Dynamic Priority Resource Allocation for Uplinks in IEEE 802.16 Wireless Communication Systems

Chih-Ming Yen, *Student Member, IEEE*, Chung-Ju Chang, *Fellow, IEEE*,
Fang-Ching Ren, *Member, IEEE*, and Jian-Ann Lai

***Abstract*—In this paper, a dynamic priority resource-allocation (DPRA) scheme is proposed for uplinks in IEEE 802.16 wireless communication systems. The DPRA scheme dynamically gives priority values to four types of service traffic based on their urgency degrees and allocates system radio resources according to their priority values. It can maximize the system throughput and satisfy differentiated quality-of-service (QoS) requirements. Furthermore, the DPRA scheme performs consistent allocation for packets of users to conform to the uplink frame structure of IEEE 802.16, to fulfill the QoS requirement, and to reduce the computational complexity. Simulation results show that the proposed DPRA scheme performs very close to the optimal method, which is by exhaustive search in system throughput, and it outperforms the conventional efficient and fair scheduling (EFS) algorithm in the performance measures such as system throughput, real-time polling service (rtPS) packet dropping rate, ratio of unsatisfied non-real-time polling service (nrtPS), and average transmission rate of the best effort (BE) service. In addition, the DPRA scheme takes only 1/1000 and 1/10 the computational times of the optimal method and the conventional EFS algorithm, respectively, thus making it more feasible for real applications.**

***Index Terms*—IEEE 802.16, priority, quality of service (QoS), radio resource allocation, uplink.**

I. INTRODUCTION

ORTHOGONAL frequency-division multiplexing has been proposed as a promising technique for future multimedia wireless communication systems due to its ability to mitigate frequency-selective fading and intersymbol interference (ISI) and its flexibility for adaptive modulation on each subcarrier. Orthogonal frequency-division multiple access (OFDMA) has been adopted for IEEE 802.16 broadband wireless access systems. Although the medium access control (MAC) signaling has been well defined in the IEEE 802.16 specifications [1], resource management and scheduling still remain as open issues. Since the wireless channel condition varies with time, adaptive resource allocation has been viewed as one of the key technologies to provide efficient utilization of the limited system resource in multiuser wireless communication systems. Furthermore, for a system with multimedia traffic provisioning, diverse quality of service (QoS) requirements should be taken into account when developing an efficient resource-allocation algorithm. Therefore, an effective resource-allocation scheme is required to exploit frequency diversity, multiuser diversity, time diversity, and QoS requirement diversity so that the overall system resource can efficiently be utilized and the QoS requirement can be guaranteed.

Subcarrier, bit, and power-allocation algorithms for multiuser OFDMA systems to maximize the overall data rate or minimize the total transmitted power under some constraints have been studied in much of the literature. Wong *et al.* [2] proposed a Lagrangian-based algorithm for minimizing the total transmission power consumption under the user's QoS requirements, which were defined by a specified data transmission rate and bit error rate (BER). However, a high computational complexity renders it impractical. To reduce the complexity, Zhang and Letaief [3] proposed a near-optimum dynamic multiuser subcarrier-and-bit allocation algorithm to maximize the overall spectral efficiency.

Many papers considered the downlink resource allocation [3]–[5], but a few papers investigated the uplink resource allocation. Resource allocation of both downlink and uplink is primarily performed by the base station (BS). Das and Mandyam [6] considered the uplink transmission of the OFDMA system and developed an efficient algorithm for the subcarrier and bit allocation of each user. The algorithm includes the power distribution over the selected set of subcarriers for every user so that the total used power is minimized. Kim *et al.* proposed a joint subcarrier and power-allocation scheme for uplink OFDMA systems to maximize the rate-sum capacity based on the Shannon capacity formula [7], where a greedy subcarrier allocation algorithm, based on a marginal rate function, and an iterative water-filling power-allocation algorithm were proposed. The scheme was shown to achieve a near-optimal solution. Jang and Lee [8] concluded that the equal power-allocation algorithm over assigned subcarriers for each user can achieve similar performance to the water-filling scheme. Hosein [9] assumed that subchannels made up of a group of contiguous subcarriers are assigned to users in unit of time slots. Furthermore, the channel state information (CSI) on the subchannels of each subscriber station (SS) is assumed to be periodically reported. 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 [10] also

Manuscript received April 28, 2008; revised August 29, 2008 and January 16, 2009. First published April 7, 2009; current version published October 2, 2009. This work was supported by the National Science Council of Taiwan under Contract NSC 97-2221-E-009-098-MY3. The review of this paper was coordinated by Dr. H. Jiang.

C.-M. Yen, C.-J. Chang, and J.-A. Lai are with the Department of Communication Engineering, College of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: cjchang@mail.nctu.edu.tw).

F.-C. Ren is with the Industrial Technology Research Institute, Hsinchu 310, Taiwan (e-mail: frank_ren@itri.org.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2009.2020328

0018-9545/$26.00 © 2009 IEEE

developed an efficient and fair scheduling (EFS) algorithm for each time slot in IEEE 802.16 OFDMA/time-division duplex systems. The EFS algorithm is designed with a fixed priority scheme that gives priorities to service traffic according to their QoS requirements. Chen and Chang proposed a dynamic uplink channel-allocation strategy for selecting a better channel for each SS, depending on the SS's signal-to-noise ratio value [11]. However, the QoS requirements and power constraint are not considered. Furthermore, an efficient uplink resource allocation for power saving in IEEE 802.16 OFDMA systems was proposed in [12]. It adaptively adjusts the modulation and coding scheme to minimize the required transmission power while guaranteeing the BER. However, multiple services and their differentiated QoS requirements are not taken in account.

In the aforementioned previous works, the QoS requirements and fairness issues are either omitted or simplified. A minimum required transmission data rate or a predefined weight that corresponds to the fixed priority scheme is usually adopted as the QoS requirements. However, with the provision of multimedia real-time traffic, the delay bound and the packet dropping rate, which is regarded as essential QoS requirements, should be included in the design of radio resource allocation for practical applications. Additionally, realistic traffic models and buffer conditions of different traffic types should be considered. Niyato and Hossain proposed a queue-aware uplink bandwidth allocation scheme for SSs with real-time polling service (rtPS) and non-real-time polling service (nrtPS) in IEEE 802.16 systems [13]. The bandwidth is allocated according to the channel quality and queue state of the traffic.

Furthermore, the tradeoff between system performance and computational complexity is an important issue. The greedy algorithm that performs symbol-by-symbol allocation can achieve an optimal solution [14], but it results in high computational complexity. According to the frame structures of DL-MAP and UL-MAP defined in IEEE 802.16 for downlink and uplink, respectively, the symbol-by-symbol allocation algorithm costs high transmission overhead. In addition, 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 [1]. Thus, an efficient and feasible resource allocation algorithm for either downlink or uplink needs to be specifically designed to meet its individual frame structures.

In this paper, we propose a dynamic priority resource allocation (DPRA) scheme for IEEE 802.16 uplink communication systems. The goal of the proposed DPRA scheme is to maximize the system throughput while satisfying various QoS requirements of multimedia traffic. Four types of service traffic for users are taken into account, including unsolicited grant service (UGS), rtPS, nrtPS, and best effort (BE) service. A priority value for every service type of each user is defined and adaptively adjusted frame by frame according to its urgency related with individual QoS requirements and buffer condition. Then, the BS will dynamically allocate the uplink subchannel, modulation order, and power to each SS according to its priority value and the CSI. Furthermore, to meet the uplink frame structures defined in IEEE 802.16 specifications and reduce the computational complexity and the transmission overhead, a consistent allocation mechanism is designed in the proposed DPRA scheme. Simulation results show that the proposed DPRA scheme performs close to the optimal method, which is by exhaustive search, in system throughput. Furthermore, it outperforms the EFS conventional algorithm [10] in system throughput and rtPS packet dropping rate. In addition, the DPRA scheme can take much less computational complexity than the optimal method and the EFS algorithm, where the DPRA scheme is just 1/1000 of the optimal method and 1/10 of the EFS algorithm.

This paper is organized as follows. The system model of the considered uplink OFDMA system is introduced in Section II. Section III presents the details of the proposed DPRA scheme. Section IV discusses the performance of the DPRA scheme, compared to the EFS [10]. Finally, conclusions are given in Section V. Two mechanisms called listen before talk (LBT) and extended quiet period (EQP) were proposed in IEEE 802.16h [4] such that IEEE 802.16 systems can coexist with existing systems in an unlicensed band. An active period (AP) is set as the holding time for secondary users. In this paper, we try to adjust AP by an effective and intelligent way for performance improvement by using fuzzy logic theory. Fuzzy logic has broadly been used in the up-to-date communication systems to mimic the operation of the human brain to solve problems with vague information [3]. It has the capabilities of soft computing and adaptation to cope with the network control problems.

II. SYSTEM MODEL

Suppose that there are N subchannels in the uplink of the IEEE 802.16 OFDMA system and that each subchannel consists of q adjacent subcarriers. There are K SSs going to communicate with one BS in one cell. Each SS can be viewed as a single user containing different service types of traffic to transmit, and each service type in an SS has its individual queue. Furthermore, based on the IEEE 802.16 uplink rectangle frame structure, traffic data are transmitted in fixed length of frames, and each frame contains L OFDMA slots. Then, the total number of resource units in each frame is $L \times N$ slots, which is in sequence from the leftmost of the top subchannel to the rightmost of the bottom subchannel.

IEEE 802.16 defines the following four service types, and each of them has different QoS requirements.

1) *UGS:* The UGS supports real-time traffic that periodically generates fixed-size data packets. Thus, the BS generally allocates a fixed amount of bandwidth for this type of service.
2) *rtPS:* It is designed to support real-time service, which generates variable-size data packets. It is a delay-sensitive traffic so that the delay requirement is an important QoS issue. The amount of bandwidth granted for this type of service needs to be dynamically determined according to its priority based on the QoS requirements and traffic models.
3) *nrtPS:* It is designed to support delay-tolerant data streams while a minimum data transmission rate is required. Furthermore, the bandwidth granted for nrtPS

needs to be dynamically determined according to its priority based on the QoS requirement and the buffer condition.

4) BE: The BE service is designed to support data streams that 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 is allocated for the BE service.

The priority value of service type $s(s \in \{\text{UGS}, \text{rtPS}, \text{nrtPS}, \text{BE}\})$ for user k, which is denoted by $\gamma_{k,s}$, is defined here in terms of the minimum number of bits required to transmit per frame. $\gamma_{k,\text{UGS}}$ remains constant in each frame since the system needs to grant a constant amount of bandwidth for UGS. $\gamma_{k,\text{rtPS}}$ and $\gamma_{k,\text{nrtPS}}$ are dynamically adjusted frame by frame so that the QoS requirements can be satisfied and the radio resource will be efficiently utilized. $\gamma_{k,\text{BE}}$ is set to be $0 \le \gamma_{k,\text{BE}} \le \gamma_{k,s}$, $s \in \{\text{UGS}, \text{rtPS}, \text{nrtPS}\}$, since there is no delay or transmission rate requirement. Usually, $\gamma_{k,\text{BE}}$ is the smallest, and $\gamma_{k,\text{UGS}}$ is the largest but not necessarily.

For rtPS, denote D_k^* as the maximum delay tolerance of user k with an rtPS head-of-line (HOL) packet and D_k as the current delay of the rtPS HOL packet of user k experienced, which is the time duration from the arrival frame of the packet to the present frame. Both D_k^* and D_k are in unit of frames. The remaining time for the rtPS HOL packet of user k before being dropped, which is denoted by ΔD_k, is given by

$$\Delta D_k \equiv D_k^* - D_k. \tag{1}$$

If $\Delta D_k \le 0$, the packet will be dropped and not considered. Therefore, $\gamma_{k,\text{rtPS}}$ is defined as

$$\gamma_{k,\text{rtPS}} = \begin{cases} B_{k,\text{rtPS}}, & \text{if } 1 \le \Delta D_k \le D_{\text{th}} \\ \frac{B_{k,\text{rtPS}}}{\Delta D_k + \log(\Delta D_k)}, & \text{if } D_{\text{th}} < \Delta D_k \end{cases} \tag{2}$$

where $B_{k,\text{rtPS}}$ is the number of residual bits of the rtPS HOL packet buffered at the queue of user k, and D_{th} is a predefined delay threshold for warning to guarantee QoS requirements. If ΔD_k is smaller than or equal to the threshold D_{th}, it means that the rtPS HOL packet of user k is very urgent and that all of the residual bits in the buffer had better finish transmission in the current frame. Otherwise, the priority value $\gamma_{k,\text{rtPS}}$ can be set lower based on the average transmission rate $B_{k,\text{rtPS}}/\Delta D_k$. We further add to its denominator a bias of $\log(\Delta D_k)$ to lessen its served transmission bits and make room for other possible high-priority users since its residual time before QoS violation is still long. Note that the larger D_{th} is, the earlier the warning and the better the QoS satisfaction of RT services will be. However, in this situation, the system will reserve or consume more resource to protect these RT services, and the system throughput will be reduced. Therefore, D_{th} should be properly set.

For nrtPS, the average transmission rate should be larger than the requirement of the minimum transmission rate, which is denoted by $R_{k,\text{nrtPS}}^*$. Denote $B_{k,\text{nrtPS}}$ as the number of residual bits of the user k's nrtPS HOL packet buffered at the current frame and ΔT_k as the maximum number of frames left for the nrtPS HOL packet of user k at the current frame so that the requirement $R_{k,\text{nrtPS}}^*$ can be fulfilled. Similarly, $\gamma_{k,\text{rtPS}}$ is designed as

$$\gamma_{k,\text{nrtPS}} = \begin{cases} a \cdot B_{k,\text{nrtPS}}, & \text{if } \Delta T_k \le T_{\text{th}} \\ \frac{a \cdot B_{k,\text{nrtPS}}}{\Delta T_k + \log(\Delta T_k)}, & \text{if } \Delta T_k > T_{\text{th}} \end{cases} \tag{3}$$

where T_{th} is a predefined threshold for nrtPS to make an obvious priority distinction, $\log(\Delta T_k)$ is a bias by the same concept as that for rtPS, and a is a weighting constant $0 < a \le 1$, which is used to depress the priority of nrtPS traffic as compared to that of rtPS traffic. Similar to D_{th}, T_{th} should properly be determined.

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

$$s_{k,n}^{(\ell)} = \sqrt{\rho_{k,n}^{(\ell)}} \cdot d_{k,n}^{(\ell)}, \quad 1 \le k \le K; \quad 1 \le n \le N \tag{4}$$

where $\rho_{k,n}^{(\ell)}$ is the power allocated to 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-quadratic-amplitude modulation (QAM) 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. Hence, the CSI is assumed to remain constant over one frame. In addition, perfect estimation of CSI on each subchannel of each user is assumed in this paper. Since in the IEEE 802.16 uplink system, the SSs only report the uplink CSI on each subchannel, the channel gain of each adjacent subcarrier that a subchannel contains is assumed to be the same. Let $h_{k,n}$ be the uplink channel gain between user k and the considered BS on subchannel n. Note that the channel gain is not a function of slot time ℓ since it remains fixed during one time frame. The received signal of user k on subchannel n at the ℓth OFDMA slot, which is denoted by $y_{k,n}^{(\ell)}$, is given by

$$y_{k,n}^{(\ell)} = h_{k,n}\sqrt{\rho_{k,n}^{(\ell)}} d_{k,n}^{(\ell)} + \sum_{k' \in K'} h_{k',n}\sqrt{\rho_{k',n}^{(\ell)}} d_{k',n}^{(\ell)} + z_{k,n}^{(\ell)} \tag{5}$$

where K' is the set of users that use the same subchannel n at the ℓth OFDMA slot in other cells, and $z_{k,n}^{(\ell)}$ is the complex white Gaussian noise of user k on subchannel n with zero mean and variance σ^2. The second term at the right-hand side of (5) is the cochannel interference from other cells. Therefore, the received signal-to-interference-plus-noise ratio (SINR) of user k on subchannel n at the ℓth OFDMA slot, which is denoted by $\text{SINR}_{k,n}^{(\ell)}$, can be obtained as [15]

$$\text{SINR}_{k,n}^{(\ell)} = \frac{\rho_{k,n}^{(\ell)} |h_{k,n}|^2}{\sum_{k' \in K'} \rho_{k',n}^{(\ell)} |h_{k',n}|^2 + \sigma^2}. \tag{6}$$

An approximated BER when using M-QAM has been given by [16]

$$\text{BER} \approx 0.2 e^{-1.5 \frac{\text{SINR}}{M-1}}. \tag{7}$$

From (7), the minimum required SINR of user k, which is denoted by SINR_k^*, can be obtained by

$$\mathrm{SINR}_k^* = -\frac{\ln\left(5\mathrm{BER}_k^*\right)}{1.5}(M-1). \tag{8}$$

Therefore, based on the BER_k^* of user k, the minimum power allocated to user k on each subcarrier of subchannel n can be obtained by

$$\rho_{k,n}^{(\ell)} = \frac{-\ln\left(5\mathrm{BER}_k^*\right)}{1.5|h_{k,n}|^2}(M-1)\sigma^2. \tag{9}$$

In addition, the allocated power on each subcarrier of subchannel n will be equally distributed. Thus, the total allocated power to user k on subchannel n, which contains q subcarriers, at the ℓth OFDMA symbol, which is denoted by $p_{k,n}^{(\ell)}$, can be obtained by

$$p_{k,n}^{(\ell)} = q \cdot \rho_{k,n}^{(\ell)}. \tag{10}$$

III. DYNAMIC PRIORITY RESOURCE ALLOCATION SCHEME

A. Problem Formulation

The goal of the proposed DPRA scheme is to maximize the overall system throughput and satisfy QoS requirements. Here, the allocation problem is first mathematically formulated into optimization equations. Let $x_{k,n}^{(\ell)}$ be the assignment variable indicating the number of bits carried by each subcarrier of subchannel n with modulation order assigned to user k at the ℓth OFDMA slot. $x_{k,n}^{(\ell)}$, with $1 \le k \le K$, $1 \le n \le N$, and $1 \le \ell \le L$ is given by

$$x_{k,n}^{(\ell)} = \begin{cases} 6, & \text{if 64-QAM} \\ 4, & \text{if 16-QAM} \\ 2, & \text{if QPSK modulation} \\ 0, & \text{if not assigned.} \end{cases} \tag{11}$$

The assignment vector for the ℓth OFDMA slot over K users and N subchannels, which is denoted by $\mathbf{x}^{(\ell)}$, is defined as

$$\mathbf{x}^{(\ell)} \equiv \left(x_{1,1}^{(\ell)}, \ldots, x_{1,N}^{(\ell)}, \ldots, x_{k,1}^{(\ell)}, \ldots x_{k,n}^{(\ell)} \right.$$
$$\left. \ldots, x_{k,N}^{(\ell)}, \ldots, x_{K,1}^{(\ell)}, \ldots, x_{K,N}^{(\ell)}\right)^T \tag{12}$$

and the assignment vector over the frame, which is denoted by $\mathbf{x}$, can be obtained by

$$\mathbf{x} = \left[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(L)}\right]. \tag{13}$$

In such a case, the total number of bits allocated to user k in one frame, which is denoted by R_k, can be calculated by

$$R_k \equiv R_k(\mathbf{x}) = \sum_{\ell=1}^{L}\sum_{n=1}^{N} q \cdot x_{k,n}^{(\ell)}. \tag{14}$$

The optimization equations for the proposed DPRA scheme is then formulated as follows.

- Objective

$$\text{Throughput maximization: } \mathbf{x}^* = \arg\max_{\mathbf{x}} \sum_{k=1}^{K} R_k(\mathbf{x}). \tag{15}$$

- Allocation Strategy
  i) First, allocate the user with the largest priority value to fulfill QoS requirements.
  ii) Consistently allocate the chosen user to conform the UL-MAP structure.
- Allocation Constraints

i) Power constraint:

$$\sum_{n=1}^{N} p_{k,n}^{(\ell)} \le p_{k,\max} \quad \forall \ell \text{ and } k. \tag{16}$$

ii) Buffer constraint:

$$R_k \le B_k \quad \forall k. \tag{17}$$

iii) Slot allocation constraint:

$$\sum_{k=1}^{K} \mathrm{sgn}\left(x_{k,n}^{(\ell)}\right) = 1 \quad \forall n \text{ and } \ell. \tag{18}$$

To fulfill the QoS requirements, allocation strategy i), which is the QoS fulfillment strategy, is to first serve the user with the highest priority value. Due to the UL-MAP structure, allocation strategy ii), which is the consistent allocation strategy, is to continuously allocate slots to the user in the same subchannel. If the allocated slot is to the end slot of the subchannel, it will continuously be allocated from the first slot of the next subchannel. On the other hand, $p_{k,\max}$ in (16) is the maximum allowable uplink transmission power of user k. For each OFDMA slot, the total transmission power of each SS should have a cap. B_k in (17) is the total number of residual bits in the buffer of all service types of user k. To not waste the system resource, the allocated bit to user k in each frame should not be larger than the total buffer occupancy. To avoid the cochannel interference, (18) indicates that each OFDMA slot can only be allocated to one user. It is a basic constraint for a single-antenna system.

B. Dynamic Priority Resource Allocation Scheme

The proposed DPRA scheme performs subchannel selection, modulation order, and power-allocation assignment for uplink users. The DPRA scheme is also designed with a *consistent allocation mechanism*. Once a subchannel with a modulation order and power is assigned to a selected user at a certain slot, the next consecutive slots at the same subchannel will consistently be given to the selected user until its required transmission of bits completes. Notice that the allocation for each user in a frame by the DPRA scheme takes execution only one time rather than symbol by symbol [7]. Consequently, the

DPRA scheme can not only meet the uplink frame structure defined in IEEE 802.16 [1] but also reduce the computational complexity and fulfill the QoS requirement.

The DPRA scheme is a heuristic algorithm that contains six steps of functions to solve the optimization problem given in (15)–(18). At the beginning, the assignment vector $\mathbf{x}^{(\ell)}$ and allocated power $p_{k,n}$ are initialized to be zero, which means that all resources are free. Denote ℓ_n as the ℓth slot of subchannel n that has been allocated and δ_k as the total number of slots allocated to user k. They are initialized to be zero too. Denote N^{free} as the set of free subchannels of the system and N^k as the set of subchannels allocated to user k. Initially, $N^{\text{free}} = \{n | 1 \leq n \leq N\}$, and $N^k = \{\phi\}$, $\forall k$. Furthermore, let Ω be the set of backlogged users whose buffers are not empty. The six steps of functions are described as follows.

Step 1—User–Subchannel Selection: The DPRA scheme selects users from the set of backlogged users Ω according to the priority values. Let Ω_h be the set of backlogged users having the service with the highest priority value, which is denoted by $\gamma_{\max}$. To achieve the goal of system throughput maximization, the DPRA scheme selects the subchannel with the best CSI in free subchannel set N^{free} for the acceptable highest modulation order. In other words, from Ω_h and N^{free}, an optimal pair of user and subchannel (k^*, n^*) would be chosen to maximize the system throughput and fulfill the QoS requirement. The function is shown below.

```
**Function**: *User–Subchannel Selection*
    $\gamma_{\max} = \max_{k \in \Omega}(\gamma_{k,s}), \forall s \in \{\text{UGS}, \text{rtPS}, \text{nrtPS}\}$
    $\Omega_h = \{k | k = \arg\max_{k \in \Omega}(\gamma_{k,s})$
            $\forall s \in \{\text{UGS}, \text{rtPS}, \text{nrtPS}\}\}$
    $(k^*, n^*) = \arg\max_{k \in \Omega_h, n \in N^{\text{free}}}(h_{k,n})$
    $N^{k^*} = N^{k^*} + \{n^*\}$                                ■
```

Step 2—Highest Modulation Order Assignment: Once an optimal pair of user and subchannel (k^*, n^*) is selected, the highest modulation order assignment and its associated power allocation given in (10) for user k^* on subchannel n^*, or the largest $x_{k^*,n^*}^{(\ell_{n^*})}$, are performed. Note that the power allocation gets its maximum power constraint, which is denoted by $p_{k^*,\max}$. The selected user k^* will be removed from the set of backlogged users if even QPSK cannot be assigned for power constraint violation. The function is given below.

```
**Function**: *Highest Modulation Order Assignment*
    while $p_{k^*,n^*}(\text{BER}_{k^*}^*, x_{k^*,n^*}^{(\ell_{n^*})} + 2) < p_{k^*,\max}$
        $x_{k^*,n^*}^{(\ell_{n^*})} = x_{k^*,n^*}^{(\ell_{n^*})} + 2$
    end while
    *if* $x_{k^*,n^*}^{(\ell_{n^*})} = 0$
        *then* $\Omega = \Omega - \{k^*\}$, $N^{k^*} = N^{k^*} - \{n^*\}$, and go to step 1.   ■
```

Step 3—Allocation Slot Calculation: The number of bits that user k^* can transmit on subchannel n^* in the first assigned slot is given by $x_{k^*,n^*}^{(\ell_{n^*})} \cdot q$ since each subchannel contains q subcarriers. Let α_{k^*} be the number of slots required for user k^* to transmit residual bits of HOL packets of all service types and $\bar{\alpha}_{k^*,n^*}$ be the number of allocation slots that the system can assign to user k^* when the allocation starts in subchannel n^*. Note that if $\gamma_{\max} > \gamma_{k,\text{BE}}$, this means that there are some more urgent services in user k^* and that its BE data will not be considered for allocation this time. Because slots $(1 \sim \ell_{n^*} - 1)$ of subchannel n^* have already been allocated to other users before, if the remaining slots $(L - \ell_{n^*} + 1)$ in subchannel n^* that can be allocated to user k^* is not sufficient, it is assumed that the available slots of the two subchannels next to subchannel n^* will be considered for allocation to user k^* due to the proposed consistent allocation mechanism.

```
**Function**: *Allocation Slot Calculation*
*if* $\gamma_{\max} > \gamma_{k,\text{BE}}$
    $$\alpha_{k^*} = \left\lceil \frac{\gamma_{k^*,\text{UGS}} + B_{k^*,\text{rtPS}} + B_{k^*,\text{nrtPS}}}{x_{k^*,n^*}^{(\ell_{n^*})} \cdot q} \right\rceil$$
*else*
    $$\alpha_{k^*} = \left\lceil \frac{\gamma_{k^*,\text{UGS}} + B_{k^*,\text{rtPS}} + B_{k^*,\text{nrtPS}} + B_{k^*,\text{BE}}}{x_{k^*,n^*}^{(\ell_{n^*})} \cdot q} \right\rceil$$
$\bar{\alpha}_{k^*,n^*} = \alpha_{k^*}$
*if* $\bar{\alpha}_{k^*,n^*} > L - \ell_{n^*} + 1$
    *if* $\bar{\alpha}_{k^*,n^*} - (L - \ell_{n^*} + 1) > L - \ell_{n^*+1}$
        *if* $\bar{\alpha}_{k^*,n^*} - (L - \ell_{n^*} + 1) - (L - \ell_{n^*+1}) < L - \ell_{n^*+2}$
            $\bar{\alpha}_{k^*,n^*} = \alpha_{k^*}$
        *else* $\bar{\alpha}_{k^*,n^*} = (L - \ell_{n^*} + 1) + (L - \ell_{n^*+1}) + (L - \ell_{n^*+2})$
    *else* $\bar{\alpha}_{k^*,n^*} = \alpha_{k^*}$
*else* $\bar{\alpha}_{k^*,n^*} = \alpha_{k^*}$.                ■
```

Step 4—Power Rechecking: If $\bar{\alpha}_{k^*,n^*} > L$, user k^* can possibly simultaneously transmit on the same slots of more than one subchannel. Since we only check the transmission power on subchannel n^* at Step 2, the power constraint should be rechecked if it is still satisfied. If the power constraint is violated, the number of slots allocated to user k^* will be decreased until the power constraint is fulfilled. The function is given below.

```
**Function**: *Power Rechecking*
$c = \lceil \bar{\alpha}_{k^*,n^*} / L \rceil$
*if* $\bar{\alpha}_{k^*,n^*} > L$
    while $p_{k^*,n^*}^{(\ell_{n^*})}(\text{BER}_{k^*}^*, x_{k^*,n^*}^{(\ell_{n^*})}) \cdot c > p_{k^*,\max}$
        $\bar{\alpha}_{k^*,n^*} = L \cdot (c - 1)$
        $c = c - 1$
    end while
end if.                                                 ■
```

Step 5—Maximum Available Slot Finding: $x_{k^*,n^*}^{(\ell_{n^*})} \cdot q \cdot \bar{\alpha}_{k^*,n^*}$ would be the total number of available bits that the system can allocate to user k^* from subchannel n^*. If it is smaller than the actual number of required bits, which is $x_{k^*,n^*}^{(\ell_{n^*})} \cdot q \cdot \alpha_{k^*}$, the QoS requirements will not be fulfilled. Thus, we will search for other subchannels and choose the one for user k^*. The function is shown below.

Function: *maximum available slot finding*

```
if $x_{k^*,n^*}^{(\ell_{n^*})} \cdot q \cdot \bar{\alpha}_{k^*,n^*} < x_{k^*,n^*}^{(\ell_{n^*})} \cdot q \cdot \alpha_{k^*}$
  if $N^{\text{free}} - N^{k^*} \neq \{\phi\}$
    $n^* = \arg\max_{n \in N^{\text{free}} - N^{k^*}} (h_{k^*,n}), N^{k^*} = N^{k^*} + \{n^*\}$
    while $p_{k^*,n^*}^{(\ell_{n^*})} (\text{BER}_{k^*}^*, x_{k^*,n^*}^{(\ell_{n^*})} + 2) < p_{k^*,\max}$
      $x_{k^*,n^*}^{(\ell_{n^*})} = x_{k^*,n^*}^{(\ell_{n^*})} + 2$
    end while
  end if
  if $x_{k^*,n^*}^{(\ell_{n^*})} > 0$, go to step 3
  else $n^* = \arg\max_{n \in N^{k^*}} (x_{k^*,n}^{(\ell_{n^*})} \cdot \bar{\alpha}_{k^*,n})$
else
  $n^* = \arg\max_{n \in N^{k^*}} (x_{k^*,n}^{(\ell_{n^*})} \cdot \bar{\alpha}_{k^*,n})$.                ■
```

Step 6—Remapping: Due to the consistent allocation, slots allocated to user k^* must sequentially be from a certain (ℓ_{n^*}) slot of subchannel n^* to slots of the next subchannel $(n^* + 1$ or $n^* + 2)$. However, slots of the next subchannel may have been allocated to other users. To fulfill the slot allocation constraint, the slots that have been allocated to other users had better shift to the neighboring available slots. This process of shifting is called "remapping." Let σ_{k^*} be the length that needs to be shifted in unit of slots for user k^*. Thus, the user besides k^* that was originally allocated at the ℓth slot on subchannel n^* will be shifted to the $(\ell + \sigma_{k^*})$th slot. Note that if $(\ell + \sigma_{k^*}) > L$, it will be shifted to the slot of the next subchannel.

Function: *remapping*

```
$x_{k^*,n^*}^{(\ell_{n^*}+1)} = x_{k^*,n^*}^{(\ell_{n^*}+2)} = \cdots = x_{k^*,n^*}^{(L)} = x_{k^*,n^*}^{(\ell_{n^*})}$
$\sigma_{k^*} = \bar{\alpha}_{k^*,n^*} - (L - \ell_{n^*} + 1)$
for $n = n^* + 1$ to $N$
  for $\ell = 1$ to $L$
    if $\sum_{k \in K} x_{k,n}^{(\ell)} > 0$
      $\bar{k} = \arg\max_{k \in K} (x_{k,n}^{(\ell)})$
      if $\ell + \sigma_{k^*} \leq L$ and $n = n^* + 1$
        $x_{\bar{k},n}^{(\ell+\sigma_{k^*})} = x_{\bar{k},n}^{(\ell)}, x_{\bar{k},n}^{(\ell)} = 0, x_{k^*,n}^{(\ell)} = x_{k^*,n^*}^{(\ell_{n^*})}$
      else $\ell + \sigma_{k^*} > L$ and $n = n^* + 1$
        if $\sum_{k \in K} x_{k,n+1}^{(\ell+\sigma_{k^*}-L)} > 0$
          $\bar{\bar{k}} = \arg\max_{k \in K} (x_{k,n+1}^{(\ell+\sigma_{k^*}-L)})$
          if $\bar{\bar{k}} = \bar{k}$
            $\ell^* = \max\{\ell | x_{\bar{k},n+1}^{(\ell)} > 0\}$
          if $\sum_{k \in K} x_{k,n+1}^{(\ell^*+1)} > 0$
            $\bar{\bar{\bar{k}}} = \arg\max_{k \in K} (x_{k,n+1}^{(\ell^*+1)})$
              $\ell^{**} = \max\{\ell | x_{\bar{\bar{\bar{k}}},n+1}^{(\ell)} > 0\}$
            $x_{\bar{\bar{\bar{k}}},n+1}^{(\ell^{**}+1)} = x_{\bar{\bar{\bar{k}}},n+1}^{(\ell^*+1)},\ x_{\bar{k},n+1}^{(\ell^*+1)} = x_{\bar{k},n}^{(\ell)},\ x_{\bar{k},n}^{(\ell)} = 0$
              $x_{k^*,n}^{(\ell)} = x_{k^*,n^*}^{(\ell_{n^*})}$
          else
            $x_{\bar{k},n+1}^{(\ell^*+1)} = x_{\bar{k},n}^{(\ell)}, x_{\bar{k},n}^{(\ell)} = 0, x_{k^*,n}^{(\ell)} = x_{k^*,n^*}^{(\ell_{n^*})}$
        else
          $\ell^{***} = \max\{\ell | x_{\bar{\bar{k}},n+1}^{(\ell)} > 0\}$
          $x_{\bar{\bar{k}},n+1}^{(\ell^{***}+1)} = x_{\bar{\bar{k}},n+1}^{(\ell+\sigma_{k^*}-L)},\ x_{\bar{k},n+1}^{(\ell+\sigma_{k^*}-L)} = x_{\bar{k},n}^{(\ell)}$
              $x_{\bar{k},n}^{(\ell)} = 0, x_{k^*,n}^{(\ell)} = x_{k^*,n^*}^{(\ell_{n^*})}$
      else
        $\bar{k} = \arg\max_{k \in K} (x_{k,n}^{(\ell)})$
        $x_{\bar{k},n+1}^{(\ell+\sigma_{k^*}-L)} = x_{\bar{k},n}^{(\ell)}, x_{\bar{k},n}^{(\ell)} = 0, x_{k^*,n}^{(\ell)} = x_{k^*,n^*}^{(\ell_{n^*})}$
    end if
  end
end
  $\Omega = \Omega - \{k^*\}$.                ■
```

Fig. 1 shows the flowchart of the proposed DPRA scheme. The DPRA scheme will continuously be executed until there are no free subchannels or no backlogged users. Note that the DPRA scheme is a kind of greedy algorithm, which can find a near-optimal solution for the optimization equations given in (15)–(18) [14]. Furthermore, the functions in Steps 3–6 are the consistent allocation customized for IEEE 802.16 systems.

IV. SIMULATION RESULTS AND DISCUSSIONS

A. Simulation Environment

In the simulations, the system parameters of the uplink OFDMA environment are set to be compatible with the IEEE 802.16 standard [1], and the scalable parameters are configured according to the suggested values in [17] and listed in Table I. Furthermore, slow fading and fast fading for wireless channels are considered. The path loss is modeled as $128.1 + 37.6 \log R(\text{dB})$ [18], where R is the distance between the BS and the SS in unit of kilometers. The shadowing model is assumed to be lognormally distributed with zero mean and a standard deviation of 8 dB. Six taps of independent Rayleigh-distributed path are adopted for the fast-fading model, and the power delay profile follows the exponential decay rule. The CSI is assumed to be invariant within a frame and varies from frame to frame.

B. Source Model and QoS Requirements

Four types of traffic corresponding to UGS, rtPS, nrtPS, and BE are considered in the simulations. The first one is voice traffic for UGS and is modeled as the ON–OFF model [19]. The ON (OFF) time is exponentially distributed with a mean of 1 (1.35) s. During the ON period, one voice packet is generated every 20 ms with a fixed size of 28 B including the payload and header. Thus, the mean data rate of voice traffic during the ON period is 11.2 Kb/s [20]. The second service type is video streaming traffic for rtPS. The video streaming consists of a sequence of video frames regularly generated with an interval of 100 ms. Each video frame is composed of eight slices, and each slice corresponds to a single packet. The size of each packet is in truncated Pareto distribution, and the interarrival time between two consecutive packets is also distributed with truncated Pareto distribution. The parameters of the video streaming model are configured according to [18] and [19], and the data rate is 64 Kb/s.

The third service type is web browsing Hypertext Transfer Protocol (HTTP) traffic for nrtPS [18]. It is modeled as a

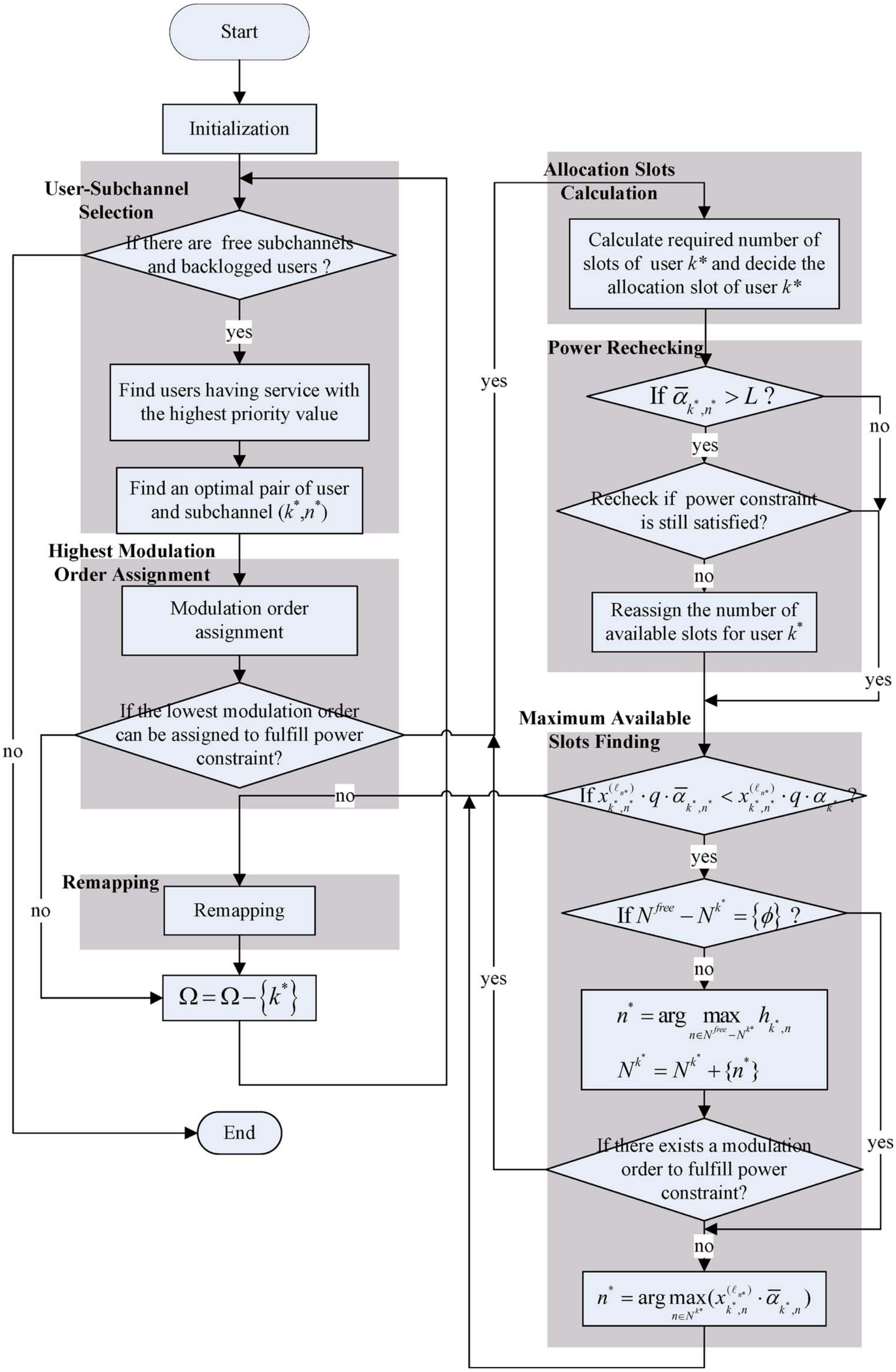


Fig. 1. Flowchart of the DPRA scheme.

sequence of page downloads, and each page consists of a sequence of packet arrivals. Each page is composed of a main object and some embedded objects, which can be divided into several packets. The maximum transmission unit of each packet is 1500 B. The interarrival time between two downloaded pages is exponentially distributed with a mean of 30 s. The sizes of both the main object and the embedded object are in truncated lognormal distribution with a mean of 10 710 and 7758 B, respectively. The last service type is File Transfer Protocol (FTP) traffic for the BE service, which is modeled as a sequence of file downloads. The size of each file is in truncated lognormal distribution with a mean of 2 MB, a standard deviation of 0.722 MB, and a maximum value of 5 MB. The interarrival time between two files is exponentially distributed with a mean of 180 s. Finally, the QoS requirements for each service type configured in the simulation are listed in Table II. Furthermore, the values of $D_{\rm th}$ and $T_{\rm th}$ given in (2) and (3) are set to be 2 and 1, respectively; the a given in (3) is set to be 1.

TABLE I
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

TABLE II
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	15 ms
	Max. allowable packet dropping rate	1%
HTTP (*nrtPS*)	Required BER	10^{-6}
	Min. transmission rate	100 Kbps
FTP (*BE*)	Required BER	10^{-6}

C. Performance Evaluation

To analyze the performance of the proposed DPRA scheme, the DPRA scheme is compared with the optimal method, which obtains the optimal solution by exhaustively solving the mathematical equations given in (15)–(18). However, when performing the consistent allocation for a user, the optimal method will choose the sort of consistent allocation that achieves the maximum system throughput. The DPRA scheme is also compared to one conventional scheme called the EFS algorithm, proposed in [10]. The EFS algorithm allocates slots to each service according to the order of UGS, rtPS, nrtPS, and BE, where UGS (BE) has highest (lowest) priority. At each slot time interval, it assigns a slot of the selected subchannel to the user with maximum transmission rate on that subchannel. If all the subchannels of a certain slot interval are exhausted, the EFS algorithm will move to the next slot interval and iteratively perform allocation slot by slot. It is an intuitive algorithm, and its performance is close to an optimal solution [10], [14].

For simplicity, based on the allocated subchannel, modulation order, and available slots, the DPRA scheme will perform consistent allocation for each service type of the selected user. Sequential slots on the selected subchannel will be allocated in order for UGS, rtPS, nrtPS, and BE until the available slots are used out. In the simulations, the number of users is varied from 4 to 40. Each user is assumed to have voice, video, HTTP, and FTP traffic. The maximum system transmission rate per frame, which is equal to 7.3728 Mb/s, can be achieved when the highest modulation order is assigned in each slot. We define the traffic intensity as the ratio of the total average arrival rate of all service types of all users over the maximum system transmission rate. In addition, the average arrival rate for voice, video, HTTP, and FTP is 4.8, 64, 14.5, and 88.9 Kb/s, respectively.

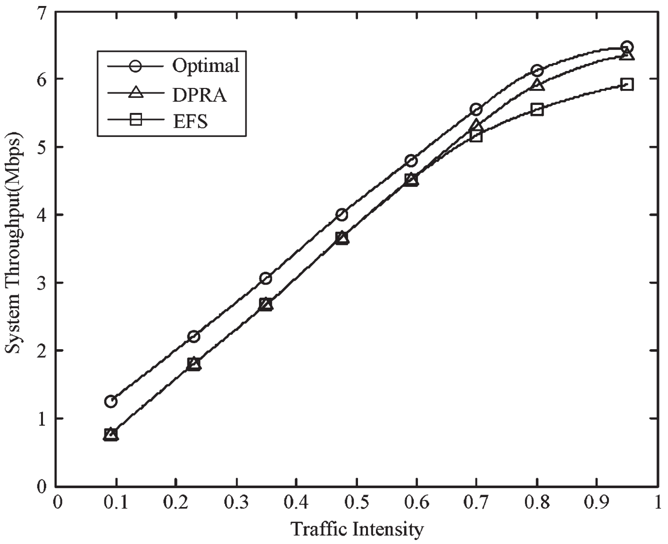


Fig. 2. System throughput.

Fig. 2 shows the system throughput versus the traffic intensity. It can be seen that the system throughput of the proposed DPRA scheme is close to that of the optimal method; the former is just less than the latter by an amount of 3.58% at a traffic intensity of 0.8. This conforms to the statement in [14]: the result of the greedy method is close to the optimal solution when the number of users is large. It can also be found that the proposed DPRA scheme achieves a higher system throughput than the EFS algorithm, particularly at higher traffic load because at high traffic intensity, the proposed DPRA scheme can dynamically adjust the priority values, and the more urgent service will be given a higher priority value and allocated more resources to avoid packet dropping. The DPRA scheme can more effectively allocate the radio resource. On the other hand, the EFS algorithm performs slot-by-slot allocation using a fixed priority scheme and can gain performance close to the optimal solution. The EFS algorithm attains a system throughput as large as that of the DPRA scheme until the traffic intensity exceeds 0.6. However, due to the lack of dynamic priority, which reflects the urgency of each service, in the EFS algorithm, more packets may be dropped, and the system throughput degrades at high traffic intensity.

Fig. 3 shows the packet dropping rate of voice traffic. The proposed DPRA scheme, the optimal method, and the EFS algorithm attain voice packet dropping rates close to zero, which fulfills the voice QoS requirement of 1%. The reason for this is that these three schemes allocate the voice traffic for UGS a constant amount of bandwidth and that the resource allocation is *a priori* to others.

Fig. 4 shows the packet dropping rate of video traffic. The video packet dropping rate of the DPRA scheme keeps smaller than the QoS requirement of dropping rate (1%) until the traffic intensity is above 0.9. On the other hand, the video packet dropping rate of the optimal method (the EFS algorithm) is guaranteed until the traffic intensity is above 0.7 (0.6). The DPRA scheme designs an appropriate dynamic priority value for each service, which is adjusted frame by frame. According to the QoS requirements of video traffic, the priority value of video packets can be promoted, and most video packets can

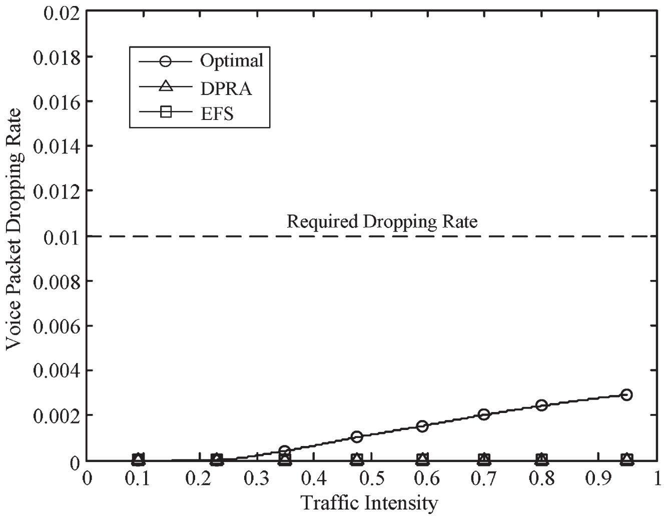


Fig. 3. Voice packet dropping rate.

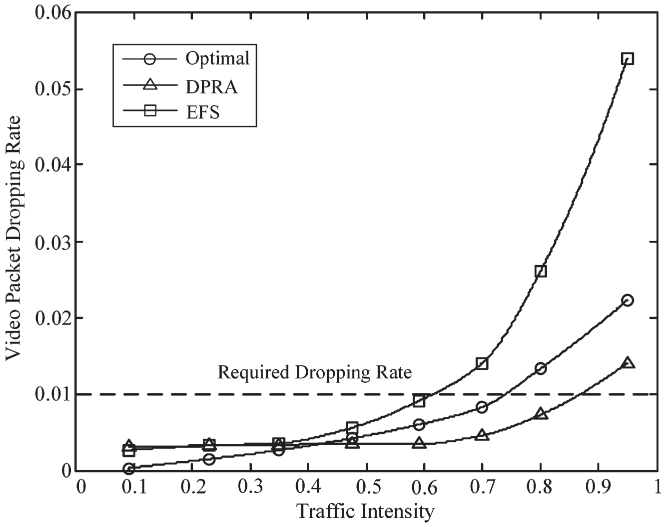


Fig. 4. Video packet dropping rate.

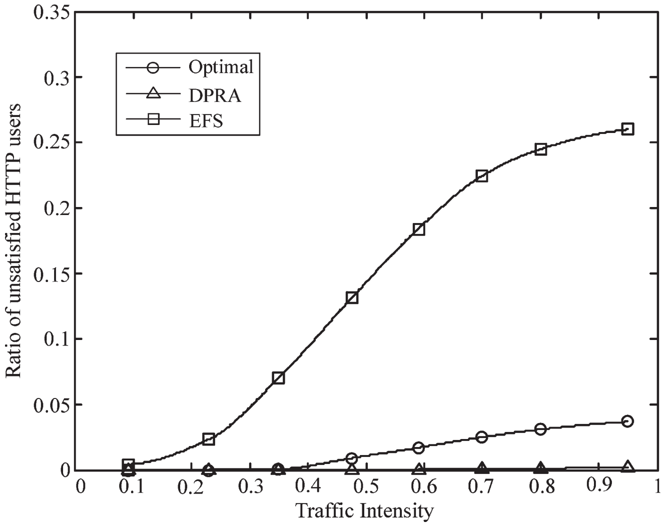


Fig. 5. Ratio of unsatisfied HTTP users.

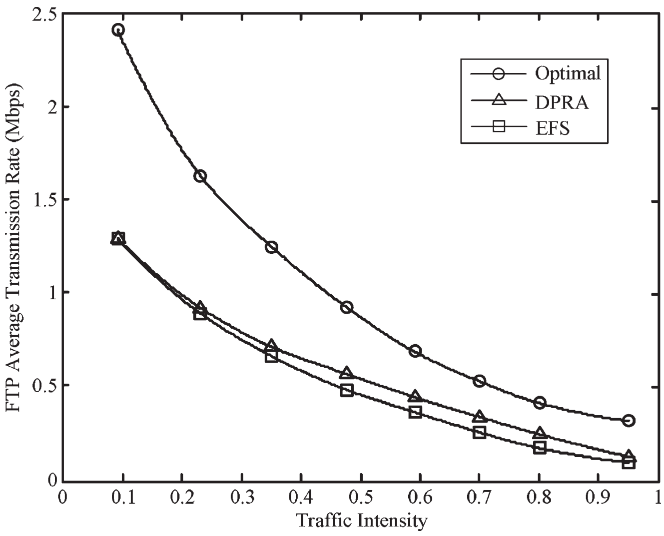


Fig. 6. FTP average transmission rate.

be served in time to avoid discarding. However, the optimal method is mainly to maximize the system throughput. Thus, sometimes, the amount of bandwidth granted to the video streaming service is not sufficient enough; the EFS algorithm allocates the system resource to video service right after finishing allocation to voice service.

Fig. 5 shows the ratio of unsatisfied HTTP users, which is defined as the number of HTTP users whose average transmission rate is less than the minimum required transmission rate, over all HTTP users. For the DPRA scheme, a high priority value will be given to the HTTP user if its average transmission rate is going to be lower than the minimum required transmission rate. Therefore, the ratio of unsatisfied HTTP users of the DPRA scheme keeps close to zero, even when the traffic intensity is high, and the minimum required transmission rate can be guaranteed. On the other hand, the optimal method is mainly to maximize the system throughput. Thus, sometimes, the minimum required transmission rate cannot be assured when the traffic intensity becomes high. The EFS algorithm is designed with a fixed priority scheme that initially assigns service traffic with priorities according to their QoS requirements. Thus, their ratios of unsatisfied HTTP users will increase with the traffic load due to the lack of enough resources allocated for each HTTP user.

Fig. 6 shows the average throughput of FTP users. For the EFS algorithm, the FTP traffic will be transmitted only when voice, video, and HTTP traffic have already been served. Thus, its average throughput is the lowest. For the DPRA scheme and the optimal method, the FTP traffic is also specified with the lowest priority value. However, the DPRA scheme (the optimal method) achieves the larger (the largest) system throughput, as illustrated in Fig. 2, thus comes the result. The optimal method outperforms the DPRA and the EFS scheme by 67.9% and 75.3% in FTP average transmission rate at a traffic intensity of 0.8, respectively.

Fig. 7 shows the average number of iterations per frame for the proposed DPRA scheme, the optimal method, and the EFS algorithm. Here, an iteration is defined as a search for an optimal pair of user and subchannel from K users and N free subchannels to be allocated with a slot. The DPRA scheme

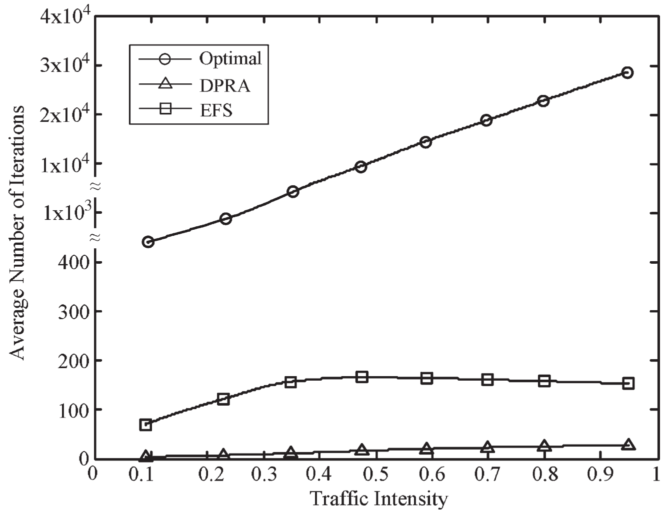


Fig. 7. Average number of iterations.

takes far fewer iterations than the EFS algorithm. It is because the DPRA scheme designs a consistent allocation mechanism, where the iteration computation for allocation to a user in each frame takes only one time, and the number of slots allocated to the user could be more than one. Therefore, the number of iterations by the DPRA scheme is greatly reduced, which is much smaller than the total number of slots, i.e., $N \times L$, in a frame. On the other hand, the EFS algorithm performs the iteration computation for allocation to a user slot by slot. It searches for an optimal pair of user and subchannel for each symbol. Furthermore, the EFS algorithm may need more than one iteration for each slot allocation to an optimal pair of user and subchannel if there is a power constraint violation. Thus, the average number of iterations per frame by the EFS algorithm could be larger than the total number of slots, i.e., $N \times L$, in a frame. Moreover, in an iteration, the DPRA and the EFS just search for a pair of user and subchannel and check the power constraint. The complexity of an iteration by the DPRA and the EFS is almost the same and equal to $O(KN)$. It can also be found that the number of iterations by the optimal method is much lager than that of the DPRA scheme. The optimal method takes more than 23 500 iterations, while the DPRA scheme needs only 23 iterations at a traffic intensity of 0.8.

V. CONCLUSION

In this paper, a DPRA scheme that performs consistent allocation has been proposed for IEEE 802.16 uplink systems with multimedia traffic. The DPRA scheme intends to maximize the system throughput and fulfill QoS requirements. It originally designs a priority value for each service type according to the urgency and QoS requirements of the traffic and dynamically adjusts it frame by frame. Simulation results have shown that the performance of the DPRA scheme is better than the conventional EFS algorithm, which performs allocation slot by slot. In addition, benefiting from the consistent allocation, the computational complexity of the DPRA scheme is much less than that of the conventional EFS algorithm. Furthermore, the system throughput of the proposed DPRA scheme is close to that of the optimal method when the traffic intensity is larger, but the computational complexity of the DPRA scheme is much less than that of the optimal solution. Therefore, we can conclude that the proposed DPRA scheme can reach throughput maximization and QoS satisfaction with a lower computational complexity and transmission overhead. The DPRA scheme would be suitable for real applications.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their suggestions that helped to improve the presentation and quality of the paper.

REFERENCES

[1] *IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std. 802.16-2004, 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 J. Sel. Areas 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, no. 5, pp. 1566–1575, Sep. 2004.

[4] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access system," *IEEE Trans. Broadcast.*, vol. 49, no. 4, pp. 362–370, Dec. 2003.

[5] N. Xu, Y. Wang, and P. Zhang, "Multiuser scheduling in downlink MIMO/OFDMA system with transmit preprocessing," in *Proc. Asia-Pacific Conf. Commun.*, Aug. 2006, pp. 1–5.

[6] S. Das and G. D. Mandyam, "An efficient sub-carrier and rate allocation scheme for M-QAM modulated uplink OFDMA transmission," in *Proc. 37th Asilomar Conf. Signals, Syst., Comput.*, 2003, vol. 1, pp. 136–140.

[7] K. Kim, Y. Han, and S. L. Kim, "Joint subcarrier and power allocation in uplink OFDMA systems," *IEEE Commun. Lett.*, vol. 9, no. 6, pp. 526–528, Jun. 2005.

[8] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.

[9] P. Hosein, "Adaptive subchannel allocation in an OFDMA-based wireless network," in *Proc. IFIP Int. Conf. Wireless Opt. Commun. Netw.*, 2006, pp. 1–5.

[10] V. Singh and V. Sharma, "Efficient and fair scheduling of uplink and downlink in IEEE 802.16 OFDMA networks," in *Proc. IEEE WCNC*, 2006, vol. 2, pp. 984–990.

[11] J. Chen and C. W. Chang, "A signal-aware uplink resource allocation strategy in IEEE 802.16 systems," in *Proc. WiMOB*, 2007, pp. 15–20.

[12] J. P. Yoon, W. J. Kim, J. Y. Baek, and Y. J. Suh, "Efficient uplink resource allocation for power saving in IEEE 802.16 OFDMA systems," in *Proc. VTC—Spring*, 2008, pp. 2167–2171.

[13] D. Niyato and E. Hossain, "Queue-aware uplink bandwidth allocation for polling services in 802.16 broadband wireless networks," in *Proc. IEEE GLOBECOM*, 2005, vol. 6, pp. 3702–3706.

[14] K. G. Murty, *Operations Research*. Englewood Cliffs, NJ: Prentice–Hall, 1995.

[15] L. M. A. Jalloul and S. P. Alex, "Coverage analysis for IEEE 802.16e/WiMAX systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4627–4634, Nov. 2008.

[16] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[17] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 WirelessMAN," *Intel Technol. J.*, vol. 8, no. 3, pp. 201–212, Aug. 2004.

[18] "Feasibility study for OFDM for UTRAN enhancement," *3rd Generation Partnership Project*. Tech. Rep., 2004-06.

[19] *Universal Mobile Telecommunication System, Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS*, UMTS Std. 30.03, 1998.

[20] CISCO Tech. Notes, *Voice over IP—per call bandwidth consumption*. Document ID 7934.

[21] IEEE 802.16m-08/004r2, *IEEE 802.16m Evaluation Methodology Document*, Jul. 3, 2008.

Chih-Ming Yen (S'06) is currently working toward the Ph.D. degree with the Department of Communication Engineering, College of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan.

His research interests include wireless networks, mobile communications, high-speed networks, communications protocol design, radio resource management, wireless communication, and network performance evaluation. He is currently focusing on research areas on resource management for orthogonal frequency division multiplexing systems.

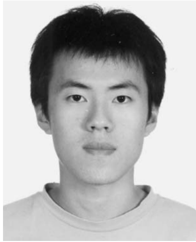
Chung-Ju Chang (S'81–M'85–SM'94–F'06) was born in Taiwan in August 1950. He received the B.E. and M.E. degrees in electronics engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 1972 and 1976, respectively, and the Ph.D. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1985.

From 1976 to 1988, he was with Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, Taiwan, as a Design Engineer, Supervisor, Project Manager, and then Division Director. He also acted as a Science and Technical Advisor for the Minister of the Ministry of Communications from 1987 to 1989. In 1988, he joined the faculty of the Department of Communication Engineering, College of Electrical and Computer Engineering, National Chiao Tung University, as an Associate Professor. He has been a Professor since 1993. He was the Director of the Institute of Communication Engineering from August 1993 to July 1995, the Chairman of the Department of Communication Engineering from August 1999 to July 2001, and the Dean of the Research and Development Office from August 2002 to July 2004. He was also an Advisor for the Ministry of Education to promote the education of communication science and technologies for colleges and universities in Taiwan from 1995 to 1999. He is acting as a Committee Member of the Telecommunication Deliberate Body, Taiwan. His research interests include performance evaluation, radio resource management for wireless communication networks, and traffic control for broadband networks.

Dr. Chang is a member of the Chinese Institute of Engineers. Moreover, he serves as an Editor for the *IEEE Communications Magazine* and an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Fang-Ching Ren (M'03) was born in Hsinchu, Taiwan. He received B.E., M.E., and Ph.D. degrees in communication engineering from the National Chiao Tung University, Hsinchu, in 1992, 1994, and 2001, respectively.

Since 2001, he has been a Protocol Design Engineer with the Industrial Technology Research Institute, Hsinchu, where he was involved in the design and development of the wideband code division multiple access chipset, WiMAX mobile multihop relay technology, and fourth-generation access technology. His current research interests include system performance analysis, protocol design, and mobile radio networks.



Jian-Ann Lai was born in Taipei, Taiwan. He received B.E. and M.E. degrees from the National Chiao Tung University, Hsinchu, Taiwan, in 2005 and 2007, respectively.

He is currently with the Department of Communication Engineering, College of Electrical and Computer Engineering, National Chiao Tung University. His research interests include radio resource management and scheduling.