THE PERFORMANCE OF OFDMA WIRELESS SYSTEMS USING PF SCHEDULING

RABIE ALMATARNEH

The Performance of OFDMA Wireless Systems using PF Scheduling

by

Rabie Almatarneh

A thesis submitted to
the School of Graduate Studies in
partial fulfillment of the requirement for the degree of
Master of Engineering.

Faculty of Engineering and Applied Science
Memorial University of Newfound

August 3, 2010

St. John's, Newfoundland and Labrador

ABSTRACT

Orthogonal Frequency Division Multiple Access (OFDMA) has gained a lot of attention from many researchers of wireless communications. OFDMA represents a promising multiple access scheme for high-data rate transmission over wireless channels as it combines the Orthogonal Frequency Division Multiplexing (OFDM) modulation and flexible and effective subcarrier allocation [1]. OFDMA inherits the favourable OFDM prosperities of high immunity to the multi-path fading and inter-symbol interference [2]-[3].

One main aspect related to OFDMA is scheduling. Scheduling in OFDMA systems is expected to manage multiple frequency subbands over time to deliver service to the system's users with specific quality requirements. Data rate, throughput, and fairness among users are key factors that specify how efficient a scheduling solution is. The Proportional Fair (PF) algorithm is an appealing scheduling scheme to improve the fairness among users without sacrificing the efficiency in terms of average throughput of the system [4].

This thesis focuses on OFDMA systems that utilize the PF algorithm to schedule the available time and frequency resources among users. The thesis studies the performance of OFDMA scheduling in terms of the throughput, fairness, and packet delay using computer simulations and analysis.

Two dimensional (frequency and time) scheduling algorithms based on the PF algorithm are proposed, evaluated using computer simulations. The proposed solutions utilize the PF criterion to achieve high system throughput while maintaining fairness among users in the system. In order to support multimedia bursty traffic, our solutions allow frequency domain sharing so that more than one user can share a subband in each time frame, where a subband is formulated by grouping a number of subcarriers. We compare

the performance of the proposed solutions with other candidate OFDMA scheduling algorithms in the literature. Results show that the proposed schemes outperform the other schemes in terms of the throughput and packet delay with comparable fairness performance.

It is worth mentioning that, to the best of our knowledge, the performance of PF scheduling scheme for OFDMA systems is not determined analytically in previous work available in the literature, and it is usually found by simulations only. Hence, we analyze the proposed PF scheduling schemes for OFDMA systems and evaluate its performance analytically. We derive closed-form expressions for the average throughput, throughput fairness index, and packet end-to-end delay. Computer simulations are used for verification. It is found that the analytical results agree very well with the results from simulations, which verifies the correctness and accuracy of the analytical solution. In our opinion, deriving analytical expressions that reflect the performance of the PF scheduling algorithm is a significant contribution because it provides deeper insight and ameliorates the understanding of the PF scheduling algorithm behaviour and mechanism for OFDMA systems. This facilitates the process of pursuing future work and further studying and developing more efficient solution for OFDMA systems.

# Acknowledgement

I would like to thank my supervisor, Dr. Mohamed Hossam Ahmed, and my co-supervisor, Dr. Octavia Dobre, for their great supervision during my M.Eng. program. Dr. Ahmed and Dr. Dobre always gave me many valuable suggestions, tremendous help, support, and excellent suggestion for my courses and research. I have learnt a lot from them and cannot thank them enough.

I also want to thank Memorial University for accepting my application on 2006 and giving me the chance to finish my graduate studies in Electrical and Computer Engineering.

I would like to thank the Faculty of Engineering members and staff for their help during my program.

I would like to thank my uncle Dr. Mansour and my parents for their support and sacrifices made while I completed this work.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMC	Adaptive Modulation and Coding
APA	Adaptive Power Allocation
BA	Base Station
BER	Bit Error Rate
CSI	Channel State Information
CAC	Connection/Call Admission Control
CP	Cyclic Prefix
cdf	cumulative distribution function
DSA	Dynamic Subcarrier Allocation
FFT	Fast Fourier Transform
IFFT	Inverse Fast Fourier Transform
MAC	Medium Access Control layer
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OFDMPF	Orthogonal Frequency Division Multiplexing Proportional Fair
PHY	Physical layer
PMP	Point-to-MultiPoint
QoS	Quality of Service
RRM	Radio Resource Management
RR	Round-Robin
SNR	Signal-to-Noise Ratio
SS	Subscriber Station

List of Symbols

i	Index of user.
j	index of subband.
N	Number of users in the system.
M	Number of subbands in the system.
J	The index of the selected user by the PF scheduler.
U	The index of user i on subband j.
$D_{i,j}(t)$	The instantaneous data rate of user i on subcarrier j.
$R_i(t)$	The time average data rate of user i on subcarrier j.
T_c	The observation window measured in time frames.
$S_{i,j}$	The selection indicator. User i selected on subband j.
P_{total}	The total power in the system.
$p_{i,j}$	The transmit power reserved for user i.
Dq_i	The minimum data rate required for user i.
$p_{i,j}$	The transmit power reserved for user i.
x_i	The resources allocated to user i.
α_i	The probability that the buffer of user i is not empty.
E[.]	The average operator.
$Avrg(.)$	The time average
I_{ij}^{N}	The probability of user i ranked N-th to be scheduled on subband j.
$f_{D_{ij}}(.)$	The probability density function.
$F_{D_{ij}}(.)$	The cumulative distribution function.
$F_{(0,1)}(.)$	The standard normal cdf with zero-mean and unity-variance.
ω_i	The average time spent in the system per packet for user i.

μ_i The service rate of user *i*.

λ_i The arrival rate of user *i*.

CHAPTER 1

Introduction

Orthogonal Frequency Division Multiple Access (OFDMA) wireless systems have gained high attention by many researchers in the wireless communications field. The OFDMA systems show very appealing features in the wireless communications market, as they can provide various advantages, such as high data rate, immunity to multipath propagation, and flexibility and efficiency in resource allocation.

Many Radio Resource Management (RRM) schemes, such as scheduling, Dynamic Subcarrier Allocation (DSA), Adaptive Power Allocation (APA) and Connection/Call Admission Control (CAC) were purposely left undefined in the wireless system standards, such as IEEE 802.16 [1], providing a hot topic for research studies. Scheduling is one of the most essential and critical resource management schemes in OFDMA systems.

The ultimate goal of scheduling schemes is to maximize the throughput and the Quality of Service (QoS) in general given a set of available resources in terms of time frames and frequency subcarriers. DSA importance arises from the need to manage the limited radio resources in terms of the allocated number of orthogonal subcarriers to each user in dynamic basis. In order to minimize the interference and power consumption in the OFDMA systems, APA can be considered. The main idea of APA is to assign dynamically each subcarrier a specific level of power. Subcarrier power level is usually determined based on the Channel State Information (CSI) in order to reduce the transmission power and minimize the interference. However, several studies show that APA has little input on the performance of OFDMA systems compared to scheduling and DSA, [4]-[5]. CAC is very critical because it decides whether to admit or reject new connection requests to prevent the system resources from being over used [4]. The CAC admits new connections and handed off connections, depending on the CSI, the loading conditions, and the available resources in order to guarantee a minimum level of QoS.

1.1 General Background

OFDMA systems strength and flexibility arise from the integration of Orthogonal Frequency Division Multiplexing (OFDM) modulation with DSA to support efficient use of the available frequency radio resources [2], [4]. In general, OFDMA systems have similar structures including the physical layer (PHY) and the Medium Access Control layer (MAC) with different variations for each standard [4]. In the first layer, PHY layer, OFDM is utilized as a modulation technique for providing a pool of parallel orthogonal narrowband subcarriers as shown in Fig. 1.1.

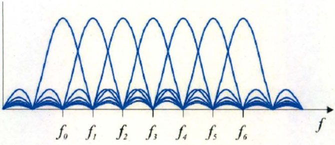

Fig. 1.1: Orthogonal Subcarriers in OFDM signal [6].

These subcarriers can be utilized by the scheduled user in the system. By using OFDM, the data steam of a user is divided into a large number of lower rate data streams, and each stream is carried by a subcarrier that is orthogonal to other subcarriers [7]-[10]. As mentioned above, The most important characteristic of OFDM is the high immunity to the multi-path frequency selective fading and the inter-symbol interference [2]. OFDMA, which is used as a multiple access scheme, allocates to each user in the system a group of the available orthogonal subcarriers. Hence, it is essential to have efficient subcarrier allocation in OFDMA systems in order to utilize these subcarriers efficiently.

The second layer, MAC layer, is responsible for supporting two network communication topologies [2], [4]: Point-to-MultiPoint (PMP) and mesh topologies as shown in Fig. 1.2. The PMP topology supports direct communications between the Subscriber Stations (SS's) to the Base Station (BS). On the other hand, the mesh topology supports communications between SS's with each other and also with the BS directly or by means of relaying. Each SS acts as a possible relay stage for other SS's in the network [2]. The MAC layer should be able to support efficient radio resource sharing between different users with different levels of QoS requirements in the system. Therefore, the provision of efficient scheduling to grant different requests from different service classes is essential for the system performance [11, 12].

RRM in OFDMA systems is a quality-aware management, in the sense that it considers the QoS requirements of existing users. Cross-layer design is implemented from the PHY and MAC layers to meet the QoS requirements. Thus, cooperation between the PHY and MAC layers should be considered [8]. The cooperation is achieved by sending the CSI from the PHY layer to the MAC layer, so that the Adaptive Modulation and Coding (AMC) can be applied on different subcarriers dynamically. The structure of OFDMA systems as shown in Fig. 1.3 is designated to control multiple users contending for available orthogonal subcarriers simultaneously. The OFDMA system contains *N* users with different data rate requirements. Subcarriers with different power levels are mapped to the requesting users in the subcarrier and power allocation block. The decision on the number of allocated subcarriers and power levels is affected by the CSI. The signals are transmitted after computing the Inverse Fast Fourier Transform (IFFT) and adding the Cyclic Prefix (CP) [2], [10]. The CP mitigates the inter-symbol interference because the CP duration exceeds the delay spread of the channel. At the receiver side, the CP is removed, and the Fast Fourier Transform (FFT) is computed. Finally the information is extracted for each user by using symbol to bit streams mapping based on the subcarrier allocation and power level information, which is available at the receiver side. This information can be provided to the receiver through feedback channels. Also, the CSI of each user should be fed-back to the transmitter in order to guarantee accurate and efficient RRM.

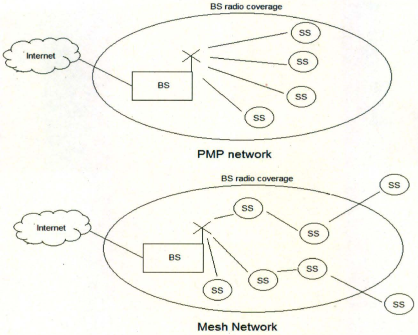

Fig. 1.2: OFDMA system topologies [2].

Generally, the OFDMA system structure contains N users and M subcarriers. The wireless signals are transmitted over multipath fading channels. A number of replicas, with different amplitudes and arrival times of the original signal are received. Besides the direct line-of-sight radio wave, there is a large number of reflected radio waves that received at the receiver with different time and phase. These multipath replicas can be added destructively or constructively depending on the relative phase of these multipath signals. Hence, the received signal experiences dynamic fluctuations (fading dips and up fading peaks). These dynamic fluctuations vary significantly from one frequency to another. In order to compensate for the dynamic changes due to multipath fading, the AMC is applied on subcarriers separately. AMC modulates and encodes the information in each subcarrier based on the Signal-to-Noise Ratio (SNR), independently. AMC improves the utilization of the OFDMA subcarriers because a subcarrier will not be used only if it appears in very bad channel conditions to all existing users in the system.

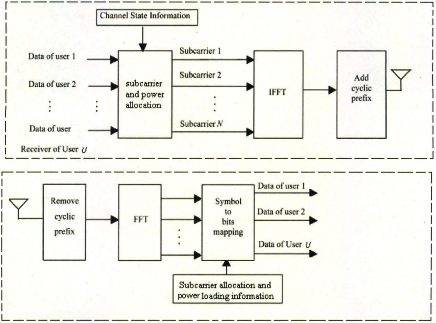

Fig. 1. 3: The structure of OFDMA system [10].

The ultimate objective of RRM in multi-user OFDMA systems is to maximize the total throughput with acceptable fairness behaviour. The secondary objective of RRM in multi-user OFDMA systems is to minimize the total transmitted power. On one hand, the throughput maximization goal must be perused under the condition of limited number of subcarriers. On the other hand, the total power that is allocated to users should not exceed the total transmit power available in the system.

OFDMA systems allow users to share time and frequency resources. Sharing resources introduces the problem of contention for resources. Hence, scheduling is a mandatory requirement while designing OFDMA systems. Scheduling manages, multiplexes, and divides available resources between competing users in the waiting queue based on a certain scheduling policy. Scheduling policy represents the behaviour of a scheduler while deciding which request to serve next, and how many resources to grant to a request. Scheduling policy is important because it impact the performance of the system including the fairness among users. There is always a trade-off between fairness and aggregate performance. For example, Round-Robin (RR) scheduler divides the available resources among users equally but with low performance. On the other hand, opportunistic algorithms grant resources to the requests that maximize the system performance sacrificing fairness among users. It is important and critical requirement to design scheduling algorithms that provide high performance with fair treatment among the requests in the waiting queue.

1.2 Proportional Fair Scheduling Algorithm

A scheduling algorithm is an event planner that switches between many users or parties that competing simultaneously and asynchronously request access to the same resource. The resources can be expressed as time of service, such as time to run a job on a computer processor or time to use a transmission channel.

In OFDMA systems, resources are divided into two categories, time and frequency. Time resources are expressed as time frames, where a time frame consists of time symbols. On the other hand, frequency resources are formulated as orthogonal subcarriers or frequency subbands, which are constructed as group of orthogonal subcarriers. Also, OFDMA systems are expected to serve multiple users requesting resources simultaneously. So, scheduling in OFDMA systems should be capable of managing and mapping orthogonal subcarriers and time frames to multiple users in order to maximize the performance of OFDMA systems.

The PF algorithm is an efficient scheduling solution that can guarantee fairness among competing users while achieving relatively high performance in terms of throughput. With this algorithm, the level of satisfaction (or starvation) of each user in the system is measured over time, and resources are assigned to users based on that. Moreover, the PF algorithm is flexible and can scale between fairness and efficiency.

The general subcarrier allocation problem can be solved by selecting user i using the following PF criterion [13]-[15],

$$J = \arg \max_{1 \leq i \leq N} D_{i,j}(t) / R_i(t-1), \tag{1.1}$$

Where J is the index of the selected user, $D_{i,j}(t)$ is the instantaneous data rate of user i on subcarrier j, and $R_i(t)$ is the time average data rate of user i on subcarrier j. $R_i(t)$ is updated for all users on each subcarrier as follows [15],

$$R_i(t) = \begin{cases} (1 - \frac{1}{T_c})R_i(t-1) & , \; i \neq k, \\ (1 - \frac{1}{T_c})R_i(t-1) + \frac{1}{T_c} D_{i,j}(t) & , \; i = k, \end{cases} \tag{1.2}$$

where T_c represents the observation window expressed in time frames. It should be noted that the value of T_c can change the trade off between the throughput and fairness. Larger values of T_c imply that the scheduling algorithm selects users in a more greedy approach by averaging $R_i(t)$ over larger number of time frames to maximize the system's throughput with the cost of less fairness among users. On the other hand, smaller T_c values indicate more fairness among users is applied by averaging $R_i(t)$ over less number of time frames at the expense of throughput degradation.

1.3 Literature Review

Two main categories can be identified while studying the RRM of OFDMA systems in the literature. The first category investigates the OFDMA systems in order to maximize the system data rate, while the second category studies the possibilities of minimizing the total power consumption:

1.3.1 Rate Aware Algorithms

One of the main goals of the RRM in OFDMA systems is to improve the efficiency and to maximize the total throughput. Many researchers studied this problem and proposed several solutions [2], [9], [13], [15]-[18]. On most of the work, the problem was identified without considering the APA problem in order to reduce the complexity of the implemented solutions. The transmission power in such solutions is simply assumed to be equal for all users in the system.

The Hungarian method is presented in literature as the optimal solution for this problem. The Hungarian algorithm is implemented in [2]-[3], [9], [13]-[14], and [18]. This solves the assignment problem between users and subbands in order to achieve the maximum possible throughput. It allocates one subband to one user in every scheduling time in a fashion that guarantees maximum profit. The Hungarian algorithm grants resources to any user regardless how bad its transmission rate is. However, the optimal solution for this problem is computationally complex, as it solves the problem with a complexity of order $O(N^3)$ [2]. Many studies proposed near-optimal solutions which provides throughput close to the optimal one, with lower computational burden [9]. In [9], a suboptimal subcarrier allocation algorithm with equal power distribution is presented. This algorithm aims to maximize the global throughput by permitting users to access subcarriers with the best channel conditions as much as possible. To attain fairness among users in [9], the user with the worst satisfaction has the highest priority to choose the subcarrier with highest transmission gain to be used in the next allocation iteration. Another suboptimal iterative solution is presented in [2] and [18]. The iterative algorithm in [18], called *max-max* scheduling algorithm, establishes a profit matrix which includes the achievable data rate on each subcarrier for each user. The algorithm iteratively finds the largest element in the matrix, and maps the corresponding user to the corresponding subcarrier. Then, the mapped user row and subcarrier column are removed from the profit matrix. This procedure is repeated until all the subcarriers are assigned to users. The iterative algorithm in [18] is very simple. However, many frequency resources can be wasted and not allocated due to removing the partially allocated subcarriers from the profit matrix, when same users do not have enough traffic to fully utilize the assigned subcarriers, particularly when they have bursty traffic.

In [3], [9], and [13]-[14], subcarrier allocation algorithms, based on the PF criterion, are presented. In [3], three subcarrier allocation algorithms are proposed. The main contribution of these algorithms is the modification of the conventional PF criterion presented in the previous section in order to meet the QoS requirements of users and attain fair treatment. The authors implemented the proposed algorithms by using different combinations of a data rate exponential weight moving average proposed in [19] and the data rate requirements of users. However, the three solutions in [3] do not effectively manage the resources in the frequency domain. All three solutions allow one user to access all frequency subcarriers at a scheduling decision. In this case, scheduling considers the only OFDM technique and does not meet the requirements and abilities of OFDMA systems. Hence, no user diversity is considered. Allowing only one user to access all the frequency subcarriers at a given time frame eliminates the chance of other users with favourable transmission rate to access and utilize excess subcarriers because the current user is blocking other users. Such behaviour severally affects bursty traffic QoS requirements. It is obvious that it is more effective to utilize OFDMA when managing OFDM subcarriers, so subcarriers can be grouped into subbands and each subband can be accessed by different user, which triggers the multiuser diversity.

1.3.2 Power Aware Algorithms

In order to approach the power-aware solution, pre-defined performance specifications are required. The power minimization problem is solved under limited total transmit power or bit error rate constrains [9], [20].

One candidate solution to this problem is the use of greedy algorithms [17], [19]. The greedy algorithm proposed in [17] is applied after candidate subcarriers are assigned to each user. It iteratively computes, for all candidate subcarriers, the additional power that is needed to load an additional bit. The power increment is applied on the subcarrier with the lowest power demands. The algorithm finishes the power allocation when all bits are assigned to subcarriers.

Another popular solution proposed in the literature is the water-filling solution [9], [21], and [22]. The water-filling algorithm maximizes the capacity of each subcarrier under the condition of limited power. The general behaviour of the water-filling algorithm is as follows. The algorithm equally distributes the limited total power among subcarriers, and then iteratively compares the assigned power to subcarriers with their required power. If the power assigned to a subcarrier is larger than its required power, the algorithm takes the redundant power and assigns it to the subcarrier whose required power is less than the assigned power. This process continues until there is no extra power can be reassigned.

In [23], a proportional fair power allocation algorithm is proposed. This algorithm is implemented with the objective of attaining better fairness than that achieved with the conventional water-filling algorithms. The authors in [23] implemented a power allocation algorithm that can measure the data rate satisfaction of the users. The rate satisfaction is reported by using a fairness metric that tracks the normalized increment of the data rate caused by incrementing the power on each subcarrier. Accordingly, the algorithm allocates more power to a subcarrier whose normalized incremental data rate is larger.

Another approach in the literature to solve the power allocation problem is given in [8], [21], and [24]. A utility functions, $U(.)$, is used to satisfy an equilibrium between the efficiency and fairness in the OFDMA systems. The main idea of implementing power allocation algorithms using the utility functions is performed by expressing the achievable throughput as a function of the transmission power $f(x)$, such that an optimal power allocation can be solved based on the utility function $U(f(x))$.

1.4 Problem Statement

OFDMA resources in terms of frequency and time are limited and expensive. OFDMA systems, such as IEEE 802.16, are expected to deliver high data rate services to multiple users [1]. Efficient scheduling algorithms can significantly improve OFDMA systems performance [4]. There have been increasing attempts to design and develop OFDMA scheduling algorithms that improve the utilization of OFDMA recourses and meet user QoS requirements.

PF scheduling algorithm is an appealing solution for scheduling OFDMA time and frequency resources. However, to the best of our knowledge, analytical studies for PF scheduling solutions proposed to schedule OFDMA systems resources have not been proposed so far in the open literature. Scheduling algorithms based on the PF criterion are only addressed and developed by computer simulations. Therefore, there is necessary to develop analytical models that represent PF scheduling solutions performance for OFDMA systems over wireless channels, where multipath fading, shadowing, and path-loss are considered.

1.5 Thesis Contribution

This thesis aims to investigate the radio resource allocation in OFDMA systems in order to introduce scheduling schemes that can improve OFDMA system performance in terms of throughput, fairness, and packet delay. We build simulation tools to analyze the performance of the proposed solutions.

We define critical factors that should be considered while developing scheduling algorithms for OFDMA systems. Considering multiuser diversity and frequency diversity is mandatory in OFDMA scheduling algorithm. Pursuing multiuser diversity and frequency diversity can dramatically improve the OFDMA system performance. Seeking short-term fairness among users in OFDMA systems dramatically decreases performance in terms of throughput without improving fairness among users. This thesis shows that performance can be improved by permitting opportunistic scheduling in the short-term while the long-term fairness among users is maintained.

This thesis proposes three scheduling algorithms for OFDMA systems, which are developed using the PF criterion. Our proposed algorithms exploit multiuser, time, and frequency diversity. The performance of the proposed scheduling algorithms is compared to other scheduling algorithms found in the literature. Our scheduling algorithms show better performance than the scheduling algorithms found in the literature in terms of throughput, fairness, and delay.

It is worthy to notice that few studies investigated the PF scheduling in OFDMA systems analytically. The literature lacks studies that provide analytical models that analyze OFDMA systems with PF scheduling. An analytical method, which is based on the Gaussian approximation of the instantaneous data rate in a Rayleigh fading environment, is used to analyze the performance of PF scheduling in [25]. However, this method is developed for single-carrier systems and is limited to the case of users with full buffers. We derive closed-form expressions for the throughput, fairness index, and average packet delay for the modified *Max-Max* scheduling algorithm in OFDMA systems under Rayleigh channel and bursty traffic conditions. We verify our analytical solution by comparing the analytical results with computer simulation results.

1.6 Thesis Organization

The organization of this thesis is as follows:

1- In Chapter 1, general background on OFDMA systems and its resources management is provided. Also, literature review that presents solutions and approaches to solve the RRM in OFDMA systems is provided.

2- In Chapter 2, the RRM in OFDMA systems is presented. Also the OFDMA system model including the wireless channel and the users' specifications are presented.

3- In Chapter 3, The Horizontal and Vertical scheduling algorithms that aim to schedule the OFDMA radio resources in joint frequency-time domains using PF criteria are proposed. The Horizontal scheduling algorithm manages the radio resources subband by subband; in other words, it allocates all the resources in a subband and then allocates all the resources in the next subband, and so on. The Vertical scheduling algorithm, manages subbands iteratively; the algorithm assigns sufficient resources from a subband to the selected users every scheduling decision. We compare our proposed algorithms with algorithms that peruse direct fairness among users, such as the Hungarian scheduling algorithm. Our algorithms show dramatical improvement in the system throughput without scarifying the fairness among users.

4- In Chapter 4, we propose a scheduling algorithm that exploits the multiuser diversity in both time and frequency domains. The proposed algorithm utilizes the PF criterion to achieve fairness among users in the system. In order to support multimedia bursty traffic, our algorithm allows more than one user to share a subband in each time frame. We also provide analytical evaluation of the performance of PF scheduling algorithm in OFDMA systems. We derive closed-form expressions for the average throughput, Jain's fairness index, and packet delay as performance metrics. We verify the correctness and accuracy of the derived closed-form expressions through simulations. Analytical and simulation results are in very good agreement, which validates our analytical performance analysis.

5- In Chapter 5, conclusions are drawn and directions for future work are presented.

CHAPTER 2

Radio Resource Management in OFDMA Systems and System Model

Introduction

OFDMA system resources in terms of frequency subcarriers and time frames are critical and limited. OFDMA systems, such as those conforming to the IEEE802.16 standard, are expected to deliver high data rate services to multiple users concurrently with varying QoS requirements. In order to meet the QoS requirements for different users, OFDMA systems should employ efficient RRM schemes. The users' QoS minimum demands need to be met under the condition of limited frequency subcarriers and time frames. Because OFDMA systems consider QoS requirements of existing users, cooperation between the PHY and MAC layers should performed [25]. The cooperation is achieved by sending the CSI from the PHY layer to the MAC layer, so that the AMC can be applied on different subcarriers dynamically [8], [21].

This chapter formulates the RRM problem in OFDMA systems. The main goal of RRM in such systems is to maximize the utilization of the limited orthogonal subcarriers and time frames. However, minimizing the transmit power consumption in OFDMA system is also another goal that can be pursued. Effective management of the available orthogonal subcarriers and time frames improves the system efficiency in terms of throughput, while minimizing the transmit power reduces the effect of the interference, which also slightly improves the system efficiency and reduces the cost of data transmission. In this chapter, we present the general throughput maximization and power minimization problems formulation.

We also present the realistic system model we consider through this thesis. The wireless channel in our model considers the path-loss, Rayleigh fading, and shadowing, and is dynamic. The channel conditions change over time from user to user depending on the evaluated path-loss, multipath fading, and shadowing. Multiple replicas of the transmitted signal reflect and scatter on objects while propagating through the channel. Each replica can add destructively or constructively at the receiver depending on its phase. Shadowing is caused by barriers that exist for a long time between the transmitter and receiver. Shadowing also changes over time, but slower than multipath fading. Additionally, we present details of other OFDMA system requirements and specifications we use throughout the thesis.

2.1 General RRM Problem Formulation in OFDMA Systems

The ultimate objective of RRM in multi-user OFDMA systems is to maximize the total throughput with acceptable fairness behaviour, and to minimize the total transmitted power. On one hand, the throughput maximization goal must be perused under the condition of limited number of subcarriers. On the other hand, the total power that is allocated to users should not exceed the total transmit power available in the system. The general formulation of the subcarrier and power allocation in the OFDMA systems can be described as follows [7], [17], and [27]:

$$Maximize: \quad \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} S_{i,j}, \tag{2.1}$$

$$Minimize: \quad \sum_{i=1}^{N} \sum_{j=1}^{M} p_{i,j} S_{i,j}, \tag{2.2}$$

subject to the following conditions:

$$C1: \sum_{i=1}^{N} S_{i,j} = 1, \quad \forall j \in \{1,2,3,\dots,M\},$$

$$C2: \sum_{j=1}^{M} S_{i,j} = 1, \quad \forall i \in \{1,2,3,\dots,N\},$$

$$C3: p_{i,j} \geq 0, \quad \forall i, j,$$

$$C4: \sum_{i=1}^{N} \sum_{j=1}^{M} S_{i,j} p_{i,j} \leq P_{total},$$

$$C5: \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} S_{i,j} \geq Dq_i .$$

where $D_{i,j}$ is the data rate of user i on subband j and $S_{i,j}$ is a selection indicator that equals 1 or 0, with $S_{i,j}$=1 meaning that the subband j is allocated to user i. $p_{i,j}$ is the transmit power reserved for user i at subcarrier j, P_{total} is the total power, and Dq_i is the minimum data rate required for user i. Due to limited number of subcarriers and transmit power in a system, conditions *C1* and *C2* are the constraints on the subcarrier allocation, while *C3* and *C4* are the constraints on transmit power on each subcarrier for all users. *C5* is placed to guarantee the target QoS requirements for all users in the system.

Solving this problem is very complex, so it is more feasible and effective to drop one of the optimization goals and solve for the other. This can be achieved by either assuming equal transmit power for all users, while perusing throughput maximization, or assuming that frequency and time resources scheduling is already performed while perusing transmit power minimization. The next two subsections present the throughput maximization and transmit power minimization problems formulation, respectively.

2.2 Throughput Maximization Problem Formulation

One of the main goals of the RRM in OFDMA systems is to improve the efficiency and maximize the total throughput. Many researchers studied this problem and proposed several solutions [3], [16]-[18], and [27]. However, in order to reduce the complexity of the implemented solutions, this was investigated without considering the power minimization. The transmission power in such solutions is simply assumed to be equal for all users in the system. The throughput maximization is formulated as a maximal bipartite matching problem as follows [17],

$$Maximize: \quad \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} S_{i,j}, \tag{2.3}$$

subject to:

$$C1: \sum_{i=1}^{N} S_{i,j} = 1, \quad \forall j \in \{1,2,3,...,M\},$$

$$C2: \sum_{j=1}^{M} S_{i,j} = 1, \quad \forall i \in \{1,2,3,...,N\},$$

$$C3: p_{i,j} = P_{total} / N, \quad \forall i, j,$$

$$C4: \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} s_{i,j} \geq Dq_i.$$

2.3 Power Minimization Problem Formulation

In order to approach the power minimization, pre-defined performance specifications are required. The power minimization problem is solved under limited total transmit power or bit error rate constrains [9]. The power allocation optimization problem under fixed subcarrier allocation is described as follows [9], [27]:

$$Minimize: \quad \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} p_{i,j} \tag{2.4}$$

subject to:

$$C1: \sum_{i=1}^{N} S_{i,j} = 1, \quad \forall j \in \{1,2,3,...M\},$$

$$C2: p_{i,j} \geq 0 \quad \forall i, j,$$

$$C3: \sum_{i=1}^{N} \sum_{j=1}^{M} p_{i,j} S_{i,j} \leq P_{total},$$

$$C4: \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} S_{i,j} \geq Dq_i.$$

2.4 OFDMA System Model

In this thesis, we consider a single cell downlink scenario for an OFDMA wireless system. The smallest data entity which the base-station can handle is a data packet. We consider fixed size packets. The traffic of the users is assumed to follow Poisson traffic model [6]. The inter-arrival time between packet requests is modeled as a random variable with an exponential distribution. The user's location in the cell is a random variable following uniform distribution.

Each subband contains a number of adjacent subcarriers, which are highly correlated in frequency domain. In time domain, the frame duration is divided into time symbols. Fig. 2.1 shows the frequency-time resources of the OFDMA system. As shown in this figure, the minimum allocable resource unit is represented by the intersection between a symbol in time domain and a subband in frequency domain.

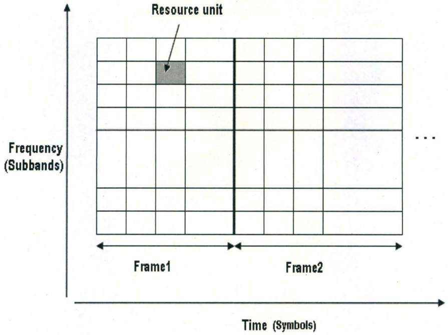

Fig. 2.1. Frequency-time OFDMA resources.

The goal is to allocate the resource units to users with time and frequency diversity in order to maximize the total system throughput with reasonable fairness among users. Moreover, AMC is used to enhance the subbands efficiency. The suitable modulation level and coding rate are decided depending on the CSI for each subband. Table 2.1 shows the AMC schemes used in this thesis along with the corresponding SNRs.

Path loss, shadowing, and small scale fading are considered in the SNR calculation. Correlated Rayleigh fading is assumed between subcarriers. In order to meet the target Bit Error Rate (BER) in the system, we consider the worst case subcarrier fading in each subband for the SNR and link budget calculations. Although the worst case subcarrier fading is considered in a subband, the overall SNR calculation does not significantly change because the fading difference between subcarriers within a subband is insignificant because the fading is highly correlated. Table 2.2 shows default simulation parameters that are used in this thesis.

Table 2.1. AMC Schemes.

Modulation format	Code rate	Bits/symbol	SNR (dB)
BPSK	1/4	1/4	-2.9
BPSK	1/2	1/2	-0.2
QPSK	1/2	1	2.2
8PSK	1/2	3/2	5.2
8PSK	2/3	2	8.4
64QAM	1/2	3	11.8
64QAM	2/3	4	15.1

Table 2.2. Simulation Parameters.

Parameter	Value
Bandwidth	20 MHz
Carrier frequency	2 GHz
Number of subcarriers	256
Number of subbands	32
Transmit power	0 dBW
Noise power	-130 dBW
Path loss exponent	4
Shadowing standard deviation	10 dB
Number of users	32
Packet size	180 bits
Frame duration	2 ms
Symbol duration	16 μs
Cell radius	1500 m

2.5 Correlated Rayleigh Fading Generation

In RRM for OFDMA systems, a frequency resource represents a group of orthogonal subcarriers to be allocated to users. We group a certain number of subcarriers into a virtual frequency subband, with the subband as the smallest entity to be allocated to a user in the frequency domain. Grouping subcarriers into subbands is more practical and reduces the computational complexity and burden of scheduling solutions when compared to considering a single subcarrier as the smallest entity. Because frequency subcarriers are relatively narrow in terms of frequency bandwidth, a user might need to access several subcarriers to transmit or receive a single data packet. Hence, it is reasonable to group the frequency subcarriers into virtual subbands to minimize computational burden.

Clustering frequency subcarriers into virtual subbands cannot be achieved arbitrarily. Because of the dynamic wireless channel, subcarriers experience different multipath fading at different times. In order to formulate a consistent frequency subband, the group of the selected subcarriers should experience similar multipath fading conditions, so that an AMC scheme can be selected for all of them. Applying the same AMC on subcarriers with different channel conditions affects the performance and reliability of the OFDMA system. Note that the AMC scheme which corresponds to the worst subcarrier channel conditions is selected. This yields a reduction in the performance because some subcarriers might experience high channel gain compared to other subcarriers in the same subband, and a lower AMC selected on that subband does not efficiently utilize those subcarriers. By applying the AMC scheme which corresponds to the best subcarrier channel conditions in the subband to the whole subband, a higher AMC scheme will be applied to subcarriers experiencing bad channel conditions, which increases the BER and reduces the OFDMA reliability.

The frequency subcarriers considered in this thesis have correlation in the frequency domain. The fading affecting the frequency subcarriers have cross correlation because of the small coherence bandwidth of the wireless channel [28]. A frequency selective Rayleigh faded channel is modeled based on [28]-[30]. The frequency selective Rayleigh subcarriers are generated with correlation between them in the frequency domain where the complex valued correlation is formulated as function of frequency separation between the subcarriers. In order to minimize the BER and improve the OFDMA system reliability, we consider the worst case subcarrier fading in each subband for the SNR and link budget calculations. Although the worst case subcarrier fading is considered in a subband while selecting an AMC scheme, the overall SNR calculation does not significantly change because the fading difference between subcarriers within a subband is insignificant because the fading is highly correlated.

2.6 Jain's Fairness Index

Jain's fairness index is utilized in this thesis to measure the fairness of the scheduling algorithms for OFDMA systems over a period of time. The Jain's fairness index is a number that expresses how fair is a resource distribution among users who contend for the same resources within one system. Theoretically, its maximum value reaches 1 when all users are assigned exactly the same share of resources, and its minimum value reaches 0 when all resources are assigned to one user only. The Jain's fairness index can be expressed as [31]

$$J(x_1, x_2, x_3, \ldots, x_N) = \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N \sum_{i=1}^{N} x_i^2}. \tag{2.5}$$

where x_i represent the resources allocated to user *i*, the number of users in the system equals *N*. The Jain's fairness index is evaluated every time x_i is updated in the system. Averaging the Jain's fairness index overtime guarantees accurate results that reflect the fairness of the system.

CHAPTER 3

Downlink Frequency-Time Scheduling Algorithms for OFDMA Systems

Introduction

In this chapter, we study the scheduling problem in the OFDMA systems. We investigate the performance and limitations of candidate OFDMA scheduling algorithms found in the literature, and we propose two OFDMA frequency-time scheduling algorithms that aim to maximize utilization of the time and frequency resources in the OFDMA systems with acceptable fairness among users competing in the OFDMA system.

The RR, Hungarian, and *Max-Max* scheduling algorithms are presented as candidate solutions for scheduling the OFDMA time and frequency resources. We study the RR algorithm as a base line for comparison purposes. The RR algorithm allocates equal proportions of resources to all users in cyclic fashion to attain equal fairness among users. The Hungarian algorithm can be utilized to solve either the cost minimization or profit maximization problems. In our case, the Hungarian algorithm is employed to solve allocation maximization problem (maximize profit). The Hungarian algorithm allocates resources (frequency subband and time symbols in the OFDMA systems case) to users in a pattern that guarantees maximum possible throughput, under certain conditions. The Hungarian algorithm maps all the available resources to all users requesting service in every admission decision in a way that guarantees the maximum global utilization. The *Max-Max* scheduling algorithm is a two dimensional matrix-based heuristic version solution of the Hungarian algorithm that aims to iteratively allocate resources to users in order to deliver near optimal performance with lower complexity. The Hungarian algorithm and its heuristic versions, such as the *Max-Max* scheduling algorithm, cause performance degradation because such solutions do not exploit multiuser diversity. Multiuser diversity means that there is likely that a user exists with a very good channel conditions at any time, which can improve the throughput in a wireless system with users faded independently. Also, such solutions pursue short-term fairness (fairness within one time frame only) by allocating all requesting users which does not improve the long-term fairness (fairness within over many time frames) among users and severely decreases the throughput. In addition, the Hungarian and *Max-Max* scheduling algorithms schedule subbands to users in a one-to-one fashion, where a subband cannot be accessed more than one user within a time frame and a user cannot access more than one subband within a time frame. Not allowing subband sharing leads to low subbands utilization because a user sometimes does not need all the resources of the subband in order to meet the QoS requirements which leaves the remaining resources (frequency subcarriers and time symbols) unutilized.

We propose two algorithms in order to solve the above issues (multiuser diversity, fairness, and subband sharing). The two scheduling algorithms are developed to schedule the OFDMA system's frequency subband and time frames with the goal of maximizing throughput with acceptable fairness among users. These two algorithms are referred to as the Horizontal scheduling and Vertical scheduling algorithms. The former manages subbands one by one, while the latter manages the subbands in a cyclic fashion. The efficiency of the proposed algorithms is investigated by computer simulations in terms of throughput and fairness expressed by the Jain's fairness index. Simulations show that the proposed algorithms outperform the one-to-one scheduling algorithms found in the literature in terms of throughput, with improved fairness.

3.1 The Round-Robin (RR) Algorithm

The RR scheduling algorithm is a simple scheduling algorithm that allocates equal portions of resources (time and/or frequency) to users in a circular order. Fig. 3.1 shows the RR scheduling system containing one server station serving multiple requests cyclically from different queues. The RR algorithm is very simple and does not have priority criterion to differentiate between users. Also, RR is not capable of sensing the requesting users' demands and conditions in terms of traffic (arrival) rate and channel conditions, respectively.

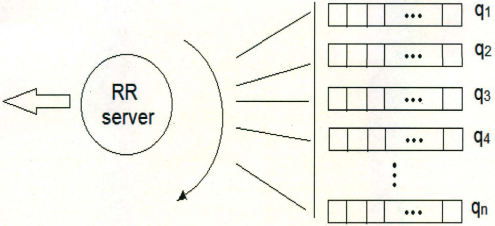

Fig. 3.1. RR scheduling system.

Therefore, the RR scheduling algorithm cannot consider different users' channel conditions and service demands while allocating resources, which severely decreases the system throughput. A large percentage of resources in terms of time symbols and frequency subbands are wasted because the RR server can allocate resources to users with bad wireless channel conditions. Accordingly, the RR scheduling algorithm does not exploit multiuser diversity. Sometimes users are not able to utilize the resources at all due to their extremely bad channel conditions. Other times a small percentage of the allocated resources can be utilized when the channel conditions allow low data rate transmission. Based on that, we conclude that the RR scheduling algorithm delivers very low throughput. Also the RR algorithm cannot guarantee equal treatment among users in wireless systems, although it allocates resources equally to users, due to the variation in the wireless channel conditions over time and users.

3.2 The Hungarian Algorithm

The Hungarian algorithm is a solution to find a maximum or minimum weight matching in bipartite graph matching problems. Given the gain matrix that expresses the possible throughput achieved by every user in subbands, the Hungarian algorithm solves for the mapping pattern that guarantees the maximum throughput as follows:

(1) If the gain matrix is not a square then add dummy columns or rows to make it a square.

(2) Subtract the entries in each row in the matrix from the maximum entry in that entire row.

(3) For each column, subtract the minimum entry in each column from every column entry.

(4) Cover all rows and columns that contain zeroes in the matrix with minimum number of vertical or horizontal lines.

(5) If the number of lines equals the size of the matrix, then solution exists. If the lines covering all of the zeroes are fewer than the size of the matrix, then find the minimum entry that is not covered by a line. Subtract it from all non-zero entries and add it to any entry lies at any intersection of the drawn lines.

(6) Iterate over the previous step until there is a solution where a solution exists at matrix entries with zero values.

We present here an example that clarifies how the Hungarian algorithm works to obtain the solution for the maximization problem. The following 4×4 matrix represents the gain matrix where the rows represent the users, the columns represent the subbands, and the values in the matrix entries reflect the achievable throughput of each user in each subband.

$$x=\begin{bmatrix}15 & 17 & 25 & 26\\ 15 & 17 & 26 & 29\\ 19 & 22 & 33 & 34\\ 21 & 25 & 35 & 40\end{bmatrix}$$

Now we apply step (2) on the matrix:

$$x=\begin{bmatrix}26-15 & 26-17 & 26-25 & 26-26\\ 29-15 & 29-17 & 29-26 & 29-29\\ 34-19 & 34-22 & 34-33 & 34-34\\ 40-21 & 40-25 & 40-35 & 40-40\end{bmatrix}$$

so the matrix will be:

$$x=\begin{bmatrix}11 & 9 & 1 & 0\\ 14 & 12 & 3 & 0\\ 16 & 13 & 2 & 0\\ 19 & 15 & 5 & 0\end{bmatrix}$$

Now we apply step (3):

$$x=\begin{bmatrix}11-11 & 9-9 & 1-1 & 0-0\\ 14-11 & 12-9 & 31-1 & 0-0\\ 16-11 & 13-9 & 21-1 & 0-0\\ 19-11 & 15-9 & 51-1 & 0-0\end{bmatrix}$$

For step (4), it is clear that two lines only are needed to cover all zeros in the matrix; one line covers the first row and the other line covers the fourth column:

$$x = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & 3 & 2 & 0 \\ 5 & 4 & 1 & 0 \\ 8 & 6 & 4 & 0 \end{bmatrix}$$

Now step (5) is applied and then a check for solution existence is carried out. We can cover all zeros by two lines only (first row and fourth column); hence we continue with step 5 by subtracting $x(3,3)$ from all the non-zero entries:

$$x = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ 4 & 3 & 0 & 0 \\ 7 & 5 & 3 & 0 \end{bmatrix}$$

The minimum number of lines that can cover the zeros is 3 (first row, third column, and fourth column), so we repeat step (5):

$$x = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 6 & 4 & 2 & 0 \end{bmatrix}$$

After checking for solution we find that the minimum number of lines covering all zeros is 3 (first row, third column, and fourth column), so we repeat step (5) again:

$$x = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 5 & 3 & 1 & 0 \end{bmatrix}$$

Now we can not cover all zeros by less than 4 lines. So a solution exists with two possibilities as follows (solution exists at the zeros locations):

$$x(1,1) + x(2,2) + x(3,3) + x(4,4) = 105$$

or

$$x(1,2) + x(2,1) + x(3,3) + x(4,4) = 105$$

The final result shows that the optimal assignment has two patterns and both of them yield a maximum profit equals 105.

As we notice from the previous example, the Hungarian algorithm guarantees resource allocation to all users in every scheduling decision. It allocates one subband to one user, which does not take into account the multiuser diversity. The Hungarian algorithm ensures optimal resource assignment under the condition of one-to-one mapping between users and subbands. However, this condition dramatically decreases the expected throughput of OFDMA systems.

Even though one-to-one mapping algorithms, such as the Hungarian method guarantee one subband to every requesting user, fair treatment among users in the OFDMA systems is not guaranteed. Pursuing short-term fairness in such systems is impractical due to different wireless channel conditions among users. It is clear that granting OFDMA resources in terms of time and frequency to users in bad channel conditions wastes the resources and decreases the system throughput without improving the fairness among users.

High computational complexity is another drawback of the Hungarian algorithm. The Hungarian algorithm is a very cumbersome algorithm to implement with complexity of $O(n^3)$ [2]. We notice from the previous example that the algorithm requires three iterations with heavy computational burden to solve a gain matrix of size 4x4 only. The simulation results we present in the simulation section prove that the Hungarian algorithm is not the best solution for OFDMA resources allocation problem because it does not fully utilize the OFDMA resources.

3.3 The *Max-Max* Scheduling Algorithm

In this section, we explain a heuristic algorithm to replace the Hungarian algorithm. As indicated in the previous section, the Hungarian algorithm is computationally expensive. Therefore, suboptimal scheduling algorithms such as *Max-Max* scheduling algorithm that deliver performance close to the Hungarian algorithm's performance with lower complexity are important [18].

The *Max-Max* scheduling algorithm establishes a gain matrix which contains the instantaneous data rate for all users on all subbands. The matrix dimension is *N×M,* with *N* represents the number of users, and *M* the number of subbands. A subband is allocated to the user with the highest achievable throughput rate; then the user's row and subband's column are deleted. This operation is repeated until all subbands are allocated to all users. Here is how the assignment problem for the matrix presented in the previous section is solved using the *Max-Max* scheduling algorithm.

By finding the maximum value in the matrix and eliminating the corresponding user and subband we obtain the following two assignment patterns. First, the algorithm chooses $x(4,4)$ and eliminate user 4 and subband 4 from the next iteration $x(3,3)$ is chosen and user 3 and subband 3 are eliminated. In the next iteration two scenarios can appear. The algorithm may choose $x(1,2)$ (or $x(2,2)$). Then, in the next iteration, the algorithm chooses $x(2,1)$ (or $x(1,1)$) depending on the previous iteration's decision.

$$x = \begin{bmatrix} \boxed{15} & 17 & 25 & 26 \\ 15 & \boxed{17} & 26 & 29 \\ 19 & 22 & \boxed{33} & 34 \\ 21 & 25 & 35 & \boxed{40} \end{bmatrix}$$

and

$$x = \begin{bmatrix} 15 & \boxed{17} & 25 & 26 \\ \boxed{15} & 17 & 26 & 29 \\ 19 & 22 & \boxed{33} & 34 \\ 21 & 25 & 35 & \boxed{40} \end{bmatrix}$$

respectively.

In the previous example the *Max-Max* scheduling algorithm provides the same performance of the Hungarian algorithm. However, this is not guaranteed all the time, especially with large matrices. Below we show another cost matrix where the Hungarian and the *Max-Max* scheduling algorithm solutions are different. The Hungarian algorithm guarantees the optimal assignment pattern under one-to-one assignment condition; hence its performance in terms of total achievable profit is higher than the *Max-Max* scheduling algorithm. The following gain matrix (*y)* is solved as follows:

$$y = \begin{bmatrix} 10 & 22 & 65 & 5 \\ 15 & 5 & 60 & 10 \\ 33 & 10 & 30 & 15 \\ 40 & 5 & 15 & 30 \end{bmatrix}$$

The Hungarian algorithm solution is:

$$y = \begin{bmatrix} 10 & \boxed{22} & 65 & 5 \\ 15 & 5 & \boxed{60} & 10 \\ \boxed{33} & 10 & 30 & 15 \\ 40 & 5 & 15 & \boxed{30} \end{bmatrix}$$

$$y\,(1,2) + y\,(2,3) + y\,(3,1) + y\,(4,4) = 145$$

Whereas the *Max-Max* scheduling algorithm solution is:

$$y = \begin{bmatrix} 10 & 22 & \boxed{65} & 5 \\ 15 & \boxed{5} & 60 & 10 \\ 33 & 10 & 30 & \boxed{15} \\ \boxed{40} & 5 & 15 & 30 \end{bmatrix}$$

$$y\,(1,3) + y\,(2,2) + y\,(3,4) + y\,(4,1) = 125$$

The Hungarian algorithm provides a total profit of 145, while the *Max-Max* scheduling algorithm provides a total profit of 125 only.

On one hand, heuristic algorithms such as the *Max-Max* scheduling algorithm reduce the computational complexity. On the other hand, such solutions inherit the Hungarian algorithm drawbacks. As in the Hungarian method case, heuristic algorithms, including the *Max-Max* scheduling algorithm, do not exploit the multiuser diversity and long-term fairness. Also such solutions apply one-to-one mapping between users and subbands, which does not improve the system fairness due to wireless channel state variation over time. Simulation results for the *Max-Max* scheduling algorithm are presented in the simulations section.

3.4 Proposed Frequency-time Scheduling Algorithms for OFDMA Systems

In this section, we propose two OFDMA scheduling algorithms that aim to schedule the OFDMA subbands and time frames in joint frequency-time domains using PF criterion in order to solve the short-term fairness problem and permit opportunistic behaviour of the scheduling solution while allocating resources to users in order to exploit the multiuser diversity. The two proposed scheduling algorithms schedule the resources to users in a many-to-many fashion where a frequency subband can be shared by more than one user. Meanwhile, a user can be assigned more than one subband during a time frame. We expect that subband sharing can significantly increase the subband utilization. We assume that it is possible to find at least one user who can utilize the excess time symbols of a subband to transmit or receive at least one packet. Allowing subband sharing should not have negative impact on the fairness among users because a subband is not shared unless the first selected user to access available resources does not have any traffic to send at least one packet. On the contrary, sharing subbands can improve the fairness because users can access the unused resources during a time frame instead of holding it by one user who does not have enough traffic to fully utilize the resources.

The first algorithm, which is referred to as the Horizontal scheduling algorithm, manages the radio resources subband by subband; in other words, it allocates all the time symbols in a subband and then allocates all the time symbols in the next subband, and so on. The second algorithm, which is referred to as the Vertical scheduling algorithm, manages subbands iteratively; the algorithm assigns sufficient number of time symbols from a subband to the selected users every scheduling decision.

The performance of the proposed Horizontal and Vertical scheduling algorithms is compared to that of previously discussed one-to-one algorithms, namely: the Hungarian, *Max-Max*, and RR scheduling algorithms. Simulation results show that the two proposed scheduling algorithms outperform the one-to-one scheduling algorithms in terms of throughput with nearly the same fairness behaviour.

Starting from the throughput maximization problem formulation presented in Chapter 2, we develop our solution in order to exploit multiuser diversity and allow subband sharing among competing users as follows [17]:

$$\max \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} \; s_{i,j}, \tag{3.1}$$

subject to:

$$C1: \sum_{i=1}^{N} s_{i,j} = 1, \quad \forall j \in \{1,2,3,\ldots,M\},$$

$$C2: \sum_{j=1}^{M} s_{i,j} = 1, \quad \forall i \in \{1,2,3,\ldots,N\},$$

$$C3: \; p_{i,j} = P_{total} / N, \quad \forall i, j,$$

$$C4: \sum_{i=1}^{N} \sum_{j=1}^{M} D_{i,j} s_{i,j} \geq Dq_i .$$

In order to overcome the throughput reduction and to guarantee fair treatment between users, we formulate the subband assignment problem as a many-to-many assignment problem. The first condition (*C*1) and the second condition (*C*2) are dropped. Hence, the subband scheduling problem is no longer maximal bipartite matching assignment. In order to guarantee fairness among users, we replace the conditions ($C1$ and C2) with a PF scheduling criterion, expressed as [15]

$$J = \arg \max_{1 \le i \le M} D_{i,j}(t) / R_{i,j}(t-1), \tag{3.2}$$

where $R_{i,j}(t)$ is updated for all users on each subband as [15]

$$R_i(n) = \begin{cases} (1-T_c^{-1})R_i(n-1), & i \ne k, \\ (1-T_c^{-1})R_i(n-1) + T_c^{-1}\sum_{j=1}^{M} D_{ij}(n), & i = k, \end{cases} \tag{3.3}$$

with $T_c \neq 0$. For the case of greedy behaviour, no observation window is being considered (T_c=0), so we consider $R_i(n)$ is equal to $D_i(n)$, $\forall$ n.

This PF criterion is used to develop the two proposed scheduling algorithms as explained in the next section.

3.5 Horizontal Scheduling Algorithm

The Horizontal scheduling algorithm allocates time symbols subband by subband within a time frame as show in Fig. 3.2. The flowchart of the Horizontal scheduling algorithm is shown in Fig. 3.3. First, the AMC levels are separately selected for users on each subband based on the SNR values. The AMC scheme is selected for each user on all subbands based on the SNR values in order to calculate the spectral efficiency. The instantaneous data rates, $D_{i,j}(t)$, are evaluated for each user $i, i = 1,2,3,...,N$, on each subband $j, j = 1,2,3,...,M,$. Based on the PF rule in equation (3.2), all users contend to access the resources (time symbols) subband by subband. The scheduler starts with the first subband, and then schedules the time symbols of the subband in the time domain to one or more users. After the subband is completed, the scheduler moves to the next subband, and so on. A subband is completed if the scheduler does not receive service requests during the frame time from any user, no time symbols remained, or the remaining time symbols in a subband cannot support at least one data packet. The wireless channel is dynamic and its conditions change over time. The achievable instantaneous data rate of each user should be determined at the beginning of every time frame. Thus, the instantaneous data rate, $D_{i,j}(t)$, is updated after finishing the current time frame.

On the other hand, the historical data rate, $R_{i,j}(t)$, is updated at the end of a time frame based on equation (3.3).

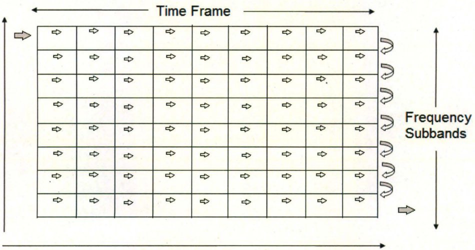

Fig. 3.2. The Horizontal scheduling algorithm allocation.

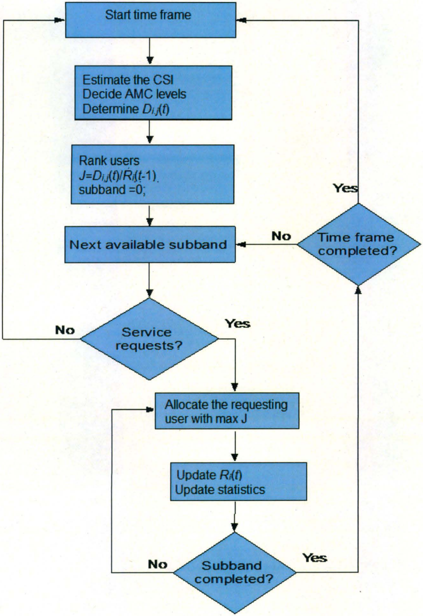

Fig. 3.3. The Horizontal scheduling algorithm flowchart.

3.6 Vertical Scheduling Algorithm

Fig. 3.4 demonstrates the allocation by the Vertical scheduling algorithm. The algorithm assigns resources subband by subband within each time symbol. First the algorithm allocates sufficient number of time symbols from each subband to the scheduled user, and then assigns time sufficient number of symbols to the next scheduled user. When the algorithm assigns one time symbol from all available subbands it moves in time and assigns time symbols to users in the next iteration.

As show in Fig. 3.5, the AMC levels are selected based on the SNR values and then the instantaneous data rate, $D_{i,j}(t)$, for all users is calculated. The vertical scheduling algorithm allocates the needed number of resource units to one user from different subbands at each scheduling decision. The scheduler starts with the first subband and continues based on an RR fashion, and then a user is chosen based on the PF criteria in equation (3.2). After scheduling one user resource demands, the scheduler proceeds to the next available resources in the subbands during the time symbol in the current iteration and chooses a user to be scheduled, and so on. If a user utilizes the entire subband within a time symbol, then he will access the subbands again in the next iteration (subbands in the next time symbol).

The instantaneous data rate, $D_{i,j}(t)$, is updated at the end of every time frame. The AMC level is adapted depending on the new channel conditions, so the instantaneous data rate, $D_{i,j}(t)$, is updated. The average data rate is updated for all users as in the Horizontal scheduling algorithm. The scheduler proceeds to the next time frame if remaining resources in all subbands cannot send or receive at least one packet, or no service requests arrived from any user. The vertical scheduling algorithm provides more frequency-time diversity than the horizontal scheduling algorithm. The Vertical scheduling algorithm switches between different subbands in frequency domain over time symbols every scheduling decision which causes more frequency diversity. On the other hand, the horizontal scheduling algorithm proceeds to different subbands in the frequency domain only when a subband is completed.

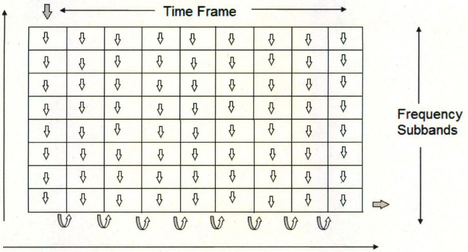

Fig. 3.4. The Vertical scheduling algorithm allocation.

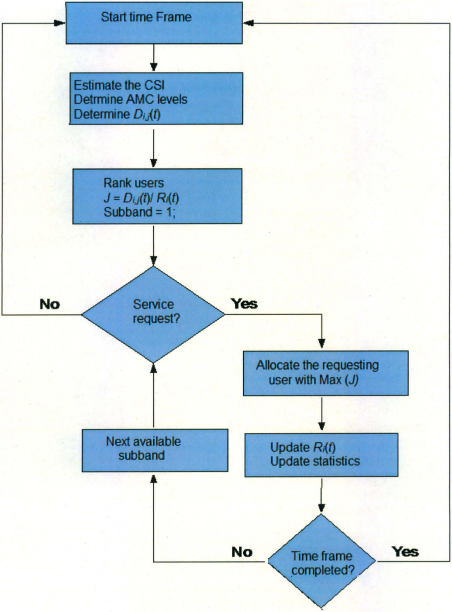

Fig. 3.5. The Vertical scheduling algorithm flowchart.

3.7 Simulation Results

The simulation results for the throughput and fairness are conducted with different values of the observation window T_c, where T_c equals 0, 1000, 3000, and 5000 time frames. T_c equals 0 means that the algorithm is greedy and no historical information is considered and the scheduling algorithm schedules users with a greedy behaviour. It is obvious that dropping the historical information will improve the throughput but decrease the fairness among users.

The throughput of the proposed Vertical and Horizontal scheduling algorithms is presented in Fig. 3.6. As we notice, the larger the observation window, T_c, the higher the throughput. Also, it is evident that the Vertical scheduling algorithm outperforms the Horizontal scheduling algorithm, due to its diversity when it schedules resources in the frequency domain, because the Vertical scheduling algorithm has more freedom to select the best subband for every user.

The Jain's fairness index is used to measure the fairness of the proposed scheduling algorithms. The fairness of the proposed scheduling algorithms is presented in terms of Jain's fairness index in Fig. 3.7 and Fig. 3.8. It is shown that both algorithms have nearly the same fairness with slight improvement of the vertical scheduling algorithm over the Horizontal scheduling algorithm. We notice that the smaller T_c the better the fairness because the PF scheduler impose more fairness among users by using less number of time frames to calculate the average data rate, $R_i(t)$, in equation (3.3). Also, it is evident that the higher the traffic load the lower the fairness because more packets will be blocked on the users' buffers.

The Hungarian, *Max-Max*, and RR scheduling algorithms are also simulated, and their performance compared with these of the proposed algorithms. The efficiency of the aforementioned algorithms in terms of throughput is presented in Fig. 3.9. The Hungarian solution outperforms the *Max-Max* scheduling algorithm in terms of throughput with the same fairness performance. The RR algorithm provides very low throughput, as this algorithm does not consider the channel state information when assigning subbands to users. Also, it is clear that the throughput of the Hungarian and the *Max-Max* scheduling algorithms is significantly lower than those of the Horizontal and Vertical scheduling algorithms. This throughput results show how the one-to-one mapping between users and subbands degrades the system throughput. It is more practical and efficient to allow opportunistic scheduling of users requests, and then compensate for users who starve for service when their channel state improves. It is obvious that scheduling users with low channel conditions does not improve the fairness, and severely impacts the performance in terms of throughput. Also the throughput reduction of one-to-one algorithms happens due to preventing subband sharing among users, which can waste considerable amount of resources.

The fairness of the three one-to-one algorithms is presented in Fig. 3.10. The Hungarian algorithm and the *Max-Max* scheduling algorithm show same the fairness behaviour, while the fairness provided by the RR is higher. Our two proposed scheduling algorithms show preferable fairness treatment to users compared to the Hungarian and the *Max-Max* scheduling algorithms due to the integration of the PF criterion, which improves the long-term fairness. Even though the RR algorithm pursues absolute fairness among users, we notice that the Jain's fairness index does not reach unity. This easily can be explained due to different wireless channel conditions from user to user which leads to different loading of the assigned subbands for each user.

Although the one-to-one algorithms guarantee a subband for each user within a time frame, the two proposed scheduling algorithms show slightly better fairness in terms of the Jain's fairness index. The proposed scheduling algorithms impose a short-term greedy behaviour, but the long-term fairness is maintained using the PF criterion. When a user i is prevented from accessing resources, his priority index, J, in equation (3.2) increases due to the weight added to his historical data rate, $R_i(t)$ in equation (3.3). Also, the proposed algorithms exploit time scheduling within each subband, which is not applicable in the one-to-one scheduling algorithms. It is evident that pursuing ultimate fairness among users in the OFDMA systems by using the one-to-one solution is not efficient and can degrade the throughput. Simulation results show that the many-to-many algorithms outperform the one-to-one algorithms in terms of throughput, without sacrificing the fair treatment among users.

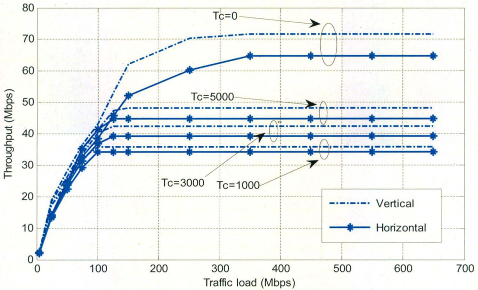

Fig. 3.6. Throughput versus traffic loads of the Vertical and Horizontal algorithms.

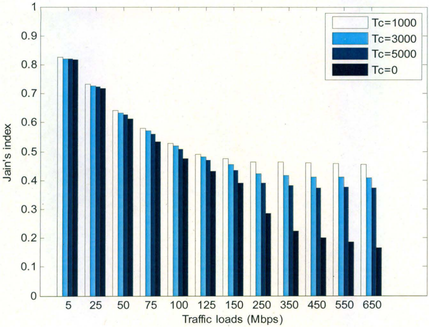

Fig. 3.7. Jain's fairness index of the Vertical algorithm.

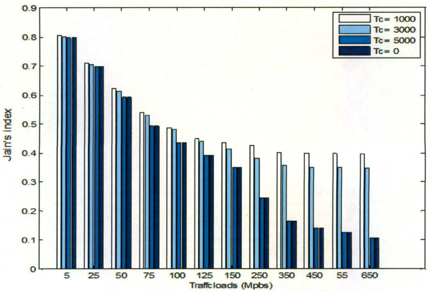

Fig. 3.8. Jain's fairness index of the Horizontal algorithm.

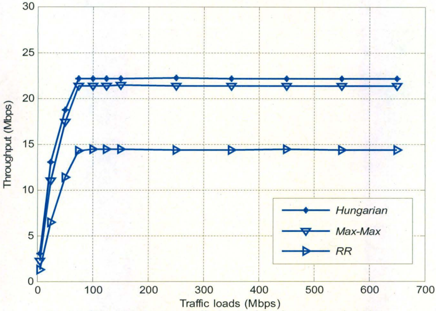

Fig. 3.9. Throughput versus traffic loads of the One-to-one algorithms.

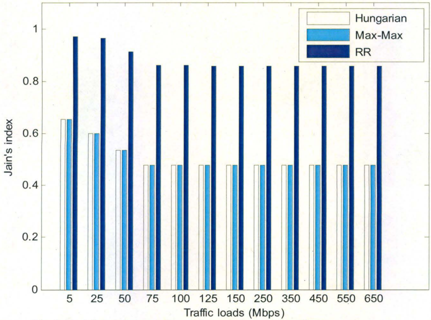

Fig. 3.10. Jain's fairness index of the one-to-one algorithms.

3.8 Conclusions

Scheduling is an important RRM component in the OFDMA systems. Dynamic subcarriers scheduling should consider both frequency and time domains in order to improve the performance. Imposing short-term fairness constrains does not improve long-term fairness behaviour of the OFDMA systems, but rather degrades the performance. In this chapter, we have proposed two joint frequency-time subband scheduling algorithms and investigated their performance in terms of throughput and fairness. Furthermore, their performance is compared with that of one-to-one scheduling algorithms.

Simulation results prove that with the one-to-one scheduling algorithms, which pursue direct fair treatment, the throughput performance is severely affected, without improving the long-term fairness. The two proposed Horizontal and Vertical scheduling algorithms outperform the one-to-one algorithms in terms of throughput and fairness among users. Also, the proposed scheduling algorithms are flexible and their behaviour can be controlled using only one parameter, namely the averaging window size, T_c. Our proposed scheduling algorithms can trade off between fairness and throughput using the averaging window, T_c.

CHAPTER 4

Performance Analysis of Proportional Fair Scheduling in OFDMA Wireless Systems

Introduction

In this chapter, a third OFDMA scheduling algorithm that utilizes the PF criterion and exploits the multiuser diversity in both time and frequency domains is proposed. Also the performance of the two dimensional (time slot and frequency subcarrier) PF scheduling algorithm for OFDMA wireless systems is analyzed analytically and by simulations.

The proposed modified *Max-Max* scheduling algorithm in this chapter aims to exploit the multiuser diversity in both time and frequency domains efficiently. Also, the proposed algorithm utilizes the PF criterion to achieve fairness among users in the OFDMA system. In order to support multimedia bursty traffic, the modified *Max-Max* scheduling algorithm allows more than one user to share a subband in each time frame. The proposed algorithm generates a gain matrix that is employed to iteratively assign the available subbands to be shared among different users concurrently. We modify the *Max-Max* scheduling algorithm in [18] in order to efficiently utilize resources in time and frequency domains taking fairness into account. The performance of our proposed algorithm is compared with other OFDMA scheduling algorithms in the literature. Results show that the proposed algorithm outperforms other algorithms in terms of throughput, fairness, and packet delay.

This chapter also analyzes the performance of the proposed modified *Max-Max* scheduling algorithm, both analytically and by simulation. Closed-form expressions for the average throughput and throughput fairness index are derived. Computer simulations are used for verification. The analytical and simulation results agree well with each other, which verifies the correctness and accuracy of the analytical expressions.

4.1 The Proposed Modified Max-Max Scheduling Algorithm

In this section, a modified *Max-Max* scheduling algorithm is proposed in order to exploit multiuser diversity, improve fairness, and meet the bursty traffic requirements for the OFDMA systems. The multiuser diversity is exploited by allowing the user that has the best channel conditions to access the system resources. Fairness between users is maintained by applying the PF criterion over time. Furthermore, the bursty traffic is dealt with by allowing subband sharing within the frame duration.

The proposed Horizontal scheduling algorithm handles the resources subband by subband. The scheduler moves from one frequency subband to another when all the time resources are utilized. The proposed Vertical scheduling algorithm shows more freedom in the frequency domain as it iterates over frequency subbands and allocates time resources to different users each single iteration. Results in Chapter 3 show that our proposed Vertical scheduling algorithm outperforms our proposed Horizontal scheduling algorithm due to more freedom in the frequency domain. This idea has motivated us to develop a scheduling algorithm with better frequency diversity than the Vertical scheduling algorithm.

The proposed modified *Max-Max* scheduling algorithm aims to provide more freedom in the frequency domain by trying to allocate time symbols of each subband to the user that maximizes the utilization on that subband. The proposed modified *Max-Max* scheduling algorithm does not iterate over frequency subbands sequentially. Instead of allowing users to compete to access a frequency subband each iteration, it allows users to compete over all available frequency subbands each iteration. The proposed *Max-Max* scheduling algorithm estimates the achievable throughput for each user on every frequency subband, and schedules users based on the PF criterion. Exploiting more freedom in the frequency domain is expected to improve the OFMDA system utilization.

Considering the efficiency and the bursty traffic requirements of the OFDMA systems, it is more suitable to allow subband sharing among users. More than one user can share the same subband. In order to exploit multiuser diversity, users with bad channel conditions are prevented from accessing subbands until their CSI is improved. However, queuing some users within a frame duration does not affect the long-term fairness because the queued users already experience bad channel conditions and cannot benefit from the available subbands anyways. With the PF criterion, the scheduler compensates for service starvation of a user when either the user is queued for a specific time or its channel conditions are improved in the next time frames.

In order to overcome the throughput reduction and guarantee fair treatment for all users, the conditions (*C1* and *C2*) in equation (3.1) in Chapter 3 are dropped. Hence, the subcarrier scheduling problem is no longer a maximal bipartite matching assignment, and a user can access more than one subband and a subband can be shared by more than one user. In order to guarantee fairness among users, we use the PF scheduling criterion described in equations (3.2) and (3.3) in Chapter 3. This PF criterion is integrated into the proposed modified *Max-Max* scheduling algorithm in order to guarantee fair treatment among users. The PF criterion compensates for the service shortage of users; also, it gives more priority to users which have better instantaneous data rates in order to maximize the throughput.

The *Max-Max* algorithm is modified as follows. First, a granted user will not be excluded in next scheduling iteration unless its queue is empty. Second, a subband will not be excluded in the next scheduling iteration unless it is fully utilized. Third, the PF criterion is implemented to evaluate ranking of users on the subbands based on the instantaneous and time average data rates.

The proposed algorithm is divided into two stages. In the first stage, the instantaneous data rates of user i and subband j at time t, $D_{i,j}(t)$, are evaluated for each user on each subband. A gain matrix that reflects the ranking of all users on all subbands is generated and updated based on the PF rule given in equations (3.2) and (3.3). In the second stage, the modified *Max-Max* scheduling algorithm is applied on the gain matrix as follows. In each iteration, the algorithm selects the user with the highest rank among all users on all subbands. After the user and subband are selected, the unused time symbols in the subband are reserved for that user depending on its traffic load and channel conditions. The scheduled user's traffic demands and the achieved throughput are updated. A user will be removed from the gain matrix if its queue is empty. A subband will be removed from the gain matrix if the remaining part cannot support at least one packet for any user. Finally, at the end of each iteration, the time average data rate, $R_i(t)$, for each user is updated based on (3.3). The algorithm terminates the iterations if all the subbands are consumed or all the users are satisfied. Moreover, at the beginning of every frame duration, the new instantaneous data rates, $D_{i,j}(t)$, are evaluated, and combined with the time average data rate, $R_i(t)$, to build a new gain matrix. The flow chart of the modified *Max-Max* scheduling algorithm is presented in Fig. 4.1.

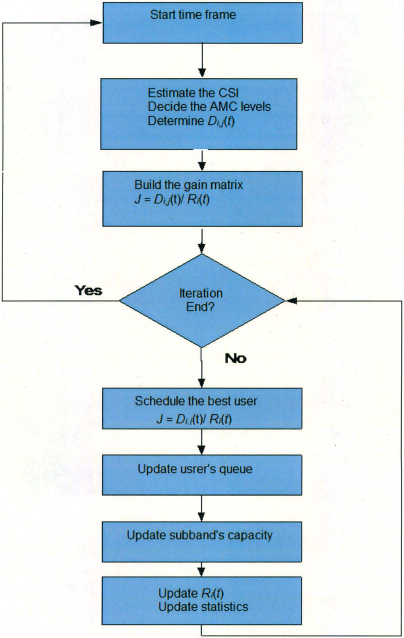

Fig. 4.1. The modified Max-Max scheduling algorithm flowchart.

4.2 Performance Analysis of Proportional Fair Scheduling in OFDMA Wireless Systems

The following sections of this chapter analyze the performance of the proposed modified *Max-Max* scheduling algorithm and evaluate its performance analytically. I derive closed-form expressions for the average throughput, average packet delay, and throughput fairness index.

An analytical method, which is based on the Gaussian approximation of the instantaneous data rate in a Rayleigh fading environment, is used to analyze the performance of PF scheduling in [25]. However, this method is developed only for single-carrier systems and limited to the case of users with full buffers. I adopt the methodology in [25] and build on it to develop an analytical solution for the PF scheduling in OFDMA systems. In the following three sections, closed-form expressions are derived for the average throughput, throughput fairness index, and mean packet delay of our proposed PF scheduling scheme for OFDMA systems.

According to the PF scheduling algorithm that I developed in the previous section for OFDMA systems, the user with the index

$$J = \arg \max_{1 \le i \le N} D_{ij}(n) / R_i(n-1), \tag{4.1}$$

is ranked first among the N users on subband j, $j=1,2,3,\ldots M$. The time average data rate is updated at the end of a time frame n-1 for each user i on all the available subbands as

$$R_i(n)=\begin{cases}(1-T_c^{-1})R_i(n-1), & i \neq k, \\ (1-T_c^{-1})R_i(n-1)+T_c^{-1} \sum\limits_{\substack{j=1 \\ j \in S_i(n)}}^{M} D_{ij}(n), & i = k,\end{cases} \tag{4.2}$$

where $S_i(n)$ represents the set of subbands assigned to user i during time frame n. T_c is the averaging window expressed in time frames. T_c controls how much historical information is taken into account when sharing the resources among multiple users, and can be chosen to achieve a desirable throughput-fairness trade-off.

Since the traffic is assumed to be bursty, the best user (chosen by (4.1)) might have empty buffer. In this case the subband assigned to the best user should be given to the second best user if it has non-empty buffer. If not, the subband is assigned to the third best users and so on, where the ranking of users is based on the same criterion used in (4.1) ($D_{ij}(n)/R_{ij}(n-1)$). As such, (4.2) is modified as follows

$$\begin{aligned} R_i(n) = (1-T_c^{-1})R_i(n-1) &+ \alpha_i T_c^{-1}\sum_{j=1}^{M} I_{ij}^1(n)D_{ij}(n) \\ &+ \alpha_i(1-\alpha_i)T_c^{-1}\sum_{j=1}^{M} I_{ij}^2(n)D_{ij}(n) \\ &+ \alpha_i(1-\alpha_i)^2 T_c^{-1}\sum_{j=1}^{M} I_{ij}^3(n)D_{ij}(n) \\ &+ \ldots + \alpha_i(1-\alpha_i)^{N-1}T_c^{-1}\sum_{j=1}^{M} I_{ij}^N(n)D_{ij}(n), \end{aligned} \tag{4.3}$$

where, $I_{ij}^c(n)$, c= 1,2,3,...N, represents a selector indicator which equals 1 if user i is ranked c-*th* on subband j, and equals 0 otherwise, and α_i is the probability that the buffer of user i is not empty. The terms in the right hand side of (4.3) represent the potential achievable throughput for a user. The first term reflects the average throughput achieved by the RR algorithm, while the remaining N terms represent the additional average throughput provided by the proposed scheduling algorithm when compared with RR. The first term (out of the remaining N terms) represents the additional average throughput when user i is ranked first and subbands assigned to it. The second term (out of the remaining N terms) reflects the additional average throughput when user i is ranked second and subbands are assigned to it because the user ranked first has empty buffer, and so on.

The proposed PF scheduling algorithm in the previous section consists of two steps: In the first step, all users in the system are ranked. A resource matrix that contains the ranking of all users on all subbands is generated based on equation (4.1). The instantaneous data rate, $D_{ij}(n)$, represents the efficiency factor, whereas the time average rate combined with T_c represents the fairness factor. As such, the ranking of the users reflects both the channel gain and users' satisfaction. In the second step, scheduling is performed based on the ranking and demands of the users on one hand, and the resource availability on the other hand. The algorithm iteratively serves the user with the highest rank among all users on all subbands.

A user will be excluded from the waiting users' list if all waiting packets are served. This algorithm allows subband sharing in time domain, where different time symbols in the subband can be utilized by different users. A subband will be eliminated from the resource matrix if the remaining resources cannot support at least one packet for any requesting user within this time-frame. This algorithm tracks the satisfaction levels of all users at the end of each time-frame by updating the time average data rate, $R_i(n)$, using equation (4.2).

4.2.1 Derivation of the Closed-Form Expressions for Throughput

It is shown that assuming a linear relationship between the instantaneous data rate, $D_{ij}(n)$, and the SNR is unrealistic under Rayleigh fading environment [25]. Actually, it is demonstrated that it is more realistic to assume that the D_{ij} follows a Gaussian distribution with mean and variance given respectively as [25]

$$E[D_{ij}] = \int_0^\infty \log(1 + SNR_{ij}\gamma)e^{-\gamma}d\gamma, \tag{4.4}$$

and

$$\sigma_{D_{ij}}^2 = \int_0^\infty \log(1+SNR_{ij}\gamma)^2 e^{-\gamma}d\gamma - (\int_0^\infty \log(1+SNR_{ij}\gamma)e^{-\gamma}d\gamma)^2, \tag{4.5}$$

where E[.] denotes the average operator and *log* is the logarithm in base 2. According to the PF algorithm presented in (4.1) and (4.2), the throughput of user *i* in subband *j* can be expressed as

$$R_i(n) = (1-\frac{1}{T_c})R_i(n-1) + \frac{1}{T_c}\sum_{j=1}^{M} I_{ij}^1(n)D_{ij}(n) + \frac{1}{T_c}\sum_{j=1}^{M} I_{ij}^2(n)D_{ij}(n)$$
$$+\frac{1}{T_c}\sum_{j=1}^{M} I_{ij}^3(n)D_{ij}(n) + \dots\dots\dots + \frac{1}{T_c}\sum_{j=1}^{M} I_{ij}^N(n)D_{ij}(n)$$

By calculating the mean value for both sides of the previous equation, we can write

$$E[R_i(n)] = (1-\frac{1}{T_c})E[R_i(n-1)] + \frac{1}{T_c}E[\sum_{j=1}^{M} I_{ij}^1(n)D_{ij}(n)] + \frac{1}{T_c}E[\sum_{j=1}^{M} I_{ij}^2(n)D_{ij}(n)]$$
$$+\frac{1}{T_c}E[\sum_{j=1}^{M} I_{ij}^3(n)D_{ij}(n)] + \dots\dots\dots + \frac{1}{T_c}E[\sum_{j=1}^{M} I_{ij}^N(n)D_{ij}(n)]$$

One can express the average achievable throughput of user *i* on all the available subbands in the time-frame *n* as

$$
\begin{aligned}
\mathrm{E}[R_i(n)] = {} & (1-T_c^{-1})\mathrm{E}[R_i(n-1)] + \alpha_i T_c^{-1}\mathrm{E}[\sum_{j=1}^{M} I_{ij}^{1}(n) D_{ij}(n)] \\
& + \alpha_i (1-\alpha_i) T_c^{-1}\mathrm{E}[\sum_{j=1}^{M} I_{ij}^{2}(n) D_{ij}(n)] \\
& + \alpha_i (1-\alpha_i)^2 T_c^{-1}\mathrm{E}[\sum_{j=1}^{M} I_{ij}^{3}(n) D_{ij}(n)] \\
& + \ldots + \alpha_i (1-\alpha_i)^{N-1} T_c^{-1}\mathrm{E}[\sum_{j=1}^{M} I_{ij}^{N}(n) D_{ij}(n)],
\end{aligned}
\tag{4.6}
$$

where I_{ij}^{N} represents the probability of user *i* ranked *N-th* to be scheduled on subband *j* is expressed as $\alpha_i (1-\alpha_i)^{N-1}$. This means that all users ranked first have empty queues, while user *i* has request in its buffer.

Under the assumption that α_i is equal for all users, as these have the same traffic rate D_{ij}, one can write

$$
E[\sum_{j=1}^{M} I^{x}{}_{ij}(n) D_{ij}(n)] = \alpha(1-\alpha)^{x-1} E[\sum_{j=1}^{M} D_{ij}(n) \,|\, I^{x}{}_{ij}(n)=1] \times P_r(I^{x}{}_{ij}(n)=1),
$$

and Now (4.6) can be re-expressed as

$$\begin{aligned}E[R_i(n)] &= (1-T_c^{-1})E[R_i(n-1)] \\ &+\alpha T_c^{-1}E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^1(n)=1]P_r(I_{ij}^1(n)=1) \\ &+\alpha(1-\alpha)T_c^{-1}E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^2(n)=1]P_r(I_{ij}^2(n)=1) \\ &+\alpha(1-\alpha)^2T_c^{-1}E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^3(n)=1]P_r(I_{ij}^3(n)=1) \\ &+\ldots+\alpha(1-\alpha)^{N-1}T_c^{-1}E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^N(n)=1]P_r(I_{ij}^N(n)=1),\end{aligned} \tag{4.7}$$

where $P_r(I_{ij}^c(n)=1)$ is the probability that user *i* is ranked *c*-th on subband *j* and during time frame *n*. Under the assumption of stationary throughput and independent subbands, one can further express (4.7) as

$$\begin{aligned}E[R_i] &= \alpha E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^1(n)=1]P_r(I_{ij}^1(n)=1) \\ &+\alpha(1-\alpha)E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^2(n)=1]P_r(I_{ij}^2(n)=1) \\ &+\alpha(1-\alpha)^2E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^3(n)=1]P_r(I_{ij}^3(n)=1) \\ &+\ldots+\alpha(1-\alpha)^{N-1}E[\sum_{j=1}^{M} D_{ij}(n)\,|\,I_{ij}^N(n)=1]P_r(I_{ij}^N(n)=1)\end{aligned} \tag{4.8}$$

Further, by applying the Bayes' theorem, (4.8) can be rewritten as

$$
\begin{aligned}
E[R_i] &= \alpha \sum_{j=1}^{M} \int_{-\infty}^{\infty} x f_{D_{ij}}(x) P_r(I_{ij}^1(n)=1 \mid D_{ij}(n)=x)\, dx \\
&+ \alpha(1-\alpha) \sum_{j=1}^{M} \int_{-\infty}^{\infty} x f_{D_{ij}}(x) P_r(I_{ij}^2(n)=1 \mid D_{ij}(n)=x)\, dx \\
&+ \alpha(1-\alpha)^2 \sum_{j=1}^{M} \int_{-\infty}^{\infty} x f_{D_{ij}}(x) P_r(I_{ij}^3(n)=1 \mid D_{ij}(n)=x)\, dx \\
&+ \ldots + \alpha(1-\alpha)^{N-1} \sum_{j=1}^{M} \int_{-\infty}^{\infty} x f_{D_{ij}}(x) P_r(I_{ij}^N(n)=1 \mid D_{ij}(n)=x)\, dx,
\end{aligned} \tag{4.9}
$$

where $f_{D_{ij}}(.)$ denotes the probability density function (pdf) of D_{ij}. By assuming independent D_{ij}, and based on the PF selection criterion presented in (4.1) and (4.2), I can determine the conditional ranking probabilities as follows

$$
\begin{aligned}
P_r(I_{ij}^1(n)=1 \mid D_{ij}(n)=x) &= \prod_{l=1, l\neq i}^{N} F_{D_{lj}}(x \frac{R_l(n)}{R_i(n)}), \\
P_r(I_{ij}^2(n)=1 \mid D_{ij}(n)=x) &= (1-F_{D_{l_1 j}}(x \frac{R_l(n)}{R_i(n)})) \prod_{\substack{l=1 \\ l\neq i, l_1}}^{N-1} F_{D_{lj}}(x \frac{R_l(n)}{R_i(n)}), \\
P_r(I_{ij}^3(n)=1 \mid D_{ij}(n)=x) &= (1-F_{D_{l_1 j}}(x \frac{R_{l_1}(n)}{R_i(n)}))(1-F_{D_{l_2 j}}(x \frac{R_{l_2}(n)}{R_i(n)})) \\
&\quad \times \prod_{\substack{l=1 \\ l\neq i, l_1, l_2}}^{N-2} F_{D_{lj}}(x \frac{R_l(n)}{R_i(n)}), \\
&\vdots \\
P_r(I_{ij}^N(n)=1 \mid D_{ij}(n)=x) &= \prod_{l=1, l\neq i}^{N} (1-F_{D_{lj}}(x \frac{R_l(n)}{R_i(n)})),
\end{aligned} \tag{4.10}
$$

where $F_{D_{ij}}(.)$ is the cumulative distribution function (cdf) of D_{ij}. By using (4.10) and the Gaussian pdf of D_{ij}, and under the assumptions that, $T_c \to \infty$, one can show that R_i is an ergodic process (such that its time average equals the statistical average). Now, by assuming $T_c \to \infty$ I have

$$\lim_{T_c\to\infty}\frac{R_l(n)}{R_i(n)}=\lim_{T_c\to\infty}\frac{R_l(n-1)}{R_i(n-1)}=\lim_{T_c\to\infty}\frac{R_l(n-2)}{R_i(n-2)}=.....=\lim_{T_c\to\infty}\frac{R_l(1)}{R_i(1)}=\lim_{T_c\to\infty}\frac{\sum_{t=1}^{n}R_l(t)}{\sum_{t=1}^{n}R_i(t)}$$

Hence,

$$\lim_{T_c\to\infty}\left[\lim_{n\to\infty}\frac{\sum_{t=1}^{n}R_l(t)}{\sum_{t=1}^{n}R_i(n)}\right]=\lim_{T_c\to\infty}\frac{Avrg(R_l)}{Avrg(R_i)}=\lim_{T_c\to\infty}\frac{R_l(n)}{R_i(n)},$$

where *Avrg*(.) is the time average. For stationary *R*, I assume that *R* is a first order ergodic, so that the time average equals the statistical average, as

$$\lim_{T_c\to\infty}\frac{\sum_{t=1}^{n}R_l(t)}{\sum_{t=1}^{n}R_i(t)}=\lim_{T_c\to\infty}\frac{E[R_l]}{E[R_i]}=\frac{E[R_l]}{E[R_i]}.$$

Now (4.10) can be re-written as

$$P_r(I_{ij}^1(n)=1\,|\,D_{ij}(n)=x)\approx\prod_{l=1,l\neq i}^{N}F_{D_{lj}(n)}\left(\frac{E[R_l]}{E[R_i]}x\right)$$

$$P_r(I_{ij}^2(n)=1\,|\,D_{ij}(n)=x)\approx(1-F_{D_{aj}(n)}(\frac{E[R_l]}{E[R_i]}x))\prod_{l=1,l\neq i}^{N-1}F_{D_{lj}(n)}(\frac{E[R_l]}{E[R_i]}x)$$

$$P_r(I_{ij}^3(n)=1\,|\,D_{ij}(n)=x)\approx(1-F_{D_{aj}(n)}(\frac{E[R_l]}{E[R_i]}x))^2\prod_{l=1,l\neq i}^{N-2}F_{D_{lj}(n)}(\frac{E[R_l]}{E[R_i]}x)$$

Now these expressions can be substituted into (4.9) so this equation can be expressed as

$$\begin{aligned}E[R_i]=&\,\alpha\sum_{j=1}^{M}\left[\int_{-\infty}^{\infty}x\times f_{D_{ij}(n)}(x)\times\prod_{l=\text{`},l\neq i}^{N}F_{D_{lj}(n)}\left(\frac{E[R_l]}{E[R_i]}x\right)dx\right]\\&+\alpha(1-\alpha)\sum_{j=1}^{M}\left[\int_{-\infty}^{\infty}x\times f_{D_{ij}(n)}(x)\left[1-F_{R_{aj}(n)}\left(\frac{E[R_l]}{E[R_i]}x\right)\right]\times\prod_{l=\text{`},l\neq i}^{N-1}F_{D_{lj}(n)}\left(\frac{E[R_l]}{E[R_i]}x\right)dx\right]\\&+\alpha(1-\alpha)^2\sum_{j=1}^{M}\left[\int_{-\infty}^{\infty}x\times f_{D_{ij}(n)}(x)\left[1-F_{R_{aj}(n)}\left(\frac{E[R_l]}{E[R_i]}x\right)\right]^2\times\prod_{l=\text{`},l\neq i}^{N-2}F_{D_{lj}(n)}\left(\frac{E[R_l]}{E[R_i]}x\right)dx\right]\end{aligned}$$

By assuming Gaussian distribution of the instantaneous traffic rate, the previous equation can be expressed as

$$
\begin{aligned}
E[R_i] = {} & \alpha\sum_{j=1}^{M}\left[\int_{-\infty}^{\infty}\left(y\sigma_{Dij}+E[D_{ij}]\right)\frac{e^{-y^2/2}}{\sqrt{2\Pi}}\times\prod_{l=`,l\neq i}^{N}F_{D_{lj}(n)}\left(\frac{E[R_l]}{E[R_i]}\times\left(y\sigma_{Dij}+E[D_{ij}]\right)\right)dy\right] \\
& +\alpha(1-\alpha)\sum_{j=1}^{M}\left[\int_{-\infty}^{\infty}\left(y\sigma_{Dij}+E[D_{ij}]\right)\frac{e^{-y^2/2}}{\sqrt{2\Pi}}\times\left[1-F_{R_{\alpha j}(n)}\left(\frac{E[R_l]}{E[R_i]}\times\left(y\sigma_{Dij}+E[D_{ij}]\right)\right)\right]\times\prod_{l=`,l\neq i}^{N-1}F_{D_{lj}(n)}\left(\frac{E[R_l]}{E[R_i]}\times\left(y\sigma_{Dij}+E[D_{ij}]\right)\right)dy\right] \\
& +\alpha(1-\alpha)^2\sum_{j=1}^{M}\left[\int_{-\infty}^{\infty}\left(y\sigma_{Dij}+E[D_{ij}]\right)\frac{e^{-y^2/2}}{\sqrt{2\Pi}}\times\left[1-F_{R_{\alpha j}(n)}\left(\frac{E[R_l]}{E[R_i]}\times\left(y\sigma_{Dij}+E[D_{ij}]\right)\right)\right]^2\times\prod_{l=`,l\neq i}^{N-2}F_{D_{lj}(n)}\left(\frac{E[R_l]}{E[R_i]}\times\left(y\sigma_{Dij}+E[D_{ij}]\right)\right)dy\right]
\end{aligned}
\tag{4.11}
$$

Now assume $\frac{E[R_l]}{E[R_i]}=\frac{E[D_l]}{E[D_i]}$, so $F_{R_{\alpha j}(n)}\left(\frac{E[R_l]}{E[R_i]}\times\left(y\sigma_{Dij}+E[D_{ij}]\right)\right)$ can be re-written as

$$
F_{R_{\alpha j}(n)}\left(\frac{E[R_l]}{E[R_i]}\times\left(y\sigma_{Dij}+E[D_{ij}]\right)\right)=F_{(0,1)}\left(\frac{E[D_{li}]\sigma_{D_{li}}}{E[R_{ij}]\sigma_{D_{ij}}}\times y\right)
$$

Furthermore, I assume proportional relationship between the mean and variance of all users in the system [25], so the previous expression can be approximated as

$$
F_{(0,1)}\left(\frac{E[D_{li}]\sigma_{D_{li}}}{E[R_{ij}]\sigma_{D_{ij}}}\times y\right)=F_{(0,1)}(y)
\tag{4.12}
$$

where $F_{(0,1)}(.)$ represents the standard normal cdf with zero-mean and unity-variance. Hence, after some manipulation, one can further express (4.11) as

$$E[R_i] = \alpha\sum_{j=1}^{M}[\sigma_{D_{ij}} \int_{-\infty}^{\infty} y \frac{e^{-y^2/2}}{\sqrt{2\pi}} F_{(0,1)}^{N-1}(y)dy + E[D_{ij}]\int_0^1 F_{(0,1)}^{N-1}(y)dF_{(0,1)}(y)]$$

$$+\alpha(1-\alpha)\sum_{j=1}^{M}[\sigma_{D_{ij}} \int_{-\infty}^{\infty} y \frac{e^{-y^2/2}}{\sqrt{2\pi}} (1-F_{(0,1)}(y))F_{(0,1)}^{N-2}(y)\, dy$$

$$+E[D_{ij}]\int_0^1 (1-F_{(0,1)}(y))F_{(0,1)}^{N-2}(y)dF_{(0,1)}(y)]$$

$$+\alpha(1-\alpha)^2\sum_{j=1}^{M}[\sigma_{D_{ij}} \int_{-\infty}^{\infty} y \frac{e^{-y^2/2}}{\sqrt{2\pi}} (1-F_{(0,1)}(y))^2 F_{(0,1)}^{N-3}(y)\, dy$$

$$+E[D_{ij}]\int_0^1 (1-F_{(0,1)}(y))^2 F_{(0,1)}^{N-3}(y)dF_{(0,1)}(y)]$$

$$+\ldots+\alpha(1-\alpha)^{N-1}\sum_{j=1}^{M}[\sigma_{D_{ij}} \int_{-\infty}^{\infty} y \frac{e^{-y^2/2}}{\sqrt{2\pi}} (1-F_{(0,1)}(y))^{N-1}\, dy$$

$$+E[D_{ij}]\int_0^1 (1-F_{(0,1)}(y))^{N-1} dF_{(0,1)}(y)], \qquad (4.13)$$

It is straightforward to show that

$$\int_0^1 F_{(0,1)}^{N-1}(y)dF_{(0,1)}(y) = \frac{1}{N}.$$

Then, one can easily find that

$$\int_0^1 (1-F_{(0,1)}(y))F_{(0,1)}^{N-2}(y)dF_{(0,1)}(y) = \frac{1}{N(N-1)},$$

and, finally, through the mathematical induction we can write

$$\int_0^1 (1-F_{(0,1)}(y))^{i-1} F_{(0,1)}^{N-i}(y)dF_{(0,1)}(y) = (i-1)!(N-i)!/N!, \qquad u=1,\ldots,N, \qquad (4.14)$$

Thus, equation (4.13) can be expressed as

$$E[R_i]=\sum_{i=1}^{N}[\alpha(1-\alpha)^{i-1}\sum_{j=1}^{M}[\sigma_{D_{ij}}\int_{-\infty}^{\infty} y\frac{e^{-y^2/2}}{\sqrt{2\pi}}(1-F_{(0,1)}(y))^{i-1}F_{(0,1)}^{N-i}(y)dy$$
$$+E[D_{ij}]\frac{(i-1)!(N-i)!}{N!}]], \tag{4.15}$$

where u represents the rank of user i on subband j. The probability of the non-empty buffer for any user, α, can be expressed based on the average throughput and the traffic rate as

$$\alpha=\frac{\lambda}{\mu}=\frac{\lambda}{E\left[R_i\right]}, \tag{4.16}$$

where μ and λ are the service and arrival rates of the user, respectively. By substituting (4.16) into (4.15), $E[R_i]$ simply becomes

$$E[R_i]=\sum_{i=1}^{N}[\frac{\lambda}{E\left[R_i\right]}(1-\frac{\lambda}{E\left[R_i\right]})^{i-1}\sum_{j=1}^{M}[\sigma_{D_{ij}}\int_{-\infty}^{\infty} y\frac{e^{-y^2/2}}{\sqrt{2\pi}}(1-F_{(0,1)}(y))^{i-1}F_{(0,1)}^{N-i}(y)\,dy$$
$$+E[D_{ij}]\frac{(i-1)!(N-i)!}{N!}]]. \tag{4.17}$$

As $E[R_i]$ represents the throughput of user i, $i= 1, 2, 3,\ldots, N$, in the system, the average throughput of the entire system can be expressed as

$$E[R]=\sum_{i=1}^{N}E[R_i]. \tag{4.18}$$

4.2.2 Derivation of the Closed-Form Expressions for the Fairness Index

System fairness while allocating resource to different users is an important requirement for successful scheduling solutions. Given a resource scheduling algorithm, a fairness index is a real number that measures how fair or unfair the system resources are shared among the competing users. Jain's fairness index is a well known quantitative metric that is widely used in wireless communications to measure fairness. Jain's fairness index can be expressed as [31]

$$J(x_1, x_2, x_3, \ldots, x_N) = \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N \sum_{i=1}^{N} x_i^2}. \tag{4.19}$$

where x_i is the amount of resources accessed by user i among N competing users.

Based on the result for the average throughput for user i, i= 1, 2, 3,..., N, as given in (4.18), it is straightforward to express the Jain's fairness index of the users' throughput as

$$J(E[R_1], E[R_2], E[R_3], \ldots, E[R_N]) = \frac{\left[\sum_{i=1}^{N} E[R_i]\right]^2}{\sum_{i=1}^{N} E[R_i]^2} \tag{4.20}$$

4.2.3 Derivation of the Closed-Form Expressions for Average Packet Delay

RRM for bursty traffic OFDMA systems is important because several services have bursty traffic. Accessing OFDMA resources by different users causes some users requests to be held in queues for certain time until a free resource is allocated. This time is referred to as the packet delay. Mean packet delay is an important metric for measuring the effectiveness of scheduling resources because services such as video and voice streaming should not exceed specific packet delay levels to meet its QoS requirements.

In this section, the average packet delay in M/G/1 queue for the proposed modified *Max-Max* scheduling algorithm is derived. I start the derivation from the proportional fairness criteria equation as

$$R_i(n) = \left[\left(\frac{T_c - 1}{T_c} \right) R_i(n-1) + \frac{1}{T_c} \sum_{i=1}^{N} \alpha (1-\alpha)^{i-1} \sum_{j=1}^{M} D_{ij}(n) I_{ij}^{i}(n) \right] \tag{4.21}$$

and $R_i^2(n)$ can be expressed as

$$R_i^2(n) = \left[\left(\frac{T_c - 1}{T_c} \right) R_i(n-1) + \frac{1}{T_c} \sum_{i=1}^{N} \alpha (1-\alpha)^{i-1} \sum_{j=1}^{M} D_{ij}(n) I_{ij}^{i}(n) \right]^2 \tag{4.22}$$

Now both sides of (4.22) are squared and averaged to express $E\left[R_i^2(n)\right]$ as

$$E\left[R_i^2(n)\right] = \left(\frac{T_c - 1}{T_c} \right)^2 E[R_i^2(n-1)] + \frac{1}{T_c^2} \sum_{i=1}^{N} \alpha^2 (1-\alpha)^{2(i-1)} E\left[\sum_{j=1}^{M} D_{ij}(n) I_{ij}^{i}(n) \right]^2 + \frac{2(T_c - 1)}{T_c^2}$$
$$E[R_i^2(n-1)] E\left[\sum_{i=1}^{N} \alpha (1-\alpha)^{i-1} \sum_{j=1}^{M} D_{ij}(n) I_{ij}^{i}(n) \right]$$

By assuming stationary R_i,

$$
\begin{aligned}
&(2T_c-1)E\left[R_i^2\right]=\sum_{i=1}^{N}\alpha^2(1-\alpha)^{2(i-1)}E\left[\sum_{j=1}^{M}D_{ij}(n)I_{ij}^{i}(n)\right]^2 \\
&2(T_c-1)E\left[\sum_{i=1}^{N}\alpha(1-\alpha)^{i-1}\sum_{j=1}^{M}D_{ij}(n)I_{ij}^{i}(n)\right]
\end{aligned}
\tag{4.23}
$$

In order to find $E\left[R_i^2\right]$, we can determine $E\left[\sum_{j=1}^{M}D_{ij}(n)I^{i}_{ij}(n)\right]^2$ as

$$
E\left[\sum_{j=1}^{M}D_{ij}(n)I^{i}_{ij}(n)\right]^2=E\left(D_{i1}^{(U_{ij})}+D_{i2}^{(U_{ij})}+D_{i3}^{(U_{ij})}+.....+D_{iM}^{(U_{ij})}\right)^2
$$

where U_{ij} is the index of user i on subband j. Now $E\left[\sum_{j=1}^{M}D_{ij}(n)I_{ij}^{i}(n)\right]^2$ can be is written as

$$
E\left[\sum_{j=1}^{M}D_{ij}(n)I_{ij}^{i}(n)\right]^2=\sum_{j=1}^{M}E\left(D_{ij}^2I_{ij}^{i}\right)+\sum_{j=1}^{M}\sum_{\substack{k=1,\\K\neq j}}^{M}E\left(D_{ij}D_{ik}I_{ij}^{i}I_{ik}^{i}\right)
$$

We further simplify $E\left[\sum_{j=1}^{M}D_{ij}(n)I_{ij}^{i}(n)\right]^2$ as

$$
\begin{aligned}
E\left[\sum_{j=1}^{M}D_{ij}(n)I_{ij}^{i}(n)\right]^2 &=\sum_{j=1}^{M}P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty}x^2f_{D_{ij}}\left(x\mid I_{ij}^{i}=1\right)\,dx \\
&+\sum_{j=1}^{M}\sum_{\substack{k=1,\\K\neq j}}^{M}P_r\left(I_{ij}^{i}=1\right)P_r\left(I_{ik}^{i}=1\right)\left(\int_{-\infty}^{\infty}xf_{D_{ij}}\left(x\mid I_{ij}^{i}=1\right)dx\right)\left(\int_{-\infty}^{\infty}xf_{D_{ik}}\left(x\mid I_{ik}^{i}=1\right)\,dx\right)
\end{aligned}
$$

Finally $E\left[\sum_{j=1}^{M}D_{ij}(n)I_{ij}^{i}(n)\right]^2$ can be written as

$$E\left[\sum_{j=1}^{M} D_{ij}(n) I_{ij}^{i}(n)\right]^2 = \sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right)dx$$
$$+\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\left(\int_{-\infty}^{\infty} x f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right)dx\right)\sum_{\substack{k=1,\\ k\neq j}}^{M} P_r\left(I_{ik}^{i}=1\right)\left(\int_{-\infty}^{\infty} x f_{D_{ik}}\left(x \mid I_{ik}^{i}=1\right)dx\right) \tag{4.24}$$

Now I simplify the first term of (4.24), $\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right)dx$

$$\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right)dx =$$
$$\sum_{i=1}^{N} \alpha^2(1-\alpha)^{2(i-1)}\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right)dx$$

Then it can be written as

$$\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right)dx =$$
$$\sum_{i=1}^{N} \alpha^2(1-\alpha)^{2(i-1)}\sum_{j=1}^{M}\int_{-\infty}^{\infty} x^2 f_{D_{ij}}(x) P_r\left(I_{ij}^{i}=1 \mid D_{ij}=x\right)dx \tag{4.25}$$

Now $P_r\left(I_{ij}^{i}=1 \mid D_{ij}=x\right)$ can be expressed as

$$P_r\left(I_{ij}^{i}=1 \mid D_{ij}=x\right)=\left(1-F_{\bar{ij}}\left(\frac{x}{R_i^2(n)}R_{\bar{l}}^2(n)\right)\right)^2 \times \prod_{l=1, l\neq i, \bar{l}} F_{D_{ij}}\left(\frac{x}{R_i^2(n)}R_l^2(n)\right)$$

For stationary first order ergodic R_i, the time average equals the statistical average, thus

$$P_r\left(I_{ij}^{U}=1 \mid D_{ij}=x\right)=\left(1-F_{\bar{ij}}\left(\frac{x}{E[R_i^2(n)]}E[R_{\bar{l}}^2(n)]\right)\right)^2 \times \prod_{l=1, l\neq i, \bar{l}} F_{D_{ij}}\left(\frac{x}{E[R_i^2(n)]}E[R_l^2(n)]\right)$$

(4.26)

By substituting (4.26) into (4.25), $\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^{2} f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right) dx$ can be written as

$$\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^{2} f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right) dx =$$

$$\sum_{i=1}^{N}\alpha^{2}(1-\alpha)^{2(i-1)}\sum_{j=1}^{M}\int_{-\infty}^{\infty} x^{2} f_{D_{ij}}(x)\left(1-F_{\bar{ij}}\left(\frac{x}{E[R_i^2(n)]}E[R_{\bar{l}}^2(n)]\right)\right)^{2}\times\prod_{l=1,l\neq i,\bar{l}} F_{D_{ij}}\left(\frac{x}{E[R_i^2(n)]}E[R_l^2(n)]\right)dx$$

It is reasonable to assume that [25]

$$\frac{E[R_l]^2}{E[R_i]^2}=\frac{E[D_{lj}]^2}{E[D_{ij}]^2},$$

Also proportional relationship between the mean and variance is assumed [25], where D_{ij} is a Gaussian with zero mean and unit variance, hence

$$F_{D_{ij}}(.)=F_{D_{ij}}\left(\frac{x}{E[R_i^2(n)]}\left(y\,\sigma_{D_{ij}}+E[D_{ij}]\right)\right)=F_{(0,1)}(y) \tag{4.27}$$

Substitute (4.27) into (4.26) and manipulate it as follows

$$\sum_{j=1}^{M} P_r\left(I_{ij}^{i}=1\right)\int_{-\infty}^{\infty} x^{2} f_{D_{ij}}\left(x \mid I_{ij}^{i}=1\right) dx =$$

$$\sum_{i=1}^{N}\alpha^{2}(1-\alpha)^{2(i-1)}\sum_{j=1}^{M}\int_{-\infty}^{\infty}\left(y\,\sigma_{D_{ij}}+E[D_{ij}]\right)^{2} f_{D_{ij}}(y)\times\left(1-F_{(0,1)}(y)\right)^{i-1}\times F_{(0,1)}^{N-i}(y)\,dx$$

Then this equation can be rewritten by solving $\left(y\,\sigma_{D_{ij}}+E[D_{ij}]\right)^{2}$

$$\sum_{j=1}^{M} P_r\left(I_{ij}^i = 1\right) \int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^i = 1\right) dx =$$

$$\sum_{i=1}^{N} \alpha^2 (1-\alpha)^{2(i-1)} \sum_{j=1}^{M} \sigma_{D_{ij}}^2 \int_{-\infty}^{\infty} y^2 f_{D_{ij}}(y) \times \left(1 - F_{(0,1)}(y)\right)^{i-1} F_{(0,1)}^{N-i}(y)\, dx$$

$$+2\sigma_{D_{ij}} E\left[D_{ij}\right] \int_{-\infty}^{\infty} y f_{D_{ij}}(y) \times \left(1 - F_{(0,1)}(y)\right)^{i-1} F_{(0,1)}^{N-i}(y)\, dx$$

$$+E\left[D_{ij}\right]^2 \int_{-\infty}^{1} y^2 f_{D_{ij}}(y) \times \left(1 - F_{(0,1)}(y)\right)^{i-1} F_{(0,1)}^{N-i}(y)\, dx$$

Finally $\sum_{j=1}^{M} P_r\left(I_{ij}^i = 1\right) \int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^i = 1\right) dx$ becomes

$$\sum_{j=1}^{M} P_r\left(I_{ij}^i = 1\right) \int_{-\infty}^{\infty} x^2 f_{D_{ij}}\left(x \mid I_{ij}^i = 1\right) dx =$$

$$\sum_{i=1}^{N} \alpha^2 (1-\alpha)^{2(i-1)} \sum_{j=1}^{M} \sigma_{D_{ij}}^2 \int_{-\infty}^{\infty} y^2 f_{D_{ij}}(y) \times \left(1 - F_{(0,1)}(y)\right)^{i-1} F_{(0,1)}^{N-i}(y)\, dx$$

$$+2\sigma_{D_{ij}} E\left[D_{ij}\right] \int_{-\infty}^{\infty} y f_{D_{ij}}(y) \times \left(1 - F_{(0,1)}(y)\right)^{i-1} F_{(0,1)}^{N-i}(y)\, dx \tag{4.28}$$

$$+\frac{(i-1) E\left[D_{ij}\right]}{N!/(N-i)!}$$

Now I solve the second part of equation (4.24) as follows

$$\sum_{j=1}^{M} P_r\left(I_{ij}^i = 1\right) \left(\int_{-\infty}^{\infty} x f_{D_{ij}}\left(x \mid I_{ij}^i = 1\right) dx \right) \sum_{\substack{k=1,\\ k \neq j}}^{M} P_r\left(I_{ik}^i = 1\right) \left(\int_{-\infty}^{\infty} x f_{D_{ik}}\left(x \mid I_{ik}^i = 1\right) dx \right) =$$

$$\sum_{i=1}^{N} \alpha^2 (1-\alpha)^{2(i-1)} \sum_{j=1}^{M} P_r\left(I_{ij}^i = 1\right) \left(\int_{-\infty}^{\infty} x f_{D_{ij}}\left(x \mid I_{ij}^i = 1\right) dx \right) \sum_{\substack{k=1,\\ k \neq j}}^{M} P_r\left(I_{ik}^i = 1\right) \left(\int_{-\infty}^{\infty} x f_{D_{ik}}\left(x \mid I_{ik}^i = 1\right) dx \right)$$

Then previous expression can be re-written as

$$\sum_{j=1}^{M} P_r\left(I_{ij}^i=1\right)\left(\int_{-\infty}^{\infty} x f_{D_{ij}}\left(x \mid I_{ij}^i=1\right) dx\right) \sum_{\substack{k=1,\\ k\neq j}}^{M} P_r\left(I_{ik}^i=1\right)\left(\int_{-\infty}^{\infty} x f_{D_{ik}}\left(x \mid I_{ik}^i=1\right) dx\right)=$$

$$\sum_{i=1}^{N}\alpha^2(1-\alpha)^{2(i-1)}\sum_{j=1}^{M}\left(\sigma_{D_{ij}}\int_{-\infty}^{\infty}\frac{ye^{-y^2/2}}{\sqrt{2\Pi}}\left(1-F_{(0,1)}(y)\right)^{i-1}F_{(0,1)}(y)^{N-i}dy+\frac{(i-1)E\left[D_{ij}\right]}{N!/(N-i)!}\right) \quad (4.29)$$

$$\sum_{\substack{k=1,\\ k\neq j}}^{M}\left(\sigma_{D_{ik}}\int_{-\infty}^{\infty}\frac{ye^{-y^2/2}}{\sqrt{2\Pi}}\left(1-F_{(0,1)}(y)\right)^{i-1}F_{(0,1)}(y)^{N-i}dy+\frac{(i-1)E\left[D_{ik}\right]}{N!/(N-i)!}\right)$$

Now I simplify the second part of equation (4.23),

$$2(T_c-1)E\left[\sum_{i=1}^{N}\alpha(1-\alpha)^{i-1}\sum_{j=1}^{M}D_{ij}(n)I_{ij}^i(n)\right]=2(T_c-1)E\left[\sum_{i=1}^{N}\alpha(1-\alpha)^{i-1}\sum_{j=1}^{M}D_{ij}(n)I_{ij}^i(n)\right]$$

$$2(T_c-1)E\left[\sum_{i=1}^{N}\alpha(1-\alpha)^{i-1}\sum_{j=1}^{M}D_{ij}(n)I_{ij}^i(n)\right]=2(T_c-1)E\left[R_i\right]\times E\left[R_i\right]=2(T_c-1)E\left[R_i\right]^2 \quad (4.30)$$

Now we substituting (4.30) and (4.31) in (4.23) yields

$$2(T_c-1)E\left[R_i\right]^2=\sum_{i=1}^{N}\alpha^2(1-\alpha)^{2(i-1)}\sum_{j=1}^{M}\sigma_{D_{ij}}{}^2\int_{-\infty}^{\infty}y^2 f_{D_{ij}}(y)\times\left(1-F_{(0,1)}(y)\right)^{i-1}F_{(0,1)}^{N-i}(y)dx$$

$$+2\sigma_{D_{ij}}E\left[D_{ij}\right]\int_{-\infty}^{\infty}y f_{D_{ij}}(y)\times\left(1-F_{(0,1)}(y)\right)^{i-1}F_{(0,1)}^{N-i}(y)dy+\frac{(i-1)E\left[D_{ij}\right]}{N!/(N-i)!}$$

$$\sum_{i=1}^{N}\alpha^2(1-\alpha)^{2(i-1)}\sum_{j=1}^{M}\left(\sigma_{D_{ij}}\int_{-\infty}^{\infty}\frac{ye^{-y^2/2}}{\sqrt{2\Pi}}\left(1-F_{(0,1)}(y)\right)^{i-1}F_{(0,1)}^{N-i}(y)dy+\frac{(i-1)E\left[D_{ij}\right]}{N!/(N-i)!}\right) \quad (4.31)$$

$$\sum_{\substack{k=1,\\ k\neq j}}^{M}\left(\sigma_{D_{ik}}\int_{-\infty}^{\infty}\frac{ye^{-y^2/2}}{\sqrt{2\Pi}}\left(1-F_{(0,1)}(y)\right)^{i-1}F_{(0,1)}^{N-i}(y)dy+\frac{(i-1)E\left[D_{ik}\right]}{N!/(N-i)!}\right)+2(T_c-1)E\left[R_i\right]^2$$

Thus it is possible to calculate the throughput variance as

$$\sigma_{R_i}^2=E\left[R_i^2\right]-E\left[R_i\right]^2 \quad (4.32)$$

and

$$E\left[E\left[R_i\right]\right]=\frac{1}{\omega_i} \tag{4.33}$$

where ω_i is the average time spent in the system per packet for user *i* and $E[R_i]$ is the throughput of user *i*. So now the packet delay in M/G/1 queue per user can be expressed as

$$\omega_i=\frac{1}{\mu_i}+\frac{\lambda_i\left(1/E^2\left[R_i\right]+\sigma_{R_i}^2\right)}{2\left(1-\frac{\lambda_i}{E\left[R_i\right]}\right)} \tag{4.33}$$

4.3 Numerical and Simulation Results

Various analytical and simulation results are presented in this section for the proposed modified *Max-Max* scheduling algorithm. Also simulation results are presented for the Orthogonal Frequency Division Multiplexing Proportional Fair (OFDMPF) scheduling algorithm proposed in [3].

The OFDMPF scheduling algorithm can be basically considered in the same category of my modified Max-*Max* scheduling algorithms. Both algorithms aim to schedule the OFDMA resources to multiple users by utilizing the PF criterion. However, my proposed modified Max-*Max* scheduling is designed in order to handle the OFDMA resources efficiently in both the frequency and time domains. The OFDMPF algorithm does not show degree of freedom in the frequency domain because all the orthogonal subbands in the frequency domain can be accessed by only one user at anytime. Hence, I expect that my proposed modified *Max-Max* scheduling algorithm will outperform the OFDMPF scheduling algorithm in terms of throughput, fairness, and packet delay. On the other hand, I compare the results of the proposed modified *Max-Max* scheduling algorithm analytical and simulation results.

The throughput of the proposed modified *Max-Max* and the OFDMPF scheduling algorithms is presented in Fig. 4.2. As expected, results in Fig. 4.2 show that the larger the observation window, T_c, the higher the throughput for both algorithms because the algorithms try to schedule users with better CSI. Both algorithms show dramatical throughput increase compared to the one-to-one algorithms presented in Chapter 3 as these two algorithms exploit multiuser diversity. However, the proposed modified *Max-Max* scheduling algorithm shows better performance compared to the OFDMPF algorithm as it efficiently utilizes the resources in the frequency domain, and can handle efficiently the bursty traffic because of the subband sharing.

The Jain's fairness index of the proposed modified *Max-Max* and OFDMPF scheduling algorithms is presented in Fig. 4.3 and Fig. 4.4, respectively. Both algorithms show approximately the same values of Jain's fairness index with slight improvement for the proposed *Max-Max* scheduling algorithm over the OFDMPF scheduling algorithm. Also, we notice that as the observation window, T_c, increases, the fairness decreases because the PF criterion gives more priority to users with higher data rates which causes other users to wait for longer time. Meanwhile, these two algorithms show slightly better fairness index compared to the Hungarian and *Max-Max* algorithms presented in Chapter 3 due to the PF criterion implementation. It is evident that imposing short-term fairness in OFDMA systems by applying the one-to-one scheduling algorithms reduces the system throughput dramatically, without improving the fairness.

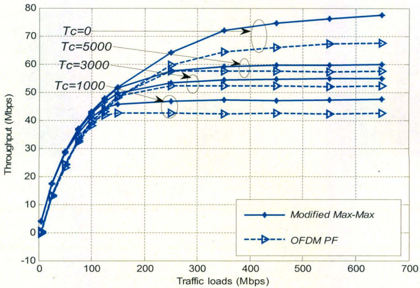

Fig. 4.2 Throughput of the proposed modified *Max-Max* and OFMDPF algorithms.

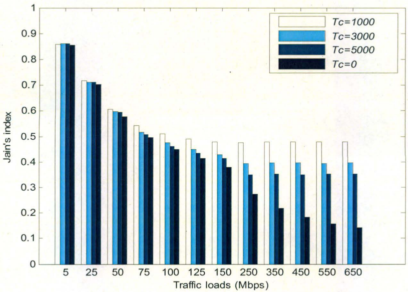

Fig. 4.3. Jain's fairness index of the proposed modified *Max-Max* algorithm.

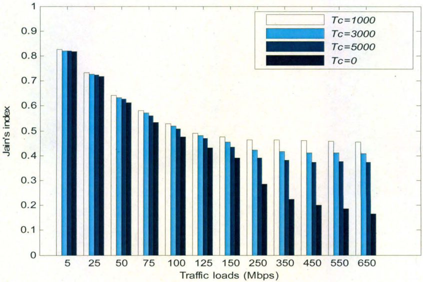

Fig. 4.4. Jain's fairness index of the OFDMPF algorithm.

Now results to validate the derived closed form expressions for the proposed PF modified *Max-Max* scheduling algorithm are presented.

Fig. 4.5 and Fig. 4.6 show the analytical and simulation results for throughput and Jain's fairness index respectively for the *Max-Max* and the throughput using simulations and the derived closed from expressions, where the observation window T_c equals 5000. The analytical results and the simulation results for the throughput in Fig. 4.5 agree very well. Both the analytical and simulation curves increase when the traffic loads increase until it reach the throughput saturation point of 60 Mbps around traffic load equals 300 Mbps.

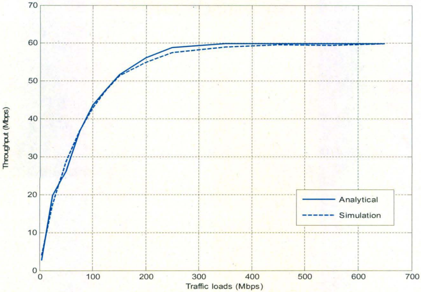

Fig. 4.5. Throughput vs. traffic loads for the proposed modified *Max-Max* algorithm analytically and by simulation.

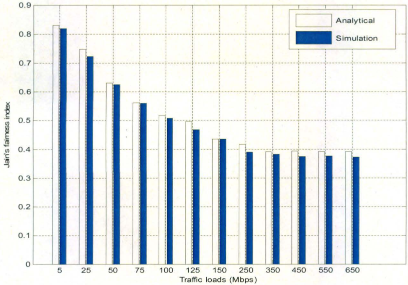

Fig. 4.6. Jain's fairness index vs. traffic loads for the proposed modified *Max-Max* algorithm analytically and by simulation.

Fig. 4.6 shows the Jain's fairness index for the proposed *Max-Max* scheduling algorithm using analysis and computer simulations where the observation window T_c equals 5000. As we can see the analytical results and the simulations agree very well. The fairness decreases as the traffic load increases until the system reaches the saturation point around traffic load equals 300 Mbps, and the Jain's fairness equals 0.4 approximately.

Fig. 4.7 and Fig. 4.8 show the throughput and Jain's fairness index respectively for the proposed modified *Max-Max* scheduling algorithm analytically and by simulations. Different number of users is considered with fixed traffic rate per user equals 10 Mbps, where the observation window T_c equals 5000. As we can see the analytical results and the simulations agree very well. As the number of users increase, the system traffic load increases so the throughput increase until it reaches the throughput reaches the saturation of 60 Mbps when the number of users equals 32 which means total traffic load equals 320 Mbps.

Fig. 4.8 shows the Jain's fairness index for different number of users with fixed traffic rate per user equals 10 Mbps for the modified *Max-Max* scheduling algorithm analytically and by simulations where the observation window T_c equals 5000. It is obvious that the analytical results and the simulations agree very well. The fairness decreases as the number of users increase where the lowest Jain's fairness index value equals 0.4 approximately when there are 32 users in the system with total traffic load equals 320 Mbps.

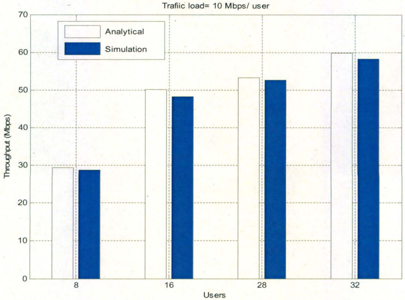

Fig. 4.7. Throughput vs. users for the proposed modified *Max-Max* algorithm analytically and by simulation.

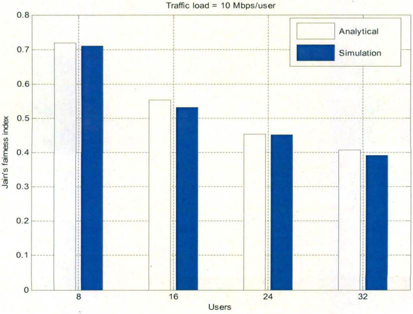

Fig. 4.8. Jain's fairness index vs. users for the proposed modified *Max-Max* algorithm analytically and by simulation.

Fig. 4.9 shows the system throughput versus different number of users for the proposed *Max-Max* scheduling algorithms analytically and by simulation where the observation window T_c equals 5000. Variable traffic rate per user is considered with fixed total traffic rate in the system equals 50 Mbps which is divided equally among the users in the system. The analytical results and the simulations agree very well. Slight throughput improvement is noticed when the number of users increases even though the traffic load is fixed. This trend can be easily explained based on the multiuser diversity. When the number of users increases the probability of scheduling users with higher channel gain will increase, which improves the overall system throughput. We notice the throughput reaches 27 Mbps when the number of users equals 32, while it reaches only 25 Mbps approximately when the number of users equals 8.

Fig. 4.10 shows the system Jain's fairness index versus different number of users for the proposed Max-Max scheduling algorithms analytically and by simulation where the observation window T_c equals 5000. Variable traffic rate per user is considered with fixed total traffic rate in the system equals 50 Mbps which is divided equally among the users in the. The analytical results and the simulation results agree very well. We notice that as the number of users increases, the fairness decreases due to the higher competition among users. When the number of users increases, the probability of more users will be blocked from accessing the system resources during a time frame will increase which negatively affects the system fairness in terms of Jain's fairness index. We notice form Fig. 4.10 the Jain's fairness index equals 0.68 when the number of users is 8 while it equals 0.62 when the number of users equal 32.

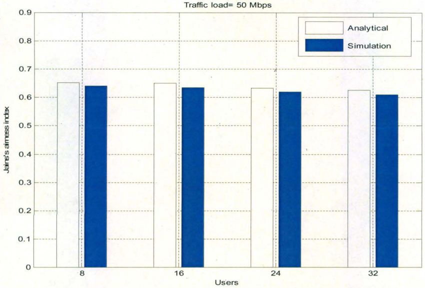

Fig. 4.9. Throughput vs. users for the proposed modified *Max-Max* algorithm analytically and by simulation.

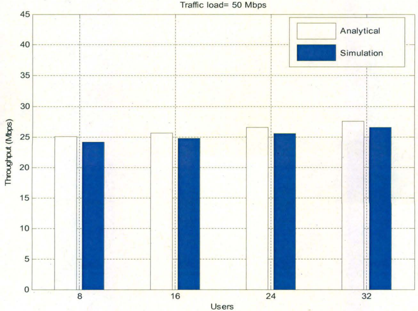

Fig. 4.10. Jain's fairness index vs. users for the proposed modified *Max-Max* algorithm analytically and by simulation. (Traffic load=50 Mbps)

Throughput and Jain's fairness index results are presented in the next six figures for different traffic loads in order to understand the behavior of the proposed *Max-Max* scheduling algorithm in OFDMA systems with different subbands number. The bandwidth is fixed and the number of subbands created using the available bandwidth is changed in terms of orthogonal subcarriers. The behavior of the OFDMA system is investigated under low (50 Mbps), moderate (250 Mbps), and high (450 Mbps) traffic loads.

Fig. 4.11, Fig. 4.12, and Fig. 4.13 show the throughput with different number of subbands for the proposed *Max-Max* scheduling algorithms analytically and by simulation under low traffic load (50 Mbps), medium traffic load (250 Mbps), and high traffic load (450 Mbps), respectively. The observation window, T_c, equals 5000. The available frequency bandwidth is divided into different number of subbands to study the behavior of the system with different numbers of subbands and subbands' size.

As we can see the analytical results and the simulations agree very well. We notice also the throughput reaches the maximum when the number of subbands equals 64 in the three figures. The throughput equals 42 Mbps, 58 Mbps, and 60 Mbps for low, medium, and high traffic rates, respectively. We notice also that the throughput decreases when the number of subbands is too small or too large. When the number of subbands equals 8, the throughput equals 14 Mbps, 50 Mbps, and 55 Mbps for low, medium, and high traffic rates, respectively. It is obvious that the number of subbands created from the available bandwidth affects the performance of the OFDMA systems in terms of throughput.

It is noticeable that moderate number of subbands guarantees best performance of OFDMA systems. The performance degradation of the OFDMA system when the number of subbands is too large and too small can be explained as follows. When the number of subbands is too small, the number of subcarriers per a subband is larger. So using the AMC for all the subcarriers based on the subcarriers with worst channel conditions will waste the resources of many subcarriers with favorable channel conditions. So it is better to group less number of subcarriers to avoid the underutilization of subcarriers with higher channel gain.

Regarding the second case where number of subbands is too large where few number of subcarriers are grouped to create a subband. The performance degradation is explained as follows. Performance degradation occurs because of the increasing of unused fractions of subbands at the end of time frames. There is high probability that a time frame finishes and there are subbands that are not fully utilized because there is no user that can benefit from the reminder of a subband to send or receive one packet. So when the number of subbands increases, the number of subbands that are not fully utilized at the end of time frames increase, so the performance decreases. We notice when the traffic rate is high (Fig. 4.13), the utilization of subbands increases, so the effect of number of subbands on the system throughput is low. The lower the traffic the more the subbands number affects the system throughput. It is obvious from Fig. 4.11 (low traffic rate) that throughput is lowest when the few number of subbands (8 subbands) are used or a large number of subbands (256 subbands) are used.

Fig. 4.14, Fig. 15, and Fig. 16 show the Jain's fairness index with different number of subbands for the proposed *Max-Max* scheduling algorithms analytically and by simulation under low traffic load (50 Mbps), medium traffic load (250 Mbps), and high traffic load (450 Mbps), respectively. The observation window, T_c, equals 5000. The available frequency subcarriers are divided into different number of subbands. Jain's fairness index simulation and analytical results agree very well. We notice the number of subbands does not affect the fairness of the system because all users suffer from the degradation of subbands utilization. In other words, the chance of accessing the resources will be affected equally for all users in the system.

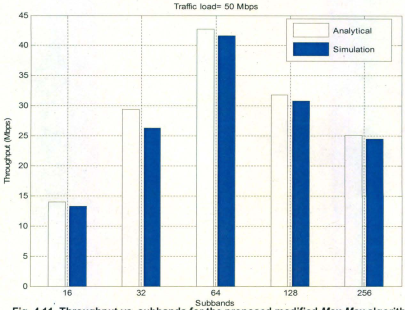

Fig. 4.11. Throughput vs. subbands for the proposed modified ***Max-Max*** algorithm analytically and by simulation. (Traffic load=50 Mbps)

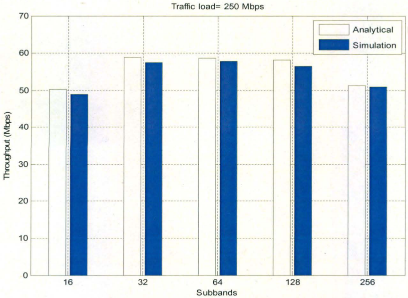

Fig. 4.12. Throughput vs. for the proposed modified ***Max-Max*** algorithm analytically and by simulation. (Traffic load=250 Mbps)

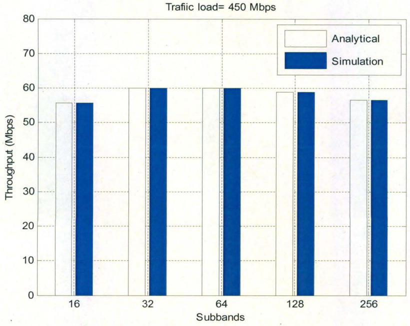

Fig. 4.13. Throughput vs. subbands for the proposed modified *Max-Max* algorithm analytically and by simulation. (Traffic load=450 Mbps)

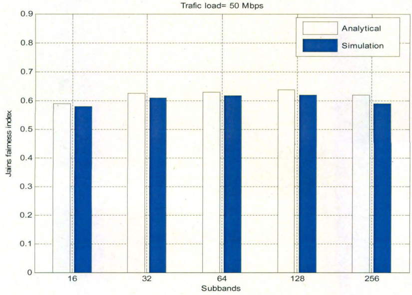

Fig. 4.14. Jain's fairness index vs. subbands for the proposed modified *Max-Max* algorithm analytically and by simulation. (Traffic load=50 Mbps)

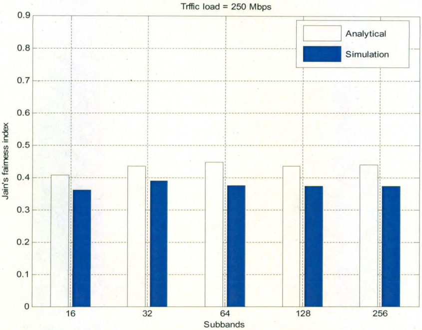

Fig. 4.15. Jain's fairness index vs. for the proposed modified *Max-Max* algorithm analytically and by simulation. (Traffic load=250) Mbps)

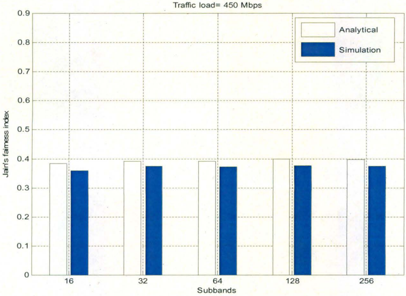

Fig. 4.16. Jain's fairness index vs. subbands for the proposed modified *Max-Max* algorithm analytically and by simulation. (Traffic load=450 Mbps)

Fig. 4.17 shows the packet delay versus traffic loads for the proposed *Max-Max* scheduling algorithm where the observation window T_c equals 5000, 3000, and 1000. It is evident that the traffic loads increase, the competition between users becomes harder, which causes more packets to wait longer time in the users queues. Also we notice that when T_c increases the packet delay increase. That can be explained as follows. When T_c increase the scheduler tries to maximize the system throughput by forcing greedy treatment among users by allocating less number of users who have favorable channel conditions. That behavior blocks more packets for requesting users, which increases the average packet delay in the system.

Fig. 4.18 shows the packet delay versus traffic loads for the proposed *Max-Max* scheduling algorithm analytically and by simulation, and the packet delay for the OFDMPF scheduling algorithms where the observation window T_c equals 5000. As we notice, the analytical curve agrees very well with the simulation curve. Also we notice slight improvement of the proposed *Max-Max* scheduling algorithm over the OFDMPF scheduling algorithm. We notice on high traffic load (650 Mbps) our proposed scheduling algorithm mean packet delay equals 3.75 seconds while the OFDMPF scheduling algorithm mean packet delay equals 3.45 seconds.

It is worth to notice that there is a small difference between the analytical and simulation results. This result difference can be explained based on the following factors:

1- Simulation time: it is expected that as the simulation increases, the analytical and simulation results converge and the gap between then decrease.

2- Analytical approximations: As per Section 4.2, some analytical approximations have been introduced while deriving the analytical model. Such approximations simplify the model in cost of minor results deviations.

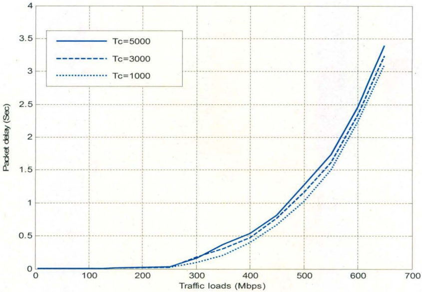

Fig. 4.17. Mean packet delay vs. traffic loads of the proposed modified *Max-Max* algorithm.

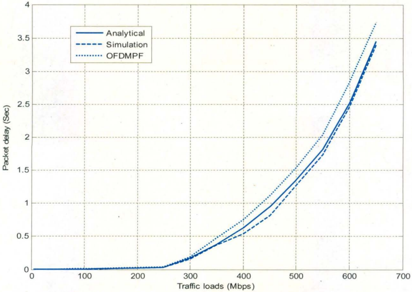

Fig. 4.18. Mean packet delay vs. traffic loads for the proposed modified *Max-Max* algorithm analytically and by simulation, and OFMDPF algorithm.

4.4 Complexity Comparison

Complexity of scheduling algorithms is an important factor that should be taken into account. We compare the complexity of our proposed modified *Max-Max* scheduling algorithm with the Hungarian scheduling algorithm and the *Max-Max* scheduling algorithm in this section.

A. Comparison with the Hungarian scheduling algorithm:

Our proposed modified *Max-Max* scheduling algorithm calculation burden can be expressed by two factors. The first factor is calculating the ranking of users on a time frame. This calculation is performed using the PF criterion. Equation 4.1 and equation 4.2 show that the calculation of users' ranking is fairly simple. The other factor that brings complexity is the fact that we introduce subband sharing among users. This requires multiple iteration of to allocate users to unutilized subband resources within a time frame. However, these iterations do don't require heavy calculations as the ranking of users is already evaluated, so the algorithm only needs to allocate the unused resources to the next best user ranked on a subband. On the other hand, the Hungarian scheduling algorithm requires iterative matrix manipulations and calculations in order to rank users on subbands (refer to Section 3.2 for calculations steps and example).

B. Comparison with the *Max-Max* scheduling algorithm:

Our proposed modified *Max-Max* scheduling algorithm involves subbands sharing where iterations are performed to fully utilize subbands within a time frame. As we stated before, these iterations are not computationally expensive as it only requires selecting the next pre-ranked user on non-fully utilized subband. It is worth to notice that this minor computational complexity of our algorithm is cost that paid to dramatically increase the system performance in terms of throughput. Fig 3.9 and Fig 4.2 show how much throughput gain our proposed algorithm can achieve over the *Max-Max* scheduling algorithm. We can notice that at the saturation state our proposed algorithm can achieve throughout up to 48Mbps when T_c=1000 and 60Mbps when T_c=5000 while the *Max-Max* scheduling algorithm can achieve throughput equals 22Mbps only.

4.5 Conclusions

In this chapter, I propose the modified *Max-Max* scheduling algorithm that allows multiple users to share a subband concurrently, in order to exploit multiuser diversity and guarantee effective scheduling of the multimedia traffic from multiple users. Simulation results prove the efficiency of this scheduling algorithm in wireless systems that involve multiple users with multimedia traffic requirements. The one-to-one solution, represented by the Hungarian, *Max-Max*, and RR scheduling algorithms, does not exploit multiuser diversity. Also, this cannot support the bursty traffic requirements from multiple users, as it pursues a direct fair treatment, which severely affects the throughput, without improving the long-term fairness among users.

In this chapter also, the PF scheduling is investigated for OFDMA wireless systems. The main contribution of this work is the analytical evaluation of the performance of PF scheduling algorithm in OFDMA systems. I derive closed-form expressions for the average throughput, Jain's fairness index, and average packet delay as the performance metrics. The PF under OFDMA systems performance is investigated for a broad range of the traffic load and the number of subbands. In addition, I verify the correctness and accuracy of the analytical solution through simulations. Analytical and simulation results are in good agreement, which validates my analytical performance analysis.

CHAPTER 5

Conclusions and Future Work

This thesis investigates the performance of scheduling solutions for OFDMA systems. OFDMA systems, such as IEEE802.16 (WiMAX), are multiuser systems that aim to support users with high data rates. As the OFDMA systems are built on OFDM, they inherit the favourable characteristics of OFDM, such as its immunity to intersymbol interference and frequency selective fading. OFDMA scheduling algorithms are expected to schedule OFDMA resources dynamically, in both time and frequency domains, among multiple users in order to improve OFDMA systems performance. The OFDMA scheduling algorithms that we proposed are able to manage radio resources for systems with multiple users and high data rate expectations, such as the WiMAX systems.

The radio channel is dynamic and is characterized by multipath fading due to scattering and reflection from terrain objects; every orthogonal OFDM subcarrier is subject to flat fading. This thesis models the dynamic radio channel by considering Rayleigh fading, shadowing, and path loss. The Rayleigh fading samples generated in this thesis are correlated in the frequency domain. The correlation between fading samples in different subcarriers is derived as a function of the frequency separation. By considering the correlation among subcarriers, grouping of adjacent subcarriers into subbands is reasonable, as grouped subcarriers experience similar channel conditions. The AMC level for a subband's subcarriers is selected based on the subcarrier that has the worst CSI in order to maintain transmission reliability. Selecting the AMC scheme based on the worst CSI is reasonable, and does not reduce the subband utilization because the subcarriers within a subband are highly correlated in the frequency domain.

The thesis investigates the design and performance of scheduling algorithms for OFDMA systems. Multiuser OFDMA systems contain resources in both time and frequency domains. Hence, exploiting diversity in both time domain and frequency domain is an important aspect for designing efficient OFDMA scheduling algorithms. Our results show that exploiting diversity in both time and frequency domain dramatically improves the frequency subband and time frame utilization. Simulation results for the Horizontal and Vertical scheduling algorithms show that the more diverse the scheduling solution in the frequency domain, the better the system performance. Also, the results for the proposed modified *Max-Max* scheduling algorithm and the OFDMPF scheduling algorithm show that frequency diversity improves the OFDMA system performance.

Scheduling radio resources in OFDMA systems with multiple users should consider another degree of freedom in terms of frequency and user diversities by allowing dynamic assignment of frequency and time to different users. Multiuser diversity is a very critical aspect in OFDMA scheduling algorithms. The thesis shows that scheduling schemes with short-term fairness constraints severely degrades the OFDMA system performance without improving the long-term fairness among users. Results prove that allowing opportunistic scheduling dramatically improves the throughput without sacrificing the long-term fairness among users.

The thesis introduces two OFDMA scheduling algorithms in the third chapter, namely, the Vertical and Horizontal scheduling algorithms. These are designated and implemented with the goal of exploiting time and frequency diversity while scheduling resources to users. Also, multiuser diversity and long-term fairness are considered. The two scheduling algorithms introduce subband sharing between users as an efficient technique for maximizing the utilization resources. Subband sharing minimizes the probability of resource wasting. The Horizontal scheduling algorithm schedules a subband's time symbols to users until all the subband symbols are occupied, then the scheduler moves to the next subband. On the other hand, the Vertical scheduling algorithm shows more flexibility in the frequency domain. It manages subbands iteratively, where the algorithm assigns sufficient resources from a subband to the selected users every scheduling decision. Both scheduling algorithms maintain long-term fairness among users. However, the Vertical scheduling algorithm is more efficient in terms of throughput and fairness due to better frequency diversity. The results for the RR, *Max-Max*, and the Hungarian algorithms prove that short term-fairness dramatically decreases the OFDMA system throughput when compared with the proposed Horizontal and Vertical scheduling algorithms.

The thesis proposes a third OFDMA scheduling algorithm in the fourth chapter. The modified *Max-Max* scheduling algorithm exploits the multiuser diversity in both time and frequency domains. Also, the proposed algorithm utilizes the PF criterion to achieve fairness among users in the system by generating profit matrix of all users on the available OFDMA subbands. The algorithm iteratively selects the user with the highest rank in the profit matrix among all users on all subbands. The modified *Max-Max* scheduling allows subband sharing when a subband is not fully utilized by another scheduled user. The simulation results of the proposed *Max-Max* scheduling algorithm and the OFDMPF scheduling algorithm prove that frequency diversity and subband sharing can improve the OFDMA systems performance in terms of throughput, fairness, and packet delay.

The thesis finally investigates analytically the performance of the PF scheduling in OFDMA wireless systems. We derive closed-form expressions for throughput, fairness, and packet delay as performance metrics. The Gaussian approximation of the instantaneous data is used in the derivation. We verify the correctness and accuracy of the analytical solution through simulations. Analytical and simulation results are in good agreement, which validates our analytical performance analysis.

We study the performance of scheduling algorithms for OFDMA systems in terms of throughput, Jain's fairness index, and packet delay under different traffic loads, number of users, and number of subbands. All simulation and analytical results agreed very well. One noticeable result is the behaviour of the scheduling algorithms for the OFDMA system under different numbers of subbands. We conclude that the number of subbands in OFDMA systems is a critical factor that affects the performance. It is evident that dividing the available subcarriers into subbands with average size can guarantee the best performance. This finding can be explained as follows. On one hand, grouping a large number of subcarriers to create a small number of subbands degrades the performance due to low correlation among subcarriers within a subband. On the other hand, dividing the available frequency resources in terms of subcarriers into large number of subbands increases the unused portions of subbands at the end of time frames due to insufficient resources to transmit a packet for any user.

As future work, we suggest the following topic:

- Extend the analytical model to consider multi-cell OFDMA systems where interference from other cells' users is taken into account
- Power control should be considered and associated with the scheduling algorithms. Note that in the thesis power is equally distributed among users.
- Our proposed scheduling algorithms are iterative algorithms. The iterative algorithms have lower complexity compared to the Hungarian algorithm. Many authors proposed iterative solutions in literature to reduce complexity of solutions. However, I did not have the chance to focus on studying the complexity mathematically during my program. This could be an interesting topic to focus on in future work.
-

REFERENCES

[1] "IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems," *IEEE, Tech. rep. 802.16*, Oct. 2004.

[2] S. Ali, K. D. Lee, and V. C. Leung, "Dynamic resource allocation in OFDMA wireless metropolitan area networks," *IEEE Wireless Communications Magazine*, vol. 14, pp. 6-13, Feb. 2007.

[3] H. J. Zhu and R. H. Hafez, "Novel scheduling algorithms for multimedia service in OFDM broadband wireless systems," *in Proc. International Conference on Communications*, pp. 772-777, July 2006.

[4] C. Huang, H. Juan, M. Lin, and C. Chang, "Radio resource management of heterogeneous services in mobile WiMAX systems," *IEEE Wireless Communications Magazine*, vol.14, pp. 20-26, Feb. 2007.

[5] C. Huang, H. Juan, M. Lin, and C. Chang, "Hierarchical Downlink Resource Management Framework for OFDMA based WiMAX Systems," *IEEE Transactions on Mobile Computing*, vol.6, pp. 621-632, June 2007.

[6] A. F. Molisch, Wireless Communications. Wiley, 2005.

[7] D. Niyato and E. Hossain, "A radio resource management framework for IEEE 802.16-based OFDM/TDD wireless mesh networks," *in Proc. International Conference*, pp. 3911-3916, June 2006.

[8] G. Song and Y. Li, "Cross layer optimization for OFDM wireless networks-Part I: Theoretic Framework," *IEEE Transaction on Wireless Communication*, vol. 4, pp. 614-624, Mar. 2005.

[9] Z. Shen, J. Andrews, and B. Evans," Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 2726-2737, Nov. 2005.

[10] A. Ghosh and D. R. Wolter, "Broadband Wireless Access with WiMAX/802.16: Current Performance Benchmarks and Future Potential", *IEEE Communication*, pp. 129-136, Feb. 2005.

[11] Z. Sun, Y. Zhou, M. Peng, and W. Wang, "Dynamic Resource Allocation with Guaranteed Diverse QoS for WiMAX System", *in Proc. International Conference*, vol.2, pp. 1347-1351, June 2006.

[12] B. Rong and Y. Qian," Integrated Downlink Resource Management for Multiservice WiMAX Networks", *IEEE Computer Society*, vol. 6, pp. 621-632, June 2007.

[13] Y. Ma," Proportional Fair Scheduling for Downlink OFDMA," *in Proc. International Communications conference*, pp. 4843-4848, June 2007.

[14] G. Huang, H. Juan, M. S. Lin, and C. Chang, "Radio resource management of heterogeneous services in mobile WiMAX systems," *IEEE Wireless Communications*, vol. 14, pp. 20-26, Feb. 2007.

[15] W. Anchun, X. Liang, Z. S. Xibin, and Y. Yan, "Dynamic resource management in the fourth generation wireless systems," *in Proc. International Conference on Communication Technology*, pp. 1095-1098, Apr. 2003.

[16] L. Badia, A. Baiocchi, A. Todini, S. Merlin, S. Pupolin, A. Zanella, and M. Zorzi, " On the impact of physical layer awareness on scheduling and resource allocation in broadband multicellular IEEE 802.16 systems," *IEEE Wireless Communications Magazine*, vol.14, pp. 36-43, Feb. 2007.

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

[18] Z. Zhang, Y. He, and E. Chong, "Opportunistic downlink scheduling for multiuser OFDM systems," *in Proc. Wireless Communications and Networking Conference*, pp. 1206-1212, Mar. 2005.

[19] H. Kim, K. Kim, Y. Han, and J. Lee, "An efficient scheduling algorithm for QoS in wireless packet data transmission," *in Proc. Personal Indoor and Mobile Radio Communications Conference*, vol. 5, pp. 2244–2248, Sep. 2002.

[20] K. D. Lee and V. C. Leung, "Fair allocation of subcarrier and power in an OFDMA wireless mesh network," *IEEE Journal of Selected Areas in Communications*, vol. 24, pp. 2051-2060, Nov. 2006.

[21] G. Song and Y. Lee, "Cross-layer optimization for OFDM wireless networks—Part II: Algorithm development," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 625-634, Mar. 2005.

[22] G. Munz, S. Pfletschinger, and J. Speidel, "An efficient waterfilling algorithm for multiple access OFDM," *IEEE Global Telecommunications Conference*, vol.1, pp. 681-685, 2002.

[23] H. Seo and B. Lee, "Proportional-fair power allocation with CDF-based scheduling for fair and efficient multiuser OFDM systems," *IEEE Wireless Communications Magazine*, vol. 5, pp. 978-983, May 2006.

[24] Ryu, B. Ryu, H. Seo, and M. Shin, "Urgency and efficiency based packet scheduling algorithm for OFDMA wireless system," *IEEE International Conference*, vol. 4, pp. 2779-2785, May 2005.

[25] E. Liu and K. K. Leung, "Proportional fair scheduling: Analytical insight under Rayleigh fading environment," *in Proc. IEEE Wireless Communications and Networking Conference*, pp. 1883 – 1888, Mar. 2008.

[26] Y. Cao, Z. L. lu, and Y. Yang, “A centralized scheduling algorithm based on multi-path routing in WiMAX mesh networks”, *Wireless Communications, Networking and Mobile Computing International Conference*, pp. 1-4, Sept. 2006.

[27] C. Y. Wong and K. B. Letaief, “Multiuser OFDM with adaptive subcarrier, bit, and power allocation”, *IEEE Selected Areas in Communications*, vol. 17, pp. 1747-1765, Oct. 1999.

[28] B. Sklar, “Rayleigh fading channels in mobile digital communication systems Part 1: Characterization,” *IEEE Communications Magazine*, pp. 90-100, July 1997.

[29] B. Sklar, “Rayleigh fading channels in mobile digital communication systems Part 2: Mitigation,” *IEEE Communications Magazine*, pp. 148-155, 1997.

[30] L. C. Tran, T. A. Wysocki, and A. Mertins, “A generalized algorithm for the generation of correlated Rayleigh fading envelopes in wireless channels,” *EURASIP Journal on Wireless Communications and Networking*, pp. 801–815, May 2005.

[31] R. Jain, D. Chiu, and W. Hawe, “A quantitative measure of fairness and discrimination for resource allocation in shared computer systems,” *Digital Equip. Corp., Littleton, MA, DEC Rep.*, DEC-TR-301, Sep. 1984.