

Delay Analysis and Comparison of OFDM-TDMA and OFDMA under IEEE 802.16 QoS Framework

Yu-Jung Chang[†], Feng-Tsun Chien[‡] and C.-C. Jay Kuo[†]

[†] Department of Electrical Engineering and Integrated Media Systems Center
University of Southern California, Los Angeles, CA 90089-2564, USA

[‡] Department of Electronics Engineering

National Chiao Tung University, Hsinchu, Taiwan, ROC

E-mail: yujungc@usc.edu and ftchien@mail.nctu.edu.tw and cckuo@sipi.usc.edu

Abstract—The delay analysis and comparison of OFDM-TDMA and OFDMA using a flow control scheme under the QoS framework of IEEE 802.16 are conducted in this work. We investigate the maximum packet delay performance by deriving delay bounds for several multiaccess modes with different subcarrier/time-slot assignment and bit allocation schemes. The analytical delay bounds provide insights into the performance difference among different multiaccess schemes, and demonstrate each scheme's capability of supporting real-time, delay-sensitive multimedia applications. The derived analytical results are verified by computer simulation.

I. INTRODUCTION

Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) are two well-known techniques for resource management in multiaccess communication systems. When combined with Orthogonal Frequency Division Multiplexing (OFDM), they are called OFDM-TDMA and OFDMA (OFDM Access), respectively. Both of them have been adopted by the IEEE 802.16 standard as two options for transmissions at the 2–11 GHz band [1]. In order to carry multimedia applications, which demand a wide range of quality-of-service (QoS) requirements, the embedment of QoS provisioning in the multiaccess systems is needed. In fact, a QoS framework in the medium access control (MAC) layer has been integrated with the multiaccess transmission systems in the IEEE 802.16 standard [2]. In this work, we consider a similar QoS framework under which we compare OFDM-TDMA and OFDMA with various scheduling methods.

A cross-layer approach is needed to examine the QoS-featured multiaccess system thoroughly, since the QoS requirements in different layers have to be treated differently. For example, there is an increasing interest in combining link-layer queueing with physical-layer transmission to analyze the delay behavior of multiaccess systems. To be more specific, the analysis of queueing delay performance for 802.16 networks was conducted in [3], [4] by considering some scheduling policies. A vacation queueing model was adopted in [5] to analyze the queueing performance of OFDM-TDMA systems.

Although previous work in [3]–[5] examined the TDMA scenario by adopting the average queueing delay as the performance measure, there has been little work on other relevant performance measures such as the delay bound and the delay violation probability, which are indicative of the worst-case

delay behavior. These performance metrics are critical to the capability of a system in support of real-time traffic (*e.g.*, voice, video). Besides, to the best of our knowledge, there is little work on the comparison of different multiaccess systems with various scheduling schemes. We aim to address these issues in this work by scrutinizing the packet delay bound performance of OFDM-TDMA and OFDMA.

To complete a meaningful packet-level analysis, we model the Rayleigh fading channel with a finite-state Markov chain [6]. It is assumed that ideal channel state information (CSI) is available at the base station. In addition, we adopt and implement a well-known flow control scheme [7] to regulate the input streams. That is, for a fast incoming stream to satisfy the imposed output rate and burstiness constraints, some packets will be dropped or buffered for later transmission by the flow control regulator. Either way would be a reasonable treatment on packets to trade for guaranteed delay, since real-time multimedia data transmission is more sensitive to delay but less to loss. With such a flow control scheme and the first-in first-out (FIFO) service strategy in place, we can derive the delay bounds deterministically or stochastically. The delay bounds provide a valuable measure in differentiating OFDM-TDMA and OFDMA in their maximum delay performance.

The rest of the paper is organized as follows. Some background information, including the QoS-aware system model, the flow control regulation scheme and the delay bound, is discussed in Sec. II. Various multiple access and resource allocation schemes are presented in Sec. III. The delay analysis for OFDM-TDMA and OFDMA is performed in Sec. IV. Simulation results are shown in Sec. V. Finally, concluding remarks are given in Sec. VI.

II. RESEARCH BACKGROUND

A. Overview of QoS-Aware System

As shown in Fig. 1(a), our comparison framework is similar to the QoS architecture of 802.16 [2]. The QoS provision is realized by first dividing application packets into two classes, *i.e.*, the premium class and the best-effort class, and then offering better QoS (*e.g.*, guaranteed delay) for the premium class. The classification can be done by examining a particular field in the packet header (*e.g.*, IPv4 TOS, type of service

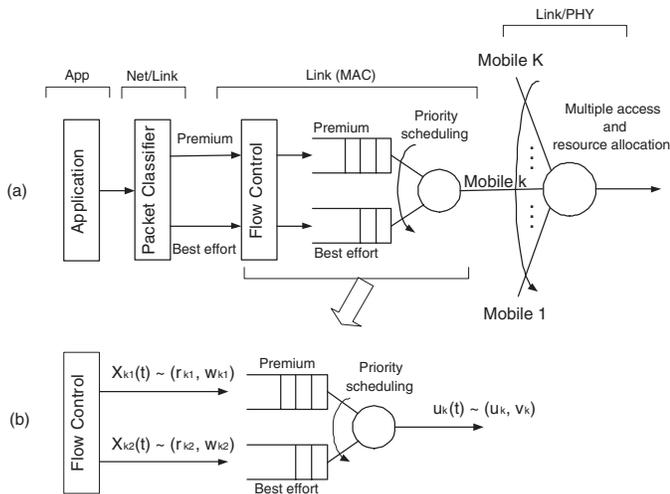


Fig. 1. (a) A cross-layer QoS-support system model. (b) A queuing system with flow-control regulated streams and preemptive priority servicing.

field). Since multimedia applications are generally delay-sensitive, bandwidth-intense and loss-tolerant, the two classes are differentiated primarily by the queuing delay [8] in our framework. In other words, the premium class contains delay-sensitive packets.

After packets are classified, a flow control scheme [7] is adopted to regulate the packet flow. For a system with guaranteed average delay, there is still no guarantee on the maximum packet delay if there is no flow control mechanism. This may degrade the real-time application performance dramatically. The flow control scheme also provides a useful tool to analyze the delay bound performance, which will be discussed in Secs. II-B and II-C. Also, as shown in Fig. 1(a), packets marked as the premium or the best effort are queued separately, and treated according to the preemptive priority scheduling. That is, the premium queue, when not empty, is always processed before the best effort queue. Furthermore, different mobile users are scheduled for service by a centralized scheduler (or particularly a multiaccess scheme), which has knowledge of users' uplink channel conditions and may perform scheduling dynamically.

B. Flow Control Regulation

We adopt the burstiness control method proposed in [7] as our flow control regulator. It is used to regulate the source stream so that the output stream is kept to meet an average rate while allowing a certain degree of burstiness. The method is described below. We use $X_{k1}(t)$ and $X_{k2}(t)$ to denote the flow-control regulated premium and best effort streams of user k , respectively, as depicted in Fig. 1(b). 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. Likewise, we have $X_{k2}(t) \sim$

(r_{k2}, w_{k2}) by following the same definition. Suppose the time varying server process, $u_k(t)$ in Fig. 1(b), satisfies similar constraints. 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

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

C. Delay Bound

With preemptive priority scheduling, the delay bound for the premium stream $X_{k1}(t)$ can be derived explicitly. In particular, if $r_{k1} + r_{k2} \leq \bar{u}_k$ (i.e., queues are stable) and the FIFO service strategy is employed in each queue, the delay for the premium stream is bounded by [7]

$$d_{pm} \leq \frac{w_{k1} + v_k}{\bar{u}_k}. \quad (4)$$

Note that only the delay performance of the premium stream is of interest. Thus, we do not analyze the delay for the best-effort stream. In addition, the delay bound in (4) refers to the delay experienced by regulated streams, notably incurred in the queues in Fig. 1(b). The waiting time inside the flow control regulator, as was analyzed in [9] before, does not take effect in the analysis and comparison of OFDM-TDMA and OFDMA. Thus, it is not of concern in this work.

Note that the server process depends on the actual multiaccess scheme used (OFDM-TDMA or OFDMA). Due to the slotted structure of OFDM, a more meaningful and useful presentation of the server process is a discrete-time random process $u_k[n]$, where n is the OFDM symbol index. This turns all the integration operations into summations, but leaves all the flow control parameters unchanged yet in different units. In particular, w_{k1} , w_{k2} and v_k are in bits; rates r_{k1} , r_{k2} and \bar{u}_k are in bits/OFDM symbol. Then, the delay bound in (4) refers to the number of OFDM symbols. We will perform delay analysis in Sec. IV based on models described in this section.

III. MULTIPLE ACCESS AND RESOURCE ALLOCATION

The multiple access mechanism shown in Fig. 1(a) is accomplished by OFDMA or OFDM-TDMA, along with a particular resource allocation method where subcarrier, time slot, and bit allocation are performed. Several different schemes, summarized in Table I, will be described in this section based on their subcarrier/time slot assignment and bit allocation schemes.

A. Subcarrier and Time-slot Assignment

OFDMA and OFDM-TDMA perform multiple access in a frequency-sharing and a time-sharing manner, respectively. In other words, OFDMA performs subcarrier assignment while OFDM-TDMA performs time-slot assignment. For both assignment tasks, two strategies are considered in this work:

TABLE I
MULTIACCESS OFDM MODES CONSIDERED IN OUR WORK.

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

static or dynamic. Dynamic (or opportunistic) assignment takes into account users' channel conditions in making decision whereas static assignment does not. For OFDMA, static subcarrier assignment divides subcarriers evenly among users and allocates a fixed interleaved set of subcarriers to users. In contrast, dynamic OFDMA assigns a subcarrier to the user with the best signal-to-noise ratio (SNR) seen at that particular subcarrier.

For OFDM-TDMA, static time-slot assignment allocates users in a round-robin fashion. In contrast, dynamic OFDM-TDMA assigns an OFDM symbol to the user with the best channel condition at that particular time slot. In OFDM-TDMA the chosen user is allocated all subcarriers exclusively.

The subcarrier SNR distribution for each multiaccess mode can be obtained as follows. For a Rayleigh fading channel, the received SNR is an exponential random variable [6], denoted by Γ , with the following probability density function (pdf)

$$g(\gamma) = \frac{1}{\gamma_0} \exp\left(-\frac{\gamma}{\gamma_0}\right), \quad \gamma \geq 0, \quad (5)$$

where γ_0 is the average SNR. This applies to all static multiaccess modes (*i.e.*, OFDMA I–II, OFDM I–II). For dynamic OFDM III, (5) applies with a higher mean, as will be verified by simulation. For dynamic OFDMA III, since a subcarrier is assigned to the user with the highest SNR, the subcarrier SNR, denoted by Γ^* , is equal to the maximum of K i.i.d. exponential random variables, where K is the number of users. The pdf of Γ^* can be derived by a standard algebraic procedure, *i.e.*,

$$g^*(\gamma) = \frac{K}{\gamma_0} \exp\left(-\frac{\gamma}{\gamma_0}\right) \left(1 - \exp\left(-\frac{\gamma}{\gamma_0}\right)\right)^{K-1}, \quad \gamma \geq 0. \quad (6)$$

B. Bit Allocation

Once resource is apportioned among users, the bit allocation scheme takes over to choose the proper type and order of modulation. We consider squared M-QAM modulations (4-, 16-, 64- and 256-QAM), and adopt both fixed and adaptive modulation (AM) methods [10]. In contrast with fixed bit allocation, where a fixed M-QAM modulation is employed throughout all subcarriers, the AM scheme first divides the whole spectrum of received SNR value into several disjoint regions. Then, if the actual subcarrier SNR falls in a particular region, a corresponding modulation method is used, as illustrated in Fig. 2. The boundary values b_0, b_1, \dots can be determined by the method presented in [10].

With pdfs in (5) and (6), we can determine for the AM scheme the statistics of I_i , the number of bits loaded onto

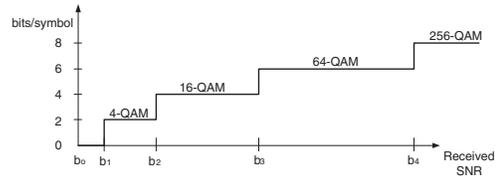


Fig. 2. The choice of (discrete-rate) adaptive modulation schemes as a function of the SNR.

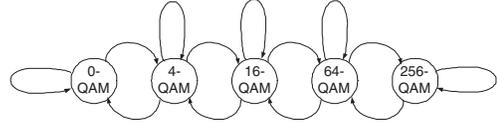


Fig. 3. A finite-state Markov chain model for the fading channel.

subcarrier i . First, we model the channel state by a finite-state Markov chain (FSMC), where each state corresponds to a certain modulation scheme as shown in Fig. 3. The steady-state probabilities for each state can be obtained by integrating pdf in (5) and (6) over each region illustrated in Fig. 2. Then, with these calculated steady-state probabilities, the first and second moments of I_i can be easily obtained, which will be used in Secs. IV-B and IV-C to derive the delay bounds.

IV. DELAY ANALYSIS

In this section, we obtain analytical packet delay bounds by evaluating \bar{u}_k and v_k in (4) for all modes listed in Table I. We first obtain the deterministic delay bounds for modes OFDMA I and OFDM I. Then, based on the approach used and the idea delineated, we further derive the probabilistic delay bounds for the rest of the modes. Note that a smaller delay bound guarantees a better worst-case delay performance.

A. OFDMA I and OFDM I

These two modes employ static multiple access, in its own fashion, and static bit allocation. A different multiaccess fashion leads to a different server process in OFDMA I and OFDM I. Fig. 4 shows the continuous-time representation of the server process while Fig. 5 shows the discrete-time representation as is the case in our simulation. The server rate for each user in OFDMA I is a constant, since subcarriers are divided evenly among users in each and every OFDM symbol. In contrast, the server rate for user k in OFDM I peaks at symbols when user k is in service while remains zero during periods other users are served.

To find the delay bound in (4), we need to derive (\bar{u}_k, v_k) . Note that, for each OFDM symbol, $N \times r$ bits can be served, where N is the total number of subcarriers with r bits loaded to each subcarrier. For OFDMA I, due to the constant server rate, we have $u_k[n] = \bar{u}_k = N \times r/K$ and $v_k = 0$, which leads to

$$d_{pm} \leq \frac{w_k 1}{\bar{u}_k}. \quad (\text{OFDMA I}) \quad (7)$$

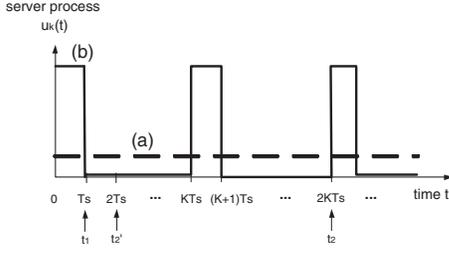


Fig. 4. Illustration of the continuous-time server process for user k : (a) OFDMA I and (b) OFDM I.

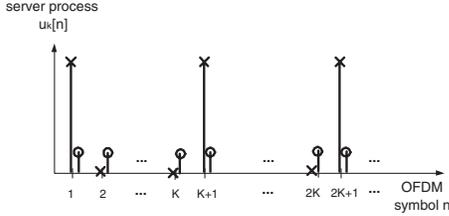


Fig. 5. Illustration of the discrete-time server process for user k with OFDMA I (o) and OFDM I (x).

For OFDM I, by the definition in (3), the average server rate is also $\bar{u}_k = N \times r/K$. Besides, the smallest v_k for which (2) is satisfied for any time instant can be found by picking proper t_1 and t_2 as depicted in Fig. 4. This in turn gives

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

from (2). By arranging the terms and the fact that $K\bar{u}_k = N \times r$, we get the smallest v_k as

$$v_k = \bar{u}_k(K - 1). \quad (8)$$

Substituting \bar{u}_k and v_k into (4) yields

$$d_{pm} \leq \frac{w_{k1}}{\bar{u}_k} + (K - 1). \quad (\text{OFDM I}) \quad (9)$$

By comparing (7) and (9), we see that OFDMA I has a smaller delay bound than OFDM I by a fixed amount of $K - 1$.

B. OFDMA II and OFDMA III

In this subsection, we first obtain the results for OFDMA III. Then, as a special case, the results for OFDMA II are readily available. We make two assumptions in deriving asymptotic analytic results. That is, we assume that the number of subcarriers, N , is big, and the subcarrier channel coefficients are i.i.d. This can be regarded as an extreme case to the real-world conditions. However, by simulation results of both i.i.d. subcarrier channel and real-world channel setups in Sec. V, we find that the analytic delay bound results coincide with the experimental results in both cases well. Thus, our theoretical analysis provides insights into performance differences among different system modes.

In OFDMA III, unlike OFDMA I, the server rate is no longer a constant. Instead, it fluctuates with user k 's instantaneous channel. Specifically, the server process $u_k[n]$ is a

discrete-time random process defined by

$$u_k[n] = \sum_{i \in \mathbb{D}_k[n]} I_i[n], \quad (10)$$

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 symbol n . We assume that subcarrier channel conditions are i.i.d., which implies that $I_i[n]$, $i \in \mathbb{D}_k[n]$ are i.i.d. The assumption, along with the opportunistic selection of users, with the probability of each user "winning" a particular subcarrier being $1/K$, 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 . Since $u_k[n]$ is the summation of $|\mathbb{D}_k[n]|$ i.i.d. random variables, as $N \rightarrow \infty$ and consequently $|\mathbb{D}_k[n]| \rightarrow \infty$, we have

$$\frac{u_k[n] - E[u_k[n]]}{\sqrt{\text{Var}(u_k[n])}} \rightarrow \mathcal{N}(0, 1), \quad (11)$$

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 the standard procedure of conditional mean and variance formulas (conditioned on $\mathbb{D}_k[n]$). By ergodicity (which can be shown by the Law of Large Numbers), we have

$$\bar{u}_k = E[u_k[n]] = \frac{N}{K} E[I_i[n]]. \quad (12)$$

By choosing t_1 and t'_2 that enclose one OFDM symbol as shown in Fig. 4, (2) reduces to

$$u_k[n] \geq \bar{u}_k - v_k.$$

Note that v_k can be obtained statistically using the asymptotic distribution of $u_k[n]$ in (11). That is, with delay violation probability P_{dv} , we can find a corresponding $v_k(P_{dv})$ satisfying

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

or equivalently,

$$v_k(P_{dv}) = -Q^{-1}(1 - P_{dv}) \cdot \sqrt{\text{Var}(u_k[n])}. \quad (13)$$

By plugging (12) and (13) into (4), and having the first two moments of $I_i[n]$, we obtain

$$d_{pm} \leq \frac{w_{k1} + v_k(P_{dv})}{\bar{u}_k}. \quad (\text{OFDMA III}) \quad (14)$$

On the other hand, with fixed and predetermined $\mathbb{D}_k[n]$, OFDMA II is actually a special case of OFDMA III. Here, we have $|\mathbb{D}_k[n]| = N/K$, which leads to the same \bar{u}_k in (12) but a new

$$\text{Var}(u_k[n]) = \frac{N}{K} \text{Var}(I_i[n]).$$

Substituting $\text{Var}(u_k[n])$ into (13) and (14) yields the delay bound for OFDMA II.

C. OFDM II and OFDM III

Again, we first obtain results for OFDM III and then for OFDM II as a special case. The server process in OFDM III can be written as

$$u_k[n] = \begin{cases} u_{k,b}[n] = \sum_{i=1}^N I_i[n], & n \in \mathbb{C}(k), \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

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

$$\bar{u}_k = E[u_k[n]] = \frac{N}{K} E[I_i[n]]. \quad (16)$$

To find v_k , we suppose that there exists an C_{idle} such that any idle period (where $u_k[n] = 0$) is of length no greater than $C_{idle} \times K - 1$ for all user k . The C_{idle} can be chosen heuristically to ensure that such a violation probability is negligibly small, and thus the statistics associated with the opportunistic assignment is unchanged.

Similar to previous analysis for OFDM I, v_k is derived as follows. First, by the definition of (2), we have

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

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

$$\frac{u_{k,b}[n] - N \cdot E[I_i[n]]}{\sqrt{N \cdot Var(I_i[n])}} \rightarrow \mathcal{N}(0, 1), \quad (17)$$

we have

$$v_k(P_{dv}) = -Q^{-1}(1 - P_{dv}) \cdot \sqrt{N \cdot Var(I_i[n])} + (2C_{idle} - 1 - \frac{1}{K})N \cdot E[I_i[n]]. \quad (18)$$

Plugging (16) and (18) into (4) gives the delay bound

$$d_{pm} \leq \frac{w_{k1} + v_k(P_{dv})}{\bar{u}_k}. \quad (\text{OFDM III}) \quad (19)$$

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

V. SIMULATION RESULTS

Computer simulation is performed in this section to verify the analysis given in Sec. IV. The Rayleigh fading channel is adopted in the simulation. It is generated by a tapped delay line (TDL) channel model with equally-spaced taps and an exponential power delay profile. The parameters chosen for the TDL channel model are given in Table II. Channel coefficients of different users are generated identically and independently. A fast fading channel based on the method described in [11] is implemented. That is, the channel varies, presumably independently, across OFDM symbols. Although it is more practical to consider a framed structure where a frame contains several OFDM symbols and the channel varies on a frame-by-frame basis, the results presented can be applied to such a setting, too. Although OFDMA typically has a larger number of subcarriers than OFDM in real-world applications, we consider an identical 512-subcarrier setting for both systems for fair comparison purpose.

TABLE II
PARAMETERS OF THE TDL CHANNEL MODEL.

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

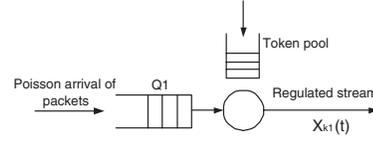


Fig. 6. A leaky bucket rate regulation scheme.

We implement the flow control regulator as shown in Fig. 6, by the “leaky bucket” scheme in [9], for which we choose the token pool size being 3000 bits and one token can serve 1 bit of data. The token arrival rates are chosen such that both the premium and the best-effort queues are stable, *i.e.*, $r_{k1} + r_{k2} \leq \bar{u}_k$. We choose a fixed packet length of 3000 bits, and the burst sizes w_{k1} and w_{k2} can be calculated to be 3001 bits. Besides, C_{idle} is chosen empirically such that the violation probability is less than 10^{-2} . For example, $C_{idle} = 4$ for $K = 4$.

In Fig. 7, we draw the analytical delay bounds derived in Sec. IV, after rounded up to the nearest integer number of OFDM symbols to reflect the real-world situations. We fix user channel SNR = 16 dB and $P_{dv} = 10^{-4}$ to obtain the bound. The SNR is chosen such that the average server rate in dynamic multiaccess modes (12) and (16) are equal to the constant server rate in static modes, for fair comparison. We see in Fig. 7 that OFDMA generally has better worst-case delay performance than OFDM. The performance gap increases as the number of users grows. The reason is because the round-robin or dynamic time-slot assignments in OFDM introduce idle period. The effect is most significant in OFDM

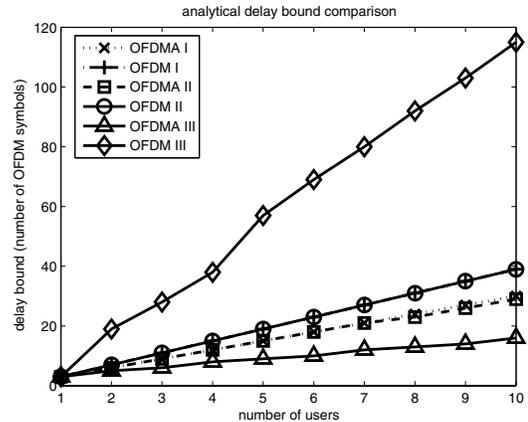


Fig. 7. Analytical delay bounds versus the number of users, K , with SNR = 16 dB.

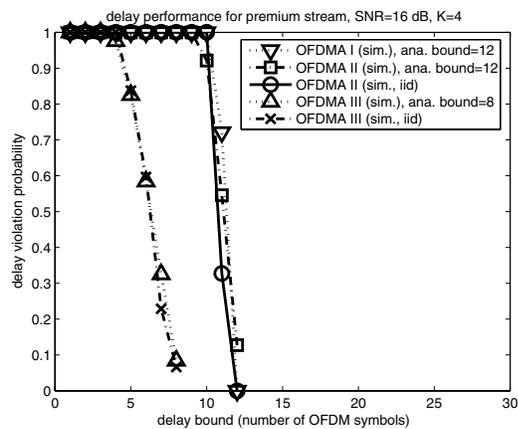


Fig. 8. Delay violation probability vs. delay bound for OFDMA modes, SNR = 16 dB and $K = 4$.

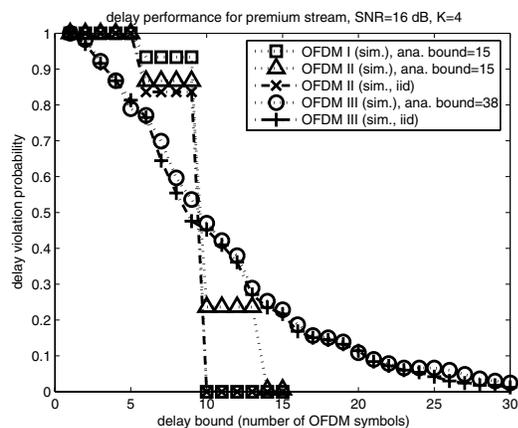


Fig. 9. Delay violation probability vs. delay bound for OFDM modes, SNR = 16 dB and $K = 4$.

III since the opportunistic assignment creates the possibility of a long idle period. Note that, when $K = 1$, all modes degenerate to the same scheme. As K increases, the delay bound increases due to the same resource shared among an increasing number of competitors.

In Figs. 8 and 9, we draw the Monte Carlo simulation curves when SNR = 16 dB and $K = 4$. These curves 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). The analytical delay bounds shown on the legend of the plots are obtained from Fig. 7 by fixing $K = 4$. We see from these curves that analytical delay bounds agree well with the actual packet maximum delay performance.

We see from Fig. 9 that the variation or the dynamic range of the packet delay for OFDM III is high. In contrast, OFDMA III in Fig. 8 has relatively small variation as shown by the sharper slope of curves. By examining Figs. 8 and 9 together, we observe that OFDM schemes generally have higher delay variation than OFDMA. This fact suggests that OFDMA is

more suitable for supporting real-time traffic, since real-time traffic may not be able to afford sensible (though perhaps only occasional) large delays.

Figs. 8 and 9 also show simulation results under the i.i.d. subcarrier channel setup. The results of the i.i.d. case are very close to its non-i.i.d. counterpart. This can be explained as follows. Although different bit allocation is performed on individual subcarriers, which affects instantaneous packet servicing, this effect is however mitigated by a number of summations in the process of calculating the overall packet delay. Therefore, 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. 8 and 9.

VI. CONCLUSION

The delay performance comparison of OFDM-TDMA and OFDMA was conducted in this work. Several OFDM modes with different multiaccess and resource allocation schemes were considered for packet delay analysis. We performed delay bound analysis based on the QoS architecture of IEEE 802.16 and a flow control scheme. The delay bounds derived, either deterministically or statistically, provide insights into performance differences among different multiaccess systems, and offer better understanding of the system's capability of supporting real-time, delay-sensitive traffic. The analytical and simulation results suggest that the opportunistic assignment can be employed more effectively in OFDMA. Specifically, dynamic OFDMA has a stronger potential of supporting real-time traffic than dynamic OFDM.

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 *IEEE MILCOM'03*, Oct. 2003, pp. 779–784.
- [3] 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.
- [4] 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.
- [5] —, "Queueing analysis of OFDM/TDMA systems," in *Proc. IEEE GLOBECOM'05*, Nov. 2005, pp. 3712–3716.
- [6] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1688–1692, Nov. 1999.
- [7] 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.
- [8] C. Dovrolis, D. Stiliadis, and P. Ramanathan, "Proportional differentiated services: Delay differentiation and packet scheduling," *IEEE/ACM Trans. Networking*, vol. 10, pp. 12–26, Feb. 2002.
- [9] 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.
- [10] M. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Wireless Personal Communications*, vol. 13, pp. 119–143, May 2000.
- [11] 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.