國 立 交 通 大 學

電信工程學系

碩 士 論 文

IEEE 802.16 上行鏈路無線系統之
動態優先次序資源分配

Dynamic Priority-based Resource Allocation
for IEEE 802.16 Uplink Wireless Systems

研究生：賴建安

指導教授：張仲儒 博士

中 華 民 國 九 十 六 年 六 月

IEEE 802.16 上行鏈路無線系統之動態優先次序資源分配

Dynamic Priority-based Resource Allocation for IEEE 802.16 Uplink Wireless Systems

研　究　生：賴建安　　　　Student：Jian-Ann Lai

指導教授：張仲儒　博士　　　　Advisor：Dr. Chung-Ju Chang

國立交通大學
電信工程學系
碩士論文

A Thesis

Submitted to Department of Communication Engineering

College of Electrical and Computer Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in

Communication Engineering

June 2007

Hsinchu, Taiwan

中華民國九十六年六月

IEEE 802.16 上行鏈路無線系統之

動態優先次序資源分配

研究生：賴建安　　　　　　　指導教授：張仲儒 博士

國立交通大學電信工程學系碩士班

摘 要

隨著多媒體資訊在無線通訊系統上的傳輸量日益增加，提供傳輸服務品質保證 (quality of service, QoS) 是一個很重要的議題。為了有效使用系統資源並且提供傳輸服務品質的保證，我們在 IEEE 802.16 上行鏈路系統提出了一個基於動態優先次序的資源分配(dynamic priority-based resource allocation, DPRA)機制。對於急迫性較高的服務，我們給予較高的優先次序值(Priority value)，使具有較高優先次序的使用者能優先被分配系統資源做傳輸。我們也會根據每一種服務在不同時間的急迫性，動態調整其優先次序。我們提出的 DPRA 機制會在子通道(subchannel), 調變方法(modulation order), 以及能量(power)三方面找尋最佳化的資源分配方法，並且對同一個使用者做一致性分配(consistent allocation)。由模擬結果顯示，我們提出的方法可以達到傳輸速率最佳化以及 QoS 的滿足，並且能減少標頭傳輸(transmission overhead)以及降低運算複雜度。

Dynamic Priority-based Resource Allocation for IEEE 802.16 Uplink Wireless Systems

Student: Jian-Ann Lai　　　　　　Advisor: Dr. Chung-Ju Chang

Department of Communication Engineering

National Chiao Tung University

Abstract

To efficiently utilize the system resource and satisfy the QoS requirements in current multimedia transmission environment, we propose a dynamic priority-based resource allocation (DPRA) algorithm for IEEE 802.16 uplink system in this thesis. The goal of DPRA algorithm is to maximize system throughput while satisfying diverse QoS requirements. Four types of multimedia traffic defined in IEEE 802.16 are considered, including unsolicited grant service (UGS), real-time polling service (rtPS), non-real-time polling service (nrtPS), and best effort (BE) service. These multimedia services are given urgency degrees via dynamic priority values, which will be calculated at the beginning of each frame. The radio resource will be allocated to users according to priority values. Also the DPRA algorithm performs consistent allocation in aspects of subchannel, modulation order, and power. Simulation results show that the proposed DPRA algorithm outperforms the conventional algorithm in terms of system throughput and satisfaction of various QoS requirements. Besides, the proposed DPRA algorithm is also designed to have lower cost of transmission overhead and complexity.

誌謝

本篇論文的完成，首先要謝謝指導教授張仲儒教授在論文上的悉心指導，以及在生涯規劃上的諄諄教誨，讓我逐漸找到屬於自己的人生目標。也要感謝下班之餘不辭辛勞趕來和我們討論論文的芳慶學長以及一直很照顧我的志明學長，給了我很多寶貴的建議，解決許多論文上的疑惑。在實驗室陪伴我們立峰學長、文祥學長、耀興學長以及詠翰學長，謝謝你們在學業以及生活上曾經給我的幫助。兩年的碩士研究生活轉瞬即逝，其中的歡笑與淚水卻讓我有無限的感動與回憶。剛進實驗室時熱情歡迎我們的學長姐家源、俊帆、煖玉、琴雅，一起同甘共苦做研究的建興、正昕、世宏、佳泓、佳璇，帶給我們許多歡樂的學弟妹們邱胤、宗利、尚樺、巧瑩、英奇、浩翔、維謙，以及辛苦的助理惟媜，因為有大家熱情的陪伴，一起出遊、吃飯、打球、唱歌、玩樂，讓枯燥的研究生活額外增添了許多活力，很開心加入 LAB 701 寬頻網路實驗室這個大家庭，也祝福實驗室的大家學業順利，事業有成！

最後，謹以此篇論文獻給我最摯愛的父母和家人，以及所有一直關心、支持我的好朋友們。

建安 謹誌

2007 年 7 月

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Orthogonal frequency division multiplexing (OFDM) has been proposed as a promising technique for future multimedia wireless communication systems due to its ability to mitigate frequency selective fading, intersymbol interference (ISI) and its flexibility for adaptive modulation on each subcarrier. With the increasing demand of wireless access to the Internet, both downlink and uplink of the wireless system have to transport a great amount of multimedia traffic. Since the wireless channel condition of each user varies with time, adaptive resource allocation has been viewed as the key technology to provide efficient utilization of the limited system resource in current multiuser wireless communication environment.

Orthogonal frequency division multiple access (OFDMA) has been adopted for IEEE 802.16 broadband wireless access (BWA) system. Although the medium access control (MAC) signaling has been well defined in the IEEE 802.16 specifications [1], resource management and scheduling remain as open issues. The adaptive resource allocation for OFDMA systems has drawn enormous research interests, not only because the transmission power and modulation order can be adapted on each subcarrier, but also the multiple access scheme can be realized through dynamic subcarrier allocation. Besides, diverse quality of service (QoS) requirements for multimedia traffic transmission should be considered when developing an efficient resource allocation algorithm. Therefore, according to the time-varying channel state information (CSI), a radio resource allocation algorithm is needed to exploit the frequency diversity, multiuser diversity and time diversity so that the overall system resource can be efficiently utilized.

It has been thought as an optimal solution that each user is allocated on subcarriers with best channel qualities, and the allocated power on selected subcarriers is based on the water-filling principle. Subcarrier, bit, and power allocation algorithms for multiuser OFDMA systems have been investigated in many literatures to maximize the overall data rate or minimize the total transmitted power under several constraints. Wong et al. [2] proposed a Lagrangian-based algorithm to minimize the total transmission power under user's QoS requirements, which were defined by a specified data transmission rate and bit error rate (BER). However, the high computational complexity renders it impractical. To reduce the complexity, a near optimum dynamic multiuser subcarrier-and-bit allocation algorithm with low complexity was proposed in [3] to maximize the overall spectral efficiency.

Many papers considered the downlink resource allocation, and a few papers investigated the uplink resource allocation. Resource allocation of both downlink and uplink is primarily performed by the base station (BS), and the granted bandwidth for uplink transmission can still be scheduled to different service types by the subscriber station (SS). Das and Mandyam [4] considered the uplink transmission of the OFDMA system and developed an efficient algorithm for subcarrier and bit allocation for each user. The algorithm also included the power distribution over the selected set of subcarriers so that the total used power is minimized. The maximization of rate-sum capacity based on Shannon capacity formula in uplink OFDMA systems was considered in [5], where a greedy subcarrier allocation algorithm based on a marginal rate function and an iterative water-filling power allocation algorithm were proposed. The algorithm proposed in [5] was shown to achieve near optimal solution in uplink. Additionally, it was also concluded for downlink and uplink that equal power allocation over selected subcarriers of each user has similar performance compared to the water-filling scheme [6].

The basic allocation unit in IEEE 802.16 OFDMA system is a subchannel, which is composed of a set of subcarriers. For both downlink and uplink case, the achievable rate of a user increases with the number of subchannels assigned, the number of time slots assigned and the fraction of power allocated. Hosein [7] assumed that subchannels made up of a group of contiguous subcarriers are assigned to SSs in unit of time slots. Also the CSI on subchannels of each SS is assumed to be reported to BS periodically. Then the optimization problem using a utility function was formulated and a practical algorithm was provided to obtain a near-optimal solution. Singh and Sharma [8] also developed an efficient and fair scheduling (EFS) algorithm for each time slot in IEEE 802.16 OFDMA/TDD system by considering service priority as QoS requirements of different SS.

Based on previous works mentioned above, the QoS requirements and fairness issues are either omitted or simplified. A minimum required transmission data rate in each OFDMA symbol or a predefined weight which corresponds to the fixed priority scheme is usually adopted when considering the QoS requirements for resource allocation. However, each user has different service types, so their traffic model and buffer condition should be considered as well. Besides, with the presence of multimedia real-time traffic, the delay requirement should be also regarded as an important QoS issue. The packet of real-time traffic will be dropped if it is not transmitted within its delay bound. Thus the packet dropping rate and packet delay should be considered as QoS requirements for transmission of real-time traffic. A queue-aware uplink bandwidth allocation scheme at SS [9] was proposed for real-time and non-real-time polling services for IEEE 802.16 broadband wireless networks. The allocated bandwidth to the polling services can be dynamically adjusted according to channel quality and queueing state of the traffic. Thus, the packet level parameters such as packet dropping rate and packet delay can be maintained at the target level.

The tradeoff between system performance and computational complexity is also an important issue. The greedy algorithm which performs symbol-by-symbol allocation can achieve optimal solution, but it will result in high computational complexity which is not practical for real systems. According to the frame structure defined for downlink and uplink transmission (DL-MAP and UL-MAP) in IEEE 802.16 specifications, the symbol-by-symbol allocation algorithm will cost high transmission overhead. Besides, most resource allocation algorithms are designed for downlink and claimed to be compatible with uplink as well. However, the downlink frame structure (DL-MAP) and uplink frame structure (UL-MAP) are differently defined in IEEE 802.16 specifications. Thus an efficient and practical resource allocation algorithm for either downlink or uplink needs to be specifically designed to meet individual frame structures.

In this thesis, we propose a dynamic priority-based resource allocation (DPRA) algorithm for IEEE 802.16 uplink wireless systems. The goal of the proposed DPRA algorithm is to maximize system throughput while satisfying various QoS requirements. A priority value for each service type of each user is given and dynamically adjusted frame by frame according to individual QoS requirements and buffer conditions. Then the BS will dynamically allocate the uplink bandwidth and power to each user according to the priority values and CSI of each user. Thus the system resource will be allocated to users with services of high priority values and good CSI. Furthermore, in order to meet the uplink frame structure (UL-MAP) defined in IEEE 802.16 specifications as shown in fig. 1.1 and reduce transmission overhead and complexity, a consistent allocation scheme is developed in the proposed DPRA algorithm, where the allocation results for each user will not need to be searched symbol by symbol in each frame. Hence the tradeoff between system performance and computational complexity is expected to be better than other greedy searching algorithms.

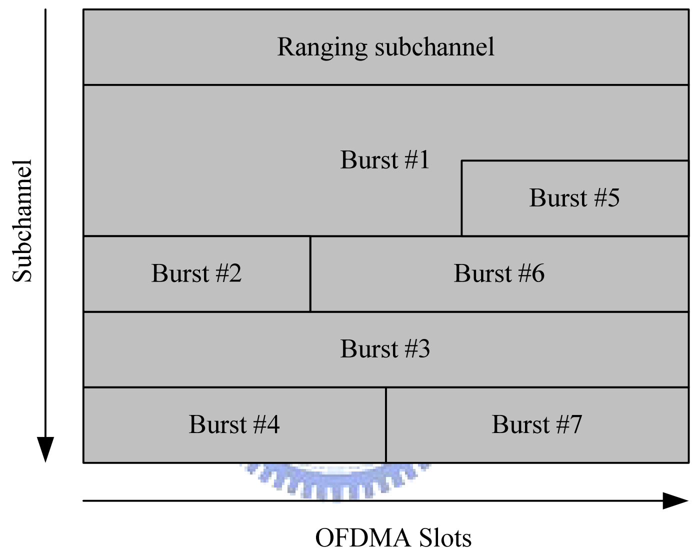

Figure 1.1: Uplink Frame Structure

Chapter 2

System Model

2.1 OFDMA System Architecture

An OFDM transmitter separates the serial symbols of input information into parallel forms and feed them into corresponding subcarriers. For discrete time signal model, inverse fast Fourier transform (IFFT) can be easily implemented as the OFDM transmitter following by a parallel-to-serial (P/S) converter. Due to the inter-carrier interference (ICI) resulting from the multipath effect, a cyclic prefix (CP) is attached at the front of the OFDM signal. Moreover, at the receiver end, at first the OFDM receiver removes the CP attached at the OFDM signal, and then converts the received serial signals into parallel forms. The received OFDM signal in parallel form is demodulated by the fast Fourier transform (FFT). Finally, after performing the FFT, the output of FFT will be converted from parallel into serial form, which will be the original pattern of the transmitted symbol of users. The major benefit of using CP for OFDM systems is to mitigate the multipath effect. When delay spread of the channel between transmitter and receiver is smaller than the length of CP, in frequency domain, the OFDM signal model of each received symbol can be formulated as the transmit symbol multiplied by the channel gain.

In this thesis, both the base station and the subscriber station are assumed to be equipped with single antenna. A set of adjacent subcarriers is grouped into an OFDM subchannel, so the basic allocation unit in frequency domain for the OFDMA system is a subchannel. Also, there are several ways for grouping subcarriers into a subchannel, however, it has been shown that grouping adjacent subcarriers would result in higher multiuser diversity, which maximizes the system throughput. Hence we will group a fixed number of adjacent subcarriers into each subchannel as defined for uplink transmission in IEEE 802.16 specifications [1]. Additionally, one subchannel along with one OFDMA slot will be the allocation unit in our consideration.

Suppose there are *N* subchannels in the OFDMA system, and we assume that each subchannel consists of *q* adjacent subcarriers, i.e. every subchannel contains the same number of adjacent subcarriers. There are *K* SSs going to communicate with one BS, and each SS can be viewed as a single user containing different type of services to be transmitted. Suppose that each service type of an SS can be viewed as one connection and each connection has its individual queuing buffer. Based on IEEE 802.16 in uplink transmission, data are transmitted in unit of frame and each frame is assumed to include *L* OFDMA slots.

The transmitted signal of user *k* on subchannel *n* at the ℓth OFDMA slot, denoted by $s_{k,n}^{(\ell)}$, is given as

$$s_{k,n}^{(\ell)} = \sqrt{\rho_{k,n}^{(\ell)}} \cdot d_{k,n}^{(\ell)}, \quad \forall k, n, \tag{2.1}$$

where $\rho_{k,n}^{(\ell)}$ is the allocated power for user *k* on subchannel *n*, and $d_{k,n}^{(\ell)}$ is the transmitted data symbol of user *k* on subchannel *n* at the ℓth slot. Note that the normalized M-QAM modulation is used so that the data symbol has unitary mean energy.

We assume that the coherence time of the wireless channel is larger than the duration of one frame duration. Hence the CSI is assumed to remain constant during one frame. The uplink CSI can be obtained by each SS's periodically reporting to the BS and perfect estimation of CSI on each subchannel of each user is assumed in this thesis. Moreover, since in IEEE 802.16 system, the SSs only report the uplink CSI on each subchannel, we won't be able to obtain the CSI on each subcarrier. We assume the channel gain of each adjacent subcarrier which a subchannel contains is the same, and hence equal power distribution on each subcarrier of the subchannel will be used. Let $h_{k,n}$ be the frequency domain uplink channel gain of user *k* on subchannel *n*. Based on the assumption mentioned above, we know that the channel gain of each subcarrier on subchannel *n* all equals to $h_{k,n}$. Note that the channel gain is not a function of slot time ℓ since it remains fixed during one frame time. Therefore, the received signal of user *k* on subchannel *n* at the ℓ th OFDMA slot, denoted by $y_{k,n}^{(\ell)}$, is given by

$$y_{k,n}^{(\ell)} = h_{k,n}\sqrt{\rho_{k,n}^{(\ell)}}\,d_{k,n}^{(\ell)} + z_{k,n}^{(\ell)}, \tag{2.2}$$

where $z_{k,n}^{(\ell)}$ is the complex white Gaussian noise of user *k* on subchannel *n* with zero mean and variance σ^2. Hence from (2.2), the received signal-to-interference -plus-noise-ratio (SINR) of user *k* on subchannel *n* at the ℓ th OFDMA slot, denoted by $SINR_{k,n}^{(\ell)}$, can be obtained as

$$SINR_{k,n}^{(\ell)} = \frac{\rho_{k,n}^{(\ell)}\left|h_{k,n}\right|^2}{\sigma^2}. \tag{2.3}$$

2.2 Overview of IEEE 802.16 Broadband Wireless Networks

The point-to-multi-point (PMP) model is adopted in this thesis, where multiple

SSs will communicate with one BS. We consider the OFDMA/TDD frame structure defined in IEEE 802.16 broadband wireless networks (BWN), where the transmission of downlink and uplink are separated in time division subframes. Since frame structure is used in IEEE 802.16 system for downlink and uplink, a frame is divided into subframes for both downlink and uplink transmission. Hence, the length of subframes can be determined dynamically by the BS and broadcasted to each SS through downlink and uplink map messages (DL-MAP and UL-MAP).

For uplink transmission in IEEE 802.16, each SS will have multiple types of service support. Besides, for each service type, it has different bandwidth request and grant mechanisms and different QoS requirements individually. IEEE 802.16 MAC supports two kinds of grant mechanisms for SS: grant per connection (GPC) and grant per SS (GPSS). In the case of GPC, each service type of a SS is regarded as one connection, and bandwidth is granted to each connection individually. On the other hand, for GPSS case, BS will grant an amount of total bandwidth for all connections to each SS. Then each SS is responsible for scheduling the bandwidth to each service type. Since the BS does not need to keep track of all connections in the GPSS mode, GPSS is more efficient for real systems.

By adopting GPSS mode for uplink transmission, we jointly consider bandwidth allocation and scheduling for each service type in BS. After the arrival of data traffic, SS will make bandwidth request to BS. The BS will dynamically allocate subchannel, time slots, and power to each SS according to uplink CSI, QoS requirements and buffer condition of each service type. Then a total amount of bandwidth allocated for each service type of the SS will be granted by BS for each SS to transmit.

2.3 Service Types

IEEE 802.16 defines the following four types of service to support real-time and non-real-time data transmission, and each of then has different QoS requirements, which will be stated below:

1) Unsolicited Grant Service (UGS):

UGS supports real-time traffic that generates fixed size of data packets periodically, such as Voice over IP (VoIP). It needs to be granted a fixed amount of bandwidth in each frame for this type of service.

2) Real-time Polling Service (rtPS):

It is designed to support real-time traffic which generates variable size of data packets, such as video streaming. It is a delay sensitive traffic so that the delay requirement is an important QoS issue for rtPS. The amount of bandwidth granted for this type of service needs to be determined dynamically based on the packet delay and dropping rate requirements.

3) Non-real-time Polling Service (nrtPS):

It is designed to support delay-tolerant data streams while a minimum data transmission rate is required, such as the HTTP traffic. The granted bandwidth for nrtPS is also determined dynamically according to the QoS requirement and buffer condition.

4) Best effort (BE):

BE service is designed to support data streams which have no QoS requirement. It will be transmitted when system resource is available. Thus the bandwidth left after serving the UGS, rtPS and nrtPS traffic will be allocated for BE service.

2.4 Power Allocation

For achieving required bit error rate (BER), power allocation is determined based on the minimum required SINR. Besides, since the allocated power on each subchannel *n* will be equally distributed on each subcarrier grouped into subchannel *n*, the allocated power to user *k* on each subcarrier of subchannel *n* all equals to $\rho_{k,n}^{(\ell)}$. An approximation of the BER when using M-QAM modulation is given by [10]:

$$BER \simeq 0.2e^{-1.5\frac{SINR}{M-1}}. \tag{2.4}$$

From (2.4), under the required BER of user *k*, which is BER_k^*, the minimum required SINR of user *k*, denoted by $SINR_k^*$, can be obtained as

$$SINR_k^* = -\frac{\ln(5BER_k^*)}{1.5}(M-1). \tag{2.5}$$

Therefore, based on the desired BER and minimum required SINR for user *k*, the allocated power to user *k* on each subcarrier of subchannel *n* can be obtained by

$$\rho_{k,n}^{(\ell)} = \frac{-\ln(5BER_k^*)(M-1)\sigma^2}{1.5\left|h_{k,n}\right|^2}. \tag{2.6}$$

Since each subchannel consists of *q* adjacent subcarriers, the total allocated power of user *k* on subchannel *n* at the ℓth OFDMA slot, denoted as $p_{k,n}^{(\ell)}$, can be obtained by summing up the allocated power on each subcarriers of subchannel n:

$$p_{k,n}^{(\ell)} = q \cdot \rho_{k,n}^{(\ell)}. \tag{2.7}$$

From (2.6) and (2.7), note that the allocated power is a function of BER_k^* and modulation order. Hence, $p_{k,n}^{(\ell)}$ will be denoted by $p_{k,n}^{(\ell)}\left(BER_k^*, x_{k,n}^{(\ell)}\right)$ if needed in the following content. Besides, the allocated power for each user should be sufficient enough so that the BER requirements can be maintained at a target level.

Chapter 3

Problem Formulation

3.1 Design Constraints

In this thesis, our goal of the resource allocation is to maximize the overall system throughput while satisfying the QoS requirements. Therefore the optimization problem would try to finding a set of assignment variables which indicate the modulation order of each OFDMA symbol such that the desired objective is achieved. We define the assignment variable as the number of bits carried by each subcarrier of subchannel n per symbol when using M-QAM modulation. Thus $x_{k,n}^{(\ell)}$ is denoted as the modulation order of user k on subchannel n for the lth OFDMA symbol and is given by:

$$x_{k,n}^{(\ell)} = \begin{cases} 0, & \text{if not assigned} \\ 2, & \text{if QPSK modulated} \\ 4, & \text{if 16-QAM modulated} \\ 6, & \text{if 64-QAM modulated} \end{cases}, \quad 1 \le k \le K,\ 1 \le n \le N,\ 1 \le \ell \le L.$$

We denote the assignment vector be the allocation results for the lth OFDMA symbol as a column vector shown below:

$$\mathbf{x}^{(\ell)} \equiv \left[x_{1,1}^{(\ell)}, \cdots, x_{1,N}^{(\ell)}, \cdots, x_{k,1}^{(\ell)}, \cdots, x_{k,N}^{(\ell)}, \cdots, x_{K,1}^{(\ell)}, \cdots, x_{K,N}^{(\ell)} \right]^T.$$

The assignment matrix of the allocation results in one frame consisting L OFDMA symbols can be represented as

$$\mathbf{x} = \left[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(L)}\right].$$

Therefore the total allocated bits to user k in the current frame, denoted as R_k, can be calculated by

$$R_k \equiv R_k(\mathbf{x}) = \sum_{\ell=1}^{L} \sum_{n=1}^{N} q \cdot x_{k,n}^{(\ell)}. \tag{3.1}$$

The uplink resource allocation scheme proposed in this thesis aims to determine a near optimal assignment matrix in one frame by maximizing the overall system throughput while satisfying the QoS requirements of each service class for every user. In the design of the allocation algorithms, there are four constrains we need to concern about, which are stated as follows:

(i) User's transmit power constraint:

The total allocated power for uplink transmission of each SS in one symbol should have a limitation. We denote $p_{k,\max}$ as the maximum allowable uplink transmission power and obtain the power constraint as following:

$$\sum_{n=1}^{N} p_{k,n}^{(\ell)} \leq p_{k,\max}, \quad \forall \ell, k. \tag{3.2}$$

(ii) Buffer constraint:

We assume that $B_{k,s}$ is the number of bits in the buffer for service class s of user k. Thus the total number of bits in the buffer of every service for user k, denoted as B_k can be obtained by:

$$B_k \equiv \sum_{s} B_{k,s}, \quad \forall k. \tag{3.3}$$

In order to make full utilization of the limited radio resource, transmission efficiency is part of our consideration. For transmission efficiency, the total allocated bit to user k in one frame, R_k, should not be larger than the total buffer occupancy of user k in this frame, B_k. We can obtain the buffer constraint as following:

$$R_k \leq B_k, \quad \forall k. \tag{3.4}$$

(iii) Slot allocation constraint:

In every single cell, each slot of every subchannel can only be allocated to one user, which is a basic constraint for the single-antenna system of one cell. Hence the slot allocation constraint can be expressed as:

$$\sum_{k=1}^{K} \operatorname{sgn}(x_{k,n}^{(\ell)}) = 1, \quad \forall n, \ell. \tag{3.5}$$

(iv) QoS fulfillment constraint:

For each type of service class introduced in chapter II, there are individual QoS requirements for each of them. In order to fulfill the QoS requirement of every user's service, we define $\gamma_{k,s}$ as the priority value for service type s of user k, which is in terms of the minimum required transmission bits per frame. Then we define the total priority value $\hat{R}_k$, which is the total minimum required transmission bits for all service types of user k in this frame. So $\hat{R}_k$ can be obtained as:

$$\hat{R}_k = \sum_{s} \gamma_{k,s}, \quad \forall k. \tag{3.6}$$

Therefore the QoS requirement constraint can be expressed as:

$$R_k \geq \hat{R}_k, \quad \forall k. \tag{3.7}$$

This means that the minimum transmission bits of user k in this frame has to be larger than $\hat{R}_k$, otherwise the QoS requirements will not be fulfilled. Intuitively, the larger the priority value $\hat{R}_k$ is, the more urgent of user k is. Hence, the more resource would be allocated to user k.

Therefore, the optimization problem for the uplink adaptive resource allocation can be formulated as following:

Objective:

$$\mathbf{x}^* = \arg\max_{\mathbf{x}} \sum_{k=1}^{K} R_k(\mathbf{x})$$

subject to the following constraints

$$\text{(i)} \sum_{n=1}^{N} p_{k,n}^{(\ell)} \le p_{k,\max}, \quad \forall \ell, k \tag{3.8}$$

$$\text{(ii)}\ R_k \le B_k, \quad \forall k$$

$$\text{(iii)} \sum_{k=1}^{K} \mathrm{sgn}(x_{k,n}^{(\ell)}) = 1, \quad \forall n, \ell$$

$$\text{(iv)}\ R_k \ge \hat{R}_k, \quad \forall k.$$

Based on the formulation in (3.8), the objective of the optimization problem is to find an optimal set of the assignment variable $\mathbf{x}^*$, which is defined as $x^* \equiv \left[\mathbf{x}^{*(1)}, \mathbf{x}^{*(2)}, \cdots \mathbf{x}^{*(L)}\right]$ such that the system throughput is maximized under the four constraints mentioned above.

3.2 Derivation of Priority Values

Since the characteristic of the traffic burst and diverse QoS requirements can vary with time, the urgency of each service type will also be different with time. Hence, the requirements of radio resource have to be adjusted frame by frame such that the QoS requirements can be satisfied and the radio resource will be utilized efficiently. Accordingly, we propose a dynamic scheme which dynamically adjusts the priority value of each service type for each user frame by frame, and then BS will allocate resource according the priority values. We define the priority value of service s for user k in unit of bits per frame as following:

$$\gamma_{k,s} = \begin{cases} \gamma_{k,UGS}, & \text{if s} \in UGS \\ \gamma_{k,rtPS}, & \text{if s} \in rtPS \\ \gamma_{k,nrtPS}, & \text{if s} \in nrtPS \\ 1, & \text{if s} \in BE, \end{cases} \tag{3.9}$$

where $\gamma_{k,UGS}$ remains a constant since it needs to be granted a constant bandwidth for its transmission. However, $\gamma_{k,rtPS}$ and $\gamma_{k,nrtPS}$ are dynamically adjusted frame by frame. The definition of priority values of UGS, rtPS, nrtPS and BE will be stated in the following.

Since the UGS supports constant-bit-rate real-time traffic, the priority value $\gamma_{k,UGS}$ is set to be a constant for each frame. In other words, the resource allocated to UGS has to be guaranteed, otherwise its QoS requirement will be violated.

Let $\hat{D}_k$ be the maximum delay tolerance of rtPS head-of-line (HOL) packet for user *k*. Also denote D_k be the current delay of the rtPS HOL packet for user *k*, which is the time interval from the arrival of the packet to the present frame. Both $\hat{D}_k$ and D_k are in unit of frame. So we can derive the remaining time for the real-time packet to be dropped is

$$\Delta D_k \equiv \hat{D}_k - D_k. \tag{3.10}$$

The derivation of ΔD_k is directly from the delay requirement of rtPS, which means that the rtPS HOL packet should complete its transmission within ΔD_k. Otherwise the delay requirement will not be satisfied and the packet will be dropped. Therefore, based on a predefined delay threshold D_{th} and the residual buffered bits of the rtPS HOL packet for user *k*, denoted by $B_{k,rtPS}$, the priority value for rtPS of user k is defined as:

$$\gamma_{k,rtPS} = \begin{cases} B_{k,rtPS}, & \text{if } \Delta D_k \le D_{th} \\ \dfrac{B_{k,rtPS}}{\Delta D_k + \log(\Delta D_k)}, & \text{if } \Delta D_k > D_{th} \end{cases}. \tag{3.11}$$

If ΔD_k is smaller than the threshold D_{th}, it means that the rtPS HOL packet of user *k* is in high urgency so that all the residual bits should be transmitted in the current frame. Otherwise, if ΔD_k is larger than the threshold D_{th}, the priority value is derived based on the averaged required transmission rate. The average required transmission rate can be derived directly from $B_{k,rtPS}$ and ΔD_k. Note that the denominator is added with a bias, $\log(\Delta D_k)$.With a larger value of ΔD_k, it means that the HOL packet of user *k* is more delay tolerable and not highly urgent. Hence the bias on the denominator will reduce the priority value of user *k* such that the user with higher priority value service will be allocated resource first.

For nrtPS, the average transmission rate should be larger than the minimum required transmission rate, $\hat{\gamma}_{k,nrtPS}$. Denote $T_{k,nrtPS}$ be the maximum tolerant transmission interval for nrtPS service of user *k* such that the minimum required transmission rate is fulfilled. Thus we can derive the following inequality,

$$\frac{B_{k,nrtPS}+B'_{k,nrtPS}}{T_{k,nrtPS}+T'_{k,nrtPS}} \geq \hat{\gamma}_{k,nrtPS}, \tag{3.12}$$

where $B_{k,nrtPS}$ is the residual buffered bits of the nrtPS HOL packet for user *k*, $B'_{k,nrtPS}$ is the total transmitted bits for nrtPS user *k* in previous time, and $T'_{k,nrtPS}$ is the total active transmission period for nrtPS user *k* in past time. Note that both $T_{k,nrtPS}$ and $T'_{k,nrtPS}$ are in unit of frames. From (3.12), we can derive $T_{k,nrtPS}$ as

$$T_{k,nrtPS} = \frac{B_{k,nrtPS}+B'_{k,nrtPS}}{\hat{\gamma}_{k,nrtPS}} - T'_{k,nrtPS}. \tag{3.13}$$

Therefore, the priority value for nrtPS of user k is given by

$$\gamma_{k,nrtPS} = \begin{cases} B_{k,nrtPS}, & \text{if } T_{k,nrtPS} \leq T_{th} \\ \dfrac{B_{k,nrtPS}}{T_{k,nrtPS}+\log(T_{k,nrtPS})}, & \text{if } T_{k,nrtPS} > T_{th} \end{cases}, \tag{3.14}$$

where T_{th} is a predefined threshold for nrtPS. Same concept as the priority value of rtPS, if $T_{k,nrtPS}$ is smaller than T_{th}, it means that the packet of nrtPS user k is highly urgent so the residual bits of its HOL packet should be transmitted in current frame. Otherwise the priority value for nrtPS user k is derived according to the average required transmission rate, denoted by $B_{k,nrtPS} / (T_{k,nrtPS} + \log(T_{k,nrtPS}))$. Note that the denominator is added with a bias value, which is $\log(T_{k,nrtPS})$. The bias is used to reduce the priority value, especially for delay tolerant user with large value of $T_{k,nrtPS}$. Thus the urgent user with higher priority value will have more chance to be allocated resource for transmission. Note that the two thresholds, D_{th} and T_{th}, defined for rtPS and nrtPS respectively, can be specified with different values to distinguish the priority of real-time from non-real-time service.

For BE service, since there is no delay or transmission rate requirement to be satisfied, its priority value should be the lowest among all services. Thus the priority value of BE service is set to one.

In summary, the higher the priority value is, the more resource will be allocated to the user such that its QoS requirement can be fulfilled. However, the service with low priority value can experience the time diversity. Note that the user with service of higher priority value will be served first in current frame, so the priority value of the served user will decrease after the current frame. Hence the user with low priority service can still be served after the users with high priority services are already served. Besides, the delay tolerant users exploit time diversity by transmitting when having good channel condition, and this will increase the overall system throughput.

Chapter 4

Dynamic Priority-based Resource Allocation Algorithm

In order to reduce the transmission overhead and meet the frame structure defined in IEEE 802.16 specifications [1], we develop a consistent resource allocation scheme. Once the modulation order and power for a certain slot in subchannel *n* are allocated for one user, we will allocate the neighboring slot of the previous allocated slot on subchannel *n* with the same modulation order and power for that user to finish transmitting its data. The user will transmit with the same power and modulation order on the same subchannel *n*, so the allocation results for that user will remain the same during its allocated slots in this frame. In other words, the allocation results of each user will be determined frame-by-frame rather than symbol by symbol. Hence this scheme is called consistent allocation and is expected to reduce the computation complexity and transmission overhead..

Since the objective of the optimization problem defined in (3.8) is to maximize the overall system throughput under some constraints, we propose a dynamic priority-based resource allocation (DPRA) algorithm and perform consistent allocation which intends to find an optimum set of assignment results. The proposed DPRA algorithm will perform subchannel selection, power allocation, modulation order assignment, and consistent allocation for the selected user. Hence the optimal user-subchannel pairs will be selected to maximize system throughput and satisfy QoS requirements.

After deriving the dynamic priority values for each service types of each user and the CSI of each user's uplink channel, BS will perform the DPRA algorithm based on the derived priority values. The procedure of the proposed DPRA algorithm will be executed through the following four major phases and each phase contains several steps to complete.

- **Phase 1 – User and subchannel selection**

For the purpose of satisfying QoS requirements of each backlogged users, we have to select users from the backlogged users set according to the priority values. So the user with the higher priority value service will be selected and allocated resource prior to other users. Besides, to achieve the goal of maximizing system throughput, we have to choose the subchannel with the highest CSI for the selected user to transmit using higher modulation order. Therefore we will select an optimal user-subchannel pair. This phase contains step 1, which the detail functions of them will be given below.

- **Phase 2 – Power constraint and available resource checking**

In order to fulfill the power constraint and maximize system throughput, we have to find the largest modulation order so that the transmission power constraint of the selected user is satisfied. Furthermore, based on the slot allocation constraint and the QoS fulfillment constraint, we will calculate the amount of required resource for the selected user and check if the amount of available resource is enough. When the amount of available resource is not sufficient for the selected user, we will search other subchannels and find the one whose available resource can satisfy the requirements of the selected user. This phase contains step 2-6 and the detail functions of them will be stated below.

- **Phase 3 – Allocation for all types of services**

To reduce the transmission overhead, we will perform consistent allocation by allocating neighboring slots of the same subchannel to each service type for the selected user. Since UGS supports constant bit rate traffic, it has the highest priority and we will start allocating available slots for UGS. Next we will allocate slots for rrPS and nrtPS according to their priority values. The service which has higher priority value means that it requires more slots. For each available slot, we will select the service among rtPS and nrtPS which requires more slots and allocate one slots for the selected service and decrease the required number of slots for the selected service by one. This is the principle we use to allocate slots for rtPS and nrtPS. So we will continuously allocate slots iteratively until rtPS and nrtPS require no more slots or there are no more available slots to allocate. After finishing allocating for UGS, rtPS and nrtPS, we will finally consider BE service. If there are remaining available slots and the buffer of BE service is not empty, we will allocate slots for BE. This phase contains step 7-9 and the function of each step will be stated below.

- **Phase 4 – Slots remapping**

Since the principle of the proposed DPRA algorithm performs consistent allocation, we may allocate from the slot of one subchannel to the slots of the next subchannel. However, some slots of the next subchannel might have been allocated to other users. This will violate the slot allocation constraint in (3.8). Therefore we have to shift other users to the neighboring available slots which have not been allocated. This function of shifting is called “remapping” and the detail function of it will be given in step 10.

An initialization step must be done before we start the DPRA algorithm and the function of initialization is given below.

Initialization:

The assignment vector $\mathbf{x}^{(\ell)}$ and allocated power $p_{k,n}$ are set to be zero, which means that no resource is allocated to any user. Denote i_n the slot index of subchannel *n* which has been allocated, and δ_k the total number of slots allocated to user *k*. They are both initialized to be zero. Denote N^{free} the set of free subchannels of the system and N^k the set of subchannels on which user *k* is allocated. They are initialized as $N^{free} = \{n \,|\, 1 \le n \le N\}$ and $N^k = \{\phi\}$, for $\forall k$. We also let Ω_k be the set of backlogged users whose buffers are not empty. The system frame has N subchannels and each subchannel has L OFDMA slots, so the total slots in the frame denoted by Φ are $L \times N$. The function of initialization is given below.

Function: *Initialization*

$$
\begin{aligned}
&\text{Set } \mathbf{x}^{(\ell)} = 0,\ p_{k,n} = 0,\ i_n = 0,\ \delta_k = 0,\ \forall k, n, \ell \\
&N^{free} = \{n \,|\, 1 \le n \le N\},\ N^k = \{\phi\},\ \forall k \\
&\Omega_k = \{k \,|\, B_k > 0,\ 1 \le k \le K\} \\
&\Phi = L \times N.
\end{aligned}
$$

After the initialization, there are ten steps in the proposed DPRA algorithm and each step represents a specific function which will be stated in detail individually as following. While there are free subchannels and backlogged users, the DPRA algorithm will be executed through step 1 to step 10 iteratively as shown below:

Function: *DPRA algorithm*

while $|N^{free}| > 0$ and $|\Omega_k| > 0$

execute step 1

execute step 2

⋮

execute step 10

end while.

- **Phase 1 – User and subchannel selection :**

Step 1) – *User-subchannel selection:*

Let Ω be the candidate users set which contains backlogged users having the service with the highest priority value among all users' services. We also denote $\gamma_{\max}$ the maximum priority vale. From the candidate users set Ω and free subchannel set N^{free}, we select an optimal pair of user and subchannel (k^*, n^*) which means user k^* has the best CSI on subchannel n^*. The function is shown below.

Function: *User-subchannel selection*

$$\gamma_{\max} = \arg\max_{k\in\Omega_k} \gamma_{k,s}, \forall s \in \{UGS, rtPS, nrtPS, BE\}$$
$$\Omega = \left\{ k \middle| k = \arg\max_{k\in\Omega_k} \gamma_{k,s}, \forall s \in \{UGS, rtPS, nrtPS, BE\} \right\}$$
$$(k^*, n^*) = \arg\max_{k\in\Omega,\ n\in N^{free}} h_{k,n}.$$

- **Phase 2 – Power constraint and available resource checking:**

Step 2) – *Power allocation:*

Once an optimal pair of user and subchannel (k^*, n^*) is selected, we perform power allocation for user k^* on subchannel n^*. We try to find the highest modulation order $x_{k^*,n^*}^{(i_{n^*})}$ which satisfies the power constraint. The maximum modulation order we consider is 64-QAM. If the power constraint cannot be fulfilled even with the lowest modulation order, the selected user k^* will be removed from the backlogged users set. The function is given below.

Function: *Power allocation*

while $p_{k^*,n^*}(BER_{k^*}^*,\ x_{k^*,n^*}^{(i_{n^*})} + 2) < p_{k^*,\max}$

$x_{k^*,n^*}^{(i_{n^*})} = x_{k^*,n^*}^{(i_{n^*})} + 2$

end while

if $x_{k^*,n^*}^{(i_{n^*})} = 0$, *then* $\Omega_k = \Omega_k - \{k^*\}$ and go to step 1.

Step 3) – *Required slot calculation:*

Since each subchannel contains q subcarriers, we denote $R_{k^*,n^*}^{(i_{n^*})}$ the allocated bits per slot for user k^* on subchannel n^* by $R_{k^*,n^*}^{(i_{n^*})} = x_{k^*,n^*}^{(i_{n^*})} \cdot q$. We also denote α_{k^*} the number of required slot for data transmission of each service type for the selected user k^*. If the maximum priority value is larger than 1, which is the priority value of set for BE service, the BE data will not be considered when calculating α_{k^*}. We will allocate the system resource for other high priority services. The function is given below.

Function: *Required slot calculation*

$$R_{k^*,n^*}^{(i_{n^*})} = x_{k^*,n^*}^{(i_{n^*})} \cdot q,\ N^{k^*} = N^{k^*} + \left\{n^*\right\}$$

$$if\ \gamma_{\max} > 1$$

$$\alpha_{k^*} = \left\lceil \frac{\gamma_{k^*,UGS} + B_{k^*,rtPS} + B_{k^*,nrtPS} + \gamma_{k^*,BE}}{R_{k^*,n^*}^{(i_{n^*})}} \right\rceil$$

$$else$$

$$\alpha_{k^*} = \left\lceil \frac{\gamma_{k^*,UGS} + B_{k^*,rtPS} + B_{k^*,nrtPS} + B_{k^*,BE}}{R_{k^*,n^*}^{(i_{n^*})}} \right\rceil$$

Step 4) – *Available slot calculation:*

Because some slots of subchannel n^* may already be allocated to other users, user k^* may not be granted as many slots as required. Therefore, besides the free slots of the selected subchannel n^*, we also consider the free slots of the following two subchannels right after subchannel n^*. Let $\bar{\alpha}_{k^*,n^*}$ be the number of slots that can be granted for user k^* when we start allocating from subchannel n^*, and it is calculated via the function shown below. Note that $(L - i_{n^*})$ denote the number of free slots of subchannel n^*.

Function: *Available slot calculation*

$$
\begin{array}{l}
\bar{\alpha}_{k^*,n^*} = \alpha_{k^*} \\
if\ \ \bar{\alpha}_{k^*,n^*} > L - i_{n^*} \\
\quad if\ \ \bar{\alpha}_{k^*,n^*} - (L - i_{n^*}) > L - i_{n^*+1} \\
\quad\quad if\ \ \bar{\alpha}_{k^*,n^*} - (L - i_{n^*}) - (L - i_{n^*+1}) < L - i_{n^*+2} \\
\quad\quad\quad\quad \bar{\alpha}_{k^*,n^*} = \alpha_{k^*} \\
\quad\quad else \quad \bar{\alpha}_{k^*,n^*} = (L - i_{n^*}) + (L - i_{n^*+1}) + (L - i_{n^*+2}) \\
\quad else \quad \bar{\alpha}_{k^*,n^*} = \alpha_{k^*} \\
else \quad \bar{\alpha}_{k^*,n^*} = \alpha_{k^*}.
\end{array}
$$

Step 5) – *Power rechecking:*

As long as $\bar{\alpha}_{k^*,n^*} > L$, user k^* will be allocated on more than one subchannel simultaneously at some slot interval. Since we only consider the power allocated on one subchannel, we should confirm whether the power constraint is still satisfied when $\bar{\alpha}_{k^*,n^*} > L$. Let c be the number of subchannel allocated to user k^*. If the power constraint is violated, we let $\bar{\alpha}_{k^*,n^*} = L \cdot (c-1)$ by decreasing the number of allocated subchannel. Thus we perform this function iteratively until the power constraint is fullfilled. The function is given below.

Function: *Power rechecking*

$$
\begin{array}{l}
c = \left\lceil \bar{\alpha}_{k^*,n^*} / L \right\rceil \\
if\ \ \bar{\alpha}_{k^*,n^*} > L \\
\quad \text{while } p_{k^*,n^*}^{(i_{n^*})}(BER_{k^*}^*,\ x_{k^*,n^*}^{(i_{n^*})}) \cdot c > p_{k^*,\max} \\
\quad\quad \bar{\alpha}_{k^*,n^*} = L \cdot (c-1) \\
\quad\quad c = c - 1 \\
\quad \text{end while.}
\end{array}
$$

Step 6) – *Maximum available slots finding:*

Note that $R_{k^*,n^*}^{(i_{n^*})} \cdot \bar{\alpha}_{k^*,n^*}$ is the number of allocated bits for user k^*. If it is smaller than the actual required value, which is $R_{k^*,n^*}^{(i_{n^*})} \cdot \alpha_{k^*}$, the QoS requirements of user k^*

may not be fulfilled. We will search other free subchannels and choose the one which has the maximum allocated bits $R_{k^*,n^*}^{(i_{n^*})} \cdot \bar{\alpha}_{k^*,n^*}$ for user k^*. Hence, the QoS requirements will tend to be satisfied. The function is shown below.

Function: *Maximum available resource finding*

$$
\begin{aligned}
& if\ R_{k^*,n^*}^{(i_{n^*})} \cdot \bar{\alpha}_{k^*,n^*} < R_{k^*,n^*}^{(i_{n^*})} \cdot \alpha_{k^*} \\
& \quad if\ N^{free} - N^{k^*} \neq \{\phi\} \\
& \quad\quad n^* = \arg \max_{n \in N^{free} - N^{k^*}} h_{k^*,n},\ N^{k^*} = N^{k^*} + \{n^*\} \\
& \quad\quad \text{while } p_{k^*,n^*}^{(i_{n^*})}(BER_{k^*}^*,\ x_{k^*,n^*}^{(i_{n^*})} + 2) < p_{k^*,\max} \\
& \quad\quad\quad x_{k^*,n^*}^{(i_{n^*})} = x_{k^*,n^*}^{(i_{n^*})} + 2 \\
& \quad\quad \text{end while} \\
& \quad\quad if\ x_{k^*,n^*}^{(i_{n^*})} > 0,\ \text{go to step 4} \\
& \quad\quad else \quad n^* = \arg\max_n \left(R_{k^*,n}^{(i_{n^*})} \cdot \bar{\alpha}_{k^*,n} \right) \\
& \quad else \\
& \quad\quad n^* = \arg\max_n \left(x_{k^*,n}^{(i_{n^*})} \cdot \bar{\alpha}_{k^*,n} \right).
\end{aligned}
$$

- **Phase 3 – Allocation for all types of services:**

Step 7) – *Allocation for UGS:*

Let α_{k^*} be the total number of available slot allocated to user k^* in this frame. First we allocate slots for UGS of user k^* from the first available slot of subchannel n^* orderly. The slots will be allocated to UGS until the buffer of UGS $\mathrm{B}_{k^*,UGS}$ is empty or there are no more available slots. If all the available slots of subchannel n^* have been allocated, we will allocate from the first slot of the next subchannel n^*+1 orderly to allocate the required slots for UGS. The slots will be allocated to UGS until the buffer of UGS $\mathrm{B}_{k^*,UGS}$ is empty or there are no more available slots. The function is given below.

Function: *Allocation for UGS*

$\alpha_{k^*} = \bar{\alpha}_{k^*,n^*}$, $i = i_{n^*}$, $n = n^*$, $x = x_{k^*,n^*}^{(i_{n^*})}$

while $\alpha_{k^*} > 0$ and $\beta_{k^*,UGS} > 0$ and $\mathrm{B}_{k^*,UGS} > 0$

if $i \le L$

$x_{k^*,n}^{(i)} = x$

else

$i = 1,\ n = n+1,\ x_{k^*,n}^{(i)} = x$

$i = i+1,\ i_n = i_n + 1,\ \alpha_{k^*} = \alpha_{k^*} - 1,\ \beta_{k^*,UGS} = \beta_{k^*,UGS} - 1$

$\delta_{k^*} = \delta_{k^*} + 1,\ \phi = \phi - 1,\ \mathrm{B}_{k^*,UGS} = \mathrm{B}_{k^*,UGS} - R_{k^*,n^*}^{(i)},\ R_{k^*} = R_{k^*} + R_{k^*,n^*}^{(i_{n^*})}$

if $i_n = L,\ then\ N^{free} = N^{free} - \{n\}$

end while.

Step 8) – *Allocation for rtPS and nrtPS:*

Following the slots allocated for UGS, then we will allocate slots for rtPS and nrtPS of user k^*. The total required transmission bits for rtPS and nrtPS is $\gamma_{k^*,rtPS} + \gamma_{k^*,nrtPS}$. As long as $\alpha_{k^*} > 0$ and $\gamma_{k^*,rtPS} + \gamma_{k^*,nrtPS} > 0$, we will allocate slots for rtPS and nrtPS of user k^* iteratively. The service among rtPS and nrtPS which has higher priority value will be allocated one slot on each iteration and decrease its priority value by $R_{k^*,n}^{(i)}$. The function is shown below.

Function: *Allocation for rtPS and nrtPS*

$$\Omega_{PS} = \{rtPS, nrtPS\}$$

while $\alpha_{k^*} > 0$ and $(\gamma_{k^*,rtPS} + \gamma_{k^*,nrtPS}) > 0$

$s^* = \arg\max_{s\in\Omega_{PS}} \gamma_{k^*,s}$

if $i \le L$

$x_{k^*,n}^{(i)} = x$

else

$i = 1,\ n = n+1,\ x_{k^*,n}^{(i)} = x$

$i = i+1,\ i_n = i_n + 1,\ \alpha_{k^*} = \alpha_{k^*} - 1,\ \gamma_{k^*,s} = \gamma_{k^*,s} - R_{k^*,n}^{(i)}$

$\delta_{k^*} = \delta_{k^*} + 1,\ \phi = \phi - 1,\ R_{k^*} = R_{k^*} + R_{k^*,n^*}^{(i_{n^*})}$

$if\ i_n = L,\ then\ N^{free} = N^{free} - \{n\}$

end while.

Step 9) – *Allocation for BE service:*

As long as there are remaining available slots for user k^* and the buffer of BE service of user k^* is not empty, we can allocate slots for BE. Otherwise, the BE data will not be transmitted in this frame and might be transmitted in the following frames whenever the high priority services have been transmitted and radio resource becomes available for BE. The function is given below.

Function: *Allocation for BE service*

while $\alpha_{k^*} > 0$ and $B_{k^*,BE} > 0$

if $i \le L$

$x_{k^*,n}^{(i)} = x$

else

$i = 1,\ n = n+1,\ x_{k^*,n}^{(i)} = x$

$i = i+1,\ i_n = i_n + 1,\ \alpha_{k^*} = \alpha_{k^*} - 1,\ \mathrm{B}_{k^*,BE} = \mathrm{B}_{k^*,BE} - R_{k^*,n}^{(i)}$

$\delta_{k^*} = \delta_{k^*} + 1,\ \phi = \phi - 1,\ R_{k^*} = R_{k^*} + R_{k^*,n^*}^{(i_{n^*})}$

$if\ i_n = L,\ then\ N^{free} = N^{free} - \{n\}$

end while.

- **Phase 4 – Slots remapping:**

Step 10) – *Remapping:*

Since we perform consistent allocation, we may allocate user k^* on the slot which has been allocated to another user and causes user-overlapped slot. Hence the slot allocation constraint will be violated. Therefore we must shift the user except k^* on the overlapped slots to the free slots, and the step is called "remapping". Let σ_{k^*} be the length of shifting in unit of slot. Thus the overlapped user $\bar{k}$ on the ℓ th slot will be shifted σ_{k^*} slots to the $(\ell+\sigma_{k^*})$ th slot with the same modulation order. Note that if $(\ell+\sigma_{k^*})>L$, user $\bar{k}$ will be shifted to the next subchannel. Note that we only consider remapping of at most two subchannels following the starting allocated subchannel n^*. The remapping function is described as following.

Function: *Remapping*

Let $j=0$

for $n=n^*+1$ to N

  for $\ell=1$ to L

    if $x_{\bar{k},n}^{(\ell)}>0$ and $\forall \bar{k}\neq k^*$

      if $\ell+\sigma_{k^*}\le L$ and $n=n^*+1$

        $x_{\bar{k},n}^{(\ell+\sigma_{k^*})}=x_{\bar{k},n}^{(\ell)},\ x_{\bar{k},n}^{(\ell)}=0,\ j=j+1$

      else if $\ell+\sigma_{k^*}>L$ and $n=n^*+1$

        $x_{\bar{k},n+1}^{(\ell+\sigma_{k^*}-L)}=x_{\bar{k},n}^{(\ell)},\ x_{\bar{k},n}^{(\ell)}=0,\ j=j+1$

      else if $n=n^*+2$

        $x_{\bar{k},n}^{(\ell+\sigma_{k^*}+j-L)}=x_{\bar{k},n}^{(\ell)},\ x_{\bar{k},n}^{(\ell)}=0$

end

$\Omega_k=\Omega_k-\{k^*\}$.

So far, we have allocated slots for each service of user k^*, so we will eliminate k^* from the backlogged users set Ω_k at the end of this step and go back to step 1 to perform allocation for other backlogged users.

Figure 4.1 shows the flow chart which summarizes the each specific function from step 1 to step 10 of the proposed allocation algorithm. Each shaded rectangular block represents each function of the steps individually. The algorithm will be activated in each frame once there are free subchannels and backlogged users. Furthermore, the proposed algorithm will continuously perform allocation until there are no free subchannels or no backlogged users.

By using the DPRA algorithm, the resource allocation problem in dimension of time, frequency, and power can be solved efficiently due to the dynamic priority values and the consistent allocation property adopted in the proposed DPRA algorithm. Besides, the consistent allocation will result in lower transmission overhead and efficient resource utilization of the system. For the worst case at each iteration of the DPRA algorithm, we search for an optimal user and subchannel pair from K users and N free subchannels, hence the complexity is $O(KN)$. Since there are totally NL slots in one frame, if we need to do iteration for each slot, the complexity of DPRA algorithm will be $O(KN^2L)$. However, due to the consistent allocation scheme used in DPRA algorithm, the allocation results for a selected user only needs to be determined once in each frame without slot-by-slot iteration. The results for the selected user will last for several OFDMA slots and result in a time burst transmission. This reduces the transmission overhead and complexity greatly so that the complexity is expected to be much smaller than $O(KN^2L)$. Thus the proposed DPRA algorithm with lower cost of complexity and transmission overhead is expected to be acceptable for real system.

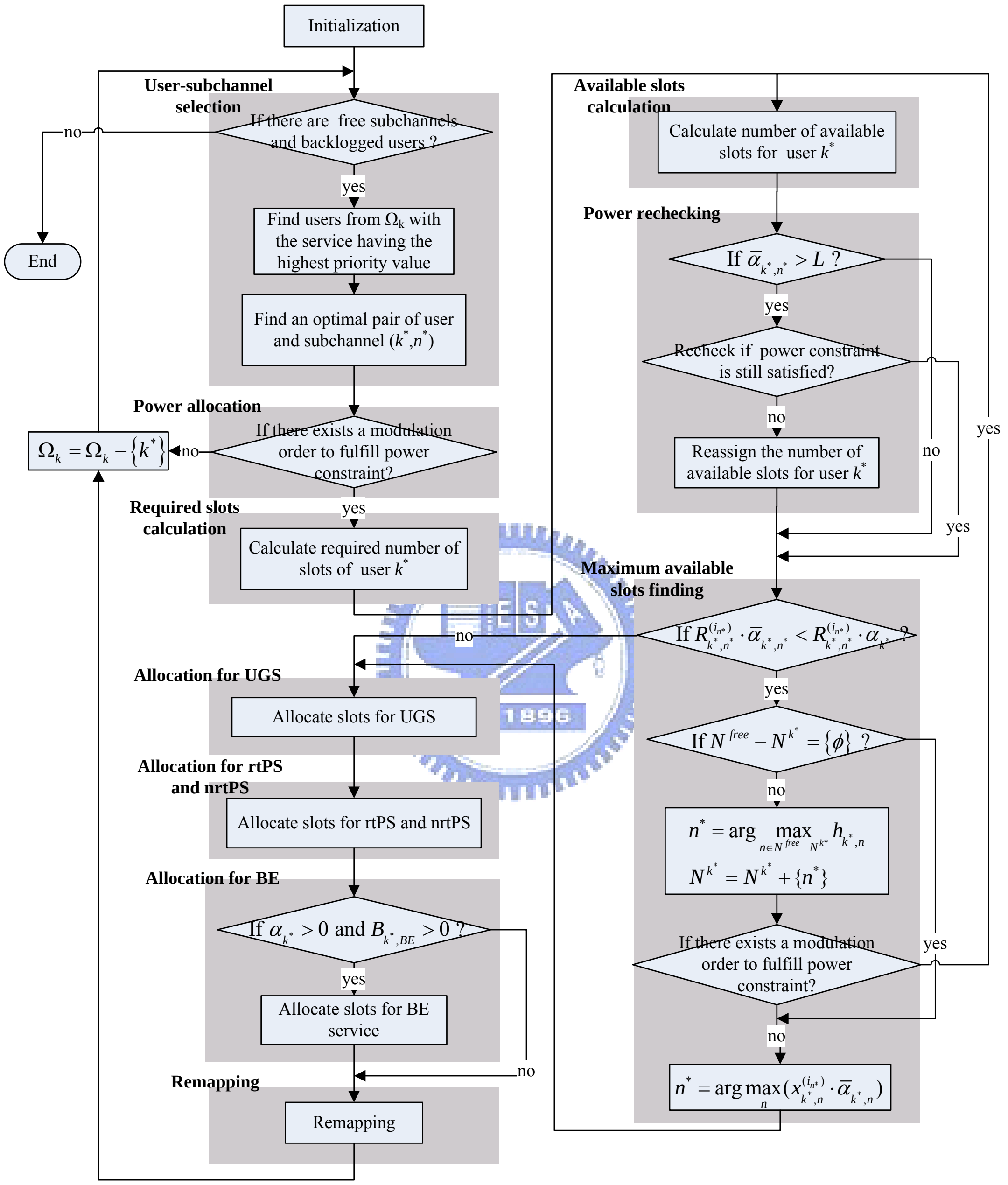

Figure 4.1: Flow Chart of the DPRA Algorithm

Chapter 5

Simulation Results and Discussion

5.1 Simulation Environment

In the simulation, the system level parameters of uplink OFDMA environment are set to be compatible with the IEEE 802.16 standard [1]. The scalable physical layer parameters are configured according to the suggested values in [11] and listed in Table 5.1. The OFDMA system bandwidth is 5 MHZ and each frame duration is 5 ms. The FFT size of 512 is adopted and 384 subcarriers out of the 512 subcarriers are user for data transmission, while the others are used for pilot tones and guard tones.

We consider large scale fading and small scale fading for the wireless fading channels. The large scale fading comes from the signal strength degradation over distance and the shadowing effect, while the small scale fading is due to multipath effect. The path loss is modeled as 128.1 + 37.6log R (dB) [12], where R is the distance between the BS and SS in unit of kilometers. Besides, the shadowing model is assumed log-normal distributed with zero mean and standard deviation of 8 dB, and the fast fading model is assumed to be Rayleigh distribution. There are six taps of multipaths and each of them has independent Rayleigh faded channel model. The power delay profile follows the exponential decay rule. The channel state is assume to be fixed within a frame and varies frame by frame with time according to the fading model stated above.

Table 5.1: System-Level Parameters

Parameters	**Value**
Cell size	1.6 km
Frame duration	5 ms
System bandwidth	5 MHz
FFT size	512
Subcarrier frequency spacing	10.9375 KHz
Number of data subcarriers	384
Number of subchannels	8
Number of data subcarriers per subchannel	48
OFDMA slot duration	102.86 μs
Number of slots for uplink transmission per frame	16
Maximum transmission power for each SS	23 dBm
Thermal noise density	-174 dBm/Hz

5.2 Source Model and QoS Requirements

Four types of traffic are evaluated in our simulation. The first one is the voice traffic for UGS. Each voice traffic is modeled as the ON-OFF model [13]. There is no packet generated during OFF period. During ON period, the voice encoder rate is 8 Kbps and a packet is generated every 20 ms. The size of each packet is 28 bytes including the payload and header. Thus the voice data rate during the ON period is

11.2 Kbps. The parameters associated with the voice traffic are configured according to CISCO VoIP [14] and listed in Table 5.2.

Table 5.2: Parameters of Voice Traffic

Parameters	**Value**
Mean ON time	1 sec
Mean OFF time	1.35 sec
Codec bit rate	8 Kbps
Packets per second	50
Payload size per packet	20 Bytes
Packet size (Payload + header)	28 Bytes
Data rate	11.2 Kbps

The second service type is the video streaming traffic for rtPS. The video streaming consists of a sequence of video frames which are generated regularly with an interval of 100ms. Each video frame is composed of eight slices, which each slice corresponds to a single packet. The size of each packet is truncated Pareto distributed and the inter-arrival time between each packet is also distributed with truncated Pareto distribution. The parameters of video streaming model configured according to [12] are listed in Table 5.3 and the source generation rate is 64 Kbps.

The third traffic type of services is the HTTP traffic for nrtPS, which the behavior of web browsing is modeled according [12]. Thus the model of HTTP traffic is a sequence of page downloads, and each page download is modeled as a sequence of packet arrivals. Each page is composed of a main object and several embedded objects. Also both of the main object and embedded objects can be divided into several packets. The inter-arrival time between each downloaded page, which represents the reading time in web browsing, is distributed with exponential distribution. The parameters associated with video streaming traffic are listed in Table 5.4. Note that the maximum transmission unit of each packet is 1500 bytes.

Table 5.3: Parameters of Video Streaming Traffic

Component	**Distribution**	**Parameters**
Inter-arrival time Between each video frame	Deterministic	100 ms
Number of packets (slices) in each video frame	Deterministic	8
Packet size	Truncated Pareto	Min. = 40 bytes, Max. = 250 bytes Mean = 100 bytes, α =1.2
Inter-arrival time between packets (slices) in a frame	Truncated Pareto	Min. = 2.5 ms, Max. = 12.5 ms Mean = 6 ms, α =1.2

The last type of service is the FTP traffic for BE service. Each FTP traffic is modeled as a sequence of file downloads. The size of each file is truncated lognormal distributed with mean 2 Mbytes, standard deviation 0.722 Mbytes, and a maximum value of 5 Mbytes. The inter-arrival time between each files is exponential distributed with mean 180 seconds. Besides, the QoS requirements of each service type configured in our simulation are listed in Table 5.5.

Table 5.4: Parameters of HTTP Traffic

Component	Distribution	Parameters
Main object size	Truncated Lognormal	Min. = 100 bytes, Max. = 2 Mbytes Mean = 10710 bytes Std. dev. = 25032 bytes
Embedded object size	Truncated Lognormal	Min. = 50 bytes, Max. = 2 Mbytes Mean = 7758 Std. dev. = 126168 bytes
Number of embedded objects per page	Truncated Pareto	Mean = 5.64, Max. = 53
Inter-arrival time between each page	Exponential	Mean = 30 sec
Packet size	Deterministic	Chop from objects with size 1500 bytes
Packet inter-arrival time	Exponential	Mean = 0.13 sec

Table 5.5: QoS Requirements of each service type

Traffic type	**Requirement**	**Value**
Voice (UGS)	Required BER	10^{-3}
	Max. delay tolerance	50 ms
	Max. allowable packet dropping rate	1%
Video (rtPS)	Required BER	10^{-4}
	Max. delay tolerance	10 ms
	Max. allowable packet dropping rate	1%
HTTP (nrtPS)	Required BER	10^{-6}
	Min. required transmission rate	100 Kbps
FTP (BE)	Required BER	10^{-6}

5.3 Performance Evaluation

We compare the DPRA algorithm with other two conventional schemes. The first one is the proportional fair (PF) scheme. The PF scheme is widely user for BS to grant an amount of bandwidth requested by each user. For each user, PF scheme will firstly aggregate the amount data requested by each service type. Then allocate the bandwidth for each user by calculating the proportional fair rate according to the system capacity. The second one is efficient and fair scheduling (EFS) algorithm proposed in [8]. The EFS allocates different subchannels to users one slot at a time. A subchannel is allocated to the user which can transmit maximum amount of data on that subchannel. If all the subchannels are exhausted at current slot, the EFS will move to the next slot. It is an intuitive algorithm but its performance is close to optimal. The EFS will allocate slot for UGS users, then rtPS, nrtPS, and BE service finally. Thus the UGS has highest priority and BE has the lowest priority in each frame.

In the simulation, the number of users is varied from 4 to 40. Each user is assumed to contain all of the four types of services to be transmitted, which are voice, video, HTTP, and FTP traffic. The maximum system transmission rate in a frame is achieved when the highest modulation order is assigned in each slot, which equals to 7.3728 Mbps. We define the traffic load as the ratio of the total average arrival rate of all service types of all users over the maximum system transmission rate. Besides, the average arrival rate of voice, video, HTTP, and FTP is 4.8 Kbps, 64 Kbps, 14.5 Kbps, and 88.9 Kbps, repectively.

The following performance metrics will be measured: (i) system throughput, (ii) packet dropping rate of UGS and rtPS users, (iii) average packet delay of UGS and rtPS users, (iv) average transmission rate of nrtPS and BE users, and (v) ratio of unsatisfied nrtPS users, which is defined as the number of nrtPS users whose average transmission rate is less than the minimum required transmission rate over all nrtPS users.

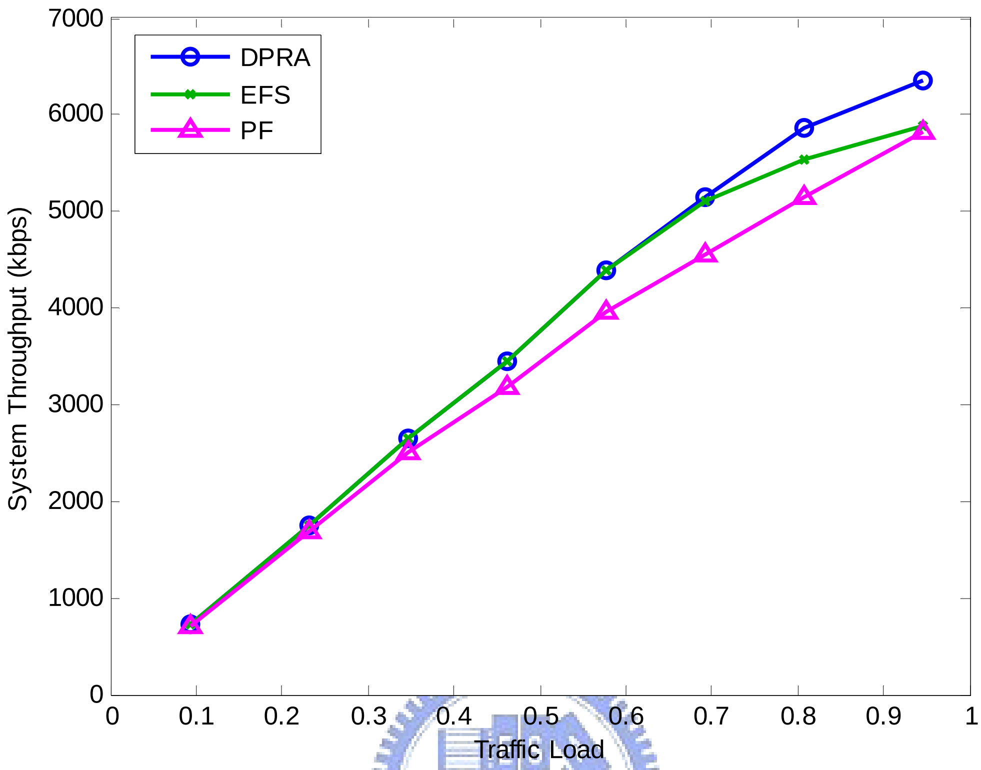

Figure 5.1: System Throughput

Figure 5.1 shows the system throughput versus the traffic load. When the traffic load is low, the system throughput of the DPRA algorithm is close to that of the PF scheme and EFS. However in high traffic load, the DPRA algorithm and EFS are both higher than the PF scheme. Recall that the dynamic priority values reflect the urgency of each service type of users. So the DPRA algorithm gives a higher priority value to the urgent service and allocated it on subchannel with good CSI. It means that the required resource for each service of each SS can be allocated more precisely, so the DPRA algorithm can allocate the radio resource more efficiently. On the other hand, the PF scheme does not consider the QoS requirements, so there are more packets of voice or video service will be dropped and results in throughput degradation. For the EFS, the throughput is almost as large as the DPRA algorithm until the traffic load approaches 1. Since the EFS performs allocation slot by slot, the system resource can also be utilized efficiently. It is shown in fig 5.1 that the proposed DPRA algorithm which performs allocation frame by frame can also reach as high system throughput as EFS algorithm.

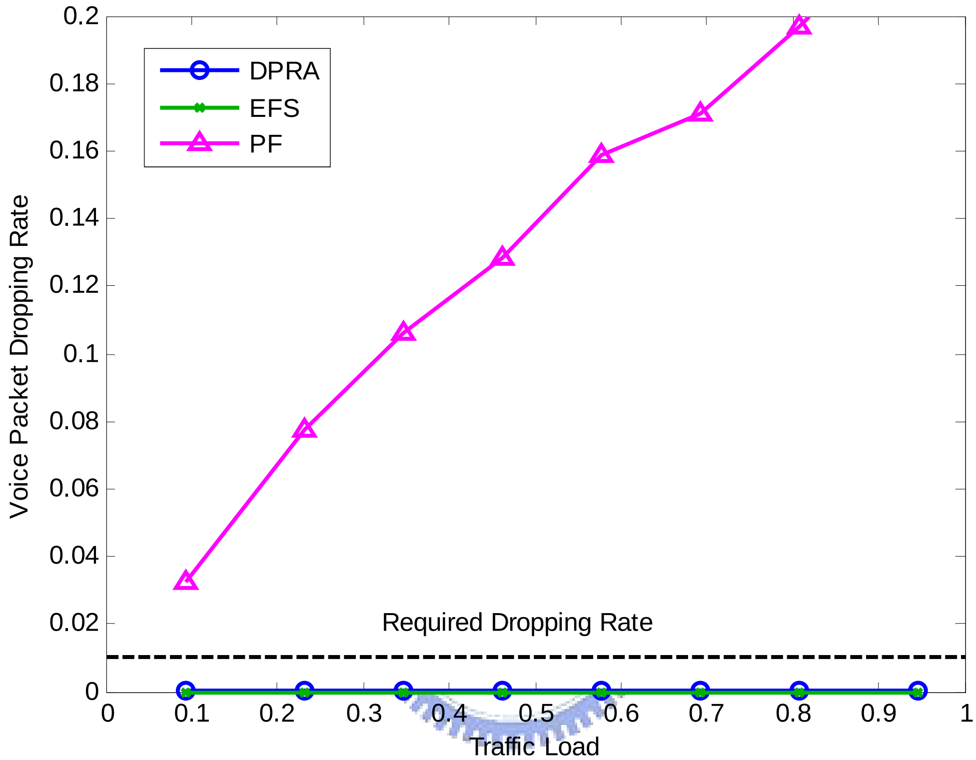

Figure 5.2: Voice Packet Dropping Rate

Figure 5.2 shows the packet dropping rate of voice users. The average dropping rate for voice packet of the DPRA algorithm and EFS is almost zero and below the required dropping rate, which is 0.01. However, the dropping rate of PF scheme increases rapidly with the traffic. Since the voice traffic belongs to UGS, it needs to be guaranteed with constant amount of resource. The PF scheme allocates resource for voice traffic according to the proportional fair rate, while the proposed DPRA algorithm guarantees constant amount of resource for the voice traffic by using the priority values. Besides , the EFS also firstly allocated resource for voice users. Therefore the QoS requirement is fulfilled in the DPRA algorithm and EFS.

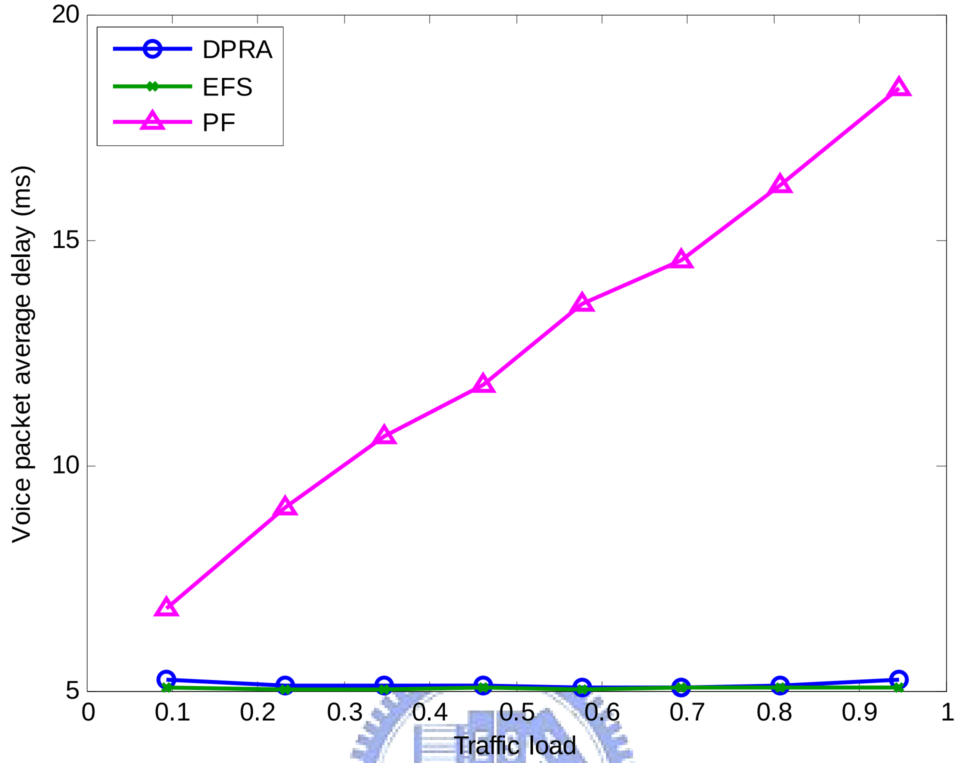

Figure 5.3: Voice Packet Average Delay

Figure 5.3 shows the average packet delay of voice users. The average delay for voice packets of the PF scheme increases with the traffic load since less resource is allocated to each user when there are many backlogged users. Note that the PF scheme always has larger packet delay than the proposed DPRA algorithm, since that most packets require more than one frame to complete transmission. When adopting PF scheme, the allocated resource is usually smaller than the required amount of resource. However, the DPRA algorithm allocates the actual required resource according to the priority value defined to fulfill QoS requirement of voice traffic, so the delay requirement is always guaranteed. Besides, the EFS always allocate resource for voice users firstly, so the voice packets will experience lower transmission delay.

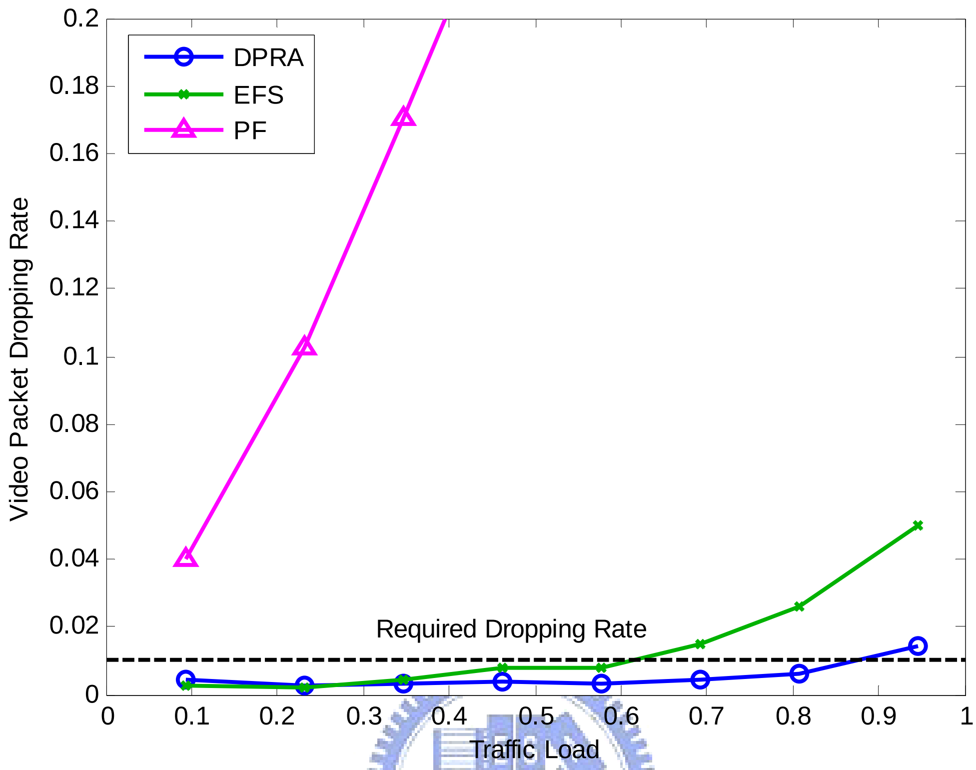

Figure 5.4: Video Packet Dropping Rate

Figure 5.4 shows the packet dropping rate of video users. The dropping rate of EFS keeps small enough to satisfy the required dropping rate until the traffic load is above 0.6. However, the dropping rate of DPRA algorithm keeps below the required dropping rate until the traffic load is above 0.9. The dropping rate of PF scheme increases rapidly with the traffic load. Since the video traffic is variable bit rate, its required resource for satisfying QoS varies in each frame. The proposed dynamic priority value is derived for each frame according to the QoS requirement, and the DPRA algorithm will allocate the required bandwidth for its HOL packet. For EFS algorithm, it will allocate resource for video users after all the voice users of UGS have been served. Resource is easily allocated for video service if the size of voice packets is not large. Therefore the dropping rate of EFS is close to the DPRA algorithm until the traffic load is above 0.6. Due to a more specifically defined priority value of DPRA, the urgency of video packet can be derived exactly by the priority value. The dropping rate is smaller than the EFS and slightly violates the required dropping rate when traffic load exceeds 0.9. Since the PF scheme does not guarantee the required resource, the allocated resource for each user will become smaller when there are many backlogged users. Therefore, DPRA has much lower packet dropping rate than PF even when traffic load is high.

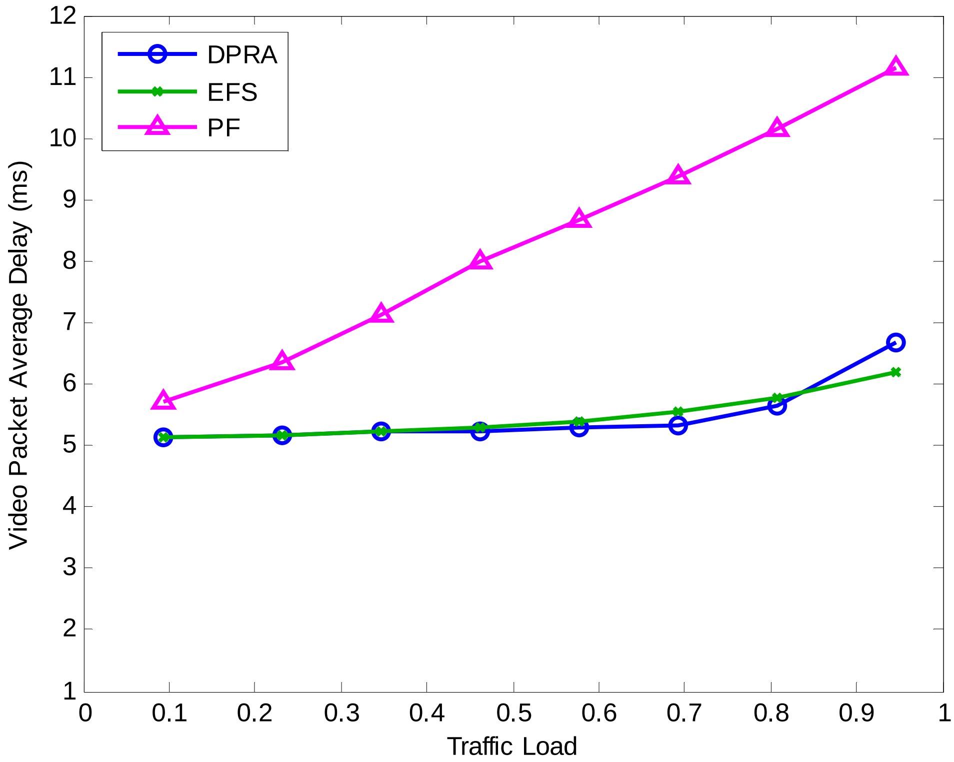

Figure 5.5: Video Packet Average Delay

Figure 5.5 shows the average packet delay of video users. The mean packet delay of DPRA algorithm and EFS increases more slowly than the PF scheme. The mean packet delay of PF scheme is also much higher than the other two schemes. When the traffic load is below 0.7, the average video packet delay for all of the three schemes satisfies the delay requirement, which is 10 ms. Note that when traffic load is above 0.7, the delay requirement of DPRA algorithm and EFS is still fulfilled, but the PF scheme increases rapidly which violates the delay requirement of video packets. Since video streaming is variable bit rate traffic, the required amount of resource for satisfying the delay requirement varies in each frame. The proposed DPRA scheme can allocate resource according to the priority value defined by considering the QoS requirement and transmit the HOL packet of the selected user. Thus the HOL packet will have more chance to be transmitted before exceeding the maximum delay tolerance. Also the EFS will allocate resource for video users prior to other service types after finish allocating voice users. Thus the packet delay of EFS is close to that of the DPRA algorithm. However, the PF only allocates the required amount of resource roughly without considering QoS, so it results in larger delay than our proposed DPRA algorithm.

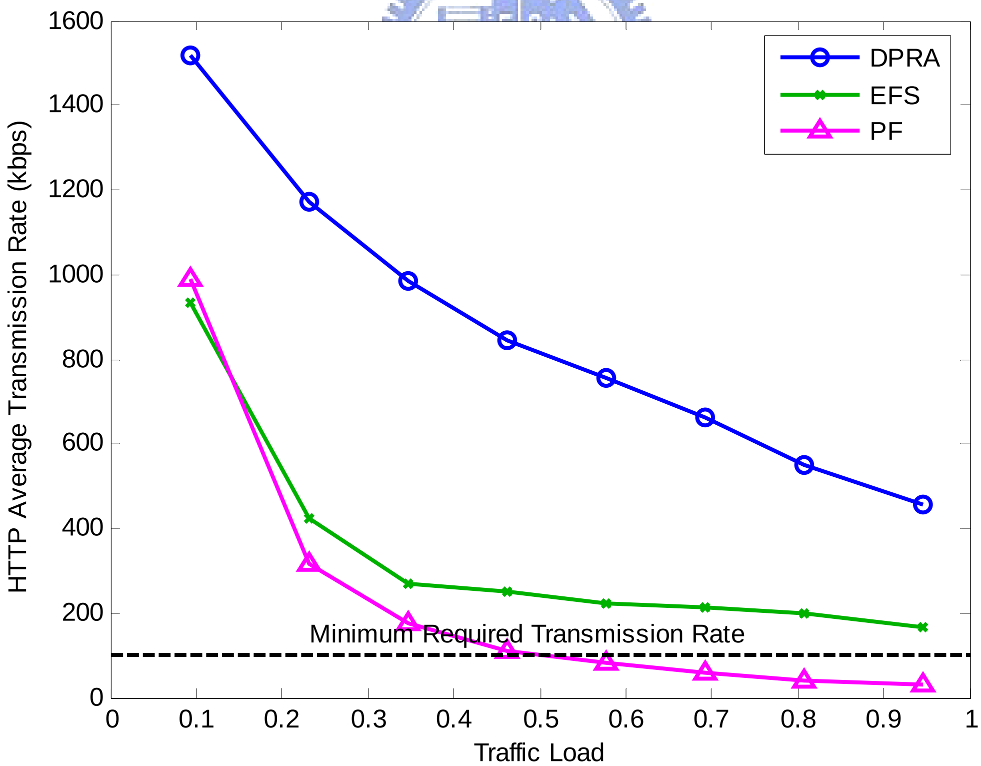

Figure 5.6: HTTP Average Transmission Rate

Figure 5.6 shows the average transmission rate of HTTP users. The average transmission rate for HTTP traffic of DPRA algorithm is always larger than the EFS and PF scheme. Besides, the average transmission rate decreases with the traffic load for all of the three cases. Note that the minimum required transmission rate for DPRA algorithm and EFS can be always guaranteed. However, the average transmission rate of PF scheme is lower than the minimum required transmission rate when traffic load is above 0.5. Since the PF scheme only considers a proportional fair bandwidth allocation without QoS requirement, a small amount of resource will be allocated to each user when traffic load is high and result in lower average transmission rate. For the EFS, the voice and video traffic are always allocated with resource prior to the HTTP traffic, hence the HTTP users have less chance to transmit their data, especially in high traffic load. For DPRA algorithm, high priority value for HTTP service will be given if the average transmission rate is below the required level. Under the circumstance that high priority value is specified, the HTTP users will be allocated more resource to satisfy the QoS requirement. Therefore it is shown in fig 5.6 that the average transmission rate of DPRA is always larger than the proposed EFS algorithm.

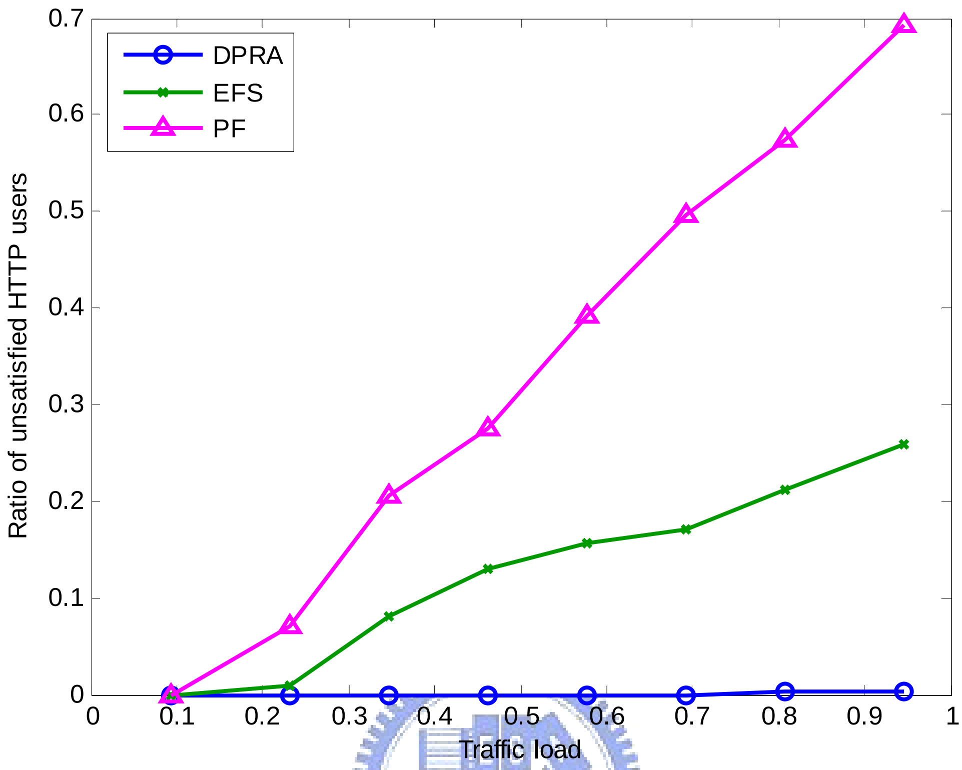

Figure 5.7: Ratio of Unsatisfied HTTP Users

Figure 5.7 shows the ratio of unsatisfied HTTP users. The DPRA algorithm gives high priority value to HTTP users with transmission rate lower than the minimum required transmission rate such that the average transmission rate of all users is guaranteed. Therefore the ratio of unsatisfied HTTP users, whose average transmission rate is below the minimum required transmission rate, keeps close to 0 even when traffic load is high. However, the PF scheme and EFS does not guarantee the minimum transmission rate. Thus the ratio of unsatisfied HTTP users will keep increasing with traffic load due to lack of enough resource allocated for each HTTP user. This result can also be seen from the fact in fig 5.6 that the average transmission of EFS and PF scheme is always less than DPRA and decreases with the traffic load. Therefore, the proposed DPRA algorithm provides acceptable average transmission rate of HTTP traffic and each HTTP user is guaranteed with minimum required transmission rate.

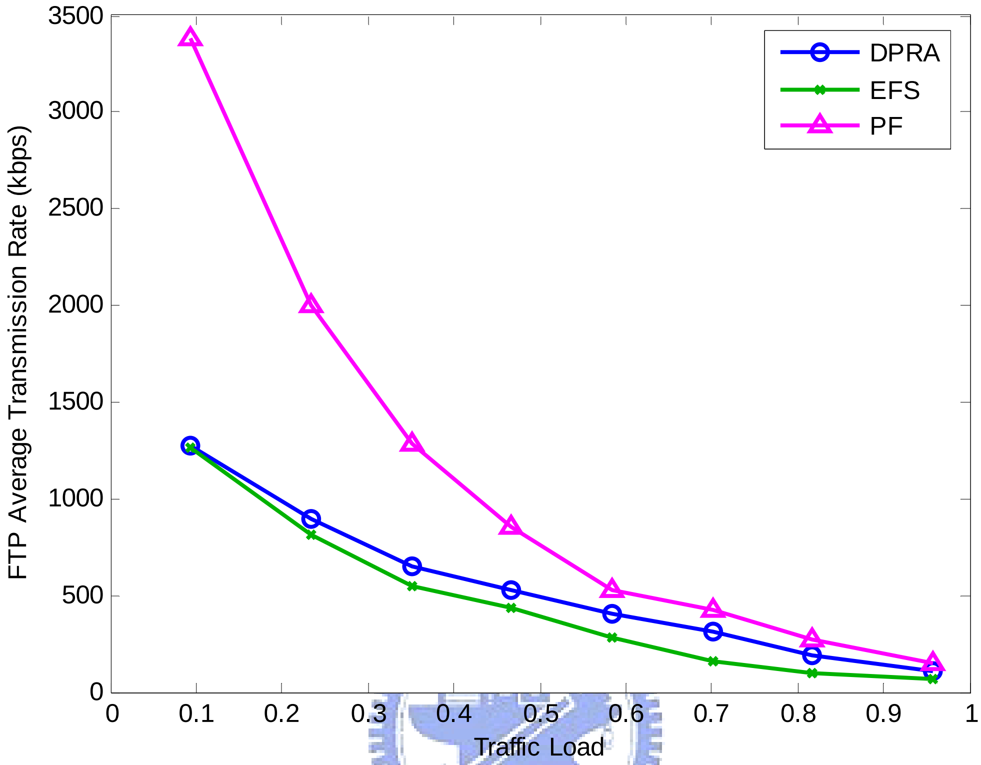

Figure 5.8: FTP Average Transmission Rate

Figure 5.8 shows the average transmission rate of FTP users. The transmission rate of DPRA algorithm for FTP service is lower than PF scheme. It is because the FTP traffic does not have any minimum required transmission rate, and hence the DPRA algorithm specifies it a lowest priority value. The DPRA algorithm will guarantee the QoS requirements for other service class with higher priority. Thus, by exploiting the time diversity, resource will be allocated to FTP traffic when high priority services are served and system resource is available. Therefore it will take longer time to transmit FTP service. For the PF scheme, the FTP service will be given resource according to proportional fair rate in each frame. Thus the average transmission rate will be higher than DPRA algorithm, but the QoS requirements for voice, video, and HTTP will not be fulfilled. For the EFS algorithm, the FTP traffic will be transmitted only when voice, video, and HTTP traffic have already been served. Thus the average transmission rate of EFS algorithm will be slightly lower than the DPRA algorithm. In summery, although the average transmission rate of FTP service for DPRA algorithm is lower than the PF scheme, the DPRA algorithm provides a worthwhile tradeoff since the QoS requirements can be highly satisfied for UGS, rtPS, nrtPS services.

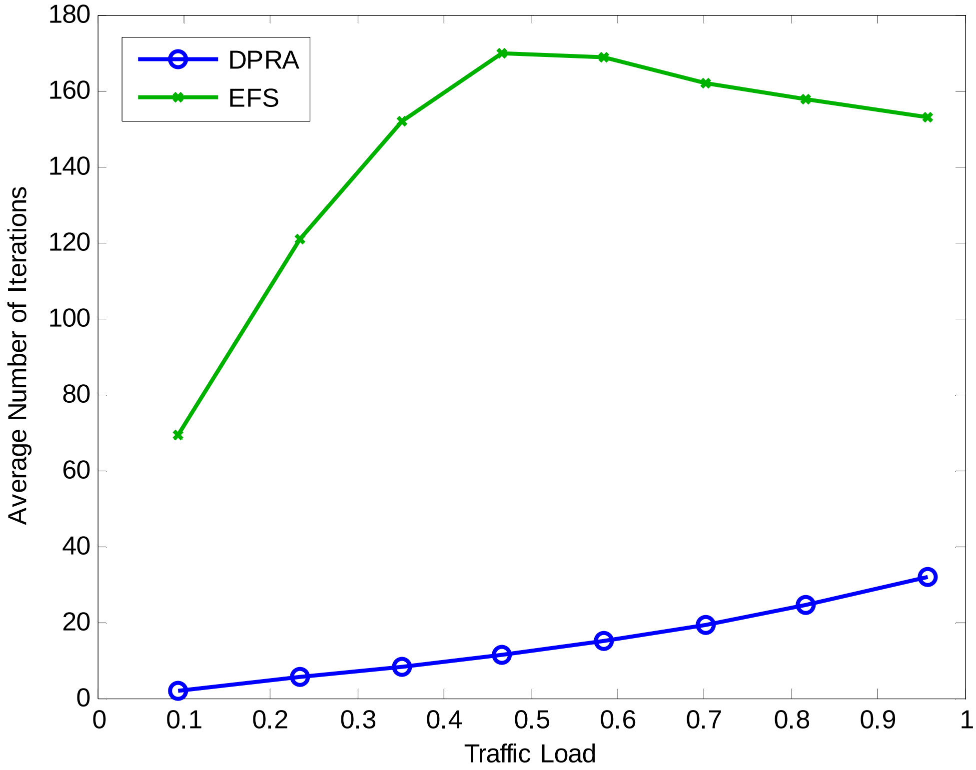

Figure 5.9: Average Number of Iterations

Figure 5.9 shows the average number of iterations per frame of the proposed DPRA algorithm and EFS algorithm. Since the DPRA algorithm performs consistent allocation, it will not need to search for optimal results in each slot for each user. The allocation results for each user only need to be determined once in each frame. However, the EFS algorithm performs slot-by-slot allocation which will search for an optimal pair of subchannel and user for each slot. Hence, when the traffic load increased, average number of iterations of the EFS algorithm will increase faster than the proposed DPRA algorithm. Besides, the average number of iterations of the DPRA algorithm is much less than the EFS algorithm. It is shown in fig. 5.9 and previous figures that the complexity of the proposed DPRA algorithm is reduced and reaches good performance in terms of system throughput and QoS requirements.

It can be summarized from the simulation results that the DPRA algorithm outperforms the PF scheme in terms of system throughput and QoS requirements. Besides, the DPRA algorithm also reaches as good performance as the EFS for UGS and rtPS, and has even better performance for nrtPS. The performance of BE service compared to the EFS is in an acceptable level since no QoS requirement is specified for BE service. As a result, the DPRA algorithm fulfill the QoS requirements for all service types and has lower cost of transmission overhead and complexity than the conventional heuristic algorithms. Hence it is believed that the DPRA will be practical for real system.

Chapter 6

Conclusions

In this thesis, a dynamic priority-based resource allocation (DPRA) algorithm which performs consistent allocation is proposed for IEEE 802.16 uplink system. By considering multimedia services transmission, including UGS, rtPS, nrtPS and BE service, a priority value is derived for each service type according to the urgency of the traffic. The high urgent service will be given a higher priority value. Then the user with service having high priority value will be allocated resource first. Also the proposed DPRA algorithm will dynamically adjust the priority values for each service type of each user frame by frame according to the QoS requirements and buffer conditions. Then subchannel allocation, modulation order assignment, and power allocation are performed by the DPRA algorithm aiming to maximize the system throughput and satisfy QoS requirements.

It is shown in simulation results that the DPRA algorithm can achieve high system throughput while satisfying the QoS requirements of each service type. It can also be noted that the performance of DPRA algorithm is very close to or even better than the conventional heuristic algorithm, which performs allocation slot by slot. Besides, benefited from the consistent allocation, the complexity of DPRA algorithm is expected to be less than the conventional heuristic algorithm. So a traffic burst can be transmitted on sequential slots of a selected subchannel. Thus the DPRA algorithm will greatly reduce the transmission overhead. Therefore we can conclude that the proposed DPRA algorithm reaches throughput maximization and QoS satisfaction at the cost of lower complexity and transmission overhead. It is believed that the DPRA algorithm will be suitable for real systems.

Bibliography

[1] IEEE Std. 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems," Oct. 2004.

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

[3] Y. J. Zhang and K. B. Letaief, "Multiuser Adaptive Subcarrier-and-Bit Allocation with Adaptive Cell Selection for OFDM Systems," *IEEE Trans. Wireless Commun.*, vol. 3, pp.1566-1575, Sep. 2004.

[4] S. Das and G. D. Mandyam, "An Efficient Sub-carrier and Rate Allocation Scheme for M-QAM Modulated Uplink OFDMA Transmission," *Thirty seventh Asilomar Conference on Signals, Systems and Computers,2003,* Vol. 1, pp.136-140, Nov. 2003.

[5] K. Kim, Y. Han, and S. L. Kim, "Joint Subcarrier and Power Allocation in Uplink OFDMA Systems," *IEEE Communication Letters*, vol. 9, No. 6, pp.526-528, June 2005.

[6] J. Jang and K. B. Lee, "Transmit Power Adaptation for Multiuser OFDM Systems," *IEEE J. Sel. Areas Commun.*, vol. 21, pp.171-178, Feb. 2003.

[7] P. Hosein, "Adaptive Subchannel Allocation in an OFDMA-Based Wireless Network," *2006 IFIP International Conference on Wireless and Optical Communication Networks*, pp. 1-5, April 2006

[8] V. Singh and V. Sharma, "Efficient and Fair Scheduling of Uplink and Downlink in IEEE 802.16 OFDMA Networks," *Proc. IEEE Wireless Communication and Networking Conference (WCNC2006)*, vol. 2, pp. 984-990, April 2006.

[9] D. Niyato and E. Hossain, "Queue-Aware Uplink Bandwidth Allocation for Polling Services in 802.16 Broadband Wireless Networks," *Proc. IEEE GLOBECOM'05*, Vol. 6, pp.3702-3706, Dec. 2005.

[10] A. J. Goldsmith and S. G. Chua, "Variable-Rate Variable-Power MQAM for Fading Channels," *IEEE Trans. Commun.*, Vol. 45, pp. 1218-1230, Oct. 1997.

[11] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 WirelessMAN," *Intel Technology Journal*, Volume 8, Issue 3, 2004.

[12] 3GPP TR 25.892, "Feasibility study for OFDM for UTRAN enhancement," 3rd Generation Partnership Project, Tech. Rep., 2004-06.

[13] Universal Mobile Telecommunication System, Selection procedures for the choice of radio transmission technologies of the UMTS, UMTS Std. 30.03, 1998.

[14] CISCO Tech Notes, "Voice Over IP – Per Call Bandwidth Consumption," Document ID 7934.

Vita

Jian-Ann Lai was born in Taipei, Taiwan. He received B.E. degree in Department of Electrical and Control Engineering and M.E. degree in Department of Communication Engineering from National Chiao-Tung University, Hsinchu, Taiwan, in 2005 and 2007, respectively. His research interests include radio resource management and wireless communication systems.