Resource Allocation in OFDMA with Time-Varying Channel and Bursty Arrivals

Somsak Kittipiyakul and Tara Javidi, *Member, IEEE*

***Abstract*— This paper considers the issue of delay optimal subcarrier allocation in OFDMA wireless networks when the arrivals and channels are stochastic. Our objective is to minimize the long-term average packet delay over multiple time epochs. In previous studies, we have shown that the optimal policy is complicated and unknown. However, based on the insights learned from a simple On-Off channel model, we provide heuristic policies that use different degrees of channel and queue state information. More importantly, these examples show how the significance of queue vs. channel state information varies with the traffic load. This is of extreme practical interest when one considers the overhead associated with channel estimation and feedback.**

I. Introduction

ORTHOGONAL frequency-division multiple access (OFDMA) achieves high spectral efficiency in multiuser environment with channel state information (CSI) by dividing the total available bandwidth to orthogonal narrow sub-bands to be shared by users in an opportunistic manner [1]. By adaptively assigning subcarriers to the best users, multi-user water-filling, aka opportunistic scheduling, achieves the instantaneous maximum throughput by taking advantage of channel diversity among users (hence, providing *multiuser diversity gain*). When each user is assumed to have an infinite amount of data in the buffer, water-filling is throughput optimal. However, when considering stochastic packet arrivals with finite rate, water-filling is not throughput optimal [15].

Although the problem of optimal real-time subcarrier allocation has been extensively studied (see [4]-[14] and references therein), our work is philosophically different from most of literature: Often, in subcarrier allocation, an “instantaneous throughput” is maximized while assuming infinitely backlogged buffers; the algorithms are run periodically, to allow updates regarding channel variations and varying number of data streams. The problem with these techniques is that the scheduler fails to anticipate the impact of its allocation decision on the future state of the system. This paper, however, considers the system’s performance under realistic arrival patterns with finite rate. In other words, we focus on minimizing the long-term average performance i.e. average queue backlog over T epochs (or equivalently minimizing delay), rather than an instantaneous optimization. We argue that, under realistic arrival patterns, maximizing instantaneous throughput (water-filling) is too myopic and show that a key factor in delay improvement is the queue backlog state information (QSI). In this sense, the present paper provides a simple explanation for a series of problems observed and solutions proposed in OFDMA systems under realistic packet arrivals and queue occupancies ([12]-[14]). In this paper, we also show that the value of QSI vs. CSI varies with the arrival statistics. This is of significant practical interest, as one considers the overhead associated with CSI estimation and feedback in a multi-carrier system.

Manuscript received April 29, 2007. The associate editor coordinating the review of this letter and approving it for publication was Dr. Xi Zhang. This work was supported in part by the Center for Wireless Communications, UCSD and UC Discovery Grant No. Com04-10176, ARO-MURI Grant No. W911NF-04-1-0224, NSF CAREER Award No. CNS-0533035, and AFOSR Grant No. FA9550-05-01-0430.

The authors are with the Department of Electrical and Computer Engineering, University of California/San Diego, La Jolla, CA 92093 USA (email: skittipi@ucsd.edu; tara@ece.ucsd.edu).

Digital Object Identifier

II. Problem Formulation and Assumptions

A. Model and Notations

Consider a downlink single-hop OFDMA system composed of a base station and N users with infinite buffers. There are K OFDM subcarriers which are time-slotted. The users are homogeneous, i.e. they see *statistically* symmetric arrival and channel connectivity processes and have the same priority. Packets of fixed size arrive stochastically for each user. To accommodate the randomness in the arrival stream, there are N queues at the base station, one for each user, to buffer the data. At the beginning of each timeslot, the assignment of subcarriers to users is made by a centralized resource manager at the base station. The resource manager has perfect knowledge of the queue backlogs and the channel states which are assumed constant during a timeslot but varying independently over timeslots (block fading model). We do not allow sharing of any subcarriers. The assignment is announced immediately to all users via a separate control channel. In this paper, we assume that random fading channel conditions can be mapped into a matrix of connectivities $[c_{ij}]$. In other words, we consider the random connectivity matrix instead of the random fading channels of the users. Figure 1 shows an example where SNR across three subcarriers is mapped to the connectivity vector $(2, 1, 0)$.

The following random vectors are used throughout the paper.

- $\mathbf{b}(n) = (b_1, \ldots, b_N)$: Backlogs (in packets) of each queue at the beginning of timeslot n.
- $\mathbf{a}(n) = (a_1, \ldots, a_N)$: Packet arrivals to each queue during timeslot n.
- $C(n) = [c_{ij}]$: the $K \times N$ stochastic *connectivity matrix* at timeslot n where c_{ij} denotes the number of packets subcarrier i can serve from queue j.

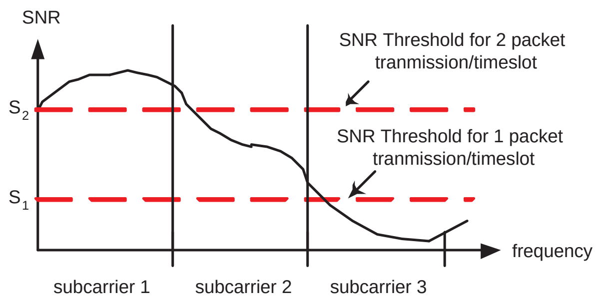

Fig. 1. Mapping of a received SNR channel profile of a user to packet capacity for each subcarrier. In this example, the user can potentially transmit 2, 1, and 0 packets on subcarriers 1, 2, and 3, respectively.

- $W(n) = [w_{ij}]$: the $K \times N$ allocation matrix at timeslot n, $w_{ij} = \begin{cases} 1 & \text{if subcarrier } i \text{ is assigned to queue } j, \\ 0 & \text{otherwise.} \end{cases}$
  We require $w_{ij} = 0$ if $c_{ij} = 0$ and $\sum_{j=1}^{N} w_{ij} \leq 1$ since sharing of subcarriers is not allowed.

III. Instantaneous Throughput Maximizing and Load-Balancing Policies

In this section, we consider two classes of server allocation policies: a class comprising of instantaneous throughput maximizing (IMT) policies and another class of load-balancing (LB) policies. As discussed previously, each class represents one of the competing goals: an IMT policy maximizes the number of packets being served now, while an LB policy maximizes the expected number of non-empty queues (hence, served packets) in the future.

Instantaneous Throughput Maximizing: An allocation $W^*(n) = [w^*_{ij}]$ is IMT if it achieves the maximum throughput at time n, i.e., for any feasible allocation $W(n) = [w_{ij}]$,

$$\sum_{j=1}^{N} \sum_{i=1}^{K} w^*_{ij} c_{ij} I_{\{b_j > 0\}} \geq \sum_{j=1}^{N} \sum_{i=1}^{K} w_{ij} c_{ij} I_{\{b_j > 0\}}. \tag{1}$$

where $I_{\{x>0\}} = 1$ if $x > 0$ and 0 otherwise.

To introduce the LB policy, we need the following definition:

Definition 1: We say $\mathbf{x} \leq_{LQO} \mathbf{y}$ ($\mathbf{x}$ is more balanced than $\mathbf{y}$) iff $ord(\mathbf{x}) \leq_{lex} ord(\mathbf{y})$ where vector $ord(\mathbf{v})$ has the ordered elements of $\mathbf{v}$ in descending order, and the relation $\leq_{lex}$ on $\mathbb{R}^N$ is the *lexicographic ordering*.

Load Balancing: An LB allocation $W^*(n) = [w^*_{ij}]$ is such that for any feasible allocation $W(n) = [w_{ij}]$,

$$[\mathbf{b} - \mathbf{1}(W^* \odot C)]^+ \leq_{LQO} [\mathbf{b} - \mathbf{1}(W \odot C)]^+ \tag{2}$$

where $\odot$ is an element-wise product, $\mathbf{1}$ is a row vector of ones, and $[\mathbf{v}]^+ = [v_1^+, \ldots, v_N^+]$ with $v_j^+ = v_j I_{\{v_j > 0\}}$.

Max Throughput Load Balancing (MTLB): is an allocation that is both IMT and LB.

Note that when $\max c_{ij} > 1$, MTLB policy might not exist [2], raising the question as to the relative importance of maximizing instantaneous throughput versus balancing. In [3], we have shown that a delay optimal policy, in general, can be a complicated mixture of IMT and LB allocations. Meanwhile, however, we have also shown that in the special case where c_{ij} only takes values 0 (OFF) or 1 (ON), 1) MTLB policy always exists and 2) is delay optimal.

In the next section, we use the obtained insights and the structure of MTLB policy for the ON-OFF channel to design several heuristic policies. We compare the performance of these policies via simulations.

IV. Heuristic Policies and Value of QSI versus CSI

1) *Algo-I* (full QSI, On-Off CSI): The subcarrier assignment uses full information about the queue lengths (full QSI) and binary (ON-OFF) information about the channel: a subcarrier is considered ON if $c_{ij} \geq c_{\text{threshold}}$[1]. Then, MTLB policy [2] is used for subcarrier allocation.

Algo-I:

- $\bar{c}_{ij} = \begin{cases} 1 & \text{if } c_{ij} \geq c_{\text{threshold}}, \\ 0 & \text{otherwise.} \end{cases}$
- For state $(\mathbf{b}/c_{\text{threshold}}, \bar{C})$ compute an MTLB allocation W^*.

2) *Algo-II* (full CSI, On-Off QSI): The subcarrier allocation uses full information about the channel (full CSI) and minimal information about the queue lengths (On-Off QSI). The subcarrier allocation considers only the queues which have some data to transmit.

Algo-II:

- Assign $W^* = [w^*_{ij}]$ such that $w^*_{ij^*(i)} = 1$ where $j^*(i) = \arg\max_{j \in \{1,\ldots,N\}} c_{ij} I_{\{b_j > 0\}}$.

3) *Algo-III* (full CSI and full QSI): Algo-III is the Maximum-Weight policy proposed in [15] where each subcarrier is assigned to the queues to achieve the highest value of $c_{ij} b_i$.

Algo-III:

- Assign $W^* = [w^*_{ij}]$ such that $w^*_{ij^*(i)} = 1$ where $j^*(i) = \arg\max_{j \in \{1,\ldots,N\}} b_j c_{ij}$.;

4) *Algo-IV* (full CSI and full QSI): Since Algo-III may over assign subcarriers to some users, and with an insight into the significance of load-balancing, we modify the known Algo-III to avoid imbalanced queues.

Algo-IV:

- $X = \{1, \ldots, K\}$;
- Loop (until stop):
  - If $X = \emptyset$, then stop;
  - $(i^*, j^*) = \arg\max_{i \in X, j \in \{1,\ldots,N\}} b_j c_{ij}$;
  - If $b_{j^*} c_{i^* j^*} > 0$, then $w^*_{i^* j^*} = 1$ else stop;
  - $b_{j^*} = b_{j^*} - c_{i^* j^*}$ and $X = X - \{i^*\}$;
- Assign $W^* = [w^*_{ij}]$.

A. Numerical Comparisons and Simulation

We consider a downlink OFDMA system in a single cell with one base station composed of $N = 32$ statistically independent and identical users and $K = 128$ subcarriers. We generate a frequency-selective channel by using 26-tap multipath with exponential intensity profile and use adaptive QAM modulation. We use the fact that there is only a very

[1]The threshold is arbitrarily chosen. It should be adjusted depending on the load of the system and the number of subcarrier and users. We assume here that the threshold is fixed in our simulation.

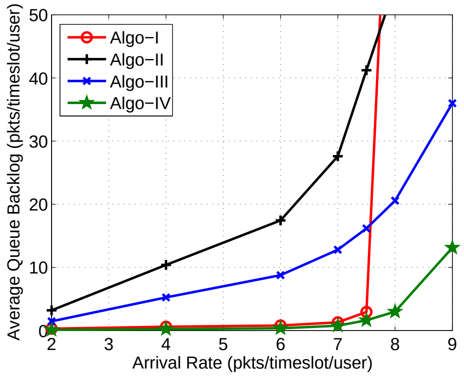

Fig. 2. Average queue backlog for bounded pareto distribution.

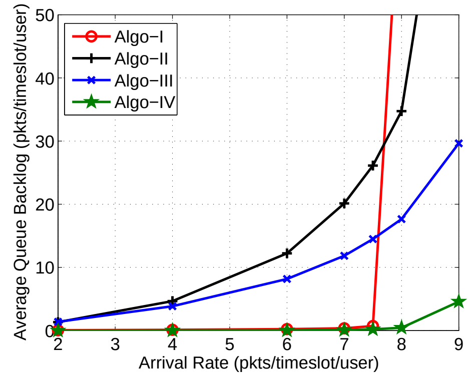

Fig. 3. Average queue backlog for the Poisson distribution.

small loss of channel capacity if a white power spectrum is used (i.e. each subcarrier receives equal power) instead of the optimal power spectrum [4]. The channel gain h_{ij} can be mapped to the number of packets per time slot, c_{ij}, that subcarrier i can potentially transmit for user j as [4]:

$$c_{ij} = \frac{D}{\beta} \max\left\{0, \left\lfloor 0.31(10\log_{10}(\frac{P|h_{ij}|^2}{KN_o}) - 6.7) \right\rfloor\right\} \quad (3)$$

where D the number of QAM symbols per channel in a timeslot, β the fixed packet length (in bits) and N_o is the noise power in the subcarrier. The parameters P, D, β and N_o are chosen such that the allocation of subcarriers over a block is equivalent to the server scheduling problem where the connectivity $c_{ij} \in \{0, 1, 2, 3\}$. All simulations are conducted over 6,000 timeslots. We consider arrivals of fixed-size packets where the number of arrivals per timeslot for each queue is a random variable having one of the two distributions: bounded pareto ($\alpha = 2$ and $x_{min}/E(x) = 1/2$) [16] and Poisson.

Figures 2 and 3 provide comparisons of the performance of the proposed algorithms under different traffic models in terms of the average total queue backlog (equivalently, in terms of the average delay by Little's Theorem). For all traffic types, Algo-IV, as expected, outperform the others, since it mimics IMT and LB policies as closely as possible. This observation is consistent with those studies in literature which take advantage of backlog information (e.g. [13] and [14]). However, the more interesting and important observation is the performance of Algo-I in the light-to-moderate traffic regime (below 6 packets/user/timeslot). Before the performance of Algo-I sees a sharp degradation reflecting the policy's low throughput, it outperforms Algo-II and Algo-III in light-to-moderate traffic, even though it does ignore much of CSI available to the other algorithms. This insight sheds light on nature of delay performance versus throughput considerations and the benefit of using queue information. When considering light-to-moderate traffic intensity (resulting in reasonable delays), the value of QSI outplays that of CSI. This is of extreme practical interest when one considers the overhead associated with CSI estimation and feedback in an OFDMA system with large number of subcarriers.

V. Conclusion and Future Research

In this paper, we considered the problem of subcarrier allocation in OFDMA system. We argued that conventional water-filling policies based on maximizing instantaneous throughput are too myopic and ignore important information about state variable (queue length). We showed by simulation that the value of CSI and QSI in optimizing the performance heavily depends on the arrival statistics. We showed that in low-to-moderate traffic regime and from a delay optimality perspective, balancing the queues is more critical than opportunistically taking advantage of CSI. The opposite becomes true in the heavy traffic regime.

References

[1] R. Van Nee and R. Prasad, *OFDM for wireless multimedia communications*, Artech House, Boston, c2000.

[2] S. Kittipiyakul and T. Javidi, "Subcarrier allocation in OFDMA systems: beyond water-filling," *2004 Asilomar Conference on Signals, Systems, and Computers*, Nov. 2004.

[3] S. Kittipiyakul and T. Javidi, "A fresh look at optimal subcarrier allocation in OFDMA systems," *IEEE Control & Decision Conference*, 2004

[4] A. Czylwik, "Adaptive OFDM for wideband radio channels," *Proc. GLOBECOM '96*, vol. 1, pp 713-718.

[5] S. Pietrzyk and G. J. M. Janssen, "Multiuser subcarrier allocation for QoS provision in the OFDMA systems," *IEEE VTC*, Sept. 2002.

[6] G. Zhang, "Subcarrier and bit allocation for real-time services in multiuser OFDM systems," 2004 IEEE International Conf. on Communications, 2004.

[7] Y. Zhang and K. Lataief, "Adaptive resource allocation and scheduling for multiuser packet-based OFDM networks," *IEEE ICC 2004*, Paris, June 2004.

[8] J. Jang, K.B. Lee, and Y.H. Lee, "Transmit power and bit allocations for OFDM systems in a fading channel," *Proc. IEEE GLOBECOM 2003*, Dec. 2003.

[9] G. Munz, S. Pfletschinger, and J. Speidel, "An efficient water-filling algorithm for multiple access OFDM," *IEEE Globecom '02*, Taipei, Taiwan, November 2002.

[10] W. Rhee and J.M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," *Proc. of Vehicular Tech. Conf. (VTC), 2000*, vol. 2, pp.1085-1089, May 2000.

[11] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks - part I: theoretical framework," *IEEE Trans. Wireless Comm.*, v.4, no.2, March 2005.

[12] J. Gross, J. Klaue, H. Karl, and A. Wolisz, "Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming," *Computer Communications*, 2004.

[13] G. Li and H. Liu, "Dynamic resource allocation with finite buffer constraint in broadband OFDMA networks," *IEEE Wireless Comm. and Networking*, v. 2, pp. 1037-1042, March 2003.

[14] G. Song, Y. Li, L. Cimini, Jr., and H. Zheng, "Joint channel-aware and queue-aware data scheduling in multiple shared wireless channels," *Globecomm 2003*, San Francisco, 2003.

[15] T. Javidi, "Rate stable resource allocation in OFDM systems: from waterfilling to queue-balancing," *Allerton Conference on Communication, Control, and Computing*, September 2004.

[16] D. von Seggern, *CRC Standard Curves and Surfaces.* Boca Raton, FL: CRC Press, p. 252, 1993