Research Article

Opportunistic Scheduling for OFDM Systems with Fairness Constraints

Zhi Zhang,[1] Ying He,[2] and Edwin K. P. Chong[1]

[1] *Department of Electrical and Computer Engineering, Colorado State University, Ft. Collins, CO 80523, USA*
[2] *QuantWorks, Inc., Oak Hill, VA 20171, USA*

Correspondence should be addressed to Edwin K. P. Chong, edwin.chong@colostate.edu

We consider the problem of downlink scheduling for multiuser orthogonal frequency-division multiplexing (OFDM) systems. *Opportunistic scheduling* exploits the time-varying, location-dependent channel conditions to achieve multiuser diversity. Previous work in this area has focused on single-channel systems. Multiuser OFDM allows multiple users to transmit simultaneously over multiple channels. In this paper, we develop a rigorous framework to study opportunistic scheduling in multiuser OFDM systems. We derive optimal opportunistic scheduling policies under three QoS/fairness constraints for multiuser OFDM systems—temporal fairness, utilitarian fairness, and minimum-performance guarantees. Our scheduler decides not only *which time slot*, but also *which subcarrier* to allocate to each user. Implementing these optimal policies involves solving a maximal bipartite matching problem at each scheduling time. To solve this problem efficiently, we apply a modified Hungarian algorithm and a simple suboptimal algorithm. Numerical results demonstrate that our schemes achieve significant improvement in system performance compared with nonopportunistic schemes.

1. INTRODUCTION

Emerging broadband wireless networks which support high-speed packet data with a different quality of service (QoS) demand more flexible and efficient use of the scarce spectral resource. In contrast to wireline networks, one of the fundamental characteristics of wireless networks is the time-varying and location-dependent channel conditions due to multipath fading. From an information-theoretic viewpoint, Knopp and Humblet showed that the system capacity can be maximized by exploiting inherent *multiuser diversity* in the wireless channel [1]. The basic idea is to schedule a single user with the best instantaneous channel condition to transmit at any one time. The technology has already been implemented in the current 3G systems, that is, 1xEV-DO [2] and high-speed downlink packet access (HSDPA) [3]. The idea has also recently been adopted in cognitive radio systems which are novel intelligent wireless communication systems providing highly reliable and efficient communications by exploiting unused radio spectrum [4, 5].

Orthogonal frequency-division multiplexing (OFDM) is a popular multiaccess scheme widely used in DVB, wireless LANs (e.g., 802.16, ETSI HIPERLAN/2), and ultra wideband (UWB) systems [6]. It is also a promising modulation scheme of choice proposed for many future cellular networks such as cognitive radio systems [7, 8]. OFDM divides the total bandwidth into many narrowband orthogonal subcarriers, which are transmitted in parallel, to combat frequency-selective fading and achieve higher spectral utilization. OFDMA, a multiuser version of OFDM, allows multiple users to transmit simultaneously on the different subcarriers [9].

Good scheduling schemes in wireless networks should *opportunistically* seek to exploit the time-varying channel conditions to improve spectrum efficiency, thereby achieving *multiuser diversity* gain. However, the potential to transmit at higher data rates opportunistically also introduces an important tradeoff between wireless resource efficiency and level of satisfaction among individual users (fairness). For example, allowing only users close to the base station to transmit at high transmission rate may result in very high throughput, but may sacrifice the transmission of other users. Such a scheme cannot satisfy the increasing demand for QoS provisioning in broadband wireless networks.

To solve this problem, Liu et al. described a framework for opportunistic scheduling to exploit the multiuser diversity while at the same time satisfying three long-term QoS/fairness constraints [10–12]. In that work, only a single user can transmit at each scheduling time. The authors of [1] show that this is optimal for single-channel systems such as TDMA. However, the same is not the case for multiple-channel systems.

In this paper, we propose an opportunistic scheduling framework for multiuser OFDM systems. We build on Liu's work by going from the single-channel to the multiple-channel case. We show how the system performance can be optimized by serving multiple users simultaneously over the different subcarriers. We focus on the downlink of an OFDM system. We derive our opportunistic scheduling policies under three long-term QoS/fairness constraints—temporal fairness, utilitarian fairness, and minimum-performance guarantees, which are similar in form to those of [12], but adapted to the setting of multiuser OFDM systems. We also state optimality conditions under each of these constraints. In particular, our scheduler decides not only *which time slot* but also *which subcarrier* to allocate to each user under the given QoS/fairness constraints. A stochastic approximation algorithm is used to calculate the control parameters online in the policies. To search over the optimal user subsets efficiently, we apply a modified bipartite matching algorithm. We also develop an efficient, low-complexity suboptimal algorithm—our experimental results illustrate that this algorithm achieves near-optimal performance.

The remainder of this paper is organized as follows. In Section 2, we discuss related work on scheduling and fairness for OFDM. The system model is described in Section 3. In Section 4, we derive opportunistic scheduling policies under various fairness constraints and prove their optimality. In Section 5, we address some implementation issues, including control parameter estimation and the assignment problem that arises in implementing these policies. An optimal algorithm and an efficient suboptimal algorithm are proposed here. In Section 6, we present the numerical results to illustrate the performance of our policies. Finally, concluding remarks are given in Section 7.

2. RELATED WORK

Wireless scheduling has attracted a lot of recent attention. The authors of [13, 14] extend the scheduling policies for wireline networks to wireless networks to provide short-term and long-term fairness bounds. However, they model a channel as being either "good" or "bad," which may be too simple in some situations. In [15–17], the authors study wireless scheduling algorithms when both delay and channel conditions are taken into account. Scheduling with short-term fairness constraints is also discussed in [10, 18].

In [19, 20], the authors present a scheduling scheme for the Qualcomm IS-856 (also known as HDR (high data rate)) system. Their scheduling scheme exploits time-varying channel conditions while satisfying a certain fairness constraint known as *proportional fairness* [21]. Although there has been considerable recent efforts on proportional fairness schedul-

ing [22–24], to the best of our knowledge, there is currently no work considering multiuser OFDM systems with the three QoS fairness constraints we mentioned above. So in this paper we will focus on these three fairness constraints.

Opportunistic scheduling exploits the channel fluctuations of users. In [22], the authors use multiple "dumb" antennas to "induce" channel fluctuations, and thus exploit multiuser diversity in a slow fading environment. The authors of [25] show that with multiple antennas, transmitting to a carefully chosen subset of users has superior performance.

The resource management problem in OFDM systems has attracted a lot of research interest [26, 27]. In [26], the authors propose an algorithm to minimize the total transmission power with minimum-rate constraints for users. Specifically, the algorithm allocates a set of subcarriers to each user and then determines the number of bits and transmission power on each subcarrier. In [27], the authors study the problem of dynamic subcarrier and power allocation with the objective to maximize the minimum of the users' data rates subject to a total transmission power constraint. All these studies show that dynamic resource allocation (in terms of bit, subcarrier, and power) schemes can achieve significant performance gains over traditional static allocations (such as TDMA-OFDM and FDMA-OFDM). However, none of the schemes described above exploit multiuser diversity. For delay-insensitive data service, we can expect to reap long-term performance benefits by exploiting multiuser diversity.

OFDM has been used in several applications in cognitive radio. To enhance spectrum efficiency, the spectrum pooling system allows a license owner to share underutilized licensed spectrum with a secondary wireless system during its idle times [8]. A preferred transmission mode of the secondary system is OFDM due to its inherent flexibility. In [28], the authors discuss the desired properties in designing physical layers of cognitive radio systems and claim that the modulation scheme based on OFDM is a natural approach that satisfies the desired properties.

Recently, there has been significant interest in opportunistic scheduling and fairness issues for multiple-channel systems [29–33]. In [31], the authors consider a total-throughput maximization problem with deterministic and probabilistic constraints for multiple-channel systems. In [33], the authors consider opportunistic fair scheduling in downlink TDMA systems employing multiple transmit antennas and beamforming.

In [34], the authors introduce cross-layer optimization for OFDM wireless networks. The interaction between the physical layer and media access control (MAC) layer is exploited to balance the efficiency and fairness of wireless resource allocation. The authors consider proportional and max-min fairness.

3. SYSTEM MODEL

In this section, we describe the system model, assumptions, notation, and formulation of the scheduling problem.

The architecture of a downlink data scheduler for a single-cell multiuser OFDM system is depicted in Figure 1.

There is a base station (transmitter) with a single antenna communicating with N mobile users (receivers). Each user has different channel conditions over different subcarriers. By inserting pilot symbols in the downlink, the users can effectively estimate the channels. Every user should report its channel-state information over every subcarrier to the base station. All the channel-state information is sent to the subcarrier and bit allocation scheduler in the base station through feedback channels from all mobile users. The scheduling decision made by the scheduler is conveyed to the OFDM transmitter. The transmitter then assigns different transmission rates to scheduled users on corresponding subcarriers. The scheduler makes decisions once every time slot based on the channel-state information and the control parameters for fairness guarantees.

We assume that the base station knows the perfect channel-state information for each user over each subcarrier. The channel conditions for different users are usually independently varying in a multiuser system. Owing to frequency-selective fading, one user may experience deep fading in some subcarriers, but relatively good in other subcarriers. By dynamically assigning users to favorable subcarriers, the overall performance of the network can be increased from the multiuser diversity. In practice, requiring "perfect" channel-state information results in significant feedback overhead burden, which might be difficult to implement. We can view our current work as providing fundamental performance bounds on what is achievable with channel feedback.

The OFDM signaling is time slotted. The length of a time slot is fixed and the channel does not vary significantly during a time slot. The length of a time slot in the scheduling policy can be different from an actual time slot in the physical layer. It depends on how fast the channel conditions vary and how fast we want to track such changes.

We assume that there is always data for each user to receive, that is, the system has *infinite backlogged* data queues. We also assume that the transmission power is uniformly allocated to all subcarriers. In principle, performance can be improved further by allocating a different power level to each subcarrier. In a system with a large number of users, this improvement could be *marginal* because of statistical effects [22].

In this paper, we will focus on scenarios with large numbers of users, or heavy-traffic systems, where the number of users is greater than the number of available OFDM subcarriers. These scenarios can be regarded as an extreme situation for OFDM. But it is important to determine the impact of a large number of users, such as in [22]. Our goal is *to maximize the system performance by exploiting the time-varying and frequency-varying channel conditions while maintaining certain QoS/fairness constraints*.

Let $i = 1, \ldots, N$ be the index of users, and $k = 1, \ldots, K$ the index of subcarriers. Following [12], let $\omega_{i,k}^t$ be the instantaneous performance value that would be experienced by user i if it were scheduled to transmit over subcarrier k at time slot t. The $\omega_{i,k}^t$ comprise an $N \times K$ matrix, denoted as ω^t. Usually, the better the channel condition of user i over subcarrier k, the larger the value of $\omega_{i,k}^t$. Throughput (in

terms of data rate bits/sec) is the most straightforward form of a time-varying and channel-condition-dependent performance measure. For convenience, the reader can think of throughput as the performance measure in this paper. However, our formulation applies in general.

Let $\overline{A} = (A_1, A_2, \ldots, A_K)$ represent a *scheduling action*, which is a vector of the indices of the users scheduled over all K subcarriers. The decision rule $\pi^t(\cdot)$, which is a function of ω^t, specifies which action should be chosen, that is, $\pi^t(\omega^t) = \overline{A}^t = (A_1^t, A_2^t, \ldots, A_K^t)$, where the value of A_k^t is the index of the user scheduled over subcarrier k at time t. We call $\pi(\cdot) = \{\pi^1(\cdot), \pi^2(\cdot), \ldots, \pi^t(\cdot), \ldots\} \in \Pi$ a *policy*, where Π is the set of all scheduling polices. Note that a policy may involve a *time-varying* rule for deciding scheduling actions. We are only interested in the so-called *feasible* policies, those that satisfy specific QoS/fairness requirements (described in the next section).

Let $U_i^T(\pi)$ be the average throughput of user i up to time T, and $R_i^T(\pi)$ the average resource consumption of user i up to time T, that is,

$$U_i^T(\pi) = \frac{1}{T} \sum_{t=1}^{T} \sum_{k=1}^{K} \omega_{i,k}^t \mathbf{1}_{\{A_k^t = i\}}, \quad i = 1, \ldots, N,$$

$$R_i^T(\pi) = \frac{1}{T} \sum_{t=1}^{T} \sum_{k=1}^{K} \mathbf{1}_{\{A_k^t = i\}}, \quad i = 1, \ldots, N, \tag{1}$$

where $\mathbf{1}_A$ is the indicator function of the event A, that is, $\mathbf{1}_A$ takes value 1 if A occurs, and is 0 otherwise.

Let $U^T(\pi) = \sum_{i=1}^{N} U_i^T(\pi)$, that is, $U^T(\pi)$ is the average overall throughput up to time T. Then we define

$$U(\pi) = \limsup_{T \to \infty} U^T(\pi), \tag{2}$$

which can be considered as the asymptotic best-case system performance of policy π.

Using the above notation, our goal can be formally stated as follows: *find a feasible policy π that maximizes the system performance $U(\pi)$ while maintaining certain QoS/fairness constraints*. In the following section, we derive optimal policies for three categories of scheduling problems, each with a unique QoS/fairness requirement.

4. OPPORTUNISTIC SCHEDULING UNDER VARIOUS FAIRNESS CONSTRAINTS

Good scheduling schemes should be able to exploit the time-varying channel conditions of users to achieve higher utilization of wireless resources, while at the same time guarantee some level of fairness among users. Fairness is central to scheduling problems in wireless systems. Without a good fairness criterion, the system performance can be trivially optimized, but might prevent some users from accessing the network resource. In this section, we will study scheduling problems under three fairness criteria for multiuser OFDM systems—temporal fairness, utilitarian fairness, and minimum-performance guarantees. These categories of fairness are adopted from [12] and are extended to multiuser OFDM systems. It turns out that the form of the optimal policies here bear a resemblance to those of [12].
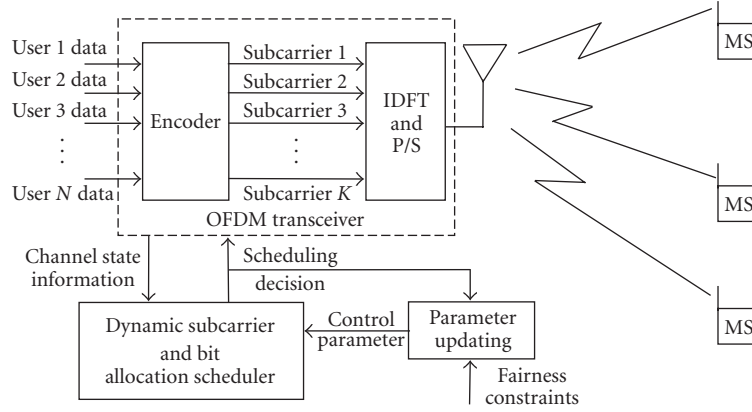
FIGURE 1: Downlink scheduling over multiuser OFDM system.

4.1. Temporal fairness scheduling

A natural fairness criterion is to give each user a certain long-term fraction of time because time is the basic resource shared among users. The problem of multiuser OFDM scheduling with temporal fairness can be expressed as

$$\max_{\pi \in \Pi} U(\pi) \quad \text{subject to} \lim_{T \to \infty} \inf R_i^T(\pi) \geq r_i, \quad (3)$$
$$i = 1, \ldots, N,$$

where r_i denotes the minimum time fraction that should be assigned to user i, with $r_i \geq 0$ and $\sum_{i=1}^{N} r_i \leq 1$. Recall that $R_i^T(\pi)$ is the average resource consumption of user i up to time T. The r_is are predetermined and serve as the prespecified fairness constraints. The value of r_i denotes the minimum fraction of time that user i should transmit over all the subcarriers in the long run, which is usually determined by the user's class, the price paid by the user, and so forth.

Define the policy π^* as follows:

$$\pi^*(\omega^t) = \arg\max_{\overline{A}^t} \left\{ \sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t = i\}} \right\}, \quad (4)$$

where the control parameters v_i^* are chosen such that

(1) $v_i^* \geq 0$, for all i;
(2) $\lim\inf_{T \to \infty} R_i^T(\pi^*) \geq r_i$, for all i;
(3) if $\lim\inf_{T \to \infty} R_i^T(\pi^*) > r_i$, then $v_i^* = 0$, for all i.

Similar to [10], we can think of $\vec{v}^* = (v_1^*, \ldots, v_N^*)$ in (4) as an "offset" or "threshold" to satisfy the temporal fairness constraints. Under this constraint, the scheduling policy schedules the "relatively best" subset of users to transmit. The subset of users selected by action $\overline{A}^t$ is "relatively best" if $\sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t = i\}}$ is maximum over all actions. If $v_i^* > 0$, then user i is an "unfortunate" user, that is, the channel conditions it experiences over all subcarriers are relatively poor. (e.g., it is far from the base station.) Hence, it has to take advantage of other users (e.g., users with $v_i^* = 0$) to

satisfy its fairness requirement. But to maximize the overall system performance, we can only give the "unfortunate" users their minimum time-fraction requirements, hence condition 3.

The policy π^* defined in (4), which represents our opportunistic scheduling policy, is optimal in the following sense.

Theorem 1. *If* $\lim_{T \to \infty} R_i^T(\pi^*)$ *exists for all i for* π^*, *then the policy* π^* *is an optimal solution to the problem defined in (3), that is, it maximizes the average OFDM system performance under the temporal fairness constraints.*

Proof. Let π be a policy satisfying the temporal fairness constraints, and let v_i^* satisfy conditions 1–3. Hence, we have

$$U(\pi) \leq U(\pi) + \sum_{i=1}^{N} v_i^* \left(\lim_{T \to \infty} \inf R_i^T(\pi) - r_i \right)$$
$$= \lim_{T \to \infty} \sup \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{k=1}^{K} \omega_{i,k}^t \mathbf{1}_{\{A_k^t = i\}}$$
$$+ \sum_{i=1}^{N} v_i^* \lim_{T \to \infty} \inf \frac{1}{T} \sum_{t=1}^{T} \sum_{k=1}^{K} \mathbf{1}_{\{A_k^t = i\}} - \sum_{i=1}^{N} v_i^* r_i$$
$$\leq \lim_{T \to \infty} \sup \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{k=1}^{K} \omega_{i,k}^t \mathbf{1}_{\{A_k^t = i\}}$$
$$+ \lim_{T \to \infty} \inf \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{k=1}^{K} v_i^* \mathbf{1}_{\{A_k^t = i\}} - \sum_{i=1}^{N} v_i^* r_i$$
$$(5)$$
$$\leq \lim_{T \to \infty} \sup \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t = i\}} - \sum_{i=1}^{N} v_i^* r_i.$$
$$(6)$$

By the definition of π^*, we have

$$\sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t = i\}} \leq \sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{(A_k^t)^* = i\}}. \quad (7)$$

Thus,

$$
\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{A_k^t = i\}}
$$
$$
\leq \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{(A_k^t)^* = i\}}. \tag{8}
$$

Therefore,

$$
U(\pi) \leq \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{k=1}^{K} (\omega_{i,k}^t + v_i^*) \mathbf{1}_{\{(A_k^t)^* = i\}} \sum_{i=1}^{N} v_i^* r_i
$$
$$
\leq U(\pi^*) + \limsup_{T \to \infty} \sum_{i=1}^{N} v_i^* R_i^T(\pi^*) - \sum_{i=1}^{N} v_i^* r_i \tag{9}
$$
$$
\leq U(\pi^*) + \sum_{i=1}^{N} v_i^* \left(\limsup_{T \to \infty} R_i^T(\pi^*) - r_i \right). \tag{10}
$$

Since $\lim_{T \to \infty} R_i^T(\pi^*)$ exists, $\limsup_{T \to \infty} R_i^T(\pi^*) = \liminf_{T \to \infty} R_i^T(\pi^*)$. Thus,

$$
U(\pi) \leq U(\pi^*) + \sum_{i=1}^{N} v_i^* \left(\liminf_{T \to \infty} R_i^T(\pi^*) - r_i \right)
$$
$$
= U(\pi^*), \tag{11}
$$

where the second part of (11) equals zero because of condition 3 on v_i^*. □

Inequalities (5), (6), (9), and (10) follow from the following properties of lim sup and lim inf [40]. If $\{x_n\}$ and $\{y_n\}$ are real sequences, we have

$$
\liminf_{n \to \infty} x_n + \liminf_{n \to \infty} y_n \leq \liminf_{n \to \infty} (x_n + y_n)
$$
$$
\leq \limsup_{n \to \infty} x_n + \liminf_{n \to \infty} y_n
$$
$$
\leq \limsup_{n \to \infty} (x_n + y_n) \tag{12}
$$
$$
\leq \limsup_{n \to \infty} x_n + \limsup_{n \to \infty} y_n.
$$

It is possible that the optimal policy is confronted with a tie between two or more users. When ties occur in the argmax in the policy, they can be broken arbitrarily.

4.2. Utilitarian fairness scheduling

In the last section, we studied the opportunistic scheduling problem for multiuser OFDM with temporal fairness constraints. In wireline networks, when a certain amount of resource is assigned to a user, it is equivalent to granting the user a certain amount of throughput. However, the situation is different in wireless networks, where the performance value and the amount of resource are not directly related. Therefore, a potential problem in wireless network is that the temporal fairness scheme has no way of explicitly ensuring that each user receives a certain guaranteed fair amount of utility. Hence, in this section, we will describe an alternative

scheduling problem that would ensure that all users get at least a certain fraction of the overall system performance.

The problem of multiuser OFDM scheduling with utilitarian fairness can be expressed as

$$
\max_{\pi \in \Pi} U(\pi) \quad \text{subject to} \liminf_{T \to \infty} U_i^T(\pi) \geq a_i U(\pi),
$$
$$
i = 1, \ldots, N, \tag{13}
$$

where a_i denotes the minimum fraction of the overall average throughput required by user i, with $a_i \geq 0$ and $\sum_{i=1}^{N} a_i \leq 1$. Recall that $U_i^T(\pi)$ is the average throughput of user i up to time T using policy π, and $U(\pi)$ is the average overall throughput. The a_i's are predetermined fairness constraints here. This constraint requires long-term fairness in terms of performance value (throughput) instead of resource consumption (time) as in Section 4.1.

We define the policy π^* as follows:

$$
\pi^*(\omega^t) = \arg\max_{\overline{A}^t} \left\{ \sum_{i=1}^{N} \sum_{k=1}^{K} (\kappa + \gamma_i^*) \omega_{i,k}^t \mathbf{1}_{\{A_k^t = i\}} \right\}, \tag{14}
$$

where $\kappa = 1 - \sum_{i=1}^{N} a_i \gamma_i^*$, and the control parameters γ_i^* are chosen such that

(1) $\gamma_i^* \geq 0$, for all i;
(2) $\liminf_{T \to \infty} U_i^T(\pi^*) \geq a_i U(\pi^*)$, for all i;
(3) if $\liminf_{T \to \infty} U_i^T(\pi^*) > a_i U(\pi^*)$, then $\gamma_i^* = 0$, for all i.

Analogous to $\vec{v}^*$ in the last section, $\vec{\gamma}^* = (\gamma_1^*, \ldots, \gamma_N^*)$ in (14) can be considered as a "scaling" to satisfy the utilitarian fairness constraints. The scheduling policy always schedules the "relatively best" subset of users to transmit. Here, the subset of users selected by action $\overline{A}^t$ is "relatively best" if $\sum_{i=1}^{N} \sum_{k=1}^{K} (\kappa + \gamma_i^*) \omega_{i,k}^t \mathbf{1}_{\{A_k^t = i\}}$ is maximum over all actions. If $\gamma_i^* > 0$, then user i is an "unfortunate" user, and its average performance value equals its minimum requirement.

The policy π^* defined in (14), which represents our opportunistic scheduling policy, is optimal in the following sense.

Theorem 2. *If $\lim_{T \to \infty} U_i^T(\pi^*)$ exists for all i for π^* defined in (14), then the policy π^* is an optimal solution to the problem defined in (13), that is, it maximizes the average OFDM system performance under the utilitarian fairness constraints.*

Proof. Let π be a policy satisfying the utilitarian fairness constraints, and let γ_i^* satisfy conditions 1–3. Hence, we have

$$
U(\pi) \leq U(\pi) + \sum_{i=1}^{N} \gamma_i^* \left(\liminf_{T \to \infty} U_i^T(\pi) - a_i U(\pi) \right)
$$
$$
= \limsup_{T \to \infty} \sum_{i=1}^{N} \kappa U_i^T(\pi) + \sum_{i=1}^{N} \gamma_i^* \liminf_{T \to \infty} U_i^T(\pi) \tag{15}
$$
$$
\leq \limsup_{T \to \infty} \sum_{i=1}^{N} (\kappa + \gamma_i^*) U_i^T(\pi),
$$

where $\kappa = 1 - \sum_{i=1}^{N} a_i \gamma_i^*$. By the definition of π^*, we get

$$\sum_{i=1}^{N} (\kappa + \gamma_i^*) U_i^T(\pi) \le \sum_{i=1}^{N} (\kappa + \gamma_i^*) U_i^T(\pi^*). \qquad (16)$$

Therefore,

$$
\begin{aligned}
U(\pi) &\le \limsup_{T \to \infty} \sum_{i=1}^{N} (\kappa + \gamma_i^*) U_i^T(\pi^*) \\
&\le U(\pi^*) + \sum_{i=1}^{N} \gamma_i^* \left(\liminf_{T \to \infty} U_i^T(\pi^*) - a_i U(\pi^*) \right) \\
&= U(\pi^*),
\end{aligned}
\qquad (17)
$$

where the second part of (17) equals zero because of condition 3 on γ_i^*. Similar to the proof of Theorem 1, the properties of lim sup and lim inf are applied here. □

4.3. Minimum-performance guarantee scheduling

So far, we have discussed two optimal multiuser OFDM scheduling policies that provide users with different fairness guarantees. However, while they satisfy a relative measure of performance (e.g., fairness), they do not consider any absolute measures such as data rate. This motivates the study of a category of scheduling problems with minimum-performance guarantees [11, 35].

The problem to maximize the OFDM system performance while satisfying each user's minimum performance requirement can be stated as

$$\max_{\pi \in \Pi} U(\pi) \quad \text{subject to} \quad \liminf_{T \to \infty} U_i^T(\pi) \ge C_i, \qquad (18)$$

$$i = 1, \ldots, N,$$

where $\vec{C} = \{C_1, C_2, \ldots, C_N\}$ is a feasible predetermined *minimum-performance requirement vector*. Feasible here means that there exists some policy that solves (18).

The QoS constraints here offer users a more direct service guarantee. For example, a user requires a minimum data rate guarantee, then the performance measure here can be data rate. Every user is guaranteed a minimum data rate, which may be more appealing from the user viewpoint. However, it can be quite difficult in practice to apply because of the difficulty to determine if a requirement vector is feasible.

Suppose $\vec{C} = \{C_1, C_2, \ldots, C_N\}$ is feasible. We define the policy π^* for the problem in (18) as follows:

$$\pi^*(\omega^t) = \arg\max_{\overline{A}^t} \left\{ \sum_{i=1}^{N} \sum_{k=1}^{K} \beta_i^* \omega_{i,k}^t \mathbf{1}_{\{A_k^t = i\}} \right\}, \qquad (19)$$

where the control parameters β_i^* are chosen such that

(1) $\beta_i^* \ge 1$, for all i;
(2) $\liminf_{T \to \infty} U_i^T(\pi) \ge C_i$, for all i;
(3) if $\liminf_{T \to \infty} U_i^T(\pi) > C_i$, then $\beta_i^* = 1$, for all i.

Note that the parameter $\vec{\beta}^* = (\beta_1^*, \ldots, \beta_N^*)$ "scales" the performance values of users, and the scheduling policy always schedules the "relatively best" subset of users to transmit. Here, the subset of users selected by action $\overline{A}^t$ is "relatively best" if $\sum_{i=1}^{N} \sum_{k=1}^{K} \beta_i^* \omega_{i,k}^t \mathbf{1}_{\{A_k^t = i\}}$ is maximum over all actions. If $\beta_i^* > 1$, then user i is an "unfortunate" user, and it is granted only its minimum-performance requirement.

The policy π^* defined in (19), which represents our opportunistic scheduling policy, is optimal in the following sense.

Theorem 3. *If $\lim_{T \to \infty} U_i^T(\pi^*)$ exists for all i for the π^* defined in (19), then the policy π^* is an optimal solution to the problem defined in (18), that is, it maximizes the average OFDM system performance under the minimum-performance guarantee constraints.*

Proof. Let π be a policy satisfying the minimum-performance guarantee constraints, and let β_i^* satisfy conditions 1–3. Hence, we have

$$
\begin{aligned}
U(\pi) &\le U(\pi) + \sum_{i=1}^{N} (\beta_i^* - 1) \left(\liminf_{T \to \infty} U_i^T(\pi) - C_i \right) \\
&\le \limsup_{T \to \infty} \sum_{i=1}^{N} \beta_i^* U_i^T(\pi) - \sum_{i=1}^{N} (\beta_i^* - 1) C_i.
\end{aligned}
\qquad (20)
$$

By the definition of π^*, we get

$$\sum_{i=1}^{N} \beta_i^* U_i^T(\pi) \le \sum_{i=1}^{N} \beta_i^* U_i^T(\pi^*). \qquad (21)$$

Therefore,

$$
\begin{aligned}
U(\pi) &\le \limsup_{T \to \infty} \sum_{i=1}^{N} \beta_i^* U_i^T(\pi^*) - \sum_{i=1}^{N} (\beta_i^* - 1) C_i \\
&\le U(\pi^*) + \sum_{i=1}^{N} (\beta_i^* - 1) \left(\liminf_{T \to \infty} U_i^T(\pi^*) - C_i \right) \\
&= U(\pi^*),
\end{aligned}
\qquad (22)
$$

where the second part of (22) equals zero because of condition 3 on β_i^*. Similar to the proof of Theorem 1, the properties of lim sup and lim inf are applied here. □

5. IMPLEMENTATION ISSUES

In this section, several implementation issues including parameter estimation and efficient policy search methods will be considered. An optimal algorithm and a low-complexity suboptimal algorithm are developed here for policy search.

5.1. Control parameter estimation

The opportunistic scheduling policies described in Section 4 involve some control parameters to be estimated online: $\vec{v}^*$ in temporal fairness, $\vec{\gamma}^*$ in utilitarian fairness, and $\vec{\beta}^*$ in the minimum-performance guarantee policy. Those parameters

Input:	an $N \times K$ nonnegative matrix $[c_{ik}]$.
Step 1:	initialization:
	(a) Append $(N - K)$ all-zero columns to the matrix.
	(b) In each row, subtract the smallest entry from every entry in that row. In each column, subtract the smallest entry from every entry in that column.
Step 2:	cover all zeros with the minimum number of (horizontal and/or vertical) lines. If the minimum number = N, go to Step 4.
Step 3:	subtract the smallest uncovered entry from every uncovered entry; add it to every intersection of lines. Go to Step 2.
Step 4:	make the assignment at zeros. If any row or column has only one 0, make that assignment. Cross out the corresponding row and column, and move to the next assignment.

ALGORITHM 1: Modified Hungarian algorithm.

are determined by the distribution of performance value matrix $\{\omega^t\}$ and the predetermined constraints. In practice, the distribution is unknown, and hence we need to estimate the control parameters.

In [12], Liu et al. give a practical *stochastic approximation* technique to estimate such parameters. The basic idea is to find the root of a unknown continuous function $f(x)$. We approach the root by adapting the weighted observation error. For example, for user i in temporal fairness scheduling, the base station updates the parameter $\vec{v}^{t+1}$ using a stochastic approximation algorithm

$$v_i^{t+1} = v_i^t - \epsilon^t \left(\sum_{k=1}^{K} \mathbf{1}_{\{A_k^t = i\}} - r_i \right), \tag{23}$$

where, for example, the step size $\epsilon^t = 1/t$. The initial estimate $\vec{v}^1$ can be set to $\vec{0}$ or some value based on the history information.

Using standard methods, it can be shown that $\vec{v}^t$ converges to $\vec{v}^*$ with probability 1 [36]. The computation burden above is $O(N)$ per time slot, where N is the number of users, which suggests that the algorithm is easy to implement online. For our OFDM scheduling schemes, we have found that this stochastic approximation algorithm also works well. For the detailed procedure, we refer the reader to [12].

5.2. Optimal user subset search methods

In our optimal OFDM policies (e.g., in the temporal fairness policy), all the "relative performance values" $(\omega_{i,k}^t + v_i^*)$, denoted c_{ik} for convenience, comprise an $N \times K$ matrix $[c_{ik}]$. Therefore, the operator $\arg\max_{\overline{A}^t}$ is to find an action $\overline{A}^t$ that indicates which K elements in $[c_{ik}]$ have the maximal sum over all K selected elements. This operator is obviously different from the $\arg\max_i$ in [12], which simply returns the index of the largest element from a vector.

It is straight forward to compute the $\arg\max$ if no hard physical limitations are considered. The operator can simply select the largest K elements. However, a common physical constraint is that in any time slot, the scheduler cannot assign two users to the same subcarrier, or two subcarriers

to the same user. Mathematically, at any time slot t, for any two subcarriers j and k, $j \neq k \Leftrightarrow A_j^t \neq A_k^t$. When this physical constraint is considered, the computation of the $\arg\max$ in the optimal policy is nontrivial. A brute-force approach is exhaustively searching over the $\binom{N}{K}$ possible assignments, which obviously has very high computational complexity. Since this optimal user subset search operation should be performed online at each slot, we need to use more efficient algorithms.

It turns out that the problem of computing the $\arg\max$ can be posed as an integer linear program (ILP) [37]:

$$\text{maximize} \sum_{i=1}^{N} \sum_{k=1}^{K} c_{ik} x_{ik} \quad \text{subject to} \sum_{i=1}^{N} x_{ik} = 1,$$
$$k = 1, \ldots, K,$$
$$\sum_{k=1}^{K} x_{ik} \le 1, \quad i = 1, \ldots, N,$$
$$x_{ik} \in \{0, 1\}, \quad c_{ik} \ge 0, \quad N \ge K, \tag{24}$$

where the decision variables x_{ik} indicate which elements to choose, and the weights c_{ik} are relative performance values defined above. This problem is called the *maximal weighted bipartite matching problem* in graph theory, or the *assignment problem* in combinatorial optimization [38].

It is interesting to see that the $\arg\max$ operator in optimal multiuser OFDM scheduling problem can be interpreted as a graph problem (U, S, E, w), where U represents the set of all users, S represents the set of all subcarriers, and E represents the set of all the feasible choices for specific users to select specific subcarriers. Each choice in E is weighted by a function $w(E)$. The problem is to find a matching $M \in E$ for U and S that maximizes the sum of the weights over all edges in M.

The Hungarian algorithm is one of many algorithms that have been devised to solve the assignment problem in polynomial time $(O(N^3))$ when $N = K$ [39]. We modify the Hungarian algorithm to solve our general *unbalanced* $(N \ge K)$ problem here by introducing a number of *slack* variables to convert the ILP problem into standard form. Note that the standard form ILP with the slack variables is algebraically equivalent to the original problem [41]. It is proven in [39]

that the Hungarian algorithm can always find the maximum assignment, that is, it is an optimal solution to this problem.

Algorithm 1 is our modified Hungarian algorithm.

Ideally, the OFDM scheduler should repeat the above procedure at every scheduling slot. However, this still poses a heavy computational burden on the base station. Hence suboptimal algorithms with lower complexity are of interest for practical implementation.

We develop a suboptimal algorithm called "*max-max*" to perform the above arg max operation with much lower complexity. This algorithm is a variation of the "*min-min*" method for task mapping in heterogeneous computing [42]. The basic idea is this: first, find the overall maximal element in the matrix $[c_{ik}]$, then assign the corresponding subcarrier to the corresponding user. Next, remove the newly assigned user-subcarrier pair from the selection table. In other words, the corresponding row and column are removed from the matrix. Continue to repeat the above procedure on the reduced matrix until all subcarriers are assigned. In the simulations in the next section, the suboptimal scheme shows near-optimal performance with a lower complexity.

6. SIMULATION RESULTS

In this section, we present numerical results to illustrate the performance of the various OFDM scheduling schemes developed in this paper. For the purpose of comparison, we also simulate two special scheduling policies. *Round-robin* [43] is a nonopportunistic scheduling policy that schedules users over all subcarriers in a predetermined order. It is simple but lacks flexibility. The round-robin policy can serve as a performance benchmark to measure how much gain results from using our opportunistic scheduling policies. The other policy for comparison is a greedy scheduling scheme that always selects the user with the maximum performance to transmit for each subcarrier at each time slot. The greedy policy will in general violate the QoS/fairness constraints, but provide an upper bound on the system performance. It is used here to expose the tradeoff between the QoS constraints for individual users and the overall system throughput. The more relaxed the fairness constraints, the higher the overall achievable throughput, therefore, the closer to what we will get to the performance of the greedy scheme.

In our simulation, we consider the downlink of a heavy-traffic single-cell OFDM system with fixed 64 subcarriers. There is one base station serving all the users in the cell. Each user suffers from multipath Rayleigh fading with the bad-urban (BU) scenario of the COST 259 channel model [44, 45], and we assume a path-loss exponent of four. Every user is assumed to be stationary or slowly moving so that the maximum Doppler shift is 20 Hz. The performance value, used by different users usually is a nondecreasing function of their SINR, and can be in various forms, such as linear functions, step functions, or S-shape functions. For simplicity, here we take all the performance values as linear functions of users' SINR (in dB). We assume that the physical limitation on scheduling discussed in Section 5.2 applies: at each time slot, no two users can be scheduled on the same subcarrier and each user is scheduled exactly one subcarrier.
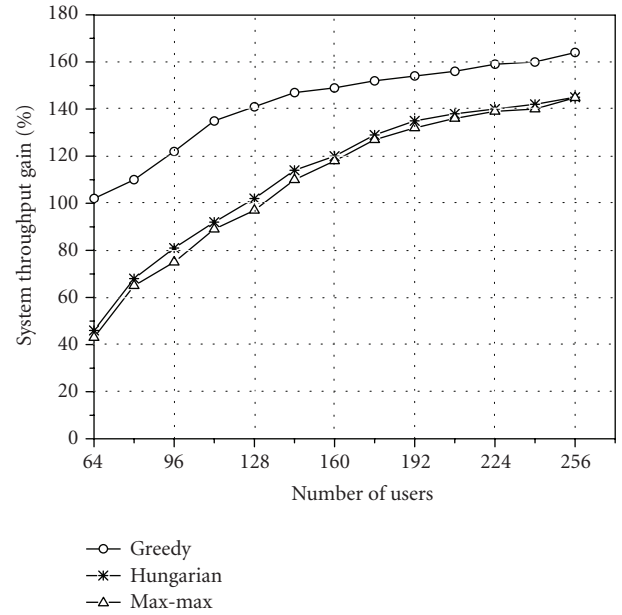

FIGURE 2: System throughput gain in the temporal fairness scheduling.

6.1. Performance gain

First, we assume the locations of all users are distributed uniformly in the cell, and examine the impact of the number of users on the average system throughput. We use the round-robin policy as the baseline, and define the system throughput gain as $(U_S - U_R)/U_R$, where U_S and U_R denote the average system throughput of a given scheduling policy and the round-robin policy, respectively.

Figure 2 shows the system throughput gain relative to round-robin from the different policies in the temporal fairness scheduling simulations. For the purpose of simulation, we assume the time-fraction assignment is done using *fair sharing*, that is, the total resources are evenly divided among the users. Therefore, if there are N users in the cell, we set $r_i = 1/N$ for all users. From Figure 2, it is evident that the system throughput gain increases with the number of users. This is reflective of the *multiuser diversity* gain. For 64 users, our optimal policy (Hungarian) achieves about 46% overall throughput gain, while the greedy policy has an improvement of 101%. This is not surprising since the greedy policy achieves the highest overall performance at the cost of unfairness among the users. The suboptimal policy (max-max) shows surprisingly near-optimal performance. Its performance gap with the optimal policy is less than 1-2%, and even smaller when we increase the number of users.

Figure 3 shows the system throughput gain relative to round-robin from the different policies in the utilitarian fairness scheduling simulations. We also assume *fair sharing* in the throughput-fraction assignment. This means we set $a_i = 1/N$ for all users in an N-user system. As expected, the increasing trend similar to Figure 2 can be also seen here. For 64 users, our optimal policy (Hungarian) achieves about 32% overall throughput gain, while the greedy policy has an
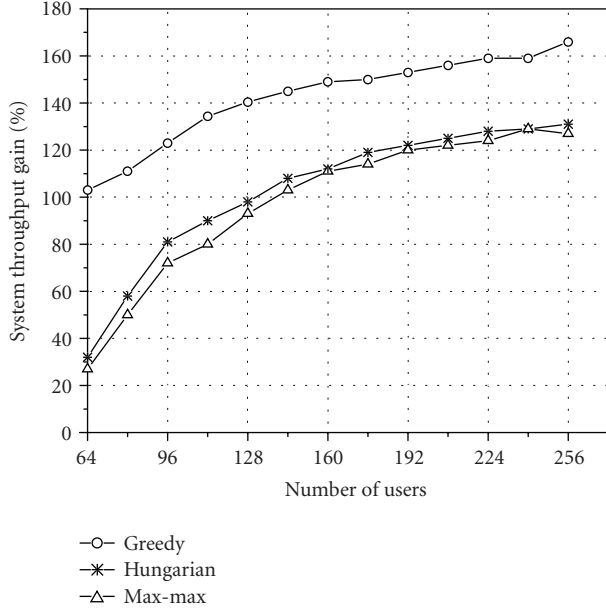
FIGURE 3: System throughput gain in the utilitarian fairness scheduling.

FIGURE 4: System throughput gain in the minimum-performance guarantee scheduling.

improvement of 102%. The suboptimal policy (max-max) also improves the system performance by 27%.

Next, we investigate the performance of the opportunistic scheduling schemes with minimum-performance guarantees. First, we run the simulation for $1,000,000$ time slots using the round-robin policy, where the resource (time) is equally distributed among all users. Then, we compute an average performance value and use it as the minimum-performance requirement for each user. It is easy to see that this minimum-performance requirement vector is feasible. Figure 4 shows the system throughput gain relative to round-robin from the different policies in the minimum-performance guarantee scheduling simulations. For 64 users, our optimal policy (Hungarian) achieves about 31% overall throughput gain, while the greedy policy (which violates the minimum-performance requirements) has an improvement of about 100%. The suboptimal policy (max-max) also performs well with 24% overall gain.

6.2. Fairness

Using the temporal fairness scheduling scenario as an example, we study the fairness among the users by applying the different policies. We use the same single-cell system with 64 subcarriers, and there are 128 users in the system. The users are divided into three "distance" groups. Users 1–48 belong to the "far" group, users 49–80 belong to the "middle" group, and users 81–128 belong to the "near" group. Obviously a user in the "near" group has a much higher probability to get a strong SINR than a user in the "far" group. We set all users to have the same minimum time-fraction requirement. Specifically, each user has a resource (time) requirement $r_i = 2/(3N)$ for an N-user system, where $\sum_i r_i = 2/3 < 1$. Therefore, the system has the freedom to as-
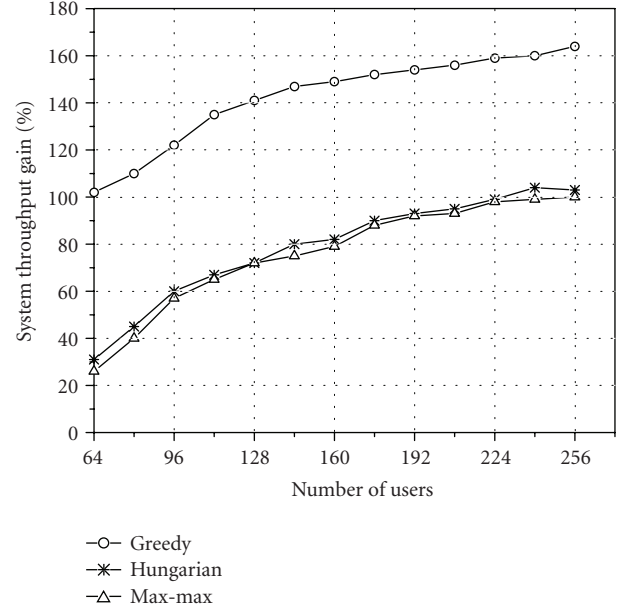
sign the remaining 1/3 portion of the resource to some "better" users (beyond their minimum requirements) to further improve the system performance.

Figure 5 indicates the amount of resource consumed by selected users in the temporal fairness scheduling simulations. The first bar represents that of round-robin, where the resource is equally shared by all users. The second bar represents our optimal policy (Hungarian). The third bar is the greedy policy. The rightmost bar shows the minimum requirements of user. The second bar is higher than the fourth bar for all the users, which indicates that our temporal fairness optimal scheduling policy meets the minimum time-fraction requirements for all users. In the greedy policy, users 1, 16, and 32 get very little resource (far below the minimum requirement line) while users 88, 96, and 128 have very large shares. As expected, the greedy algorithm is heavily *biased* though it achieves the highest overall performance.

In the following, we simply check the fairness among the users with utilitarian fairness and minimum-performance guarantee scheduling. We use the same cellular system and user group settings as temporal fairness.

In Figure 6, we show the average performance values of selected users in the utilitarian fairness scheduling simulations. The preset performance requirements of the selected users 1, 16, 32, 56, 64, 88, 96, and 128 are $[0.001, 0.002, 0.001, 0.003, 0.003, 0.004, 0.005, 0.005]$. The values represent the minimum fraction of overall average performance for individual users.

In Figure 7, we show the average performance values of selected users in the minimum-performance guarantee scheduling simulations. Similar to the previous section, we first run a round-robin simulation, then use the obtained average performance as minimum-performance requirement for each user. From the figure, we see that our optimal
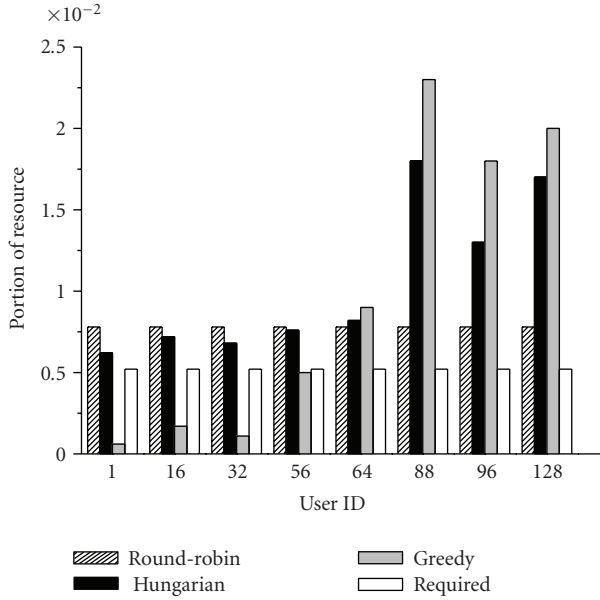
FIGURE 5: Portion of resource shared by users in the temporal fairness scheduling.
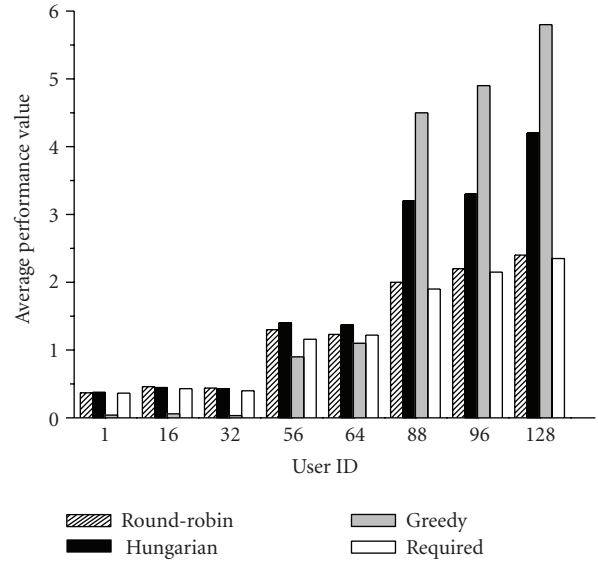

FIGURE 6: User average performance in the utilitarian fairness scheduling.

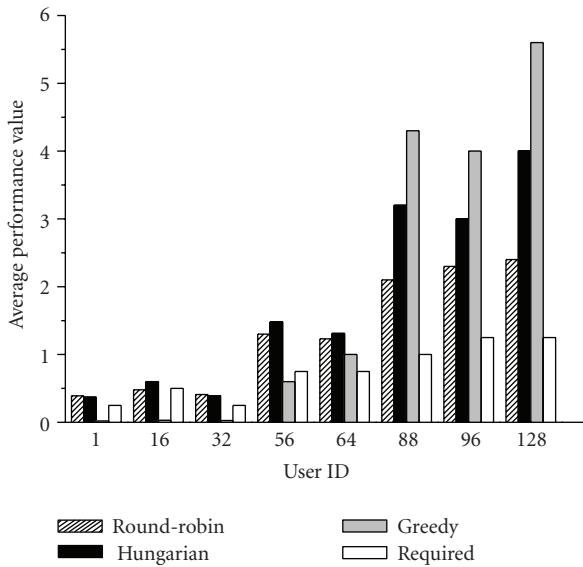FIGURE 7: User average performance in the minimum-performance guarantee scheduling.

scheduling policy (Hungarian) meets all the requirements and outperforms round-robin policy everywhere.

In summary, the simulation results show that using our OFDM opportunistic scheduling policies, the system can achieve significant performance gains over the nonopportunistic round-robin policy while satisfying the various QoS/fairness requirements. Also, the low-complexity suboptimal policy shows near-optimal performance in every scenario.

7. CONCLUSIONS

Opportunistic transmission scheduling is a promising technology to improve spectrum efficiency by exploiting time-varying channel conditions. We investigated the application of opportunistic scheduling in multiuser OFDM systems, which dynamically allocates resource in both temporal and spectral domains. Optimal scheduling policies were presented and proven to be optimal under the temporal fairness, utilitarian fairness, and minimum-performance QoS constraints. We developed optimal and suboptimal algorithms to implement these optimal policies efficiently. The simulation showed that the schemes achieve improvements of about 30%–140% in network efficiency compared with a scheduling scheme that does not take into account channel conditions.

Scheduling problems with multiple mixed QoS/fairness constraints will be interesting to tackle as future work and is definitely of practical interests. For example, a user might ask for both minimum temporal fraction and minimum performance guarantees. Or a user might be constrained by both maximum and minimum requirements of wireless resource. We also plan to investigate the significant feedback overhead involved in assuming perfect channel-state information feedback in OFDM systems, especially in fast fading channels. Scenarios with relatively small numbers of users in the system will also be explored. That means two or more subcarriers could be available for each user. The effects of finite-length data arrival queues or explicit delay requirement for certain users also will be studied. The application of multiple-channel opportunistic scheduling for MAC layer QoS control in cognitive radio systems will be considered in our future work.

ACKNOWLEDGMENTS

REFERENCES

[1] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proceedings of the IEEE International Conference on Communications*, vol. 1, pp. 331–335, 1995.

[2] Qualcomm, "1xEV: 1xEVolution, IS-856 TIA/EIA standard, airlink overview," White Paper Rev. 7.2, Qualcomm, USA, 2001.

[3] S. Parkvall, E. Dahlman, P. Frenger, P. Beming, and M. Persson, "The high speed packet data evolution of WCDMA," in *Proceedings of the IEEE Vehicular Technology Conference*, vol. 3, pp. 2287–2291, 2001.

[4] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[5] C. Peng, H. Zheng, and B. Y. Zhao, "Utilization and fairness in spectrum assignment for opportunistic spectrum access," *Mobile Networks and Applications*, vol. 11, no. 4, pp. 555–576, 2006.

[6] R. Prasad, *Universal Wireless Personal Communications*, Artech House, Boston, Mass, USA, 1998.

[7] B. Bing, Ed., *Wireless Local Area Networks: the New Wireless Revolution*, John Wiley & Sons, New York, NY, USA, 2002.

[8] T. A. Weiss and F. K. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE Communications Magazine*, vol. 42, no. 3, pp. S8–S14, 2004.

[9] J. Chuang and N. Sollenberger, "Beyond 3G: wideband wireless data access based on OFDM and dynamic packet assignment," *IEEE Communications Magazine*, vol. 38, no. 7, pp. 78–87, 2000.

[10] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 10, pp. 2053–2064, 2001.

[11] X. Liu, E. K. P. Chong, and N. B. Shroff, "Transmission scheduling for efficient wireless resource utilization with minimum-performance guarantees," in *Proceedings of the IEEE Vehicular Technology Conference*, vol. 2, pp. 824–828, October 2001.

[12] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Computer Networks*, vol. 41, no. 4, pp. 451–474, 2003.

[13] S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," *IEEE/ACM Transactions on Networking*, vol. 7, no. 4, pp. 473–489, 1999.

[14] T. Ng, I. Stoica, and H. Zhang, "Packet fair queueing algorithms for wireless networks with location-dependent errors," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '98)*, vol. 3, pp. 1103–1111, 1998.

[15] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150–153, 2001.

[16] S. Shakkottai and A. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR," Tech. Rep., Bell Laboratories, USA, 2000.

[17] P. Liu, R. Berry, and M. Honig, "Delay-sensitive packet scheduling in wireless networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, vol. 3, pp. 1627–1632, March 2003.

[18] S. S. Kulkarni and C. Rosenberg, "Opportunistic scheduling policies for wireless systems with short term fairness constraints," in *Conference Record / IEEE Global Telecommunications Conference*, vol. 1, pp. 533–537, December 2003.

[19] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and A. Viterbi, "CDMA/HDR: a bandwidth-efficient high-speed wireless data service for nomadic users," *IEEE Communication Magazine*, vol. 38, no. 7, pp. 70–77, 2000.

[20] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proceedings of the IEEE Vehicular Technology Conference*, vol. 3, pp. 1854–1858, May 2000.

[21] F. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, 1997.

[22] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.

[23] M. Hu, J. Zhang, and J. Sadowsky, "Traffic aided opportunistic scheduling for wireless networks: algorithms and performance bounds," *Computer Networks*, vol. 46, no. 4, pp. 505–518, 2004.

[24] Z. Shen, J. Andrews, and B. Evans, "Optimal power allocation in multiuser OFDM systems," in *Proceedings of the IEEE Global Telecommunications Conference*, vol. 1, pp. 337–341, Las Vegas, Nev, USA, 2003.

[25] H. Viswanathan and K. Kumaran, "Rate scheduling in multiple antenna downlink wireless systems," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 645–655, 2005.

[26] 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 Selected Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, 1999.

[27] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proceedings of the IEEE Vehicular Technology Conference*, vol. 2, pp. 1085–1089, May 2000.

[28] D. Čabrić and R. W. Brodersen, "Physical layer design issues unique to cognitive radio systems," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '05)*, vol. 2, pp. 759–763, September 2005.

[29] G. Song and Y. Li, "Adaptive resource allocation based on utility optimization in OFDM," in *Conference Record / IEEE Global Telecommunications Conference*, vol. 2, pp. 586–590, December 2003.

[30] J. Zhu, B. Bing, Y. Li, and J. Xu, "An adaptive subchannel allocation algorithm for OFDM-based wireless home networks," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '04)*, pp. 352–356, 2004.

[31] Y. Liu and E. Knightly, "Opportunistic fair scheduling over multiple wireless channels," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '98)*, vol. 2, pp. 1106–1115, 2003.

[32] H. Seo and B. G. Lee, "Proportional-fair power allocation with CDF-based scheduling for fair and efficient multiuser OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 978–983, 2006.

[33] C. Li and X. Wang, "Adaptive opportunistic fair scheduling over multiuser spatial channels," *IEEE Transactions on Communications*, vol. 53, no. 10, pp. 1708–1717, 2005.

[34] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks—part I: theoretical framework & part II: algorithm development," *IEEE Transaction on Wireless Communications*, vol. 4, no. 2, pp. 614–634, 2005.

[35] M. Andrews, S. Borst, F. Dominique, et al., "Dynamic bandwidth allocation algorithms for high-speed data wireless networks," Tech. Rep., Bell Laboratories, USA, 2000.

[36] H. J. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*, Springer, New York, NY, USA, 1997.

[37] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Upper Saddle River, NJ, USA, 1982.

[38] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms and Applications*, Prentice-Hall, NJ, USA, 1993.

[39] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistic Quarterly*, vol. 2, pp. 83–97, 1955.

[40] H. L. Royden, *Real Analysis*, Macmillan, New York, NY, USA, 2nd edition, 1968.

[41] E. K. P. Chong and S. H. Żak, *An Introduction to Optimization*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2001.

[42] T. D. Braun, H. J. Siegel, N. Beck, et al., "A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems," *Journal of Parallel and Distributed Computing*, vol. 61, no. 6, pp. 810–837, 2001.

[43] E. L. Hahne and R. G. Gallager, "Round robin scheduling for fair flow control in data communication networks," in *Conference Record—International Conference on Communications*, pp. 103–107, May 1986.

[44] L. M. Correia, *Wireless Flexible Personalized Communications, COST 259: European Co-operation in Mobile Radio Resource*, John Wiley & Sons, New York, NY, USA, 2001.

[45] G. Stüber, *Principles of Mobile Communication*, Kluwer Academic, Norwell, Mass, USA, 2nd edition, 2001.

[46] Z. Zhang, Y. He, and E. K. P. Chong, "Opportunistic downlink scheduling for multiuser OFDM systems," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '05)*, vol. 2, pp. 1206–1212, New Orleans, La, USA, March 2005.