Design of Fair Weights for Heterogeneous Traffic Scheduling in Multichannel Wireless Networks

Mehri Mehrjoo, *Member, IEEE*, Mohamad Khattar Awad, *Member, IEEE*, Mehrdad Dianati, *Member, IEEE*, and Xuemin (Sherman) Shen, *Fellow, IEEE*

Abstract—Fair weights have been implemented to maintain fairness in recent resource allocation schemes. However, designing fair weights for multiservice wireless networks is not trivial because users' rate requirements are heterogeneous and their channel gains are variable. In this paper, we design fair weights for opportunistic scheduling of heterogeneous traffic in orthogonal frequency division multiple access (OFDMA) networks. The fair weights determine each user's share of rate for maintaining a utility notion of fairness. We then present a scheduling scheme which enforces users' long term average transmission rates to be proportional to the fair weights. The proposed scheduler takes the advantage of users' channel state information and the inherent flexibility of OFDMA resource allocation for efficient resource utilization. Furthermore, using the fair weights allows flexibility for realization of different scheduling schemes which accommodate a variety of requirements in terms of heterogeneous traffic types and user mobility. Simulation based performance analysis is presented to demonstrate efficacy of the proposed solution in this paper.

Index Terms—Resource allocation, opportunistic fair scheduling, orthogonal frequency division multiple access (OFDMA).

I. INTRODUCTION

NEW generations of broadband wireless access standards, such as IEEE 802.16, deploy orthogonal frequency division multiple access (OFDMA) mechanism to improve service provisioning and to overcome fading channel impairments. Deliberate resource scheduling of OFDMA networks facilitates different quality of service (QoS) provisioning, efficient utilization of limited resources available at the base station (BS), and maintaining of fairness among users. This objective can be achieved by proper implementation of fair opportunistic scheduling schemes.

A pure opportunistic scheduling scheme allocates resources (i.e., sub-carriers, rate, power) to users with the highest channel gain. This scheme is shown to be throughput-optimal [1]; however, it results in unfair resource allocation [2], when significant discrepancies exist among the average quality of channels for different users. In particular, this scheme allocates the most network resources to the users with strong channel gains, starving the less fortunate users.

Opportunistic fair scheduling schemes for multi-carrier networks have been appeared in the literature, recently. In [3], an opportunistic fair scheduler for code division multiple access (CDMA) networks is proposed. The scheduling process is decoupled into a network throughput maximization process and a fairness process which can ensure probabilistic or deterministic fairness. The probabilistic fairness maintains the differences among users' long-term throughput within a limited range with a bounded probability, and the deterministic fairness guarantees equal long-term throughput among users. However, it remains unclear how the expected differences are defined to achieve fairness. In addition, maintaining equal long-term throughput among users is not desirable and efficient in a scenario with heterogeneous user traffic. An opportunistic scheduler for OFDMA networks, which maintains temporal fairness or "utilitarian" fairness, is introduced in [4]. Whereas in temporal fairness, a certain long-term portion of time is allocated to each user, in "utilitarian" fairness a portion of the overall average throughput is allocated to each user. Temporal fairness criterion maintains resource fairness, i.e., resources are allocated on an equal time duration basis, which does not ensure performance (throughput) fairness in wireless networks. The portion of the overall average throughput that should be allocated to a user is pre-specified. A scheduling scheme based on high data rate (HDR) scheduling is presented for OFDMA networks in [5]. The HDR scheduling is a proportional fair scheduling scheme which has been proposed for scheduling data packets in $CDMA2000 1x$ evolution [6]. Although this scheme attempts to reduce the complexity of the scheduling scheme by clustering the sub-carriers into sub-bands, it is not proved how this scheduler can maintain proportional fairness. Enforcing fairness through weighting factors is formulated in [7] for OFDMA networks. The focus is on proposing a linear complexity approach for sub-carrier assignment and rate allocations problems.

The above scheduling schemes consider fairness provisioning among users with homogeneous rate requirements. However, the problem becomes complicated when users are heterogeneous in terms of traffic requirements and have non-concave utility functions. Resource scheduling for heterogeneous types of traffic needs to be revisited to optimally utilize resources while maintaining fairness among users. In addition, the existing schemes address the complexity of the scheduling

M. Mehrjoo is with the Faculty of Electrical and Computer Engineering, University of Sistan and Baluchestan, Iran (e-mail: mehrjoo@ece.usb.ac.ir).

M. K. Awad is with the American University of Kuwait, Hawalli 32056 Kuwait, P.O.Box 44129 (e-mail: mohamad@ieee.org).

M. Dianati is with the Faculty of Engineering and Physical Sciences, University of Surrey, UK, GU2 7XH (e-mail: m.dianati@surrey.ac.uk).

X. (S.) Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, N2L 3G1 (e-email: xshen@bbcr.uwaterloo.ca).

for multi-carrier networks assuming that the fair weights for different traffic types are known; however, finding proper weights is a non-trivial open problem that is tackled in this paper.

We design fair weights and accordingly propose an opportunistic fair scheduling of heterogeneous traffic. We design the fair weights for the users with different service rate (more precisely, utility) requirements, where the fair weights represent the fair proportions among long term achieved rates of users taking into account users' channel gain differences and heterogeneity of their data traffic. The fair weights are then used for opportunistic fair scheduling in the downlink of an OFDMA wireless network. The proposed scheduling scheme intends to maintain long-term fair allocation of sub-carriers and transmission power according to the weighting factors. In specific, we propose a modular scheduler consisting of a fairness module and a resource allocation module to separate fair weight computation from the resource allocation part. The fairness module executes a fairness scheme that generates a set of fair weights associated with users. These weights are periodically computed and fed into the resource allocation module to reduce the complexity of the computation inside the scheduling module which needs to make fast carrier and power allocation decisions. The resource allocation module allocates OFDMA sub-carriers and power based on users' instantaneous channel gains and fair weights. The fairness module performs in parallel with the resource allocation module, which allows for simple and fast scheduling. Numerical results show the proposed scheduling scheme is effective in maintaining fairness and achieving multi-user diversity gain for time variant channel and users with heterogeneous service demands.

The remainder of the paper is organized as follows. The system model and mathematical formulations of the OFDMA resource allocation and fairness modules are presented in Section II. Formulated optimization problems for the scheduling are solved in Section III. Numerical results are given and discussed in Section IV. Finally, the concluding remarks are given in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATIONS

The proposed scheduler consists of a fairness module and an OFDMA resource allocation module with N_u users and N_c sub-carriers, as shown in Fig. 1. The notations i and j are user and sub-carrier indexes, respectively.

The fairness module includes the Fair Weight block, which computes the set of fair weights, W_i, and the Transmission History block, which computes the moving averages of transmitted rates, R_i, for $i = 1, ..., N_u$. The Fair Weight block computes the fair weights, based on a fairness criteria, users' large scale channel state information (CSI), and users' utilities. The Transmission History block updates the exponentially weighted moving average (EWMA) of transmitted rate to user i, R_i, at the beginning of each scheduling interval n as given by

$$R_i(n) = (1 - \frac{1}{T_c})R_i(n-1) + (\frac{1}{T_c})r_i(n-1), \quad (1)$$

where r_i is the transmitted rate to user i, and T_c is a time constant that determines the rate of exponential decay of the
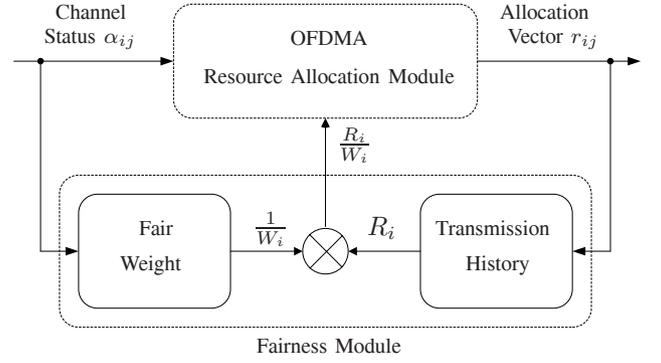

Fig. 1. Architecture of the proposed scheduler.

impacts of old samples. A larger T_c results in rapid decay, and the converse is true. In addition, T_c can be considered as the width of the moving average window so that when T_c is large, averaging is performed over a large numbers of scheduling intervals [6]. EWMA emphasizes more on the recent data by definition. This technique is desirable in the sense that the fairness scheme attempts to compensate for unfairness of recent allocations as much as possible.

The OFDMA resource allocation module determines allocated rates to users based on the values of a_{ij} and $\frac{R_i}{W_i}$. Diversity gain is achieved by taking into account users' instantaneous CSI, a_{ij}, and fairness is compensated for by considering $\frac{R_i}{W_i}$ as a measure of fairness deficit. When $\frac{R_i}{W_i}$ is close to 1, the average transmitted rate to user i is close to its fair value, determined by W_i. On the other hand, $\frac{R_i}{W_i} << 1$ or $\frac{R_i}{W_i} >> 1$ mean starvation and overallocation of user i, respectively. In either case, the scheduler should compensate for the occurred unfairness in the upcoming scheduling intervals.

We formulate the OFDMA resource allocation problem and the fair weight design problem as two separate optimization problems. The OFDMA resource allocation, described in subsection II-A, is an optimization problem where its objective function and constraints model the scheduling scheme and OFDMA specifications. Similarly, we present an optimization problem that considers users' heterogeneous rate requirements and CSI to compute proportional fair weights in subsection II-B.

A. OFDMA Resource Allocation

We consider the downlink of a single BS and multiple users located in one hop neighborhood from the BS. Users' backlogged traffic, buffered in separate queues at the BS, is scheduled at the beginning of each downlink frame consisting of N_s OFDM symbols. The BS assigns OFDM sub-carriers to users and allocates a fraction of its power, P_{BS}, to each sub-carrier at each scheduling instance. Relevant system parameters in our model are defined in Table I.

Without loss of generality, we assume that the spectral density of noise and sub-carriers bandwidth are equal to one. Thus, the maximum achievable rate to user i on sub-carrier j of symbol n, denoted by r_{ijn}, is

$$r_{ijn} = \log_2\left(1 + \alpha_{ijn}p_{ijn}\right). \quad (2)$$

TABLE I
LIST OF SYMBOLS

Symbol	Description
N_u	number of users in the network
N_c	number of OFDMA sub-carriers
N_s	number of OFDMA symbols in the downlink frame
N_f	number of OFDMA symbols in fair weights computations
i	user index belongs to $\mathcal{N}_u := \{1, 2, ..., N_u\}$
j	sub-carrier index belongs to $\mathcal{N}_c := \{1, 2, ..., N_c\}$
n	symbol index belongs to $\mathcal{N}_s := \{1, 2, ..., N_s\}$
R_i	average transmitted rate to user i
W_i	fair weight of user i
P_{BS}	the BS total power budget
α_{ijn}	channel gain of user i on sub-carrier j of OFDMA symbol n
p_{ijn}	required power by user i on sub-carrier j of OFDMA symbol n to transmit r_{ijn}
r_{ijn}	achievable rate by user i on sub-carrier j of OFDMA symbol n

Total allocated power to the sub-carriers of each OFDMA symbol is limited by P_{BS}, i.e.,

$$\sum_{i=1}^{N_u} \sum_{j=1}^{N_c} p_{ijn} \leq P_{BS} \quad \forall n \in \mathcal{N}_s. \tag{3}$$

Implementation of OFDMA requires exclusive allocation of a sub-carrier to a single user. This constraint can be represented by

$$r_{\hat{i}jn} \cdot r_{ijn} = 0 \quad \forall \hat{i} \in \mathcal{N}_u, i \neq \hat{i}, \; \forall j \in \mathcal{N}_c, \; \forall n \in \mathcal{N}_s. \tag{4}$$

Constraint (4) implies that if sub-carrier j is assigned to user $\hat{i}$, i.e., $r_{\hat{i}jn} \neq 0$, the allocated rate to every other user on sub-carrier j of OFDMA symbol n must be zero.

The trade-off between throughput and fairness can be tuned by the objective function. As different notions of fairness in long or short term basis are expected in practice, the objective function is defined in such a way to be adjusted for different fairness demands. We consider $\frac{R_i}{W_i}$ as the fairness tuning term in the objective function. Recall that W_i is the fair allocation rate to user i, and R_i is the moving average of the transmitted rate to user i. Depending on the length of the averaging window, T_c, the scheduler reaction time to the unfairness can be tuned. For a short average window length, R_i approaches zero very fast, if user i does not receive any service rate for multiple consecutive scheduling intervals. In other words, a short average window length forces the scheduler to strive for a short term fairness; so, the scheduler rate allocation policy will be in favor of fairness provisioning rather than throughput maximization. On the other hand, a long average window length for R_i allows the scheduler to compensate for the fairness in longer time and take advantage of users' channel diversity to improve the overall system throughput. Thus, an opportunistic fair scheduler aims to find a rate allocation set,

$\{r_{ijn}\}$, for $i = 1, ..., N_u$, $j = 1, ..., N_c$, and OFDM symbols $n = 1, ..., N_s$, such that

$$\max \sum_{n=1}^{N_s} \sum_{j=1}^{N_c} \sum_{i=1}^{N_u} \left(\frac{r_{ijn}}{\frac{R_i}{W_i}} \right). \tag{5}$$

Accordingly, the probability of assigning sub-carrier j to user i increases when the achievable transmission rate of user i on sub-carrier j is high or the average transmitted rate to user i is smaller than its fair weight. Different degrees of performance trade-off between throughput and fairness can be obtained and optimized [8] in this scheme, which is out of the scope of this paper.

The objective function (5) along with constraints (2), (3), and (4), represent opportunistic fair scheduling as an optimization problem, denoted by $(P1)$ in the following:

$$P_1: \quad \max_{r_{ijn}} \sum_{n=1}^{N_s} \sum_{j=1}^{N_c} \sum_{i=1}^{N_u} \left(\frac{r_{ijn}}{\frac{R_i}{W_i}} \right) \tag{6}$$

$$\text{s.t} \quad \sum_{i=1}^{N_u} \sum_{j=1}^{N_c} \frac{2^{r_{ijn}} - 1}{\alpha_{ijn}} \leq P_{BS} \quad \forall n \in \mathcal{N}_s, \tag{7}$$

$$r_{\hat{i}jn} \cdot r_{ijn} = 0 \quad \forall \hat{i} \in \mathcal{N}_u, i \neq \hat{i}, \tag{8}$$

$$\forall j \in \mathcal{N}_c, \; \forall n \in \mathcal{N}_s, \tag{9}$$

$$r_{ijn} \geq 0 \quad \forall i \in \mathcal{N}_u, \; \forall j \in \mathcal{N}_c, \; \forall n \in \mathcal{N}_s. \tag{10}$$

Problem P_1 is solved in each scheduling interval to obtain allocated rate to users on all sub-carriers for each OFDMA symbol. Since CSI may not be received accurately at the BS, we have proposed an approach in [9] to account for CSI inaccuracy resulted from estimation error and feedback delay. In practice, providing CSI of each sub-carrier over all symbols of each scheduling interval results in large messaging overhead on the reverse feedback channel. Because of the correlation among CSI of a sub-carrier over consecutive symbols, the CSI of each sub-carrier is assumed to be constant for all symbols over a scheduling interval. Accordingly, index n representing symbols of each scheduling interval can be dropped; thus, P1 can be reduced to a simpler optimization problem, denoted by P_2:

$$P_2: \quad \max_{r_{ij}} \sum_{j=1}^{N_c} \sum_{i=1}^{N_u} \left(\frac{r_{ij}}{\frac{R_i}{W_i}} \right) \tag{11}$$

$$\text{s.t} \quad \sum_{i=1}^{N_u} \sum_{j=1}^{N_c} \frac{2^{r_{ij}} - 1}{\alpha_{ij}} \leq P_{BS}, \tag{12}$$

$$r_{\hat{i}j} \cdot r_{ij} = 0 \quad \forall \hat{i} \in \mathcal{N}_u, i \neq \hat{i}, \; \forall j \in \mathcal{N}_c, \tag{13}$$

$$r_{ij} \geq 0 \quad \forall i \in \mathcal{N}_u, \; \forall j \in \mathcal{N}_c, \tag{14}$$

where r_{ij} represents allocated rate to user i on sub-carrier j of all symbols in each scheduling interval.

B. Fair Weight Design

Several factors should be taken into account for designing fair weights. First, users' average channel status often varies with time, so, the resource allocation should be frequently updated. Second, users' resource requirements depend on their traffic types. Third, fairness criteria are not unique, and
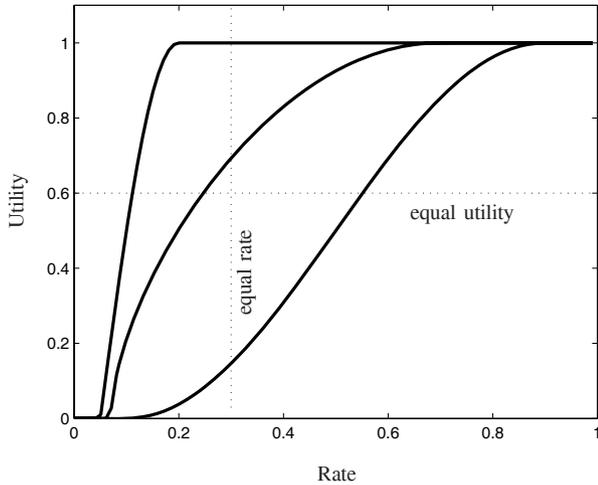
Fig. 2. Comparison between equal rate and equal utility allocation.

different fairness criteria, such as *proportional*, $\alpha-fair$, or *maxmin* fairness [10], [11] can be applied to users' rates or utilities.

Wireless channel suffers from fast and slow channel variations. Fast variations are highly unpredictable, so they are not considered in a long term resource allocation scheme. We look at the long trend of wireless channel, happened over multiple frames, and design the fair weights accordingly. These fair weights are used in resource allocation for the next multiple frames. However, to adapt to the channel variations, the fair weights are periodically updated, dependent to channel statistics.

To allocate resources based on users' traffic types, utility-based allocation can be employed. Fig. 2 shows the utilities of three different applications. The dotted line, labeled "equal rate", illustrates that equal rate allocation does not provide equal user satisfaction. On the other hand, equal allocation of utilities, which is interpreted as equal users' satisfaction, utilizes the network resources more efficiently [12]. Thus, we consider utility fairness instead of rate fairness.

The fair weights are determined based on the notion of utility proportional fairness where the allocated resources are proportional to the users' demands. Consider a bounded set of N_u users' feasible utility subset U_k, $\mathcal{U} = \{U_k | U_k = \{u_{k1}, u_{k2}, ..., u_{kN_u}\}\}$, where u_{ki} is the user i's utility. Utility proportional fairness is defined as follows [13].

Definition 2.1: A set of utilities $U_x = \{u_{x1}, u_{x2}, \ldots, u_{xN_u}\}$ is *utility proportional fair* if for any feasible utility set $U_y = \{u_{y1}, u_{y2}, \ldots, u_{yN_u}\}$, the sum of proportional changes in their utilities is non-positive, i.e.,

$$\sum_{i=1}^{N_u} \frac{u_{yi}(r_{yi}) - u_{xi}(r_{xi})}{u_{xi}(r_{xi})} \leq 0. \tag{15}$$

Note that u_{ki} is a function of allocated rate. Therefore, U_k is utility proportional fair if the set of rates $\{r_{k1}, r_{k2}, \ldots, r_{kN_u}\}$ is found to satisfy (15). A straightforward way to obtain a utility proportional fair allocation $U_k \in \mathcal{U}$ is to maximize $\sum_i \log(u_{ki})$ over the convex set of

feasible allocations $\mathcal{U}$:

$$\max_k \quad \mathcal{F} = \sum_i \log(u_{ki}). \tag{16}$$

Accordingly, a set of rates, denoted by $\{w_{ij}\}$, which is utility proportional fair can be obtained by solving the following optimization problem:

$$P_3: \quad \max_{w_{ij}} \quad \mathcal{F} \tag{17}$$

$$\text{s.t} \quad \sum_{i=1}^{N_u} \sum_{j=1}^{N_c} \frac{2^{w_{ij}} - 1}{g_{ij}} \leq P_{BS}, \tag{18}$$

$$w_{ij} \geq 0 \quad \forall i \in \mathcal{N}_u, \quad \forall j \in \mathcal{N}_c. \tag{19}$$

Problem P_3 has a power constraint similar to (2) and (3), except g_{ij}, very slow varying channel gains[1], are replacing slow or fast varying channel gains α_{ijn}. Also, the exclusive sub-carrier assignment restriction, constraint (4), is relaxed because this problem is solved for fair weights regardless of specific sub-carrier assignments, i.e., the sub-carriers can be shared. Once the fair rate allocations, w_{ij}, are computed, they are valid for $N_f \geq N_s$ symbols, where the value of N_f depends on the frequent g_{ij} variations. As fairness is maintained per user basis, w_{ij} are summed up over the number of sub-carriers, N_c. Therefore, the utility proportional fair rate allocation to user i, i.e., the user i's fair weight is:

$$W_i = \sum_{j=1}^{N_c} w_{ij}. \tag{20}$$

The fair weights are normalized to represent the rate fractions that users should receive over a long time with respect to the other users, i.e., $\sum_{i \in \mathcal{N}_u} W_i = 1$. The OFDMA resource allocation module may allocate more or less rate than the fair rate to each user in each scheduling instance. However, it attempts to maintain the following equalities over a long time [14]:

$$\frac{R_1}{W_1} = \frac{R_2}{W_2} = \cdots = \frac{R_{N_u}}{W_{N_u}}. \tag{21}$$

If the scheduler allocates the available resources to users such that the set of aggregate transmitted rates to users is proportional to the set of fair weights, W_i, i.e., equation (21) is satisfied, the scheduling scheme is utility proportional fair.

III. OFDMA RESOURCE ALLOCATION AND FAIR WEIGHT DESIGN SOLUTIONS

Problem P_2 needs to be solved in every scheduling interval, while P_3 is solved only when its input parameters are changed. Problems P_2 and P_3 are non-convex optimization problems in general, and finding their optimal solutions is nontrivial [15]. Problem P_2 is non-convex because of its discrete feasible region, while P_3 is non-convex because of the non-convex utility functions in the objective function. The efficiency in solving a non-convex problem strongly depends on how non-convexity of the problem is treated. Therefore, we apply

[1]A channel with pathloss and shadowing effects is considered as a very slow varying fading channel, when multipath and Doppler effects are added to the aforementioned effects, the channel is considered as a slow or a fast fading channel.

two different approaches to treat the non-convexity of each problem:

- First, we use a Lagrange dual decomposition method to solve P_2. The method does not guarantee an optimal solution, but it can efficiently obtain near optimal solution(s) with a practical number of sub-carriers [16]. The adaptation of Lagrange dual decomposition method hinges on the results reported in [17] that the duality gap[2] vanishes as the number of sub-carriers increases.

- Second, an interior point method is applied to solve P_3 because the objective function is the sum of users' utilities which can be non-linear functions of users' rates, and interior point methods can efficiently solve non-linear optimization problems [18].

A. Solution of the OFDMA Resource Allocation Problem P_2

If $\delta_i = W_i/R_i$, the objective function of problem P_2 is a maximization of $\sum_{i=1}^{N_u} \left(\delta_i \sum_{j=1}^{N_c} r_{ij} \right)$. Constraints (13) and (14) form the domain $\mathcal{D}$ over which the Lagrangian of P_2 can be defined as

$$\mathcal{L}\left(\{r_{ij}\},\lambda\right) = \sum_{i=1}^{N_u}\sum_{j=1}^{N_c} \delta_i r_{ij} - \lambda\left(\frac{2^{r_{ij}}-1}{\alpha_{ij}} - P_{BS}\right), \quad (22)$$

where λ is the Lagrange multiplier. The dual problem of P_2, can be expressed as

$$\min_{\lambda}\ \max_{\{r_{ij}\}\in\mathcal{D}}\ \mathcal{L}\left(\{r_{ij}\},\lambda\right). \quad (23)$$

From the solution of the dual problem, the set of rate allocations r_{ij}, $\forall i$ & j, can be determined. The optimization problem (23) is a minimization problem with one scalar variable λ that can be solved by an iterative algorithm (Algorithm 1). In each iteration of Algorithm 1, the set of r_{ij} that maximizes $\mathcal{L}$ is determined by solving N_c decomposed problems of rate allocation to sub-carriers. As allocation of sub-carriers to users are independent, the optimization problems (24) can be solved in parallel to obtain allocated rate to sub-carriers.

$$\max_{\{r_{ij}\}\in\mathcal{D}}\ \sum_{i=1}^{N_u} \delta_i r_{ij} - \lambda\left(\frac{2^{r_{ij}}-1}{\alpha_{ij}}\right)\ \forall j = 1\cdots N_c. \quad (24)$$

When adaptive modulation is used, allocated number of bits to each sub-carrier is a discrete variable that can be chosen from the bit loading vector of the modulation technique [19]. Accordingly, the solution of problem (24) is determined by searching over the domain $\mathcal{D}$. The search is performed in real-time because the size of the domain $\mathcal{D}$ is confined by the number of modulation levels, users, and sub-carriers.

B. Solution of the Fair Weight Design Problem P_3

For notational simplicity, a solution of P_3 is denoted by a rate allocation vector $\mathbf{w}$:

$$\mathbf{w} = [w_{11}, w_{12}, \ldots, w_{1N_c}, \ldots, w_{N_u 1}, \ldots, w_{N_u N_c}]^T, \quad (25)$$

where w_{ij} represents allocated rate to user i on sub-carrier j and $w_i = \sum_{j=1}^{N_c} w_{ij}$ is allocated rate to user i. We form a vector $\mathbf{c}(\mathbf{w})$ of the inequality constraints (18) and (19), and

---

[2]The difference between the primal optimal and dual optimal solutions

---

Algorithm 1 Solution algorithm for the dual problem of P_2

Input: N_u, N_c, P_{BS}, α_{ij}, δ_i, *bit loading set*
Result: r_{ij}
begin
  Setting up and initialization:
  Set $h = 1$, $\epsilon = 1$, $Exit\_flag = 1$, $\lambda_{h-1} = \lambda_h = 0$.
  Solve (24) for r_{ij}.
  Compute $\Delta p = P_{BS} - p_{ij}$.
  if $\Delta p > 0$ **then**
    return r_{ij}.
  else
    while $Exit\_flag > 1e-5$ **do**
      if $\Delta p > 0$ **then**
        $\epsilon = 0.99 * \epsilon$.
        $\lambda_h = \lambda_{h-1}$.
        $\Delta p_h = \Delta p_{h-1}$.
      else
        $\lambda_{h-1} = \lambda_h$.
        $\Delta p_{h-1} = \Delta p_h$.
      end
      $\lambda_h = \lambda_h + |\epsilon * \Delta p|$.
      Solve (24) for r_{ij}.
      Update Δp.
      $Exit\_flag = \lambda_h - \lambda_{h-1}$.
      $h = h + 1$.
    end
  end
  return r_{ij}.
end

---

convert the inequality constraints to equality constraints by associating a positive slack variable to each constraint. Denote the $(2N_u + 1)N_c$ vector of slack variables by $\mathbf{s}$. Hence, P_3 is converted to the following minimization problem:

$$P_4: \quad \min_{\mathbf{w}} \ -\sum_i \log(u_{ki}(\mathbf{w})) \quad (26)$$

$$\text{s.t} \quad \mathbf{c}(\mathbf{w}) - \mathbf{s} = 0, \quad (27)$$

$$\mathbf{s} \geq 0. \quad (28)$$

To find an approximation for a local optimum of a nonlinear problem, the interior point algorithm solves a series of perturbed Karush-Kuhn-Tucker (KKT) conditions of P_4:

$$\nabla \mathbf{u}(\mathbf{w}) - \mathbf{A}^T(\mathbf{w})\mathbf{z} = 0, \quad (29)$$

$$\mathbf{c}(\mathbf{w}) - \mathbf{s} = 0, \quad (30)$$

$$\mathbf{S}\mathbf{z} = \mu\mathbf{e}, \quad (31)$$

$$\mathbf{s} \geq 0, \quad \mathbf{z} \geq 0, \quad (32)$$

with $\mathbf{e} = (1, 1, ..., 1)^T$ and $\mu > 0$. In the perturbed KKT conditions, $\mathbf{S}$ is a diagonal matrix with diagonal elements given by vector $\mathbf{s}$, and vector $\mathbf{z}$ contains $(2N_u+1)N_c$ Lagrange multipliers used in the definition of the Lagrangian function of P_4:

$$\mathcal{L}(\mathbf{w}, \mathbf{s}, \mathbf{z}) = \log(\mathbf{u}(\mathbf{w})) - \mathbf{z}^T(\mathbf{c}(\mathbf{w}) - \mathbf{s}). \quad (33)$$

The matrix $\mathbf{A}$ in (29) is the Jacobian matrix of $\mathbf{c}(\mathbf{w})$. Interior point methods begin with an initial interior point in the feasible region that satisfies perturbed KKT conditions for some μ and proceeds to find another interior point that satisfies perturbed KKT conditions for a smaller value of μ. As the algorithm evolves, μ decreases, and consequently the solution of the perturbed KKT conditions approaches the solution of the KKT

conditions, where $\mu = 0$. It is expected that after several iterations the solution will converge to a point that satisfies the KKT conditions of the problem [18].

In each iteration of the interior point method, the directions and lengths of steps from one interior point to another are updated based on the first and second order gradients of objective function and constraints. At each iteration, step direction for each of the variables **w**, **s**, and **z**, i.e., $\mathbf{b} = [\mathbf{b_w}, \mathbf{b_s}, \mathbf{b_z}]^T$, are computed by solving the following linear system of equations:

$$\begin{pmatrix} \nabla^2_{\mathbf{ww}}\mathcal{L} & 0 & -\mathbf{A}^T(\mathbf{w}) \\ 0 & \mathbf{Z} & \mathbf{S} \\ \mathbf{A}(\mathbf{w}) & -\mathbf{I} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{b_w} \\ \mathbf{b_s} \\ \mathbf{b_z} \end{pmatrix} = \\ \begin{pmatrix} \nabla_{\mathbf{w}}\mathbf{u}(\mathbf{w}) - \mathbf{A}^T(\mathbf{w})\mathbf{z} \\ \mathbf{Sz} - \mu\mathbf{e} \\ \mathbf{c}(\mathbf{w}) - \mathbf{s} \end{pmatrix}, \quad (34)$$

Here, **Z** denotes the diagonal matrix whose diagonal elements are given by vector **z**. After obtaining step directions, the length of step in each direction, step length, denoted with α_s^{max} and α_z^{max}, are specified as:

$$\boldsymbol{\alpha_s}^{max} = \max\left\{\alpha \in (0,1] : \mathbf{s} + \alpha\mathbf{b_s} \geq (1-\tau)\,\mathbf{s}\right\}, \quad (35)$$
$$\boldsymbol{\alpha_z}^{max} = \max\left\{\alpha \in (0,1] : \mathbf{z} + \alpha\mathbf{b_z} \geq (1-\tau)\,\mathbf{z}\right\}, \quad (36)$$

where $\tau \in (0,1)$. A small value of τ forces **s** and **z** to approach zero very quickly, so a large value of τ close to one, e.g., $\tau = 0.995$, is usually chosen. The new interior point, slack variables, and Lagrange multipliers, $(\mathbf{w}^+, \mathbf{s}^+, \mathbf{z}^+)$, are determined with the information of step directions and step lengths accordingly:

$$\mathbf{w}^+ = \mathbf{w} + \alpha_s^{max}\mathbf{b_w}, \quad (37)$$
$$\mathbf{s}^+ = \mathbf{s} + \alpha_s^{max}\mathbf{b_s}, \quad (38)$$
$$\mathbf{z}^+ = \mathbf{z} + \alpha_z^{max}\mathbf{b_z}. \quad (39)$$

For the next iteration, μ is updated to a smaller value, $\mu^+ < \mu$, via a linear method:

$$\mu^+ = \sigma\mu \qquad \sigma \in (0,1). \quad (40)$$

Since $\sigma < 1$, μ approaches to zero over several iterations. However, choosing a very small σ or a very large σ will cause faster or slower convergence, respectively. Although fast convergence is always desired, it may force some algorithm parameters, such as **s** and **z**, to approach zero too quickly, which degrades the performance of the algorithm, e.g., the offered solution may be infeasible or far from optimality.

The interior point algorithm is terminated when a stopping criterion is satisfied. In this work, an initial value of $\mu_0 = 1$ has been chosen, and when μ approaches a very small value or the change in allocated weight vector, **w**, is negligible, the algorithm stops. Algorithm 2 presents a summary of the interior point algorithm used in our simulation.

C. Complexity of Proposed Approach

The decomposition of (23) into N_c equations (24) reduces the problem's exponential complexity to a linear one in terms of N_c [17]. The solution of (24) is obtained by a heuristic search method because of the non-convexity of the domain $\mathcal{D}$. The size of $\mathcal{D}$ is confined by the number of modulation

---

Algorithm 2 The interior point algorithm for P_4

Input: $N_u, N_c, P_{BS}, B, \alpha, \textit{users' utilities}, \boldsymbol{w_{initial}}, \boldsymbol{s_0}, \mu_0, \tau, \sigma$
Result: **w**
begin
  Setting up and initialization:
  Choose $\boldsymbol{w_{initial}}$ and compute $\mathbf{s_0} > 0$.
  Choose $\mu_0 > 0$ and compute $\mathbf{z_0} > 0$ accordingly.
  Set parameters $\tau \in (0,1)$ and $\sigma \in (0,1)$.
  Set $k = 0$.
  while $Exit\_flag == 0$ **do**
    Solve (34) to obtain step direction $\mathbf{b} = [\mathbf{b_w}, \mathbf{b_s}, \mathbf{b_z}]^T$.
    Compute $\boldsymbol{\alpha_s}^{max}$, and $\boldsymbol{\alpha_z}^{max}$ using (35) and (36).
    Compute $(\mathbf{w}^{k+1}, \mathbf{s}^{k+1}, \mathbf{z}^{k+1})$ using (37) to (39).
    Set $\mu^{k+1} \leftarrow \mu^k$ and $k \leftarrow k+1$.
    Compute $Exit\_flag$.
  end
  return w.
end

---

levels, users, and sub-carriers, denoted by Q, N_u, and N_c, respectively. In each iteration of the *while* loop in Algorithm 1, a set with QN_u size is searched for rate allocation to each sub-carrier. If the *while* loop requires N_{while} iterations to converge, then the set of equations (24) will be solved in $N_c Q N_u N_{while}$ iterations. Whereas solving the equation (23) with an exhaustive search requires searching over a set of size $(QN_u)^{N_c}$.

Problem P_3 is required to be solved only when the network characteristics, such as users' average channel gains or the number of admitted users to the network, change. The scheduling scheme starts with default fair weights, e.g., all equal to one, and updates the fair weights with the ones obtained by solving P_3 during the first iteration of the scheduling scheme.

IV. NUMERICAL RESULTS

Performance of the opportunistic fair scheduling scheme is evaluated in this section. Performance metrics are throughput and fairness index which are compared with those of a pure opportunistic and a proportional fair scheduling scheme. In addition, the scheduling scheme performance for diverse channel status and utility functions are investigated.

To compare the performance in terms of fairness, a fairness metric needs to be defined first. Gini fairness index, which is an inequality measure of resource sharing, measures deviation from equations (21) for each scheduler. Let the total allocated rate to user i over the simulated intervals be symbolized by $\tilde{R}_i$. We examine the inequality among the set of proportions $\mathbf{v} = \{v_i \mid v_i = \tilde{R}_i/W_i\}$ by Gini fairness index, G_{FI}, defined as follows:

$$G_{FI} = \frac{1}{2M^2\bar{\mathbf{v}}} \sum_{x=1}^{N_u} \sum_{y=1}^{N_u} |\mathbf{v}_x - \mathbf{v}_y|, \quad (41)$$

where

$$\bar{\mathbf{v}} = \frac{\sum_{i=1}^{N_u} \mathbf{v}_i}{N_u}. \quad (42)$$

The Gini fairness index takes a value between 0 and 1. A rate allocation is perfectly fair if $G_{FI} = 0$. A high value of G_{FI} indicates higher unfairness among the proportions.

The wireless channel is simulated to experience both frequency selective and large-scale fading [20], [21]. The users

receive six Rayleigh distributed multipath signals [22]–[24]. The real and imaginary components of the received signals to different users are generated from an uncorrelated multidimensional Gaussian distribution with zero mean and an identity covariance matrix. The large-scale fading is distance dependent and follows the inverse-power law [20]:

$$|\gamma_{ij}|^2 = D_i^{-\kappa}|\alpha_{ij}|, \tag{43}$$

where D_i is the distance between the BS and user i in meters, κ is path loss exponent, and γ_{ij} is path loss of user i on sub-carrier j. The numerical values of the wireless channel used in the simulation are: Doppler frequency= 30 Hz, and $\kappa = 2$.

The network supports users with non-concave and concave utility functions, respectively. The users' utility functions are expressed by equation (44) [25], where **r** denotes allocated rate to the user, $l1$ and $l2$ are lower and upper rate thresholds, and k controls the convexity of the utility function. The function is concave for $k < 1$ and convex for $k > 1$. ($k = 0.7$, $l1 = 1$, $l2 = 800$) and ($k = 2$, $l1 = 10$, $l2 = 600$) have been chosen for concave and non-concave utility functions, respectively.

$$Util_i(r_i) = \begin{cases} 0 & r_i \leq l_1, \\ sin^k\left(\frac{\pi}{2}\frac{r_i-l_1}{l_2-l_1}\right) & l_1 < r_i \leq l_2, \\ 1 & r_i > l_2. \end{cases} \tag{44}$$

The simulated network consists of the BS, with total power equals to 20 Watt, located at the center of the cell with 800m radius, that transmits accumulated traffic in its queues to the users over 64 sub-carriers. The value of $T_c = 1000$ (symbols) is chosen in this work, which is equivalent to 100 frames or 1000 symbols in IEEE 802.16[3] standard (downlink length= $1msec$, symbol length= $80\mu sec$ for $5MHz$ channel) [26].

We implement the performance evaluation in two steps. First, we compare the opportunistic fair scheme with a pure opportunistic and a proportional fair scheme when the users are randomly distributed in the network and have the same concave utility functions. Then, we investigate the opportunistic fair scheduling scheme performance for different scenarios where users have different pathloss, mobility, and utility functions.

Fig. 3 and Fig. 4 show the normalized transmitted data, versus 1000 of scheduling intervals (symbols), for a pure opportunistic, a proportional fair, and the opportunistic fair scheduling scheme. As no fairness constraint exists for pure opportunistic scheme, its normalized transmitted data outperforms the ones of the opportunistic fair and proportional fair schemes. While the former schemes have almost the same amount of transmitted data, their fairness indexes are different. The opportunistic fair scheme allocates the resources based on the computed fair weights, so the averages of transmitted rates (R_i) deviations from the fair weights, i.e., its fairness indexes, are less than the ones of the proportional fair scheme.

In the second stage of the performance evaluation, we consider the three scenarios shown in Fig. 5 to evaluate the performance for large scale channel variations and heterogeneous users' utility functions. In the first and the second scenarios, Fig. 5-(a) and Fig. 5-(b), the traffic is homogeneous,
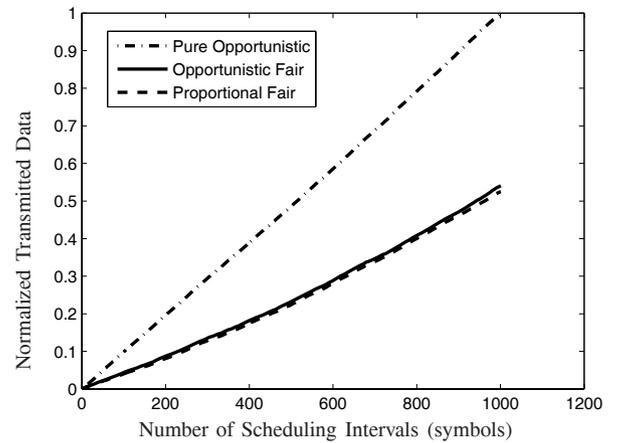

Fig. 3. Normalized transmitted data for a pure opportunistic, a proportional fair, and the opportunistic fair scheduling schemes.
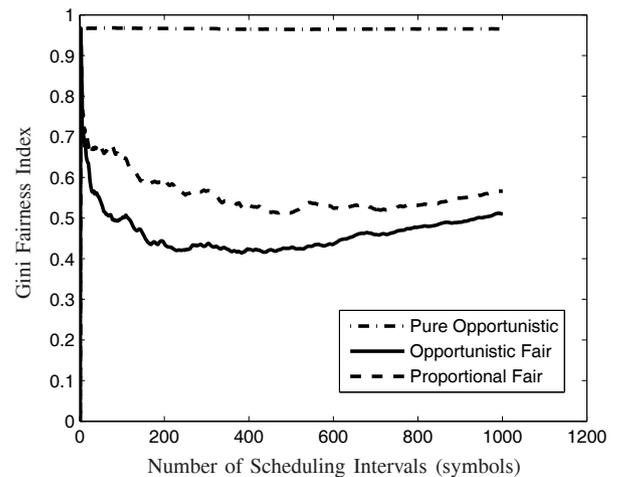

Fig. 4. Gini fairness index for a pure opportunistic, a proportional fair, and the opportunistic fair scheduling schemes.

and we show the effect of channel gain variations on the scheduling performance. In the third scenario, Fig. 5-(c), we show the scheduling performance when users have heterogeneous traffic, i.e., users with non-concave and a concave utility functions exist.

A. Fixed Users

In the first scenario, there are 16 users, half of them are uniformly located on a circle with 50 meters radius, and the other half are located on the cell edge at equal angular distance. Users have diverse channel gains due to path loss and distance fading. We investigate the effect of multiuser diversity on throughput and fairness performance of the scheduling schemes using this scenario.

Fig. 6 shows total throughput of the network versus the number of users for the opportunistic and opportunistic fair scheduling schemes. As the opportunistic scheduling assigns a sub-carrier to a user with the highest channel gain, its throughput is the upper bound. The opportunistic fair scheduling achieves lower throughput than opportunistic scheduling

---

[3]known as Worldwide Interoperability for Microwave Access (WiMAX)
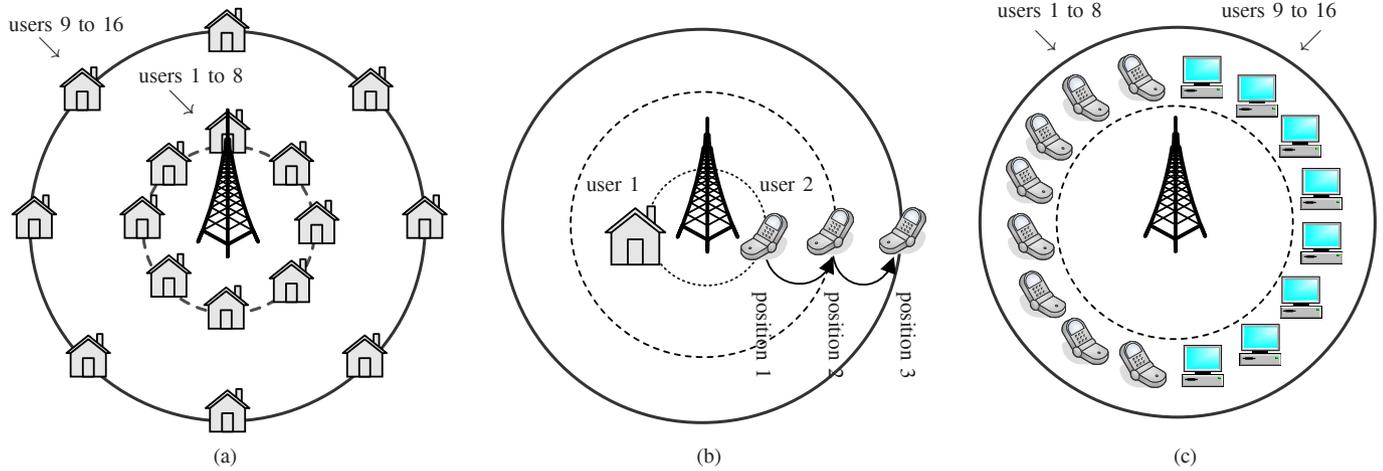
Fig. 5. Simulated scenarios: (a) fixed users, (b) a fixed user and a mobile user, (c) users with heterogeneous traffic.
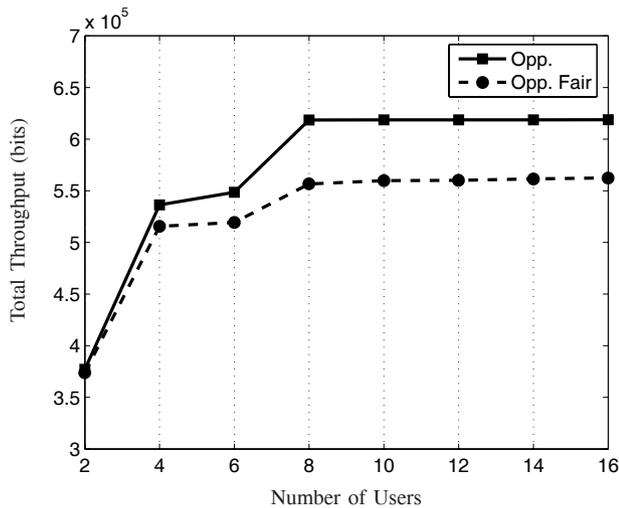

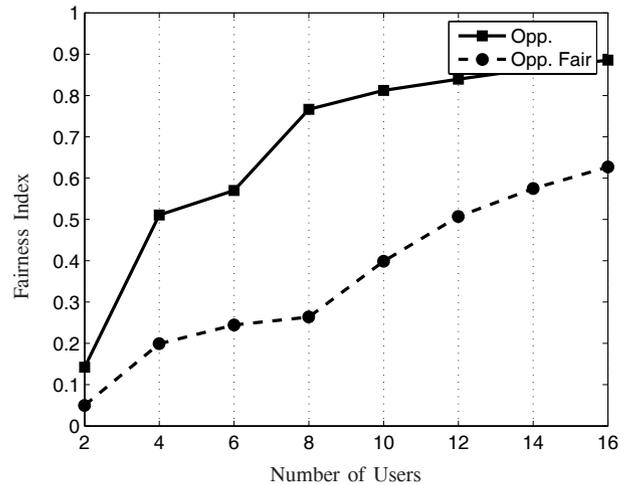Fig. 6. Total network throughput of the first scenario.

Fig. 7. Fairness index of the first scenario.

because in some scheduling intervals it assigns a number of sub-carriers to users who have not been supported for a long time, irrespective to their channel gain. Both scheduling schemes exploit multi-user diversity as more users join the inner circle, i.e., when the number of users increases from 1 to 8 in Fig. 6. Users 9 to 16 are far from the BS and their channel gains are always much lower than the users located on inner circle, so they do not increase multi-user diversity gain and the throughput remains almost constant when these users join the network.

Fig. 7 shows the Gini fairness index of the first scenario. The fairness index of opportunistic and opportunistic fair scheduling increases as the number of users increases. Increasing user diversity has an adverse effect on fairness. However, this effect is moderated in the opportunistic fair scheduling especially at low spatial diversity, i.e., users 1 to 8.

B. A Fixed and a Mobile User

In the second scenario, a fixed user and a mobile user that moves away from the BS are considered. At first, users 1 and 2 are located close to the BS with the same distance. Then, user 2 moves away from the BS toward the edge of the cell. We investigate the adaptivity of the opportunistic fair scheduling in capturing the network status variations using this scenario.

Fig. 8 shows the throughput of user 1 and user 2 at three positions for opportunistic and opportunistic fair scheduling schemes. The throughput of opportunistic fair scheduling has been shown for two different time constants, T_c, of the exponentially weighted moving average. As user 2 moves away from the BS and its average channel gain drops, the opportunistic scheduling allocates less rate to it and finally ignores it when it is very far. On the other hand, the opportunistic fair scheduling scheme, which intends to allocate proportional rates to the fair weights, allocates more rate to user 2 than the ones of opportunistic allocation. In comparison to schedul-
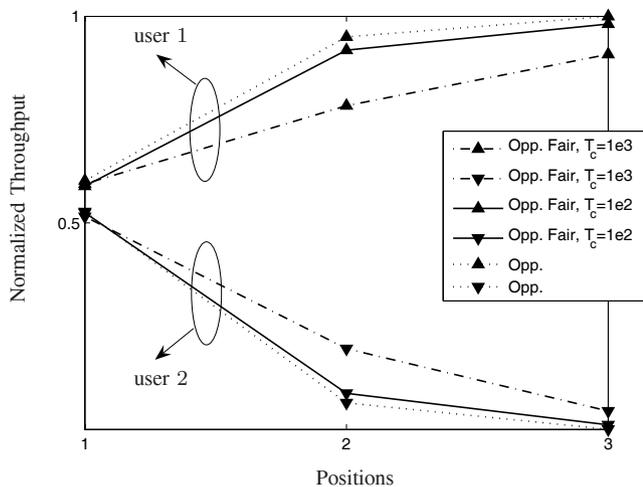
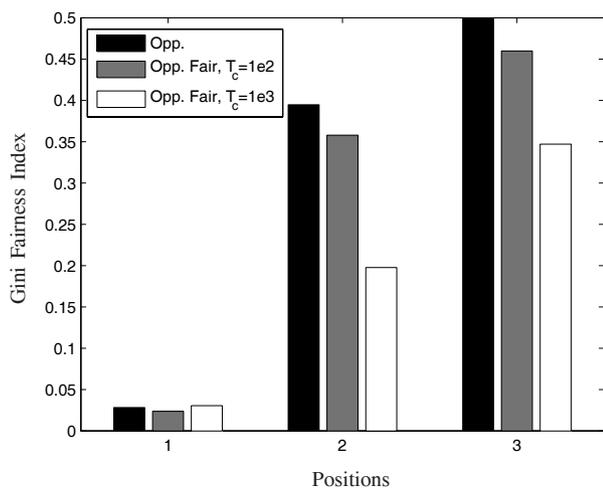Fig. 8. Users throughput at different positions of the second scenario.

Fig. 10. Utilities of users 1 to 8 versus time for opportunistic and opportunistic fair scheduling schemes.

Fig. 9. Fairness performance of the second scenario.

ing schemes with large T_c, the opportunistic fair scheduling scheme with small T_c is less effective in compensating the effect of bad channel gain of user 2 as it moves away from the BS. This can be explained as follows. A smaller number of scheduling intervals is considered and compensated for in the fairness scheme when T_c is small. Therefore, the scheduler has shorter time to compensate for the unfairness.

Fig. 9 shows the Gini fairness index of the opportunistic and opportunistic fair scheduling with two different T_c in the second scenario. When both users are close to the BS and their channels are almost similar, unfairness of opportunistic scheduling is not observed. However, as user 2 moves and its channel condition degrades, the opportunistic fair scheduling treats it more fairly than the opportunistic scheduling, so the fairness index of the opportunistic scheduling deteriorates when user 2 is at positions 2 and 3. Opportunistic fair scheduling with larger T_c outperforms the one with smaller T_c in terms of fairness.

The performance study of the second scenario indicates that the opportunistic fair scheduling can capture the net-
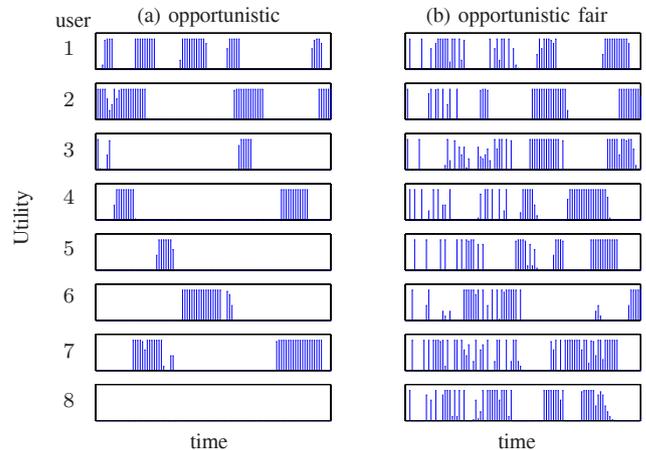
work changes and adapt the fairness scheme accordingly. The adaptivity of the scheme can be adjusted by controlling the transmission history duration which is one of the components of the fairness module. Furthermore, the trade-off between fairness and throughput can be adjusted similarly.

C. Heterogeneous Users

In the third scenario, all 16 users are within the same distance from the BS, on a circle with 50 meters radius, but they run two different applications with different utility functions. The first group of users, users 1 to 8, have a non-concave utility function, and the second group of users, users 9 to 16, have a concave utility function.

The utilities of users 1 to 8 versus time, when their traffic is scheduled by opportunistic and opportunistic fair scheduling schemes, are represented in Fig. 10-a and Fig. 10-b, respectively. The figures show that, first, opportunistic scheduling ignores few users with low channel gains over the simulation intervals, such as user 8 in Fig. 10-a. This fact causes severe unfairness in service provisioning when user diversity is high. Second, the rate allocations and hence the users' utilities for opportunistic scheduling is highly interrupted in time compared to those of opportunistic fair scheduling. Although scheduling elastic traffic, with concave utility, is not sensitive to service provisioning delay, inelastic traffic, with non-concave utility, should be scheduled within least possible of service delays. Therefore, opportunistic scheduling is not appropriate for inelastic traffic service provisioning.

Furthermore, the aggregate users' utilities shown in Table II for both scheduling schemes demonstrate that the improvement in resource utilization or in the users' satisfaction of received service, represented by sum of the users' utilities, is higher for opportunistic fair scheduling than that of the opportunistic scheduling scheme. Moreover, the aggregate utilities of users with non-concave utilities are higher than that of the users with concave utilities. The reason is that the gradient of the non-concave utility function is higher than the gradient of the concave utility function at lower rates.

Table II
AGGREGATE UTILITIES OF THE SCHEDULING SCHEMES

Scheduling Scheme	$\sum_{i=1}^{8} U_i(r_i)$	$\sum_{i=9}^{16} U_i(r_i)$	$\sum_{i=1}^{16} U_i(r_i)$
Opportunistic	161.4702	141.0092	302.4793
Opportunistic Fair	306.2989	196.6107	502.9096

Therefore, for the same allocated rate, the non-concave utility is larger than the concave utility.

V. CONCLUSIONS

Fair weights have been designed for scheduling heterogeneous traffic in the downlink of OFDMA networks. We adopt the utility proportional fair criteria, design a set of fair weights associated with users, and propose an opportunistic fair scheduling which allocates the resources according to the fair weights. The proposed scheduler is adaptive because the fair weights can be modified dynamically when the network characteristics change due to mobility of users, admitting a new user, or changing the fairness policy of the network service provider. In addition, it reduces service interruption for real-time traffic which is sensitive to long service delays. In our further works, we will investigate various optimal strategies to detect changes in users' average channel gains which trigger the computation of the fair weights.

VI. APPENDIX

The mathematical representations of $\nabla^2_{\mathbf{ww}}\mathcal{L}$ and $\nabla_{\mathbf{w}}f(\mathbf{w})$, required by the interior point algorithm and depended on users' utility functions, are presented in the appendix.

The objective function of P_4, based on utility functions (44), is given by:

$$f(\mathbf{w}) = - \ log(Util_1(w_1)) - \ldots - log(Util_{N_u}(w_{N_u})). \quad (45)$$

Accordingly, $\nabla_{\mathbf{w}}f(\mathbf{w}) = \left(\frac{\partial f}{\partial w_{11}}, \ldots, \frac{\partial f}{\partial w_{N_u N_c}}\right)^T$ is computed as follows:

$$\nabla_{\mathbf{w}}f(\mathbf{w}) = - \begin{pmatrix} \frac{1}{Util_1(w_1)}\frac{\partial Util_1(w_1)}{\partial w_{11}} \\ \vdots \\ \frac{1}{Util_1(w_1)}\frac{\partial Util_1(w_1)}{\partial w_{N_u N_c}} \\ \vdots \\ \frac{1}{Util_{N_u}(w_{N_u})}\frac{\partial Util_{N_u}(w_{N_u})}{\partial w_{N_u 1}} \\ \vdots \\ \frac{1}{Util_{N_u}(w_{N_u})}\frac{\partial Util_{N_u}(w_{N_u})}{\partial w_{N_u N_c}} \end{pmatrix}, \quad (46)$$

where, for $j = 1, \cdots, N_c$, and $\theta = \frac{\pi}{2}\frac{w_i - l_1}{l_2 - l_1}$:

$$\frac{\partial Util_{\breve{i}}}{\partial w_{ij}} = \begin{cases} \frac{-k\pi}{2(l_2 - l_1)}\frac{cos(\theta)}{sin(\theta)}, & if \ \ i = \breve{i}, \\ 0, & otherwise. \end{cases} \quad (47)$$

To obtain $\nabla^2_{\mathbf{ww}}\mathcal{L}$, first $\nabla^2_{\mathbf{ww}}f(\mathbf{w})$ and $\nabla^2_{\mathbf{ww}}\mathbf{c}(\mathbf{w})$ are computed:

$$\nabla^2_{\mathbf{ww}}f(\mathbf{w}) = - \begin{pmatrix} \mathbf{G}(w_1) & 0_{(N_c,N_c)} & \cdots & 0_{(N_c,N_c)} \\ 0_{(N_c,N_c)} & \mathbf{G}(w_2) & \cdots & 0_{(N_c,N_c)} \\ \vdots & \vdots & & \vdots \\ 0_{(N_c,N_c)} & 0_{(N_c,N_c)} & \cdots & \mathbf{G}(w_{N_u}) \end{pmatrix}, \quad (48)$$

where

$$\mathbf{G}(w_i) = \begin{pmatrix} \frac{\partial^2 Util_i}{\partial w_{i1}\partial w_{i1}} & \cdots & \frac{\partial^2 Util_i}{\partial w_{i1}\partial w_{iN_c}} \\ \frac{\partial^2 Util_i}{\partial w_{i2}\partial w_{i1}} & \cdots & \frac{\partial^2 Util_i}{\partial w_{i2}\partial w_{iN_c}} \\ \vdots & & \vdots \\ \frac{\partial^2 Util_i}{\partial w_{iN_c}\partial w_{i1}} & \cdots & \frac{\partial^2 Util_i}{\partial w_{iN_c}\partial w_{iN_c}} \end{pmatrix}. \quad (49)$$

$0_{(N_c,N_c)}$ is a $N_c \times N_c$ matrix with all zero entries. The second partial derivatives of the utility functions required for calculating $G(w_i)$ functions are:

$$\frac{\partial^2 Util_i}{\partial w_{\breve{i}j}\partial w_{ij}} = \begin{cases} \frac{k\pi^2}{4(l_2-l_1)^2}\left(\frac{1}{sin^2(\theta)}\right), & if \ \ i = \breve{i}, \\ 0, & otherwise, \end{cases} \quad (50)$$

for $\breve{j}$ and $j \in \{1, \cdots, N_c\}$.

In problem P_4, $\mathbf{c}(\mathbf{w})$ is represented by:

$$\mathbf{c}(\mathbf{w}) = \begin{pmatrix} \sum_{j=1}^{N_c} w_{1j} - \frac{l_1+l_2}{2} \\ \vdots \\ \sum_{j=1}^{N_c} w_{N_u j} - \frac{l_1+l_2}{2} \\ P_{BS} - \sum_{j=1}^{N_c}\sum_{i=1}^{N_u}\frac{2^{w_{ij}}-1}{\alpha_{ij}} \\ w_{11} \\ \vdots \\ w_{N_u N_c} \end{pmatrix}. \quad (51)$$

Accordingly, $\nabla^2_{\mathbf{ww}}\mathbf{c}(\mathbf{w})$ for calculating $\nabla^2_{\mathbf{ww}}\mathcal{L}$ can be obtained by:

$$\nabla^2_{\mathbf{ww}}\mathbf{c}(\mathbf{w}) = (ln(2))^2 \begin{pmatrix} \frac{2^{w_{11}}}{\alpha_{11}} & 0 & \cdots & 0 \\ 0 & \frac{2^{w_{12}}}{\alpha_{12}} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{2^{w_{N_u N_c}}}{\alpha_{N_u N_c}} \end{pmatrix}. \quad (52)$$

REFERENCES

[1] R. Knopp and P. A. Humblet, "Information capacity and power control in single cell multiuser communications," in *Proc. IEEE International Conf. Commun. (ICC'95)*, Seattle, USA, June 1995.

[2] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," 2003.

[3] Y. Liu and E. Knightly, "Opportunistic fair scheduling over multiple wireless channels," in *Proc. IEEE International Conf. Comput. Commun. (INFOCOM'03)*, San Francisco, USA, Apr. 2003.

[4] Z. Zhang, Y. He, and E. K. P. Chong, "Opportunistic scheduling for OFDM systems with fairness constraints," *EURASIP J. Wireless Commun. Netw.*, vol. 8, no. 3, pp. 275-277, 2008.

[5] A. Wang, L. Xiao, S. Zhou, X. Xu, and Y. Yao, "Dynamic resource management in the fourth generation wireless systems," in *Proc. International Conference on Communication Technology (ICCT'03)*, Beijing, China, Apr. 2003.

[6] A. Jalali and R. Padovani, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. IEEE Veh. Technol. Conf. (VTC 2000-Spring)*, Tokyo, Japan, Spring 2000.

[7] I. C. Wong and B. L. Evans, "Optimal downlink OFDMA resource allocation with linear complexity to maximize ergodic rates," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 962-971, 2008.

[8] H. T. Cheng and W. Zhuang, "An optimization framework for balancing throughput and fairness in wireless networks with QoS support," in *Proc. 3rd International Conf. Quality Service Heterogeneous Wired/Wireless Netw. (QShine'06)*, Waterloo, Ontario, Canada, 2006.

[9] M. Awad, V. Mahinthan, M. Mehrjoo, X. Shen, and J. Mark, "Downlink resource allocation for OFDMA-based multiservice networks with imperfect CSI," in *Proc. IEEE International Conf. Commun. (ICC'09)*, Dresden, Germany, June 2009.

[10] F. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *J. Operational Research Society*, vol. 49, pp. 237-252, 1998.

[11] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 556-567, 2000.

[12] M. Dianati, X. Shen, and K. Naik, "Coperative fair scheduling for the downlink of CDMA cellular networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1749-1760, 2007.

[13] T. Zhang, Z. Zeng, and Y. Qiu, "A subcarrier allocation algorithm for utility proportional fairness in OFDM systems," in *Proc. IEEE Veh. Technol. Conf. (VTC'8)*, 2008, pp. 1901-1905.

[14] M. Mehrjoo, M. Dianati, X. Shen, and K. Naik, "Opportunistic fair scheduling for the downlink of IEEE 802.16 wireless metropolitan area networks," in *Proc. 3rd International Conf. Quality Service Heterogeneous Wired/Wireless Netw. (QShine)*, Waterloo, Canada, Aug. 2006.

[15] M. Mehrjoo, S. Moazeni, and X. Shen, "A new modeling approach for utility-based resource allocation in OFDM networks," in *Proc. IEEE International Conf. Commun. (ICC'08)*, Beijing, China, May 2008.

[16] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. IEEE International Symp. Inf. Theory (ISIT'06)*, Seattle, USA, July 2006.

[17] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310-1322, 2006.

[18] J. Nocedal and S. Wright, *Numerical Optimization*. Springer, 2006.

[19] M. Mehrjoo, M. K. Awad, and X. Shen, "Resource allocation for OFDM-based WiMAX," in *WiMAX Network Planning and Optimization*. CRC Press - Taylor and Francis, 2009.

[20] J. W. Mark and W. Zhuang, *Wireless Communications and Networking*. Upper Saddle River, NJ: Pearson Education International, 2003.

[21] J. Cai, X. Shen, and J. Mark, "Downlink resource management for packet transmission in OFDM wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1688-1703, 2005.

[22] M. K. Awad, K. T. Wong, and Z. Li, "An integrated overview of the open literatures empirical data on the indoor radiowave channels delay properties," *IEEE Trans. Antennas Propagation*, vol. 56, no. 5, pp. 1451-1468, 2008.

[23] I. T. U. (ITU), Recommendation ITU-R M.1225, guidelines for evaluation of radio transmission technologies for IMT-2000, std., 1997.

[24] R. Jain, Apr. 2010, "Channel models: a tutorial." [Online]. Available: accessedonApril20,2010,www.cs.wustl.edu/jain/cse574-08/ftp/channelmodeltutorial.pdf

[25] D. M. Chiu and A. S. Tam, "Fairness of traffic controls for inelastic flows in the Internet," *Comput. Netw.*, vol. 51, no. 11, pp. 2938-2957, 2007.

[26] IEEE, IEEE 802.16e-2005, IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, Std., Feb. 2006.

Mehri Mehrjoo (M'06) received the B.A.Sc. and the M.A.Sc. degrees from Ferdowsi University, Mashad, Iran, and Ph.D. from the University of Waterloo, Waterloo, Canada in 1993, 1996, and 2008, respectively. From 2008 to 2009, she has been a post-doctoral fellow at the University of Waterloo. Currently, she is an assistant professor in the Department of Telecommunications, University of Sistan and Baluchestan, Zahedan, Iran. Her research interests are in the areas of resource allocation, performance analysis of wireless protocols, vehicular communication.

Mohamad Khattar Awad (S'02, M'09), earned the B.A.Sc in electrical and computer engineering (communications option) from the University of Windsor (Windsor, Ontario, Canada) in 2004 and the M.A.Sc and Ph.D. in electrical and computer engineering from the University of Waterloo (Waterloo, Ontario, Canada) in 2006 and 2009, respectively. From 2004 to 2009 he was a Research Assistant in the Broadband Communications Research Group (BBCR), University of Waterloo.

Currently, he is an Assistant Professor of Electrical and Computer Engineering at the American University of Kuwait, Kuwait. He received the Ontario Research & Development Challenge Fund Bell Scholarship in 2008 and 2009, and the University of Waterloo Graduate Scholarship in 2009. His research interest includes wireless communications, wireless networks resource allocation and acoustic vector-sensor signal processing.

Mehrdad Dianati is a Lecturer (Assistant Professor) at the Department of Electronic Engineering of the University of Surrey in United Kingdom. He received his BSc from Sharif University of Technology, Iran, his MSc from K.N.Toosi University of Technology, Iran, and his Ph.D. from University of the Waterloo in Canada in Electrical and Computer Engineering. He has also several years of industrial experience as a software/hardware developer and technical manager.

Xuemin (Sherman) Shen received the B.Sc.(1982) degree from Dalian Maritime University (China) and the M.Sc. (1987) and Ph.D. degrees (1990) from Rutgers University, New Jersey (USA), all in electrical engineering. He is a Professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo, Canada. Dr. Shen's research focuses on resource management in interconnected wireless/wired networks, UWB wireless communications networks, wireless network security, wireless body area networks and vehicular ad hoc and sensor networks. He is a co-author of three books, and has published more than 400 papers and book chapters in wireless communications and networks, control and filtering. Dr. Shen served as the Technical Program Committee Chair for IEEE VTC'10, the Symposia Chair for IEEE ICC'10, the Tutorial Chair for IEEE ICC'08, the Technical Program Committee Chair for IEEE Globecom'07, the General Co-Chair for Chinacom'07 and QShine'06, the Founding Chair for IEEE Communications Society Technical Committee on P2P Communications and Networking. He also served as a Founding Area Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS; Editor-in-Chief for *Peer-to-Peer Networking and Application*; Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY; *Computer Networks*; and *ACM/Wireless Networks*, etc., and the Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, *IEEE Wireless Communications*, *IEEE Communications Magazine*, and *ACM Mobile Networks and Applications*, etc. Dr. Shen received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004 and 2008 from the University of Waterloo, the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. Dr. Shen is a registered Professional Engineer of Ontario, Canada, an IEEE Fellow, and a Distinguished Lecturer of IEEE Communications Society.