A Utility-Based TMCR Scheduling Scheme for Downlink Multiuser MIMO-OFDMA Systems

Chih-Ming Yen, Chung-Ju Chang, *Fellow, IEEE*, and Li-Chun Wang, *Senior Member, IEEE*

Abstract—In this paper, a utility-based throughput maximization and complexity-reduction (U_TMCR) scheduling scheme is proposed for downlink multiuser multiple-input–multiple-output orthogonal frequency-division multiple-access (MIMO-OFDMA) systems. The U_TMCR scheme allocates subchannels, antenna sequence, and modulation order to multimedia users with goals not only to maximize system throughput under quality-of-service (QoS) guarantee but to reduce computational complexity as well. Based on the channel quality and the QoS requirements of each user, the U_TMCR scheme designs a utility function for every user and formulates the scheduling into an optimization problem of overall system utility function subject to system constraints. It also contains a heuristic TMCR algorithm to *efficiently* solve the optimization problem. Simulation results show that the U_TMCR scheme achieves system throughput very close to the optimal solution by exhaustive search and higher than conventional schemes such as adaptive radio resource allocation (ARRA) and cross-layer design of packet scheduling (CDPS) by about 8% and 21%, respectively. The U_TMCR scheme also has a QoS satisfaction ratio that is better than the ARRA and CDPS schemes. Moreover, the U_TMCR scheme can reduce computational complexity. Generally, the total number of allocation trials of the U_TMCR scheme in a frame is smaller than that of the ARRA scheme by 6.25%–29.2%.

Index Terms—Computational complexity, multiple-input–multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM), quality-of-service (QoS) requirement, scheduling, throughput, utility function.

I. INTRODUCTION

MULTIPLE-INPUT–multiple-output-based orthogonal frequency-division multiplexing (MIMO-OFDM) gives a solution for future wireless communication since it can help to achieve high system capacity and provide transmit/receiver diversity for reliable communication link. Resource management for multiuser OFDM (MU-OFDM) systems has recently been investigated [1]–[5], where topics were focused on transmission power allocation, subcarrier allocation, bit allocation, or adaptive modulation and coding. The design goal is to maximize system capacity, minimize total transmission power, provide fairness, or guarantee quality-of-service (QoS) requirements.

Wong *et al.* proposed an optimal algorithm of subcarrier, bit, and power allocation for the MU-OFDM system to minimize the total transmission power using the Lagrangian multiplier [1]. However, the computational complexity is too high to make it feasible. Yin and Hui studied a so-called low-complexity algorithm, where the total bit rate was maximized, subject to users' rates and total power constraints by decoupling an NP-hard combinatorial problem into two steps [2]. The first step determines the required power and the number of subcarriers for each user; the second step then assigns the subcarrier and the number of bit for each user. However, the computation is still too complicated because of the Hungarian algorithm. Bala and Cimini proposed a low-complexity resource-allocation algorithm to minimize the total transmit power using the Lagrangian multiplier [3]. This algorithm was based on the idea of linearizing the original problem by transmitting data at the same rate on each subchannel. The computational complexity of this simplification method is still too high. Zhang and Letaief presented a two-step reduced-complexity subcarrier-bit-and-power-allocation algorithm, which first loosens the rate constraints and allocates subcarriers to maximize the throughput and afterward adjusts the subcarrier allocation step by step to satisfy the rate constrains [5]. This iterative algorithm still takes too much time to solve the problem, which is infeasible for real-time applications.

Lee *et al.* proposed an optimal opportunistic scheduler to maximize the total utility of a wireless system [6]. Neely presented a utility theory and delay tradeoffs for general data networks with stochastic channel condition and random arrival traffic [7]. Lei *et al.* proposed a packet-scheduling algorithm in the downlink of the orthogonal frequency-division multiple-access (OFDMA) system for mixed services [8]. These utility-based scheduling schemes [6]–[8] were proposed and optimized for single-input–single-output OFDMA systems, and they cannot be directly applied to MIMO-OFDMA systems. Another utility-based scheduling was proposed for single-user MIMO-OFDM system [9], and similarly, it cannot be extended to multiuser MIMO-OFDMA systems due to users' correlation. In addition, these utility-based scheduling schemes mapped the system resource (bandwidth or power) or performance measures (data rate or delay) to the corresponding utility and optimized the established utility system, where the utility function is either concave or convex. They then used the utility theory to solve the scheduling problem. However, they did not consider the QoS requirements of different service types, and they took only the minimum transmission rate or only the delay

time as the QoS requirement. The computational complexity is too high as well.

For MIMO-OFDM systems, the computational complexity on radio resource scheduling for downlink multiuser exponentially increases with the number of subcarrier, multiuser, transmitting antenna, and receiving antenna. Fuchs *et al.* [10] proposed a low-complexity space–time–frequency scheduling for MIMO systems with space division multiple access (SDMA). Hara *et al.* [11] proposed spatial scheduling with interference cancellation in multiuser MIMO systems. They focused on maximizing the system throughput but did not consider the QoS requirement. Xu *et al.* [12] presented a two-step multiuser scheduling algorithm for system throughput maximization with reduced complexity in a downlink MIMO/OFDMA system with transmitting preprocessing. They decoupled the multiuser scheduling problem into frequency and spatial domains. The preprocessing scheme decomposes the multiuser downlink MIMO channel into multiple parallel independent single-user MIMO channels (like OFDMA). However, the number of simultaneously transmitted users is restricted by the number of transmitting antennas. The computational complexity of the scheduling algorithm is still too high. Hu *et al.* proposed a low computational complexity scheduling algorithm for a downlink multiuser MIMO-OFDM system in [13]. The more the number of transmitting/receiving antenna and users is, the higher the system capacity will be. This results from space diversity and multiuser diversity. However, the scheduling algorithm did not consider the user demand and the QoS requirements; also, its symbol-by-symbol allocation is not suitable for current communication systems.

Yu *et al.* proposed a QoS guarantee resource-allocation algorithm for multiuser MIMO-OFDMA system, which is called cross-layer design of packet scheduling (CDPS) [14]. It serves users by considering fixed priority of service traffic. The priority orders of the service traffic are real-time (RT) service, nonreal-time (NRT) service, and best-effect service. The real-time traffic can be served in time at low traffic intensity, whereas the transmission rate of nonreal-time traffic is too low to fulfill the requirement rate. Tsai *et al.* proposed a dynamic priority scheduling scheme for the downlink OFDMA/SDMA system, which is called adaptive radio resource allocation (ARRA) scheme [15]. It gives high priority to urgent users and dynamically adjusts the priority of users frame by frame. However, the ARRA scheme adjusts the priority only depending on the time to expiration but not giving the clear differentiation of the RT service from the NRT service. This may result in that the RT service may have a high possibility to be served after the nonreal-time traffic at high traffic intensity.

This paper proposes a utility-based throughput maximization and complexity reduction (U_TMCR) scheduling scheme for downlink multiuser MIMO-OFDMA systems. The U_TMCR scheme first designs a utility function to adjust the priority of the user. The utility function consists of a QoS monitoring function and a radio resource function, where the QoS monitoring function indicates the service priority and the degree of urgency of the user, and the radio resource function shows the spectrum efficiency if it is used by the user. Subsequently, the U_TMCR scheme formulates the scheduling problem into

utility-based optimization equations with an objective of maximizing the overall system utility value under two system constraints. Finally, the U_TMCR scheme proposes a heuristic TMCR algorithm to solve the utility-based scheduling problem. The heuristic TMCR algorithm is a greedy method to maximize the overall system utility value. It contains two functions. The first function is a maximum utility allocation (MUA) function, which finds the most appropriate subchannel for the user and determines the receiving antenna by a multiuser antenna selection (MAS) scheme [16], [17] according to the highest individual utility value. The second function is a consistent allocation (CSA) function, which generalizes the allocation result determined by the MUA function on the following consecutive OFDMA symbols according to the urgency value such that the users' QoS requirements can be fulfilled and the computational complexity can be reduced.

Simulation results show that the proposed U_TMCR scheme can achieve system throughput very close to the optimal solution of the optimization equations by exhaustive search and higher than those of the ARRA scheme and the CDPS scheme by an amount of about 8% and 21%, respectively. In addition, the overall QoS satisfaction degree by the U_TMCR scheme is higher than those by the ARRA scheme and the CDPS scheme. The packet-dropping probability of RT services can be kept lower than the requirement by the U_TMCR scheme, but those by the ARRA and the CDPS schemes would violate the QoS requirement at high traffic intensity. Moreover, the U_TMCR scheme reduces the computational complexity. Generally, the total number of allocation trials of the U_TMCR scheme in a frame is reduced by 6.25%–29.2% compared with the ARRA scheme.

This paper is organized as follows: The system model of the considered downlink multiuser MIMO-OFDMA system is described in Section II. Section III presents the U_TMCR scheduling scheme. Section IV discusses the performance of the U_TMCR scheduling scheme. Finally, conclusions are given in Section V.

II. SYSTEM MODEL

Consider a downlink multiuser MIMO-OFDMA system with the proposed U_TMCR scheduling scheme. As shown in Fig. 1, assume that the system has K active mobile users, each mobile user k has Q_k receiving antennas, and the base station (BS) has N subchannels and Q transmitting antennas. The time is divided into frames, and there are L OFDMA downlink symbols per frame. At the beginning of a frame, the U_TMCR scheme computes the individual utility value for each user with QoS requirements over all subchannels and the receiving antenna. It allocates subchannels, antenna sequence, and modulation order to users based on these utility values.

By the allocation order from the U_TMCR scheduling scheme, the subchannel allocation block chooses the subchannel, the antenna allocation block assigns the antenna sequence, and the adaptive modulation block selects the modulation order to the user. Subsequently, the complex symbols at the output of the adaptive modulation block are transformed into the time-domain samples by inverse fast Fourier transform (IFFT). After
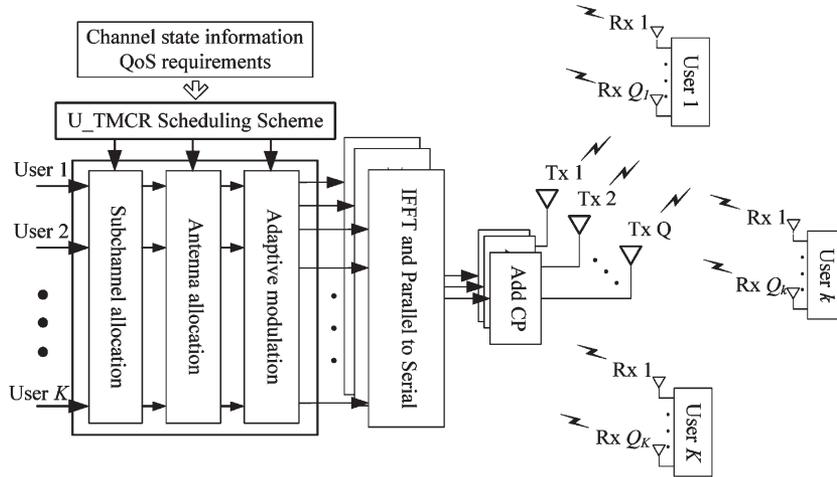
Fig. 1. System configuration of the downlink MIMO-OFDMA system.

IFFT and cyclic prefix (CP) addition, the transmission signal is then passed through different frequency-selective fading channels to different users. The CP is the guard interval, which is added to ensure orthogonality between the subchannels. If the length of the CP is longer than the maximum time dispersion, then the intersymbol interference is mitigated.

A. System Assumptions

In the multiuser MIMO-OFDMA system, assume that the MIMO channel on each subchannel is regarded as flat fading, and correlations among the receiving antennas of mobile users are negligible. The general channel matrix of the MIMO channel on a subchannel, which is denoted by $\mathbf{H}$, is expressed as

$$\mathbf{H}^T = [\mathbf{H}_1, \ldots, \mathbf{H}_k, \ldots, \mathbf{H}_K] \tag{1}$$

where $\mathbf{H}^T$ is the transpose vector of $\mathbf{H}$, $\mathbf{H}_k$ is the channel response of user k, $\mathbf{H}_k = [H_{q,r}]_{Q \times Q_k}$, $1 \leq k \leq K$, and $H_{q,r}$ is the impulse response between the transmitting antenna q at BS and the receiving antenna r at user k.

To achieve orthogonality among the selected users' MIMO channels, we adopt the multiuser antenna selection scheme [16], [17]. The MAS scheme first chooses one receiving antenna from mobile users, which can obtain a maximum multiple-input single-output channel capacity between the receiving antenna and the transmitting antennas at BS. Then, it selects the next receiving antenna from the remains by the same way. Note that to guarantee orthogonality, the next selected receiving antenna must not reduce the capacity of the previously selected receiving antennas. Repeat the foregoing process until the number of receiving antennas is equal to Q, which is the number of transmitting antennas. Consequently, the virtual MIMO channel matrix between the transmitting antennas at BS and the selected receiving antennas at chosen users on a subchannel can be constructed as

$$\overline{\mathbf{H}}^T = \left[\overline{\mathbf{H}}_1, \ldots, \overline{\mathbf{H}}_{k'}, \ldots, \overline{\mathbf{H}}_Q\right]_{Q \times Q} \tag{2}$$

where $\overline{\mathbf{H}}_{k'}$ is the $Q \times 1$ vector of impulse response between the transmitting antennas at BS and the selected receiving antenna at chosen user k', $1 \leq k' \leq Q$.

The zero-forcing beamforming [16] weight matrix on the subchannel, which is denoted by $\mathbf{W}$, can then be calculated by

$$\mathbf{W} = \overline{\mathbf{H}}^H \left(\overline{\mathbf{H}}\,\overline{\mathbf{H}}^H\right)^{-1} \mathbf{D} \tag{3}$$

where $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_k, \ldots, d_K)$ is the diagonal matrix that keeps the transmitting power unchanged after beamforming, and $\overline{\mathbf{H}}^H$ means the Hermitian transpose of $\overline{\mathbf{H}}$ [18]. The diagonal elements d_k are defined as

$$d_k = 1 \left/ \sqrt{\left[(\overline{\mathbf{H}}\,\overline{\mathbf{H}}^H)^{-1}\right]_{k,k}}\right. . \tag{4}$$

Because it is concluded that the equal power allocation algorithm over assigned subcarriers for each user can achieve performance similar to the water-filling scheme [19], in this paper, the equal power allocation is adopted. Therefore, the SNR of user k at the receiving antenna r on the subchannel can be given by

$$\mathrm{SNR}_{k,r} = \frac{P_{\mathrm{total}} d_k^2}{NQ\sigma_0^2} \tag{5}$$

where P_{total} is the total power of the system, and σ_0^2 is the noise power.

B. Urgency Values

There are three kinds of service classes provided for users: RT service, NRT service, and best-effort (BE) service. For RT services, the requirements of delay bound (D^*), bit error rate (BER^*), and dropping probability (P_D^*) are considered. For NRT services, the requirements of minimum transmission rate $(R_{\min}^*)$ and BER^* are taken into account. For BE service, only BER^* is required. In this paper, there are two kinds of RT services (i.e., voice and video), i.e., one kind of NRT service [i.e., hypertext transfer protocol (HTTP)] and one kind of BE service (i.e., FTP).

For the fulfillment of the users' QoS requirement, an urgency value for a user is defined. The urgency value of a user with RT or NRT service indicates how much of its tolerance time is left. The tolerance time of a user is defined as the residual time that the head-of-line (HOL) packet of the user at the present frame

remains before its violation of QoS requirements. The smaller (larger) the tolerance time, the higher (lower) the urgency value. Denote $\zeta_{\mathrm{RT}}(k)$ and $\zeta_{\mathrm{NRT}}(k)$ as the urgency values of user k with RT and NRT services, respectively, and define them as

$$\zeta_{\mathrm{RT}}(k) = \begin{cases} 1 + \frac{D^*}{D^* - F(k)}, & \text{if } 0 \leq F(k) < D^* \\ 0, & \text{if } F(k) \geq D^* \end{cases} \quad (6)$$

$$\zeta_{\mathrm{NRT}}(k) = \begin{cases} 1 + \frac{D^*}{F^*(k) - F(k)}, & \text{if } 0 \leq F(k) < F^* \\ 1 + D^*, & \text{if } F(k) \geq F^* \end{cases} \quad (7)$$

where $F(k)$ is the integer number of frames that the HOL packet of user k has experienced at the beginning of the present frame, and $F^*(k)$ is the maximum number of delay (unit of frames) that the HOL packet of user k with NRT service can have to achieve its minimum rate requirement $R_{\min}^*$. Note that $F^*(k)$ can be regarded as a requirement value in unit of time equivalent to $R_{\min}^*$ in units of bit rate. The numerators of fractions in (6) and (7) are set to be the same, making maximum values of $\zeta_{\mathrm{RT}}(k)$ and $\zeta_{\mathrm{NRT}}(k)$ equal to $1 + D^*$; the denominators in (6) and (7) are the differences just presenting the tolerance time for the RT and NRT services, respectively. For the RT service, if $F(k)$ is not less than D^*, then the RT packet will be dropped, resource scheduling is not considered, and thus, $\zeta_{\mathrm{RT}}(k) = 0$. For the NRT service, F^* can be expressed as

$$F^*(k) = \left\lceil \frac{L_{\mathrm{NRT}}(k)}{R_{\min}^*} \right\rceil + B(k) \quad (8)$$

where $L_{\mathrm{NRT}}(k)$ is the payload size of the NRT packet of user k, $B(k)$ is a balance factor of user k, and $\lceil x \rceil$ denotes the smallest integer larger than x. The balance factor $B(k)$ is used to achieve the goal that user k with the NRT service could be scheduled with an overall average rate close to $R_{\min}^*$. $B(k)$ is adaptively adjusted packet by packet and is set as

$$B(k) = \max \left\{ F_p^*(k) - F_p(k), 0 \right\} \quad (9)$$

where $F_p^*(k)$ $(F_p(k))$ is the $F^*(k)$ $(F(k))$ of the last NRT packet of user k. Equation (9) denotes that if the transmission time of the last NRT packet of user k is less than $F_p^*(k)$, then $B(k)$ will be a positive integer to allow the current NRT packet take longer to transmit but with the minimum rate requirement still satisfied.

The BS provides one individual FIFO queue for each service class of every active user. The packet of RT services will be dropped if the delay of the packet exceeds D^*. The packet of NRT services or BE services is allowed to be queued and delayed without being dropped if its buffer occupancy is not overflowing. Retransmission due to erroneous transmission of packets is not considered here.

III. UTILITY-BASED THROUGHPUT MAXIMIZATION AND COMPLEXITY REDUCTION SCHEDULING SCHEME

The U_TMCR scheme first designs a utility function for each user, then formulates the scheduling problem into optimization equations of overall system utility maximization, and finally solves the optimization equations by a heuristic TMCR algorithm for the MU-MIMO-OFDMA system.

A. Utility Function

Denote $U(k, n, r)$ as the individual utility function of user k on subchannel n using receiving antenna r at OFDMA symbols of a frame $1 \leq k \leq K$, $1 \leq n \leq N$, $1 \leq r \leq Q_k$. $U(k, n, r)$ is defined as a *radio resource function* of user k on subchannel n at the lth OFDMA symbol using receiving antenna r, which is denoted by $R(k, n, r)$, weighted by its QoS monitoring function at the same frame, which is denoted by $M(k)$. It is given by

$$U(k, n, r) = R(k, n, r) \times M(k). \quad (10)$$

The radio resource function $R(k, n, r)$ indicates the number of bits eligibly carried for user k on subchannel n with receiving antenna r per symbol at a frame. $R(k, n, r)$ has been obtained in [20, eq. (20)] as

$$R(k, n, r) = n_s \times \log_2 \left(1 - \mathrm{SNR}_{k,r} \times 1.5 / \ln(5 \times \mathrm{BER}^*)\right) \quad (11)$$

where n_s is the number of subcarriers in a subchannel. $\mathrm{SNR}_{k,r}$ has been given in (5) such that the orthogonality among the selected users' channels is obtained. A larger $R(k, n, r)$ denotes user k having a better channel quality, and more radio resources should be allocated to the user to achieve a higher system throughput. The larger the $R(k, n, r)$ is, the more system throughput that can be achieved.

The QoS monitoring function $M(k)$ is defined as

$$M(k) = \begin{cases} \gamma_{\mathrm{RT}} \times \zeta_{\mathrm{RT}}(k), & \text{if user } k \text{ is with RT service} \\ \gamma_{\mathrm{NRT}} \times \zeta_{\mathrm{NRT}}(k), & \text{if user } k \text{ is with NRT service} \\ \gamma_{\mathrm{BE}}, & \text{if user } k \text{ is with BE service} \end{cases} \quad (12)$$

where γ_{RT}, γ_{NRT}, and γ_{BE} are default priority constants for RT, NRT, and BE services, respectively, and $\zeta_{\mathrm{RT}}(k)(\zeta_{\mathrm{NRT}}(k))$ is the urgency value of user k with RT (NRT) service, which is given in (6) and (7). It would be obvious that $\gamma_{\mathrm{RT}} > \gamma_{\mathrm{NRT}} > \gamma_{\mathrm{BE}}$ since the RT (BE) service is the most (least) time-constraint service. The larger $M(k)$ is, the more urgently the user k will be served. This will guarantee the QoS requirement.

B. Problem Formulation

Assume that the allocation matrix $\mathbf{A}$ is given by $\mathbf{A} = [A_1^T, \ldots, A_l^T, \ldots, A_L^T]$. A_l is the allocation vector at the lth OFDMA symbol denoted by $A_l = [\delta_l(1, 1, 1), \ldots, \delta_l(1, 1, Q_1), \ldots, \ldots, \delta_l(k, n, r), \ldots, \delta_l(k, n, Q_k), \ldots, \delta_l(1, N, 1), \ldots, \delta_l(K, N, Q_K)]$, where $\delta_l(k, n, r)$, $1 \leq k \leq K$, $1 \leq n \leq N$, $1 \leq r \leq Q_k$ is an allocation indictor given by

$$\delta_l(k, n, r) = \begin{cases} 1, & \text{if subchannel } n \text{ is allocated to user } k \\ & \text{with receiving antenna } r \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

The U_TMCR scheme formulates the scheduling problem of the utility-based radio resource allocation in a frame for
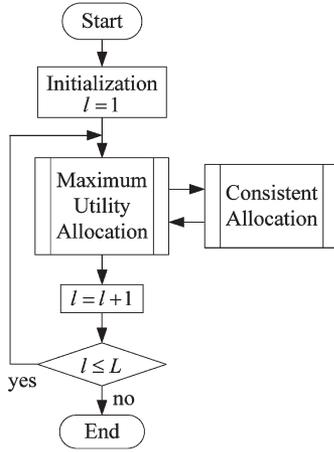
Fig. 2. Flowchart of the heuristic TMCR algorithm.

the downlink MIMO-OFDMA system as

$$\mathbf{A}^* = \arg\max_{\mathbf{A}} \left\{ \sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{r=1}^{Q_k} [U(k,n,r) \times \delta_l(k,n,r)] \right\} \tag{14}$$

subject to system constraints

$$\sum_{k=1}^{K} \sum_{r=1}^{Q_k} \delta_l(k,n,r) \leq Q \qquad \forall n \tag{15}$$

$$\sum_{l=1}^{L} \delta_l(k,n,r) \leq \left\lceil \frac{\xi_{\text{HOL}}^k}{\sum_{n=1}^{N} \sum_{r=1}^{Q_k} R(k,n,r) \times \delta_l(k,n,r)} \right\rceil \qquad \forall k \tag{16}$$

where the term inside the bracket of (14) is the overall system utility, and ξ_{HOL}^k in (16) is the residual bits of the HOL packet of user k. The constraint of (15) means that each subchannel can be allocated to Q users at most for the same symbol time, and the constraint of (16) implies that the bits allocated to each user cannot be larger than the residual bits of its HOL packet to avoid wasting resources. Notice that maximizing the overall system utility value in (14) can achieve the high system throughput and can guarantee QoS.

C. Heuristic TMCR Algorithm

Finally, the U_TMCR scheme contains a heuristic TMCR algorithm to solve the radio resource-scheduling problem given in (14)–(16). Because the optimal method by exhaustive search [21] to find an optimal solution takes too much time, which is infeasible to realize, the heuristic TMCR algorithm is proposed to maximize throughput and to reduce complexity. It can also conform to the allocation map structure defined in the standard [22]. Fig. 2 shows the flowchart of the heuristic TMCR algorithm. The heuristic TMCR algorithm mainly contains two functions: MUA and CSA. It finds an allocation map of a frame, allocates the radio resource from the first symbol of each subchannel, and schedules the allocation until there is no unallocated OFDMA symbol in the frame. At the initialization, the heuristic TMCR algorithm calculates the individual utility

value $U(k,n,r)$ given in (10) by using the MAS scheme [16], [17], $1 \leq k \leq K$, $1 \leq n \leq N$, $1 \leq r \leq Q_k$. Then, it uses the MUA function to schedule the allocation according to the individual utility value. It also uses the CSA function to perform continuous allocation to the user, which is determined by the MUA function, by an appropriate number of symbols for QoS fulfillment. In the heuristic TMCR algorithm, a possible allocation of a subchannel and a receiving antenna to a user is named as an allocation trial; the process to find the right channel and the right receiving antenna to the right user is called allocation iteration. Notice that the heuristic TMCR algorithm is a greedy method, and the result of the greedy method is close to the optimal solution when the number of user is large [21]. Furthermore, its computational complexity is much smaller than that of the exhaustive search.

MUA Function: The MUA function finds an optimal allocation indictor $\delta_l(k,n,r)$, which can have the largest individual utility value for the lth OFDMA symbol. The MUA function finds the highest individual utility value, which is denoted by $U(k^*,n^*,r^*)$, and then assigns the user k^* with the receiving antenna r^* on the subchannel n^* and checks whether the number of allocated users is equal to Q [comply with the constraint of (15)]. When the subchannel is assigned, the next allocated user cannot interfere with the previously assigned user. The MUA adopts the MAS scheme to guarantee orthogonality among mobile users and to recalculate the radio resource function $R(k,n^*,r)$ of the subchannel. Afterward, the MUA function will call the CSA function to accomplish resource allocation for user k^*. Repeat the next highest individual utility value user allocation and QoS requirement fulfillment until there is no unallocated user or no free subchannel in the lth OFDMA symbol. The pseudocode of the MUA function is given as follows:

Function: [MUA]

Step 1: Find $(k^*,n^*,r^*) = \underset{k \in \Omega, n \in S_{\text{free}}^l, r \in \Psi_{n,k}^l}{\arg\max} (U(k,n,r))$

Step 2: Set $\delta_l(k^*,n^*,r^*) = 1$, $S_{k^*} = S_{k^*} \cup \{n^*\}$, $\Psi_{n,k}^l = \Psi_{n,k}^l \setminus \{r^*\}$
  if $\sum_{k=1}^{K} \sum_{r=1}^{Q_k} \delta_l(k,n^*,r) = Q$
    $S_{\text{free}}^l = S_{\text{free}}^l \setminus \{n^*\}$
  else recalculate $R(k,n^*,r)$ for all $k \in \Omega, r \in \Psi_{n^*,k}^l$

Step 3: Call the CSA function

Step 4: Find $(k^*,n^*,r^*) = \underset{k=k^*, n \in S_{\text{free}}^l, r \in \Psi_{n,k^*}^l}{\arg\max} (U(k,n,r))$,
  and go to step 2

Step 5: Repeat Step 1–Step 3, if $\{S_{\text{free}}^l\} \neq 0$ and $\{\Omega\} \neq 0$ ∎

Note that in the foregoing pseudocode, S_{free}^l is the set of available subchannels that can be allocated at the lth OFDMA symbol, Ω is the set of unallocated users, $\Psi_{n,k}^l$ is the unallocated antenna of user k on subchannel n at the lth OFDMA symbol, and S_k is the set of subchannels that are allocated to user k.

CSA Function: The CSA function performs continuous allocation to fulfill the QoS requirement and to reduce the computational complexity and conforms to the allocation map structure

defined by the specification [22]. To satisfy the user's QoS requirement, the CSA function first determines an appropriate number of transmission symbols required for the allocated user selected by the MUA function and then performs CSA for the user. Assume that the selected user is user k. The required number of transmission symbols for user k, which is denoted by τ_k, is according to its degree of urgency $\zeta(k)$, which was given in (6) and (7) if the user k is with RT (NRT) service. The more urgent user would be allocated more transmission symbols to make the service fulfill its QoS requirement as much as possible. τ_k is designed as

$$
\tau_k = \begin{cases} 1, & \text{if } 1 < \zeta(k) \leq \frac{\zeta_{\max}(k)}{4} \\ \left\lfloor \frac{\tau_{k,\max}}{4} \right\rfloor, & \text{if } \frac{\zeta_{\max}(k)}{4} < \zeta(k) \leq \frac{\zeta_{\max}(k)}{2} \\ \left\lfloor \frac{2\tau_{k,\max}}{3} \right\rfloor, & \text{if } \frac{\zeta_{\max}(k)}{2} < \zeta(k) \leq \frac{3\zeta_{\max}(k)}{4} \\ \tau_{k,\max}, & \text{if } \frac{3\zeta_{\max}(k)}{4} < \zeta(k) \leq \zeta_{\max}(k) \end{cases} \quad (17)
$$

where $\zeta_{\max}(k)$ is the maximum value of $\zeta(k)$, and $\tau_{k,\max}$ is the maximum transmission symbols required by the current packet of user k. $\zeta_{\max}(k) = 1 + D^*$, and $\tau_{k,\max}$ can be expressed as

$$
\tau_{k,\max} = \left\lceil \frac{\xi_{HOL}^k}{\sum\limits_{n \in S_k} \sum\limits_{r=1}^{Q_k} R(k,n,r) \times \delta(k,n,r)} \right\rceil. \quad (18)
$$

If the unallocated OFDMA symbols of the allocated subchannel are smaller than τ_k, then the CSA function will go to Step 4 of the MUA function. The MUA function will assign another new subchannel to the user according to the utility value, add the subchannel to the S_k, and recalculate τ_k based on (17) and (18). Repeat the foregoing steps until τ_k is smaller than the unallocated OFDMA symbols of the allocated subchannel or there is no longer a free subchannel. Then, the CSA function generalizes the allocation to the same user in the following consecutive τ_k (remaining) unallocated OFDMA symbols and removes the user k from the unallocated users. Thus, the constraint of (16) can be complied, and the computation complexity of the U_TMCR can be reduced. The pseudocode of the CSA function is given as follows:

Function: [CSA]
Step 1: Calculate τ_{k^*}
Step 2: if $\tau_{k^*} > L - l + 1$
         if $\{S_{\text{free}}^l\} \neq 0$
             Go to Step 4 of the MUA function
         else
             Set $\delta_{l+1}(k^*,n,r^*) = \cdots = \delta_L(k^*,n,r^*) = 1$,
             $\forall n \in S_{k^*}$
             $\Omega = \Omega \setminus \{k^*\}$
             Go to Step 5 of the MUA function
         else
             Set $\delta_{l+1}(k^*,n,r^*) = \cdots = \delta_{l+\tau_{k^*}-1}(k^*,n,r^*) = 1$,
             $\forall n \in S_{k^*}$
             $\Omega = \Omega \setminus \{k^*\}$
             Go to Step 5 of the MUA function   ∎

TABLE I
SYSTEM-LEVEL PARAMETERS

Parameters	Value
Cell size	1.6 km
Frame duration	5 ms
System bandwidth	5 MHz
FFT size	512
Subcarrier frequency spacing	10.9375 KHz
Number of data subcarriers	384
Number of subchannels	8
Number of receiving antennas	2
Number of transmitting antennas	2
Number of data subchannel	8
Number of data subcarriers per subchannel	48
Number of slots for downlink transmission per frame	24
Maximum transmission power BS	43 dBm
Thermal noise density	-174 dBm/Hz

TABLE II
QoS REQUIREMENTS OF EACH SERVICE

Service / Requirement	voice (RT)	Video (RT)	HTTP (NRT)	FTP (BE)
Required BER (BER^*)	10^{-3}	10^{-4}	10^{-6}	10^{-6}
Maximum Packet Delay Tolerance (D^*)	40ms	10ms	N/A	N/A
Maximum Packet Dropping Ratio (P_D^*)	1%	1%	N/A	N/A
Minimum Required Transmission Rate ($R_{\min}^*$)	N/A	N/A	100kbps	N/A

N/A: Not Applicable

Note that the CSA function extends the allocation result spreading over the following consecutive OFDMA symbols, and this can reduce the system overhead for the U_TMCR scheme because the CSA function makes the user using the same subchannel with the same modulation order on several continuous OFDMA symbols in a frame. Therefore, the allocation map just needs to record information, which includes subchannel number and transmission period. For conventional allocation schemes, a user may be allocated on different subchannels in adjacent OFDMA symbols, which needs a large number of overhead bits to record it. Note that the TMCR algorithm immediately uses the CSA function to generalize the allocation result after the MUA function finds an allocation indictor. Therefore, the urgent user has a high probability to have long length for generalization and easily fulfill the QoS requirement.

IV. SIMULATION RESULTS

In the simulations, the system-level parameters of the downlink multiuser MIMO-OFDMA system are set to be compatible with the IEEE 802.16 standard [23], and the scalable parameters are configured according to the values suggested in [24] and listed in Table I. The path-loss model is modeled as $128.1 + 37.6 \log(R)$ dB, where R is the distance between the BS and the user in kilometers [25]. Log-normal shadowing is assumed with zero mean and standard deviation of 8 dB. The QoS requirements of each service are listed in Table II.

Each user is assumed to have only one service type. The voice service is modeled as the ON–OFF model, in which

Fig. 3.   System throughput.

Fig. 4.   (a) Voice packet-dropping rate. (b) Video packet-dropping rate.
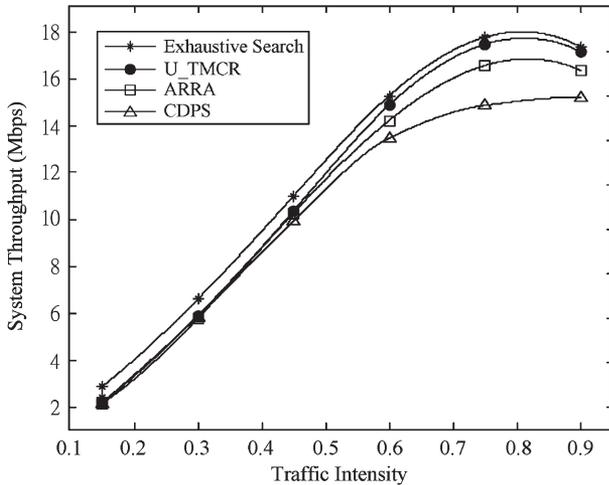
lengths of ON (OFF) period follow an exponential distribution with means of 1.0 (1.5) s [25]. The video data are assumed to arrive at a regular interval of 100 ms, with each frame decomposed into eight slices (packets), and the size of a packet is distributed in a truncated Pareto distribution [26], in which there are delay intervals between two consecutive packets of a frame, which denote the encoding delay at the video encoder. These intervals are modeled by a truncated Pareto distribution. The HTTP of NRT service is modeled as the behavior of Web browsing. Thus, the HTTP traffic is modeled as a sequence of page downloads, and each page download is modeled as a sequence of packet arrivals. The interval between two consecutive page downloads, representing the reading time in Web browsing, is distributed in an exponential distribution. For detailed parameters of video and HTTP traffic models, see [26]. The FTP traffic of BE user is modeled as a sequence of file downloads. The size of a file is distributed in a truncated log-normal distribution with mean of 2 MB, standard deviation of 0.722 MB, and maximum value of 5 MB. In addition, the interval between files is distributed in an exponential distribution with mean of 180 s. The average arrival data rates of voice, video, HTTP, and FTP are assumed to be 4.8, 64, 14.5, and 88.9 kb/s, respectively. The traffic intensity is defined as the ratio of the total average arrival data rate of all service types of all users over the maximum system transmission rate. Thus, the traffic intensity varies from 0.15 to 0.90 as the number of users varies from 80 to 480.

In the following, the proposed U_TMCR scheme is compared with the exhaustive search method [21], the CDPS [14] scheme, and the ARRA [15] scheme in multiuser MIMO-OFDMA systems for performance analysis. The exhaustive search method is the search that exhaustively searches the optimal solution of the optimization equations given in (14)–(16). The CDPS scheme and the ARRA scheme are modified to fulfill the MIMO-OFDMA system.

Fig. 3 shows the system throughput versus the traffic intensity for the proposed U_TMCR scheme, the exhaustive search method, the CDPS scheme in [14], and the ARRA scheme in [15]. It can be seen that the system throughput of the proposed U_TMCR scheme is very close to that of the exhaustive search
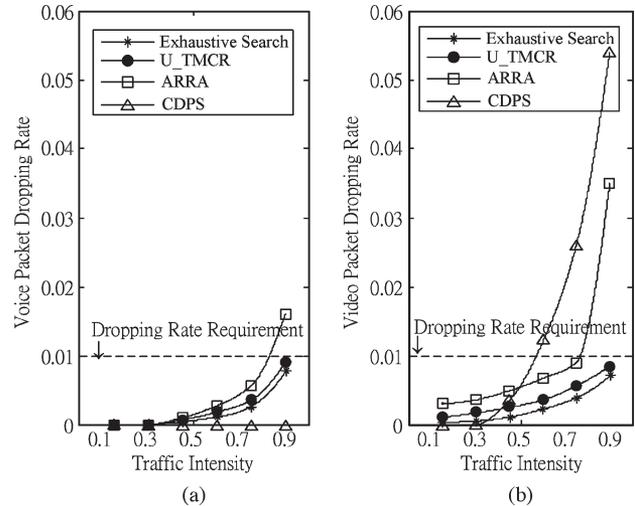
method; the former is less than the latter by only an amount of 2.12% at traffic intensity of 0.75. The reason is that the heuristic U_TMCR is a kind of greedy method [21]. In addition, at high traffic intensity, the proposed U_TMCR scheme can achieve system throughput higher than the ARRA scheme and the CDPS scheme by an amount of 8% and 21%, respectively, at traffic intensity of 0.75. It is because the U_TMCR scheme allocates the resource according to the utility function given in (10), where a user with a high individual utility value is the user who is urgent and with the good channel quality, whereas the services order of users in the ARRA scheme is only based on the urgency that denotes the residual time to expiration. This would make the U_TMCR scheme attain higher system throughput than the ARRA scheme. In addition, the U_TMCR scheme designs the maximum urgency values of the RT and NRT packets to be the same but sets various priority constants for them in the QoS monitoring function to further differentiate the services so that the RT packet, which is with the same urgency as the NRT packet, will have a higher probability to be served. The ARRA scheme adaptively adjusts the priority of users according to just the time to expiration. There are possibilities that the priority value of the NRT packet may be larger than that of the RT packet. Thus, some packets of the RT traffic will be squeezed and not be served in time at high traffic intensity. These RT packets will be dropped, resulting in system throughput decrement for the ARRA scheme. On the other hand, the CDPS scheme serves users by fixed priority, and thus, the multiuser gain cannot obviously appear. Additionally, when the traffic intensity increases, the probability of serving video packets would become small, and the video packet-dropping rate intuitively increases. This makes the system throughput of the CDPS scheme the smallest.

Fig. 4(a) and (b) shows the voice and video packet dropping rates, respectively, where the dropping rate requirements (1%) are also given. It can be seen that the U_TMCR performs quite close to the exhaustive search solution in the performance measures of the two RT packet-dropping rates. The U_TMCR (ARRA) scheme keeps the voice packet-dropping rate lower than the voice packet-dropping probability requirement of 1%
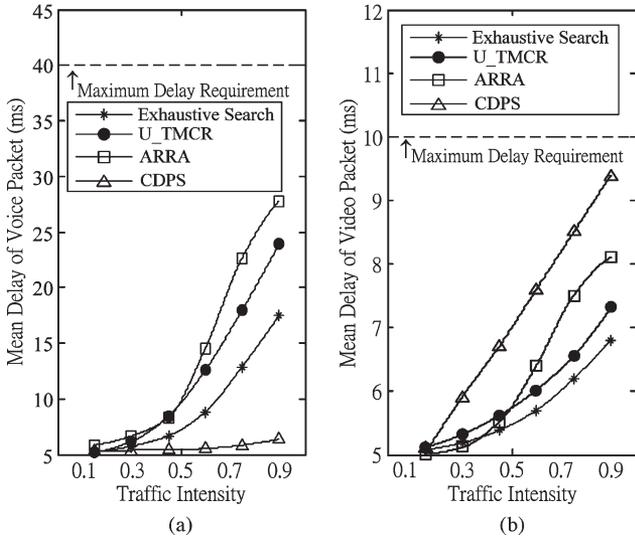
Fig. 5.   (a) Mean delay of voice packet. (b) Mean delay of video packet.

Fig. 6.   Guaranteed ratio of HTTP packets.

Fig. 7.   BE throughput.

until the traffic intensity is above 0.9 (0.8), whereas the CDPS scheme attains voice packet-dropping rates always close to zero. In addition, the U_TMCR scheme keeps the video packet-dropping rate lower than the video packet-dropping probability requirement of 1% until the traffic intensity is near 1.0, whereas that of the CDPS (ARRA) scheme violates the requirement at traffic intensity of 0.56 (0.76). The reasons are quite similar to what we give for the system throughput comparison among the U_TMCR, ARRA, and CDPS schemes in Fig. 3.

Fig. 5(a) and (b) shows the mean voice and video packet delays, respectively, where the maximum packet-delay requirements (D^*) are also included. It can be seen that the mean RT packet delays by the U_TMCR scheme are larger than the exhaustive search method. The CDPS scheme employs the fixed priority; thus, it has the smallest voice packet delay but the largest video packet delay. Both the U_TMCR scheme (exhaustive search method) and the ARRA scheme exploit the dynamic priority and make full use of the voice packet-delay tolerance. The three schemes delay the voice packets to serve more video packets in time. Thus, the delays of the video packet of these three schemes are smaller than that of the CDPS scheme. It can also be found that the voice and video packet delays by the U_TMCR scheme are smaller than by the ARRA scheme at traffic intensity of over 0.5. The reason is that the urgent packet of the RT user in the U_TMCR scheme has the highest priority, whereas the priority of the urgent packet of the RT user in the ARRA scheme may be smaller than that of NRT packet; the urgent RT users in the U_TMCR scheme have higher probability to be served in time.

Fig. 6 illustrates the guaranteed ratio for HTTP packets. Unlike the RT traffic, packets of the HTTP traffic will not be dropped but will still wait for service when the requirement of minimum transmission rate $R^*_{\min}$ cannot be kept. It can be observed that the guaranteed ratio of the U_TMCR scheme is close to the exhaustive search method, and it is much higher than that of the CDPS scheme but lower than that of the ARRA scheme. As with the same reason previously given, by the CDPS scheme, the priority of the HTTP packets is the third,
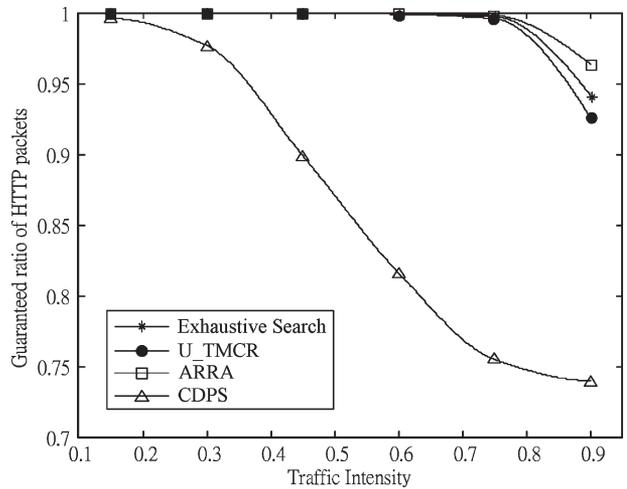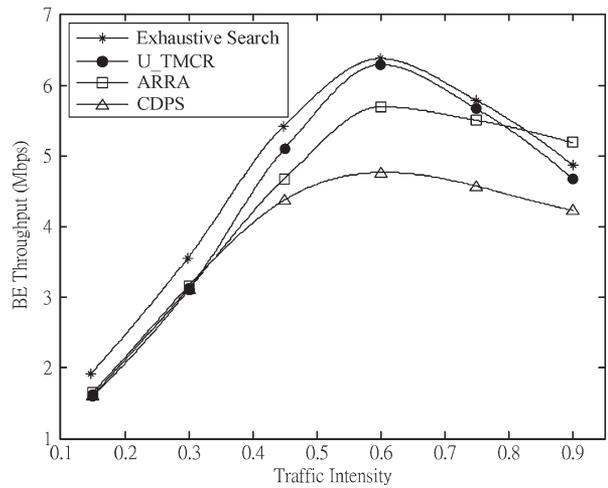
which means that HTTP packets have to wait until all real-time packets are served. By the U_TMCR scheme (the exhaustive search method) and the ARRA scheme, since the priority of users is dynamically adjusted frame by frame, it can avoid more resources always being occupied by RT traffic. Thus, more HTTP packets can be guaranteed to be served with the minimum transmission rate. Additionally, the ARRA scheme further gives the NRT service to override the RT service to be first served more possibly than the U_TMCR scheme when the requirement of minimum transmission rate $R^*_{\min}$ is going to be violated. Thus, the ARRA scheme has the largest guaranteed ratio for HTTP packets.

Fig. 7 shows the BE throughput versus the traffic intensity. It can be found that the U_TMCR scheme has BE through-put lower than the exhaustive search method, as the system throughput shown in Fig. 3; the U_TMCR scheme (the exhaustive search method) and the ARRA scheme outperform the CDPS scheme in the BE throughput. The BE service in the U_TMCR scheme (the exhaustive search method) still has more chances to transmit. The reason is that the design of utility function in (10) can achieve higher multiuser gain and make the resource allocation more efficiently. Furthermore, it

could make the RT users be delayed near the delay bound and the NRT users with good channel quality be served with high probability. It can also be found that the BE throughput of the U_TMCR scheme (the exhaustive search method) increases until the traffic intensity is 0.6 and then decreases. It is because the U_TMCR scheme (the exhaustive search method) will give the RT services more resource to satisfy their QoS requirements when the traffic intensity is higher, which can seen in Fig. 4. For the CDPS scheme, the FTP traffic will be transmitted only when voice, video, and HTTP traffic have already been served. Thus, its average throughput of BE is the lowest.

Finally, the computational complexity among the exhaustive search method and the U_ TMCR, ARRA, and CDPS schemes is compared. The computation complexity of a scheme depends on the number of allocation trials in an allocation iteration and the total number of allocation iterations needed for the scheme. Theoretically, the worst-case computational complexity for the four schemes would be $O(LKN^2Q_M)$, where Q_M is the maximal number of receiving antenna Q_k. However, either the CSA function in the U_TMCR scheme or the generalization function in the ARRA scheme continues allocation to the same user in several of the following OFDMA symbols. Thus, the complexity of the U_TMCR and ARRA schemes would greatly be reduced by at least L times over the exhaustive search method and the CDPS scheme in the worst case.

However, the U_TMCR scheme has a lower computational complexity than the ARRA scheme in realistic operations. The U_TMCR scheme uses the CSA function to generalize the allocation result immediately after the MUA function has allocated a subchannel to a user. Thus, if a user needs more than one subchannel to transmit, then the U_TMCR scheme performs allocation for the user in just *one* iteration and removes the user from the unallocated user set. On the other hand, the ARRA scheme extends the results after all subchannels of the OFDMA symbol are allocated. Thus, the ARRA scheme needs *more* number of allocation iteration, which is equal to the number of required subchannels of the user, and it removes the user until the QoS requirements of the user are satisfied. Therefore, the U_TMCR scheme needs the number of allocation iteration to allocate the resource and the number of allocation trials in the following allocation iteration smaller than the ARRA scheme. Take one example. Assume that users are with packet length uniformly distributed between 72 and 576 B. The number of allocation trials in a frame by the U_TMCR scheme is reduced by 6.25%–29.2% compared with the ARRA scheme. The larger the packet length is, the larger the reduction of computation complexity by the U_TMCR scheme would be.

We show in Fig. 8 the computation time of the four scheduling schemes, where the dotted line denotes the frame time (5 ms). It can be observed that the exhaustive search takes the largest computation time; the computation times of the U_TMCR and ARRA schemes are less than one frame time, and they increase until the traffic intensity is 0.45 and then decrease. It is because, as described in the last two paragraphs, the exhaustive search method searches all the possible combinations that the CSA function in the U_TMCR scheme and the generalization function in the ARRA scheme work effectively. However, the benefit of the CSA or the generalization function
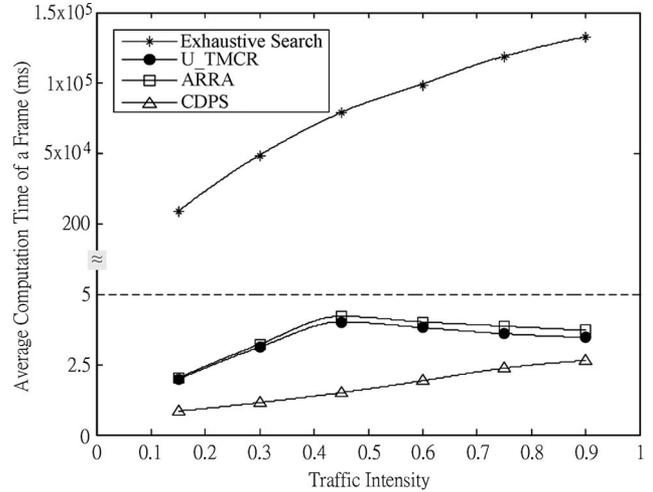
Fig. 8.   Average computation time for a frame allocation.

is not obvious when the traffic intensity is low, and it needs more allocation iterations to allocate L symbols of a frame. When the traffic intensity is high, the CSA function can consistently allocate more OFDMA symbols, and the computation times are decreased. In addition, as expected, the computation time of the U_TMCR scheme is lower than that of the ARRA scheme. The improvement is at 7.2% at a traffic intensity of 0.75. Since the CDPS scheme performs fixed priority allocation, the number of allocation trials in an allocation iteration is one quarter of that of the other three schemes, and the number of allocation iteration of the other three schemes is more than one quarter of that of the CDPS scheme, resulting in the computation time of the CDPS scheme being the smallest in reality.

V. CONCLUSION

In this paper, a U_TMCR scheduling scheme has been proposed for downlink multiuser MIMO-OFDMA systems, where the radio resource allocation to multimedia users includes subchannel allocation, modulation order assignment, and receiving antenna. The goals of the U_TMCR scheme are to maximize system throughput, fulfill QoS requirements, and lessen the computational complexity while considering multiple service classes, such as RT, NRT, and BE services. The proposed U_TMCR scheme designs a utility function, formulates the utility-based scheduling problem, and solves the problem by a heuristic TMCR algorithm. For RT (NRT) service, the value of the utility function is dynamically adjusted to maximize the spectrum efficiency and to fulfill the delay requirement (minimum transmission rate) and the BER requirement. The heuristic TMCR algorithm includes a MUA function to maximize the overall utility value and a CSA function to fulfill the QoS requirements of users and reduce the computational complexity.

Simulation results show that the performance of the proposed U_TMCR scheme is very close to that of the exhaustive search method, and the proposed U_TMCR scheme outperforms the conventional ARRA scheme [15] and the CDPS scheme [14] in system throughput by an amount of 8% and 21.5%, respectively. The U_TMCR scheme can sustain the users' QoS

requirement up to the traffic intensity of 0.9, whereas the ARRA (CDPS) scheme cannot guarantee QoS requirements after a traffic intensity of 0.8 (0.55). The overall QoS satisfaction by the U_TMCR scheme is higher than that by the ARRA and CDPS schemes. This is because the U_TMCR scheme adjusts the individual utility value of every user by considering the users' service priority, urgency degree, and channel quality, frame by frame. In addition, the U_TMCR scheme gives the RT (NRT) users with larger packet delay (lower average transmission rate) a higher priority to quickly obtain the resource and clearly differentiates the RT and NRT users. Moreover, the U_TMCR scheme can reduce the number of allocation iterations. The total number of the allocation trials in a frame by the U_TMCR scheme is reduced by 6.25%–29.2% compared with the ARRA scheme.

REFERENCES

[1] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[2] H. Yin and L. Hui, "An efficient multiuser loading algorithm for OFDM-based broadband wireless system," in *Proc. IEEE CLOBECOM*, 2000, vol. 1, pp. 103–107.

[3] E. Bala and L. J. Cimini, Jr., "Low-complexity and robust resource allocation strategies for adaptive OFDMA," in *Proc. VTC—Fall*, 2005, vol. 1, pp. 176–180.

[4] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access system," *IEEE Trans. Broadcast.*, vol. 49, no. 4, pp. 367–370, Dec. 2003.

[5] Y. J. Zhang and K. B. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1566–1575, Sep. 2004.

[6] N. H. Lee, J. G. Choi, and S. Bahk, "Opportunistic scheduling for utility maximization under QoS constraints," in *Proc. PIMRC*, 2005, pp. 1818–1822.

[7] M. J. Neely, "Super-fast delay tradeoffs for utility optimal fair scheduling in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1489–1501, Aug. 2006.

[8] H. Lei, L. Zhang, X. Zhang, and D. Yang, "A packet scheduling algorithm using utility function for mixed services in the downlink of OFDMA systems," in *Proc. VTC—Fall*, 2007, pp. 1664–1668.

[9] Z. Chen, K. Xu, F. Jiang, Y. Wang, and P. Zhang, "Utility based scheduling algorithm for multiple services per user in MIMO OFDM system," in *Proc. ICC*, 2008, pp. 4734–4738.

[10] M. Fuchs, G. D. Galdo, and M. Haardt, "Low-complexity space-time-frequency scheduling for MIMO systems with SDMA," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2775–2784, Sep. 2007.

[11] Y. Hara, L. Brunel, and K. Oshima, "Spatial scheduling with interference cancellation in multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 893–905, Mar. 2008.

[12] N. Xu, Y. Wang, and P. Zhang, "Multiuser scheduling in downlink MIMO/OFDMA system with transmit preprocessing," in *Proc. Asia-Pacific Conf. Commun.*, Aug. 2006, pp. 1–5.

[13] Y. Hu, C. Yin, and G. Yue, "Multiuser MIMO-OFDM with adaptive antenna and subcarrier allocation," in *Proc. VTC—Spring*, 2006, vol. 6, pp. 2873–2877.

[14] J. Yu, Y. Cai, Y. Ma, D. Zhang, and Y. Xu, "A cross-layer design of packet scheduling and resource allocation for multiuser MIMO-OFDM system," in *Proc. Inf. Commun. Signal Process. Conf.*, Dec. 2007, pp. 1–5.

[15] C. F. Tsai, C. J. Chang, F. C. Ren, and C. M. Yen, "Adaptive radio resource allocation for downlink OFDMA/SDMA systems with multimedia traffic," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1734–1743, May 2008.

[16] J. Kim, S. Park, J. H. Lee, J. Lee, and H. Jung, "A scheduling algorithm combined with zero-forcing beamforming for a multiuser MIMO wireless system," in *Proc. IEEE VTC—Fall*, 2005, pp. 211–215.

[17] J. Zhang, G. Liu, and W. Wang, "Multiuser antenna selection for zero forcing beamforming based MIMO OFDMA," in *Proc. IEEE APCC*, 2007, pp. 107–110.

[18] B. Noble and J. W. Danlel, *Applied Linear Algebra,* 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[19] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.

[20] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[21] K. G. Murty, *Operations Research.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[22] WiMAX forum, "WiMAX system evaluation methodology," V.1.0, Tech. Rep., Jan. 2007.

[23] *IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std. 802.16-2004, Oct. 2004.

[24] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 Wireless-MAN," *Intel Technol. J.*, vol. 8, no. 3, pp. 202–212, 2004.

[25] 3GPP TR 25.892, "Feasibility study for OFDM for UTRAN enhancement," 3rd Generation Partnership Project, Tech. Rep., 2004-06.

[26] *Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS*, UMTS Std. 30.03, 1998.

Chih-Ming Yen was born in Tainan, Taiwan. He received the Ph.D. degree in communication engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 2010.

He is currently with the Department of Communication Engineering, College of Electrical Engineering and Computer Science, National Chiao Tung University. His interest includes wireless networks, mobile communications, high-speed networks, communications protocol design, and network performance evaluation. He is currently focusing on resource management for orthogonal frequency-division multiplexing systems. His research interests also include radio resource management and wireless communication.

Chung-Ju Chang (S'81–M'85–SM'94–F'05) was born in Taiwan in August 1950. He received the B.E. and M.E. degrees in electronics engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 1972 and 1976, respectively, and the Ph.D. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1985.

From 1976 to 1988, he was a Design Engineer, Supervisor, Project Manager, and then Division Director with Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, Taiwan. From 1987 to 1989, he acted as a Science and Technical Advisor for the Minister of the Ministry of Communications. In 1988, he joined the Faculty of the Department of Communications Engineering, College of Electrical and Computer Engineering, National Chiao Tung University, as an Associate Professor, where he has been a Professor since 1993. He was the Director of the Institute of Communication Engineering from August 1993 to July 1995, Chairman of the Department of Communication Engineering from August 1999 to July 2001, and the Dean of the Research and Development Office from August 2002 to July 2004. He is currently a Chair Professor with the Department of Electrical Engineering, National Chiao Tung University. In addition, he was an Advisor for the Ministry of Education to promote the education of communication science and technologies for colleges and universities in Taiwan during 1995–1999. He is acting as a Committee Member of the Telecommunication Deliberate Body, Taiwan. His research interests include performance evaluation, radio resource management for wireless communication networks, and traffic control for broadband networks.

Dr. Chang is a member of the Chinese Institute of Engineers and the Chinese Institute of Electrical Engineers. He serves as an Editor for the IEEE COMMUNICATIONS MAGAZINE and as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Li-Chun Wang (S'92–M'96–SM'06) received the B.S. degree in electrical engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 1986, the M.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1988, and the M.Sc. and Ph.D. degrees in electrical engineering from Georgia Institute of Technology, Atlanta, in 1995 and 1996, respectively.

From 1990 to 1992, he was with Chunghwa Telecom, Taoyuan, Taiwan. In 1995, he was with Northern Telecom, Richardson, TX. From 1996 to 2000, he was a Senior Technical Staff Member with the Wireless Communications Research Department, ATT Laboratories. In August 2000, he was an Associate Professor with the Department of Electrical Engineering, National Chiao Tung University, where he has been a Full Professor since August 2005. He is the holder of three U.S. patents and has three patents pending. His research interests include cellular architectures, radio network resource management, and cross-layer optimization for cooperative and cognitive wireless networks.

Dr. Wang was a corecipient of the Jack Neubauer Best Paper Award from the IEEE Vehicular Technology Society in 1997.