2006

Radio resource allocation in OFDMA system

Lin, Jian

http://knowledgecommons.lakeheadu.ca/handle/2453/3313

Radio Resource Allocation in OFDMA System

by

Jian Lin

B.S., Huazhong University of Science and Technology (1992)

Submitted to the Department of Electrical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science

at the

LAKEHEAD UNIVERSITY

December 2005

Signature of Author ....
Department of Electrical Engineering
2 December 2005

Certified by ....
Zhiwei Mao
Assistant Professor, Electrical Engineering
Thesis Supervisor

Accepted by ....
Dr. H. T. Saliba
Dean, Faculty of Engineering

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-15628-5
Our file Notre référence
ISBN: 978-0-494-15628-5

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Radio Resource Allocation in OFDMA System

by

Jian Lin

Submitted to the Department of Electrical Engineering
on 2 December 2005, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

In this thesis, orthogonal frequency division multiple access (OFDMA) system is considered. Assuming perfect knowledge of instantaneous channel conditions for all users, we propose resource allocation algorithms to minimize the total transmission power subject to the constraints of the requirements of each user's data rate and bit error rate (BER) which is referred to as margin adaptive (MA) problem, or to maximize the overall spectral efficiency while simultaneously satisfying the requirements of each user's data rate, BER and base station (BS)'s transmission power constraint which is referred to as rate adaptive (RA) problem. By converting the above problems into linear integer programming problems, a branch-and-bound method based optimal algorithm and a fast suboptimal algorithm are proposed. The proposed branch-and-bound method based optimal algorithm offers the same optimal performance as full-search algorithm with remarkably reduced computational complexity. The proposed suboptimal algorithm, which is based on the formulation with constraints considered and greedy approach, can be used to solve both MA and RA optimization problems by satisfying the constraints one by one without any bit loading or transmission power distribution assumptions. Compared with other suboptimal methods, the performance of this suboptimal algorithm is close to the optimal one with even lower computational complexity.

Index Terms-adaptive modulation, frequency selective fading channel, multi-access communication, multiuser channel, channel capacity, orthogonal frequency division multiple access (OFDMA), power control, resource management.

Thesis Supervisor: Zhiwei Mao
Title: Assistant Professor, Electrical Engineering

Contents

List of Figures

Chapter 1

Introduction

With the rapid development of wireless communication systems in recent years, the need for high-speed data transmission has increased. One of the major problems in wideband high-speed data transmission in wireless communication systems is intersymbol interference (ISI) which comes from multipath propagation. There are many methods proposed to combat ISI, among which orthogonal frequency division multiplexing (OFDM) is a promising solution.

OFDM is not a new technique. It was first proposed and implemented in the 1960's, but it was not widely used until its all-digital implementation using the fast Fourier transform (FFT) was proposed. After that, its attractive features were unraveled and there arised widespread interests for adoption in various single-user and multiple access communication standards [1]. The basic idea of OFDM is to divide the transmitted bit stream into different substreams and send them simultaneously over different orthogonal subchannels centered at different subcarrier frequencies. The number of the substreams is chosen to make the symbol duration on each substream much greater than the delay spread of the channel, or equivalently, to make the bandwidth of each subchannel less than the coherence bandwidth of the channel [2]. Therefore, each subchannel undergoes frequency flat fading and ISI is eliminated. Today, OFDM is a major candidate for the fourth generation (4G) wireless applications with significant potential performance enhancements over existing wireless technologies.

Orthogonal frequency division multiple access (OFDMA) is a promising multi-access technique for high-speed data transmissions over wireless channels. In OFDMA system, the subchannels that appear to be in deep fades for one user may not be in deep fades for others.

Therefore spectral efficiency can be improved, or equivalently, transmission power can be reduced by dynamically allocating subchannels among the users. For OFDMA systems, in order to efficiently utilize radio resources, users' data need to be divided and distributed among subchannels in an optimal manner. Channel bandwidth and transmission power constitute two primary 'resources,' the efficient utilization of which provides the motivation for the search for radio resource allocation schemes [3].

One primary objective of the radio resource allocation schemes is to minimize the overall transmission power required to achieve users' quality of service (QoS) requirements; the other objective is to maximize users' overall data throughput at a given transmission power level, while at the same time satisfying each user's QoS requirements. The QoS requirements here refer to two aspects: one is the bit error rate (BER) performance and the other is data rate requirement. In other words, we want to minimize the total transmission power or to maximize the overall spectral efficiency while simultaneously satisfying the requirements on each user's data rate, BER and transmitter's transmission power constraint.

In many papers, it has been shown that adaptive modulation can achieve significant performance improvement in OFDM systems with the knowledge of instantaneous channel information. In OFDM systems with adaptive modulation, the subchannels with large channel gains carry more bits, while the subchannels in deep fades carry less or even zero bits. However, in OFDMA systems, users are distributed over different orthogonal subchannels. This is the so-called multiuser diversity technique. As different subchannels may experience different fades for different users, the transmission power levels or numbers of bits allocated to different subchannels for different users must be changed accordingly to achieve the overall optimal result. The problem of radio resource allocation has been studied in many recent papers. A dynamic Lagrangian-based power allocation scheme and an IP method based subchannel and bit allocation scheme are proposed in [4] and [5] respectively. Although both schemes can achieve very good performance, they are prohibitively complex and not suitable for real time applications. In [6], a real time subchannel allocation scheme is proposed to minimize the overall transmission power, where each subchannel is loaded with an equal number of bits. And in [7], another scheme is proposed to maximize the overall data throughput with the assumption that power is uniformly distributed on each subchannel. Both [6] and [7] get acceptable performance and low computational complexity; however, they are suboptimal under certain assumptions and their results are far from optimal ones. Furthermore, all the schemes mentioned above are applicable only to one of the two aspects of the radio resource allocation problem, i.e., to minimize the overall transmission power with the constraint of users' QoS requirements which is referred to as the margin adaptive (MA) problem, or to maximize the spectral efficiency subject to users' QoS requirements and the total transmission power constraints which is referred to as the rate adaptive (RA) problem [8]. None of the above mentioned schemes can be applied to both aspects of the radio resource allocation problem.

In this thesis, we consider an optimal resource allocation scheme based on branch-and-bound method [9] and a suboptimal fast bits feeding algorithm which can be used to solve both MA and RA problems. The above problems are first formulated as linear optimization problems with integer variables. The solutions to these problems are the optimal resource allocation schemes which reflect the corresponding relationship between users and subchannels. The subchannels are adaptively distributed among users in a frame of symbol duration and the users' modulation levels are adaptively selected in order to maintain a target QoS and power constraint. Once the resource allocation information is determined, the allocation information will be sent to users via a designated control channel so that they can extract their own data correspondingly. The channel characteristics of all the transmitter-to-receiver links can be estimated by the base station transmitter. In a practical system, there will always exist some estimation error and estimation delay. In this thesis, to simplify our discussion, we assume that the base station transmitter has perfect knowledge of instantaneous channel characteristics of all links. It is also assumed that the channels vary very slowly and the assignment is done once every several frames of symbol duration.

The rest of this thesis is organized as follows. Chapter 2 provides the overview of the wireless communication systems. Multipath propagation and the formulation of ISI are also explained in this chapter. Chapter 3 explains some fundamentals of OFDM and OFDMA. In Chapter 4, a constraint considered linearized formulation of the radio resource allocation problems is discussed first. Then, the full-search algorithm which is a basic optimal method is discussed with the constraint considered formulation. A branch-and-bound method based resource allocation algorithm is also proposed, which offers the same performance as the optimal one achieved by using exhaustive full-search algorithm with significantly reduced computational complexity. Finally, based on the constraints considered formulation, an efficient suboptimal algorithm is proposed which utilizes constant transmission power matrix and does not need any bit loading or transmission power distribution assumptions which are the basis of previous suboptimal algorithms. Simulation results are shown in this chapter as well to compare the performance and complexity of the proposed algorithms relative to some existing algorithms. Finally, conclusions and recommendations for future studies are given in Chapter 5.

Chapter 2

Fundamentals of Wireless Communication Systems

Wireless communication systems transmit video, images, text and data, not only analog signals but also digital signals which are obtained directly from data signals or by digitizing analog signals. The demand for wireless communication systems has been growing rapidly since Guglielmo Marconi first demonstrated radio's ability to provide continuous contact with ships sailing the English channel in 1897, two years after he invented radio transmission [10]. Wireless communications is one of the fastest growing fields in the engineering world. As one of the most active areas of present technology development, wireless communications is enjoying a fast growth period.

Although wireless communications can be retrospected to pre-industry age by using smoke signals, torch signaling and such other methods, the modern wireless communications was invented in 1895 when Marconi demonstrated the first radio transmission from the Isle of Wight to a tugboat 18 miles away a few decades after the telephone was invented [2]. In 1915 wireless voice transmission between New York and San Francisco was first established when radio and telephony were converged. In 1946 public mobile telephone service was introduced in 25 cities across the United States [2]. Growth in the mobile communications field first came slowly because these initial systems used a central transmitter to cover an entire metropolitan area. The solution to this capacity problem was not even conceived until AT&T Bell Laboratories developed the cellular concept in the 1960's and 1970's [11]. The wireless communications era was born with the invention and development of highly reliable, miniature, solid-state radio frequency hardware in the 1970's [10].

Wireless communications today covers a very wide array of applications. Cellular telephony is developed very rapidly at present time. Among other technologies are wireless piconetworking (as exemplified by the Bluetooth radio-on-a-chip) and other personal area network (PAN) systems (e.g., the IEEE 802.15 family of standards), wireless local area network (LAN) systems (exemplified by the IEEE 802.11 and HiperLAN families of standards, called WiFi systems), wireless metropolitan area network (MAN) systems (exemplified by the IEEE 802.16 family of standards, called WiMax systems), other wireless local loop (WLL) systems, and a variety of satellite systems.

These wireless technologies provide a basis for a very rich array of applications. They are supported by a number of transmission and channel-assignment techniques, including frequency-division multiple access (FDMA), time-division multiple access (TDMA), code-division multiple access (CDMA), OFDM and OFDMA.

2.1 Model of A Digital Wireless Communication System

Consider a discrete time wireless channel with noise. The simplified digital wireless communication system model, which sends an input message from a transmitter via wireless channel to a receiver, is illustrated in Figure 2-1.

The message may be a human voice, a television picture or data. If the message is nonelectrical or analog, it must be converted into a digital signal. The destination is the unit to which the message is communicated.

The source encoder maps the digital signal from the information source into another digital signal. This mapping is unique, and the objective is to eliminate or reduce redundancy so as to provide an efficient representation of the output of the information source. Since this mapping is pre-known by the source decoder, the decoder simply performs the inverse mapping and thereby delivers to the destination a reproduction of the original digital source output. The primary benefit of source coding is to reduce the required bandwidth.

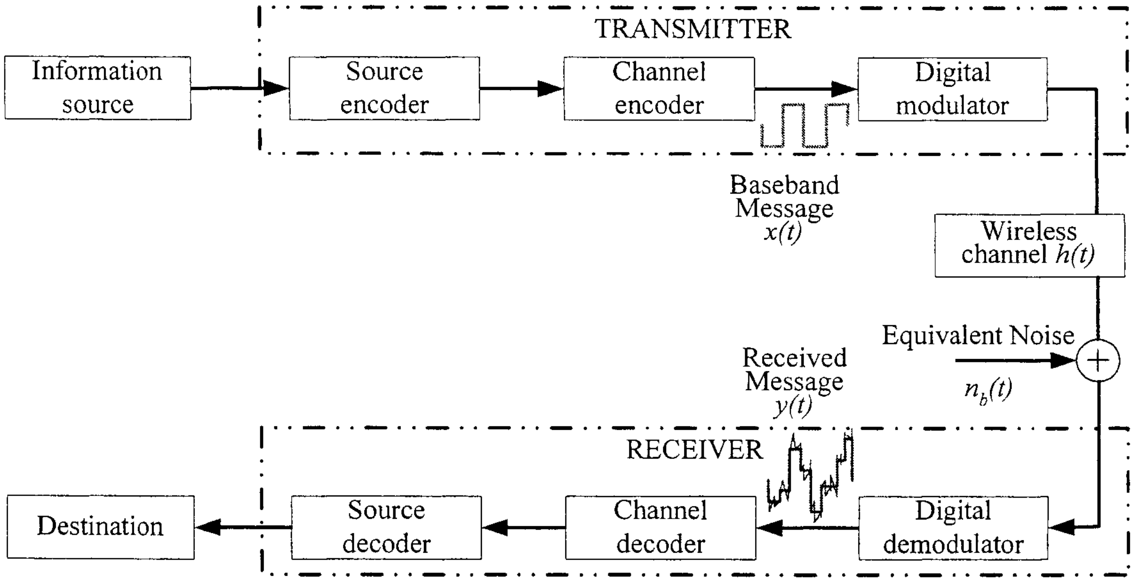


Figure 2-1: Block diagram of a simplified digital wireless communication system.

The channel encoder maps the incoming digital signal into an effective channel input. It introduces controlled redundancy to minimize the affect of the channel noise. However, this controlled redundancy is known by both of channel encoder and channel decoder and the mapping is known by both sides as well. The combination of the encoder and decoder provides reliable communication over a wireless channel with noise. In some case, we may put source encoder and channel encoder together, and source decoder and channel decoder together respectively. Naturally, the source encoding always performs before channel encoding and source decoding always performs after channel decoding as shown in Figure 2-1.

The digital modulator provides efficient transmission for the signal over the wireless channel. It modulates the incoming digital signal onto a sinusoidal carrier wave with fixed frequency limits imposed by the wireless channel. The digital demodulator performs the inverse of modulation, resuming the incoming digital signal of the system.

Channels are classified into hardwire channel and softwire channel. Wireless channel here is softwire channel [12]. Normally wireless channel takes air or vacuum as transmission medium. Digital modulator, wireless channel and digital demodulator shown in Figure 2-1 consist a discrete wireless channel [3], the input and output signals of which are in discrete form.

Practically, the electrical waveforms of the baseband message before and after the discrete wireless channel are quite different. For example, in Figure 2-1, the baseband message before the channel is $x(t)$, while after the discrete wireless channel, the received signal is $y(t)$ which is out of shape and far from the original information signal $x(t)$. $y(t)$ is formulated as

$$y(t) = x(t) \otimes h(t) + n_b(t) \tag{2.1}$$

In (2.1), $h(t)$ is the wireless channel impulse response, $n_b(t)$ is the equivalent noise. The received signal $y(t)$ is the convolution of the original information signal $x(t)$ and the impulse response of the wireless channel $h(t)$, superposed by equivalent noise $n_b(t)$.

Let T represent the time interval between successive observations of signals. Letting $t = t_n$ where $t_n = nT$ with n being an integer, time waveforms may be equivalently expressed as a sequence on n in the discrete form. That is, (2.1) may be replaced by the following [13]:

$$\begin{aligned} y(n) &= x(n) \otimes h(n) + n_b(n) \\ &= x(D)h(n-D) + \sum_{l \neq D, l=-\infty}^{\infty} x(n)h(n-l) + n_b(n) \end{aligned} \tag{2.2}$$

where D is a proper delay. The first term in (2.2) is produced by the properly delayed nth transmitted bit. This term can be correctly decoded by decoder without any extra efforts. The third term is noise. The second term represents the residual effect of all other transmitted bits on the decoding of the nth bit. It is the so called ISI. ISI is caused by multipath propagation in bandlimited channels. It distorts the transmitted signal, causing bit demodulation/detection errors at the receiver. It has been recognized as the major obstacle to high speed data transmission over wireless channels.

2.2 Wireless Multipath Channel

The performance of a wireless channel restricts the performance of a wireless communication system. Unlike wired channel which is stationary, wireless channel is extremely random because it varies due to surroundings. If there is no obstruction between transmitter and receiver in the transmission path, a line-of-sight electromagnetic wave reaches the receiver directly from

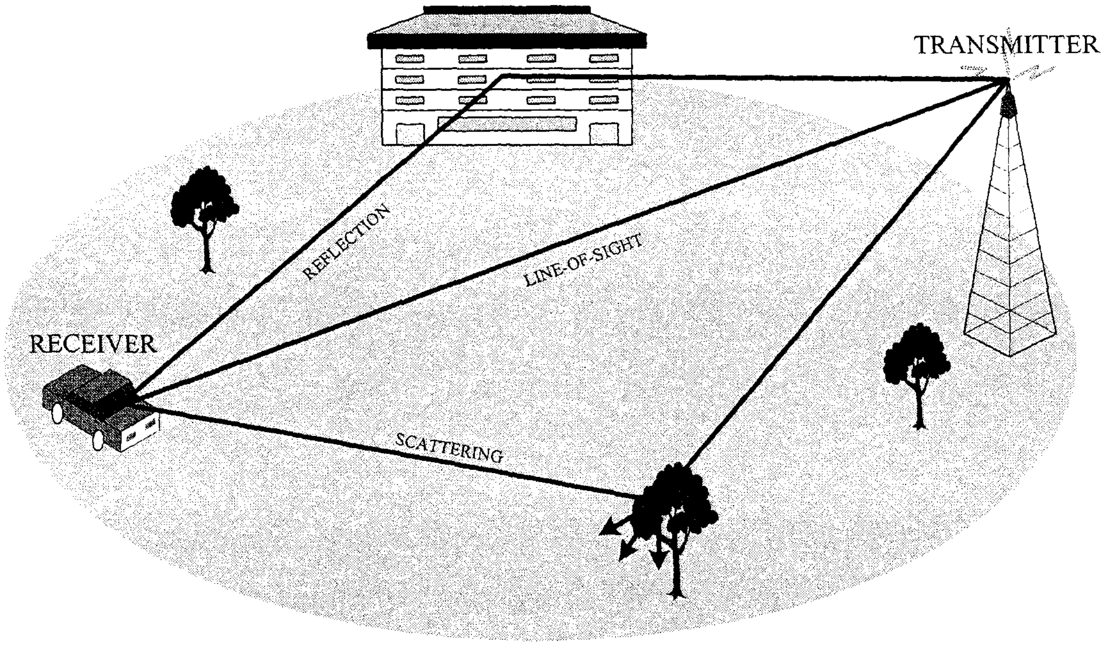


Figure 2-2: Diagram to illustrate multipath propagation in a wireless channel

the transmitter. Even so, the receiver may still receive two or even more versions of the same transmitted signal from the surroundings at slightly different times. These multipath waves combine at the receiver to give a resultant signal which is far from the original one and may vary wildly in amplitude and phase.

There are three basic propagation mechanisms impacting propagation in a wireless channel, which are reflection, diffraction and scattering. Reflection occurs when signal encounters a surface that is large relative to the wavelength of the signal. Diffraction occurs at the edge of an impenetrable body that is large compared to wavelength of radio wave. Scattering occurs when incoming signal hits an object whose size is in the order of the wavelength of the signal or less.

Figure 2-2 is an example of multipath propagation. The transmitter is at the top of the tower, while the receiver is in the car. The receiver gets a line-of-sight electromagnetic wave since there is no obstruction in the transmission path. Meanwhile, it receives reflected signal from a building and scattered signal from a tree. If there is an obstruction between the transmitter and receiver in the transmission path, the receiver may receive diffracted signal instead of the line-of-sight signal. Due to multiple transmission paths, the electromagnetic waves travel along different paths of varying lengths and the receiver receives the same signal via different paths at different delay. These multipath waves interact at the receiver causing fluctuations in signal strength, thereby inducing signal distortion. The received signal power varies instead of fading gradually with time. This is the so called multipath propagation.

Different environments produce different delay profiles and delay profiles can be viewed as the impulse response of the channel [14]. Figure 2-3 shows one of the delay profiles for the hilly terrain environment per GSM standard [15] with 12 identifiable paths having delays from 0.1 to 20 microseconds.

The multipath delay axis τ of the impulse response is discretized into equal time delay segments called excess delay bins. Each bin has a time delay width $\triangle\tau$ equal to $\tau_{i+1} - \tau_i$. τ_0 represents the first arriving signal at the receiver, which is set to be 0. For the convenience of discussion, $\tau_1 = \triangle\tau$, $\tau_i = i\triangle\tau$, for $i = 0$ to $N - 1$, where N is the total number of possible excess delays. Any signals received within the ith bin are represented by a single resolvable multipath component having delay τ_i. The number of the delay bins determines the time delay resolution of the channel model. The time resolution $\triangle\tau$ determines the bandwidths of the transmitted signals which can be analyzed by the model. That is, the model may be used to analyze such transmitted signals whose bandwidths are less than $1/2\triangle\tau$. The time resolution of the above example in Figure 2-3 is $0.1\mu s$. That means this model can be used to analyze transmitted signals having bandwidths less than $5MHz$.

2.3 The Formation of ISI

As electromagnetic waves reflect off hills, buildings, vehicles and other obstacles, they establish different transmission paths from transmitter to receiver. The received signal at the output of a wireless channel is a superposition of multiple signals arriving from different paths. It is distorted and dispersed, because the signal from each path represents a version of the transmitted signal with different delay and attenuation. The transmitted symbols within the multipath channel with these different time delays encounters time dispersion, resulting in ISI.

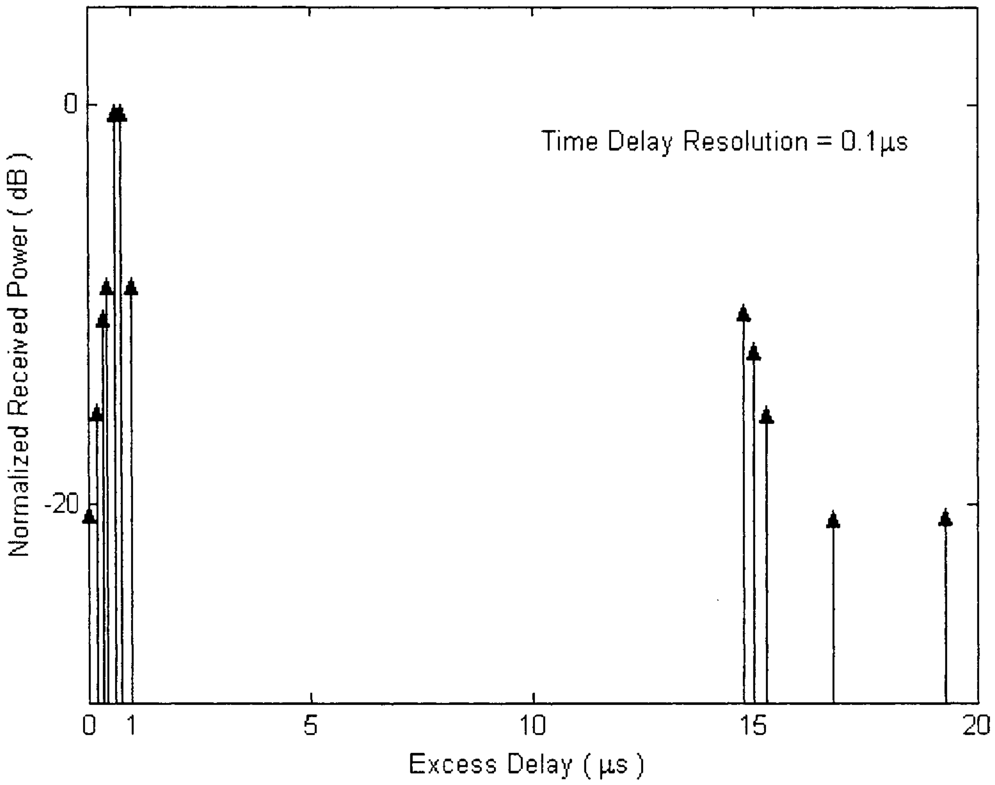


Figure 2-3: Hilly terrain delay profile for GSM (12 tap).

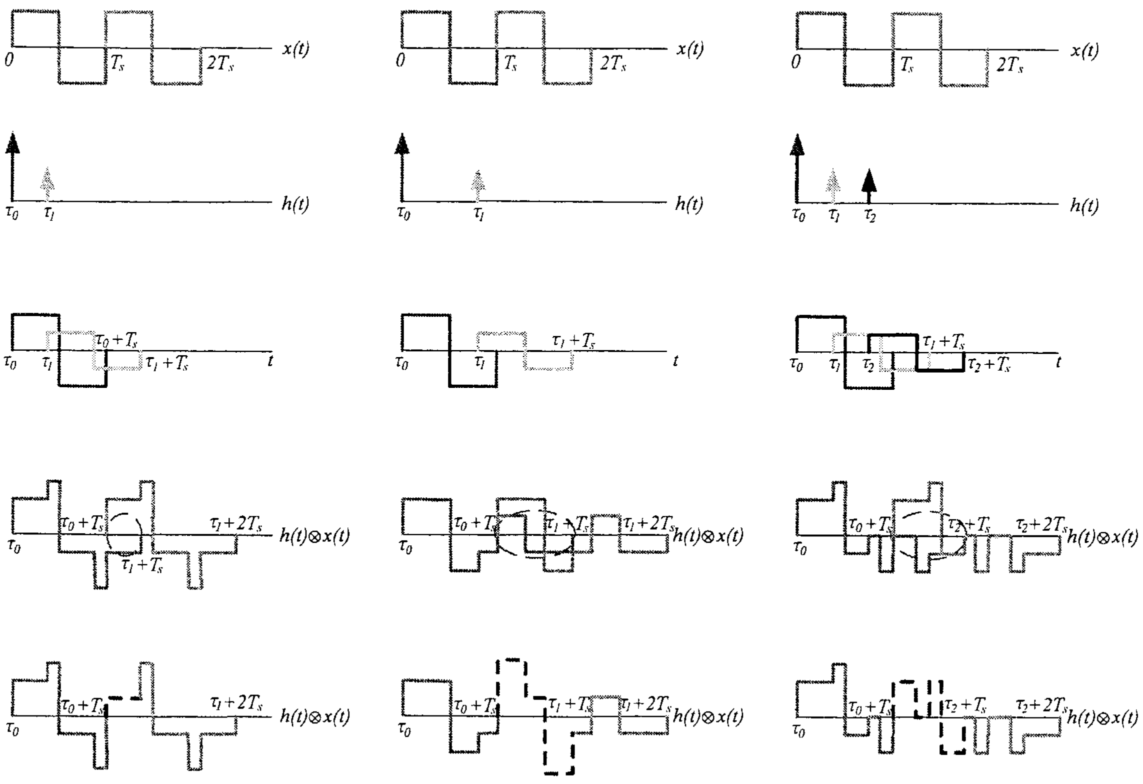
Considering a channel response in time domain without noise, simplified examples in Figure 2-4 show how time dispersion influences the received signal regardless of noise. Figure 2-4a) shows how symbols go through a channel with two time-delayed pulses. Figure 2-4b) demonstrates influence from a channel with two time-delayed pulses but the delays are longer than that in Figure 2-4a). Figure 2-4c) illustrates influence from a channel with three time-delayed pulses. In Figure 2-4 flat-top pulses shown in the top row are fed into the three channels shown in the second row respectively. The third row shows copies of the first pulse due to the channel impulse response. The fourth row shows two pulses after the channel. The dashed circles indicate the ISI regions. The first pulse spreads into the second one causing it to interfere with the neighboring pulse. The dashed curves in the bottom row are the result of superposition of both pulses, while the rest are from each pulse individually.

Compared to the channel in Figure 2-4a), the channel in Figure 2-4b) has longer excess delay. As the excess delay prolongs, the interference region extends. The channel in Figure 2-4c) has one more delay. As the number of channel delays increases, one more copy of delayed pulse falls into interference region, expanding the interference region with stronger disorder. It is easy to conclude from Figure 2-4 that longer excess delay extends the interference region and more channel multipath delays produce more copies of the same symbol with different delays and attenuation inducing stronger interference. ISI is related to both symbol duration and channel impulse response.

2.4 Frequency Flat and Frequency Selective Fading Channels

In a wireless channel, a signal may travel from transmitter to receiver over multiple paths. This phenomenon of multipath propagation causes fluctuations in the received signal's amplitude, phase and even angle of arrival, which is so-called multipath fading [16]. Based on the multipath time delay spread, fading can be further classified as flat fading or frequency selective fading [10].

If the wireless channel has a constant gain and linear phase response over a bandwidth which is larger than the bandwidth of the transmitted signal, the signal undergoes flat fading. Under these conditions, the multipath delay spread of the wireless channel impulse response



a) a two-multipath channel

b) a two-multipath channel with a longer $\triangle\tau$ than that in a)

c) a three-multipath channel

Figure 2-4: A diagram to demonstrate the formulation of ISI.

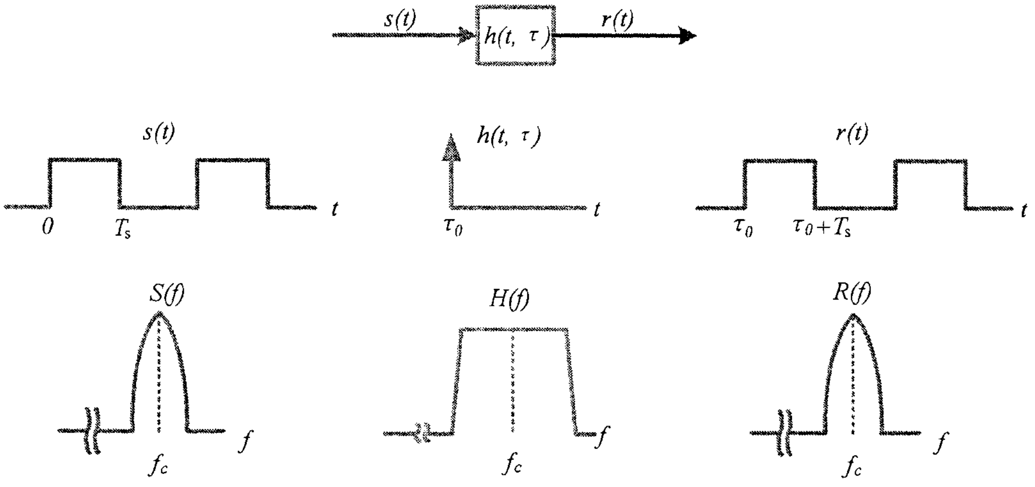
is less than the reciprocal bandwidth of the transmitted signal. The received signal has only amplitude fluctuations due to the variations in the channel gain over time caused by multipath propagation. And the spectral characteristics of the transmitted signal remain intact at the receiver.

Figure 2-5a) depicts a flat fading channel and its effect on a signal in both the time domain and the frequency domain. In a flat fading channel, the time delay spread of the channel is much less than the reciprocal bandwidth of the transmitted signal. $h_b(t, \tau)$ approximately has no excess delay, as τ could be zero compared to the symbol period T_s of the signal. Over time, only the amplitude of the signal is affected, but the spectrum of the transmission is preserved. Flat fading channels are also known as narrowband channels since the bandwidth of the signal $s(t)$ is narrower than the channel coherence bandwidth B_c.

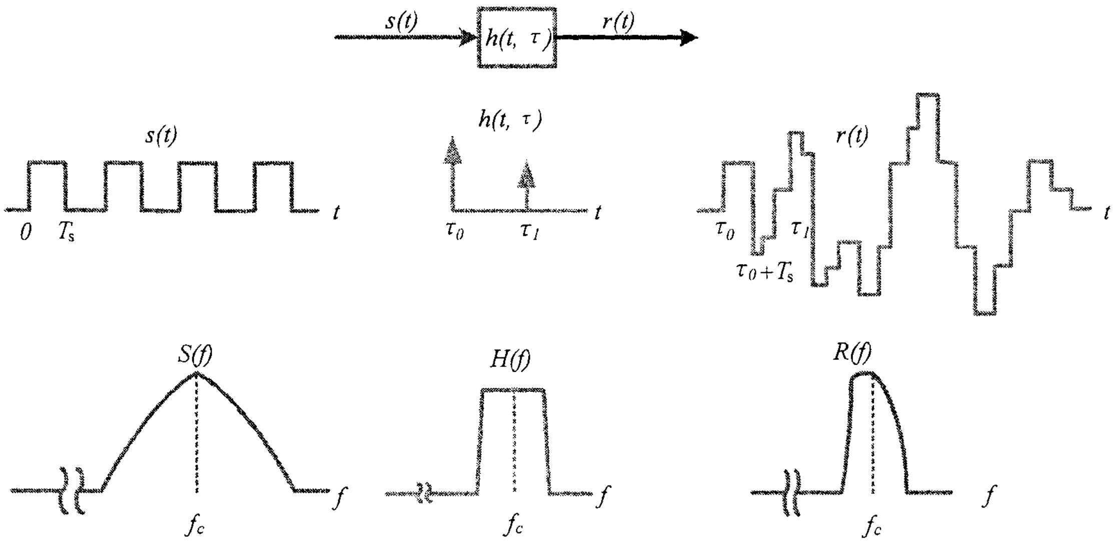
If the wireless channel has a constant gain and linear phase response over a bandwidth which is smaller than that of the transmitted signal, the signal undergoes frequency selective fading. In this case, the multipath delay spread of the wireless channel impulse response is greater than the reciprocal bandwidth of the transmitted signal. The received signal is distorted and dispersed, because it consists of multiple attenuated and delayed versions of the transmitted signal. This leads to time dispersion of the transmitted symbols, resulting in ISI.

Figure 2-5b) shows a frequency selective fading channel and its effect. The output signal suffers much distortion, as shown in both time and frequency domains. Not only the amplitude but also the phase of the signal fluctuate. Frequency selective fading is caused by multipath delays which approach or exceed the symbol period of the transmitted signal. In a frequency selective fading channel, the time delay spread of the channel is much larger than the reciprocal bandwidth of the transmitted signal. Frequency selective fading channels are also known as wideband channels, since the bandwidth of the transmitted signal is greater the channel coherence bandwidth B_c.

Coherence bandwidth B_c is a statistical measure of a frequency range over which the channel passes all spectral components with approximately equal gain and linear phase [10]. Thus, the coherence bandwidth represents a frequency range over which frequency components have a strong potential for amplitude correlation. That is, a signal's spectral components in that range are affected by the channel in a similar manner.



a) Flat fading channel characteristics



b) Frequency-selective fading channel characteristics

Figure 2-5: Flat-fading and frequency-selective fading characteristics.

If coherence bandwidth is defined as the frequency interval over which the channel's complex frequency transfer function has a correlation of at least 0.9, the coherence bandwidth is approximately given as [17]:

$$B_c \approx \frac{1}{50\delta_\tau} \tag{2.3}$$

A more popular approximation of B_c corresponding to a bandwidth interval having a correlation of at least 0.5 is as following [10]:

$$B_c \approx \frac{1}{5\delta_\tau} \tag{2.4}$$

where δ_τ is rms delay spread, which is defined as the square root of the second central moment of the power delay profile. That is, δ_τ is given as [10]:

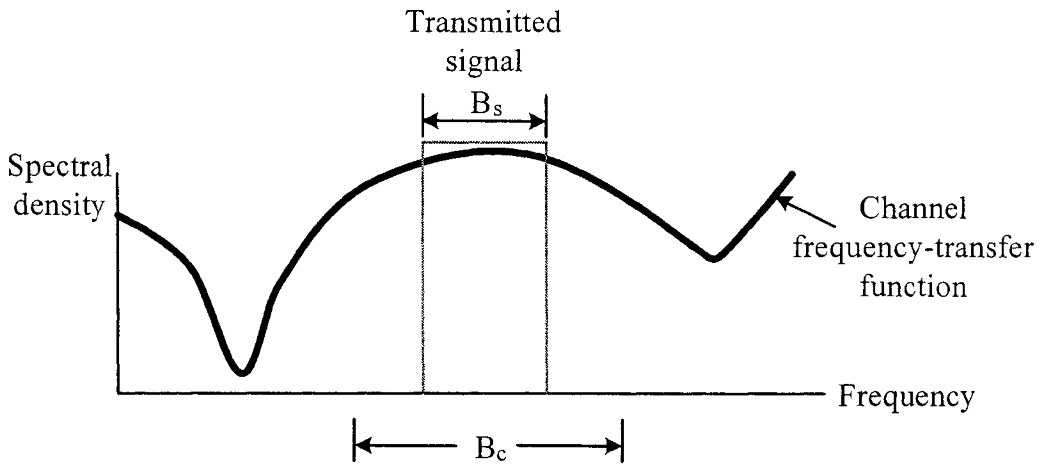
$$\delta_\tau = \sqrt{\overline{\tau^2} - (\overline{\tau})^2} \tag{2.5}$$

where $\overline{\tau}$ is the mean excess delay, $(\overline{\tau})^2$ is the mean squared delay, $\overline{\tau^2}$ is the second moment of the excess delay.

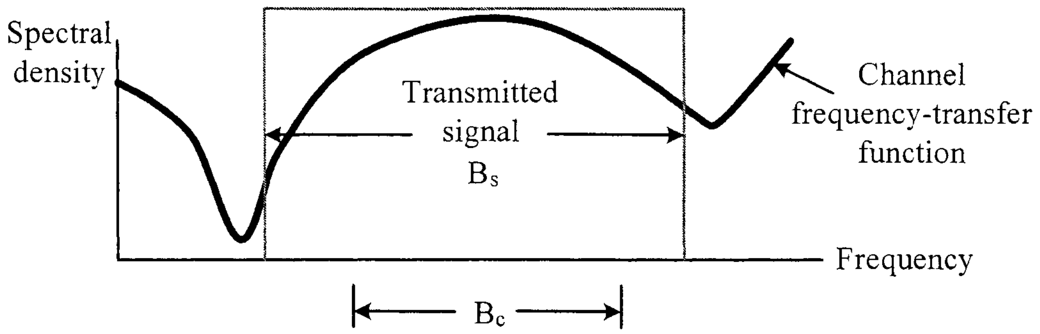
The time delay spread and coherence bandwidth of a channel are related to its multipath characteristics. It is important to note that neither B_c nor δ_τ depends on signal rate. A system's signal rate only affects its transmission bandwidth.

Whether a channel is flat fading or frequency selective fading depends on the relationship between the channel coherence bandwidth B_c and the transmitted signal bandwidth B_s. A channel may be flat fading to one transmitted signal but be frequency selective fading to another transmitted signal. Figure 2-6 shows two examples with the same typical multipath channel but two different transmitted signals.

A channel is referred to as flat fading if $B_s < B_c$. All of the signal's spectral components will be affected by the channel in the similar manner. This is illustrated in Figure 2-6a), which features the frequency response of a wireless channel superimposed by the spectral density of a transmitted signal. This channel is a flat fading channel since $B_s < B_c$. All of the spectral components of the transmitted signal will be affected in approximately the same way. Flat-fading does not introduce channel-induced ISI distortion, but performance degradation can still be expected due to the loss in SNR whenever the signal is in fading.



a) Typical flat fading case ($B_c > B_s$)



b) Typical frequency-selective fading case ($B_c < B_s$)

Figure 2-6: Relationships between the channel frequency-transfer function and transmitted signals.

Figure 2-6b) shows the frequency response of the same wireless channel superimposed by the spectral density of a wide transmitted signal, where $B_s > B_c$. Some of the signal's spectral components, falling outside the coherence bandwidth, are affected differently compared to those components contained within the coherence bandwidth. Whenever a signal's spectral components are not all affected equally by the channel, frequency selective fading distortion occurs. This channel is frequency selective fading.

Hence, the channel coherence bandwidth, B_c, sets an upper limit on the transmission rate that can be used without incorporating an equalizer in the receiver. In order to avoid channel induced ISI distortion, the channel is required to exhibit flat fading.

Chapter 3

Fundamentals of OFDM and OFDMA

3.1 Fundamentals of OFDM

Within the last few years the demand for wireless communication systems has drastically increased. There is not only a much greater amount of demand for wireless connections, but also much more interest in broadband connection. In broadband systems, the signal is scattered and reflected from objects in the environment, components of the signal arriving at the receiver are spread out over a longer period of time than they are desired. It causes uneven delays in the signal arrival time. In such an environment, ISI caused by multipath propagation in bandlimited channels distorts the transmitted signal, causing bit demodulation/detection errors at the receiver. ISI has been recognized as the fundamental performance limitations to high speed data transmission over wireless channels.

Reducing the negative impact of time dispersive channel can be accomplished by OFDM. OFDM is a promising technique to combat intersymbol interference. It spreads data streams over a large number of parallel narrowband flat fading subchannels without ISI. It also can achieve higher bandwidth efficiency than a single carrier scheme [18]. OFDM has recently received considerable interest for its advantages in high bit-rate transmissions over frequency selective fading channels [19].

The basic idea of OFDM is to divide the wideband channel into many parallel orthogonal narrowband flat fading subchannels [20]. The transmitted bit stream is divided into different substreams and sent simultaneously over different orthogonal subchannels centered at different subcarrier frequencies. Each substream has much lower data rate than the original one and each subchannel has a bandwidth much smaller than the coherence bandwidth of the channel. Therefore, each subchannel undergoes flat fading and ISI due to multipath propagation is eliminated.

As a multicarrier transmission technique, OFDM was first proposed by Chang in 1966 [21] for dispersive fading channels. This technique was first used in several high frequency (HF) military systems such as KINEPLEX, ANDEFT and KATHRYN [22]. Due to improvements in digital signal processing (DSP) and Very Large Scale Integrated circuit (VLSI) technologies, the initial obstacles of OFDM implementations such as massive complex computation and high speed memory was removed. The use of FFT algorithm further eliminated arrays of sinusoidal generators and coherent demodulation required in parallel data systems, making the implementation of OFDM cost-effective [20]. In the 1980's, OFDM has been studied for high-speed modems, digital mobile communications and high-density recording. In 1985, a cellular mobile radio system was proposed by Cimini, which used a pilot tone for stabilizing carrier, clock frequency control and trellis coding [23]. In 1990's, OFDM has been exploited for wideband data communications over mobile radio FM channels, high-bit-rate digital subscriber lines (HDSL), asymmetric digital subscriber lines (ADSL), very high-speed digital subscriber lines (VHDSL), digital audio broadcasting (DAB) and HDTV terrestrial broadcasting. Digital satellite services became popular then and the application of OFDM to the satellite mobile channels was proposed by Fernando and Rajatheva in 1998 [24].

As an unquestionable proof of its maturity, OFDM is a standard of the European DAB [25] as well as digital video broadcast (DVB) scheme [26]. It also constitutes a credible proposal for the recent third-generation mobile radio standard competition in Europe. Finally, OFDM is selected as the transmission technique of the high performance local area network's (HIPERLAN) as well as becoming part of the IEEE 802.11 Wireless Local Area Network (WLAN) standard and broadband local area networks [27]. After about forty years of research and development, OFDM has been widely implemented in high speed digital communications. Today, OFDM is a major contender for 4G wireless applications with significant potential performance

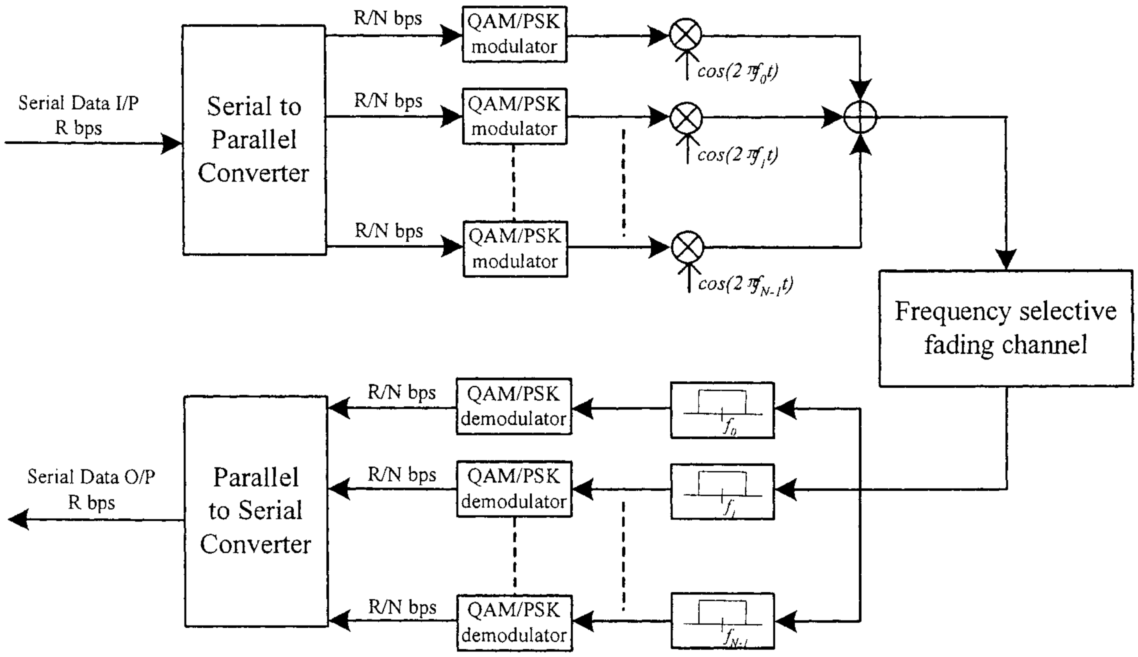


Figure 3-1: Block diagram of a simplified OFDM system.

enhancements over existing wireless technologies.

Figure 3-1 illustrates the structure of an OFDM system. The incoming serial data is first converted from serial to parallel by splitting up into N substreams. These substreams are modulated independently and sent simultaneously by many subcarriers over a frequency selective fading channel. At the receiver, they are filtered, demodulated and converted from parallel into serial to form the original data. Since N is sufficiently large, the data rate of each substream is very low and the bandwidth of each subchannel is much less than the coherence bandwidth of the channel. Therefore, each subchannel is a flat fading channel and ISI can be eliminated. The frequency selective fading channel can be approximated by the sum of many flat fading subchannels as well. As an example, Figure 3-2 is the approximation of the typical multipath channel in Figure 2-6. The original channel is divided into N subchannels with each having a bandwidth $B_N << B_C$.

In OFDM systems N has to be sufficiently large to ensure that each subchannel has a bandwidth much smaller than the coherence bandwidth of the channel. It can be determined

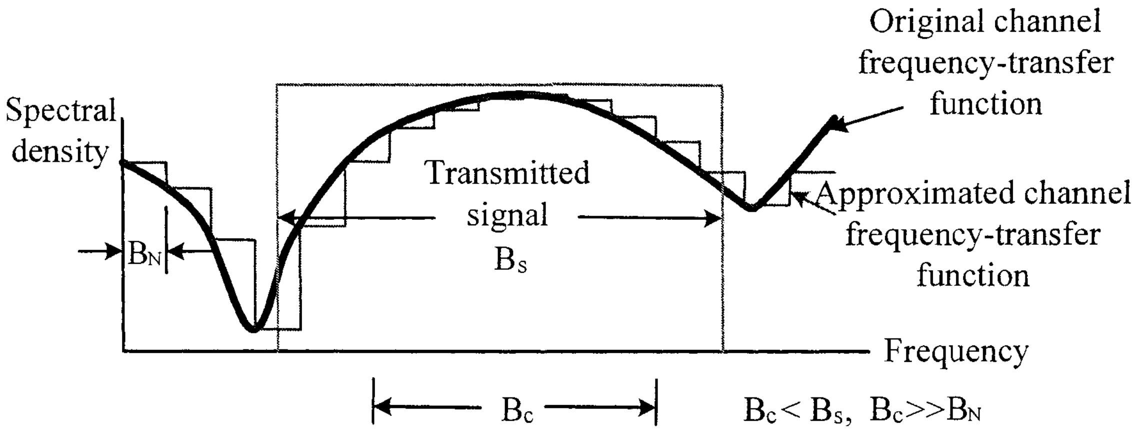
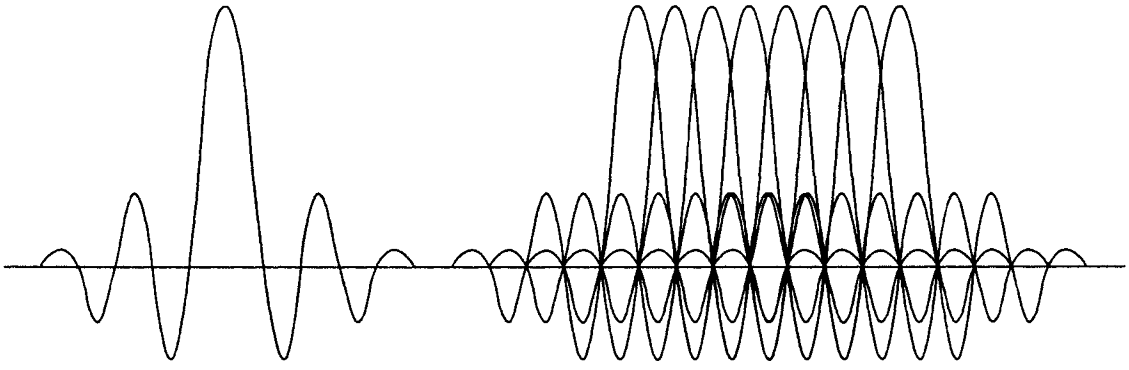


Figure 3-2: An example of frequency division.

based on the channel bandwidth, data throughput and useful symbol duration. N corresponds to the number of complex points being processed in FFT. It is a trade-off between computational complexity and algorithm precision. A properly set N maintains a low level computational complexity for radio resource allocation schemes, while the system achieves an acceptable performance. For HDTV applications, the number of subcarriers are in the range of several thousands, so as to accommodate the data rate and guard interval requirement.

In OFDM systems, in order to correctly receive and demodulate the signals, the many subcarriers are spaced apart mathematically orthogonally. It is possible to arrange the sidebands of the individual subcarriers overlapping to improve the spectral efficiency of OFDM. The signals can still be received without adjacent subcarrier interference.

Consider an OFDM system with baseband channel bandwidth B (passband bandwidth $2B$) and channel coherence bandwidth B_c. The baseband subchannel bandwidth $B_N = B/N << B_c$, where N is sufficiently large. This insures flat fading on each subchannel. The baseband subcarriers are $\{cos(2\pi f_i t), i = 1, 2, ...\}$. They form a set of orthonormal bases on the interval $[0, T]$, where T is the OFDM symbol period and $f_i = i/T$. $f = 1/T$ is the minimal frequency separation between orthogonal subcarriers over $[0, T]$ [2]. If raised cosine pulses $T = 0.5(1+\beta)/B_N$ with $\beta = 1$ are used, we get subcarrier separation $f = 1/T = B_N$. Because the passband bandwidth of each subchannel is $2B_N$, the passband subchannels in the system overlap, illustrated in



a) A single carrier. b) 8 overlapping subcarriers.

Figure 3-3: Example of overlapping OFDM spectrum.

Figure 3-3. OFDM allows the spectrum of each tone to overlap. And since they are orthogonal, they do not interfere with each other. We can still separate the different subcarriers.

Due to the requirement for separated modulators and demodulators on each subchannel, it is too complex and expensive for most systems to implement OFDM until the development of FFT and inverse FFT (IFFT) and improvements of DSP and VLSI technologies. Figure 3-4 illustrates the process of an FFT based OFDM system. The incoming serial data is first converted from serial to parallel and grouped into x bits each to form a complex number. The complex numbers are modulated in a baseband fashion by the IFFT. And converted back to serial data for transmission. A guard interval is inserted between symbols to avoid ISI caused by multipath distortion. The discrete symbols are converted to analog and lowpass filtered for radio frequency (RF) up-conversion. The receiver performs the inverse process of the transmitter. One tap equalizer is used on each subchannel to correct channel distortion. The tap coefficients of the filter are calculated based on channel information [22].

The key components of an OFDM system are the IFFT at the transmitter and FFT at the receiver. These operations perform reversible linear mappings between N complex data symbols and N complex OFDM symbols. The number of multiplications required by an N-point FFT is only on the order of $NlogN$ rather than N^2 as in a straightforward computation. Due to this fact, an OFDM system typically requires fewer computations per unit time than an equivalent

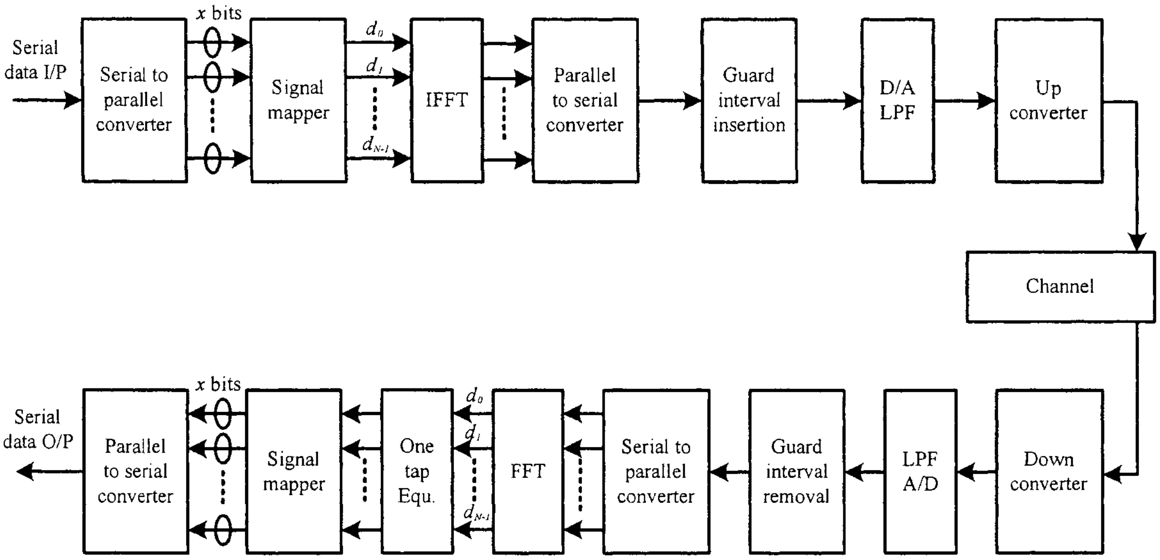


Figure 3-4: Block diagram of an IFFT based OFDM system.

system with equalization. Transmission of data in the frequency domain using an FFT, as a computationally efficient orthogonal linear transformation, results in robustness against ISI in the time domain [2].

3.2 Fundamentals of OFDMA

As a multi-carrier modulation technique, OFDM is very immune to impulse noise and frequency selective fading due to the increased symbol interval [18]. The performance of OFDM systems can be further improved by adaptive modulation with the knowledge of instantaneous channel transfer functions. In OFDM systems with adaptive modulation, in order to improve the system performance the subchannels with large channel gains carry more bits, while the subchannels in deep fades carry less or even zero bits [4]. If a user encounters deep fades in some frequency ranges, the transmitter has to spend more power on feeding substreams into corresponding subchannels. Sometimes these subchannels have to be dropped because of efficiency of the system. The radio bandwidth is being wasted.

It is noted that the subchannels that appear to be in deep fades for one user may not be

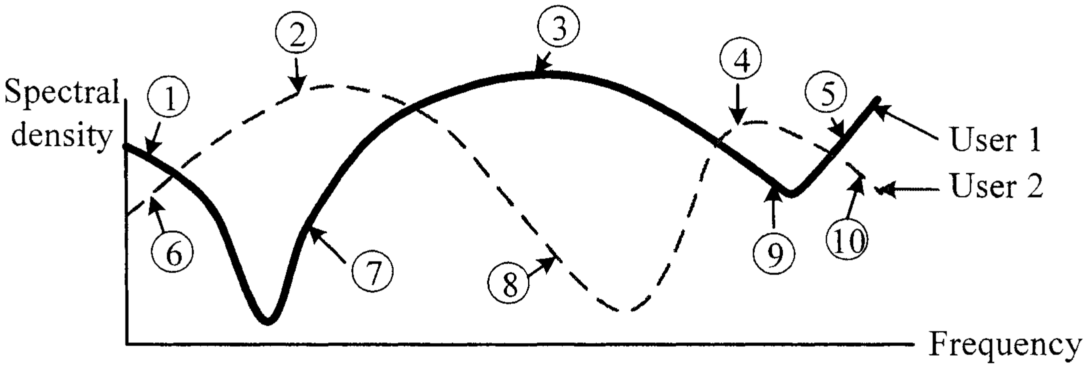


Figure 3-5: Principle of OFDMA

in deep fades for others. This is illustrated in Figure 3-5. There are two users in this system. Region 7 and 9 are in fade to user 1 but not to user 2, while regions 6, 8 and 10 are in fade to user 2 but not to user 1. It implies that the combined channel frequency transfer function of the system may take the regions 1 through 5 and there are not bad subchannels in a multiuser system. In this way, the spectral efficiency can be improved, or equivalently, transmission power can be reduced. Based on this observation, OFDM based multiuser technique can be used to enhance the performance of the system. Multiuser OFDM was first presented by Wahlqvist [28] who suggested one possible implementation in 1996. It has been known that the spectral efficiency of the multiuser OFDM system is much higher than that of the single user OFDM system.

OFDMA is a promising multiple access technique for multiuser high-speed data transmissions over wireless channel. In OFDMA systems, users' data are divided and sent simultaneously, users are distributed over different orthogonal subchannels. This is the so called multiuser diversity technique. Since different subchannels may experience different fades for different users and each user's performance requirements have to be satisfied, in order to efficiently utilize radio resources, the transmission power levels or numbers of bits allocated to different subchannels for different users must be changed accordingly to achieve the overall optimal result.

Figure 3-6 is the block diagram of a simplified OFDMA system downlink with a radio resource allocation scheme. The considered channel is divided into N subchannels. There are K users transmitting information simultaneously. The K users' data are divided, modulated

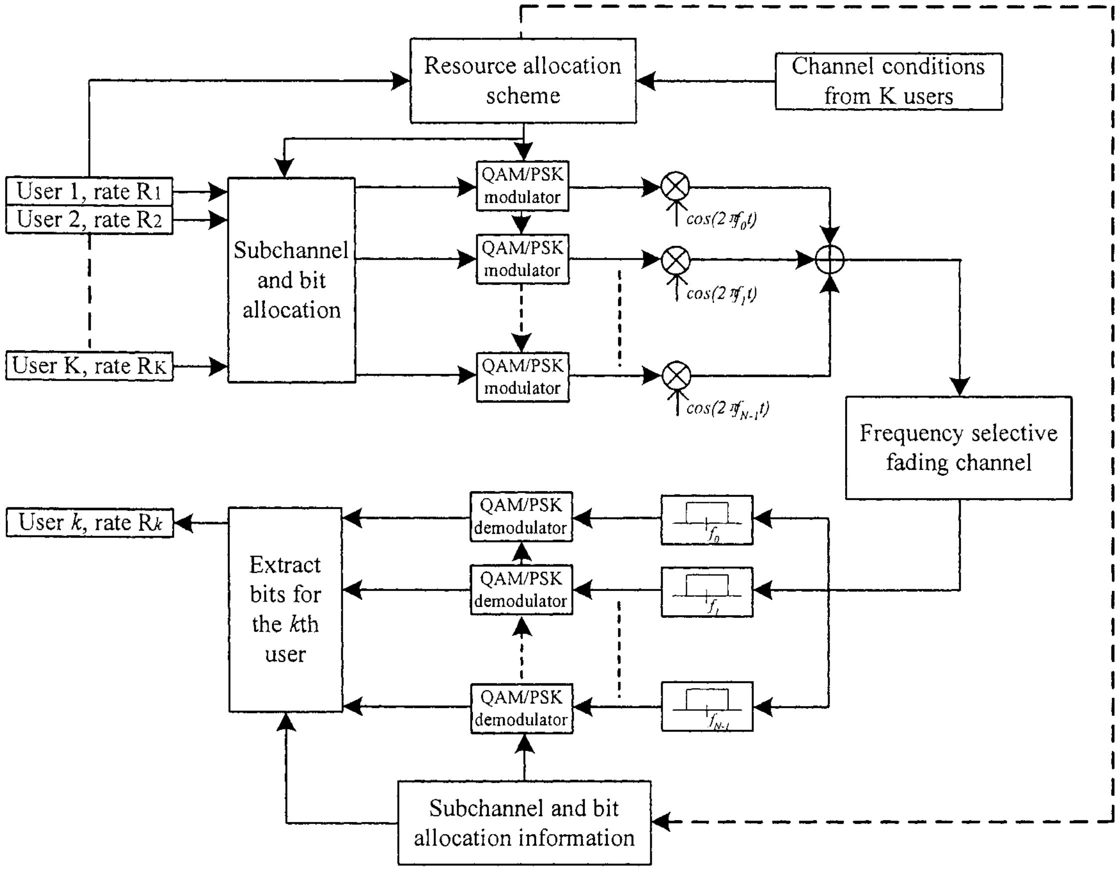


Figure 3-6: Block diagram of a simplified OFDMA system downlink with a radio resource allocation scheme.

according to the radio resource allocation scheme and fed into the N orthogonal subchannels simultaneously. At the receiver side, all data from N subchannels are demodulated. The bit allocation information is sent to the corresponding users so that they can extract their own data via a designated control channel. In OFDMA systems distinct subchannels are assigned to distinct users in an efficient manner.

Chapter 4

Radio Resource Allocation in OFDMA Systems

The demand for mobile wireless communication systems is growing rapidly. This demand, combined with limited availability of radio frequency spectrum, calls for communication networks that can operate with high bandwidth efficiency. The design of efficient signal modulation and demodulation techniques for physical layer transmission is utmost important in this regard. And the future growth of consumer-based mobile and portable communication systems will be tied more closely to transmission power, radio spectrum allocations and regulatory decisions which affect or support new or extended services, as well as to consumer needs and technology advances in signal processing, access and network areas. These demands in future wireless services make radio resource more and more invaluable.

Channel bandwidth and transmission power constitute two primary "radio resources" [3]. Unfortunately, these two traditional resources are among the most severely limited deployment of modern wireless networks. Radio bandwidth is very tight with regard to useful radio spectrum, and transmission power is restricted because not only basestations but also mobiles or other transmitters have limited power supply. These two resources are simply not growing or improving at rates that can support anticipated demands for wireless systems.

The efficient utilization of these two "radio resources" provides the motivation for the search for spectrally efficient schemes. One primary objective of the spectrally efficient schemes is to minimize the overall transmission power required to achieve users' QoS requirements; the other objective is to maximize users' overall data throughput at a given transmission power level, while at the same time satisfying each user's QoS requirements. These optimization problems are referred to as the margin adaptive (MA, the former) and rate adaptive (RA, the latter) optimization problems respectively [8]. The QoS requirements here refer to two factors. One is the BER performance and the other is data rate requirement. In other words, we want to minimize the total transmission power or to maximize the overall spectral efficiency while simultaneously satisfying the requirements on each user's data rate, BER and transmitter's transmission power constraint.

A great deal of active researches have been devoted to the problem of resource allocation recently. Previous researches on OFDMA systems have shown that spectral efficiency can be improved by bit and power allocation at the transmitter. A dynamic Lagrangian-based power allocation scheme and an IP method based subchannel and bit allocation scheme are proposed in [4] and [5] respectively. Although both schemes can achieve very good performance, they are prohibitively complex and not suitable for real time applications. In [6], a real time subchannel allocation scheme is proposed to minimize the overall transmission power, where each subchannel is loaded with an equal number of bits. And in [7], another scheme is proposed to maximize the overall data throughput with the assumption that power is uniformly distributed on each subchannel. Both [6] and [7] get good performance and low computational complexity; however, they are suboptimal under certain assumptions and can not achieve optimal results. Furthermore, all the schemes mentioned above are applicable only to one of the two aspects of the resource allocation problem, i.e., either to MA problem or to RA problem. None of them can be applied to both aspects of the resource allocation problem.

In this chapter, MA and RA problems are formulated as linear optimization problems with integer variables. The solutions to these problems are the resource allocation schemes which reflect the corresponding relationship between users and subchannels. To simplify our discussion, we assume that the base station transmitter has perfect knowledge of instantaneous channel characteristics of all links. The channel characteristics of all the transmitter-to-receiver links can be estimated by the basestation transmitter, although in a practical system, there always exist some estimation error and estimation delay. It is also assumed that the channels vary very slowly and the assignment is done once every several frames of symbol duration.

In this chapter, full-search algorithm which is a basic optimal method is discussed based on the formulation with constraint considered. A branch-and-bound method based optimal resource allocation algorithm is proposed. It offers the same performance as the optimal one achieved by using exhaustive full-search algorithm, but with significantly reduced computational complexity. Based on the formulation with constraint considered, an efficient suboptimal algorithm is proposed which utilizes the constant transmission power matrix and does not need any bit loading [6], transmission power distribution assumptions [7] or intensive computation [5].

4.1 System Model and Problem Formulation

In a single user OFDM system, the data stream is divided into multiple substreams to be transmitted over different orthogonal subchannels centered at different subcarrier frequencies. In a multiuser environment, users' data will be distributed over different orthogonal subchannels. A typical multiuser adaptive OFDMA system studied in this thesis with downlink transmission is shown in Figure 3-6. The wireless downlink channel of this system is divided into N subchannels. N is large enough to ensure that each subchannel has a bandwidth much smaller than the coherence bandwidth of the channel, so that each subchannel undergoes flat fading. It is also assumed that there are K users and the kth user has a data rate requirement of R_k $(k = 1, 2, \cdots, K)$.

In the transmitters of OFDMA systems, by using the channel information, the data streams from K users are fed into the subchannels and bit allocation block to allocate bits from different users to different subchannels. Depending on the number of bits assigned to a subchannel, the adaptive modulator will choose a corresponding modulation level, and the transmission power will be adjusted correspondingly to satisfy each user's QoS requirements. The resource allocation information will also be sent to corresponding users via a designated control channel so that they can extract their own data.

It is assumed that the greatest possible number of bits that can be transmitted by any subchannel in a frame of symbol duration is M. $\alpha_{k,n}$ is denoted as the magnitude of the channel gain of the kth user over the nth subchannel. The single-sided noise power spectral density (PSD) level N_0 is assumed to be unit (i.e., $N_0 = 1$) for all subchannels and is the same for all users. QAM modulation scheme is applied.

At the receiver side, the required power to support c bits/symbol at a given BER P_e is [4]

$$P_r(c) = \frac{N_0}{3}\left[Q^{-1}\left(\frac{P_e}{4}\right)\right]^2 (2^c - 1) \tag{4.1}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$.

It is easy to see that $P_r(c)$ increases as c increases and $P_r(0) = 0$.

At the transmitter side, in order to achieve the required BER at the receiver, the transmission power $P_{k,n}$ assigned to the nth subchannel for the kth user must be

$$P_{k,n} = \frac{P_r(c_{k,n})}{\alpha_{k,n}^2} \tag{4.2}$$

where $c_{k,n}$ is the number of bits which are assigned over the nth subchannel for the kth user.

Margin Adaptive (MA) Optimization

The MA problem minimizes the overall transmission power with the constraint of users' QoS. In other words, we want to minimize the overall transmission power while the resultant BER should not be higher than a target one and every user's data rate should not be lower than its rate requirement.

Mathematically, the MA problem can be expressed as [4]

$$P_T = \min_{c_{k,n}\in D}\sum_{n=1}^{N}\sum_{k=1}^{K}\frac{P_r(c_{k,n})}{\alpha_{k,n}^2} \tag{4.3}$$

subject to:

$C1 : \sum_{n=1}^{N} c_{k,n} \geq R_k, k = 1, 2, \cdots, K$

$C2$: If there exists k' with $c_{k',n} \neq 0$, then $c_{k,n} = 0$, $\forall k' \neq k$, $k = 1, 2, \cdots, K, n = 1, 2, \cdots, N$

In (4.3) $D = \{0, 1, 2, \cdots, M\}$ is the set of all possible values for $c_{k,n}$. $c_{k,n} = 0$ indicates that the kth user does not use the nth subchannel to transmit any information. The constraint $C1$ is the data rate requirement with R_k being the kth user's required data rate. $C2$ ensures that each subchannel is only occupied by at most one user at any instance of time [29].

Rate Adaptive (RA) Optimization

The RA problem maximizes the spectral efficiency subject to users' QoS and the total transmission power constraints. This problem is mathematically expressed as [6]

$$R = \max_{c_{k,n} \in D} \sum_{n=1}^{N} \sum_{k=1}^{K} c_{k,n} \tag{4.4}$$

subject to:

$C1 : \sum_{n=1}^{N} c_{k,n} \geq R_k, k = 1, 2, \cdots, K$

$C2$: If there exists k' with $c_{k',n} \neq 0$, then $c_{k,n} = 0$, $\forall k' \neq k$, $k = 1, 2, \cdots, K$, $n = 1, 2, \cdots, N$

$C3 : \sum_{n=1}^{N} \sum_{k=1}^{K} P_{k,n} \leq P$, $k = 1, 2, \cdots, K$, $n = 1, 2, \cdots, N$

Similar to the case in (4.3), $D = \{0, 1, 2, \cdots, M\}$ in (4.4). The constraints $C1$ and $C2$ have the same meaning as those in (4.3). $C3$ is the total transmission power constraint. $P_{k,n}$ is the transmission power assigned to the nth subchannel for the kth user. P is the overall restricted transmission power.

The MA and RA optimizations are nonlinear because $P_r(c)$ in (4.1) is nonlinear. In next section, we will convert the nonlinear MA and RA optimization problems into linear MA and RA optimization problems with integer variables, since $c_{k,n}$ takes only integer values. In the following discussion, boldface lower case letters represent vectors and boldface upper case letters represent matrices.

4.2 Integer Programming Problem Formulation

As mentioned in the former section, $c_{k,n} \in D = \{0, 1, 2, \cdots, M\}$. We represent $c_{k,n}$ by a subchannel-user-bit vector $\mathbf{x}_{k,n}$ which is defined to indicate the number of bits assigned over the nth subchannel for the kth user, and $\mathbf{x}_{k,n}$ is expressed as

$$\mathbf{x}_{k,n} = \begin{bmatrix} x_{k,n,1} \\ x_{k,n,2} \\ \vdots \\ x_{k,n,M} \end{bmatrix}, \ \mathbf{x}_{k,n} \in \left\{ \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\},$$

$$k = 1, 2, \cdots, K,\ n = 1, 2, \cdots, N,\ \mathbf{x}_{k,n} \in \{0, 1\}^{M \times 1} \tag{4.5}$$

In (4.5), $x_{k,n,m}$ $(m = 1, 2, \cdots, M)$ is the bit-allocation indicator which can only be one or zero. It can be seen that at most one component of $\mathbf{x}_{k,n}$ can be one. If $x_{k,n,m} = 1$, then $x_{k,n,m'} = 0$, for all $m' \neq m$. $x_{k,n,m} = 1$ means that the nth subchannel is assigned to the kth user to transmit m bits during one symbol duration. The position of one in the subchannel-user-bit vector indicates the number of bits carried by the nth subchannel for the kth user. If all the components are zeros, the nth subchannel is not assigned to the kth user.

Based on the definition of the subchannel-user-bit vector $\mathbf{x}_{k,n}$, we define a bit-allocation matrix $\mathbf{X}$ which is resource allocation scheme in matrix form given as

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{1,1} & \mathbf{x}_{1,2} & \cdots & \mathbf{x}_{1,N} \\ \mathbf{x}_{2,1} & \mathbf{x}_{2,2} & \cdots & \mathbf{x}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{K,1} & \mathbf{x}_{K,2} & \cdots & \mathbf{x}_{K,N} \end{bmatrix},\ \mathbf{X} \in \{0, 1\}^{KM \times N} \tag{4.6}$$

$\mathbf{x}_K^k = \begin{bmatrix} \mathbf{x}_{k,1}^T & \mathbf{x}_{k,2}^T & \cdots & \mathbf{x}_{k,N}^T \end{bmatrix}^T$ corresponds to the kth user. It shows how the bits of the kth user are distributed over N subchannels. Each user can occupy more than one subchannel at the same time. $\mathbf{x}_N^n = \begin{bmatrix} \mathbf{x}_{1,n}^T & \mathbf{x}_{2,n}^T & \cdots & \mathbf{x}_{K,n}^T \end{bmatrix}^T$ corresponds to the nth subchannel. It shows how the nth subchannel is occupied by K users. Since each subchannel can be occupied by no more than one user at the same time, the nth subchannel vector has $(KM+1)$ possibilities shown below. The position of one in $\mathbf{x}_N^n$ indicates how this subchannel is occupied by the user and how many bits are transmitted in one symbol duration.

$$\begin{aligned} \mathbf{x}_N^n &= \begin{bmatrix} \mathbf{x}_{1,n} \\ \mathbf{x}_{2,n} \\ \vdots \\ \mathbf{x}_{K,n} \end{bmatrix},\ \mathbf{x}_N^n \in \left\{ \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}, \\ n &= 1, 2, \cdots, N,\ \mathbf{x}_N^n \in \{0, 1\}^{KM \times 1} \end{aligned} \tag{4.7}$$

We define a bit-allocation vector $\mathbf{x}$ shown as below whose length is NKM, which is the resource allocation scheme in vector form..

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_N^1 \\ \mathbf{x}_N^2 \\ \vdots \\ \mathbf{x}_N^N \end{bmatrix}, \mathbf{x} \in \{0,1\}^{NKM\times 1} \tag{4.8}$$

We now define a subchannel-user-power vector $\mathbf{p}_{k,n}$ below, whose product with subchannel-user-bit vector $\mathbf{x}_{k,n}$ is the consumed transmission power in order to send m bits over the nth subchannel for the kth user if $x_{k,n,m} = 1$.

$$\mathbf{p}_{k,n} = \begin{bmatrix} p_{k,n,1} \\ p_{k,n,2} \\ \vdots \\ p_{k,n,M} \end{bmatrix}, k = 1,2,\cdots,K,\ n = 1,2,\cdots,N;\ \mathbf{p}_{k,n} \in C^{M\times 1} \tag{4.9}$$

Referring to (4.2), $p_{k,n,m}$ in (4.9) is defined below with m indicating the number of bits.

$$p_{k,n,m} = \frac{P_r(m)}{\alpha_{k,n}^2},\ k = 1,2,\cdots,K,\ n = 1,2,\cdots,N,\ m = 1,2,\cdots,M \tag{4.10}$$

Based on the definition of subchannel-user-power vector $\mathbf{p}_{k,n}$, for the whole system we define a power matrix $\mathbf{P}$ below corresponding to bit-allocation matrix $\mathbf{X}$,

$$\mathbf{P} = \begin{bmatrix} \mathbf{p}_{1,1} & \mathbf{p}_{1,2} & \cdots & \mathbf{p}_{1,N} \\ \mathbf{p}_{2,1} & \mathbf{p}_{2,2} & \cdots & \mathbf{p}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{p}_{K,1} & \mathbf{p}_{K,2} & \cdots & \mathbf{p}_{K,N} \end{bmatrix}, \mathbf{P} \in C^{KM\times N} \tag{4.11}$$

In (4.11), $\mathbf{p}_N^n = \begin{bmatrix} \mathbf{p}_{1,n}^T & \mathbf{p}_{2,n}^T & \cdots & \mathbf{p}_{K,n}^T \end{bmatrix}^T$ corresponds to the nth subchannel. In this thesis, since we choose bit-allocation vector $\mathbf{x}$ in subchannel vector form, correspondingly, for the whole system we define power vector $\mathbf{p}$ in subchannel form below, which is constant with length NKM as

$$\mathbf{p} = \begin{bmatrix} \mathbf{p}_N^1 \\ \mathbf{p}_N^2 \\ \vdots \\ \mathbf{p}_N^N \end{bmatrix}, \mathbf{p} \in C^{NKM \times 1} \tag{4.12}$$

In the rest of this section, based on the definition of bit-allocation vector $\mathbf{x}$ and power vector $\mathbf{p}$, we convert the nonlinear MA and RA optimization problems into linear optimization with integer variables.

4.2.1 Margin Adaptive (MA) Optimization

Based on the fact that the channels vary very slowly, $\alpha_{k,n}$, which is denoted as the magnitude of the channel gain of the kth user over the nth subchannel, is assumed invariant. Therefore, the power vector $\mathbf{p}$ is a constant. The MA optimization problem in (4.3) can be converted into the following linear optimization problem.

$$\min_{\mathbf{x}} \mathbf{p}^T \mathbf{x} \tag{4.13}$$

subject to:

$$\mathbf{A}_u \cdot \mathbf{x} \geq \mathbf{r} \tag{4.14}$$

$$\mathbf{A}_c \cdot \mathbf{x} \leq \mathbf{c} \tag{4.15}$$

$$x_i \in \{0, 1\}, \; i = 1, 2, \cdots, MKN \tag{4.16}$$

Where x_i donates the ith component of the vector $\mathbf{x}$, $\mathbf{A}_u$, $\mathbf{r}$, $\mathbf{A}_c$, and $\mathbf{c}$ are given by the following equations.

$$\mathbf{A}_u = \begin{bmatrix} \mathbf{a}_{K1} & \mathbf{a}_{K1} & \cdots & \mathbf{a}_{K1} \\ \mathbf{a}_{K2} & \mathbf{a}_{K2} & \cdots & \mathbf{a}_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_{KK} & \mathbf{a}_{KK} & \cdots & \mathbf{a}_{KK} \end{bmatrix}, \mathbf{A}_u \in Z^{K \times MKN} \tag{4.17}$$

$$\left\{\begin{array}{ll} \mathbf{a}_{K1} = \left[\begin{array}{cccc} \mathbf{a}_u & \mathbf{0}_u & \cdots & \mathbf{0}_u \end{array} \right] & \mathbf{a}_{K1} \in Z^{1\times MK} \\ \mathbf{a}_{K2} = \left[\begin{array}{cccc} \mathbf{0}_u & \mathbf{a}_u & \cdots & \mathbf{0}_u \end{array} \right] , & \mathbf{a}_{K2} \in Z^{1\times MK} \\ \cdots & \vdots \\ \mathbf{a}_{KK} = \left[\begin{array}{cccc} \mathbf{0}_u & \mathbf{0}_u & \cdots & \mathbf{a}_u \end{array} \right] & \mathbf{a}_{KK} \in Z^{1\times MK} \end{array}\right. \tag{4.18}$$

$$\mathbf{a}_u = \left[\begin{array}{cccc} 1 & 2 & \cdots & M \end{array} \right], \ \mathbf{a}_u \in Z^{1\times M} \tag{4.19}$$

$$\mathbf{0}_u = \left[\begin{array}{cccc} 0 & 0 & \cdots & 0 \end{array} \right], \ \mathbf{0}_u \in \{0\}^{1\times M} \tag{4.20}$$

$$\mathbf{r} = \left[\begin{array}{cccc} R_1 & R_2 & \cdots & R_K \end{array} \right]^T, \ \mathbf{r} \in C^{K\times 1} \tag{4.21}$$

$$\mathbf{A}_c = \left[\begin{array}{cccc} \mathbf{1}_c & \mathbf{0}_c & \cdots & \mathbf{0}_c \\ \mathbf{0}_c & \mathbf{1}_c & \cdots & \mathbf{0}_c \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_c & \mathbf{0}_c & \cdots & \mathbf{1}_c \end{array} \right], \mathbf{A}_c \in \{0,1\}^{N\times MKN} \tag{4.22}$$

$$\mathbf{1}_c = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array} \right], \ \mathbf{1}_c \in \{1\}^{1\times MK} \tag{4.23}$$

$$\mathbf{0}_c = \left[\begin{array}{cccc} 0 & 0 & \cdots & 0 \end{array} \right], \ \mathbf{0}_c \in \{0\}^{1\times MK} \tag{4.24}$$

$$\mathbf{c} = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array} \right]^T, \ \mathbf{c} \in \{1\}^{N\times 1} \tag{4.25}$$

In this formulation, the bit-allocation vector $\mathbf{x}$ is given by (4.8), and the power vector $\mathbf{p}$ is given by (4.12). $\mathbf{p}^T\mathbf{x}$ is the system overall transmission power which we are going to minimize. BER requirement is satisfied in the definition of power vector $\mathbf{p}$. The product of the user-bit matrix $\mathbf{A}_u$ and the bit-allocation vector $\mathbf{x}$ returns K users' data rates. $\mathbf{r}$ is K users' data requirement with R_k being the kth user's data requirement. Therefore, the constraint (4.14) is the data rate requirements. The product of the subchannel matrix $\mathbf{A}_c$ and the bit-allocation vector $\mathbf{x}$ shows occupation status of N subchannels. Therefore, (4.15) regulates N subchannels' occupation status. Since each subchannel can only be occupied by at most one user at the same time, $\mathbf{c}$ has to be composed of N ones.

4.2.2 Rate Adaptive (RA) Optimization

Based on the expression of $\mathbf{a}_u$ in (4.19), we define $\mathbf{b}_{fv}$ as

$$\mathbf{b}_{fv} = \begin{bmatrix} \mathbf{a}_u & \mathbf{a}_u & \cdots & \mathbf{a}_u \end{bmatrix}^T, \ \mathbf{b}_{fv} \in Z^{MKN \times 1} \tag{4.26}$$

$\mathbf{b}_{fv}$ is composed of KN $\mathbf{a}_u$s and its product with the bit-allocation vector $\mathbf{x}$ is the system overall data throughput.

The RA optimization problem in (4.4) can be reformulated as

$$\max_{\mathbf{x}} \mathbf{b}_{fv}^T \mathbf{x} \tag{4.27}$$

subject to:

$$\mathbf{p}^T \mathbf{x} \leq P \tag{4.28}$$

$$\mathbf{A}_u \cdot \mathbf{x} \geq \mathbf{r} \tag{4.29}$$

$$\mathbf{A}_c \cdot \mathbf{x} \leq \mathbf{c} \tag{4.30}$$

$$x_i \in \{0, 1\}, \ i = 1, 2, \cdots, MKN \tag{4.31}$$

In this formulation, (4.28) is the power constraint where $\mathbf{p}^T\mathbf{x}$ is the system overall transmission power and it should not be greater than the given total system transmission power P.

Based on this integer programming problem formulation, the exhaustive full-search algorithm is discussed first in the following section. Then a branch-and-bound method based optimal algorithm and an efficient suboptimal algorithm are proposed.

4.3 Methods To Solve the IP Problems

OFDMA efficiently improves the link performance in wireless communication systems. A great deal of researches have been devoted to the problem of resource allocation in OFDMA systems already. Some optimal or suboptimal resource allocation algorithms have been proposed. Due to the nature of nonlinear optimization, most of them require intensive computation, the assumptions of equal bit loading or equal transmission power distribution over different subchannels.

In this section, full-search algorithm, a basic optimal method, is discussed first. It is based on the linearized formulation in Section 5.2 with constraint considered. A branch-and-bound method based optimal resource allocation scheme is proposed as well. The proposed algorithm achieves the same optimal performance as that by using full-search algorithm but with a much lower computational complexity. Finally, a suboptimal algorithm is developed. This suboptimal algorithm achieves closer results to the optimal ones comparing to other suboptimal algorithms, without any bit loading or transmission power distribution assumptions, while it requires much less computational complexity than optimal algorithms.

4.3.1 Full-search Algorithm

In order to get optimal solutions, the basic method to solve integer programming problem is full-search algorithm. The essence of full-search algorithm is to enumerate all the possibilities. For the linearized MA and RA optimization problems, all possibilities of the bit-allocation vector $\mathbf{x}$ will be examined in the full-search algorithm. For each variable vector satisfying all the constraints, the values of the MA and RA cost functions are calculated. The variable vectors corresponding to the minimum of the MA cost function or the maximum of the RA cost function are the required resource allocation solution schemes.

Full-search algorithm is the simplest but the most time consuming method due to its essence of enumeration. Even though, a proper formulation of the optimal problem can still dramatically reduce its computational complexity. In the system considered, there are KMN components in the bit-allocation vector $\mathbf{x}$. Traditionally, each component would be treated as 0 or 1. It implies that there would be 2^{KMN} possible $\mathbf{x}$s to examine. From (4.7) and (4.8), it is easy to

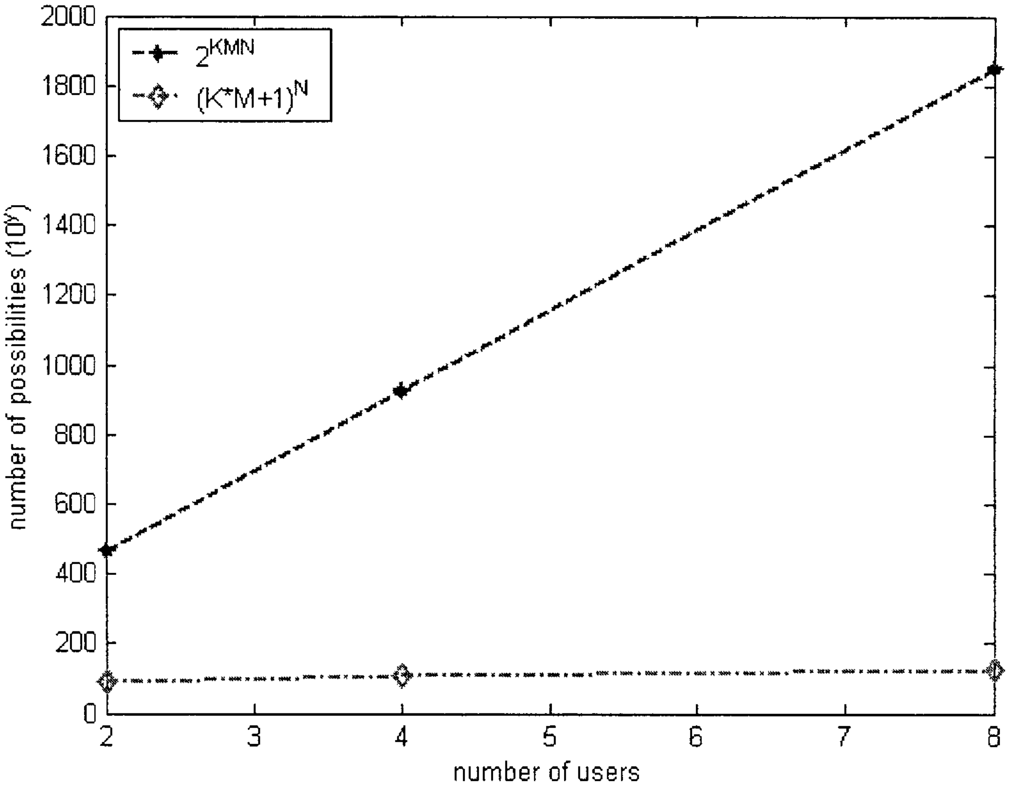


Figure 4-1: Comparison of the number of possibilities.

know that the bit-allocation vector $\mathbf{x}$ is composed of N subchannel vector $\mathbf{x}_N^n$ with $(MK + 1)$ possibilities each, since each subchannel cannot be occupied by more than one user at the same time. The number of possibilities of the bit-allocation vector $\mathbf{x}$ are $(MK + 1)^N$. It is much lower than the number of possibilities of the traditional method which is 2^{KMN}. The comparison of the number of possibilities of the two methods is illustrated in Figure 4-1 for a system with $N = 64, M = 12$. With the number of users increasing, the number of possibilities for the formulation with constraint considered to enumerate become lower and lower relative to the traditional method. Therefore, in the following discussion, when full-search method is mentioned, it is always referred to the method with $(MK + 1)^N$ possibilities. However, the complexity of this full-search algorithm for the wireless resource allocation problem is still exponential, which makes full-search algorithm not a real time method.

4.3.2 Branch-and-Bound Method Based Optimal Algorithm

Branch-and-bound is a general algorithmic method for finding optimal solutions to various optimization problems without exhaustive enumeration [9]. It doesn't expand all search nodes, rather, it determines which node to expand and when an optimal solution has been found. The essence of the branch-and-bound approach is to prune in the total enumeration tree at the node that the optimal solution cannot occur in any of its descendents. If enough branches of the tree are pruned in this way, the computational complexity is able to be reduced to a computationally manageable size. The optimal solution is achieved by keeping the best optimal cost function value found so far. If this cost function value can't be further updated, it is the final cost function value to the problem and the solutions corresponding to this cost function value are the optimal solutions [30].

For a problem having N nodes with P possibilities each, the branch-and-bound method starts from the root problem which is the original problem with complete feasible region. A temporary cost function value V with a temporary solution $\mathbf{x}$ is first found. The bounding procedure is applied to the root problem. If each of the N nodes of $\mathbf{x}$ falls into one of the P possibilities, that means the bounds match, then $\mathbf{x}$ and V are optimal results to the root problem and the procedure terminates. Otherwise, one of the N nodes which are not among the P possibilities is divided into P branches which generate subproblems of the root problem.

These subproblems together cover the whole feasible region. The algorithm is applied recursively to the subproblems. The branches with the best cost function value at their levels are expanded and followed prior to the other branches at the same level, generating a tree of subproblems, until all the N nodes are assigned among the P possibilities and a temporary cost function value V_{tem} to the original problem which is also optimal to a subproblem is found. V_{tem} is a feasible solution to the full problem, but not necessarily globally optimal. Since it is feasible, it can be used to develop the rest of the search tree. Any branches with worse cost function value than V_{tem} are pruned, any branches with better cost function value than V_{tem} are developed until all the N nodes are assigned among the P possibilities. If the new temporary cost function value is better, V_{tem} is updated and it is used to examine the rest branches. The search proceeds until all nodes have been solved or pruned, therefore, V_{tem} is the final cost function value to the problem and the solutions corresponding to this cost function value are the optimal solutions [31].

Branch-and-bound method gets optimal results with low complexity. It is not a heuristic or approximating procedure, but an exact optimizing procedure that finds an optimal solution [32]. Branch-and-bound method belongs to the class of implicit enumeration methods. It was first proposed by A. H. Land and A. G. Doig in 1960 for linear programming [33] and has been widely used in Artificial Intelligence and Operations Research.

Branch-and-bound approach is an efficient optimal method to solve IP problem with low computational complexity. In this section, a branch-and-bound method based optimal algorithm on radio resource allocation is presented. Since LINPROG of MATLAB is applied to solve LP, in order to apply branch-and-bound method, the radio resource allocation problems have to be first formulated into linear form. By removing the integer constraints, the original integer linear programming problem is relaxed to the following LP form [34].

$$\min_{\mathbf{x}} \mathbf{f}^T \mathbf{x} \tag{4.32}$$

subject to:

$$\mathbf{A} \cdot \mathbf{x} \leq \mathbf{b} \tag{4.33}$$

$$\mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub} \tag{4.34}$$

where $\mathbf{f}$, $\mathbf{x}$, $\mathbf{b}$, $\mathbf{lb}$, and $\mathbf{ub}$ are vectors and $\mathbf{A}$ is matrix.

For **MA optimization**,

$$\mathbf{f} = \mathbf{p} \tag{4.35}$$

$$\mathbf{A} = \begin{bmatrix} -\mathbf{A}_u \\ \mathbf{A}_c \end{bmatrix} \tag{4.36}$$

$$\mathbf{b} = \begin{bmatrix} -\mathbf{r} \\ \mathbf{c} \end{bmatrix} \tag{4.37}$$

while $\mathbf{lb}$ is a column vector composed of MKN 0s, and $\mathbf{ub}$ is of MKN 1s.

For **RA optimization**,

$$\mathbf{f} = -\mathbf{b}_{fv} \tag{4.38}$$

$$\mathbf{A} = \begin{bmatrix} -\mathbf{A}_u \\ \mathbf{A}_c \\ \mathbf{p}^T \end{bmatrix} \tag{4.39}$$

$$\mathbf{b} = \begin{bmatrix} -\mathbf{r} \\ \mathbf{c} \\ P \end{bmatrix} \tag{4.40}$$

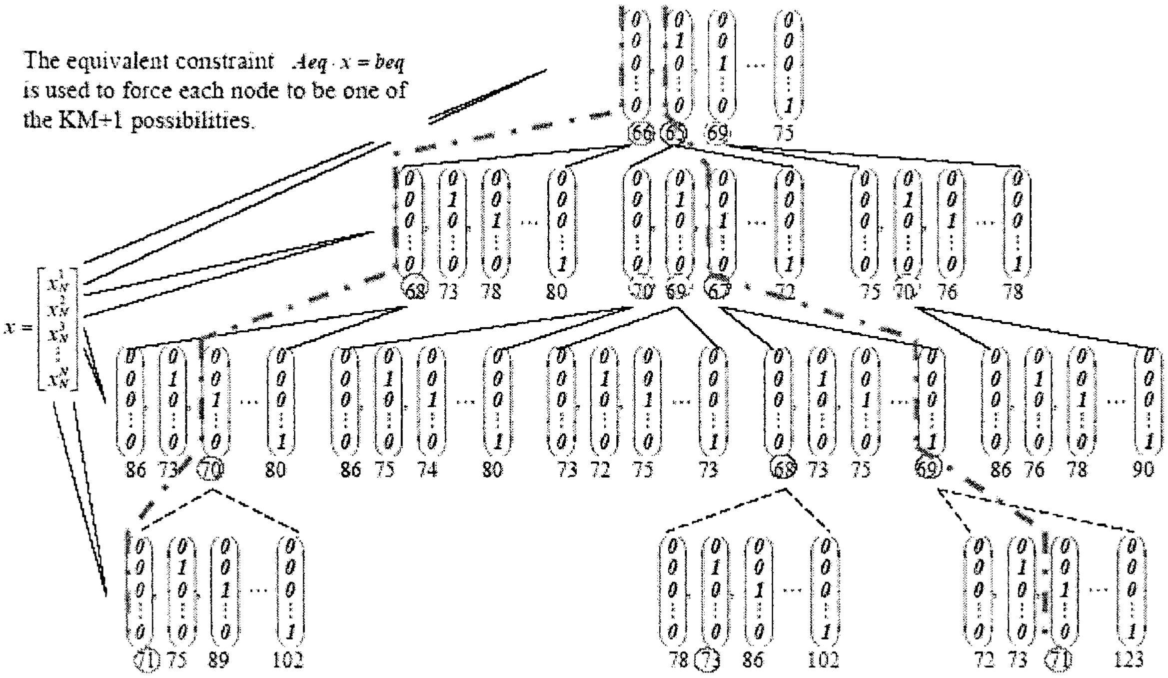
while $\mathbf{lb}$ is a column vector composed of MKN 0s, and $\mathbf{ub}$ is of MKN 1s.

Based on the general branch-and-bound approach, the proposed branch-and-bound method based optimal algorithm contains a node stack *ACTIVE*, a scalar *UPPER*, which is the minimum feasible cost value obtained so far, and the corresponding solution *CurrentBest*. The proposed algorithm uses the equivalent constraint [34] to set each node to be one of the $(KM+1)$ possibilities. Each node is labelled with the level of the node k and the lower bound z_k. The proposed branch-and-bound method based optimal algorithm for both MA and RA problems is summarized as follows [35]:

```
Initialization: $k = 0$, $UPPER = 10^{20}$, $z_k = 10^{10}$, $CurrentBest = null$, and $ACTIVE = null$
Step 1: Set $k = k + 1$, and let $z_k = z_{k-1}$.
Step 2: If $z_k \geqslant UPPER$, drop this node and go to Step 6. Otherwise, set $i = 1$.
Step 3: Generate the node in level $k$ such that
        $\begin{cases} \text{if } i < KM+1, \text{ then } [\mathbf{x}_N^k]_i = 1,\ [\mathbf{x}_N^k]_{j \neq i} = 0; \\ \text{if } i = KM+1, \text{ then } \mathbf{x}_N^k = \mathbf{0}. \end{cases}$
Step 4: Solve the subproblem of the MA or RA LP relaxation problem with $\mathbf{x}_N^1, \cdots, \mathbf{x}_N^k$ set already.
        If the subproblem is infeasible, drop this node and go to Step 5.
        If the subproblem is feasible, calculate the cost value and denote it as $z_k$.
          4a) If $z_k \geqslant UPPER$, drop this node and go to Step 5.
          4b) If $z_k < UPPER$ and $k < N$, append this node to the end of $ACTIVE$ and store $k$, $z_k$, and $\mathbf{x}_N^1, \cdots, \mathbf{x}_N^k$ associated with this node.
          4c) If $z_k < UPPER$ and $k = N$, update $UPPER = z_k$ and $CurrentBest = \mathbf{x}$.
Step 5: $i = i + 1$. If $i \leqslant KM + 1$, go to Step 3. Otherwise, go to Step 6.
Step 6: If ACTIVE is not empty, pick the node from the end of $ACTIVE$, and set $k$, $z_k$, and $\mathbf{x}_N^1, \cdots, \mathbf{x}_N^k$ to the stored values associated with this node and go to Step 1.
Step 7: Stop and output the solution $CurrentBest$.
```

This procedure can get bit allocation scheme for both MA and RA optimization problems. For MA, the optimal cost function value is the minimal transmission power. For RA, the negative optimal cost function value is the maximal data throughput.

In Figure 4-2, an example is shown to explain the proposed algorithm. This is an integer linear problem requiring minimal cost function value. There are N nodes with $(KM + 1)$ possibilities each. The order of the nodes $x_N^1, \cdots, x_N^N$ which are going to be examined is assumed from x_N^1 to x_N^N. The tool used to solve linear problem is LINPROG of MATLAB [34]. The possibilities of each node are regulated by the equivalent constraint of the LP problem. The temporary optimal cost function values V_{br} corresponding to all first level branches are found



The equivalent constraint $Aeq \cdot x = beq$ is used to force each node to be one of the KM+1 possibilities.

Figure 4-2: An example to demonstrate the process of the proposed branch-and-bound method based optimal algorithm.

by first only assign x_N^1 to be one of the $(KM+1)$ possibilities (branches). They are listed under corresponding branches in Figure 4-2. The branch-and-bound algorithm develops the branch with the least V_{br} first. Since in this example 65 is the least, the corresponding branch is developed first. In this branch, by assigning $\mathbf{x}_N^2$ to be one of the $(KM+1)$ possibilities, the branch corresponding to 67 is developed secondly, since 67 is the least at its level so far. So on so forth, until all nodes are assigned, a temporary cost function value V_{tem} 73 is achieved. V_{tem} is used to develop the rest of the search tree by comparing $V_{tem} = 73$ to all found V_{br} from node x_N^N approaching node x_N^1. At one level upper, the branch with V_{br} 69 is developed, since 69 is less than 73; the other branches are pruned, since they are not less than 73 so that they can not develop any results less than 73. By developing the branch with 69, 71 is then achieved, which is less than 73. V_{tem} is updated to be 71. Then 71 is used to develop the rest of the search tree. At the second level, the branches with 69 and 70 are developed, none returns a better result than 71. At the first level, the branch with 66 returns 71, while the branch with 69 doesn't return better result. So far all the nodes have been examined, then the updated V_{tem} 71 becomes the global optimal solution to this problem. The corresponding branches are the solutions to this problem. There are maybe more than one branch corresponding to the optimal cost function value. For a practical application, one solution is enough.

4.3.3 Fast Bits Feeding Algorithm

Branch-and-bound is a systematic method for solving optimization problems. If applied carefully, it can lead to algorithms that run reasonably fast on average. However, it is still quite slow for some application. It often leads to exponential time complexities in the worst case. In order to further reduce computation load, suboptimal algorithms have been developed. As alternative suboptimal algorithms are cheaper and easier to implement, although their results are a little bit far from the optimal ones, they are still acceptable. Most suboptimal algorithms so far are based on the nonlinear formulation and under certain bit loading or transmission power distribution assumption. For instance, in [6], a real time subchannel allocation scheme is proposed to minimize the overall transmission power, while it is assumed that each subchannel is loaded with an equal number of bits. In [7], another scheme is proposed to maximize the overall data throughput with the assumption that power is uniformly distributed on each subchannel. [6] or

[7] is used for only one aspect of the resource allocation problem. Although the scheme based on linearized formulation in [5] gets better performance and can be used for both aspects of the optimization problem, it is extremely complicated for real time applications. Based on the linearized formulation in Section 4.2, a fast bits feeding algorithm is proposed in this section which can be used to solve both MA and RA optimization problems by feeding one bit a time with least incremental transmission power to satisfy the constraints. This fast bits feeding algorithm outperforms the previous suboptimal algorithms with better performance but without the above equal bit loading assumption, equal transmission power distribution assumption or intensive computation.

Defined by (4.9) (4.10) and (4.11), the constant power matrix **P** contains all the information regarding the characteristics of all subchannels, all possibilities of bits assigned to different subchannels and the corresponding power levels for different number of bits loading over different subchannels. For instance, the component $p_{k,n,m}$ represents the consumed transmission power to assign m bits from the kth user over the nth subchannel. The columns of power matrix **P** correspond to subchannels.

The algorithm here is to pick up the required components to satisfy the constraints of (4.3) or (4.4) for MA or RA problem respectively. The indices of the chosen components form the resource allocation scheme.

For **MA problem**, the process is first to distribute all the subchannels from subchannel pool A to all users without assigning any bits and one user can only be assigned one best subchannel at one time. Secondly, bits are assigned one by one to the assigned subchannels S_k with the least incremental transmission power, restricted by the assumed maximal bits M which can be assigned to a subchannel. Once a user's data rate requirement is satisfied, any subchannels assigned to this user without any bits assigned are released to the subchannel pool A. Then all the subchannels in the subchannel pool A are distributed again to the other users, until all users' data rate requirements are satisfied. This process ensures the constraints of MA problem satisfied. Matrix **X** represents the resource allocation scheme here. The fast bits feeding algorithm for MA problem is summarized as follows:

```
Initialization: $A = \{1, \ldots, N\}$, $U = \{1, \ldots, K\}$, $\mathbf{X} = \mathbf{0} \in Z^{K \times N}$, $S_k = \phi, \forall k \in \{1, \ldots, K\}$.
Step 1: Find $n$ and $k$ satisfying:
            $p_{k,n,1} \leqslant p_{k',n',1}, \forall k,\ k' \in U,\ n,\ n' \in A$,
            set $S_k = S_k + \{n\}$, $U = U - \{k\}$, $A = A - \{n\}$.
Step 2: 2a) If $A \neq \phi$, $U \neq \phi$, go to Step 1,
        2b) If $A \neq \phi$, $U = \phi$, set $U = \{1, \ldots, K\}$, go to Step 1.
        2c) Otherwise, set $U = \{1, \ldots, K\}$, $\triangle p_{k,n,0} = p_{k,n,1}, \forall k \in U,\ n \in S_k$.
Step 3: $[n, k] = argmin_{k \in U, n \in S_k} \triangle p_{k,n,x_{k,n}}, \forall k \in U$,
            set $x_{k,n} = x_{k,n} + 1$,
        If $x_{k,n} = M$, set $S_k = S_k - \{n\}$,
            otherwise, evaluate $\triangle p_{k,n,x_{k,n}} = p_{k,n,x_{k,n}+1} - p_{k,n,x_{k,n}}$.
Step 4: 4a) If $\sum_{n=1}^{N} x_{k,n} < R_k, \forall k \in U$, go to Step 3,
        4b) If $\sum_{n=1}^{N} x_{k,n} = R_k$, for any $k \in U$,
                find all $n \in S_k$, such that, $x_{k,n} = 0$,
                then set $A = A + \{n\}$, $S_k = S_k - \{n\}$.
Step 5: Set $U = U - \{k\}$.
        5a) If $U = \phi$, go to Step 8,
        5b) If $U \neq \phi$, $A = \phi$, go to Step 3.
        5c) Otherwise, go to Step 6.
Step 6: Find $n$ and $k$ satisfying:
                $p_{k,n,1} \leqslant p_{k',n',1}, \forall\ k, k' \in U,\ \ n,\ n' \in A$
                set $S_k = S_k + \{n\}$, $A = A - \{n\}$.
Step 7: If $A = \phi$, go to Step 3,
        otherwise, go to Step 6.
Step 8: Stop and output the least transmission power $P_T = \sum_{k=1,n=1}^{K,\ N} p_{k,n,x_{k,n}}$.
```

RA optimization problem is similar to MA problem mathematically with one more constraint. Since subchannel occupation status constraint and users' data rate requirements constraint are satisfied by solving MA problem, for RA problem, in order to satisfy the transmission power constraint, all the subchannels in the subchannel pool A have to be distributed to the best users. Bits are assigned one by one to the assigned subchannels S_k with the least incremental transmission power, restricted by the assumed maximal bits M which can be assigned to a subchannel. The transmission power is evaluated until the transmission power constraint is satisfied. Matrix $\mathbf{X}$ returns the resource allocation scheme. The solution is $R = \sum_{k=1,n=1}^{K,\,N} x_{k,n}$. The fast bits feeding algorithm for RA problem is shown below:

```
Initialization: $A = \{1, \dots, N\}$, $U = \{1, \dots, K\}$, $\mathbf{X} = \mathbf{0} \in Z^{K \times N}$, $S_k = \phi, \forall k \in \{1, \dots, K\}$.
Step 1: Find $n$ and $k$ satisfying:
        $p_{k,n,1} \leqslant p_{k',n',1}$, $\forall k,\ k' \in U,\ n,\ n' \in A$,
        set $S_k = S_k + \{n\}$, $U = U - \{k\}$, $A = A - \{n\}$.
Step 2: 2a) If $A \neq \phi$, $U \neq \phi$, go to Step 1,
        2b) If $A \neq \phi$, $U = \phi$, set $U = \{1, \dots, K\}$, go to Step 1.
        2c) Otherwise, set $U = \{1, \dots, K\}$, $\triangle p_{k,n,0} = p_{k,n,1}$, $\forall k \in U,\ n \in S_k$.
Step 3: $[n, k] = argmin_{k \in U, n \in S_k} \triangle p_{k,n,x_{k,n}}$, $\forall k \in U$,
        set $x_{k,n} = x_{k,n} + 1$,
        If $x_{k,n} = M$, set $S_k = S_k - \{n\}$,
        otherwise, evaluate $\triangle p_{k,n,x_{k,n}} = p_{k,n,x_{k,n}+1} - p_{k,n,x_{k,n}}$.
Step 4: 4a) If $\sum_{n=1}^{N} x_{k,n} < R_k$, $\forall k \in U$, go to Step 3,
        4b) If $\sum_{n=1}^{N} x_{k,n} = R_k$, for any $k \in U$,
            find all $n \in S_k$, such that, $x_{k,n} = 0$,
            then set $A = A + \{n\}$, $S_k = S_k - \{n\}$.
Step 5: Set $U = U - \{k\}$.
        5a) If $U = \phi$, go to Step 8,
        5b) If $U \neq \phi$, $A = \phi$, go to Step 3.
        5c) Otherwise, go to Step 6.
Step 6: Find $n$ and $k$ satisfying:
            $p_{k,n,1} \leqslant p_{k',n',1}$, $\forall\ k, k' \in U,\ \ n,\ n' \in A$
            set $S_k = S_k + \{n\}$, $A = A - \{n\}$.
Step 7: If $A = \phi$, go to Step 3,
        otherwise, go to Step 6.
```

```
Step 8: Set $U = \{1, \dots, K\}$, $P_T = \sum_{k=1,n=1}^{K,\,N} p_{k,n,x_{k,n}}$.
    If $P_T > P$, stop and output an error message.
    otherwise,
        If $A \neq \phi$, find $n$ and $k$ satisfying:
            $p_{k,n,1} \leqslant p_{k',n',1}, \ \forall\, k, k' \in U, \ \ n, n' \in A$
            set $S_k = S_k + \{n\}, \ A = A - \{n\}$.
        until $A = \phi$.
Step 9: $[n, k] = argmin_{k \in U, n \in S_k} \triangle\, p_{k,n,x_{k,n}}, \ \forall k \in U$,
    set $x_{k,n} = x_{k,n} + 1$, $P_T = P_T + p_{k,n,x_{k,n}}$
    If $x_{k,n} = M$, set $S_k = S_k - \{n\}$,
     otherwise, evaluate $\triangle p_{k,n,x_{k,n}} = p_{k,n,x_{k,n}+1} - p_{k,n,x_{k,n}}$.
Step 10: If $P_T < P$, go to Step 9,
    If $P_T = P$, output the maximal throughput $R = \sum_{k=1,n=1}^{K,\,N} x_{k,n}$, and stop,
    If $P_T > P$, set $x_{k,n} = x_{k,n} - 1$,output the maximal throughput $R = \sum_{k=1,n=1}^{K,\,N} x_{k,n}$,
    and stop.
```

4.4 Simulation Results

4.4.1 System Parameters

An OFDMA system with N subchannels is considered in this thesis. It is assumed that the greatest possible number of bits that can be transmitted by any subchannel in one symbol duration is M. It is also assumed that there are K users and the kth user has a data rate requirement of R_k $(k = 1, 2, \cdots, K)$. The parameters used in the simulations are listed in the following table.

Subchannels number	Users number	Maximal loaded bits (bits/symbol)	Required BER	Data throughput (bits/symbol)	Power constraint (jouls/symbol)
$N = 64$	$K = 2, 4, 8$	$M = 12$	$Pe = 10^{-4}$	$\sum_{k=1}^{K} R_k = 256$	$P = 100$

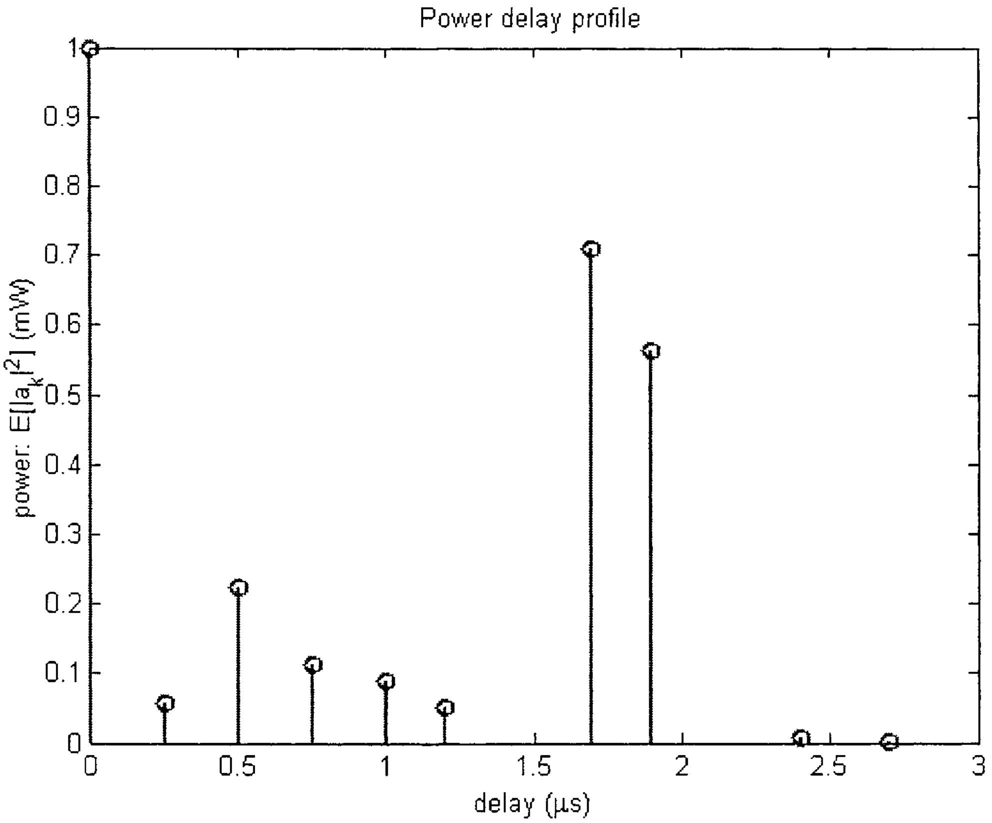


Figure 4-3: An example of the power delay profile of a 10-multipath wireless channel.

The wireless channels for simulation in the thesis are Rayleigh fading channels with 10 multipaths. During the simulation, 100 channels are independently generated. Figure 4-3 and Figure 4-4 are examples of the impulse response and frequency response of a 10-multipath channel in this thesis respectively. The results presented in this thesis are the average of 100 trials.

4.4.2 Performance of Algorithms

Since the proposed branch-and-bound method based algorithm is optimal, it achieves the same results as full-search algorithm. The simulation results of the proposed branch-and-bound method based algorithm and fast bits feeding algorithm are shown in Figure 4-5 and Figure 4-6

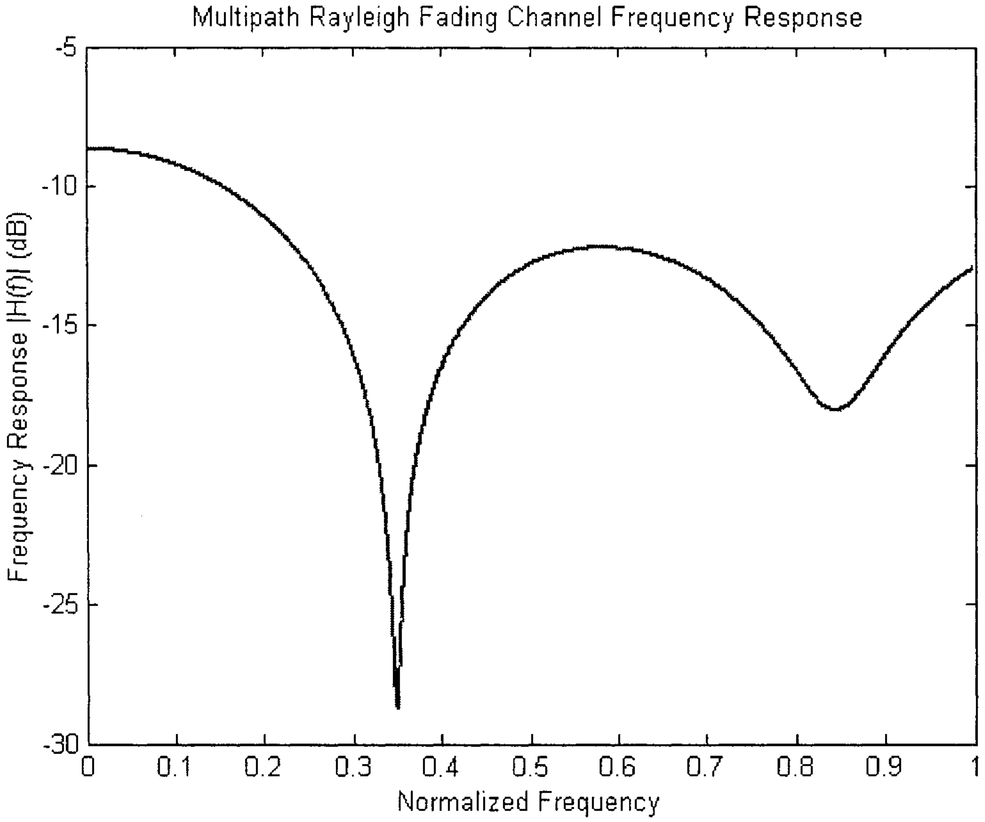


Figure 4-4: An example of the frequency response of a 10-multipath normalized Rayleigh fading channel.

for MA and RA problems respectively.

For comparison, the algorithms in [5], [6] (constructive initial allocation plus iterative improvement) and [7] are simulated as well and the simulation results are shown in Figure 4-5 and Figure 4-6 respectively. From Figure 4-5 and Figure 4-6, it can be seen that with the number of users increasing the total transmission power decreases under limited system data throughput and the maximum data throughput increases under the restricted system transmission power. This is the so-called multiuser diversity among users in different locations.

The simulation results also show that the performance of fast bits feeding algorithm is very close to that of optimal algorithm. In fact, the difference between the results from the fast bits feeding algorithm and that from optimal algorithm is less than 2%. For RA problem, fast bits feeding algorithm provides much better performance than the algorithms in [5] and [7]. For MA problem, fast bits feeding algorithm is much better than the algorithm in [6] (constructive initial allocation plus iterative improvement). For MA problem, fast bits feeding algorithm is a little bit worse than the algorithm in [5], the difference is less than 0.1%. However, as will be indicated in Section 4.4.3, fast bits feeding algorithm has a much lower computational complexity than the algorithm in [5].

4.4.3 Comparison of Computational Complexity

The proposed branch-and-bound method based algorithm is an optimal method for finding optimal radio resource allocation scheme without exhaustive enumeration. It achieves the same results as full-search algorithm but with much lower computational complexity. The computational complexities of the proposed branch-and-bound method based algorithm and full-search algorithm for both MA and RA problems are compared in Figure 4-7. For MA problem, when user number $K = 2$, the computing time of the proposed branch-and-bound method based algorithm is 25.0% of that of full-search algorithm; when $K = 4$, it is 21.5%; when $K = 8$, it is only 8.4%. For RA problem, full-search algorithm consumes about 15% more computing time than that in solving MA problem, since there are more steps to consider with one more constraint; but the computing time of the proposed branch-and-bound method based algorithm is less than 2% more than that in solving MA problem, since the procedure solving RA problem is the same as solving MA problem but with larger size matrix.

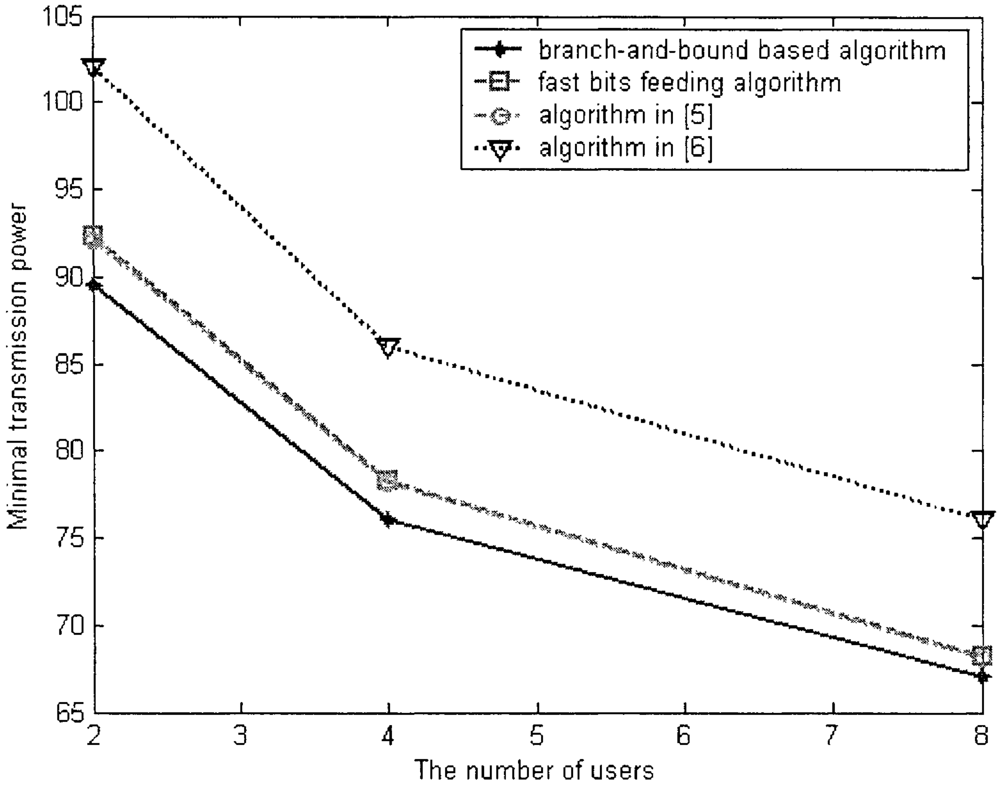


Figure 4-5: The minimal system transmission power (MA problem).

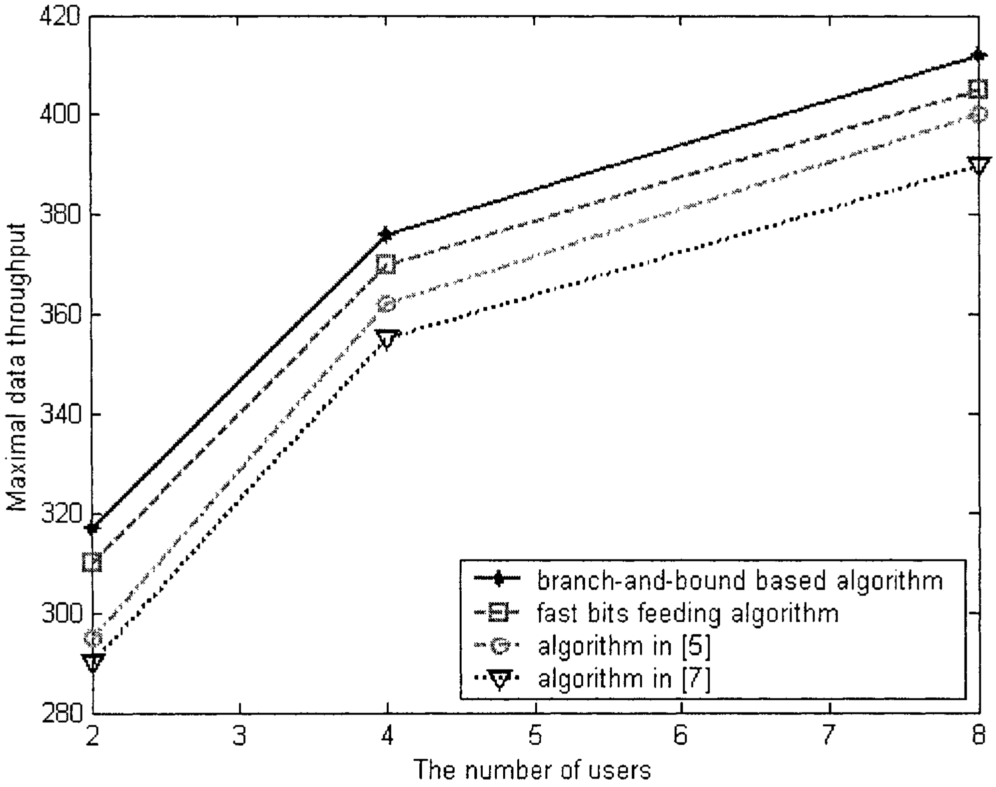


Figure 4-6: The maximal system data throughput (RA problem).

The computing time of the proposed branch-and-bound method based algorithm, fast bits feeding algorithm and the above mentioned suboptimal algorithms for both MA and RA problems are compared in Figure 4-8 and 4-9 respectively. The computing time of fast bits feeding algorithm is less than 10^{-6} of that of the branch-and-bound based method, and less than 10^{-3} of that of the algorithm in [5] for both MA and RA problems, which is less than 0.03 second. The computing time of fast bits feeding algorithm and [6] (constructive initial allocation plus iterative improvement) for MA problem, and the computing time of fast bits feeding algorithm and [7] for RA problem are very close. Fast bits feeding algorithm takes less than 0.5% longer to solve RA problem than to solve MA problem, in order to feed more bits to satisfy more constraints.

From the comparison, it is easy to conclude that the proposed branch-and-bound method based algorithm is optimal and provides a practical comparison guideline. It can also be conclude that the fast bits feeding algorithm proposed in this section is cheap, easy to implement and suitable for real time application with low computational complexity and acceptable results.

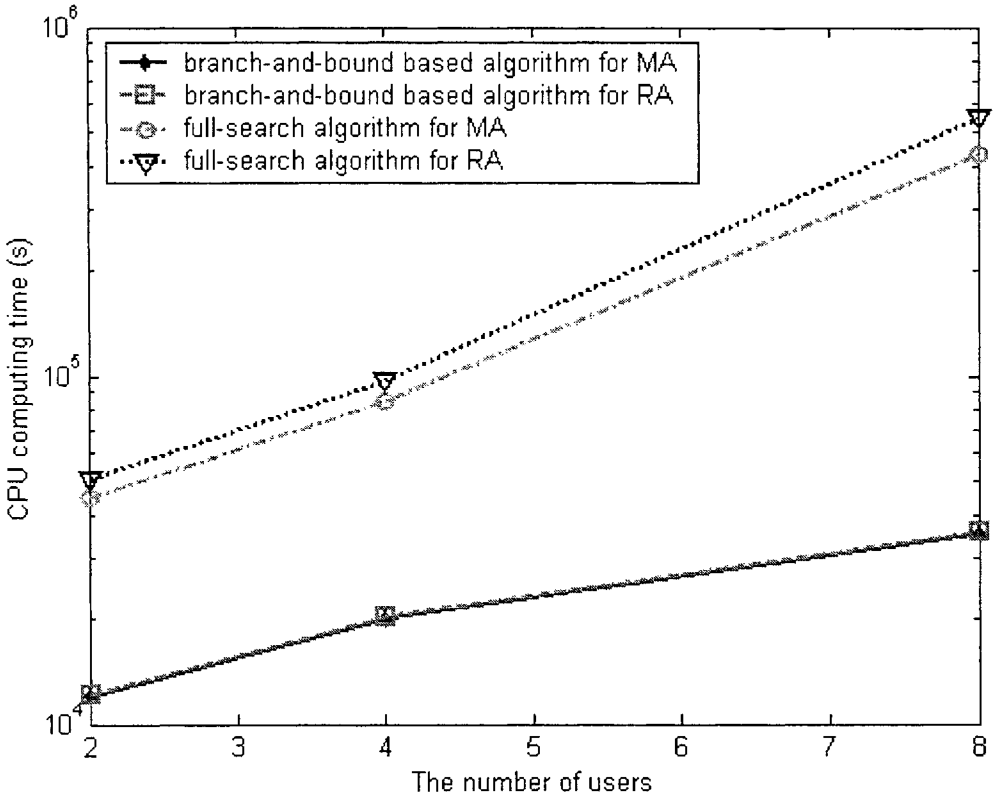


Figure 4-7: Computational complexity comparison of full search algorithm and the proposed branch-and-bound based optimal algorithm.

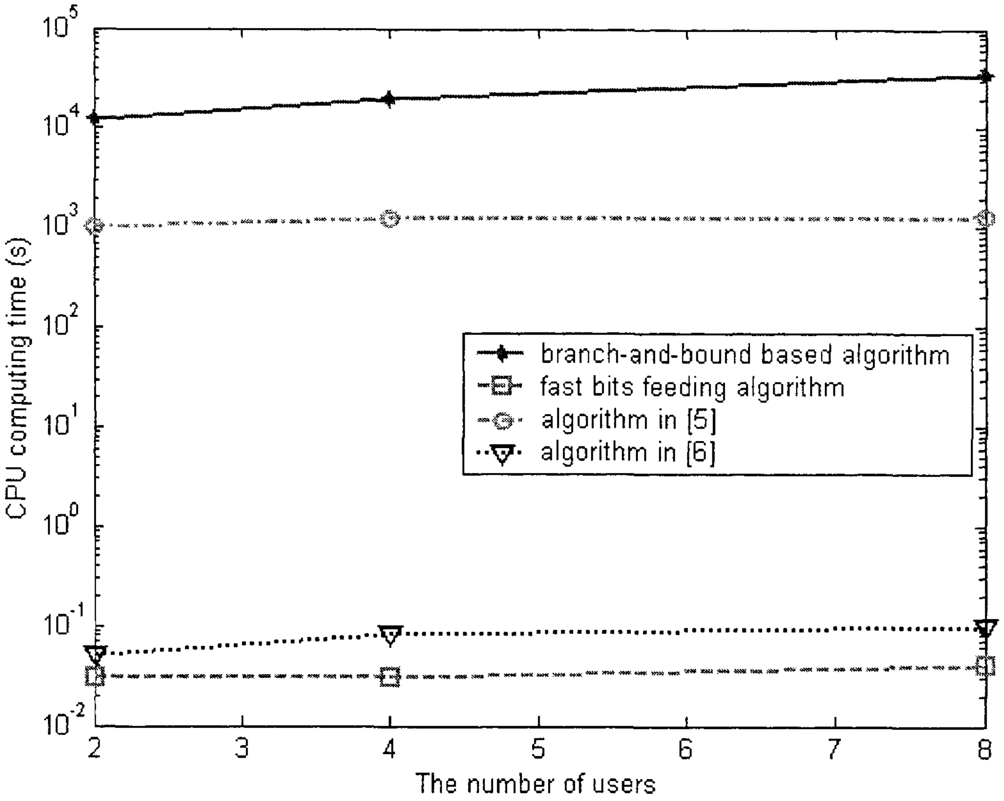


Figure 4-8: Computational complexity comparison of fast bits feeding algorithm, the algorithms in [5] and [6] for MA problem.

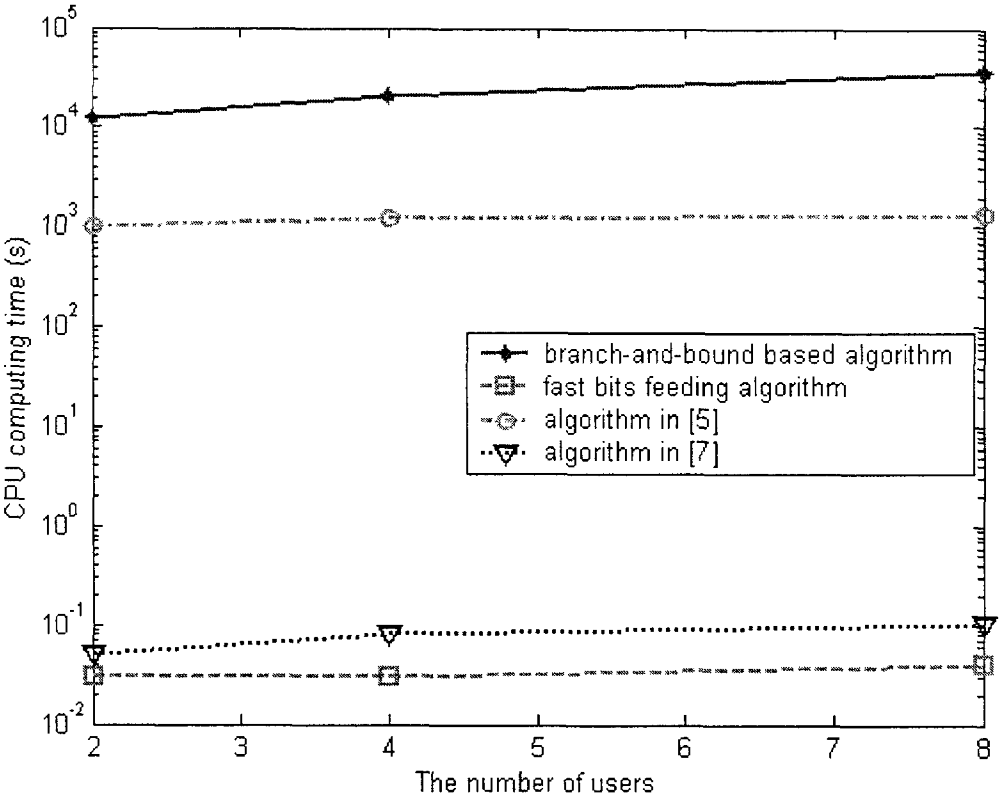


Figure 4-9: Computational complexity comparison of fast bits feeding algorithm, the algorithms in [5] and [7] for RA problem.

Chapter 5

Conclusions and Future Studies

In this thesis, the nonlinear optimization problems associated with the MA and RA problems in OFDMA systems are converted into linear optimization problems with integer variables. By using the formulation with constraint considered, the computational complexity of the full-search algorithm can be reduced. A branch-and-bound method based optimal resource allocation algorithm is proposed. Based on LP method, the proposed branch-and-bound algorithm solves the MA and RA optimization problem without exhaustive enumeration. It offers the same performance as the optimal one achieved by using exhaustive full-search algorithm, but with significantly reduced computational complexity. Based on the formulation with constraint considered, an efficient suboptimal fast bits feeding algorithm is proposed based on the constant transmission power matrix. The proposed fast bits feeding algorithm does not need any bit loading or transmission power distribution assumptions which are the basis of many previous suboptimal algorithms. Simulation results show that the proposed suboptimal algorithm outperforms some other suboptimal methods in the literature.

In this thesis the node checking order of branch-and-bound method based optimal algorithm is assumed from 1 to N which is not necessary to be optimal. If an optimal node order can be applied in this method, the optimal solution can be found in the first few branches and all redundant branches can be pruned at the most beginning. It is my future study.

Bibliography

[1] ZhengdaoWang and Georgios B. Giannakis, "Wireless Multicarrier Communications," *IEEE Signal Processing Magazine,* vol. 47, no. 3, pp. 29-48, May 2000.

[2] Andrea Goldsmith, *Wireless Communications*, Cambridge: Cambridge University Press, September 2005.

[3] Simon Haykin, *Digital communications*, Toronto: John Wiley & Sons Inc, 1998.

[4] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747-1758, Oct. 1999.

[5] Inhyoung Kim, Hae Leem Lee, Beomsup Kim and Y. H. Lee. "On the use of linear programming for dynamic subchannel and bit allocation in multiuser OFDM," *Global Telecommunications Conference, IEEE*, 25-29 Nov. 2001 pp: 3648-3652.

[6] 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," *Proc. VTC'99*, vol. 2, pp. 1124-1128, 1999.

[7] Yingjun Zhang and K. B. Letaief, "Multiuser Subcarrier and Bit Allocation along with Adaptive Cell Selection for OFDM Transmission," *IEEE International Conference on Communications*, no. 1, pp. 861-865, April 2002.

[8] J. M. Cioffi, *Lecture Notes for Advanced Digital Communications*, Stanford, CA: Stanford Univ. Press, 1997.

[9] C. H. Papadimitriou and K. Steiglitz, "*Combinatorial optimization: algorithms and complexity*," Englewood Cliffs, NJ: Prentice-Hall Inc., 1982.

[10] Theodore S. Rappaport, *Wireless Communications Principles and Practice*, 2nd Edition, Upper Saddle River: Prentice Hall, Inc., 2002.

[11] V. H. MacDonald, "The Cellular Concept," *The Bell Systems Technical journal*, vol. 58, no. 1, pp. 15-43, January 1979.

[12] Leon W. Couch, *Digital and Analog Communication Systems*, Upper Saddle River: Prentice Hall, Inc., 1997.

[13] Inseop Lee and W. K. Jenkins, "Adaptive Canceler-equalizer for Digital Communication Channels," *mwscas*, vol. 58, no. 3, pp. 153, 1998.

[14] Kal Kalbasi, "Simulation of Interference and Diversity Scenarios Using Propagation Models," *IEEE Asilomar Conf. on Signals, Systems, and Computers*, vol. 62, no. 5, pp. 328, June 2001

[15] GSM 05.05 Recommendation, *Radio Transmission and Reception.*

[16] Bernard Sklar, *Digital Communications: Fundamentals and Applications*, 2nd Edition, Upper Saddle River: Prentice Hall, Inc., 2001.

[17] W. C. Y. Lee, *Mobile Cellular Telecommunications Systems*, New York: McGraw-Hill Publications, 1989.

[18] R. Van Nee and R. Prasad, *OFDM for Wireless Multimedia Communication*, Artech House Publisher, 2000.

[19] John A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.

[20] S.B. Weinstein and P.M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 628-634, Oct. 1971.

[21] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Systems Technical Journal*, vol. 46, pp. 1775–1796, December 1966.

[22] W.Y. Zou and Y. Wu, "COFDM : An Overview," *IEEE transactions on broadcasting*, Vol. 41, No. 1, pp 1-8, March 1995.

[23] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.,vol.* COM-33, pp. 665-675, July 1985.

[24] W.A.C Fernando and R. M. A. P Rajatheva, "Performance of COFDM for LEO satellite channels in global mobile communications," *IEEE conference proceedings VTC*, Ottawa, Canada, Vol. 1, pp. 412-416, 1998.

[25] "Radio Broadcasting System, Audio Broadcasting (DAB) to Mobile, Portable, and Fixed Receivers," *Eur. Telecommun. Stand.* Inst., Sophia- Antapolis, Valbonne, France, ETS 300 401, 1995–1997.

[26] "Digital Broadcasting System Television, Sound and Data Services; Framing Structure, Channel Coding, and Modulation Digital Terrestrial Television," *Eur. Telecommun. Stand.* Inst., Sophia-Antapolis, Valbonne, France, ETS 300 744, 1996.

[27] "Broadband Radio Access Networks (BRAN); High Performance Radio Local Area Networks (HIPERLAN) Type 2; System Overview," Eur. Telecommun. Stand. Inst., Sophia-Antapolis, Valbonne, France, ETR 101 683 114, 1999.

[28] M. Wahlqvist, C. Östberg, J. Beek, O. Edfors, P. Börjesson, "A Conceptual Study of OFDM-based Multiple Acess Schemes," *Technical Report Tdoc* 117/96, *ETSI STC SMG2* meeting no 18, Helsinki, Finland, May 1996.

[29] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE Journal of Selected Areas of Communications*, pp. 1747-1758, Oct. 1999.

[30] *Dictionary of Algorithms and Data Structures*, Paul E. Black, ed., U.S. National Institute of Standards and Technology, http://www.nist.gov/dads/, June 2005.

[31] Perry Gray, William Hart, Laura Painton, Cindy Phillips, Mike Trahan, John Wagner and Sandia National Laboratories, "*A survey of global optimization methods*," http://www.cs.sandia.gov/opt/survey/branch-and-bound.html, March 10, 1997.

[32] Ömer Benli, "The branch-and-bound approach," Industrial Engineering Applications and Practice, *International Journal of Industrial Engineering*, http://benli.bcc.bilkent.edu.tr/~omer/research/bb.html, 1999.

[33] A. H. Land and A. G. Doig, "An automatic method for solving discrete programming problems," *Econometrica*, Vol.28, pp. 497-520, 1960.

[34] MATLAB, Version 6.1.0.450, Release 12.1, Copyright 1984–2001, The Mathworks, Inc, May 18, 2001.

[35] Z. Mao, X. M. Wang, and J. Lin, "Fast optimal radio resource allocation in OFDMA system based on branch-and-bound method," *Proc. IEEE PacRim Conf. on Communications, Computers, and Signal Processing'2005*, pp. 348-351, Aug. 2005.