QoS-Driven Asynchronous Uplink Subchannel Allocation Algorithms for Space-Time OFDM-CDMA Systems in Wireless Networks

Xi Zhang and Jia Tang
Networking and Information Systems Laboratory
Department of Electrical Engineering
Texas A&M University, College Station, TX 77843, USA
Email: {*xizhang, jtang*}*@ee.tamu.edu*

***Abstract*— In order to support the diverse Quality of Service (QoS) requirements for differentiated data applications in broadband wireless networks, advanced techniques such as space-time coding (STC) and orthogonal frequency division multiplexing (OFDM) are implemented at the physical layer. However, the employment of such techniques evidently affects the subchannel-allocation algorithms at the medium access control (MAC) layer. In this paper, we propose the QoS-driven cross-layer subchannel-allocation algorithms for data transmissions over asynchronous uplink space-time OFDM-CDMA systems. We mainly focus on QoS requirements of maximizing the best-effort throughput and proportional bandwidth fairness, while minimizing the upper-bound of scheduling delay. Our extensive simulations show that the proposed infrastructure and algorithms can achieve high bandwidth fairness and system throughput while reducing scheduling delay over wireless networks.**

***Index Terms*— QoS, cross-layer design, subchannel-allocation, wireless networks, space-time spreading, OFDM, fairness, throughput, delay.**

I. Introduction

The increasing demand for wireless network services such as the wireless Internet access, mobile computing, and wireless communications motivates an unprecedented revolution in the wireless broadband access [1]. This presents great challenges in designing the wireless networks since the wireless channel has a significant impact on supporting the various Quality of Service (QoS) requirements for different users.

A number of promising schemes are developed at the physical layer to overcome the impact of wireless channels. Among them, space-time (ST) processing using multiple-input-multiple-output (MIMO) architecture emerges as one of the important technical breakthroughs in wireless communications [2]-[6]. Besides, the combination of the widely employed code division multiple access (CDMA) with orthogonal frequency division multiplexing (OFDM), called OFDM-CDMA, takes the advantages of these two techniques and also receives a great deal of research efforts [6]-[9]. Clearly, employment of the integrated design combining space-time processing and OFDM-CDMA can achieve the integrated diversities from spatial, temporal, frequency, and code domains, which will result in significant improvements in supporting QoS for differentiated data users with different QoS requirements over wireless networks.

Space-Time OFDM-CDMA provides us with not only the diversities, but also the multiple access control, such that multiple users can be assigned into a single subchannel, where they distinct themselves from each other by different signature sequences. As a result, how to allocate the resources efficiently in meeting the various QoS requirements becomes increasingly critical. While there have already been a large body of literature on both space-time processing and OFDM-CDMA, the impact of such architectures on resource allocations at MAC layer, its bandwidth fairness, throughput, and delay analysis, and cross-layer optimizations, have received relatively much less attention. Therefore, it is important to develop a cross-layer scheme to integrate the resource allocation at the MAC layer and the multi-antenna infrastructure implemented at the physical layer.

In this paper, we propose the QoS-driven cross-layer subchannel-allocation algorithms for asynchronous uplink space-time OFDM-CDMA systems targeting at differentiated data transmission applications. The QoS requirements that we focus on include transmission reliability, bandwidth fairness, system throughput, and scheduling delay. Specifically, the proposed algorithms are based on the delay-fairness-driven scheduling, proportional bandwidth fairness, and best-effort throughput. Also, we conduct extensive simulations to evaluate the performance of the proposed algorithms. Our simulation results show that the proposed infrastructures and algorithms can significantly improve the bandwidth fairness and system throughput, while achieving more efficient resource allocations over wireless networks.

The paper is organized as follows. Section II describes the space-time OFDM-CDMA system model. Section III proposes the subchannel-allocation algorithms. Section IV evaluates and compares the various performance metrics through simulations and numerical solutions. The paper concludes with Section V.

II. System Model

We consider the BPSK modulation-based uplink from a mobile user to the basestation in a packet-cellular network with

This work was supported in part by the National Science Foundation CAREER Award under Grant ECS-0348694.

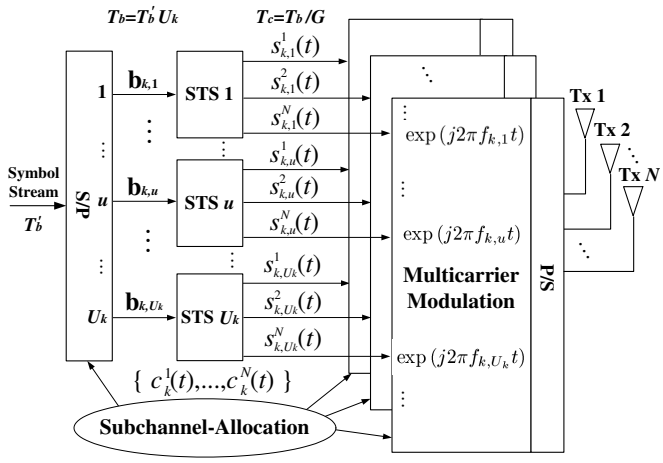


Fig. 1. Block diagram of the kth mobile user's uplink transmitter.

M antennas at the basestation (BS) and N antennas at each mobile user. Let K denote the total number of mobile users and U the total number of subcarriers or subchannels,[1] which are to be assigned to the K mobile users. The index set of all K users is defined as $\Omega \triangleq \{1, 2, ..., K\}$ and the U subcarrier frequencies are denoted by the set of $\{f_u\}_{u=1}^{U}$. In addition, at any instantaneous time point, only the users which have been assigned with bandwidth resources can transmit data packets. These users assigned with bandwidth resources are defined as *active* users. Denote the index set of active users by Φ. Clearly, $\Phi \subseteq \Omega$.

A. Mobile Uplink Transmitter Model

The kth mobile user's transmitter structure of our proposed space-time OFDM-CDMA system is shown in Fig. 1. In this paper, we consider the asynchronous uplink, where different users transmit data asynchronously to the basestation, and thus we cannot use the synchronous model as used for downlink from the basestation. As shown in Fig. 1, using the Serial-to-Parallel (S/P) converter, a block of $U_k \times N$ BPSK symbols each with bit duration of T_b' is converted to U_k parallel sub-streams. Each of U_k sub-streams consists of N bits, which are denoted by $\mathbf{b}_{k,u} = \left(b_{k,u}^1\ b_{k,u}^2\ ...\ b_{k,u}^N\right)^T$, where $(\cdot)^T$ represents the transpose of $(\cdot)$ and $u \in \{1, 2, ..., U_k\}$. The bit duration T_b after S/P conversion becomes $T_b = T_b' U_k$. The value $U_k (U_k \leq U)$ is determined by our proposed subchannel-allocation algorithms, which will be described in Section III with more details. Then, each $\mathbf{b}_{k,u}$ is space-time spread (STS) [3] using the spreading code given by $\mathbf{c}_k = (c_k^0\ c_k^1\ ...\ c_k^{G-1})^T$, where $c_k^g \in \{\pm 1\}$, $g \in \{0, 1, ..., G-1\}$ and G denotes the spreading gain of the code. The chip duration T_c of the spreading code satisfies $T_c = T_b/G = T_b' U_k/G$. The waveform expression $c_k(t)$ of the spreading code corresponding to $\mathbf{c}_k$ is determined by

$$c_k(t) = \frac{1}{\sqrt{G}} \sum_{g=0}^{G-1} c_k^g\, p(t - gT_c), \quad 0 \leq t < T_b \tag{1}$$

[1]We use the terms "subchannel" and "subcarrier" interchangeably in the following discussions.

where $p(t)$ is a normalized rectangular chip waveform which has the finite duration $[0, T_c)$. In this paper, we focus on a specific subset of the general STS schemes investigated in [3], by which the N-bit data is coded, spread, and allocated to N transmit antennas, and then transmitted by N time intervals. This kind of STS schemes can provide the maximal transmit diversity without demanding extra spreading codes and thus be considered as *attractive* schemes [6]. Denote the corresponding space-time block coding square matrix of the uth sub-stream by

$$\mathbf{B}_{k,u} = \begin{pmatrix} b_{k,u}^{1,1} & b_{k,u}^{1,2} & \cdots & b_{k,u}^{1,N} \\ b_{k,u}^{2,1} & b_{k,u}^{2,2} & \cdots & b_{k,u}^{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k,u}^{N,1} & b_{k,u}^{N,2} & \cdots & b_{k,u}^{N,N} \end{pmatrix} \begin{matrix} \rightarrow \text{space} \\ \\ \downarrow \ \text{time} \end{matrix} \tag{2}$$

where the rows and columns of matrix $\mathbf{B}_{k,u}$ consist of $\mathbf{b}_{k,u}$ with different signs and sequences according to the orthogonal design rules [4]. The spreading code vector $\mathbf{c}_k(t)$ used for STS can be expressed as

$$\mathbf{c}_k(t) = \left(c_k^1(t)\ c_k^2(t)\ ...\ c_k^N(t)\right) \tag{3}$$

where $c_k^i(t)$, $i \in \{1, 2, ..., N\}$ has the finite duration $[0, NT_b)$, which is given by

$$c_k^i(t) = \begin{cases} c_k(t - iT_b + T_b), & (i-1)T_b \leq t < iT_b \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Using Eqs. (2) and (3), STS can be expressed as

$$\mathbf{s}_{k,u}(t) \triangleq \left(s_{k,u}^1(t)\ s_{k,u}^2(t)\ ...\ s_{k,u}^N(t)\right) = \mathbf{c}_k(t)\mathbf{B}_{k,u}. \tag{5}$$

Following STS, the U_k data streams of each antenna are transmitted simultaneously by modulating U_k different subcarriers, which can be implemented by the operation of IFFT. The frequency spacing Δ between any of the adjacent subcarriers $\{f_u\}_{u=1}^{U}$ satisfies $\Delta = 1/T_c$, guaranteeing the orthogonal subcarrier condition. The selection of which U_k out of U subcarriers are used to transmit data is also determined by our proposed subchannel-allocation algorithms. Denote the U_k subcarrier central frequencies assigned to the kth user by $\{f_{k,u}\}_{u=1}^{U_k}$, the transmitted signal $x_{k,n}(t)$ from the nth transmit antenna of the kth user to the basestation within a block-interval $[\tau_k, \tau_k + NT_b)$ can be expressed by

$$x_{k,n}(t) = \sqrt{\frac{P}{NU}} \sum_{u=1}^{U_k} \sum_{i=1}^{N} b_{k,u}^{i,n} c_k^i(t - \tau_k) e^{j2\pi f_{k,u}(t - \tau_k)} \tag{6}$$

where τ_k represents the transmission delay of the kth user which satisfies $0 \leq \tau_k < T_b$; P denotes the maximum transmission power, which can be achieved when $U_k = U$; the coefficient $\sqrt{P/(NU)}$ indicates that the maximum transmission power is independent of the total numbers of transmit antennas and subcarriers; $b_{k,u}^{i,n}$ is given by Eq. (2) and $c_k^i(t)$ is given by Eq. (4), respectively. Clearly, the larger the number U_k of subcarriers assigned to the kth user, the higher the throughput can be achieved.

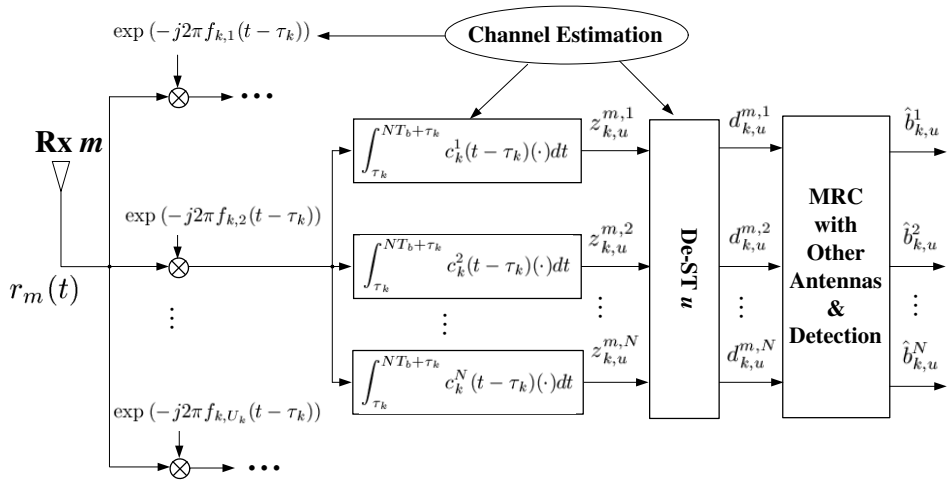


Fig. 2. Uplink data-/signal-flow received by the mth receive antenna at the basestation for the kth user.

B. Uplink Channel Model

We assume that the Rayleigh fading channel is frequency-selective, but the delay-spreads T_m of the channel satisfy $T_m \ll T_c$ such that each subchannel conforms to the flat fading. In addition, the channel is assumed to be *quasi-static*, i.e., the fading coefficients are invariant over a block-interval and vary from one block to another. Thus, during each block-interval, the fading coefficient of the uth subcarrier $h_{k,u}^{n,m}(t)$, between nth transmit antenna of the user and the mth receive antenna of the basestation, can be denoted by $h_{k,u}^{n,m}[i]$, where $i \in \{1, 2, ...\}$ is the discrete time index for the ith block interval. The time-varying channel can be modeled by an auto-regressive (AR) process [9] as follows

$$h_{k,u}^{n,m}[i] = \xi_k h_{k,u}^{n,m}[i-1] + \nu_{k,u}^{n,m}[i] \tag{7}$$

where ξ_k is determined by the kth user's Doppler velocity and $\nu_{k,u}^{n,m}[i]$ is a zero-mean independent identically distributed (i.i.d.) complex-Gaussian variable. In Section II and Section III we will focus on the discussion within a block interval. Therefore, we drop the time index i for convenience. Assume that $\{h_{k,u}^{n,m} \,|\, \forall k, u, m, n\}$ are independent[2] identically distributed (i.i.d.) complex-Gaussian variables with zero-mean and variance of $\sigma/2$ per dimension, i.e., $\sigma = E\left(|h_{k,u}^{n,m}|^2\right)$.

Assuming perfect power control, the received signal $r_m(t)$ received at the mth receive antenna at the basestation is given (see Fig. 2) by

$$\begin{aligned} r_m(t) &= \sqrt{\frac{P}{NU}} \sum_{k\in\Phi} \sum_{n=1}^{N} \sum_{u=1}^{U_k} \sum_{i=1}^{N} h_{k,u}^{n,m} b_{k,u}^{i,n} c_k^i(t-\tau_k) \\ &\quad \cdot e^{j2\pi f_{k,u}(t-\tau_k)} + w_m(t) \end{aligned} \tag{8}$$

where $w_m(t)$ denotes the complex Additive White Gaussian Noise (AWGN) at the mth receive antenna with zero-mean and double-sided power-spectral density of $N_0/2$.

[2]Theoretically, the coefficients between different subcarriers are not independent. However, this independence assumption is valid if we employ frequency-interleaving operations [6][9]. In this paper, we omit the frequency-interleaving to simplify the presentation.

C. Uplink Receiver Model at the Basestation

The schematic for the mth receive antenna at the basestation of our proposed space-time OFDM-CDMA system is shown in Fig. 2, where we focus on the decoding scheme for the kth user within a block interval $[\tau_k, \tau_k + NT_b)$. In addition, we assume that the channel state information (CSI) and the transmission delay of each user can be perfectly estimated by the basestation.

Performing the inverse operation of the transmitter, the received signal $r_m(t)$ at the mth antenna is split into U_k sub-streams by demodulating U_k different subcarriers $\{e^{-j2\pi f_{k,u}(t-\tau_k)}\}_{u=1}^{U_k}$. Then, each sub-stream correlates with the kth user's referenced waveforms $\{c_k^i(t-\tau_k)\}_{i=1}^{N}$ during $[\tau_k, \tau_k + NT_b)$ to obtain correlation outputs $\mathbf{z}_{k,u}^m = (z_{k,u}^{m,1}\ z_{k,u}^{m,2}\ ...\ z_{k,u}^{m,N})^T$. Then, the space-time decoding (De-ST) is employed to obtain N decision variables $\mathbf{d}_{k,u}^m = (d_{k,u}^{m,1}\ d_{k,u}^{m,2}\ ...\ d_{k,u}^{m,N})^T$, corresponding to the original transmitted N bits expressed by $\mathbf{b}_{k,u} = (b_{k,u}^1\ b_{k,u}^2\ ...\ b_{k,u}^N)^T$. Following space-time decoding, all decision variables $\{\mathbf{d}_{k,u}^m\}_{m=1}^{M}$ obtained from M receive antennas are combined together

$$\mathbf{d}_{k,u} = \sum_{m=1}^{M} \mathbf{d}_{k,u}^m, \quad \forall k \in \Phi, \forall u = 1, 2, ..., U_k \tag{9}$$

which represents the procedure of Maximum Ratio Combining (MRC). Based upon decision variables given in Eq. (9), the receiver makes the decisions of the transmitted bits (see Fig. 2) by

$$\hat{\mathbf{b}}_{k,u} \triangleq \left(\hat{b}_{k,u}^1\ \hat{b}_{k,u}^2\ ...\ \hat{b}_{k,u}^N\right)^T = \text{sgn}\left[\text{Re}(\mathbf{d}_{k,u})\right] \tag{10}$$

where $\text{sgn}(\cdot)$ is the signum function and $\text{Re}(\cdot)$ denotes the real part of $(\cdot)$. Denote the index-set of active users allocated in the uth subchannel by Φ_u. It is easy to see that

$$\Phi = \bigcup_{u=1}^{U} \Phi_u. \tag{11}$$

We demonstrate [10] that the SINR of decoding signals for the kth user ($k \in \Phi_u$) at the uth subchannel can be expressed by Eq. (12), which is shown at the bottom of this page, where $\tau_{k,l} = \tau_k - \tau_l$ and $\rho_{k,l}(\tau_{k,l})$ denotes the correlation factor between spreading codes $c_k(t-\tau_k)$ and $c_l(t-\tau_l)$.

III. Subchannel-Allocation Algorithms

In order to design the subchannel-allocation algorithms for data transmissions, we mainly focus on QoS requirements for guaranteeing *transmission reliability*, maximizing *system throughput*, optimizing *proportional bandwidth fairness*, and minimizing *maximum scheduling delay*. In the rest of this section, we will discuss each of these issues in detail.

A. Optimizing the SINR-Threshold

Due to the nature of CDMA technique employed in our system, subchannel can be reused by multiple users. This introduces the following tradeoff. On one hand, in order to increase the system throughput, we need to assign as many users as possible into a single subchannel. The larger the number of users assigned into a subchannel, the higher the system throughput can be achieved. On the other hand, each user experiences co-subchannel interferences from other users within the same subchannel. When the number of users within the same subchannel becomes large, the interferences increase and users' SINRs decrease. As a result, the Bit-Error Rate (BER) at physical layer and corresponding packet-loss/-error rate at higher network-protocol layers increases. To characterize the tradeoff between error rate and throughput, we need to introduce a pre-determined SINR threshold denoted by γ. To ensure the transmission reliability, each user that is assigned with the uth subcarrier, $u \in \{1, 2, ..., U\}$, must satisfy SINR $> \gamma$ requirement, which can be expressed as

$$\text{SINR}_{k,u} \geq \gamma, \quad \forall k \in \Phi_u. \tag{13}$$

In order to guarantee the transmission reliability QoS for data transmissions, the upper-layer protocol must employ some error-control methods, e.g., ARQ based protocol, to recover the data error/loss caused at the physical layer. When we choose a too small SINR threshold γ, the system throughput in terms of the number of users per subcarrier is high, but the number of retransmissions will be increased due to random loss; when we choose a too large SINR threshold γ, the number of retransmissions will be reduced, but the system throughput drops because more subchannel admission requests are rejected. This implies that there is the optimal SINR threshold γ that can maximize the system goodput, which is defined as the rate the packets are transmitted without retransmissions. When satisfying SINR $\geq \gamma$ requirement, the BER P_b of physical layer using BPSK modulation is upper-bounded by

$$P_b \leq Q\left(\sqrt{2\gamma}\right) \tag{14}$$

where $Q(\cdot)$ denotes the Q-function defined by

$$Q(x) \triangleq \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \tag{15}$$

The corresponding packet error rate P_{pkt} is upper-bounded by

$$P_{pkt} \leq 1 - (1 - P_b)^L \tag{16}$$

where L denotes the packet size. When applying Selective-Repeat ARQ, the system goodput $\widetilde{R}$ can be expressed as

$$\widetilde{R} = \overline{R}(1 - P_{pkt}) \tag{17}$$

where $\overline{R}$ denotes the system throughput. Finally, the optimal SINR threshold γ_{opt} can be derived by

$$\gamma_{opt} = \arg\max_{\gamma} \widetilde{R}. \tag{18}$$

B. The Proportional Bandwidth Fairness

For data transmission over wireless networks, different mobile users usually have different bandwidth capacities due to hardware or power facilities and service priorities due to the pricing schemes. This differentiated QoS in bandwidth can be achieved by employing the fairness-design criterion. The fairness problems have been widely studied in literatures [11][12]. In this paper, we develop our algorithms based on the *proportional bandwidth fairness*, where the allocated bandwidth to different users is proportional to their different bandwidth capacities and service priorities. Thus, the users with the higher bandwidth capacities or priorities, which are characterized by their maximum bandwidth capacities $U_k^{\max}$ (the maximum number of subcarriers assigned to the kth user), will receive the larger bandwidth-resource allocations.

Let the kth user's maximum bandwidth requirement be $U_k^{\max}$ ($U_k^{\max} \leq U$). To formally define the fairness factor, we introduce the *aggressive factor* α_k by:

$$\alpha_k \triangleq \frac{U_k}{U_k^{\max}}, \quad k \in \Omega \tag{19}$$

$$\text{SINR}_{k,u} = \frac{\left(\frac{PT_b}{NU}\right)\left(\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{k,u}^{n,m}|^2\right)^2}{\left(\frac{PT_b}{NU}\right)\sum_{m=1}^{M}\sum_{p=1}^{N}\sum_{q=1}^{N}\sum_{l\in\Phi_u, l\neq k}\left[\rho_{k,l}(\tau_{k,l})|h_{k,u}^{p,m}h_{l,u}^{q,m}|^2\right]^2 + \left(\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{k,u}^{n,m}|^2\right)N_0} \tag{12}$$

where $0 \leq \alpha_k \leq 1$. Then, we can define *fairness factor* ϕ [11] by:

$$\phi \triangleq \frac{\left(\sum_{k\in\Omega}\alpha_k\right)^2}{K\sum_{k\in\Omega}\alpha_k^2} \tag{20}$$

where the cardinality of $\|\Omega\| = K$ (the total number of users) and α_k is given by Eq. (19). It is easy to see that $0 \leq \phi \leq 1$. The perfectly fair allocation has the maximum fairness factor $\phi = 1$, which is attained when $\alpha_k = \alpha_j, \forall k \neq j$, i.e., the bandwidth allocated to all users are equally proportional to their maximum bandwidth capacity $U_k^{\max}$. The worst bandwidth-allocation fairness has $\phi = 1/K \to 0$ as $K \to \infty$ when only one user occupies all the bandwidth resources. Obviously, the more the number of inactive users, the lower the bandwidth fairness attained.

Our proposed subchannel-allocation algorithm can be divided into three steps, which are called *allocation-scheduling*, *initial allocation*, and *dynamic allocation*, respectively. In Section III-C and III-D, we first describe our initial and dynamic allocation algorithms. Then, in Section III-E we present our allocation-scheduling scheme. All algorithms are exerted each time the system receives its CSI information.

C. Initial Allocation Algorithm

Let the jth user ($j \in \Omega, j \notin \Phi_u$) be the candidate which attempts to transmit data using the uth subcarrier. Since the basestation knows the information and the statistical characteristics about the channel, we can pre-compute the SINR of decoding signals for each user at the uth subcarrier using Eq. (12). The jth user is admitted to use the uth subcarrier if and only if

$$\begin{cases} \text{SINR}_{j,u} \geq \gamma \\ \text{SINR}_{k,u} \geq \gamma, \quad \forall k \in \Phi_u \end{cases} \tag{21}$$

where $\text{SINR}_{j,u}$ and $\text{SINR}_{k,u}$, as the special cases of Eq. (12), can be expressed by Eqs. (22) and (23), respectively, which are shown at the bottom of this page.

If Eq. (21) is satisfied, the jth user is qualified to be assigned to the uth subcarrier. Then, it becomes a new active user in Φ_u. Otherwise, if any of the SINR in Eq. (21) is lower than the threshold γ, the jth user is rejected to use the uth subcarrier.

Our initial allocation algorithm attempts to "activate" as many users as possible under the constraint that these users' SINR $> \gamma$. Specifically, the U subcarriers are sequentially tested for each user. Once a user is successfully assigned a single subcarrier, it becomes an active user. Then, the basestation stops join-in-testing for this user and starts searching bandwidth for other users. The sequence of testing users is based on our allocation-scheduling scheme, which will be described in Section III-E for details. The users which cannot be assigned with any subcarrier are not allowed to transmit data during a block-interval. But, they will apply for data transmission when the next subchannel-allocation procedure is performed. We define these users within a block-interval as *inactive users*. It is clear that the index set of inactive users is $\Omega\backslash\Phi$. As a result, every active (inactive) user which is accepted (rejected) to transmit data is assigned one (zero) subcarrier during the initial allocation algorithm, which can be expressed as

$$\alpha_k = \frac{U_k}{U_k^{\max}} = \begin{cases} \frac{1}{U_k^{\max}}, & k \in \Phi \quad \text{(active user, } U_k = 1\text{)} \\ 0, & k \in \Omega\backslash\Phi \quad \text{(inactive user, } U_k = 0\text{).} \end{cases} \tag{24}$$

D. Dynamic Allocation Algorithm

The initial allocation algorithm maximizes the number of active users, but it considers neither optimizing the proportional bandwidth fairness nor maximizing the system throughput, which are the goals of our dynamic allocation algorithm. The key idea of our dynamic allocation is to assign all the leftover bandwidth resources after the initial allocation to active users based on the proportional fairness criterion [see Eq. (20)] defined in Section III-B.

Let Θ denote the index set of active users which cannot be assigned with extra bandwidth resources. At the beginning of dynamic allocation algorithm, Θ is initialized as an empty set $\Theta = \emptyset$. If the ith user belongs to Θ ($i \in \Theta \subseteq \Phi$), it implies that:

Case I: The ith user has achieved its maximum bandwidth capacity, i.e., $U_i = U_i^{\max}$; or

$$\text{SINR}_{j,u} = \frac{\left(\frac{PT_b}{NU}\right)\left(\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{j,u}^{n,m}|^2\right)^2}{\left(\frac{PT_b}{NU}\right)\sum_{m=1}^{M}\sum_{p=1}^{N}\sum_{q=1}^{N}\sum_{k\in\Phi_u}\left[\rho_{k,j}(\tau_{k,j})|h_{k,u}^{p,m}h_{j,u}^{q,m}|^2\right]^2 + \left(\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{j,u}^{n,m}|^2\right)N_0} \tag{22}$$

$$\text{SINR}_{k,u} = \frac{\left(\frac{PT_b}{NU}\right)\left(\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{k,u}^{n,m}|^2\right)^2}{\left(\frac{PT_b}{NU}\right)\sum_{m=1}^{M}\sum_{p=1}^{N}\sum_{q=1}^{N}\sum_{l\in\Phi_u\cup\{j\},l\neq k}\left[\rho_{k,l}(\tau_{k,l})|h_{k,u}^{p,m}h_{l,u}^{q,m}|^2\right]^2 + \left(\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{k,u}^{n,m}|^2\right)N_0} \tag{23}$$

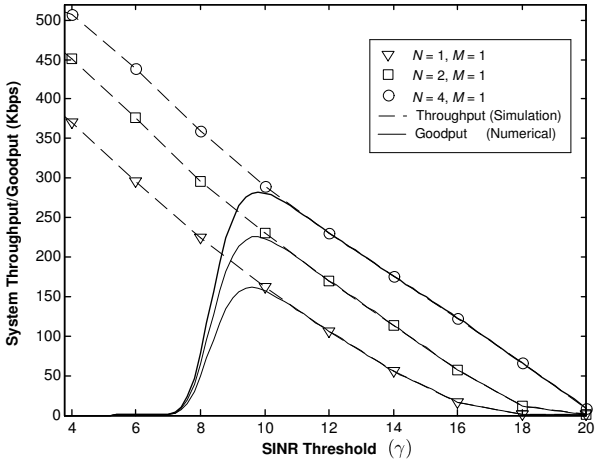


Fig. 3. System throughput $\overline{R}$ and goodput $\widetilde{R}$ versus SINR threshold γ.

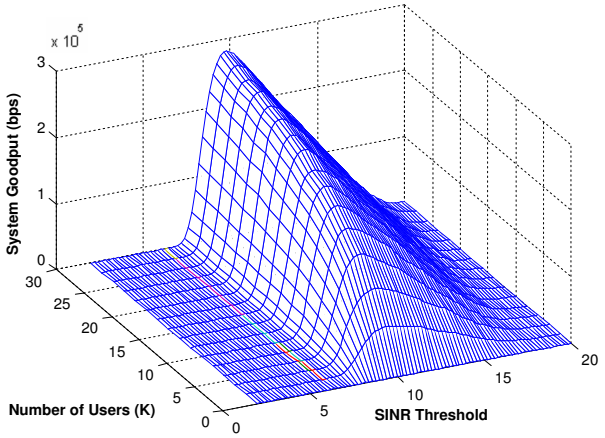


Fig. 4. System goodput $\widetilde{R}$ versus SINR threshold γ and number K of users.

Case II: There is no leftover bandwidth available for the ith user.

The dynamic allocation algorithm searches for the kth user ($k \in \Phi$), which has the *minimum* aggressive factor α_k defined by Eq. (19) and can be assigned with extra bandwidth. The index of this user can be obtained by:

$$k = \arg \min_{k \in \Phi \setminus \Theta} \{\alpha_k\} \tag{25}$$

where Φ is the index set of active users obtained by the initial allocation algorithm.

If the kth user does not achieve its maximum bandwidth capacity, i.e., $U_k < U_k^{\max}$, the algorithm at the basestation attempts to assign an extra subcarrier to it based on SINR criterion given by Eq. (21). The attempt can succeed or fail. If it succeeds, the basestation updates the aggressive factor $\alpha_k$of the kth user and selects the new user which has the minimum α_k using Eq. (25). If it fails, the basestation adds the kth user to Θ (**Case II**).

If the kth user has achieved its maximum bandwidth capacity, i.e., $U_k = U_k^{\max}$, the basestation also adds it to Θ (**Case I**). The dynamic allocation algorithm repeats until $\Theta = \Phi$.

E. Delay-Fairness-Driven Scheduling

Re-examining the subchannel-allocation algorithm described above, one can observe that when the initial allocation is applied, the user which is tested earlier receives the higher priority to become the active user than the later tested ones. This is due to the fact that SINR criterion is easier to be satisfied by the earlier tested users. If we employ a Round-Robin-based scheme testing all K users, the later attempted users are more likely to be rejected for transmitting data in each subchannel-allocation. As the result, their scheduling delays will be longer than the earlier attempted users. This problem gets more serious when the number K of users becomes larger. In the worst case, some users even cannot transmit any data at all. To solve this problem, we propose the delay-fairness driven scheduling scheme to minimize the upper-bound of scheduling delay. Specifically, the scheduling is applied periodically to adjust the sequence of mobile users' attempts. During each interval of scheduling, the scheduling delays of all K users are recorded accumulatively and sorted. The complexity of the scheduling varies, depending on different scheduling algorithms used. The user which experiences larger delay will be tested earlier in initial allocations during the next interval of scheduling. Therefore, the scheduling delays will be fairly shared by all users, and thus this scheduling is called *delay-fairness driven* scheduling.

IV. Simulation and Numerical Evaluations

We investigate the performance of the proposed system infrastructure and subchannel-allocation algorithms by simulations. In the simulations, the chip duration is set to $T_c = 7.8125\mu$s and spreading gain G set to 16. Thus, the bit duration is $T_b = T_c G = 125\mu$s and bit rate per subcarrier is 8Kbps. We set the total number U of subcarriers equal to $U = 8$. Therefore, the total bandwidth of the system is $U/T_c = 1.024$MHz. The packet size is set to $L = 8000$bits and the SNR per bit γ_b is set to 10dB at the basestation. The subchannel-allocation-scheduling is executed every 10ms. The spreading codes are random signature sequences. The correlation factor between different asynchronous random spreading sequences is approximated as a Gaussian random variable with zero-mean and variance of $1/(3G)$ [7]. Doppler frequency is set to 20Hz for all users. Since we mainly focus on transmit diversity in this paper, the number of receive antenna at the basestation is set to $M = 1$ for all simulations.

Fig. 3 plots the simulated system throughput and numerically solved goodput [using Eq. (17)] versus SINR threshold γ. The plots of throughput are obtained through simulation experiments and the results of goodput are based on the numerical solutions for the analyses derived in Section III-A. The total number K of mobile users is set to 15. As expected, the system throughput decreases when SINR threshold γ increases. However, the goodputs have the peak value when γ is around 10dB, which represents the optimal SINR threshold $\gamma_{opt}(\approx 10\text{dB})$ for $K = 15$. We can also observe that the space-time infrastructure can significantly improve the system throughput and goodput, but the optimal SINR threshold

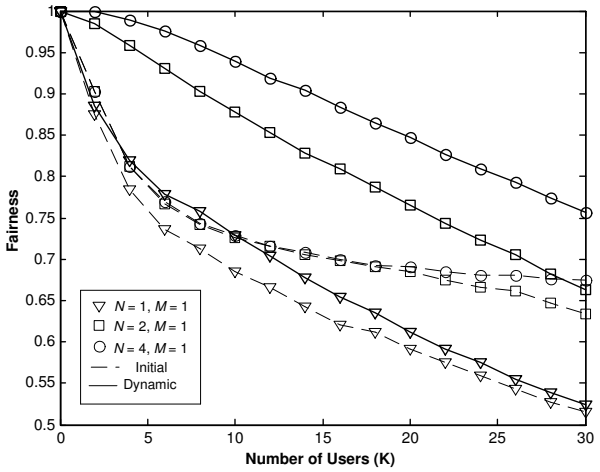


Fig. 5. Fairness ϕ versus number K of users.

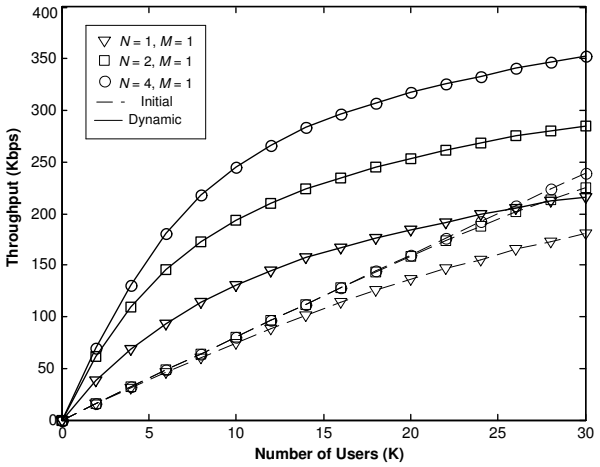


Fig. 6. Throughput $\overline{R}$ versus number K of users.

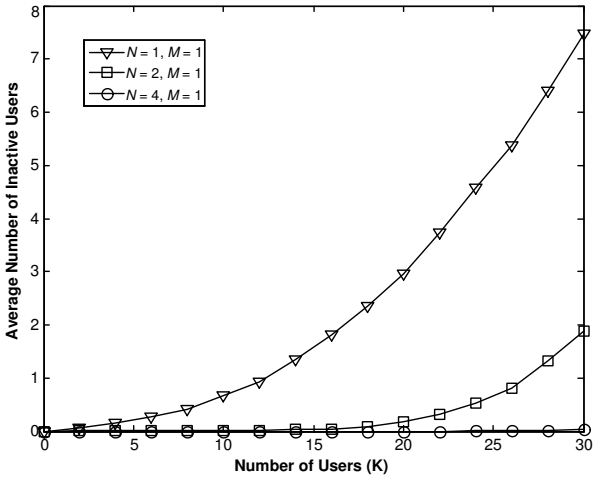


Fig. 7. Average number of inactive users versus total number of users.

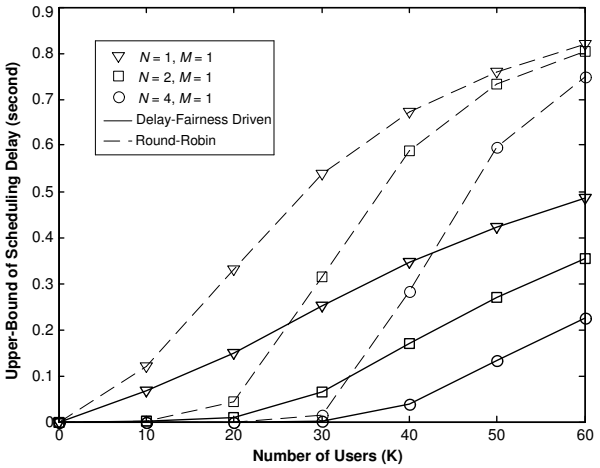
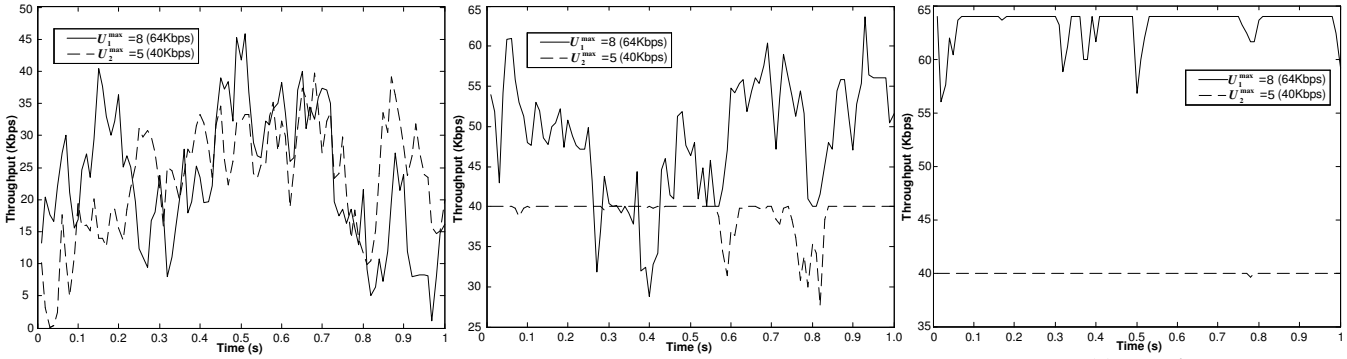


Fig. 8. Maximum delay (second) versus number of users.

$\gamma_{opt} \approx 10$dB virtually does not change regardless of what different antenna combinations are used.

To further investigate the influence of total number K of mobile users on the optimal SINR threshold γ_{opt}, Fig. 4 plots the numerically solved system goodput against two independent variables SINR threshold γ and the total number K of mobile users. The number of transmit antennas is set to $N = 2$. As shown in Fig. 4, the system goodput attains the maximum when $\gamma \approx 10$dB which is virtually independent of the total number K of mobile users. From Fig. 3 and Fig. 4, we can conclude that the optimal SINR threshold γ_{opt} can be pre-determined as a fixed value, regardless of what number of transmit antennas are used and what number of users are employed. Since the maximum goodput can be achieved, and the goodput approaches to the throughput when $\gamma \approx 10$dB, thus, we set the SINR threshold γ to 10dB and only investigate the system throughput ($\approx$ goodput when $\gamma = 10$dB) in the following simulations for convenience.

Applying the initial and dynamic allocation algorithms, the simulated fairness and throughput versus the number K of mobile users with different number of transmit antennas are presented in Fig. 5 and Fig. 6, respectively. The bandwidth capacities $U_k^{\max}$ are randomly chosen between 1 and U. We observe that the bandwidth fairness decreases, while the throughput increases, as K gets larger. This is expected since the larger number of users, the higher spectrum-efficiency of CDMA system, and thus the higher system throughput. On the other hand, the larger number of mobile users makes the bandwidth fairness more difficult to be achieved, decreasing the fairness factor. Figs. 5 and 6 also show that after exerting the initial allocation, algorithm, the dynamic allocation can evidently improve the bandwidth fairness and system throughput. When the number of users is relatively small, there are sufficient leftover bandwidth resources after the initial allocation. Therefore, the dynamic allocation can easily assign these resources to active users based on the fairness criterion. Hence, both fairness and throughput can be significantly improved. When the number of users becomes larger, there are less leftover bandwidth resources after the initial allocation. Thus, the improving rate slows down. Fig. 5 and Fig. 6 also show that the space-time OFDM-CDMA infrastructure at physical layer can significantly increase the fairness and the throughput of the system as the number of transmit antenna increases. For example, both fairness and throughput can be increased more than 40% by using 4 transmit antennas instead of 1 transmit



(a) $N = 1$ (b) $N = 2$ (c) $N = 4$

Fig. 9. Subchannel-allocation experiments for two users with different bandwidth capacities, $U_1^{\max} = 8, U_2^{\max} = 5$.

antenna when 30 mobile users are in the system.

Fig. 7 plots the simulated average number of *inactive* mobile users versus the total number K of users with the different combinations in the number of antennas. From Fig. 7 we find the average number of inactive mobile users increases when K gets larger, which agrees with the observations in Fig. 3 since the more inactive users, the lower fairness attained (see Section III-B). We can also observe that the space-time infrastructure significantly reduce the number of inactive users. For example, for total number of $K = 30$ users, the average number of inactive users approaches to zero when $N = 4$ transmit antennas are employed as compared to the case where the average number of inactive mobile users > 7, when only $N = 1$ transmit antenna is implemented. This verifies that the space-time structure provides higher QoS than the system without the space-time infrastructure.

Fig. 8 investigates the performance of our delay-fairness-driven scheduling scheme by simulations. We can see from Fig. 8 that the delay-fairness-driven scheduling scheme can significantly reduce the upper-bound of the scheduling delay as compared to Round-Robin based scheduling. For instance, the upper-bound of scheduling delay is about 0.23 second using delay-fairness-driven scheduling as compared to 0.73 second using Round-Robin scheduling for $K = 60$ mobile users with $N = 4$ transmit antennas for each mobile user. Also, the more transmit antennas, the lower upper-bound of scheduling delay.

Finally, Fig. 9 plots the simulated subchannel-allocation results for two differentiated users during a period of 1 second with different number of antennas. The bandwidth capacity of the 1st user is $U_1^{\max} = 8$, i.e., 64Kbps. Also, the bandwidth capacity of the 2nd user is $U_2^{\max} = 5$, i.e., 40Kbps. Fig. 9 shows that the space-time infrastructures significantly improve the proportional bandwidth fairness and throughput performance. When $N = 1$, we even cannot distinguish which user has the higher bandwidth requirement as shown in Fig. 9(a). When $N = 2$, Fig. 9(b) shows that the user which has the smaller $U_k^{\max}$ nearly achieves its bandwidth capacity but the user with larger $U_k^{\max}$ still starves for bandwidth resources. When $N = 4$, as illustrated by Fig. 9(c), both users closely approach to their bandwidth capacities. Thus, Fig. 9 shows that the space-time infrastructure offers a great deal of advantages for supporting differentiated applications.

V. Conclusions

We proposed and analyzed the subchannel-allocation algorithms based upon Quality of Service requirements of maximizing fairness and throughput while minimizing upper-bound of scheduling delay for the space-time OFDM-CDMA-based wireless networks. Also, we conducted extensive simulation experiments to evaluate the performance of the proposed algorithms. Our simulation results show that the proposed algorithms can significantly improve the proportional bandwidth fairness and system throughput while reducing scheduling delay. Furthermore, the space-time infrastructure can achieve more efficient bandwidth allocations over finite resources in the wireless networks.

References

[1] T. Rappaport, A. Annamalai, R. Buehrer, and W. Tranter, "Wireless Communications: Past Events and a Future Perspective," *IEEE Comm. Magazine*. May, 2002, pp. 148-161.

[2] D. Gesbert, M. Shafi, D. Shiu, P. Smith, and A. Naguib, "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems," *IEEE J. Select Areas Comm*., vol. 21, no. 3, April 2003, pp. 281-302.

[3] B. Hochwald, T. Marzetta, and C. Papadias, "A Transmitter Diversity Scheme for Wideband CDMA Systems Based on Space-Time Spreading," *IEEE J. Select Areas Comm*., Vol. 19, Jan. 2001, pp. 48-60.

[4] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-Time Block Codes from Orthogonal Designs," *IEEE Trans. Inform. Theory*, Vol. 45, No. 5, Jul. 1999, pp. 1456-1467.

[5] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications", *IEEE J. Select Areas Comm*., vol. 16, no. 8, Oct. 1998, pp. 1451-1458.

[6] L-L Yang and L. Hanzo, "Space-Time Spreading Assisted Broadband MC-CDMA," *IEEE 55th VTC*, Vol. 4, 2002, pp. 1881-1885.

[7] L-L Yang and L. Hanzo, "Performance of Generalized Multicarrier DS-CDMA Over Nakagami-m Fading Channels," *IEEE Trans. Commun*., vol. 50, No. 6, Jun. 2002, pp. 956-966.

[8] J. Li, P. Fan, and Z. Cao, "Space-time Spreading in Forward Links of the Multicarrier DS CDMA System," *Intl. Conf. Info-tech and Info-net*, vol. 2 , Oct. 2001, pp. 285-290.

[9] D. Kalofonos, M. Stojanovic, and J. Proakis, "Performance of Adaptive MC-CDMA Detectors in Rapidly Fading Rayleigh Channels," *IEEE Trans. Wireless Comm*., Vol. 2, No. 2, Mar. 2003, pp. 229-239.

[10] J. Tang and X. Zhang, "Subchannel-Allocation Algorithms and Performance Analysis for Space-Time OFDM-CDMA Based Systems in Wireless Networks," *Technical Report*, Texas A&M University.

[11] D-M Chiu and R. Jain, "Analysis of the Increase and Decrease Algorithms for Congestion Avoidance of Computer Networks," *Computer Networks and ISDN Systems*, 17 (1989), pp. 1-14. [12]

[12] S. Rao and S. Vasudevan, "Resource Allocation and Fairness for Downlink Shared Data Channels," *IEEE WCNC* 2003, vol. 2, March 2003, pp. 1049-1054.