

# Cross-layer QoS Analysis of Opportunistic OFDM-TDMA and OFDMA Networks

Yu-Jung Chang, Feng-Tsun Chien, *Member, IEEE*, and C.-C. Jay Kuo, *Fellow, IEEE*

**Abstract**—Performance analysis of multiuser orthogonal frequency division multiplexing (OFDM-TDMA) and orthogonal frequency division multiple access (OFDMA) networks in support of multimedia transmission is conducted in this work. We take a cross-layer approach and analyze several quality-of-service (QoS) measures that include the bit rate and the bit error rate (BER) in the physical layer, and packet average throughput/delay and packet maximum delay in the link layer. We adopt a cross-layer QoS framework similar to that in IEEE 802.16, where service classification, flow control and opportunistic scheduling with different subcarrier/bit allocation schemes are implemented. In the analysis, the Rayleigh fading channel in the link layer is modeled by a finite-state Markov chain, and the channel state information (CSI) is assumed to be available at the base station. With the M/G/1 queueing model and flow control results, our analysis provides important insights into the performance difference of these two multiaccess systems. The derived analytical results are verified by extensive computer simulation. It is demonstrated by analysis and simulation that OFDMA outperforms OFDM-TDMA in QoS metrics of interest. Thus, OFDMA has higher potential than OFDM-TDMA in supporting multimedia services.

**Index Terms**—quality of services (QoS), multiple access, OFDM, OFDMA, cross-layer analysis, opportunistic scheduling.

## I. INTRODUCTION

MULTIMEDIA delivery is one of the key objectives of next-generation wireless networks. Its success relies on how the underlying network can support different QoS requirements demanded by a variety of multimedia applications. A significant challenge is posted since multimedia applications have very diverse characteristics in terms of physical measures such as bandwidth and delay. Furthermore, it is desirable that the underlying network can serve multiple users and meet their individual QoS requirement. All of these call for a QoS-provisioning broadband network in conjunction with proper multiple access schemes such as Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA). Recently, OFDM-based networks in combination with TDMA and FDMA have become a popular choice for

such an endeavor. The IEEE 802.16 standard, for instance, has adopted OFDM-TDMA and OFDMA (OFDM-FDMA) as two transmission schemes at the 2–11 GHz band [1]. In addition, a QoS framework in the medium access control (MAC) layer has also been integrated with the multiaccess transmission systems in the IEEE 802.16 standard [2].

Multiuser diversity provided by opportunistic scheduling [3] has been incorporated in multiuser OFDM-TDMA and OFDMA networks recently. Its effect on QoS-provisioning is an interesting yet challenging topic. On one hand, by allocating resources to users with better channel quality, the opportunistic scheduling scheme can maximize the overall system throughput [3]. On the other hand, it may degrade other QoS metrics such as delay, since users are suspended from transmission when their channels are poor. The impact of opportunistic scheduling on QoS provisioning is investigated in the course of comparison of OFDM-TDMA and OFDMA in this work. The bit error rate (BER) performance of OFDM-TDMA and OFDMA with multiuser diversity has been studied in previous work, *e.g.*, [4], [5]. Specifically, uncoded and coded systems with opportunistic OFDMA were shown to outperform those with static OFDM-TDMA by 3 dB and 7 dB at  $\text{BER} = 10^{-3}$  in [4] and [5], respectively. The BER performance of OFDM-TDMA and OFDMA was also compared in [6] without considering multiuser diversity.

While the BER analysis can be used to characterize the physical layer performance, it is not sufficient to reflect other QoS metrics such as packet throughput and delay in the link layer. The fact that QoS requirements should be treated differently in different layers suggests a cross-layer approach for QoS provisioning and analysis. In fact, the cross-layer approach has been applied to the design and analysis of QoS-featured multiaccess systems by a few researchers recently. For example, the analysis of queueing delay for 802.16 networks was conducted in [7], [8] by combining link-layer queueing with physical-layer transmission. A vacation queueing model was adopted in [9] to analyze the link-layer queueing performance of OFDM-TDMA systems with round-robin scheduling. A queueing model for OFDMA systems was used in [10] to design a scheduling scheme that balances multiuser diversity and queueing delay.

Although the packet-level analysis has been conducted for OFDM-TDMA or OFDMA in [7]–[10], there are a few open issues to be addressed. We discuss these issues and point out our contributions below. First, an analytical framework to account for both OFDM-TDMA and OFDMA systems to facilitate their comparison is missing. By generalizing results in [11], [12], we propose a framework to achieve this goal here. Second, performance evaluation of 802.16 has been

Manuscript received May 15, 2006; revised November 25, 2006. This research was funded by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative agreement No. EEC-9529152. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation.

Yu-Jung Chang and C.-C. Jay Kuo are with the Integrated Media Systems Center and the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564, USA (e-mails: yujungc@usc.edu and cckuo@sipi.usc.edu).

Feng-Tsun Chien is with the Department of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. (e-mail: ftchien@mail.nctu.edu.tw).

Digital Object Identifier 10.1109/JSAC.2007.070503.

conducted primarily by simulation in the past, *e.g.*, [13], [14]. We conduct an analysis to demonstrate that OFDMA outperforms OFDM-TDMA in terms of several QoS metrics. This is consistent with the trend of the latest IEEE Standard 802.16e-2005 [15], which adopts OFDMA as its principal multiaccess scheme. Third, although packet average delay and maximum delay are useful link-layer performance measures for non-real-time (*e.g.*, file transfer and web browsing) and real-time (*e.g.*, voice and video) applications, respectively, most previous work has focused on the packet average delay. The packet maximum delay and the delay violation probability will be examined in this work. Finally, the performance of two well-known scheduling strategies, namely, the round-robin scheduling and the opportunistic scheduling, is examined for OFDM-TDMA and OFDMA networks so as to understand their pros and cons in the context of QoS provisioning.

Our approach to physical and link layer analysis is simply stated as follows. The ideal channel state information (CSI) is assumed to be available at the base station. This is often achieved by feeding the estimated channel information from the receiver back to the transmitter through a control channel. The Rayleigh fading channel is modeled by a finite-state Markov chain [16] to translate the effect of the physical layer to higher layers. Specifically, the channel effect is manifested in the link layer as a time-varying server with the M/G/1 queueing model for the packet throughput-delay analysis. To analyze the packet maximum delay, we adopt and implement a well-known flow control scheme [17] as part of the proposed QoS framework. With such a flow control scheme, we can derive delay bounds based on network calculus results for different scheduling and rate adaptation schemes. All the aforementioned performance metrics provide valuable measures in differentiating OFDM-TDMA and OFDMA in their QoS-provision performance. It is finally concluded that OFDMA has a higher potential in meeting the requirements of multimedia delivery.

The rest of this paper is organized as follows. The system model, which consists of the QoS-aware framework, the flow control regulation scheme, multiple access and resource allocation schemes, is discussed in Sec. II. The physical and link layer analysis for OFDM-TDMA and OFDMA is presented in Sec. III. Simulation results are shown and discussed in Sec. IV. Finally, concluding remarks are given in Sec. V.

## II. SYSTEM MODEL

The proposed cross-layer QoS framework is shown in Fig. 1(a), where QoS provision is achieved by packet categorization and service differentiation. Specifically, packets are classified into the premium and the best-effort two classes and better service is granted to the premium class. Particularly, a delay-sensitive service is offered to the premium class by the preemptive priority scheduling mechanism, where premium packets maintain a higher priority for processing in our framework. The reason to choose delay as a main performance metric for service differentiation is that the delivery of many multimedia applications are delay-sensitive. As shown in Fig. 1(b), a flow control scheme is adopted on top of service differentiation to regulate the end-to-end packet delay of the premium and the best-effort classes. The proposed QoS

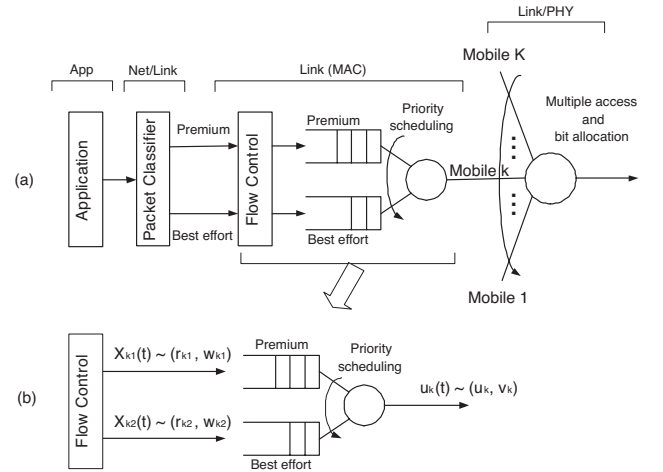


Fig. 1. (a) A cross-layer QoS-support system model and (b) a queueing system with flow-control regulated streams and preemptive priority servicing, shown for mobile user  $k$ .

framework is a simplification of those used in DiffServ [18] and the 802.16 MAC protocol [2].

### A. Flow Control Regulation

A flow control regulator [17] is adopted to process real-time multimedia data so as to keep the delay bound and arrival constraints [19]. As depicted in Fig. 1(b),  $X_{k1}(t)$  and  $X_{k2}(t)$  are the flow-control regulated premium and best-effort streams of user  $k$ , respectively. We write  $X_{k1}(t) \sim (r_{k1}, w_{k1})$  if, for any  $t_1 \leq t_2$ ,

$$\int_{t_1}^{t_2} X_{k1}(t) dt \leq r_{k1}(t_2 - t_1) + w_{k1}, \quad (1)$$

where  $r_{k1}$  is the predefined average rate of the stream and  $w_{k1}$  is the allowed burst degree. In other words, the input stream has to be regulated so that the output stream,  $X_{k1}(t)$ , can meet the imposed rate and burstiness constraints. Likewise, we have  $X_{k2}(t) \sim (r_{k2}, w_{k2})$ .

Suppose that the time-varying server process,  $u_k(t)$ , in Fig. 1(b) conforms to a similar but slightly different constraint. That is, we write  $u_k(t) \sim (\bar{u}_k, v_k)$  if, for any  $t_1 \leq t_2$ ,

$$\int_{t_1}^{t_2} u_k(t) dt \geq \bar{u}_k(t_2 - t_1) - v_k, \quad (2)$$

where  $v_k$  is the service lag and  $\bar{u}_k$  is the average service rate defined by

$$\bar{u}_k = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t u_k(\tau) d\tau, \quad (3)$$

with probability one. We will show in Sec. III-C that the server process  $u_k(t)$ , which is prescribed by actual multiaccess and scheduling schemes, satisfies (2) asymptotically, and parameters  $\bar{u}_k$  and  $v_k$  can be derived analytically.

### B. Multiple Access and Scheduling

The multiple access scheme in Fig. 1(a) is accomplished by OFDMA or OFDM-TDMA along with subcarrier/time slot assignment and bit allocation mechanisms. Several different schemes to be examined are summarized in Table I. Note that

TABLE I  
MULTIACCESS OFDM MODES CONSIDERED IN THIS WORK

OFDMA Mode	Subcarrier Assignment	Bit Allocation
OFDMA I	static	fixed
OFDMA II	static	adaptive modulation
OFDMA III	dynamic	adaptive modulation
OFDM-TDMA Mode	Time-slot Assignment	Bit Allocation
OFDM I	static	fixed
OFDM II	static	adaptive modulation
OFDM III	dynamic	adaptive modulation

we use OFDM *time slot* and OFDM *symbol* interchangeably in this paper.

OFDMA and OFDM-TDMA perform multiple access in a frequency-sharing and a time-sharing manner, respectively. Specifically, OFDMA performs subcarrier assignment while OFDM-TDMA performs time-slot assignment, both statically or dynamically. The difference between static (or round-robin) and dynamic (or opportunistic) assignments lies in whether users' channel conditions are considered. For OFDMA, static assignment allocates an equal, fixed and interleaved set of subcarriers to users while dynamic assignment allocates each subcarrier to the user with the best signal-to-noise ratio (SNR). Likewise, for OFDM-TDMA, static assignment allocates fixed and alternate time slots to users (*i.e.*, round robin) whereas dynamic assignment assigns a time slot to the user with the best channel condition. Besides, the chosen user is allocated all subcarriers exclusively in OFDM-TDMA.

The subcarrier SNR distribution for each multiaccess mode can be obtained as a basis for analysis in Sec. III. For a Rayleigh fading channel, the received SNR, denoted by  $\Gamma$ , is exponentially distributed with the following probability density function (pdf) [16]:

$$g_{\Gamma}(\gamma) = \frac{1}{\gamma_0} \exp\left(-\frac{\gamma}{\gamma_0}\right), \quad \gamma \geq 0, \quad (4)$$

where  $\gamma_0$  is the average SNR. Note that Eq. (4) holds for subcarrier SNR in a multicarrier system as well [20]. Eq. (4) applies to all modes in Table I except for OFDMA III up to a difference in mean  $\gamma_0$  which will be verified by simulation. In OFDMA III, recall that the dynamic assignment scheme assigns a subcarrier to the user with the best SNR at that subcarrier. Thus, when  $K$  homogeneous users are considered, the post-assignment subcarrier SNR,  $\Gamma^*$ , is distributed according to the maximum of  $K$  independent and identically distributed (i.i.d.) random variables  $\Gamma_1, \dots, \Gamma_K$ , which represent the received SNR of users  $1, \dots, K$  and follow the pdf in (4). To derive the pdf of  $\Gamma^*$ , we first obtain the cumulative distribution function (cdf):

$$\begin{aligned} P\{\Gamma^* \leq \gamma\} &= P\{\max(\Gamma_1, \dots, \Gamma_K) \leq \gamma\} \\ &= (1 - \exp(-\frac{\gamma}{\gamma_0}))^K, \quad \gamma \geq 0. \end{aligned}$$

Then, by differentiating the cdf, we have

$$g_{\Gamma^*}(\gamma) = \frac{K}{\gamma_0} \exp\left(-\frac{\gamma}{\gamma_0}\right) (1 - \exp(-\frac{\gamma}{\gamma_0}))^{K-1}, \quad \gamma \geq 0. \quad (5)$$

### C. Bit Allocation

The multiaccess scheme apportions resource among users while the bit allocation scheme chooses the type and order of

modulation for each user. We consider squared  $M$ -QAM modulations with  $M = 2^{2r}$ ,  $r = 1, \dots, r_m$ , where  $r_m$  determines the highest modulation allowed. Both fixed and (discrete-rate) adaptive modulation (AM) methods are considered, where only the AM methods take channel conditions into account in adaptive bit allocation [21]. The AM methods are described below. A tight BER approximation for squared  $M$ -QAM is given by [21]:

$$P_b = 0.2 \exp\left(-\frac{3\beta}{2(M-1)}\right), \quad (6)$$

where  $\beta$  is the channel SNR. With continuous-rate adaptation, bit rate  $R_b^c$  is given by the following capacity expression:

$$R_b^c = \log_2 M = \log_2\left(1 + \frac{1.5}{-\ln 5P_b}\beta\right). \quad (7)$$

Note that (7) is obtained directly by the rearrangement of (6).

Discrete-rate adaptation confines the bit rate to integer values (more precisely to  $2r$ ,  $r = 0, \dots, r_m$ ), which is described as follows. First, the set of possible received SNR (*i.e.*, the nonnegative real line) is partitioned into  $r_m + 1$  disjoint regions  $\mathbb{R}_0, \dots, \mathbb{R}_{r_m}$  by boundary points  $b_0, b_1, \dots, b_{r_m+1}$ , where  $\mathbb{R}_r$  is the interval  $[b_r, b_{r+1})$  for  $r = 0, 1, \dots, r_m$  and  $b_0 < b_1 < \dots < b_{r_m+1}$  with  $b_0$  and  $b_{r_m+1}$  set to 0 and  $\infty$ , respectively. Second, the boundary points are determined by

$$b_r = -\frac{2}{3}(\ln 5P_b)(2^{2r} - 1), \quad r = 1, 2, \dots, r_m. \quad (8)$$

Last, when the received SNR falls in  $\mathbb{R}_r$  and the information is successfully fed back to the transmitter,  $2r$  bits are loaded to the corresponding subcarrier. Note that the channel is too poor to support any order of modulation when the SNR falls in  $\mathbb{R}_0$ . This leads to the following bit rate expression of discrete-rate adaptation:

$$R_b^d = \begin{cases} 2\lfloor \frac{1}{2} \log_2(1 + \frac{1.5}{-\ln 5P_b}\beta) \rfloor, & \text{if } \beta < b_{r_m}, \\ 2r_m, & \text{if } \beta \geq b_{r_m}, \end{cases} \quad (9)$$

where  $\lfloor x \rfloor$  represents the largest integer that is less than or equal to  $x$ . Note that for any  $P_b$  and  $\beta$ ,  $R_b^d \leq R_b^c$ .

## III. PERFORMANCE ANALYSIS AND COMPARISON

Multiaccess schemes in Table I are analyzed in this section based on the system model introduced in Sec. II. The physical layer performance is considered in Sec. III-A while the link layer performance metrics are examined in Secs. III-B and III-C.

### A. Bit Rate and BER Analysis

For the bit rate and BER analysis, we focus on AM-based modes (*e.g.*, OFDMA II–III, OFDM II–III) since non-AM modes are trivial special cases. The theoretical bit rate upper bounds are derived under the assumption of continuous-rate adaptation. However, such bounds also hold for the discrete-rate adaptation case.

For fixed  $P_b$  and  $\beta$ , the bit rate per subcarrier with continuous-rate adaptation is given in (7). With the distribution

of  $\beta$  shown in (4) for OFDMA II and OFDM II-III, the bit rate corresponding to these modes is bounded by

$$\begin{aligned} R_{b1} &= E_{\Gamma}[\log_2 M(\Gamma)] \leq \log_2 E_{\Gamma}[M(\Gamma)] \quad (10) \\ &= \log_2 \int_0^{\infty} M(\gamma) g_{\Gamma}(\gamma) d\gamma \\ &= \log_2 \left(1 + \frac{1.5}{-\ln 5P_b} \gamma_0\right). \end{aligned}$$

The total bit rate  $R_{t1}$  (in the unit of bits/sec) can be obtained by scaling with the total number of subcarriers  $N$  and OFDM symbol time  $T_s$  as

$$R_{t1} \leq \frac{N}{T_s} \log_2 \left(1 + \frac{1.5}{-\ln 5P_b} \gamma_0\right). \quad (11)$$

Similarly, the total bit rate for OFDMA III,  $R_{t2}$ , is obtained by replacing  $\Gamma$  by  $\Gamma^*$  in (10) and using (5). That is, we have

$$R_{t2} \leq \frac{N}{T_s} \log_2 \left(1 + \frac{1.5}{-\ln 5P_b} \gamma_0 \left(\sum_{k=1}^K \frac{1}{k}\right)\right). \quad (12)$$

For the BER analysis, note that the AM scheme described in Sec. II-C will lead to comparable BER performance for all OFDMA and OFDM modes due to the predetermined  $P_b$  in place. A more informative comparison in BER can be achieved by fixing a target bit rate for all modes. Towards this end, we adopt a method based on [22] to add and subtract bits from proper subcarriers successively until the target bit rate is achieved. That is, when the actual bit rate is smaller (or larger) than the target bit rate, additional bits are added to (or subtracted from) subcarriers such that the error probability is increased as small as possible (or decreased as far as possible). In other words, additional bits are successively added to (or subtracted from) the subcarrier where the difference between capacity ( $R_b^c$  in (7)) and assigned discrete-rate bits ( $R_b^d$  in (9)) is maximal (or minimal).

The equivalence between “maximizing the difference between capacity and the assigned bits” and “minimizing the bit error rate” is intuitive and it can be formally proved. The proof is however omitted here due to the space limit. This operation to increase (or decrease) the BER when bits are added (or subtracted) will be confirmed by simulation in Sec. IV. Furthermore, the proposed bit rate adjustment method may not be the best solution due to its computational complexity. The main purpose for us to adopt this method is to demonstrate a more meaningful BER comparison by fixing the bit rate. In a realistic system design, if no target bit rate is established as a requirement, such a manipulation is unnecessary.

### B. Packet Average Throughput and Delay Analysis

The packet-level average throughput and delay performance is analyzed in this subsection. By delay, we refer exclusively to queueing delay. To complete delay analysis, a proper queueing model for both OFDM-TDMA and OFDMA is needed, and the M/G/1 model [23] is adopted for this purpose. We simplify Fig. 1(a) by including the premium and the best effort in one composite queue. This simplification does not compromise our comparison. In fact, the M/G/1 results as well as the analysis given here are readily extensible to priority queues [23]. To

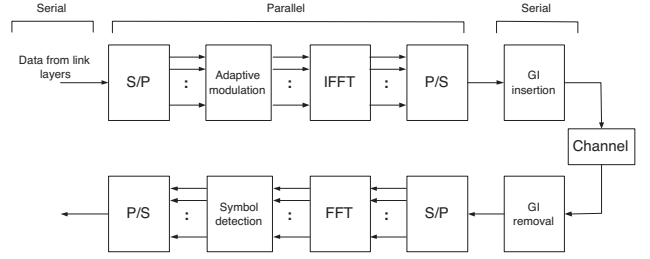


Fig. 2. A typical OFDM transmission system.

fit the M/G/1 model, the Poisson packet arrival process is assumed to replace the regulated streams as the input to the queue.

The average delay  $\bar{W}$  in the M/G/1 model is given by [23]:

$$\bar{W} = \frac{\lambda E[S^2]}{2(1 - \lambda E[S])}, \quad (13)$$

where  $S$  is the packet service time and  $\lambda$  is the Poisson packet arrival rate. Note that in the case of infinite-sized queues as assumed in this work, the throughput is proportional to the  $\lambda$  value. Therefore, we only need to determine  $S$  up to the second moment to complete the throughput-delay relation in (13).

Since a packet is served by subcarrier allocation in OFDM/OFDMA systems, we consider the service time (and thus delay) being measured in the unit of the number of subcarriers. To be more specific,  $S$  is defined to be the number of subcarriers  $N_s$  satisfying

$$S = \left\{ N_s : U \triangleq \sum_{i=1}^{N_s} I_i = \alpha \right\},$$

where  $U$  is the number of bits loaded in  $N_s$  subcarriers,  $I_i$  is the number of bits loaded to subcarrier  $i$ , which is identically (but not necessarily independently) distributed over all  $i$ , and  $\alpha$  is the fixed packet size in bits. The idea can be explained by a typical OFDM transmission system shown in Fig. 2, where data streams can be grouped in a parallel or a serial manner. The “time unit” in serial and parallel groupings are samples and subcarriers, respectively. With the serial-to-parallel (S/P) and the parallel-to-serial (P/S) conversions, the conventional notion of delay in the serial grouping translates to the equivalent delay measured in the number of subcarriers in the parallel grouping, and vice versa. This fits both OFDM-TDMA and OFDMA since we are concerned with the average delay over all users and packets.

We evaluate the first two moments of  $S$  in the following. For the first moment, we have

$$\begin{aligned} E[U|U] &= E \left[ \sum_{i=1}^{N_s} I_i \middle| U \right] \\ &= E \left[ E \left[ \sum_{i=1}^{N_s} I_i \middle| N_s, U \right] \middle| U \right] \\ &= E \left[ N_s \cdot E[I_i|U] \middle| U \right] \\ &= E[I_i|U] \cdot E[N_s|U]. \end{aligned}$$

Conditioning on  $U = \alpha$ , we have  $S = N_s$  by definition. It

follows that

$$E[S] = \frac{\alpha}{E[I_i|U = \alpha]}. \quad (14)$$

The second moment can be derived similarly. That is,

$$\begin{aligned} E[U^2|U] &= E\left[\left(\sum_{i=1}^{N_s} I_i\right)^2 \middle| U\right] \\ &= E\left[E\left[\left(\sum_{i=1}^{N_s} I_i\right)^2 \middle| N_s, U\right] \middle| U\right] \\ &= E\left[N_s \cdot E[I_i^2|U] + \sum_{i \neq j} E[I_i I_j|U] \middle| U\right] \end{aligned} \quad (15)$$

The cross-term in (15) represents the cross-correlation among subcarriers, which depends on the coherence bandwidth of the channel. Without further knowledge or assumption about the channel, we may proceed in deriving a lower bound. By the Schwartz inequality and the fact that  $I_i$  and  $I_j$  are nonnegative, we have

$$E[I_i I_j] \leq \sqrt{E[I_i^2]E[I_j^2]} = E[I_i^2], \quad i \neq j,$$

and consequently,

$$\sum_{i \neq j} E[I_i I_j|U] \leq N_s(N_s - 1)E[I_i^2|U]. \quad (16)$$

By substituting (16) into (15), we obtain

$$E[U^2|U] \leq E[I_i^2|U] \cdot E[N_s^2|U].$$

Again, conditioning on  $U = \alpha$ , we have  $S = N_s$  so that

$$E[S^2] \geq \frac{\alpha^2}{E[I_i^2|U = \alpha]}. \quad (17)$$

To complete the analysis in (14) and (17), we need to obtain the first two moments of  $I_i$ , *i.e.* the number of bits loaded to subcarrier  $i$ . Apparently,  $I_i$  depends on the AM scheme in (9), the channel condition given in (4) or (5) and the selected multiaccess mode. To proceed, we model the channel state in the link layer by a finite-state Markov chain, where each state corresponds to a certain channel condition  $(\mathbb{R}_0, \dots, \mathbb{R}_{r_m})$  and a particular modulation scheme. The state “0” corresponds to  $\mathbb{R}_0$  where no bits are transmitted. The transition probability, which was derived before (*e.g.*, [16]), is not of our concern here. Our interest is to calculate the steady-state probabilities  $\pi_r$ ’s, which can be obtained by integrating pdfs in (4) or (5) over disjoint regions  $\mathbb{R}_r$ ’s. That is,

$$\pi_r = \int_{b_r}^{b_{r+1}} g_{\Gamma}(\gamma) d\gamma \text{ or } \int_{b_r}^{b_{r+1}} g_{\Gamma^*}(\gamma) d\gamma, \quad (18)$$

where  $r = 0, 1, \dots, r_m$ . With these  $\pi_r$ ’s, the first and second moments of  $I_i$  can be obtained by

$$\begin{aligned} E[I_i] &= \sum_{r=0}^{r_m} \pi_r \cdot (2r), \\ E[I_i^2] &= \sum_{r=0}^{r_m} \pi_r \cdot (2r)^2. \end{aligned} \quad (19)$$

Note that the condition on  $U = \alpha$  does not change the moments of  $I_i$ . Finally, by substituting (19) in (14) and (17),

and then (14) and (17) in (13), we obtain the theoretical throughput-delay lower bound curves. This result will be verified by computer simulation in Sec. IV.

### C. Packet Maximum Delay Analysis

We obtain analytical packet delay bounds for all modes listed in Table I in this subsection. A result from network calculus is described below, which will be used in analysis later. The delay for the premium stream in Fig. 1(b) is bounded by [17]

$$d_{pm} \leq \frac{w_{k1} + v_k}{\bar{u}_k}, \quad (20)$$

if queues are stable (*i.e.*,  $r_{k1} + r_{k2} \leq \bar{u}_k$ ) and the first-in-first-out (FIFO) service strategy is employed for each queue. Note that only the delay performance of the premium stream is of our interest. Also, the delay bound in (20) refers to delay experienced by the regulated streams. The waiting time inside the flow control regulator as analyzed in [24] is not of our concern.

Recall that the definition in Sec. II-A is tied with continuous processes. However, the slotted structure of OFDM/OFDMA suggests that the server process be represented by a discrete-time random process  $u_k[n]$ , where  $n$  is the OFDM time slot index. This change turns all integrations into summations in Sec. II-A while leaving flow control parameters unchanged but in different units. That is,  $w_{k1}$ ,  $w_{k2}$  and  $v_k$  are in the unit of bits and rates  $r_{k1}$ ,  $r_{k2}$  and  $\bar{u}_k$  are in the unit of bits per time slot. Then, the delay bound in (20) is in the unit of the number of OFDM time slots. Note that there is no discrepancy between the units used in this subsection and those in Sec. III-B, where the number of subcarriers is used, since the number of subcarriers can be translated to an equivalent number of OFDM symbols, and vice versa. The most remarkable difference is that, instead of averaging over all users and packets in presenting the average delay results in Sec. III-B, we must treat each packet and each user individually here since the measure of delay bounds is not an average.

In the following, we obtain deterministic delay bounds for static modes such as OFDMA I and OFDM I, and then derive probabilistic delay bounds for the remaining modes that use dynamic allocation. Note that a smaller delay bound guarantees better worst-case delay performance.

1) *OFDMA I and OFDM I*: These two modes employ static multiple access and bit allocation yet in a different fashion. Recall from Fig. 1 that the server process  $u_k(t)$  is prescribed by actual multiaccess and bit allocation, thus resulting different server processes of OFDMA I and OFDM I as depicted in Fig. 3 for the continuous-time representation and Fig. 4 for the discrete-time representation. From these figures, we observe that the server rate of OFDMA I is a constant, since subcarriers are divided evenly among users in each OFDM time slot. In contrast, in OFDM I, user  $k$ ’s server rate peaks at time slots when user  $k$  is in service and remains zero during periods in which other users are served.

To obtain the delay bound in (20), our task is to find the pair  $(\bar{u}_k, v_k)$  that defines the server process ( $w_{k1}$  is controlled and known in the design of a flow-control regulator). Let  $N$

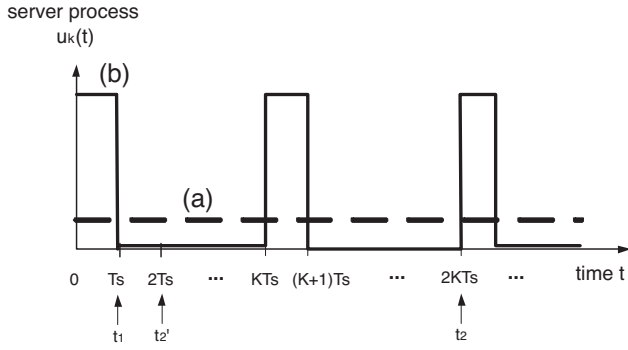


Fig. 3. Illustration of the continuous-time server process for user  $k$  with a total of  $K$  users: (a) OFDMA I and (b) OFDM I.

be the total number of subcarriers and  $r$  the (fixed) number of bits loaded to each subcarrier. Therefore, in each OFDM time slot  $N \times r$  bits are served. OFDMA I has a constant server rate, *i.e.*,

$$u_k[n] = \bar{u}_k = N \times r / K \text{ and } v_k = 0.$$

Consequently, we have

$$d_{pm} \leq \frac{w_{k1}}{\bar{u}_k}. \quad (\text{OFDMA I}) \quad (21)$$

For OFDM I, the average server rate  $\bar{u}_k$  is the same as that in OFDMA I according to (3). To obtain  $v_k$ , we need to examine (2) carefully. First, we observe that a sufficiently large  $v_k$  in the RHS will make (2) always hold since the LHS of (2) is non-negative. However, an arbitrarily large value of  $v_k$  is not useful in deriving the delay bound in (20). Thus, we want to find a smallest  $v_k$  for which (2) is satisfied for *any* choice of  $t_1$  and  $t_2$ . As depicted in Fig. 3, our choice of  $t_1$  and  $t_2$  is the pair that spans the widest among any  $t_1$  and  $t_2$  values that enclose a “peak-rate” time slot of OFDM. Since  $t_2 - t_1 = 2K - 1$  is the largest in this case, the associated  $v_k$  will guarantee that (2) holds also for any other choices of  $t_1$  and  $t_2$ . With such  $t_1$  and  $t_2$  and (2), we have

$$N \times r \geq \bar{u}_k(2K - 1) - v_k.$$

By arranging the terms and the fact that  $N \times r = K\bar{u}_k$ , we have the smallest  $v_k$  as

$$v_k = \bar{u}_k(K - 1). \quad (22)$$

Substituting  $\bar{u}_k$  and  $v_k$  into (20) yields

$$d_{pm} \leq \frac{w_{k1}}{\bar{u}_k} + (K - 1), \quad (\text{OFDM I}) \quad (23)$$

where  $\bar{u}_k = N \times r / K$ . By comparing (21) and (23), we see that OFDMA I has a smaller delay bound than OFDM I by a fixed amount of  $K - 1$ . The physical interpretation of this result is that the  $K - 1$  idle time slots in OFDM I account for the  $K - 1$  extra delay bound.

2) *OFDMA II and OFDMA III*: To obtain delay bounds for OFDMA III and, as a special case, OFDMA II, we make two assumptions in the asymptotic analysis. The number of subcarriers,  $N$ , is large, and the subcarrier channel coefficients are i.i.d. These may be regarded as an ideal approximation to real-world situations. However, as will be demonstrated by simulation in Sec. IV, the theoretically derived delay bounds

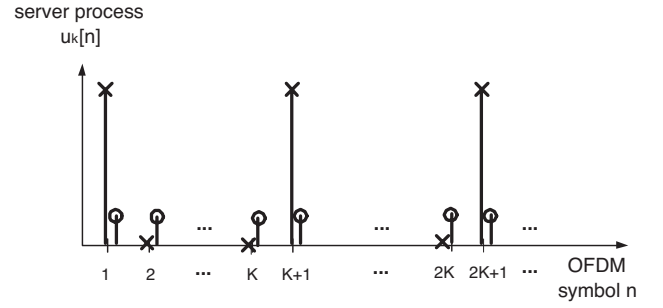


Fig. 4. Illustration of the discrete-time server process for user  $k$  with a total of  $K$  users: (o) OFDMA I and (x) OFDM I.

coincide well with experimental results in both the i.i.d. and the real-world channel setups. Thus, our analysis does provide insights into performance differences among different modes.

OFDMA III, unlike OFDMA I, has a time-varying server process due to the fluctuation of user channels. Specifically, the server process  $u_k[n]$  is a discrete-time random process defined by

$$u_k[n] = \sum_{i \in \mathbb{D}_k[n]} I_i[n], \quad (24)$$

where  $I_i[n]$  is the number of bits loaded onto subcarrier  $i$ , and  $\mathbb{D}_k[n]$  is the set of subcarriers assigned to user  $k$ , both at OFDM time slot  $n$ .  $I_i[n]$ ,  $i \in \mathbb{D}_k[n]$ , are i.i.d. by our aforementioned assumption. Besides, due to opportunistic selection of users, the probability of each user “winning” a particular subcarrier is  $1/K$ , which leads to the binomially-distributed  $|\mathbb{D}_k[n]|$ , denoted by  $|\mathbb{D}_k[n]| \sim B(N, 1/K)$ , where  $|x|$  is the cardinality of set  $x$ . Given  $u_k[n]$  as the summation of  $|\mathbb{D}_k[n]|$  i.i.d. random variables, along with the assumption of big  $N$ , we have (as  $N \rightarrow \infty$  and consequently  $|\mathbb{D}_k[n]| \rightarrow \infty$ )

$$\frac{u_k[n] - E[u_k[n]]}{\sqrt{\text{Var}(u_k[n])}} \rightarrow \mathcal{N}(0, 1), \quad (25)$$

where  $\mathcal{N}(0, 1)$  is the standard Gaussian distribution according to the Central Limit Theorem, and

$$\begin{aligned} E[u_k[n]] &= \frac{N}{K} E[I_i[n]], \\ \text{Var}(u_k[n]) &= \frac{N}{K} \text{Var}(I_i[n]) + \frac{N}{K} \left(1 - \frac{1}{K}\right) (E[I_i[n]])^2, \end{aligned}$$

which can be obtained by applying the standard conditional mean and variance procedures to (24) (conditioned on  $\mathbb{D}_k[n]$ ). Furthermore, by ergodicity (which can be shown by the Law of Large Numbers) we have

$$\bar{u}_k = E[u_k[n]] = \frac{N}{K} E[I_i[n]]. \quad (26)$$

To obtain  $v_k$ , we resort to the definition in (2) and a proper choice of  $t_1$  and  $t'_2$  in Fig. 3. Then, (2) reduces to

$$u_k[n] \geq \bar{u}_k - v_k. \quad (27)$$

Second, we define the delay violation probability  $P_{dv}$  as the probability that the delay bound in (20) is violated. Since  $w_{k1}$  and  $\bar{u}_k$  are known or derived in (26), the violation in (20) may only occur when  $v_k$  is violated, *i.e.*, (27) fails to hold. Then, with  $P_{dv}$  being the outage probability in (27) and by



TABLE II  
THE TDL CHANNEL MODEL PARAMETERS

rms delay spread ( $\tau_{rms}$ )	1 $\mu$ s
tap spacing ( $T$ )	175 ns
number of taps ( $L$ )	20
max delay spread ( $\tau_{max}$ )	3.325 $\mu$ s (= 175 ns $\times$ (20-1))

plugging the asymptotic distribution of  $u_k[n]$  into (27), it is straightforward to show

$$Q\left(\frac{-v_k(P_{dv})}{\sqrt{Var(u_k[n])}}\right) = 1 - P_{dv},$$

or, equivalently,

$$v_k(P_{dv}) = -Q^{-1}(1 - P_{dv}) \cdot \sqrt{Var(u_k[n])}. \quad (28)$$

Finally, by substituting (26) and (28) into (20), we obtain

$$d_{pm} \leq \frac{w_{k1} + v_k(P_{dv})}{\bar{u}_k}. \quad (\text{OFDMA III}) \quad (29)$$

OFDMA II, which is a special case with fixed  $|\mathbb{D}_k[n]| = N/K$ , has the same  $\bar{u}_k$  in (26) but a different  $Var(u_k[n]) = \frac{N}{K} Var(I_i[n])$ . Substituting  $Var(u_k[n])$  into (28) and then (29) yields the delay bound for OFDMA II.

3) *OFDM II and OFDM III*: We first obtain results for OFDM III and then OFDM II as a special case. The server process in OFDM III can be written as

$$u_k[n] = \begin{cases} u_{k,b}[n] = \sum_{i=1}^N I_i[n], & n \in \mathbb{C}(k), \\ 0, & \text{otherwise,} \end{cases} \quad (30)$$

where  $\mathbb{C}(k)$  is the set of time slots assigned to user  $k$ . The average server rate is easily obtained by the definition in (3) as

$$\bar{u}_k = E[u_k[n]] = \frac{N}{K} E[I_i[n]]. \quad (31)$$

To find  $v_k$ , we follow the analysis for OFDM I with one additional assumption; namely, there exists  $C_{idle}$  such that any idle period (when  $u_k[n] = 0$  for user  $k$ ) is, up to a negligible violation probability, of length no greater than  $C_{idle} \times K - 1$  time slots. The value of  $C_{idle}$  is chosen empirically to ensure the statistics associated with the opportunistic assignment is unchanged.

Similarly to the analysis for OFDM I,  $v_k$  is derived with judicious choice of  $t_1$  and  $t_2$  applied to (2), *i.e.*,

$$u_{k,b}[n] \geq \bar{u}_k[2(C_{idle} \cdot K - 1) + 1] - v_k.$$

Then, due to the asymptotic distribution of  $u_{k,b}[n]$ ,

$$\frac{u_{k,b}[n] - N \cdot E[I_i[n]]}{\sqrt{N \cdot Var(I_i[n])}} \rightarrow \mathcal{N}(0, 1), \quad (32)$$

we have

$$v_k(P_{dv}) = -Q^{-1}(1 - P_{dv}) \cdot \sqrt{N \cdot Var(I_i[n])} + (2C_{idle} - 1 - \frac{1}{K})N \cdot E[I_i[n]]. \quad (33)$$

Substituting (31) and (33) into (20) gives the delay bound

$$d_{pm} \leq \frac{w_{k1} + v_k(P_{dv})}{\bar{u}_k}. \quad (\text{OFDM III}) \quad (34)$$

All the above derivation also applies to OFDM II by setting  $C_{idle} = 1$  due to the static round-robin time-slot assignment in OFDM II.

TABLE III  
PARAMETERS OF THE OFDM SYSTEM

OFDM symbol time ( $T_s$ )	100.8 $\mu$ s
useful symbol time ( $T_b$ )	89.6 $\mu$ s
guard time ( $T_g$ )	11.2 $\mu$ s
FFT size ( $N_{FFT}$ )	512
sample time ( $T$ )	175 ns (= 89.6 $\mu$ s/512)

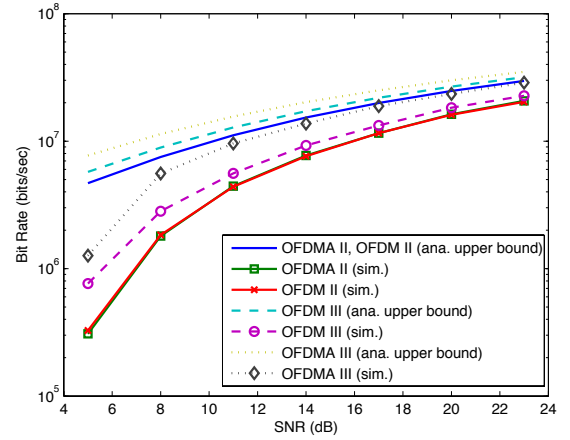


Fig. 5. Comparison of maximum supportable bit rates with  $K = 4$ .

#### IV. SIMULATION RESULTS

In this section, we perform computer simulation to verify the analysis given in Sec. III and make an extensive comparison of the link and the physical layer performance of OFDM-TDMA and OFDMA. The Rayleigh fading channel is adopted and generated by a tapped delay line (TDL) channel model with equally-spaced taps and an exponential power delay profile. The parameters for the TDL channel model are given in Table II. We consider homogeneous users whose channel coefficients are i.i.d. A practical framed structure is considered where, unless otherwise noted, each frame contains ten OFDM symbols. A fast fading channel based on the method described in [3] is implemented. That is, the channel varies, presumably independently, across OFDM frames but remains the same within a frame. We restrict the squared  $M$ -QAM modulations to 4-, 16-, 64- and 256-QAM (*i.e.*,  $r_m = 4$ ). Besides, although OFDMA typically has a larger number of subcarriers than OFDM in real-world applications, we use an identical set of parameters for both systems for fair comparison. The parameters of the OFDM system are borrowed from [25] and summarized in Table III.

We first examine the total bit rate that can be supported by each mode. Both analytical upper bounds from (11) and (12) and empirical results are shown in Fig. 5. Note that target BER  $P_b = 10^{-3}$  is set for bit allocation. We observe that, among these four modes, OFDMA III achieves the best performance due to multiuser diversity and frequency diversity that arise from dynamic allocation. Likewise, OFDM III outperforms OFDMA II and OFDM II due to dynamic time-slot allocation. Both OFDMA II and OFDM II perform static multiple access which leads to the same share of resource of each user on the average. Consequently, we observe comparable curves in Figs. 5 and 6 for OFDMA II and OFDM II, where the difference is contributed by randomness in bit allocation.

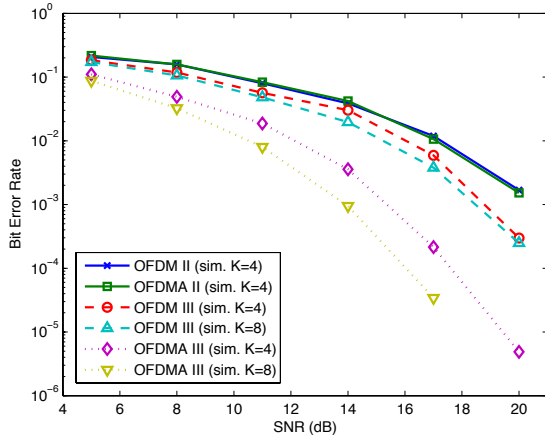


Fig. 6. Comparison of BER results with  $K = 4$  and 8.

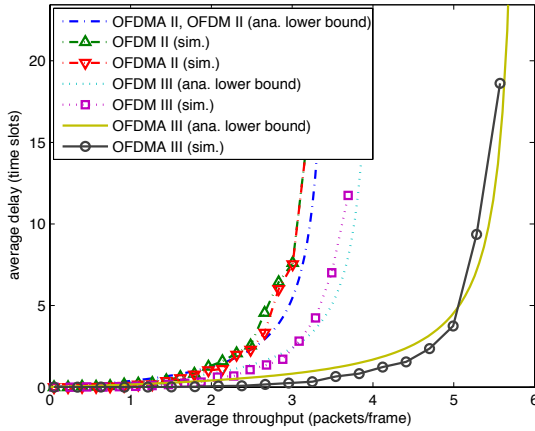


Fig. 7. Comparison of the packet throughput and delay results with SNR = 16 dB and  $K = 4$ .

Fig. 6 shows the uncoded BER performance when the bit rate is fixed at 4 bits/subcarrier (or roughly 20 Mbits/sec). For  $K = 4$ , we see that OFDMA III outperforms OFDM III by 3.5 dB, and OFDM III outperforms OFDMA II and OFDM II by 2 dB when  $\text{BER} = 10^{-3}$ . This result is slightly different from that of [4] due to different settings. We also observe an additional multiuser diversity gain contributed by more users with  $K = 8$ . Note that the bit adding and subtracting operations described in Sec. III-A renders BER above  $P_b$  at the low SNR region since extra bits are added to make up 4 bits/subcarrier.

We choose a fixed packet length of  $\alpha = 3000$  bits and SNR = 16 dB in presenting the packet throughput-delay performance curves shown in Fig. 7. We see that the performance differences are translated from the physical-layer transmission schemes and they are consistent with the results in Figs. 5 and 6. In particular, given a fixed delay, OFDMA III achieves the largest throughput and, consequently, the best packet-level throughput-delay performance. As the throughput value approaches capacity, all schemes suffer from a significant amount of delay.

Note that although there is no notion of frames in the analytical M/G/1 model, a frame should be considered in the finite granularity implementation. In the simulation, we

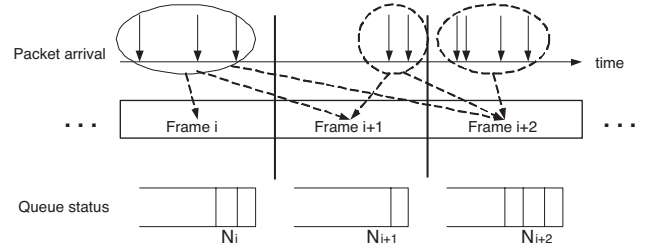


Fig. 8. Illustration of the exhaustive service system and queueing in the link layer simulation in Fig. 7.

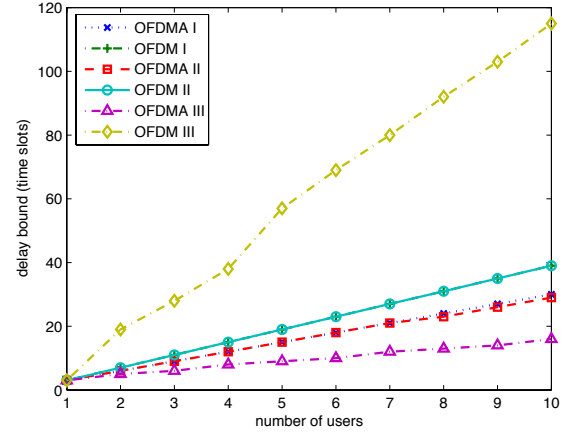


Fig. 9. Analytical delay bounds versus the number of users,  $K$ , with SNR = 16 dB.

consider a practical exhaustive service system where only packets that arrive before or during the current frame can be served in the current frame as shown in Fig. 8. Packets are multiplexed and served in the frame sequentially in both time and frequency. Besides, the number of packets in a queue is counted frame by frame and averaged to give a finite granularity approximation of  $\bar{N}$ , the time-average number of packets in the queue. With  $\bar{N}$  and throughput  $\lambda$  controlled to get different points in the plot, we calculate the average delay by Little's Law [23]:  $\bar{W} = \bar{N}/\lambda$ .

Due to the exhaustive service assumption and finite granularity approximation,  $\bar{N}$  is nearly zero at a low throughput value, resulting  $\bar{W} \approx 0$ . This accounts for the fact that the simulation curve does not obey the analytical lower bound curve when the throughput is low. It is worthwhile to point out that the throughput-delay result here is an average over all users and packets. The result of the individual delay incurred to each user packet is presented next.

We implement the flow control regulator by the "leaky bucket" scheme in [24]. We choose a token pool of size of 3000 tokens. Each token can serve 1 bit of data, from which the burst sizes  $w_{k1}$  and  $w_{k2}$  can be calculated to be 3001 bits. The token arrival rates, which determine the average regulated rates  $r_{k1}$  and  $r_{k2}$ , are chosen such that queues are stable. Moreover,  $C_{idle}$  is found empirically to ensure the violation probability less than  $10^{-2}$ ; for example,  $C_{idle} = 4$  (or, respectively, 5) for  $K = 4$  (or, respectively, 8). We fix user channel SNR = 16 dB and  $P_{dv} = 10^{-4}$  to obtain the analytical delay bounds in Sec. III-C, which, after rounded up to the nearest integer number of time slots,



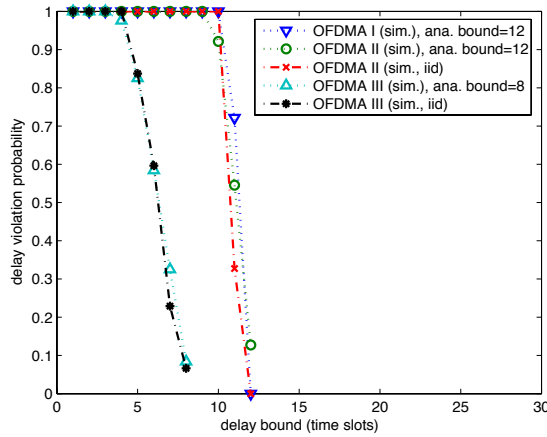


Fig. 10. The delay violation probability vs. the delay bound for OFDMA modes with SNR = 16 dB and  $K = 4$ .

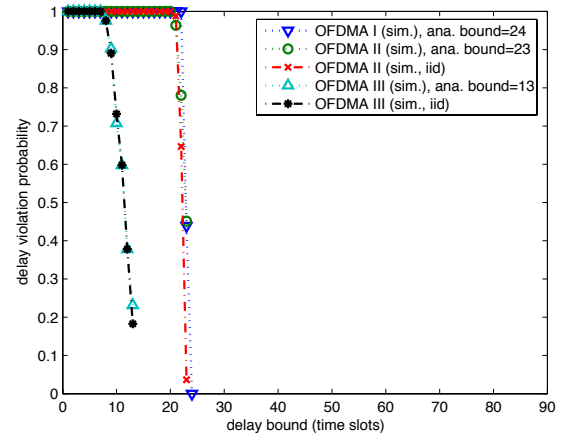


Fig. 12. The delay violation probability vs. the delay bound for OFDMA modes with SNR = 16 dB and  $K = 8$ .

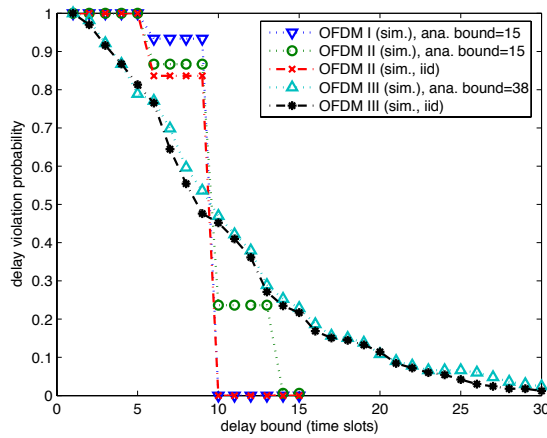


Fig. 11. The delay violation probability vs. the delay bound for OFDM modes with SNR = 16 dB and  $K = 4$ .

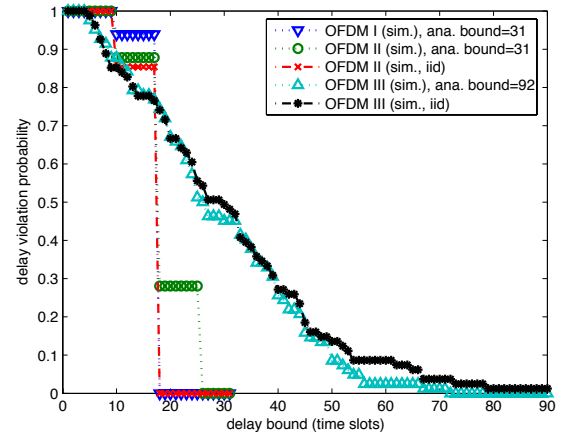


Fig. 13. The delay violation probability vs. the delay bound for OFDM modes with SNR = 16 dB and  $K = 8$ .

are drawn in Fig. 9 for the premium stream. We see from Fig. 9 that OFDMA modes generally have better worst-case delay performance than OFDM modes. The idle periods introduced by round-robin or opportunistic time-slot assignments in OFDM account for the performance gap. The opportunistic assignment, in particular, creates the possibility of a long idle period which explains for the largest delay bounds of OFDM III. Note that, when  $K = 1$ , all modes degenerate to the same scheme. As  $K$  increases, the delay bound increases because the same resources are shared among an increasing number of competitors.

In Figs. 10 and 11, we draw the Monte Carlo simulation curves when SNR = 16 dB and  $K = 4$ . We see that the analytical delay bounds, obtained from Fig. 9 by fixing  $K = 4$  and shown in the legend of both plots, agree well with the actual packet maximum delay performance. Figs. 10 and 11 provide knowledge of the packet delay distribution by presenting the percentage of packets (y-axis) that experience delay higher than a particular value (x-axis). It is observed that OFDM schemes generally have higher delay variation than OFDMA. For example, OFDM III in Fig. 11 has a large dynamic range in the packet delay while OFDMA III has a much smaller range as shown by the sharper slope of curves in Fig. 10. This observation suggests that OFDMA

is more suitable for supporting real-time traffic, since real-time traffic is vulnerable to large delay. Figs. 10 and 11 also show simulation results under the setting of i.i.d. subcarrier channels. The results of the i.i.d. case are very close to its non-i.i.d. counterparts. This may be explained as follows. Although these two settings produce different bit allocation results on individual subcarriers, the difference is mitigated by a number of summations in the process of calculating the overall packet delay. To conclude, as far as the link layer delay analysis is concerned, the i.i.d. assumption facilitates the derivation of analytical delay bounds that work well for both i.i.d. and real-world setups as confirmed by Figs. 10 and 11.

Simulation results for SNR = 16 dB and  $K = 8$  are shown in Figs. 12 and 13 to demonstrate the effect of user number  $K$ . As compared with the case of  $K = 4$  in Figs. 10 and 11, an increased individual delay is observed due to a larger number of users. Overall, we observe consistent results for  $K = 4$  and  $K = 8$ , and the discussion made for  $K = 4$  holds for  $K = 8$  in Figs. 12 and 13 as well.

## V. CONCLUSION

Performance analysis and comparison of OFDM-TDMA and OFDMA centered on scheduling with cross-layer consideration were conducted. Several OFDM/OFDMA modes

with different multiaccess and resource allocation schemes were considered along with an analytical framework based on the QoS architecture of IEEE 802.16. The analysis and simulation offered a thorough understanding of the system's capability of supporting multimedia delivery from a cross-layer viewpoint involving both link and physical layers. The analytical and empirical results suggest that dynamic OFDMA has a stronger potential to support multimedia transmission than dynamic OFDM-TDMA. It is also observed that the opportunistic assignment can be employed more effectively in OFDMA.

## REFERENCES

- [1] C. Eklund, R. B. Marks, K. L. Stanwood, and S. Wang, "IEEE standard 802.16: A technical overview of the WirelessMAN air interface for broadband wireless access," *IEEE Commun. Mag.*, vol. 40, pp. 98–107, June 2002.
- [2] K. Wongthavarawat and A. Ganz, "IEEE 802.16 based last mile broadband wireless military networks with quality of service support," in *Proc. IEEE MILCOM'03*, Oct. 2003, pp. 779–784.
- [3] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1277–1294, June 2002.
- [4] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1747–1758, Oct. 1999.
- [5] H. Rohling and R. Grunheid, "Performance comparison of different multiple access schemes for the downlink of an OFDM communication system," in *Proc. IEEE VTC'97*, May 1997, pp. 1365–1369.
- [6] S. Kaiser, "MC-FDMA and MC-TDMA versus MC-CDMA and SS-MC-MA: performance evaluation for fading channels," in *Proc. IEEE 5th International Symposium on Spread Spectrum Techniques and Applications*, Sept. 1998, pp. 200–204.
- [7] R. Iyengar, P. Iyer, and B. Sikdar, "Delay analysis of 802.16 based last mile wireless networks," in *Proc. IEEE GLOBECOM'05*, Nov. 2005, pp. 3123–3127.
- [8] D. Niyato and E. Hossain, "Queue-aware uplink bandwidth allocation for polling services in 802.16 broadband wireless networks," in *Proc. IEEE GLOBECOM'05*, Nov. 2005, pp. 3702–3706.
- [9] —, "Queueing analysis of OFDM/TDMA systems," in *Proc. IEEE GLOBECOM'05*, Nov. 2005, pp. 3712–3716.
- [10] G. Song, Y. Li, J. Leonard J. Cimini, and H. Zheng, "Joint channel-aware and queue-aware data scheduling in multiple shared wireless channels," in *Proc. IEEE WCNC'04*, Mar. 2004, pp. 1939–1944.
- [11] Y.-J. Chang, F.-T. Chien, and C.-C. J. Kuo, "Performance comparison of OFDM-TDMA and OFDMA with cross-layer consideration," in *Proc. IEEE VTC'06*, Sept. 2006.
- [12] —, "Delay analysis and comparison of OFDM-TDMA and OFDMA under IEEE 802.16 QoS framework," in *Proc. IEEE GLOBECOM'06*, Nov. 2006.
- [13] C. Hoymann, "Analysis and performance evaluation of the OFDM-based metropolitan area network IEEE 802.16," *Computer Networks*, vol. 49, pp. 341–363, June 2005.
- [14] Y. M. Ki, E. S. Kim, and D. K. Kim, "Downlink scheduling and resource management for best effort service in TDD-OFDMA cellular networks," *Lecture Notes in Computer Science*, pp. 316–329, 2004.
- [15] 802.16 Working Group. [Online]. Available: <http://www.ieee802.org/16>
- [16] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1688–1692, Nov. 1999.
- [17] R. L. Cruz, "A calculus for network delay, part I: Network elements in isolation," *IEEE Trans. Inform. Theory*, vol. 37, pp. 114–131, Jan. 1991.
- [18] L. L. Peterson and B. S. Davie, *Computer Networks*, 3rd ed. San Francisco: Morgan Kaufmann, 2003.
- [19] J.-Y. Le Boudec, "Application of network calculus to guaranteed service networks," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1087–1096, May 1998.
- [20] G. Song and Y. G. Li, "Cross-layer optimization for OFDM wireless networks—Part I: Theoretical framework," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 614–624, Mar. 2005.
- [21] M. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Wireless Personal Communications*, vol. 13, pp. 119–143, May 2000.
- [22] A. Czylik, "Adaptive OFDM for wideband radio channels," in *Proc. IEEE GLOBECOM'96*, Nov. 1996, pp. 713–718.
- [23] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1991.
- [24] M. Sidi, W.-Z. Liu, I. Cidon, and I. Gopal, "Congestion control through input rate regulation," *IEEE Trans. Commun.*, vol. 41, pp. 471–477, Mar. 1993.
- [25] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 WirelessMAN," *Intel Technology Journal*, vol. 8, pp. 201–212, Aug. 2004.



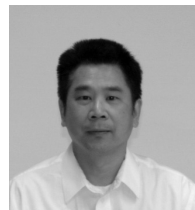
**Yu-Jung Chang** received the B.S. degree from the National Tsing Hua University, Taiwan, in 2000, the M.S. degree from the National Chiao Tung University, Taiwan, in 2002, and is currently pursuing his Ph.D. degree at the University of Southern California (USC), Los Angeles, all in Electrical Engineering.

Since January 2005, he has been a Research Assistant in EE department at USC. His research interests include cross-layer analysis and design of OFDM and OFDMA networks; resource allocation, QoS provisioning, and performance evaluation for these networks.



**Feng-Tsun Chien** (S'02-M'05) received the B.S. degree from the National Tsing Hua University, Hsinchu, in 1995, the M.S. degree from the National Taiwan University, Taipei, in 1997, and the Ph.D. degree from the University of Southern California, Los Angeles, in 2004, respectively, all in Electrical Engineering.

He joined the Department of Electronics Engineering of the National Chiao Tung University, Hsinchu, in July 2005, as an Assistant Professor. His current research interests include signal processing aspects on communications, cross-layer considerations for OFDM and OFDMA systems, multicarrier CDMA, and MIMO-OFDM systems.



**C.-C. Jay Kuo** (S'83-M'86-SM'92-F'99) received the B.S. degree from the National Taiwan University, Taipei, in 1980 and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1985 and 1987, respectively, all in Electrical Engineering. Dr. Kuo was Computational and Applied Mathematics (CAM) Research Assistant Professor in the Department of Mathematics at the University of California, Los Angeles, from October 1987 to December 1988. Since January 1989, he has been with the University of Southern California,

where he is currently Professor of Electrical Engineering, Computer Science and Mathematics and Director of the Signal and Image Processing Institute. His research interests are in the areas of digital signal and image processing, multimedia compression, communication and networking technologies. Dr. Kuo has guided about 80 students to their Ph.D. degrees and supervised 15 postdoctoral research fellows. He is a co-author of about 120 journal papers, 700 conference papers and 7 books.

Dr. Kuo is a Fellow of IEEE and SPIE and a member of ACM. He is Editor-in-Chief for the *Journal of Visual Communication and Image Representation*, and Editor for the *Journal of Information Science and Engineering* and the *EURASIP Journal of Applied Signal Processing*. He was on the Editorial Board of the *IEEE Signal Processing Magazine* in 2003–2004. He served as Associate Editor for *IEEE Transactions on Image Processing* in 1995–98, *IEEE Transactions on Circuits and Systems for Video Technology* in 1995–1997 and *IEEE Transactions on Speech and Audio Processing* in 2001–2003. He received the National Science Foundation Young Investigator Award (NYI) and Presidential Faculty Fellow (PFF) Award in 1992 and 1993, respectively.