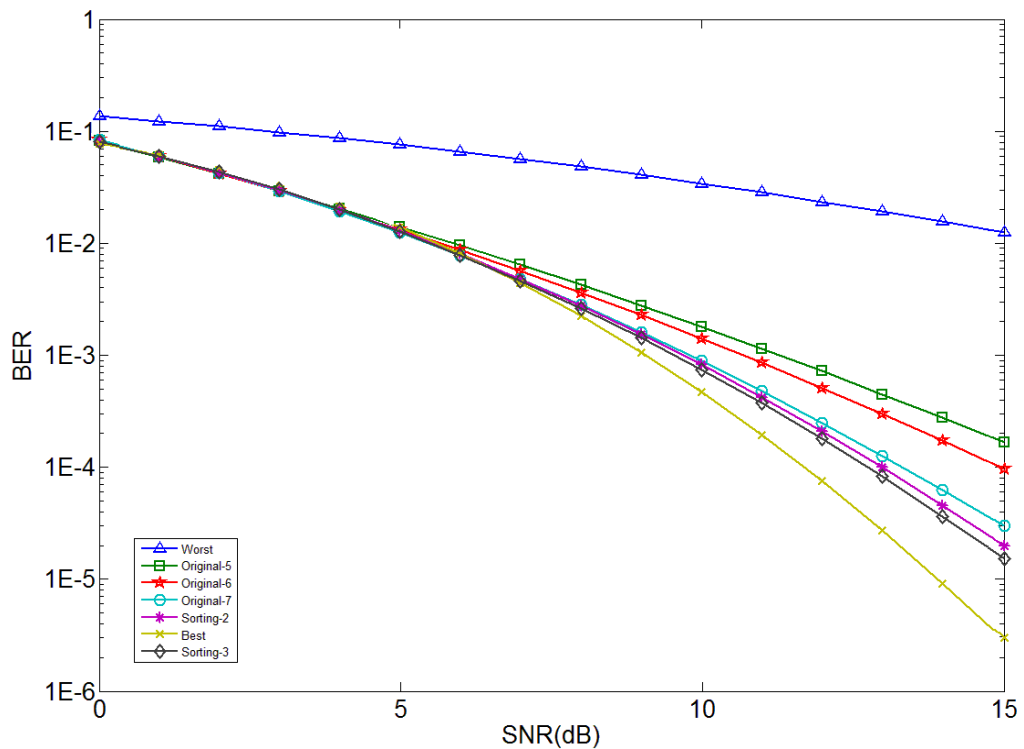
Figure 5.7   Performance of DCT feedback compression algorithms under channel

condition C

It can be seen from Figure 5.8 that performances of the two sorting DCT algorithms and three original DCT algorithms are similar under channel condition D. This is because the frequency selectivity is weak, and correlations between sub-carriers are large. Original DCT algorithm is able to achieve relatively good compression performance. Although the data has not been sorted, more DCT coefficients are fed back which enhances its performance. Also in this figure, compression rate of algorithm Sorting2 is 50%. Compression rate of algorithm
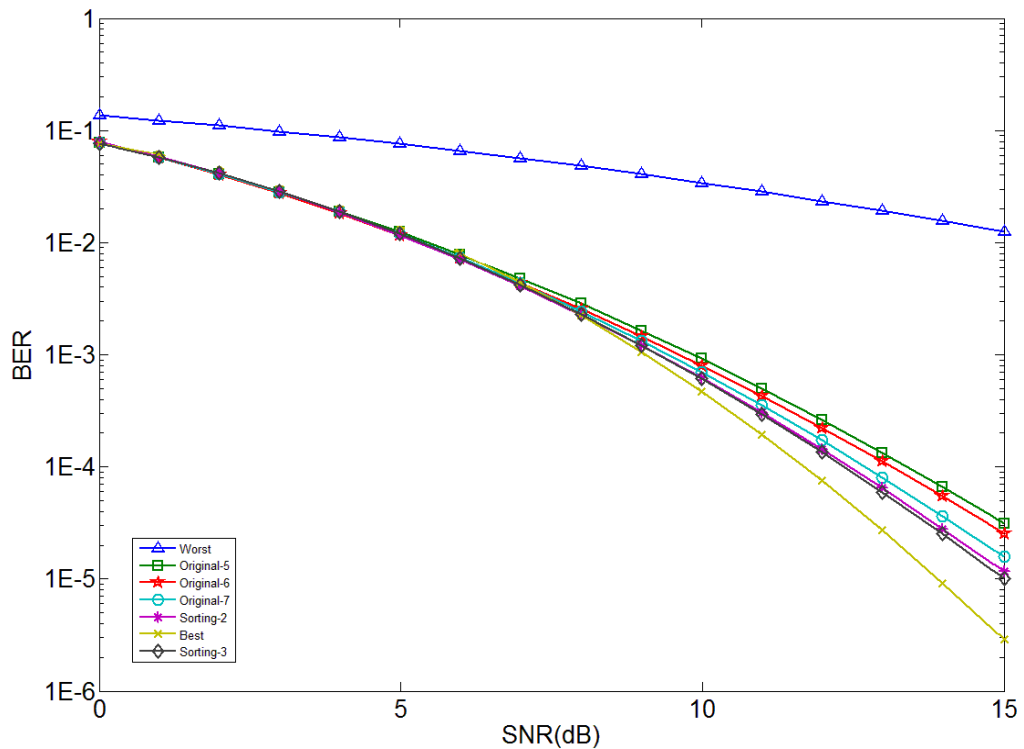
Figure 5.8   Performance of DCT feedback compression algorithms under channel condition D

Original5 is 62.5%. It is clear that if we further reduce the frequency selectivity of the channel, i.e. the channel tends to flat fading, then performance of sorting DCT algorithm will be lower than the original DCT algorithms, because the sequence information overhead $\lceil \log_2 N! \rceil$ cannot be reduced.

5.4 Summary

In this chapter, we first analyze the limited feedback algorithms and channel correlation in frequency domain and time domain. After introduction of three existing lossless compression algorithms and two lossy compression algorithms, and their application in OFDM systems' feedback compression, a DCT compression algorithm based on sorting is proposed. Appropriate data block size for this algorithm is analyzed which helps reducing the feedback overhead of the system.

CHAPTER VI   CONCLUSIONS

6.1 Dissertation Summary

The work mainly researched bit and power allocation for OFDM systems, joint resource allocation, user scheduling, and limited feedback problem in multiuser OFDM systems.

Chapter 3 shows research on resource allocation in OFDM system. First, bit and power allocation algorithms are analyzed for single-user OFDM systems. These algorithms are usually more complex. We analyze the characteristics of the optimal bit and power allocation scheme in OFDM systems. Then a low complexity bit and power allocation algorithm is designed in accordance with these characteristics. The results show that the proposed algorithm can effectively reduce the computational complexity while still ensure the system performance. Second, resource allocation for multiuser OFDM system is studied in this chapter. Real systems with user fairness problem are considered. In these systems, each user has a minimum transfer rate requirement. Lagrange multiplier method is derived and a low complexity sub-carrier, power, and bit allocation algorithm is proposed. This algorithm has lower computational complexity and results in performance that is comparable to that of the existing algorithms.

In Chapter 4, first, both conventional and fair user scheduling schemes are introduced and analyzed. Then variations of proportional fair scheduling scheme are

proposed and analyzed. Simulations on performance are performed and show the significant advantages of proposed proportional fair schemes over the conventional ones. The proposed schemes improve system throughput and delay time. Normalized adaptive proportional fair scheme achieves higher throughput without sacrificing fairness which makes it a better one in terms of efficiency and fairness.

One application of the widely studied compression algorithms in recent years is feedback compression in wireless communication systems. Chapter 5 introduces existing compression algorithms including three lossless compression algorithms and two lossy compression algorithms, and their application in OFDM systems. Then a DCT feedback compression algorithm based on sorting is proposed. This algorithm uses sorting algorithm to increase the correlation between feedback CQI of frequency selective channel. Then a suitable size of data block is analyzed and calculated. So the feedback overhead of the system is successfully reduced.

6.2 Future Work

In existing resource allocation scheme for OFDM systems, a number of algorithms introduce efficient and low complexity resource allocation schemes for different user fairness. However, how to find reasonable user fairness criteria and guidelines according to real system requirements, and how to design an efficient resource allocation scheme, are still worth further study. In existing literatures, mostly description of user fairness are maximize user's minimum rate, proportional fairness, and setting a minimum transfer rate requirement for each user. However, in the real

system, due to the different type of wireless communication services provided by individual users, system fairness should be considered from the following three aspects: user's type of service, such as data service, multimedia service, etc., user and type of service priority, as well as overall system performance. Especially when users are carrying different types of services, their requirements for delay and reliability of wireless transmission are different. Therefore, future research on more reasonable and practical user fairness criteria for this situation is necessary. And a resource allocation scheme can be designed under this criteria for wireless OFDM systems.

For MIMO-OFDM systems, researchers currently focus on how to design a reasonable resource allocation and user scheduling scheme in time, frequency and spatial three dimensions. This dissertation studies the resource allocation and multiuser scheduling scheme in OFDM systems. In future research, schemes in this dissertation can be extended to multiuser MIMO-OFDM systems. Compared with MIMO system, each user in MIMO-OFDM systems experiences different multi-antenna channel gain in different sub-carriers. Thus user scheduling and resource allocation become more complex. In this case, the user scheduling scheme needs to consider the MIMO channel characteristics of the same user on different sub-carriers, and the orthogonality of different users on the same sub-carrier. Following this direction, research can be started on user scheduling and resource allocation scheme for multiuser MIMO-OFDM systems. The limited feedback algorithm considered in this dissertation exploits correlation of feedback in frequency

domain and spatial domain. How to utilize the correlation in time domain to save even

more system feedback overhead is to be further studied.

In 4G standard, the standard for scheduling and limited feedback scheme for

multiuser MIMO systems is still not determined. Therefore, in the existing framework

of LTE wireless transmission, how to design effective user scheduling and limited

feedback schemes for multiuser MIMO system is also worthy of concern.

LIST OF REFERENCES

1 H. Yang, "A road to future broadband wireless access: MIMO-OFDM based air interface," IEEE Communication Magazine, vol. 43, no. 1, pp. 53-60, Jan. 2005.

2 M. Torabi, W. Ajib, and D. Haccoun, "Multiuser scheduling for MIMO-OFDM systems with continuous-rate adaptive modulation," IEEE Wireless Communications and Networking Conference, pp. 946-951, Mar./Apr. 2008.

3 K. Yungsoo, J. J. Byung, and C. Jauehak, "Beyond 3G: vision, requirements, and enabling technologies," IEEE Communication Magazine, vol. 41, pp. 120-124, Mar. 2003.

4 S. Uskela, "Key concepts for evolution toward beyond 3G networks," IEEE Transactions on Wireless Communication, vol. 10, pp. 43, 48, Feb. 2003.

5 M. Munoz and C. G. Rubio, "A new model for service and application convergence in B3G/4G networks," IEEE Personal Communication, vol. 2, pp. 6-12, Oct. 2004.

6 J. A. C. Bingham, "Multicarrier modulation for data transmission：an idea whose time has come," IEEE Communication Magazine, vol. 28, pp. 5-14, May 1990.

7 Z. Han and K. J. R. Liu, Resource allocation for wireless networks: basics, techniques, and applications, New York: Cambridge University Press, pp. 178-179, 2008.

8 T. M. Cover and J. A. Thomas, Elements of information theory, New York: Wiley& Sons, 1991.

9 J. K. Cavers, "Variable rate transmission for Rayleigh fading channels," IEEE Transactions on Communication, vol. 20, pp. 15-22, Feb. 1972.

10 W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," IEEE Transactions on Communication, vol. 43, pp. 2223-2230, Jul. 1995.

11 A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," IEEE Transactions on Information Theory, vol. 43, pp. 1986-1992, Nov. 1997.

12 A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels," IEEE Transactions on Communication, vol. 45, pp. 1218-1230, Oct.

1997.

13 T. Keller and L. Hanzo, "Adaptive modulation techniques for duplex OFDM transmission," IEEE Transactions on Vehicular Technology, vol. 49, pp. 1893-1906, Sep. 2000.

14 D. Hughes Hartogs, "Ensemble modem structure for imperfect transmission media," U. S. Patents No. 4679227, Jul. 1987.

15 P. S. Chow, J. M. Cioffi, and A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," IEEE Transactions on Communication, vol. 43, pp. 773-775, Feb. 1995.

16 F. H. Fischer and J. B. Huber, "A new loading algorithm for discrete multitone transmission," Global Telecommunications Conference, 1996. 'Communications: The Key to Global Prosperity, pp. 724-728, Nov. 1996.

17 A. Leke and J. M. Cioffi, "A maximum rate loading algorithm for discrete multitone modulation systems," Global Telecommunications Conference, 1996. 'Communications: The Key to Global Prosperity, vol. 3, pp. 1514-1518, Nov. 1996.

18 A. Fasano, "On the optimal discrete bit loading for multicarrier systems with constraints," 57th IEEE Semiannual Vehicular Technology Conference, vol. 2, pp. 915-919, Apr. 2003.

19 M. Lei, P. Zhang, H. Harada, and H. Wakaua, "An adaptive power distribution algorithm for improving spectral efficiency in OFDM," IEEE Transactions on Broadcasting, vol. 50, pp. 347-351, Sep. 2004.

20 B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," IEEE Transactions on Communication, vol. 48, pp. 23-27, Jan. 2000.

21 A. M. Wyglinski, F. Labeau, and P. Kabal, "Bit loading with BER-constraint for multicarrier systems," IEEE Transactions on Wireless Communication, vol. 4, pp. 1383-1387, Jul. 2005.

22 C. Y. Wong, R. S. Chen, K. B. Latief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," IEEE Journal on Selected Areas in Communication, vol. 17, pp. 1747-1758, Oct. 1999.

23 J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," IEEE Journal on Selected Areas in Communication, vol. 21, pp. 171-178, Feb. 2003.

24 C. Y. Wong, C. Y. Tsui, R. S. Cheng, and K. B. Letaief, "A real-time subcarrier allocation scheme for multiple access downlink OFDM transmission," 50th IEEE Vehicular Technology Conference, vol. 2, pp. 1124-1128, 1999.

25 Y. F. Chen, J. W. Chen, and C. P. Li, "A fast suboptimal subcarrier, bit, and power allocation algorithm for multiuser OFDM-based systems," IEEE International Conference on Communications, vol. 4, pp. 3212-3216, 2004.

26 K. D. Choe, Y. J. Lim, and S. K. Park, "Subearrier adaptation for multiuser OFDM systems," IEEE Global Telecommunications Conference, vol. 2, pp. 1230-1233, 2004.

27 W. Rhee and J. M. Ciofii, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," 51st IEEE Vehicular Technology Conference, vol. 2, pp. 1085-1089, 2000.

28 Z. K. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," IEEE Transactions on Wireless Communication, vol. 4, pp. 2762-2737, Nov. 2005.

29 Z. Hart, Z. Ji, and K. J. R. Liu, "Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions," IEEE Transactions on Communication, vol. 53, pp. 1366-1376, Aug. 2005.

30 Y. J. Zhang and K. B. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems," IEEE Transactions on Wireless Communication, vol. 3, pp. 1566-1575, Sep. 2004.

31 I. E. Telatar, "Capacity of multi-antenna Gaussian channels," European Transactions on Telecommunications, vol. 10, pp. 585-595, Nov./ Dec. 1999.

32 N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple access and broadcast channel," IEEE Transactions on Theory, vol. 50, pp. 768-783, May 2004.

33 S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality achievable rates and sum rate capacity of Gaussian MIMO broadcast channel," IEEE Transactions on Information Theory, vol. 49, pp. 2658-2668, Oct. 2003.

34   H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," International Symposium on Information Theory, pp. 174-174, 2004.

35   M. H. M. Costa, "Writing on dirty paper," IEEE Transactions on Information Theory, vol. 29, pp. 439-441, May 1983.

36   Q. H. Spencer, A. Lee, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," IEEE Transactions on Signal Processing, vol. 52, pp. 461-471, Feb. 2004.

37   K. K. Wong, R. D. Murch, and K. B. Latief, "A joint-channel diagonalizaiton for multiuser MIMO antenna systems," IEEE Transactions on Wireless Communication, vol. 2, pp. 773-786, July 2003.

38   L. U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," IEEE Transactions on Wireless Communication, vol. 3, pp. 20-24, Jan. 2004.

39   T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," IEEE Journal on Selected Areas in Communication, vol. 24, pp. 528-541, Mar. 2006.

40   J. Kim, S. Park, J. H. Lee, J. Lee, and H. Jung, "A scheduling algorithm combined with zero-forcing beamforming for a multiuser MIMO wireless system," IEEE Vehicular Technology Conference, vol. 1, pp. 211-215, Sep. 2005.

41   Z. Shen, R. Chen, J. G. Andrews, R. W. H. Jr., and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," IEEE Transactions on Signal Processing, vol. 54, pp. 3658-3663, Sep. 2006.

42   G. Primolevo, O. Simeone, and U. Spagnolini, "Towards a joint optimization of scheduling and beamfoming for MIMO downlink," IEEE International Symposium On Spread Spectrum Techniques and Applications, pp. 493-497 Aug. 2006.

43   H. Lee, M. Shin, and C. Lee, "An eigen-based MIMO multiuser scheduler with partial feedback information," IEEE Communication Letters, vol. 9, pp. 328-330, Apr. 2005.

44   T. Kang and H. Kim, "An efficient random beamforming method based on beam subset selection," IEEE 18th International Symposium on Personal, Indoor and

Mobile Radio Communications, pp. 1-5, Sep. 2007.

45 M. Kountouris, T. Salzer, and D. Gesbert, "Scheduling for multiuser MIMO downlink channels with ranking-based feedback," EURASIP Journal on Advances Signal Processing, vol. 2008, no. 14, pp. 1-14, 2008.

46 T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna broadcast channels with limited feedback and user selection," IEEE Journal on Selected Areas in Communication, vol. 25, pp. 1478-1491, Sep. 2007.

47 K. Jagannathan, S. Borst, P. Whiting, and E. Modiano, "Scheduling of multiantenna broadcast systems with heterogeneous users," IEEE Journal on Selected Areas in Communication, vol. 25, pp. 1424-1434, Sep. 2007.

48 C. Swannack, G. W. Womell, and E. Uysal-Biyikoglu, "MIMO broadcast scheduling with quantized channel state information," IEEE International Symposium on Information Theory, pp. 1788-1792, July 2006.

49 M. Maddah-Ali, M. Ausari, and A. Khandani, "Broadcast in MIMO systems based on a generalized QR decomposition: signaling and performance analysis," IEEE Transactions on Information Theory, vol. 54, pp. 1124-1138, Mar. 2008.

50 B. Bandemer and S. Visuri, "Capacity-based uplink scheduling using long-term channel knowledge," IEEE International Conference on Communications, pp. 785-790, Jun. 2007.

51 D. Gesbert, L. Pittman, and M. Kountourm, "Transmit correlation-aided scheduling in multiuser MIMO networks," IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, pp. 249-252, May 2006.

52 M. Kountouris, D. Gesbert, and L. Pittman, "Transmit correlation-aided opportunistic beamforming and scheduling," in European Signal Processing Conference, Florence, Italy, pp. 4-8, Sept. 2006.

53 D. Hammarwall, M. Bengtsson, and B. Ottersten, "Beamforming and user selection in SDMA systems utilizing channel statistics and instantaneous SNR feedback," IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3, pp. 113-116, Apr. 2007.

54 N. Jindal, "MIMO broadcast channels with finite-rate feedback," IEEE Transactions on Information Theory, vol. 32, pp. 5054-5060, Nov. 2006.

55 H. Samra, H. Sun, and Z. Ding, "Capacity and linear precoding for packet

retransmissions," IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3, pp. 541-544, Mar. 2005.

56 H. Sun, J. H. Manton, and Z. Ding, "Progressive linear precoder optimization for MIMO packet retransmissions," IEEE Journal on Selected Areas in Communication, vol. 24, pp. 448-456, Mar. 2006.

57 Z. Shi, H. Sun, C. Zhao, and Z. Ding, "Linear precoder optimization for ARQ packet retransmissions in centralized multiuser MIMO uplinks," IEEE Transactions on Wireless Communication, vol. 7, pp. 736-745, Feb. 2008.

58 V. Raghavan and V. V. Veeravalli, "On quantized multiuser beamforming in spatially correlated broadcast channels," IEEE International Symposium on Information Theory, pp. 2041-2045, Jun. 2007.

59 B. Mielczarek and W. A. Krzymiefi, "Influence of CSI feedback errors on capacity of linear multiuser MIMO systems," IEEE Vehicular Technology Conference, pp. 2043-2047, Apr. 2007.

60 S. Catreux, V. Erceg, and D. Gebert, "Adaptive modulation and MIMO coding for broadband wireless data networks," IEEE Communication Magazine, vol. 40, no. 6, pp. 108-115, 2002.

61 A. Pandharipande, "Adaptive modulation for MIMO-OFDM systems," IEEE Vehicular Technology Conference, pp. 1266-1270, 2004.

62 Chang-Jun Ahn, "Channel ranking based joint symbols detection for MQRD-PCM/MIMO-OFDM," Consumer Communications and Networking Conference, pp. 1-5, 2009.

63 H. J. Zhu et al., "Scheduling schemes for multimedia service in wireless OFDM systems," IEEE Wireless Communications, vol. 14, pp. 99-105, Oct. 2007.

64 V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs", IEEE Transactions on Information Theory, vol. 45, no. 5, pp. 1456-1467, Jul. 1999.

65 H. Shin and J. Lee, "Performance analysis of space-time block codes over keyhole Nakagami-m fading channels," IEEE Transactions on Vehicular Technology, vol. 53, no. 2, pp. 351-362, Mar. 2004.

66 M. Torabi et al., "Multiuser scheduling for MIMO-OFDM systems with continuous-rate adaptive modulation," Wireless Communications and Networking

Conference, pp. 946-951, Mar. 2008.

67 F. Wang et al., "IEEE 802.16e System performance: analysis and simulations," 16th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 900-904, 2005.

68 A. Jalali, R Padovani and R. Pankaj, "Data throughput of CDMA-HDR, a high efficiency high data rate personal communication wireless system," IEEE 51st IEEE Vehicular Technology Conference, pp. 1854-58, May 2000.

69 P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," IEEE Transactions on Information Theory, vol. 48, no. 6, pp. 1277-1294, Jun. 2002.

70 C. H. Rentel et al., "Comparative forward link traffic channel performance evaluation of HDR and 1XTREME systems," IEEE Vehicular Technology Conference, vol. 1, pp. 160-164, 2002.

71 H. Kim et al., "An Efficient Scheduling Algorithm for QoS in Wireless Packet Data Transmission," 13th International Symposium on Personal, Indoor and Mobile Radio Communications, vol. 5, pp. 2244-2248, 2002.

72 V. P. G. Jimenez and A. G. Armada, "An adaptive MIMO-OFDM system：design and performance evaluation," 3rd International Symposium on Wireless Communication Systems, pp. 809-813, Sep. 2006.

73 V. P. G. Jimenez et al., "Methods for compression of feedback in adaptive multi-carrier 4G schemes," Wireless Personal Communications, vol. 47, no. 1, pp. 101-112, 2007.

74 B. Varadarajan et al, "Efficient channel quality feedback schemes for OFDMA systems with different schedulers," 69th IEEE Vehicular Technology Conference, pp. 1-5, Apr. 2009.

75 T. Lestable and M. Battelli, "LZW adaptive bit loading," IEEE International Symposium on Advances in Wireless Communications, Sep. 2002.

76 H. Nguyen, J. Brouet, V. Kumar, and T. Lestable, "Compression of associated signaling for adaptive multicarrier systems," 59th IEEE Vehicular Technology Conference, vol. 4, pp. 1916-1919, May 2004.

77 H. Nguyen, Hang and T. Lestable, "Compression of bit loading power vectors for adaptive multi-carrier systems," 47th Midwest Symposium on Circuits and

Systems, vol. 3, pp. iii-243-6, 2004.

78   H. Nguyen Nguyen and T. Lestable, "Network signaling compression for bit loading," IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, vol. 2, pp. 1170-1174, 2005.

79   K. Karhunen, "Über lineare methoden in der Wahrscheinlichkeitsrechnung," Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys., no. 37, pp. 1-79, 1947.

80   M. Loeve, Probability theory II (Graduate Texts in Mathematics), 4th ed., New York: Springer-Verlag, pp. 46, May 1978.

81   T. Eriksson and T. Ottosson, "Compression of feedback in adaptive OFDM based systems using scheduling," IEEE Communications Letters, vol. 11, pp. 859-861, Nov. 2007.

82   LG Electronics, "Further analysis on DCT based CQI reporting scheme," RI-062575, 2006.

83   Electronics LG, "Analysis on DCT based CQI reporting Scheme," R1-063174, 2006.

84   Huawei, "Overhead reduction of Best-M based CQI reporting," R1-063086, 2006.

85   Huawei, "System level evaluation of CQI compression schemes for E-UTRA," R1-071417, 2007.

86   Panasonic, "DCT partitioning for CQI reporting," R1-072796, 2007.

87   D. Taubman and M. Marcellin, JPEG2000: image compression, fundamentals, standards and practice, Dordrecht: Kluwer Academic Publishers, 2001.

88   A. Haghighat, Z. Lin, and G. Zhang, "Haar compression for efficient CQI feedback signaling in 3GPP LTE systems," IEEE Wireless Communications and Networking Conference, pp. 819-823, 2008.

89   InterDigital Communications Corporation, "Performance comparison of distributed Haar based compression for CQI feedback in E-UTRA," R1-072783, 2007.

VITA

NANSONG WU

EDUCATION

2008 to 2012
Ph.D.c Electrical and Computer Engineering
Florida International University

2006 to 2007
M. S. Electrical Engineering
Florida International University

2000 to 2004
B. S. Communication Engineering
Beijing University of Aeronautics and Astronautics

PUBLICATIONS AND PRESENTATIONS

N. Wu, Y. Fang, Y. Du, and K. Yen, "Multiuser Scheduling Schemes for MIMO-OFDM Systems", FAS Annual Meeting, March, 2010.

Y. Du, N. Wu, Y. Fang, and K. Yen, "A Comparison of different contrast functions of FastICA Algorithms in a MIMO-OFDM System", FAS 74th Annual Meeting, March, 2010

Y. Fang, N. Wu, Y. Du, K. Yen, "A Queue Management Algorithm Based on Network Processors", The 8th Int'l Conf. on IMS, Kunming, China, July, 2009.

Y. Fang, K. Yen, N. Wu, Y. Du, "An Improved Packet Scheduling Algorithm Based on Network Processors", The 2nd International Conference on Information Technology Convergence and Services (ITCS 2010) Cebu, Philippines, August 11 - 13, 2010, Accepted on May 12, 2010.

Y. Du, Y. Fang, N. Wu and K. K. Yen, " Performance Analysis for Complex FastICA Algorithms in MIMO OFDM Systems" ,The 2nd International Conference on Future Computer and Communication, Wuhan, May, 2010, Accepted on April 06, 2010.

Y. Fang, K. Yen, A. Caballero, N. Wu, "A Multidimensional Packet Classification Algorithm Based on Network Processors", The 12th WSEAS International Conference on AUTOMATIC CONTROL, MODELLING & SIMULATION (ACMOS '10) Catania, Sicily, Italy, May 29-31, 2010, Accepted on April 10, 2010.

Y. Fang, K. Yen, A. Caballero, N. Wu, "Design and Realization of Transportation Information Acquisition System Based on 3G and DSP", Journal of E-business Technology and Strategy, Vol.6 No.1, pp.61-69, March 2010.

Y. Fang, K. Yen, A. Caballero, N. Wu, "Design of a New Transportation Information Acquisition System Based on 3rd Generation Mobile Communication Technology", Proc. of the 5th Int'l Conf. on Information, Kyoto, Japan, November, 2009.

MANUSCRIPTS IN PREPARATION

N. Wu, Y. Fang, K. Yen, "Compression of Channel State Information for MIMO-OFDM System"

N. Wu, Y. Fang, K. Yen , "Efficient Adaptive Resource Allocation for MIMO-OFDM System with Rate Constraints"