Scheduling Algorithms for Quality of Service Aware OFDMA Wireless Systems

Farrokh R. Farrokhi, Masoud Olfat, Mehdi Alasti, and K. J. Ray Liu
Electrical and Computer Eng. Department, University of Maryland, College Park, MD 20742

***Abstract*— In this paper, we investigate scheduling in multiuser OFDMA systems. The objective is to maximize the resource utilization in a network, while a required Quality of Service (QoS) per user is satisfied. A notion of revenue maximization is used to determine the optimal OFDMA subcarrier allocation to different users based on their required QoS. The optimal solution for the problem is presented and its performance and complexity are studied. An iterative low complexity algorithm is also proposed to achieve near optimal performance.**

I. Introduction

OFDMA is an efficient multiaccess scheme for broadband wireless access. In OFDM systems a user data stream is split into a number of lower rate substreams and each is modulated separately on one of the orthogonal subcarriers. In OFDMA, on the other hand, each transmitter is dynamically assigned to a subset of subcarriers. This capability enables the network to perform a flexible resource allocation with the goal of increasing the overall network throughput under varying traffic loads, channel conditions and multiuser interference. This leads to significant improvement in system throughput and spectral efficiency when the allocation of subcarriers to different users is performed carefully [1], [2]. On the other hand, the network has to maintain a required Quality of Service (QoS) for each user, which is not necessarily in line with maximizing the network's total throughput. In traditional OFDMA systems, a fixed subset of subcarriers in consecutive time slots are assigned to each user according to a static assignment [3]. In order to increase resource utilization in the network, a scheduler has to adjust the allocation of subcarriers to users based on their demands and link conditions with reasonable performance and low complexity.

The main objective in this paper is to allocate a subset of subcarriers to users such that the QoS is satisfied for each one, and at the same time the overall network throughput is maximized. In [4], for a single carrier system, we introduced a revenue model based on Service Level Agreement (SLA) that relates system throughput and QoS. An SLA is a contract between a user and the network that includes the requested QoS, pricing for the service provided and a penalty when the agreement is violated by the network. In this model, the scheduler is rewarded when total throughput is maximized and penalized when the QoS for a user is violated. We have shown that in order to maximize the throughput and meet QoS, the revenue function has to be maximized. In this work we extend that methodology to OFDMA systems. It should be mentioned that the solution for multicarrier systems is more complicated since the scheduler has to perform resource allocation in both time and frequency domain and unlike time slot allocation, subcarrier allocation affects the cost function for all users. Here, we investigate the optimal solution for this problem. Then we propose iterative scheduling algorithms that can achieve near optimal performance with significantly lower complexity than that of the optimal algorithm.

This paper is organized as follows: System model is introduced in Section II. The proposed scheduling algorithms are presented in Section III. In Section IV the performance of those algorithms are compared through numerical studies. Finally, Section V concludes the paper.

II. System Model

A single cell multiuser OFDMA system with N users and M subcarriers is considered (see Figure 1), in which the scheduler is able to assign any subset of subcarriers to different users for each OFDM symbol.Traffic from different users is directed to their assigned queues and each queue is served according to its user's QoS. Let r_n denote the rate reserved by user n ($n = 1, 2, \cdots, N$), i.e., assigned to queue n. It is well known that the QoS of a user can be translated into a minimum guaranteed rate (i.e., r_n) through the notion of effective bandwidth [5]. We assume that the queues of all users are backlogged, so they have packets to transmit at all times. Also, we assume that at each time slot a subcarrier can be assigned to only one user.

The maximum achievable rate per unit bandwidth at the m^{th} subcarrier for the n^{th} user is given by

$$g_n^m = \log_2(1 + |H_n^m|^2 P_n^m/(N_n^m\Gamma)), \qquad (1)$$

where H_n^m, P_n^m and N_n^m are the n^{th} user channel response, transmit power, and noise power at the m^{th} subcarrier, respectively, and Γ is the SNR gap [6].

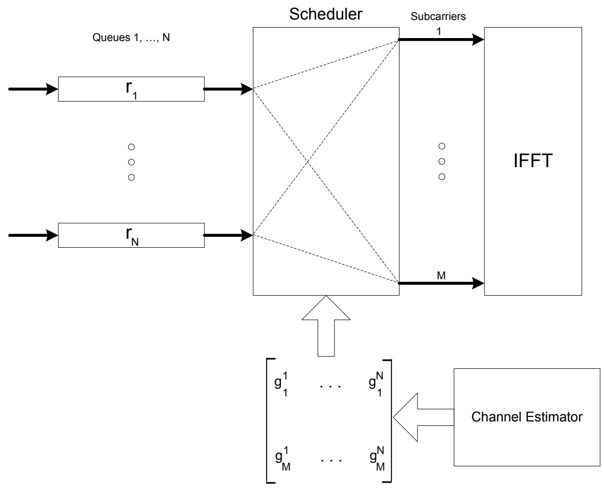

Fig. 1. System block diagram.

We assume that the base station has the knowledge of channel condition for each user, in other words, g_n^m is known at every time slot for all users and subcarriers [7]. We denote the set of subcarriers assigned to user n at time t by $S_n(t)$. Obviously, the total rate assigned to user n at time t is $\sum_{m \in S_n(t)} g_n^m(t)$.

III. Scheduling Algorithms

We introduce a revenue model where the network charges users based on the throughput it provides for them, and is penalized if the QoS defined in SLA for any user is violated. We assume that for the service received by user n at time t, i.e., $\sum_{m \in S_n(t)} g_n^m(t)$, the network charges the user by $\alpha_n \sum_{m \in S_n(t)} g_n^m(t)$. Here, α_n is the rate at which the n^{th} user is charged and is defined in its SLA.

To probe the QoS delivered to users, we define a credit for each user that indicates the amount of service the network has provided to the user [4], [8]. The credit for user n is denoted by $C_n(t)$ that evolves as follows:

$$C_n(t) = C_n(t-1) + r_n - \sum_{m \in S_n(t)} g_n^m(t)$$

where r_n and $\sum_{m \in S_n(t)} g_n^m(t)$, are the user reserved rate and the received service, respectively. We measure the QoS provided to users by their credits. If the network has provided a reserved rate to a user, that user's credit is close to zero, and if a user has not received the requested rate, its credit is high [4]. In this case, the network is penalized by $f[C_n]$, where $f[.]$ is a real, positive, convex and continuous function [4]. The penalty function that has a significant role in the performance of the SLA-based algorithm, is chosen in such a way that a user with negative credit does not penalize the system since this user has received its requested QoS; therefore, $f[x] = 0$ for $x \le 0$. One special example for $f[.]$ is:

$$f[x] = \begin{cases} \gamma x^2 & \text{if } x > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where γ is a positive number.

Let us denote $d_n(t)$ to be the scheduler revenue from user n at time t. Here, we perform scheduling in one OFDM symbol period, and maximize the total revenue at time t using a greedy algorithm. Hereafter, without loss of generality, we drop the index of time (t) in our discussions. Let us assume that C_n is the credit accumulated by user n before reassigning the subcarriers at the current OFDM sysmbol. The updated credit after assigning all subcarriers is given by $C_n + r_n - \sum_{m \in S_n} g_n^m$. As a result, we obtain:

$$d_n = \alpha_n \sum_{m \in S_n} g_n^m - f\left[C_n + r_n - \sum_{m \in S_n} g_n^m\right] \tag{3}$$

The total revenue of the network is given by $D = \sum_{n=1}^{N} d_n$. At any time slot, the optimal SLA-based scheduler, knowing r_n, C_n $(n = 1, \cdots, N)$ and also g_n^m $(n = 1, \cdots, N, m = 1, \cdots, M)$, assigns subcarriers to the users such that the total income (D) is maximized.

In the following, we present the optimal solution for subcarrier allocation in OFDMA systems and also an iterative suboptimal scheduling algorithm with a near optimal performance.

A. Optimal Solution

An exhaustive search among all possible assignments can achieve the optimal solution that maximizes the total revenue, $D = \sum_{n=1}^{N} d_n$. There are M subcarriers that can be assigned to N different users. Therefore, the total number of assignments is N^M. The set of all possible assignments can be illustrated by scheduling tree as shown in Figure 2. The leaf labeled with 1 shows the choice of allocating all subcarriers to user 1, in leaf N^M all are allocated to user N and in leaf 2, carriers $1 \ldots M-1$ are assigned to user 1 and carrier M to user 2. Other leaves are labeled accordingly. The exhaustive search algorithm evaluates the revenue for each leaf of the scheduling tree (all the possible assignments) and selects the one that achieves the maximum revenue.

If the average complexity of evaluating d_n is bounded from above by L, the complexity of performing the exhaustive search will be LN^{M+1}. Since the complexity of the algorithm grows exponentially with the number of subcarriers, exhaustive search may not be a practical solution. However, its performance can be used as a

 0-7803-8794-5/04/$20.00

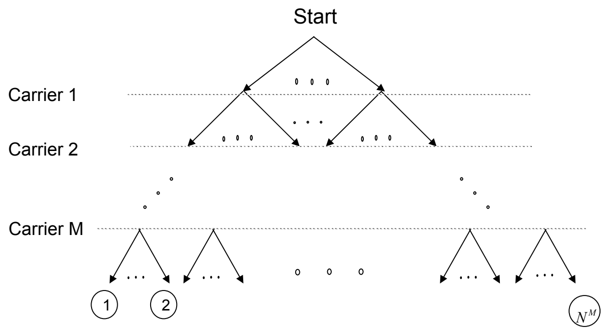

Fig. 2. Scheduling tree of subcarrier allocation.

reference point for the other algorithms. In the next sections, we propose a lower complexity algorithm with close to optimal performance.

B. Iterative Algorithm (IA)

As the simplest subcarrier assignment algorithm, we can assign subcarriers to users, one at a time. Assume that at the k^{th} step, there is a pool of subcarriers left. We define the revenue at step k (d_n^k) similar to (3) by replacing S_n with S_n^k, the set of subcarriers assigned to user n at the end of the k^{th} step, defined as

$$d_n^k = \alpha_n \sum_{m \in S_n^k} g_n^m - f\left[C_n + r_n - \sum_{S_n^k} g_n^m\right]$$

Also, assume that at the k^{th} step, a subcarrier denoted by m_k is to be assigned. The best user for this subcarrier is determined by:

$$\begin{aligned} \hat{n}_k &= \arg\max_n \{\alpha_n \sum_{m \in S_n^{k-1}} g_n^m + \alpha_n g_n^{m_k} \\ &- f[C_n + r_n - \sum_{m \in S_n^{k-1}} g_n^m - g_n^{m_k}]\} \end{aligned} \tag{4}$$

The above Successive Assignment (SA) algorithm is summarized as:

- Start with a subcarrier, calculate the revenue for each user and find the best user with the best revenue according to (4).
- Assign the subcarrier to the best user, and remove the subcarrier from the pool.
- Proceed to the next subcarrier.

The subcarrier selection order can be random or fixed. However, we have observed through numerical studies that the performances in both cases are similar. In other words, the performance of SA is independent of subcarrier assignment order. While having low complexity, this scheme performs far from the optimal solution since assignments in the future steps would change the revenue for the current assignment. In the following we modify the sequential assignment algorithm to achieve close to optimal solution. For this purpose, starting from an initial assignment (which can be obtained by a fixed, random or sequential assignment), we repeat the subcarrier assignment and refine the set of subcarriers assigned to each user in order to maximize revenue until they converge. The Iterative Assignment (IA) algorithm steps for the k^{th} step (assignment for the k^{th} subcarrier) are as follows:

1) Reassign the k^{th} subcarrier to the locally optimum user. In this step the revenues for all users are checked for a possible assignment. The subcarriers will be assigned to the users that maximizes the total revenue. That is,

$$\begin{aligned} \hat{n} &= \arg\max_n \{\alpha_n \sum_{m \in S_n^{k-1}} g_n^m + \alpha_n g_n^{m_k} \\ &- f[C_n + r_n - \sum_{m \in S_n^{k-1}} g_n^m - g_n^{m_k}]\} \end{aligned}$$

2) Assume that in the previous iterations (or initial assignment for the first iteration), the k^{th} subcarrier has been assigned to another user, say $\tilde{n}$. Then the assignments are updated as

$$\begin{cases} S_n^k = S_n^{k-1} & n \neq \tilde{n} \text{ or } n \neq \hat{n}, \\ S_{\tilde{n}}^k = S_{\tilde{n}}^{k-1} - \{k\} & , \\ S_{\hat{n}}^k = S_{\hat{n}}^{k-1} + \{k\} & . \end{cases} \tag{5}$$

3) Update the revenues for affected users, $\tilde{n}$, and $\hat{n}$.
4) Repeat steps 1-3 for the remaining subcarriers until all of the subcarriers are reassigned.
5) Repeat steps 1-4 until the revenue does not increase anymore.

In the above algorithm the revenue will increase at each step of the algorithm. On the other hand, total revenue has an upper bounds which is that of the optimal assignment. That is, in the above algorithm subcarrier assignment converges to a fixed point. Depending on the initial assignments, the algorithm may converge to a local or the global optimum. In order to improve the performance, we start the iteration from different and random initial assignments, and then pick the fixed point with the maximum overall revenue. As the number of random initial points is increased, the probability that the algorithm achieves the globally optimal assignment increases. We have observed that if in the first step of each iteration, reassignment of carriers is performed in random order, the convergence speed is increased.

We will show through numerical studies that this algorithm achieves near optimal performance with much lower complexity than that of exhaustive search. If we assume that we repeat the iterative algorithm from Q different initial points, and the average number of iterations to reach a fixed point is P, and the average complexity of evaluating the revenue for one user is L, the overall complexity of this scheme is $QPNML$. Considering typical numbers for N and M, this value is much less than LN^{M+1} for the exhaustive search method.

In order to reduce the complexity of subcarrier assignments, we can bundle subcarriers into a number of clusters. For instance, a total of M subcarriers can be bundled into M/k clusters, each with k subcarriers. Then, the algorithms presented in this paper can be applied to each cluster rather than subcarriers. By clustering, we can reduce the complexity at the expense of system performance. However, by proper choice of clusters performance degradation can be minimized for a given complexity.

IV. Simulations Results

We evaluate the performance of the proposed algorithms in a single cell system where four users ($N = 4$) are randomly distributed in the cell, and each user can be assigned to any of $M = 32$ subcarriers. We consider a multipath channel model with $R = 4$ distinct paths. The channel response for the n^{th} user can be represented as:

$$h_n(t) = \sqrt{G_n} \sum_{r=0}^{R-1} \alpha_n^r \delta(t - \tau_n^r), \quad (6)$$

where G_n includes log-normal shadow fading and path loss, τ_n^l and α_n^l denote the l^{th} path delay and fading, respectively. Each path fading is assumed to have a complex Gaussian distribution, so the received signal amplitude has a Rayleigh distribution. Channel frequency response can be represented simply by the Fourier transform of $h_n(t)$ sampled at each subcarrier center frequency, mf_c, where f_c is the subcarrier separation:

$$H_n^m = \sqrt{G_n} \sum_{l=0}^{L-1} \alpha_n^l e^{(-j2\pi m f_c \tau_n^l)}. \quad (7)$$

We assume that the path loss and shadowing for different paths of the same link are the same and any difference can be absorbed in the fading coefficients. In addition, the path loss and shadow fading can be compensated by a power control mechanism.

Numerical results from our simulations reveal that the Iterative Algorithm (IA) performs close to the optimal exhaustive search algorithm. Figure 3 shows the cumulative distribution functions (CDF) of the total revenue for the exhaustive search algorithm along with

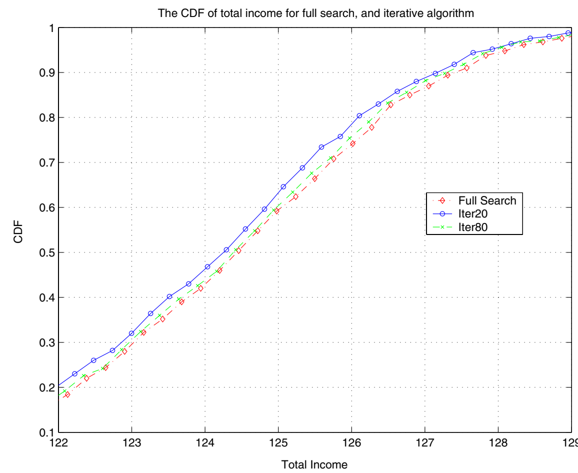

Fig. 3. CDF's of the optimal exhaustive search algorithm, iterative algorithm with 20 and 80 iterations

IA with 20 and 80 iterations. As shown in this figure, the performance of IA is close to the optimal one even with 20 iterations. Therefore, from now on, we can consider IA is as the reference algorithm.

For performance evaluation, we consider a fixed assignment (FA) scheduler, sequential assignment (SA), and IA. In the fixed assignment, we assume that the network assigns a set of subcarriers to each user, and this assignment stays unchanged over the course of communications. In the other algorithms, the subcarrier assignment is performed for one time slot, and changes from one time slot to another. Obviously, FA and the exhaustive search have the lowest and the highest complexities, respectively.

The total throughput versus the network load is shown in Figure 4. As it is illustrated in this figure, the IA achieves the maximum throughput (the expectation of maximum link capacity, $E\{\sum_{m=1} \max(g_n^m)\}$). FA achieves close to the average capacity of the system, $E\{g_n^m\}$. At low network loads, SA achieves the maximum throughput. However, its performance drops very fast as the load is increased. This is because this algorithm assigns the subcarriers independent of future assignments; therefore, the assignments at the early stages of the algorithm limit the performance of the later stages.

To present QoS, we measure the minimum assigned relative rate. If R_n is the assigned rate to user n, and r_n is the reserved rate by that user, we define the *minimum assigned relative rate* over all users as $\eta = \min_n\{\frac{R_n}{r_n}\}$. This value can be considered as the measure of QoS; to support QoS for all users, we want $\eta \geq 1$. This value is displayed versus the network load in Figure 5. Again IA satisfies QoS requirement for almost all loading values,

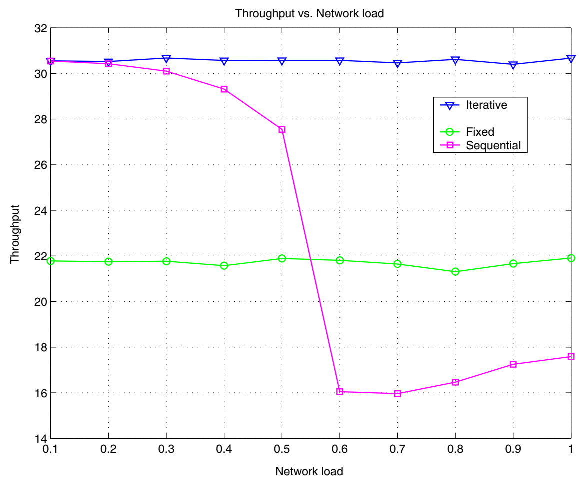

Fig. 4. Throughput vs. load

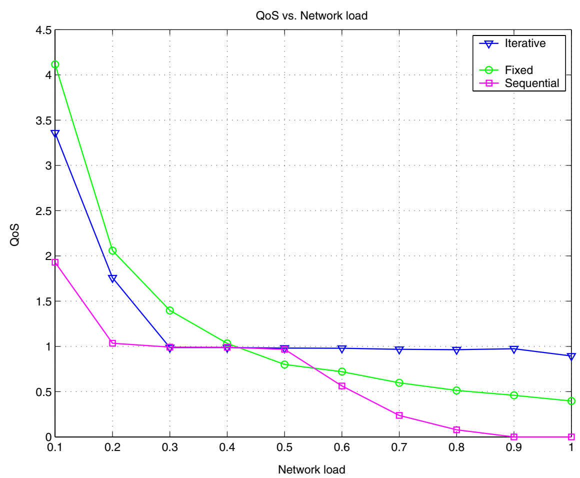

Fig. 5. Minimum assigned relative rate vs. network load

while the SA and FA fail to meet the QoS requirement for large loadings.

The total revenue is depicted versus network load in Figure 6. The revenue for IA is higher than that of the other algorithms since its throughput is the highest and the penalty for violating the QoS is the lowest. The performance of SA and FA drop significantly as the network load increases.

V. Conclusion

In this paper, we presented scheduling algorithms that maximize OFDMA system throughput for QoS sensitive users. We have used a notion of revenue maximization to balance throughput optimization and QoS. The OFDM subcarriers are allocated to different users based on their channel conditions and required rates, and also the total revenue of the system. An optimal solution and

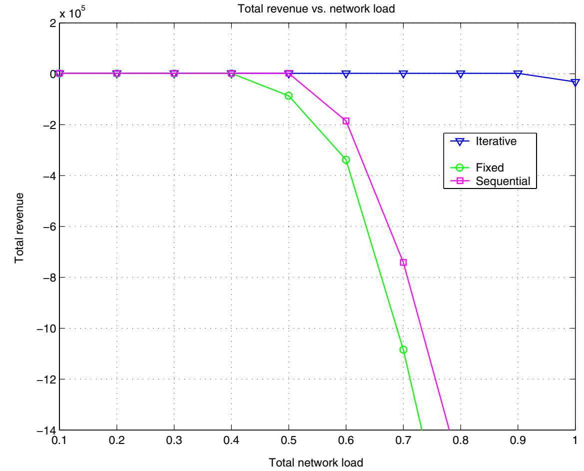

Fig. 6. Total revenue vs. load

its performances and complexities are studied. We have proposed low complexity iterative sequential scheduling algorithms, and through numerical studies we have shown that it would achieve near optimal performance. Also we have shown through revenue optimization, a scheduler can achieve a fair trade-off between the QoS and total throughput.

References

[1] I. Koutsopoulos and L. Tassiulas, " Channel State-Adaptive Techniques for Throughput Enhancement in Wireless Broadband Networks," *Proceedings of IEEE INFOCOM*, pp 757-766, 2001.

[2] W. Rhee and J. Cioffi, Increase in capacity of multiuser OFDM system using dynamic subchannel allocation, in Proc. Vehicular Technology Conference (VTC), 2000, pp. 1085 1089.

[3] IEEE 802.16 standard committee "IEEE Standard for Local and metropolitan area networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems", *http://standards.ieee.org/getieee802/download/802.16-2001.pdf, IEEE Std 802.16-2001*

[4] M. Alasti, F. R. Farrokhi, M. Olfat, K. J. Ray Liu, "Service Level Agreement (SLA) Based Scheduling Algorithms for Wireless Networks," *to appear in International Conference on Communications, ICC, 2004*.

[5] R. Guerin, H. Ahmadi, and M. Naghsineh, "Equivalent capacity and its application to bandwidth allocation in high-speed networks," *IEEE Journal of Selected Areas in Communications,*, Vol. 9, No. 7, September 1991.

[6] J. A. C. Bingham and J. M. Cioffi and P. S. Chow," A Practical Discrete Multitone Tranceiver Loading Algorithm for Data Transmission over Spectrally Shaped Channels", *IEEE Trans. On Communications,* pp 773-775, Feb., 1995.

[7] J. Gross and F. Fitzek, Channel state dependent scheduling policies for an OFDM physical layer using a binary state model, Tech. Rep. TKN-01-009, Telecommunication Networks Group, Technische Universitat Berlin, June 2001.

[8] A. C. Kam, K. Y. Siu, "Linear Complexity Algorithms for Bandwidth Reservations and Delay Guarantees in Input Queued Switches with no Speedup," *Proc. International Conference on Network Protocols 98*, Austin TX, Oct. 1998, pp. 2-11.

