Abstract—This paper presents the novel resource allocation scheme for multiclass services in the downlink OFDMA system, in which the frame-based transmission with multiple subchannels and multiple time slots is used. As system utilization and fairness are necessary, the scheme is based on Proportional Fairness (PF) utility function. Two main service classes are considered. One is the guaranteed service class. The other is the non-guaranteed service class. The scheme guarantees the minimum bit rate for the former class. The PF-based optimization problem with minimum bit rate requirement is formed and solved by using the Lagrange multiplier method with relaxed constraints. The novel PF-based utility function is derived and used for this class. The non-guaranteed class is served after the guaranteed class by using the conventional PF utility function. Extensive simulation with user mobility and finite backlog is conducted. The results show that the proposed scheme is able to satisfy the minimum rate requirement for the guaranteed class with moderate system utilization. Furthermore, within the same class, the scheme can provide very high throughput fairness.

I. INTRODUCTION

The Orthogonal Frequency Division Multiple Access (OFDMA) technique introduces great enhancement in reliability over the frequency selective radio link. It is used in many new generation wireless networks such as IEEE 802.16 standard series or WiMAX. Such networks are aimed to serve the mixture of traffic types from various applications ranging from text transfer to voice and video streaming. Multiple service classes are defined by IEEE 802.16e [1] to support various types of traffic. Those classes are UGS, erTPS, rtPS, nrtPS, and BE. Though classes have various requirements, the first 4 classes (UGS, erTPS, rtPS, nrtPS) have one common constraint on minimum bit rate guarantee. On the other hand, Best Effort (BE) class is served on the space available basis. Multiple service classes resource allocation is considered in [2] and [3]. [2] defines various utility functions based on Proportional Fairness (PF) [4] for multiple classes. PF utility function is multiplied by the weight function based on the waiting time and/or queue size. Hence, there are different utility functions and each utility function is for each class. In [3], multiple service classes are differentiated by different bit rate requirement. Additionally, minimum bit rate requirement is considered in [5] and [6]. [3] considers the similar approach as [5] and [6]. Their approach allocates transmit power and subchannels in each time slot with multiple subchannels. Not only the computational complexity is increased by power allocation algorithm, but also making decision by each slot is not suitable because the real system (such as WiMAX) transmits series of frames with hundreds time slots.

In our design, we consider frame-by-frame resource allocation. A frame consists of multiple subchannels and multiple time slots. We aim to exploit multichannel multiuser diversity in order to efficiently utilize the frequency band, i.e. to transmit as many bits as possible. Spectral efficiency can be obtained by exploiting independent fading characteristics of the users. Such gain can be obtained by assigning resources to the user who has the ‘best’ channel gain such as the maximum carrier-to-interference (maxC/I) scheme [7]. However, maxC/I worsens the user’s throughput when that user is in a deep fade. Hence, being fair to all users is important. Proportional Fairness [8], [4] was proposed to trade throughput with fairness. In addition, we also aim to provide class-based service guarantee. We mention earlier that the first 4 classes (UGS, erTPS, rtPS, nrtPS) have the common requirement on minimum bit rate guarantee, and the BE class does not have any requirement. Therefore, we define 2 main service classes. One is the guaranteed class. We guarantee the minimum bit rate for the guaranteed class. We form the optimization problem based on PF and minimum bit rate requirement. We use the Lagrange multiplier method with relaxed constraints to solve the optimization problem. Furthermore, we propose a fast cross-layer approach to model the multichannel multiuser diversity into two-dimensional matrix. The matrix-based scheduling integrates the solution of the optimization problem to the constraint requirement. After the scheduler serves the guaranteed class, the non-guaranteed class is served by using the conventional PF utility function. Detailed explanation of our system model is in Section II. The proposed scheme as well as the optimization problem formulation and the solution is described in Section III. We explain our integrated cross-layer matrix-based approach in Section IV. Section V is about performance evaluation and results. Finally, Section VI concludes this paper.

978-1-4244-4148-8/09/$25.00 ©2009
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2009 proceedings.
II. SYSTEM MODEL

Our scheme considers the downlink part of a multiuser frame-based OFDM network with a cell diameter of $D$. The downlink part composes of $M$ subchannels. The network has one base station (BS) serving $N_g$ active guaranteed mobile stations (MSs) and $N_{ng}$ active non-guaranteed MSs. We may use the term ‘mobile station’ and ‘user’ interchangeably. The BS contains the centralized scheduler and the queue for each traffic destined to each MS. The network works in a Time Division Duplex (TDD) mode. A frame is used as a transmission unit. Frame duration is constant. A frame is further divided into a downlink subframe and an uplink subframe. Time axis of a downlink subframe is divided into $K$ time slots which has constant duration. Figure 1 illustrates a frame in our model.

![An OFDMA TDD frame.](image)

In our model, frame duration is assumed to be small so that the channel is time-invariant during one frame duration, which is also a scheduling epoch. By using adaptive modulation and coding (AMC), the maximum number of bits per symbol per Hz can be expressed as a function of SNR (signal-to-noise ratio) and the target BER (bit error rate). The capacity of each subchannel is estimated by using the Shannon capacity expression. We denote $R_{m,n}(t)$ as the achievable rate of user $n$ at frame $t$ if subchannel $m$ is assigned. The achievable rate can be estimated by

$$R_{m,n}(t) = \Delta B \cdot \log_2 \left(1 + \alpha \cdot \frac{SNR_{m,n}(t)}{\delta} \right).$$

(1)

$\Delta B$ is the frequency bandwidth and $SNR_{m,n}(t)$ is the SNR perceived by user $n$ on subchannel $m$ at frame $t$. Since Shannon’s channel capacity theorem does not consider the forward error correction scheme and the target bit error rate (BER), $\delta$ is a parameter to bridge the gap between Shannon capacity and practical forward error correction scheme and $\alpha$ is a constant BER [9]. These parameters can be expressed as shown in

$$\delta = \frac{\ln(5 \cdot BER_n)}{1.5}.$$  

(2)

$BER_n$ is the bit error rate requirement by user $n$. Furthermore, $SNR_{m,n}(t)$ is a function of the transmission power and the channel gain, and can be estimated by

$$SNR_{m,n}(t) = \frac{P_{m,n} \cdot g_{m,n}}{N_0 \cdot \Delta B}.$$  

(3)

Note that $p_{m,n}$ and $g_{m,n}$ are the transmission power allocated to user $n$ on subchannel $m$, and the channel gain perceived by user $n$ on subchannel $m$, respectively. $N_0$ is the noise spectral power density. According to [10], dynamic power allocation brings only a negligible improvement in total throughput, but incurs tremendous complexity and immense feedback data requirement. For this reason, we choose to employ static power allocation, that is, in the downlink, the total transmission power is constant and uniformly distributed over all subchannels. At each scheduling epoch, once the network resources are allocated to users, the control information and map messages are sent at the beginning of the frame to indicate which subchannel and time slots are allocated to which user.

III. RESOURCE ALLOCATION SCHEME FOR GUARANTEED AND NON-GUARANTEED CLASSES

In our scheme, we consider 2 main classes. Users (or user’s applications) have to register to either one of these 2 classes.

A. Guaranteed Class

We aim to guarantee the minimum bit rate requirement for this class. We also aim to maximize the total system utilization and to be fair among users within the same class. We employ proportional fairness which has been proved to achieve a tradeoff between system throughput and fairness and to maximize the summation of logarithmic function of average rate [8]. In a single carrier system, the PF utility function follows

$$i_{m,n}(t) = \frac{R_{m,n}(t)}{I_n(t)}. $$

(4)

$i_{m,n}(t)$ is the PF utility function of user $n$ via subchannel $m$. $I_n(t)$ is an average bit rate of user $n$. [11] extended PF to a multicarrier system and proved the maximization objective function for PF in a multicarrier system, which is

$$\prod_{n \in U_c} \left(1 + \frac{\sum_{m \in C_n} R_{m,n}}{(t_c - 1)T_n} \right).$$  

(5)

$U_c$ is a set of selected users and $C_n$ is a set of subcarriers that are allocated to user $n$. $t_c$ is the exponential moving average weight factor. Note that in (5), we omit the notation of frame $t$. By using PF objective function in (5), we can form our optimization problem with constraints on the minimum bit rate requirement. We denote $R_{min,n}$ as a minimum bit rate requirement submitted by user $n$, and $c_{m,n}$ as an integer that indicates the number of time slots of subchannel $m$ that are allocated to user $n$. The optimization problem is formulated as

$$\text{maximize} \prod_{n=1}^{N_g} \left(1 + \frac{\sum_{m=1}^{M} R_{m,n} \cdot c_{m,n}}{(t_c - 1)T_n} \right)$$  

subject to

$$\sum_{n=1}^{N_g} c_{m,n} \leq K; \quad \forall m = 1, 2, \ldots, M; \quad c_{m,n} \in \mathbb{Z} \quad \text{and} \quad c_{m,n} \geq 0,$$

(7)
\[ R_{\text{min}} \leq \sum_{m=1}^{M} R_{m,n} \cdot c_{m,n}; \quad \forall n = 1, 2, \ldots, N_g. \]  

Note that in this context user \( n \) belongs to the guaranteed group \( (N_g) \). (7) indicates that the integer allocation indicator cannot be more than \( K \) time slots of each subchannel. (7) also allows time multiplexing of multiple users in the same subchannel by using separated time slots but the summation of allocation indicators in one subchannels cannot exceed \( K \). (8) is the constraint on minimum bit rate requirement.

We use the technique adopted from [12] to relax the constraint \( c_{m,n} \in \mathbb{Z} \) to \( c_{m,n} \in \mathbb{R} \), i.e., \( c_{m,n} \) belongs to a set of real numbers. This relaxation means that the time duration of one subchannel in one frame is infinitely divisible. The optimization problem can be rewritten as

\[
\text{minimize} \quad -\sum_{n=1}^{N_g} \log \left( 1 + \frac{\sum_{m=1}^{M} R_{m,n} \cdot c_{m,n}}{(t_c - 1)T_n} \right) 
\]

subject to

\[
\sum_{n=1}^{N_g} c_{m,n} \leq K; \quad \forall m = 1, 2, \ldots, M; \quad c_{m,n} \in \mathbb{R} \text{ and } c_{m,n} \geq 0,
\]

\[ R_{\text{min}} \leq \sum_{m=1}^{M} R_{m,n} \cdot c_{m,n}; \quad \forall n = 1, 2, \ldots, N_g. \]

By using the method of Lagrange multipliers, we obtain the Lagrangian

\[
L = -\sum_{n=1}^{N_g} \log \left( 1 + \frac{\sum_{m=1}^{M} R_{m,n} \cdot c_{m,n}}{(t_c - 1)T_n} \right) + \sum_{m=1}^{M} \mu_m \left[ \sum_{n=1}^{N_g} c_{m,n} - K \right] + \sum_{n=1}^{N_g} \lambda_n \left( \sum_{m=1}^{M} R_{m,n} \cdot c_{m,n} - R_{\text{min}} \right),
\]

where \( \mu_m \) \((m = 1, 2, \ldots, M)\) and \( \lambda_n \) \((n = 1, 2, \ldots, N_g)\) are the Lagrangian multipliers for the constraints in (10) and (11) respectively.

We differentiate (12) with respect to \( c_{m,n} \) and obtain

\[
\frac{\partial L}{\partial c_{m,n}} = -\frac{R_{m,n}}{(t_c - 1)T_n + \sum_{m=1}^{M} R_{m,n} \cdot c_{m,n}} + \mu_m + \lambda_n R_{m,n}.
\]

The necessary conditions for \( c_{m,n}^* \) which is the optimal solution should satisfy

\[
\frac{\partial L}{\partial c_{m,n}} \bigg|_{c_{m,n}^*} = \begin{cases} > 0 & \text{if } c_{m,n}^* = 0 \\ = 0 & \text{if } 0 < c_{m,n}^* < K \\ < 0 & \text{if } c_{m,n}^* = K \end{cases}
\]

(14) can be analyzed as

\[
c_{m,n}^* = 0 \quad \text{if } \mu_m > H_{m,n}(\lambda_n), \\
c_{m,n}^* < K \quad \text{if } \mu_m = H_{m,n}(\lambda_n), \\
c_{m,n}^* = K \quad \text{if } \mu_m < H_{m,n}(\lambda_n).
\]

\[ H_{m,n}(\lambda_n) \text{ is defined as}
\]

\[
H_{m,n}(\lambda_n) = \left( \frac{1}{(t_c - 1)T_n + \sum_{m=1}^{M} R_{m,n} c_{m,n}} - \lambda_n \right) R_{m,n}.
\]

Because \( \sum_{m=1}^{M} R_{m,n} c_{m,n}^* \) indicates the allocation of down-link resource to user \( n \), and we aim to satisfy each user with the minimum bit rate, hence, we use the observation approach [13]. Our observation is that \( \sum_{m=1}^{M} R_{m,n} c_{m,n} \) is the assigned bit rate, and it should satisfy the minimum rate requirement \( R_{\text{min}} \). Hence, we substitute \( \sum_{m=1}^{M} R_{m,n} c_{m,n} \) in (18) with \( R_{\text{min}} \), and obtain the suboptimal solution as

\[
H_{m,n}(\lambda_n) = \frac{1}{(t_c - 1)T_n + R_{\text{min}}} - \lambda_n R_{m,n}
\]

By this suboptimal solution, each subchannel should be allocated to the user with the largest \( H_{m,n}(\lambda_n) \). However, this solution contains \( \lambda_n \), which should be optimum of each user. Searching for the optimal \( \lambda_n \) for each user is computation time-consuming. Hence, we temporarily drop \( \lambda_n \) and yield

\[
H_{m,n} = \frac{R_{m,n}}{(t_c - 1)T_n + R_{\text{min}}}
\]

In (20), \( R_{\text{min}} \) only represents the minimum bit rate requirement, but does not represent the actual possible bit rate due to the backlog. Hence, we propose

\[
h_{m,n} = \frac{R_{m,n}}{(t_c - 1)T_n + \min R_{\text{min}}, Q_{\text{Kavaian}, T K}}
\]

as it takes the actual backlog into consideration. The minimum bit rate requirement lies between the submitted value \( (R_{\text{min}}) \) and the actual minimum possible bit rate (according to queue size and channel capacity). In (21), \( \min(x, y) \) returns the smaller value of either \( x \) or \( y \), and \( T_K \) is duration of one time slot. \( Q_{\text{Kavaian}} \) represents the number of available time slots of subchannel \( m \). \( Q_{n} \) is the current queue size (bits) of user \( n \).

Note that \( \lambda_n \) is introduced into the Lagrangian (12) due to the constraint on the minimum bit rate requirement (8), (11). We propose the cross-layer method that considers satisfying this constraint, as well as integrates the suboptimal solution and multichannel multiuser diversity together. At every scheduling epoch (frame), the scheduler creates the utility matrix for the active guaranteed users. The matrix’s rows represent subchannels and columns represent active guaranteed users. An example of a utility matrix is shown as

\[
\begin{bmatrix}
& \text{user}_1^\prime & \text{user}_2^\prime & \ldots & \text{user}_{N_g}^\prime \\
\text{subch}_1 & h_{1,1} & h_{1,2} & \ldots & h_{1,N_g} \\
\text{subch}_2 & h_{2,1} & h_{2,2} & \ldots & h_{2,N_g} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\text{subch}_M & h_{M,1} & h_{M,2} & \ldots & h_{M,N_g} \\
\end{bmatrix}_{M \times N_g}
\]

Due to limited space, the steps of our cross-layer matrix-based approach is summarized by the pseudo code in Algorithm 1. We briefly describe our algorithm.
selects the user who has the highest utility function and serves that user according to $B_{\min,n}$, which is the minimum number of bits per frame according to the minimum bit rate requirement ($R_{\min,n}$). This approach satisfies minimum bit rate constraint of the selected users. The number of required time slots is calculated according to $B_{\min,n}$ and the selected subchannel’s capacity. The algorithm allocates required slots ($K_{req,m,n^\star}$). Then, $B_{\min,n}$ is reduced and the matrix is updated accordingly. A satisfied user is deleted from the set of candidates, e.g. the corresponding column is deleted from the matrix. A subchannel that runs out of available time slots is also deleted from the matrix. Hence, matrix size shrinks. Algorithm ends when matrix size is 0 or all active users’ queues are served. If a frame still has remaining slots and there are non-empty queues, we serve the non-guaranteed users.

### Algorithm 1: Cross-layer Approach for Guaranteed Class

**Initialize:** $h_{m,n} \in H_{M \times N_g}$ /* Create a matrix $H_{M \times N_g}$ */

$B_{\min,n} = R_{\min,n} \times T_{frame}$ /* $T_{frame}$ = frame time */

/* Calculate minimum no. of bits due to $R_{\min,n}$ */

$B_{\min,n} \in B_{\min list}$ /* and keep in a list */

**while** count($h_{m,n} \in H_{M \times N_g}$) > 0 **do**

$h_{m,n} \leftarrow \arg \max h_{m,n}$

$K_{req,m,n^\star} = \min(B_{\min,n} - Q_n^\star) / T_K$ = slot time */

/* Calculate required no. of slots due to $R_{\min,n}$ */

if $K_{req,m,n^\star} \leq K_{\text{avail}}$ then

/* $K_{\text{avail}}$ = available slots of subchannel $m^\star$ */

allocate $K_{req,m,n^\star}$

*K_{\text{avail}} = K_{req,m,n^\star} / Update $K_{\text{avail}}$ */

for all $n = n^\star$, delete ($h_{m,n} \in H_{M \times N_g}$) delete ($B_{\min,n} \in B_{\min list}$)

update $T_n$.

else if $K_{req,m,n^\star} > K_{\text{avail}}$ then

allocate $K_{\text{avail}}$

*K_{\text{avail}} = 0 / Update $T_n$ */

for all $n = n^\star$, update ($h_{m,n} \in H_{M \times N_g}$)

end if

if $K_{\text{avail}} = 0$ then

for all $m = m^\star$, delete ($h_{m,n} \in H_{M \times N_g}$)

end if

end while

### B. Non-guaranteed Class

For the non-guaranteed class, we aim to achieve system utilization and fairness. We employ PF utility function (4) in our scheme. We can see that (4) is equivalent to (20) when $R_{\min,n}$ is removed (or equals to 0) because ($t_c - 1$) is constant. When there are remaining resources after serving the guaranteed class, the scheduler creates the utility matrix for the active non-guaranteed users. An example of a utility matrix for the non-guaranteed class is

$\begin{bmatrix}
\text{user}_1 & \text{user}_2 & \ldots & \text{user}_{N_g} \\
\text{subch}_1 & i_{1,1} & i_{1,2} & \ldots & i_{1,N_g} \\
\text{subch}_2 & i_{2,1} & i_{2,2} & \ldots & i_{2,N_g} \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\text{subch}_{M} & i_{M,1} & i_{M,2} & \ldots & i_{M,N_g}
\end{bmatrix}_{M \times N_g}$

We use the similar cross-layer matrix-based approach as our previous work [14], which is summarized in Algorithm 2.

### Algorithm 2: Cross-layer Approach for Non-guaranteed Class

**Initialize:** $i_{m,n} \in I_{M \times N_{n^\star}}$ /* Create a matrix $I_{M \times N_{n^\star}}$ */

**while** count($i_{m,n} \in I_{M \times N_{n^\star}}$) > 0 **do**

$i_{m,n} \leftarrow \arg \max i_{m,n}$

$K_{req,m,n^\star} = \frac{Q_n^\star}{T_K}$ /* Calculate required no. of slots due to $Q_n^\star$ */

if $K_{req,m,n^\star} \leq K_{\text{avail}}$ then

/* $K_{\text{avail}}$ = available slots of subchannel $m^\star$ */

allocate $K_{req,m,n^\star}$

*K_{\text{avail}} = K_{req,m,n^\star} / Update $K_{\text{avail}}$ */

for all $n = n^\star$, delete ($i_{m,n} \in I_{M \times N_{n^\star}}$)

update $T_n$.

else if $K_{req,m,n^\star} > K_{\text{avail}}$ then

allocate $K_{\text{avail}}$

*K_{\text{avail}} = 0 / Update $T_n$ */

for all $n = n^\star$, update ($i_{m,n} \in I_{M \times N_{n^\star}}$)

end if

if $K_{\text{avail}} = 0$ then

for all $m = m^\star$, delete ($i_{m,n} \in I_{M \times N_{n^\star}}$)

end if

end while

Our cross-layer approach schedules a batch of time slots at a time. Therefore, the scheduling can be done faster than in the traditional approach, in which the scheduler considers slot by slot of each subchannel.

In order to provide fairness, the past received service is considered. For both classes, the portion or received service is taken into account by using $T_n$. $T_n$ is updated every time when user $n$ is scheduled by using

$$T_n = \begin{cases} 
(1 - \frac{1}{t_c}) T_n + \frac{1}{t_c} \cdot \frac{R_{req,m,n}}{K_{req,m,n}} \cdot R_{m,n}(t), & \text{if } Q_n \text{ is served} \\
(1 - \frac{1}{t_c}) T_n, & \text{if } Q_n \text{ is not served} \\
T_n, & \text{if } Q_n \text{ is empty (inactive).}
\end{cases}$$

(22)

## IV. PERFORMANCE EVALUATION AND DISCUSSION

Our model consists of one base station serving 48 MSs. MSs are divided into 4 groups, i.e. each group has 12 MSs. The first group registers to be guaranteed with the minimum bit rate of 0.75 Mbps. The second group registers to be
guaranteed with 1 Mbps and the third group registers to be guaranteed with 1.25 Mbps. The fourth group is the non-guaranteed group. Each MS has downlink traffic with the same long-term average rate. The traffic is generated using the ON-OFF traffic model. The ON duration is created with an exponential distribution with a mean of 0.01 second. The OFF duration is exponentially distributed with a mean of 0.03 sec. During the ON period, packets are generated with an exponentially distributed interarrival time. We control the average load by adjusting the mean interarrival time. Packet size is exponentially distributed around a mean of 180 bytes. Any packets larger than 1,500 bytes are set to 1,500 bytes. The cell coverage area is assumed to be circled with a diameter \( D \) of 1 km. Frame duration is 5 millisec and two thirds of the frame is for downlink data (3.33 msec). The path loss model and log-normal shadowing conforms to the 3GPP TR 25.848 technical report [15]. The model also considers noise and channel bandwidth effective SNR. Simulation parameters are shown in Table I. The simulation is run for 20,000 frames. MSs’ initial positions are uniformly distributed within a cell, i.e. the initial distance from the BS is between zero and \( D^2 \). MS’s initial speed is uniformly distributed within \([0-30]\) km/hr. MSs keep moving with their initial directions; either toward the BS or away from the BS. They keep moving with the same speed as the initial speed. When MS reaches the cell border, it rebounds back into the cell with the speed that is redetermined by using the same distribution, and with the direction toward the BS.

### Table I
SIMULATION PARAMETERS

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Number of subchannels</td>
<td>4</td>
</tr>
<tr>
<td>Number of subcarriers/subchannel</td>
<td>48</td>
</tr>
<tr>
<td>Number of slots/downlink subframe</td>
<td>80</td>
</tr>
<tr>
<td>Total transmit power</td>
<td>10 W</td>
</tr>
<tr>
<td>Shadowing (( \sigma ))</td>
<td>10 dB</td>
</tr>
<tr>
<td>Bit error rate</td>
<td>(10^{-5})</td>
</tr>
<tr>
<td>Exponential moving average weight factor (( t_c ))</td>
<td>100</td>
</tr>
</tbody>
</table>

We examine our scheme’s fairness and system utilization. The performance is compared with some other schemes. The maxC/I scheme always allocates resource to users who have the highest channel gain, in order to achieve the highest possible system utilization. On the other hand, the max-min scheme [16] focuses on fairness by allocating resource to users who have the worst channel gain. Fig. 2 shows the throughput fairness by using Jain’s fairness index [17]. The higher Jain’s fairness index, the better system’s fairness. In Fig. 2, maxC/I’s fairness index is the lowest and max-min’s is the highest (almost equal to 1 regardless of system load). PF’s fairness index is in between maxC/I’s and max-min’s. In our scheme, we calculate the fairness index of each group individually. The fairness indices of the guaranteed groups (0.75, 1, and 1.25 Mbps) of our scheme is very high (close to 1) and is nearly insusceptible to the system load. Fig. 3 shows more details of fairness indices of the guaranteed groups. The first guaranteed group (guaranteed 0.75 Mbps) has slightly higher fairness index than the other groups that request higher minimum bit rate. This is because for smaller request, the system has enough resource to accommodate, and hence, fairness can be better. Nonetheless, the fairness index of the guaranteed group is very high. In Fig. 2, the fairness index of the non-guaranteed group is still high when the system load is light. However, when the system load increases, the fairness index of the non-guaranteed group becomes lower. This is because the non-guaranteed group is served with the remaining resource after the guaranteed group. When the guaranteed group has more traffic, the non-guaranteed group receives less allocated resource, and hence, it is more difficult to provide fairness.
Fig. 4 shows the system utilization. System utilization is evaluated in the form of the number of scheduled bits per second per Hz in order to show spectral utilization. MaxC/I attains the highest system utilization. Our scheme’s utilization is comparable to that of PF and slightly lower when the system load increases. However, in comparison with PF, our scheme considers not only providing fairness but also providing minimum bit rate guarantee.

Fig. 5 shows the average output bit rate (among users of the same group) versus the user’s input bit rate. In our simulation, all users have the same input bit rate, regardless of their group. We are interested to see whether our scheme can differentiate and provide the output bit rate according to different minimum bit rate requests. We can see from Fig. 5 that the users’ output bit rate starts to be different when the input rate is around 0.8 Mbps. The average output bit rate of the users who request for 0.75 Mbps stays around 0.75 - 0.8 Mbps. Their output rate becomes lower when other users who request higher bit rate guarantee have more input bit rate. The output bit rate of the users who request 1 Mbps is around 0.92 - 1 Mbps. The output bit rate of the users who request 1.25 Mbps minimum bit rate is around 1.25 Mbps when their input rate reaches 1.53 Mbps. In addition, Fig. 5 shows that maxC/I and max-min schemes do not respect the minimum bit rate requests at all. Fig. 6 shows the average bit rate of each group when no users are in the non-guaranteed class. Each group has 16 guaranteed users.

We propose the resource allocation scheme for multicell service in an OFDMA downlink system. The frame-by-frame resource allocation scheme considers 2 classes. We guarantee minimum bit rate for the users of the guaranteed class by using PF-based optimization with minimum bit rate constraint integrating with the matrix-based method that exploits multi-channel multiuser diversity. The non-guaranteed class is served by using the conventional PF function. Simulation shows that the proposed scheme achieves very high fairness, and can differentiate and satisfy the required minimum bit rate as well.

V. Conclusion

We propose the resource allocation scheme for multicell service in an OFDMA downlink system. The frame-by-frame resource allocation scheme considers 2 classes. We guarantee minimum bit rate for the users of the guaranteed class by using PF-based optimization with minimum bit rate constraint integrating with the matrix-based method that exploits multi-channel multiuser diversity. The non-guaranteed class is served by using the conventional PF function. Simulation shows that the proposed scheme achieves very high fairness, and can differentiate and satisfy the required minimum bit rate as well.

REFERENCES