An FDD Wideband CDMA MAC Protocol with Minimum-Power Allocation and GPS-Scheduling for Wireless Wide Area Multimedia Networks

Xudong Wang, *Member*, *IEEE*

Abstract—In this paper, a frequency division duplex (FDD) wideband code division multiple access (CDMA) medium access control (MAC) protocol is developed for wireless wide area multimedia networks. In order to reach the maximum system capacity and guarantee the heterogeneous bit error rates (BERs) of multimedia traffic, a minimum-power allocation algorithm is first derived, where both multicode (MC) and orthogonal variable spreading factor (OVSF) transmissions are assumed. Based on the minimum-power allocation algorithm, a multimedia wideband CDMA generalized processor sharing (GPS) scheduling scheme is proposed. It provides fair queueing to multimedia traffic with different QoS constraints. It also takes into account the limited number of code channels for each user and the variable system capacity due to interference experienced by users in a CDMA network. To control the admission of real-time connections, a connection admission control (CAC) scheme is proposed, in which the effective bandwidth admission region is derived based on the minimum-power allocation algorithm. With the proposed resource management algorithms, the MAC protocol significantly increases system throughput, guarantees BER, and improves QoS metrics of multimedia traffic.

Index Terms—Wideband CDMA, FDD, MAC protocol, CAC, CDMA GPS, minimum-power allocation, BER.

1 Introduction

It is known that there is no single wireless network that can provide global coverage of wireless communication. Instead, various wireless networks, based on technologies that are already deployed or still under development, are vertically and horizontally organized in a hierarchy to constitute the next generation wireless network. Considering such heterogeneity, different MAC protocols are required in various wireless networks. Among the multiple radio transmission technologies for the next generation wireless network, although some are still under investigation, wideband code-division multiple-access (CDMA) has been chosen as the basic access technology [1], [2]. Wideband CDMA can be categorized into pure wideband CDMA and wideband time-division (TD) CDMA. Pure wideband CDMA uses frequency division duplex (FDD) to organize the uplink and downlink transmissions, while wideband TD-CDMA uses time division duplex (TDD). One fundamental property of FDD wideband CDMA is that the downlink output power at the base station is shared by all mobile users [3]. If a mobile user needs more power, all other mobile users will have less power. Considering an indoor and some microcell environments, a mobile user tends to experience very high path loss due to building penetration and wall losses. Thus, indoor or microcell users have a negative effect on the overall system performance of FDD wideband CDMA system. Furthermore, FDD wideband CDMA is well-suited for paired-bands. However, in picocell and microcell environments, the applications generate highly asymmetric traffic. Paired bands in the FDD wideband CDMA system will significantly degrade the spectral efficiency. As a consequence, it is advantageous to apply the FDD wideband CDMA to macrocell environments.

In this paper, an FDD wideband CDMA MAC protocol, which is a type of *two-time-scale* resource allocation schemes [4], is proposed for wireless wide area multimedia networks. At the fast time-scale, channel is assumed fixed and minimum power level of each code channel is determined by satisfying the given target signal-to-noise-interference-ratios (SINRs) of multimedia traffic. Thus, power allocation considered here is different from the linear receiver CDMA power control solutions [5]. In order to simplify the problem of how the interference-limited system capacity affects packet scheduling for multimedia traffic, received (instead of transmit) power level is considered in the minimum-power allocation algorithm. How to derive the transmit power level from the received power level is not discussed in this paper, but this can be accomplished through a closed-loop power control algorithm [6]. As a consequence, the minimum-power allocation algorithm derived here is different from those power control algorithms surveyed in [4]. However, the *monotonicity* law [4] can still be applied to the minimum-power allocation algorithm. Furthermore, rather than using multislot operation [8], a user of FDD wideband CDMA networks can have multiple codes, each with orthogonal variable spreading

- *The author is with the Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332. E-mail: wxudong@ece.gatech.edu.*

Manuscript received 27 Jan. 2003; revised 6 Aug. 2003; accepted 29 Dec. 2003; published online 1 Dec. 2004.
For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-0001-0103.

factors (OVSF). Thus, the minimum-power allocation algorithm considers both multicode (MC) and OVSF operations, which distinguishes it from other power allocation schemes [4], [7].

At the slow time scale, a wideband CDMA generalized processor sharing (GPS) scheduling scheme is proposed based on the minimum-power allocation algorithm. A similar concept is proposed in [9] for hybrid CDMA/TDMA networks. However, due to the flexibility of resource units in FDD wideband CDMA, i.e., both MC and OVSF operations, the new GPS scheduling scheme is much more difficult to derive and also involves an issue of code allocation to a user. Additionally, the limited number of codes for a user must be considered in the new scheduling scheme.

To enhance the performance of the MAC protocol, a CAC algorithm is derived. The concept of effective bandwidth is adopted in order to make the CAC algorithm applicable to bursty traffic. However, the system capacity of wideband CDMA is variable due to interference, so the effective bandwidth-based admission region must be derived on the basis of the minimum-power allocation algorithm, which is different from the algorithm in [10]. A simple call admission rule is suggested in [9], but it does not consider the situation of bursty traffic. Besides, this rule cannot be applied to a CDMA networks with MC operation. Power control is connected with CAC via a notion of active link protection in [11], [12]. However, the system model does not include either MC or OVSF operation. In addition, the CAC algorithm does not address the issue of bursty traffic.

A few MAC protocols have been proposed for FDD wideband CDMA networks. An uplink MC CDMA system architecture is proposed in [13] to support heterogeneous traffic with diverse QoS requirements. The power allocation algorithm does not consider both MC and OVSF operations. In addition, the scheduling scheme for non-real-time traffic is based on first-in first-out (FIFO) queueing and round-robin queueing. In [14], a proposal for an RLC/MAC protocol for wideband CDMA is presented. How to allocate resources to different services is not considered in this proposal. In [15], the power allocation algorithm for the bucket regulator only considers the OVSF operation. Furthermore, how the capacity estimation algorithm is applied to the token bucket traffic regulator is not investigated. For UMTS/IMT-2000 based on wideband CDMA, the performance of a multiple access protocol for integration of variable bit rate multimedia traffic is analyzed in [16], where a packet reservation multiple access (PRMA)-like MAC protocol is assumed. However, power control, which is important to resource management of a CDMA network, is not considered in [16]. Although the third Generation Partnership Project (3GPP) standardization committee has released a specification on the MAC protocol for wideband CDMA [17], development of the resource allocation algorithms for the MAC protocol for wideband CDMA is still an open problem. The access scheme developed in [18] is only applicable to the voice/data traffic transmissions over the common packet channel (CPCH).

The paper is organized as follows: The overall MAC protocol is described in Section 2. In Section 3, the minimum-power allocation algorithm is derived for wideband CDMA system. The multimedia wideband CDMA GPS scheduling scheme is developed in Section 4, and the CAC scheme is presented in Section 5. The overall MAC protocol is evaluated through simulations in Section 6. It is also compared with another FDD mode CDMA MAC protocol in Section 7. The paper is concluded in Section 8.

2 The Wideband CDMA MAC Protocol

In this paper, an FDD mode wideband CDMA system is considered. The following transport channels are used in the MAC protocol: 1) *Random access channel (RACH)* is used by mobile terminals to send control packets, 2) *Broadcast control channel (BCCH)* conveys system information from the base station to mobile terminals, and 3) *Dedicated channel (DCH)* is a point-to-point channel used to transmit data from mobile terminals to the base station or vice versa. The different transport channels are multiplexed in the code division. A DCH can have variable transmission rates depending on the spreading factor, and the basic transmission rate of the DCH corresponds to the maximum spreading factor used in this channel. As a consequence, a variable-length packet is accommodated in a DCH channel. In order to simplify the segmentation of packets from the link access control (LAC) layer to the MAC layer, a fixed-length packet, called radio link control (RLC) packet data unit (PDU) as defined in [1], [14], is used. The size of the fixed-length RLC packet l_r relates to the basic transmission rate r_b of a DCH according to $l_r = r_b \cdot t_{fr}$, where t_{fr} is the frame length. When a packet is generated in the LAC layer of a mobile terminal, it is segmented into multiple RLC PDUs. These RLC PDUs may be transmitted in one MAC frame or several MAC frames, depending on the number of DCHs available for the mobile terminal and the transmission rates of these DCHs. In order to reduce the complexity of a mobile terminal, the number of DCHs that can be used by the mobile terminal is limited. The limited number varies with the service type. For example, a mobile terminal transmitting video traffic needs to have several DCHs, while a mobile terminal transmitting voice traffic is satisfied with one DCH.

2.1 The MAC Protocol

The operation procedures of the MAC protocol is shown in Fig. 1, where only uplink transmission is depicted. The downlink transmission has a similar but simpler procedure, thanks to the broadcast nature of downlink. The MAC protocol in Fig. 1 supports both real-time and non-real-time services. When a mobile terminal wants to support real-time service, it needs to send a *connection request* in the RACH. Once this request is received at the base station, an effective-bandwidth CAC scheme, which is based on minimum-power allocation, is used to check the admission of the connection request. If the answer is positive, the connection request is accepted and the terminal is ready to transmit real-time traffic. However, how the packets of this connection are transmitted in each frame is determined by the wideband CDMA GPS scheduling scheme. When a mobile terminal wants to deliver non-real-time service, no admission control is used. Whenever packets become available in this terminal, they are ready to be transmitted as long as the resources have been allocated by the base station.

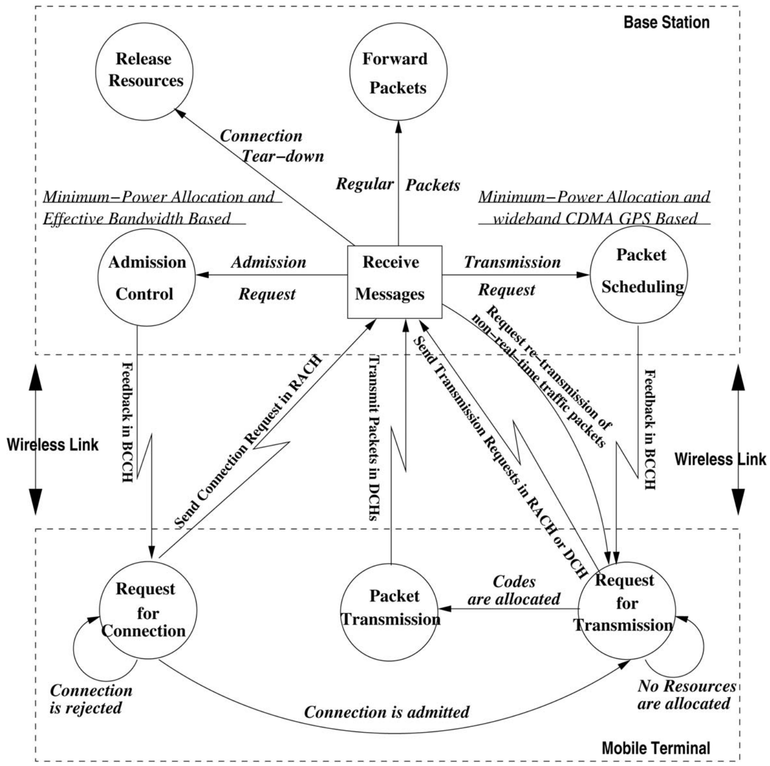


Fig. 1. The operation procedures of the MAC protocol.

Packet transmission in a real-time connection and a non-real-time traffic flow follows the same procedure, as shown in Fig. 1. In a mobile terminal, when a packet in the network layer is generated, its virtual arrival time and virtual departure time are determined by the wideband CDMA GPS scheduling scheme, as will be proposed in Section 4. The priority of packet transmission is determined according to the virtual departure time. The packets in all traffic flows are scheduled for transmission from the highest priority towards the lowest one, until the capacity is exhausted or no packet waits for scheduling. Via the minimum-power allocation based wideband CDMA GPS scheduling scheme, the interference-sensitive CDMA system capacity reaches its maximum value. After scheduling is finished in a frame, the allocated OVSF codes and their received power levels for each traffic flow have been determined. They are sent back from the base station to mobile terminals through the BCCH. Such feedback information is the major cause of overhead in the scheduling scheme. However, such overhead is limited in BCCH and does not affect DCHs because only codes and power levels need to be transmitted frame by frame. After the feedback information is received by mobile terminals, the transmitted power level of a DCH is determined based on the received power level and estimated channel gain. Packets are then transmitted in a DCH by using allocated OVSF and desired transmitted power level. When packets are received at the base station, if errors are detected and cannot be corrected in non-real-time traffic, selective-repeat ARQ is used to retransmit the erroneous packets.

2.2 Compatibility with Standard Systems

When the proposed FDD wideband CDMA MAC is applied to a standard system such as UMTS, additional issues need to be considered:

1. *Synchronization of uplink transmissions*. FDD wideband CDMA in UMTS is an asynchronous network. Thus, uplink transmissions in different cells are not synchronized, which causes larger intercell interference than that in a synchronized network. However, the uplink transmissions of different users in the same cell are still synchronized frame by frame.
2. *Associated control channels for DCHs*. In UMTS, one control channel is required for DCHs of a user in order to maintain these connections. In order to minimize interference from associated control channels, the power allocation algorithm for DCHs must also minimize power levels in control channels.
3. *Constraint on MC operation*. MC operation in UMTS is permitted only for DCHs with maximum transmission rate.

In this paper, we do not intend to propose a MAC protocol fully compatible with UMTS systems. We assume the effect of increased interference from intercell asynchronized transmissions as well as from control channels are neglected. We also assume MC operation is permitted for all DCHs as long as the *utilization* and *complexity* criteria [8] are satisfied. How to make the MAC protocol fully compatible with UMTS specifications is subject to future research.

3 Minimum-Power Allocation Algorithm for Multimedia Traffic

As described in Section 2.1, both the effective-bandwidth-based CAC scheme and the wideband CDMA GPS scheduling scheme are based on the minimum-power allocation algorithm. The objective of this algorithm is, given a number of code channels of different users with heterogeneous

BER requirements, to find the minimum received power level of each code channel such that the heterogeneous BER values of different users are satisfied. Compared to other related work [15], [7], [19], the algorithm considers all the following features of FDD wideband CDMA: 1) Multiple service types with heterogeneous BER requirements are supported and 2) both MC operation and OVSF operation are taken into account when allocating resources. However, MC operation is not considered in [15], [7], while OVSF operation is not included in [19], where only one service type is supported.

In order to guarantee the target BER, the received power level must be controlled at a target value. This includes two tasks when multimedia traffic is supported. First, the target received power level of each service type is determined so that required SINRs of all service types are satisfied (the relationship between SINR and BER is established in the Appendix). Second, a closed or open-loop power control algorithm [6] is used to maintain the target received power level. How to design a closed or open-loop power control algorithm is an independent problem from the MAC protocol.

To derive the minimum-power allocation algorithm, we consider a cell in FDD mode wideband CDMA wireless networks. We assume that there are K type of services supported in the cell and that service type k requires SINR to be γ_k in order to have desired BER. Of service type k, it is assumed that there are N_k users. The set of DCHs allocated to user n_k is denoted by a vector $C^{n_k} = [C_1^{n_k}, \cdots, C_M^{n_k}]$, which must be chosen from an OVSF code tree with M levels of orthogonal codes, and the SF of mth level is $G_m = 2^{m-1}, m = 1, 2, \cdots, M$. Thus, the transmission rate of the a DCH using a code at mth level is $r_m = W/G_m$, where W is the bandwidth of the wideband CDMA system. In addition, the overall transmission rate of user n_k is $r_{n_k} = \sum_{m=1}^{M} C_m^{n_k} \frac{W}{G_m}$. $P^{n_k} = [P_1^{n_k}, \cdots, P_M^{n_k}]$ denotes the received power levels that corresponds to DCHs of C^{n_k}.

Considering a specific user n_ζ with service type ζ, one of its DCHs at the νth level of the OVSF code tree experiences interference $I_\nu^{n_\zeta}$ at the receiver of the base station. $I_\nu^{n_\zeta}$ consists of two components: One is the interference from DCHs of other users in the same system, denoted by I_I, and the other is the noise, denoted by N_o. Thus, the SINR of one of the νth level DCHs, denoted by $\gamma_\nu^{n_\zeta}$, can be described as

$$\gamma_\nu^{n_\zeta} = \frac{P_\nu^{n_\zeta}/r_\nu}{(I_I + N_o)/W}, \tag{1}$$

where $\nu \in \{1, \cdots, M\}$, $\zeta \in \{1, \cdots, K\}$, $n_\zeta \in \{1, \cdots, N_\zeta\}$, and r_ν is the transmission rate of a DCH at the νth level. Interference I_I is contributed by the power levels of DCHs of all users except those of user n_ζ, so

$$I_I = \underbrace{\sum_{k=1}^{K}\sum_{n_k=1}^{N_k}\sum_{m=1}^{M} C_m^{n_k} P_m^{n_k}}_{\textit{power levels of all users}} - \underbrace{\sum_{l=1}^{M} C_l^{n_\zeta} P_l^{n_\zeta}}_{\textit{power levels of user } n_\zeta}. \tag{2}$$

Combining (1) and (2) and considering $\gamma_\nu^{n_\zeta} \geq \gamma_\zeta$ yield

$$\frac{P_\nu^{n_\zeta} G_\nu}{\sum_{k=1}^{K}\sum_{n_k=1}^{N_k}\sum_{m=1}^{M} C_m^{n_k} P_m^{n_k} - \sum_{l=1}^{M} C_l^{n_\zeta} P_l^{n_\zeta} + N_o} \geq \gamma_\zeta, \tag{3}$$

where $G_\nu = W/r_\nu$. To minimize the power levels of each DCH, the equality in (3) must hold. This yields

$$P_\nu^{n_\zeta} \frac{G_\nu}{\gamma_\zeta} = \sum_{k=1}^{K}\sum_{n_k=1}^{N_k}\sum_{m=1}^{M} C_m^{n_k} P_m^{n_k} - \sum_{l=1}^{M} C_l^{n_\zeta} P_l^{n_\zeta} + N_o. \tag{4}$$

Note that (4) is satisfied for all $\nu \in \{1, \cdots, M\}$ of user n_ζ. Thus, for the first level DCHs of user n_ζ, the right side of (4) is the same as that for the νth level channel. Thus,

$$P_\nu^{n_\zeta} G_\nu = P_1^{n_\zeta} G_1, \tag{5}$$

i.e., $P_\nu^{n_\zeta} = P_1^{n_\zeta} \frac{G_1}{G_\nu}$. According to this equation, all power levels in (4) can be represented by the power level $P_1^{n_\zeta}$ or $P_1^{n_k}$. Defining Γ_{n_k} as $\sum_{m=1}^{M} \frac{G_1}{G_m} C_m^{n_k}$ and Γ_{n_ζ} as $\sum_{l=1}^{M} \frac{G_1}{G_l} C_l^{n_\zeta}$, and from (5), (4) becomes

$$\left(\frac{G_1}{\gamma_\zeta} + \Gamma_{n_\zeta}\right) P_1^{n_\zeta} = \sum_{k=1}^{K}\sum_{n_k=1}^{N_k} \Gamma_{n_k} P_1^{n_k} + N_o. \tag{6}$$

It should be noted that (6) is also satisfied for any user n_k, i.e., the left side of (6) can be $\left(\frac{G_1}{\gamma_k} + \Gamma_{n_k}\right) P_1^{n_k}$. Thus,

$$\left(\frac{G_1}{\gamma_\zeta} + \Gamma_{n_\zeta}\right) P_1^{n_\zeta} = \left(\frac{G_1}{\gamma_k} + \Gamma_{n_k}\right) P_1^{n_k}, \tag{7}$$

i.e., $P_1^{n_k} = \frac{\frac{G_1}{\gamma_\zeta} + \Gamma_{n_\zeta}}{\frac{G_1}{\gamma_k} + \Gamma_{n_k}} P_1^{n_\zeta}$. Putting this into (6) yields

$$P_1^{n_\zeta} = \frac{N_o}{\left(\frac{G_1}{\gamma_\zeta} + \Gamma_{n_\zeta}\right)\left(1 - \sum_{k=1}^{K}\sum_{n_k=1}^{N_k} \frac{\Gamma_{n_k}}{\frac{G_1}{\gamma_k} + \Gamma_{n_k}}\right)}. \tag{8}$$

According to the definition of Γ_{n_k} and $r_{n_k} = \sum_{m=1}^{M} C_m^{n_k} \frac{W}{G_m}$, $\Gamma_{n_k} = G_1 \frac{r_{n_k}}{W}$ and $\frac{\Gamma_{n_k}}{\frac{G_1}{\gamma_k} + \Gamma_{n_k}} = \frac{1}{1 + \frac{W}{\gamma_k r_{n_k}}}$. Putting these results back into (8), then

$$P_1^{n_\zeta} = \frac{N_o/G_1}{\left(\frac{1}{\gamma_\zeta} + \frac{r_{n_\zeta}}{W}\right)\left(1 - \sum_{k=1}^{K}\sum_{n_k=1}^{N_k} \frac{1}{1 + \frac{W}{\gamma_k r_{n_k}}}\right)}. \tag{9}$$

Equation (9) gives the minimum-required power level of one of the first level DCHs of user n_ζ. For DCHs on other levels of the OVSF code tree of user n_ζ, this equation also holds, i.e.,

$$P_m^{n_\zeta} = \frac{N_o/G_m}{\left(\frac{1}{\gamma_\zeta} + \frac{r_{n_\zeta}}{W}\right)\left(1 - \sum_{k=1}^{K}\sum_{n_k=1}^{N_k} \frac{1}{1 + \frac{W}{\gamma_k r_{n_k}}}\right)}, \tag{10}$$

as long as an mth level channels is allocated to user n_ζ. Since each DCH has power constraint, it is required that the received power level of one of the mth level DCHs be less than P_m^{max}. Thus,

$$1 - \sum_{k=1}^{K}\sum_{n_k=1}^{N_k} \frac{1}{1 + \frac{W}{\gamma_k r_{n_k}}} \geq \frac{N_o/G_m}{P_m^{max}\left(\frac{1}{\gamma_\zeta} + \frac{r_{n_\zeta}}{W}\right)}$$

for all $m \in \{1, \cdots, M\}$ and $\zeta \in \{1, \cdots, K\}$. Thus,

$$\sum_{k=1}^{K}\sum_{n_k=1}^{N_k} \frac{1}{1 + \frac{W}{\gamma_k r_{n_k}}} \leq 1 - \max_{\substack{m=1,\cdots,M \\ \zeta=1,\cdots,K}} \frac{N_o/G_m}{P_m^{max}\left(\frac{1}{\gamma_\zeta} + \frac{r_{n_\zeta}}{W}\right)}. \tag{11}$$

If

$$\Delta \equiv \max_{\substack{m=1,\cdots,M \\ \zeta=1,\cdots,K}} \frac{N_o/G_m}{P_m^{max}\left(\frac{1}{\gamma_\zeta}+\frac{r_{n_\zeta}}{W}\right)}, \tag{12}$$

(11) becomes

$$\sum_{k=1}^{K}\sum_{n_k=1}^{N_k}\frac{1}{1+\frac{W}{\gamma_k r_{n_k}}} \leq 1-\Delta, \tag{13}$$

which must be satisfied in order to have minimum-power allocation for each DCH and satisfy BER of each user in a wideband CDMA system.

When multiple cells are considered, the intercell interference exists in (2). However, this does not change the formula of the minimum-power allocation algorithm, except that Δ in (12) will be increased due to the contribution of the intercell interference.

In (13),

$$\frac{1}{1+\frac{W}{\gamma_k r_{n_k}}}$$

can be viewed as the *normalized transmission rate* of user n_k whose service type is k, transmission rate in a frame is r_{n_k}, and required SINR is γ_k. Thus, (13) means that the overall normalized transmission rates of all users in a frame cannot exceed $1-\Delta$, which is called the *normalized system capacity*. Since the constraint in (13) guarantees minimum power allocation for each user, the interference among users is minimized. Thus, an efficient packet scheduling scheme needs to be developed based on the constraint in (13).

In a wireless network, bandwidth tends to be variable due to fading and user mobility. Such variable bandwidth can also be captured by the normalized capacity $1-\Delta$. The reason is as follows: Given a maximum allowable transmitted power level, the receiver power level P_m^{max} in (12) is variable with fading and user mobility, which in turn causes Δ variable. However, the variable Δ does not impact resource management schemes that will be derived in Sections 4 and 5 because none of them needs a fixed value of Δ all the time. They just assume that Δ is fixed during one MAC frame. Actually, this assumption is practical because the channel gain during one MAC frame is a fixed snapshot value, although it varies from one MAC frame to another. This technique has been widely used [20], [21].

4 GPS Scheduling Scheme for Wideband CDMA

When packets in a frame are available for transmission, a scheme is necessary to schedule the packets of different users with heterogeneous QoS and BER requirements. Such a scheduling scheme must satisfy several constraints in order to achieve high performance in the FDD mode wideband CDMA system:

1. *Power/BER constraint*. When packets are transmitted in a wideband CDMA frame, they must satisfy the BER requirement and have minimum-power allocation in order to achieve maximum capacity. Therefore, the constraint in (13) must be satisfied in the scheduling scheme.
2. *QoS Requirement*. The scheduling scheme must support heterogeneous QoS requirements of multimedia traffic. Thus, the scheduling scheme must be fair and provide QoS guarantees to multimedia traffic.
3. *Code channel constraint*. Not all DCHs on the OVSF code tree can be used simultaneously by a user. One reason is that the DCHs used simultaneously must be orthogonal. The other reason is that the number of DCHs available to a user is generally limited, which reduces the complexity of a mobile terminal by reducing the transceiver units [22]. However, since scrambling codes are user-specific [1], code blocking existing in the downlink [22] does not occur in the uplink of wideband CDMA systems.

Having these constraints in mind, we propose a new scheduling scheme, called wideband CDMA GPS scheduling, in this section. The significant feature of GPS is that it treats various traffic types differently according to their QoS requirements [23]. GPS also assumes that multiple traffic flows with variable traffic rates can be served simultaneously. This was considered as a drawback of GPS because the classical packet-based systems are TDMA-based and, thus, do not permit parallel packet transmissions. However, in a CDMA system, it is natural to simultaneously serve multiple traffic flows with variable transmission rates. Thus, GPS helps to design a high performance scheduling scheme for CDMA systems. In [9], a GPS-based scheduling scheme is proposed for hybrid CDMA/TDMA systems. The scheme in [9] assumes that only OVSF transmission is used. The scheduling scheme under this assumption is easy to derive because the power/BER constraint has a simpler form and no code channel constraint exists. However, in FDD mode wideband CDMA systems, both MC and OVSF transmissions need to be considered. Thus, the wideband CDMA GPS scheduling scheme designed here is much more general and flexible to support multimedia traffic.

The FDD mode wideband CDMA does not have a TDMA frame, so scheduling schemes proposed in [24] cannot be adopted.

The wideband CDMA GPS scheduling scheme is operated as follows:

1. *Determine the virtual finishing time of MAC packets*. When a LAC PDU of a mobile terminal arrives, the base station should be informed of the arrival time of this packet. The virtual finishing time of this LAC PDU is determined by the base station. Since an LAC PDU is segmented into several RLC PDUs, RLC PDUs belonging to the same LAC PDU have the same virtual finishing time.
2. *Serve RLC PDUs according to virtual finishing times*. To provide QoS guarantees for multimedia traffic, the smaller the virtual finishing time of RLC PDUs, the higher the priority. Moreover, two constraints need to be considered.

- *Check system capacity*. The power constraint in (13) must be checked. As long as this constraint is satisfied, capacity is still available.
- *Check code channel constraint*. Since each mobile terminal has a limited number of DCHs, an algorithm is necessary to select appropriate DCHs and transmit as many packets as possible by using these DCHs.

3. *Calculate received power levels*. After the code channels has been assigned to each mobile terminal, the received power level of each assigned code channel is calculated according to (10).

4.1 Determining Virtual Finishing Time

As in Section 3, a cell in a FDD mode wideband CDMA network can be considered as a queueing system with capacity $1-\Delta$. Denote $\Omega_i(t_1,t_2)$ as the amount of traffic of session i served in the time interval $(t_1,t_2]$, and $\omega_i(t)$ as the work rate of a session, i.e., $\omega_i(t)=\frac{d}{dt}\Omega_i(0,t)$. ϕ_i is a positive number associated with session i. According to the definition of GPS and its work conserving characteristics [23], the GPS for a cell of FDD mode wideband CDMA network must have the following two features: 1) the work rate of each backlogged session i is guaranteed to be

$$\Omega_i=\frac{\phi_i}{\sum_{j\in A}\phi_j}(1-\Delta), \tag{14}$$

where A is the set of all the accepted sessions in the system; 2) fair resource sharing is guaranteed, i.e., for any two backlogged sessions i and j, $\frac{\omega_i(t)}{\omega_j(t)}=\frac{\phi_i}{\phi_j}$. Thus,

$$\omega_i(t)=\frac{\phi_i}{\sum_{j\in\beta(t)}\phi_j}(1-\Delta), \forall i\in\beta(t), \tag{15}$$

where $\beta(t)$ is the set of all backlogged sessions.

For the mth LAC PDU of session n_k, assume it arrives at $a^m_{n_k}$, starts service at $S^m_{n_k}$, and finishes service at $d^m_{n_k}$. According to the definition of *normalized transmission rate* in Section 3, $\omega_{n_k}(t)=\frac{1}{1+\frac{W}{\gamma_k r_{n_k}}}$ during $(S^m_{n_k},d^m_{n_k}]$. Combining this result with (15) when $i=n_k$, then $\frac{1}{1+\frac{W}{\gamma_k r_{n_k}}}=\frac{\phi_{n_k}}{\sum_{j\in\beta(t)}\phi_j}(1-\Delta)$, i.e.,

$$\begin{aligned} r_{n_k}&=\frac{\frac{W}{\gamma_k}}{\frac{\sum_{j\in\beta(t)}\phi_j}{\phi_{n_k}(1-\Delta)}-1},\\ &=\frac{W}{\gamma_k}\frac{\phi_{n_k}(1-\Delta)}{\sum_{j\in\beta(t)}f_j\phi_j}, \end{aligned} \tag{16}$$

where $n_k\in\beta(t)$ and $f_j=1$ if $j\neq n_k$; otherwise, $f_j=\Delta$.

For any given busy period $(t_1,t_2]$ in GPS, the virtual time $v(t)$ of session i is defined as [25]:

$$v(t_2)-v(t_1)=\frac{\Omega_i(t_2,t_1)}{\Omega_i}, \forall i\in\beta(t_1,t_2), \tag{17}$$

where $v(0)=0$. Thus, the virtual time of the mth LAC PDU of session n_k is

$$v(d^m_{n_k})-v(S^m_{n_k})=\frac{\Omega_{n_k}(S^m_{n_k},d^m_{n_k})}{\Omega_{n_k}}. \tag{18}$$

Considering that $\omega_{n_k}(t)=\frac{1}{1+\frac{W}{\gamma_k r_{n_k}}}$ during $(S^m_{n_k},d^m_{n_k}]$, $\Omega_{n_k}(S^m_{n_k},d^m_{n_k})$ can be described as

$$\begin{aligned} \Omega_{n_k}(S^m_{n_k},d^m_{n_k})&=\int_{S^m_{n_k}}^{d^m_{n_k}}\omega_{n_k}(t)dt,\\ &=\frac{d^m_{n_k}-S^m_{n_k}}{1+\frac{W}{\gamma_k r_{n_k}}},\\ &=\frac{L^m_{n_k}}{r_{n_k}+\frac{W}{\gamma_k}}, \end{aligned} \tag{19}$$

where $L^m_{n_k}=r_{n_k}(d^m_{n_k}-S^m_{n_k})$ is the length of the mth LAC PDU of session n_k. Putting (16) into (19) yields

$$\Omega_{n_k}(S^m_{n_k},d^m_{n_k})=\frac{L^m_{n_k}}{\frac{W}{\gamma_k}\frac{\sum_{j\in\beta(t)}\phi_j}{\sum_{j\in\beta(t)}f_j\phi_j}}. \tag{20}$$

Thus, (18) becomes

$$v(d^m_{n_k})-v(S^m_{n_k})=\frac{L^m_{n_k}}{\frac{W}{\gamma_k}\frac{\sum_{j\in\beta(t)}\phi_j}{\sum_{j\in\beta(t)}f_j\phi_j}\Omega_{n_k}}. \tag{21}$$

From (14), $\Omega_{n_k}=\frac{\phi_{n_k}}{\sum_{j\in A}\phi_j}(1-\Delta)$. Thus, (21) becomes

$$v(d^m_{n_k})-v(S^m_{n_k})=\frac{L^m_{n_k}}{\frac{W}{\gamma_k}\frac{\sum_{j\in\beta(t)}\phi_j}{\sum_{j\in\beta(t)}f_j\phi_j}\frac{\phi_{n_k}(1-\Delta)}{\sum_{j\in A}\phi_j}}. \tag{22}$$

Because $v(S^m_{n_k})$ is defined to be $v(S^m_{n_k})=\max\{v(d^{m-1}_{n_k}), v(a^m_{n_k})\}$ [25], $v(d^m_{n_k})$ is thus derived as

$$v(d^m_{n_k})=\max\{v(d^{m-1}_{n_k}),v(a^m_{n_k})\}+\frac{L^m_{n_k}}{\frac{W}{\gamma_k}\frac{\sum_{j\in\beta(t)}\phi_j}{\sum_{j\in\beta(t)}f_j\phi_j}\frac{\phi_{n_k}(1-\Delta)}{\sum_{j\in A}\phi_j}}, \tag{23}$$

where $v(a^m_{n_k})$ is determined from

$$\frac{dv(t)}{dt}=\frac{\sum_{j\in A}\phi_j}{\sum_{j\in\beta'(t)}\phi_j}.$$

It should be noted that $\beta'(t)$ is the set of all backlogged sessions before the arrival of the mth LAC PDU. Equation (23) determines the virtual finishing time of a LAC PDU once it arrives. Such virtual time is used as a priority for fair queueing of packets. In a wideband CDMA frame, all packets are serviced from the queue with the smallest virtual finishing time toward the one with the largest virtual finishing time, until system capacity is exhausted.

4.2 Checking System Capacity

When packets of different mobile terminals are serviced in a frame, the overall transmission rates cannot exceed the system capacity. Since the system capacity in a CDMA system is interference-related, the capacity must be determined by considering the heterogeneous BER requirements of different mobile terminals. As derived in Section 3, if (13) is satisfied, BER requirements of different mobile terminals can be satisfied by using minimum power levels. Therefore, (13) can be used as a criterion to check if the system capacity is exhausted. The system capacity is available unless the constraint in (13) is violated.

4.3 Checking Code Channel Constraint

In order to decrease the complexity of the transceiver of mobile terminals, the maximum number of DCHs for user n_k, denoted by M_{n_k}, is normally much less than the available codes on the OVSF tree. Thus, in the wideband CDMA GPS scheduling scheme, this constraint must be always checked so that it is not violated.

Given M_{n_k} DCHs for user n_k and an OVSF tree with M levels of codes, the available transmission rates for such a user need to be determined by considering the following two factors: 1) the total number of DCHs is not larger than M_{n_k} and 2) the codes of two DCHs cannot be on the same path to the root of the OVSF tree. One approach to this problem is to check if a transmission rate can be decomposed into $l(l \leq M_k)$ smaller transmission rates such that each of these smaller transmission rates is $2^{i-1}(i = 1, \cdots, M)$ times of the basic transmission rate r_b, i.e., the transmission rate of a DCH using the Mth level code on the OVSF code tree. As an example, the set of available transmission rates when $M_{n_k} = 2$ and $M = 7$ is $\{1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 16, 17, 18, 20, 24, 32, 33, 34, 36, 40, 48, 64\}$, where the transmission rates are in unit of r_b. Such an approach can be used offline to determine the set of available transmission rates. Thus, the computation complexity of wideband CDMA GPS scheduling scheme will be approximately in the same order as that of a classical GPS scheme. This helps to maintain the practicality of the wideband CDMA GPS scheme.

Suppose the number of packets scheduled by the wideband GPS scheduling scheme for user n_k is g_{n_k}. Then, the transmission rate corresponding to g_{n_k} packets must be $g_{n_k} r_b$ since one packet needs a basic transmission rate. If such a transmission rate lies in the set of available transmission rates for user n_k, all the scheduled packets can be transmitted. Otherwise, the maximum transmission rate that is less than $g_{n_k} r_b$ is used. Suppose this allowed transmission rate is $h_{n_k} r_b$, then $(g_{n_k} - h_{n_k})$ packets cannot be served in the current frame due to the constraint of available code channels. However, since some packets are not served, a certain amount of system capacity is not utilized by user n_k. Therefore, under this situation, the wideband CDMA GPS scheduling scheme needs to be performed again until system capacity is fully utilized or all packets are served.

5 The Effective Bandwidth-Based CAC Algorithm

An effective-bandwidth-based CAC algorithm is developed by considering the minimum-power allocation algorithm. In contrast, the CAC algorithms in [10] do not consider minimum-power allocation. In [26], the CAC scheme only applies to OVSF CDMA networks and does not consider bursty traffic. However, the CAC algorithm proposed here is applicable to bursty traffic since effective bandwidth concept is adopted.

5.1 The CAC Algorithm

In each MAC frame of a wideband CDMA system, suppose there are N_k connections of service type k, and the transmission rate of connection n_k is r_{n_k}. According to the minimum-power allocation algorithm, (13) must be satisfied. Assume that

$$R_{n_k} = \frac{1}{1 + \frac{W}{\gamma_k r_{n_k}}}$$

is the *normalized transmission rate* of user n_k. From (13), we have

$$\sum_{k=1}^{K} \sum_{n_k=1}^{N_k} R_{n_k} \leq 1 - \Delta, \tag{24}$$

i.e., the overall *normalized transmission rates* cannot exceed the *normalized system capacity*.

In general, non-real-time connections are accepted without admission control, so the CAC algorithm is only applied to real-time connections. In order to take into account the contribution of non-real-time traffic to the overall *normalized transmission rates*, we reserve a minimum *normalized transmission rate*, denoted by R_{nrt}, for non-real-time traffic. Therefore, for all real-time connections, the following constraint must be satisfied:

$$\sum_{k=1}^{K} \sum_{n_k=1}^{N_k} R_{n_k} \leq 1 - \Delta - R_{nrt}. \tag{25}$$

During the life time of the connection n_k, R_{n_k} is a random variable because r_{n_k} varies from frame to frame. Thus, a satisfaction factor α is used to evaluate the probability that (25) is satisfied, i.e., for $0 < \alpha \leq 1$, if

$$\Pr\left(\sum_{k=1}^{K} \sum_{n_k=1}^{N_k} R_{n_k} \leq 1 - \Delta - R_{nrt}\right) > \alpha, \tag{26}$$

then (25) is satisfied with probability α.

Given a satisfaction factor α, the admission region of real-time connections can be determined based on (26). In what follows, Gaussian approximation [10] is used to derive the admission region.

Of the same service type k, different connections are independent and follow the same traffic characteristics. Considering the connection n_k, denote the mean and variance of R_{n_k} as μ_k and σ_k^2, respectively. According to the central limit theorem, when N_k is large, $\sum_{n_k=1}^{N_k} R_{n_k}$ can be approximated by a Gaussian random variable $\mathcal{G}_k$ with mean and variable equal to $N_k \mu_k$ and $N_k \sigma_k^2$, respectively. Since $\{\mathcal{G}_k, k = 1, \cdots, K\}$ are independent Gaussian random variables, $\sum_{k=1}^{K} \mathcal{G}_k$ can also be approximated by a Gaussian random variable $\mathcal{G}$ whose mean and variance are $\sum_{k=1}^{K} N_k \mu_k$ and $\sum_{k=1}^{K} N_k \sigma_k^2$, respectively. Thus, (26) becomes

$$\Pr(\mathcal{G} < 1 - \Delta - R_{nrt}) > \alpha. \tag{27}$$

According to the characteristics of Gaussian random variable, (27) is satisfied if and only if

$$\frac{1 - \Delta - R_{nrt} - E[\mathcal{G}]}{\sqrt{Var[\mathcal{G}]}} \geq \beta, \tag{28}$$

where $E[\mathcal{G}]$ and $Var[\mathcal{G}]$ are the mean and the variance of $\mathcal{G}$, respectively, and β is defined by

$$\frac{1}{\sqrt{2\pi}}\int_{\beta}^{\infty} e^{-t^2/2}dt = 1-\alpha. \tag{29}$$

With $E[\mathcal{G}] = \sum_{k=1}^{K} N_k\mu_k$ and $Var[\mathcal{G}] = \sum_{k=1}^{K} N_k\sigma_k^2$ and after some algebra, (28) becomes

$$\sum_{k=1}^{K} N_k\mu_k + \beta\sqrt{\sum_{k=1}^{K} N_k\sigma_k^2} \leq 1-\Delta-R_{nrt}, \tag{30}$$

which determines an admission region $(N_1, \cdots, N_K)$. When a new connection arrives, its admission depends on whether or not the new N_k (increased by one) satisfies (30); its effective bandwidth does not need to be explicitly determined.

5.2 Runtime Issues of the CAC Algorithm

In the CAC algorithm proposed in Section 5.1, the mean μ_k and variance σ_k^2 of the *normalized transmission rate* of a new arrival connection is assumed to be known in order to determine whether or not the new connection can be accepted. This assumption is also used in other effective-bandwidth-based CAC algorithms. However, when the CAC algorithm is implemented in a real system, a practical method must be used to determine the values of μ_k and σ_k^2, which is not a trivial task. Here, a scheme is proposed to approximately calculate μ_k and σ_k^2 of service type k.

As discussed in Section 4.3, if the transmission rate of connection n_k is nonzero, it must lie in the set of available transmission rates for connection n_k. Suppose the set of available transmission rates is $\{r_{n_k}^1, \cdots, r_{n_k}^j, \cdots, r_{n_k}^J\}$, then the sample set of the transmission rates is

$$\{r_{n_k}^1, \cdots, r_{n_k}^j, \cdots, r_{n_k}^J, r_{n_k}^{J+1}\},$$

where $r_{n_k}^{J+1} = 0$ representing no transmission in a frame by connection n_k. We denote φ_j as the probability that the transmission rate of user n_k is $r_{n_k}^j$. Thus, $\sum_{j=1}^{J+1}\varphi_j = 1$, and μ_k and σ_k^2 of R_{n_k} are calculated according to

$$\begin{aligned}\mu_k &= \sum_{j=1}^{J} R_{n_k}^j\varphi_j,\\ &= \sum_{j=1}^{J}\frac{1}{1+\frac{W}{\gamma_k r_{n_k}^j}}\varphi_j\end{aligned} \tag{31}$$

and

$$\sigma_k^2 = \sum_{j=1}^{J}\left(\frac{1}{1+\frac{W}{\gamma_k r_{n_k}^j}}\right)^2\varphi_j - \mu_k^2, \tag{32}$$

respectively. Since $r_{n_k}^{J+1} = 0$, it is not included in (31) and (32).

To user n_k, the set of available transmission rates $\{r_{n_k}^1, \cdots, r_{n_k}^j, \cdots, r_{n_k}^J\}$ can be determined by the method in Section 4.3. The probability φ_j corresponding to $r_{n_k}^j$ is related to both the traffic characteristics of user n_k and the wideband CDMA GPS scheduling scheme. To find the accurate value of the probability φ_j is impractical. However, based on experiments, we found that φ_j can be approximated by a discrete gamma random variable with mean and variance equal to the average rate and variance, respectively, of the transmission rate of user n_k. After the probability φ_j is determined, μ_k and σ_k^2 can be easily calculated from (31) and (32), respectively.

6 Performance Evaluation

6.1 Traffic Models

Six types of traffic are considered in the simulation.

1. *Voice*. The duration of a voice connection is exponentially distributed with the average equal to 180.0s. The traffic of each connection follows the traffic model in [27]. The average length of talkspurts and gaps are 1.00s and 1.35s, respectively. When a connection is in the talkspurt, the average length of minispurts and gaps are 0.235s and 0.050s, respectively.
2. *Audio*. The duration of an audio connection is also exponentially distributed with the average equal to 180.0s. The bit rate is 128 kbps.
3. *CBR video*. The duration of a CBR video connection is exponentially distributed with the average equal to 360.0s. The bit rate is 220 kbps.
4. *VBR video*. The duration of a VBR video connection is exponentially distributed with the average equal to 180.0s. The traffic of a connection is simulated according to the model in [28]. Duration of each state of the model is also exponentially distributed with the average equal to 160 msec. The traffic rate in each state is obtained from a truncated exponential distribution with the maximum and minimum bit rates equal to 120 and 420 kbps, respectively.
5. *Computer data*. The length of a computer data message is exponentially distributed with the mean size equal to 30 kbytes.
6. *Email*. An empirical size distribution of email messages is shown in Fig 2. It is obtained after analyzing the size of more than 2,500 email messages [24]. In Fig. 2, the mean size of an email message is 3,387 bytes.

6.2 System Parameters

6.2.1 Input Parameters

System parameters in the simulation include: frame length $t_{fr} = 10$ msec, the level of the OVSF tree $M = 7$, the basic transmission rate $r_b = 19.5$ kbps, the length of a RLC packet $l_r = 195$ bits, system bandwidth $W = 5$ MHz, satisfaction factor $\alpha = 0.990$, and the minimum capacity for non-real-time traffic $R_{nrt} = 0.02$. To capture fading, Δ must be variable with dynamic channel characteristics. However, to simplify experiments and focus on the mechanism of the MAC protocol, we assume that Δ is fixed with a value of 0.0005. In Table 1, parameters related to multimedia traffic are listed, which include time out value of a packet t_{out}, BER values BER, the SINR in dB, the maximum number of codes M_{n_k} for each user, and the relative composition p_c of call arrival rates. In Table 1, N_p is the message size of non-real-time traffic.

6.2.2 Output Parameters

Such parameters include the average packet delay d_p, packet loss ratio l_p, throughput t_r, and call blocking

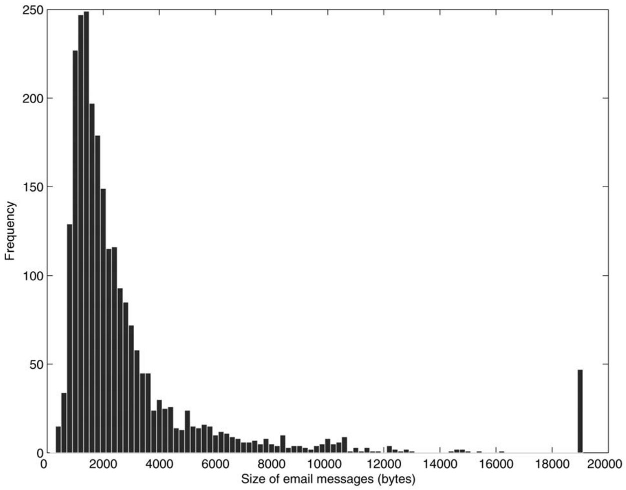


Fig. 2. Size distribution of e-mail messages.

TABLE 1
Parameters of Multimedia Traffic

Service	t_{out}	BER	SINR	M_{n_k}	p_c
Voice	2	10^{-3}	2.54	1	60
CBR Audio	4	10^{-4}	2.72	2	5
CBR Video	4	10^{-5}	2.91	3	6
VBR Video	6	10^{-6}	3.08	3	9
Data	$2N_p$	≈ 0	2.75	1	15
Email	$50N_p$	≈ 0	2.75	1	5

probability b_c. The average packet delay consists of three components, i.e., $d_p = d_r + d_a + d_t$, where d_r is the average time of successfully sending a packet transmission request, d_a is the queueing time before a code channel is allocated to a packet, and d_t is the transmission time of a packet after the code channel is allocated. In the simulation, the code channels in RACH are assumed to have a large number, so the request transmission is collision-free. Thus, the average value of d_r is equal to half a frame because the generation time of a request is uniformly distributed in a frame. Since pure CDMA is used in FDD mode wideband CDMA, a packet is transmitted within a whole frame. Thus, d_t is equal to the frame length. l_p of a connection is defined as $l_p = \frac{N_l}{N_l+N_t}$, where N_l is the number of lost packets due to timeout, and N_t is the number of packets being successfully transmitted. r_t is defined as the total packets being transmitted in a frame. b_c of a service type is defined as the $b_c = \frac{C_b}{C_b+C_a}$, where C_b and C_a are the number of blocked calls and accepted calls, respectively, of a service type.

6.3 Numerical Results

6.3.1 Experiments without CAC

Three performance metrics such as average packet delay, packet loss ratio, and throughput are used. To show the performance versus traffic load, different experiments are performed by varying the call arrival rates. The average packet delay versus the call arrival rate is shown in Fig. 3.

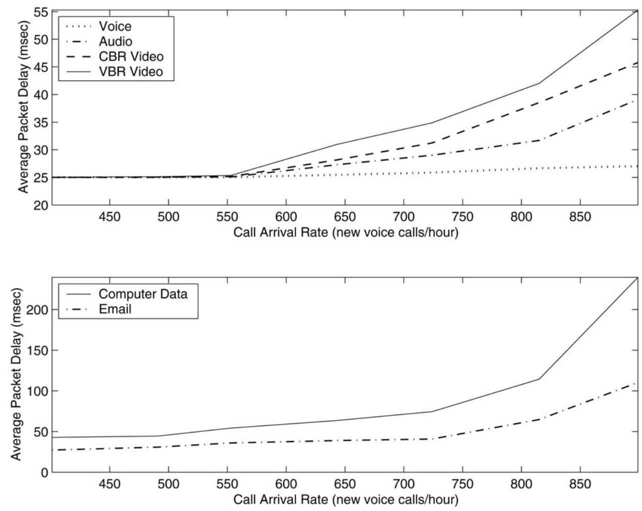


Fig. 3. Average packet delay versus call arrival rate without CAC.

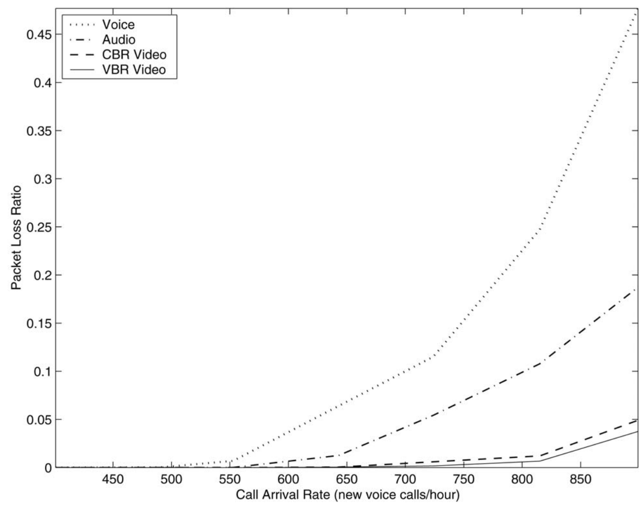


Fig. 4. Packet loss ratio versus call arrival rate without CAC.

The results of packet loss ratio and throughput are shown in Fig. 4 and Fig. 5, respectively.

As shown in Fig. 3, the average packet delay of each service type increases as the traffic load becomes high. The average packet delay of non-real-time traffic is much larger than that of real-time traffic because: 1) the virtual finishing times of non-real-time packets are much larger and 2) no timeout occurs in non-real-time traffic. Among the service types of real-time traffic, voice has the smallest average packet delay. The average packet delay of audio is smaller than that of video traffic because the bit rate of a video connection is larger. VBR video has the largest average packet delay due to its bursty characteristics. Although only average packet delay is shown in Fig. 3, it should be noted that the bound of packet delay of each service type is guaranteed. The reason is that each service type has a certain timeout value and, thus, the delay of each packet cannot exceed this value.

In Fig. 4, the packet loss ratio of voice traffic is much larger than those of other service types, which is reasonable because voice traffic is less sensitive to packet loss than other types of traffic. As shown in Fig. 5, the increment of throughput becomes lower and lower when the traffic load

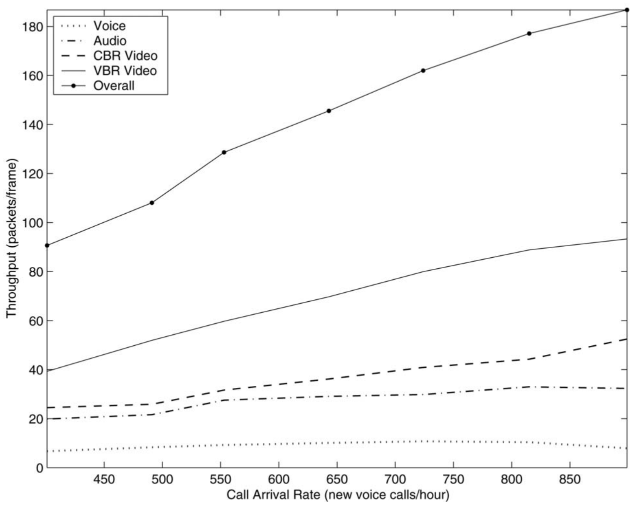


Fig. 5. Throughput versus call arrival rate without CAC.

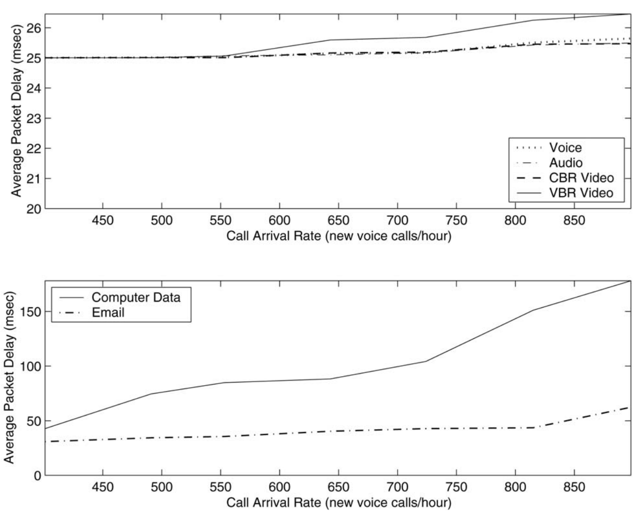


Fig. 6. Average packet delay versus call arrival rate with CAC.

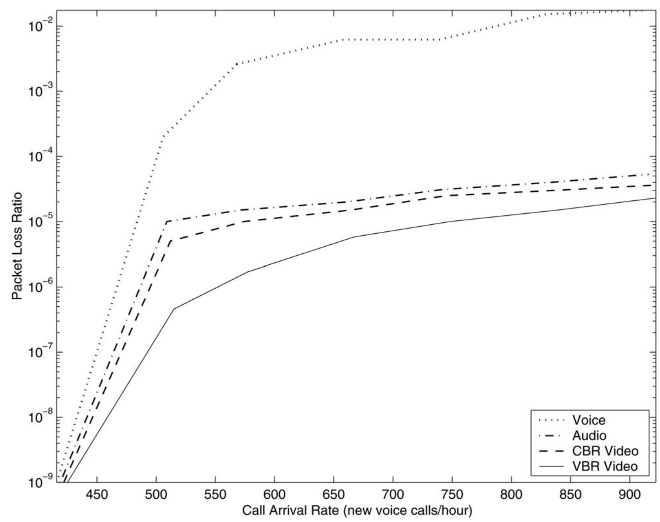


Fig. 7. Packet loss ratio versus call arrival rate with CAC.

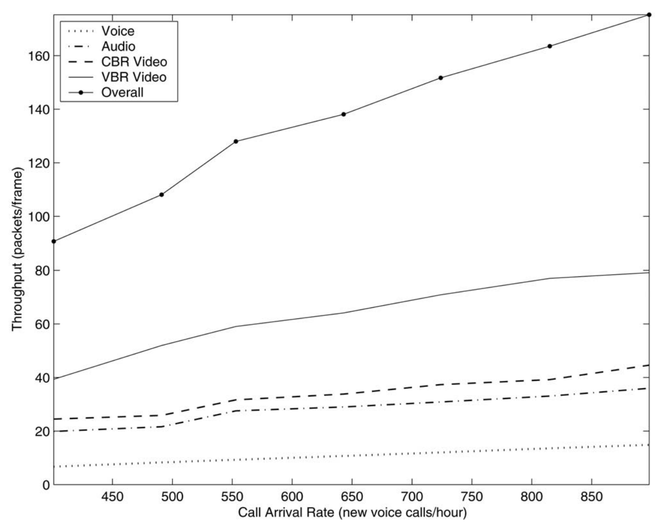


Fig. 8. Throughput versus call arrival rate with CAC.

increases. The reason is that the maximum system capacity is gradually approached as traffic load increases and, thus, the packet loss becomes higher and higher. This can be illustrated by the comparison between the throughput in Fig. 5 and the packet loss ratio in Fig. 4.

When the packet arrival rate is very high, the packet loss ratio of a service type becomes too large to be acceptable, as shown in Fig. 4. In order to resolve this issue, CAC must be used so that some connections are blocked to ensure that the packet arrival rate in the system does not exceed the maximum system capacity.

6.3.2 Experiments with CAC

In this experiment, the CAC scheme proposed in Section 5 is employed to admit connections of real-time traffic. With CAC, the average packet delay, packet loss ratio, and throughput are shown in Figs. 6, 7, and 8, respectively. In addition, the connection blocking probability of each real-time service type is shown in Fig. 9.

As illustrated by the comparison between Figs. 3 and 6, the average packet delay is greatly decreased. For real-time traffic, the average packet delay is almost decreased to 25 msec. Comparisons between the results in Figs. 7 and 4 show that the CAC algorithm greatly reduces packet loss ratio of real-time traffic. The reason for such performance improvement is that CAC ensures that the system capacity is not exceeded. Thus, it is guaranteed that packets of all real-time connections can be transmitted after a short queueing delay. It should be noted that the minimum value of average packet delay is 25 msec because the average packet delay at least consists of half a frame of request delay, one frame of queueing delay, and one frame of transmission delay.

As shown in Fig. 8, the system throughput is lower than that in Fig. 5. The reason is that the rejected connections reduce overall offered load in the system. However, the throughput reduces by less than 10 percent for two orders of magnitude of reduced packet loss ratio. This reflects that the approximation scheme for μ_k and σ_k^2 proposed in Section 5.2 achieves a satisfactory accuracy. Although there may exist a better approximation method to achieve lower throughput reduction without increasing packet loss ratios, to develop such a method is out of the scope of this paper. As shown by the connection blocking probability in Fig. 9,

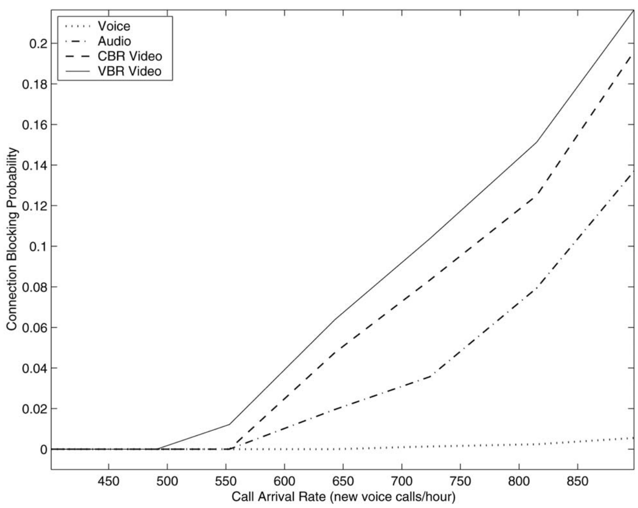


Fig. 9. Connection blocking probability versus call arrival rate.

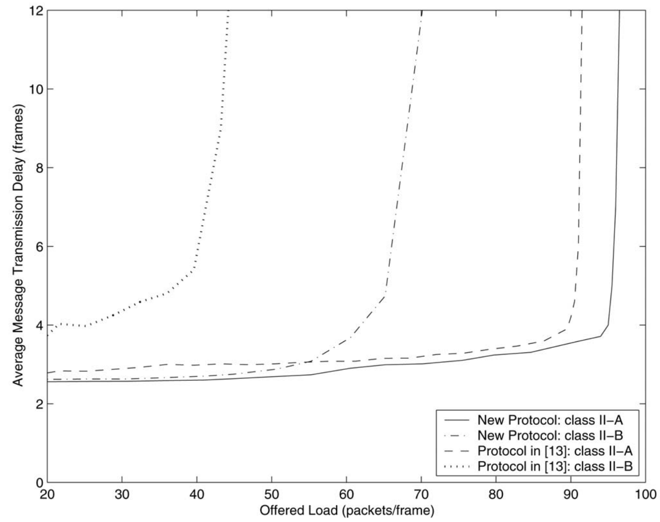


Fig. 10. Average message transmission time versus offered load.

connections with higher traffic rate are easier to be rejected. Thus, the CAC scheme is fair to different service types.

7 Comparisons

The protocol proposed in [13] and the new MAC protocol proposed in this paper have similar features. For example, multimedia traffic with diverse QoS requirements can be supported by both protocols. Moreover, power allocation to a code channel is considered in the CAC and scheduling schemes. However, in the new MAC protocol, the minimum-power allocation algorithm and the wideband CDMA GPS scheduling scheme are used. Therefore, interference-sensitive system capacity can be utilized more efficiently, and packets of different services can be fairly serviced according to their heterogeneous BER and QoS requirements. In order to show such advantages, the new MAC protocol is compared with that in [13].

Assumptions, traffic models, and system parameters are summarized as follows, which are same as those in [13]:

1. There are two types of non-real-time traffic (i.e., the class II traffic in [13]). Class II-A traffic is delay sensitive, while class II-B is delay-tolerable.
2. The number of packets in each message of class II-A is geometrically distributed with an average of 2. Same distribution is used for class II-B, but the average message size is equal to 18 packets.
3. Fifty mobiles are generated for each traffic type. In each mobile, the message generation rate is a Poisson process. The average generation rate of class II-A is $0.9\lambda/50$, while that of class II-B is $0.1\lambda/50$. Thus, each traffic type has the equal load (i.e., 1.8λ), and the total offered load is $0.9\lambda \cdot 2 + 0.1\lambda \cdot 18 = 3.6\lambda$.
4. Fading is not considered in signaling and control channels.
5. The system bandwidth is equal to 128×8 kbps, i.e., 1.024 Mbps.
6. The average message transmission delay consists of three components: request access delay, queueing delay, and transmission delay.
7. A large number of request access code channels are used so that request collision rarely occurs. In the simulation, the number is equal to 25, which achieves almost free collision.
8. Fifty out of 107 bandwidth units are reserved for class II traffic, i.e., bandwidth units (57) for real-time connections do not change throughout the simulation.

In [13], no result of connection blocking probability was reported. In order to have a reference to compare the two protocols, in our experiments we assume no connection blocking occurs for real-time traffic in the protocol proposed in [13]. Moreover, the overall peak rate of all real-time connections are required to be less than 57 by the protocol in [13]. Therefore, the actual traffic load of real-time connections is very low, especially when VBR traffic exists. Due to the peak rate allocation for real-time connections in [13], packet loss ratio is zero and the average packet delay is less than half a MAC frame.

When the same traffic load is applied to our protocol, the CAC in our protocol does not block any real-time connections either, because our protocol better utilize the bandwidth. Moreover, our protocol also achieves a zero packet loss ratio and an average packet delay of less than two and half MAC frames for real-time connections. The two additional MAC frames in our protocol are due to the time of getting a packet transmission permission in the scheduling scheme.

Thus, in our experiments, the two protocols achieve the same performance (except for the two-MAC-frame difference in the average packet delay). In order to avoid repetitive figures, we do not show comparisons of the two protocols for real-time connections, and only the results of class II traffic are compared between the two protocols.

The average message transmission delay versus the offered load is shown in Fig. 10. In our protocol, although a wideband CDMA GPS scheduling scheme is used, the average message transmission delay of two traffic types is different. The reason is that a larger weight (i.e., ϕ_i in (23)) is assigned to class II-A traffic than that to class II-B traffic.

As shown in Fig. 10, for class II-B traffic, the new MAC protocol achieves a lower average message transmission delay. When the traffic load is lower than 40 packets/frame, the difference of the average message transmission delay is approximately between 1.5 and 3.5 frames. However, when the offered load is between 45 and 65 packets/frame, the average message transmission delay achieved by the protocol in [13] is more than 10 frames larger than that of the new protocol. For class II-A traffic, when the offer load is less than 90 packets/frame, the average message transmission delay of the new protocol is a little smaller than that achieved by the protocol in [13]. However, when the offered load is between 90 and 95 packets/frame, the average message transmission delay of the new MAC protocol is more than 10 frames smaller. The main reasons for the better performance achieved by the new protocol are as follows:

1. *Minimum-power allocation algorithms*. The new MAC protocol uses minimum-power allocation for each traffic type. Although the protocol in [13] allocates different power levels to different traffic type, minimum-power allocation is not considered. Therefore, given the same offered load, the new protocol can transmit more packets in a frame.
2. *Fair scheduling schemes*. A wideband CDMA GPS scheduling scheme is used in the new MAC protocol to allocate the code channels to mobile terminals. The transmission order of packets are determined based on their virtual arrival times. Thus, packets of class II-B are not necessarily transmitted later than those of class II-A traffic. However, in [13], higher priority is always given to class II-A traffic. Thus, the message transmission delay of class II-B traffic is large and increases abruptly when the offered load is as low as 45 packets/frame.
3. *Better retransmission mechanism*. In the protocol in [13], although the packets to be retransmitted are put into a FIFO queue and have higher priority than packets in the round-robin queue, packets in the FIFO queue of class II-B traffic is still lower than packets in the FIFO and the round-robin queues of class II-A traffic. Thus, for class II-B traffic, the packets to be retransmitted have a large queueing delay. In our protocol, a packet to be retransmitted participates in scheduling according to its virtual arrival time. Thus, this packet does not need to be transmitted later than a packet of class II-A traffic.

8 Conclusion

In this paper, an FDD wideband CDMA MAC protocol was proposed for wireless wide area multimedia networks. Computer simulations showed that high performance was achieved by the MAC protocol. It is a promising MAC protocol for wireless wide area multimedia networks. Considering the technical heterogeneity of next generation wireless networks, an adaptive MAC protocol may be required in the future in order to support the global roaming of a wireless mobile terminal. However, the resource management schemes proposed in this paper can still be adopted by the adaptive MAC protocol.

The proposed MAC protocol did not aim to exactly match all requirements of UMTS systems and 3GPP specifications. How to tailor it into a protocol that is fully compatible with standard systems is subject to future investigation.

Appendix

In a wireless channel, fading and noise exist. Thus, error control and power control are generally used to satisfy BER of multimedia traffic. Given a BER of a service type, an error control scheme reduces the required SINR of received signals. In this paper, the wireless channel is assumed to have Rayleigh fading and additive Gaussian noise. As used in [13], packets are sequentially encoded with a CRC encoder, an RS encoder, and a convolutional encoder. At the receiver side, for real-time traffic, a convolutional/RS/CRC decoder is used to perform error-correction decoding. However, for non-real-time traffic, a convolutional/RS decoder is used to perform error-correction decoding, and a CRC decoder is used for error detection. When an error cannot be corrected by the convolutional/RS decoder, it will be detected by the CRC decoder. In this situation, a selective-repeat ARQ scheme is used to retransmit erroneous packets. In such a way, a low BER value of non-real-time traffic can be achieved without using very high SINR. In this paper, binary phase shifting keying (BPSK) is assumed as the modulation scheme. Under these assumptions, the relationship between BER and SINR can be derived as follows:

1. *Real-time traffic*. At the receiver, convolutional, RS, and CRC decoders are used sequentially to correct the errors in packets. The bit-error-probability P_e^{rt} at the output of the last decoder (CRC decoder) is [13]:

$$P_e^{rt} = P_b \sum_{j=c+1}^{J} \frac{j}{J} \binom{J}{j} P_s^{j-1} (1 - P_s)^{J-j}, \qquad (33)$$

where P_b is the bit error probability at the output of convolutional decoder in a Rayleigh fading environment and is upper-bounded by

$$\frac{1}{2} \sum_{d=d_{free}}^{\infty} b_d \frac{1}{(1+\gamma_{rt})^d}.$$

γ_{rt} is the SINR of a real-time service type, d_{free} is the free distance of the convolutional code, and b_d is the number of nonzero information bits on all weight-d paths on the trellis code tree of the convolutional code. (J, L, q) is the code used in RS coding, and its error correction probability c is $\lfloor (J-L)/2 \rfloor$. P_s is the symbol error rate at the input of RS decoder and upper-bounded by bP_b, where $b = log_2(q)$.
2. *Non-real-time traffic*. Only the convolutional and RS decoders are used to correct errors in packets. The residual errors are detected by the CRC decoder. Thus, the probability P_r of packet retransmission triggered by the CRC decoder is approximately equal to the bit error probability at the output of the RS decoder, i.e., P_r is given as [13]

$$P_r = \sum_{j=c+1}^{J} \binom{J}{j} P_s^j (1 - P_s)^{J-j}. \qquad (34)$$

Suppose the transmission rate of a non-real-time mobile terminal is r_{nrt} and its allowed normalized transmission rate is R_{nrt}. Thus, according to the definition of *normalized transmission rate* in Section 3,

$$r_{nrt} = \frac{W}{\left(\frac{1}{R_{nrt}} - 1\right)\gamma_{nrt}},$$

where γ_{nrt} is the required SINR of non-real-time traffic. Thus, the throughput ρ of the mobile terminal is $(1 - P_r)r_{nrt}$, i.e.,

$$\rho = (1 - P_r)\frac{W}{\left(\frac{1}{R_{nrt}} - 1\right)\gamma_{nrt}}. \quad (35)$$

P_r increases as γ_{nrt} decreases, so there is an optimum SINR γ^*_{nrt} that achieves maximum throughput of the mobile terminal. Such an optimum SINR is determined by maximizing $(1 - P_r)/\gamma_{nrt}$.

Suppose $1/2$ convolutional coding with $d_{free} = 10$ and $(256, 240, 256)$ RS coding with error correction capability $t = 8$ are used, then the SINR values corresponding to the typical BERs of different services are listed in Table 1.

ACKNOWLEDGMENTS

The author would like to thank Dr. Ian F. Akyildiz for his invaluable advice and suggestions. This work is supported by the State of Georgia YAMACRAW project (E21-105).

REFERENCES

[1] E. Dahlman, P. Beming, J. Knutsson, F. Ovesjöo, M. Persson, and C. Roobol, "WCDMA—The Radio Interface for Future Mobile Multimedia Communications," *IEEE Trans. Vehicular Technology,* vol. 47, no. 4, pp. 1105-1117, Nov. 1998.

[2] M. Haardt, A. Klein, R. Koehn, S. Oestreich, M. Purat, V. Sommer, and T. Ulrich, "The TD-CDMA Based UTRA TDD Mode," *IEEE J. Selected Areas Comm.,* vol. 18, no. 8, pp. 1375-1385, Aug. 2000.

[3] H. Andersson, R.S. Karlsson, P. Larsson, and P. Wikstrom, "Improving System Performance in a WCDMA FDD Network Using Indoor Base Stations," *Proc. IEEE Vehicular Technology Conf.,* vol. 1, pp. 467-471, Oct. 2002.

[4] S.V. Hanly and D. Tse, "Power Control and Capacity of Spread-Spectrum Wireless Networks," *Automatica,* vol. 35, no. 12, pp. 1987-2012, Dec. 1999.

[5] D. Tse and S. Hanly, "Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity," *IEEE Trans. Information Theory,* vol. 45, no. 2, pp. 641-657, Mar. 1999.

[6] L. Song, N.B. Mandayam, and Z. Gajic, "Analysis of an Up/Down Power Control Algorithm for the CDMA Reverse Link under Fading," *IEEE J. Selected Areas Comm.,* vol. 19, no. 2, pp. 277-286, Feb. 2001.

[7] A. Sampath, P.S. Kumar, and J.M. Holtzman, "Power Control and Resource Management for a Multimedia CDMA Wireless System," *Proc. Sixth IEEE Int'l Symp. Personal, Indoor, and Mobile Radio Comm.,* pp. 21-25, 1995.

[8] Third Generation Partnership Project; Technical Specification Group Radio Access Network; Radio Resource Management Strategies (Release 4), 3GPP TR 25.922, v 4.0.0, Mar. 2001.

[9] M.A. Arad and A. Leon-Garcia, "A Generalized Processor Sharing Approach to Time Scheduling in Hybrid CDMA/TDMA," *Proc. IEEE INFOCOM'98 Conf.,* pp. 1164-1171, Apr. 1998.

[10] J.S. Evans and D. Everitt, "Effective Bandwidth-Based Admission Control for Multiservice CDMA Cellular Networks," *IEEE Trans. Vehicular Technology,* vol. 48, no. 1, pp. 36-46, Jan. 1999.

[11] N. Bambos, S.C. Chen, and G.J. Pottie, "Radio Link Admission Algorithms for Wireless Networks with Power Control and Active Link Quality Protection," *Proc. IEEE INFOCOM'95 Conf.,* pp. 97-104, Apr. 1995.

[12] N. Bambos, "Toward Power-Sensitive Network Architecture in Wireless Communications: Concepts, Issues, and Design Aspects," *IEEE Personal Comm.,* vol. 5, no. 3, pp. 50-59, June 1998.

[13] S. Choi and K.G. Shin, "An Uplink CDMA System Architecture with Diverse QoS Guarantees for Heterogeneous Traffic," *IEEE/ACM Trans. Networking,* vol. 7, no. 5, pp. 616-628, Oct. 1999.

[14] C. Roobol, P. Beming, J. Lundsjo, and M. Johansson, "A Proposal for an RLC/MAC Protocol for Wideband CDMA Capable of Handling Real Time and Non Real Time Services," *Proc. IEEE Vehicular Technology Conf.,* pp. 107-111, 1998.

[15] L. Carrasco and G. Femenias, "W-CDMA MAC Protocol for Multimedia Traffic Support," *Proc. IEEE Vehicular Technology Conf.,* pp. 2193-2197, 2000.

[16] R. Fantacci and S. Nannicini, "Multiple Access Protocol for Integration of Variable Bit Rate Multimedia Traffic in UMTS/IMT-2000 Based on Wideband CDMA," *IEEE J. Selected Areas Comm.,* vol. 18, no. 8, pp. 1441-1454, Aug. 2000.

[17] Third Generation Partnership Project; Technical Specification Group Radio Access Network; MAC Protocol Specification (Release 4), 3GPP TR 25.321, v 4.0.0, Mar. 2001.

[18] J.-W. So and D.-H. Cho, "Access Schemes for Integrated Voice/Data Transmissions over Common Packet Channel in 3GPP," *IEEE Comm. Letters,* vol. 5, no. 2, pp. 46-48, Feb. 2001.

[19] Z. Liu, M.J. Karol, M.E. Zarki, and K.Y. Eng, "Channel Access and Interference Issues in Multi-Code DS-CDMA Wireless Packet (ATM) Networks," *ACM/Baltzer Wireless Networks (WINET),* vol. 2, pp. 173-193, 1996.

[20] S.V. Hanly, "An Algorithm for Combined Cell-Site Selection and Power Control to Maximize Cellular Spread Spectrum Capacity," *IEEE J. Selected Areas Comm.,* vol. 13, no. 7, pp. 1332-1340, Sept. 1995.

[21] H. Haas and S. McLaughlin, "A Dynamic Channel Assignment Algorithm for a Hybrid TDMA/CDMA-TDD Interface Using the Novel TS-Opposing Technique," *IEEE J. Selected Areas Comm.,* vol. 19, no. 10, pp. 1831-1846, Oct. 2001.

[22] T. Minn and K.-Y. Siu, "Dynamic Assignment of Orthogonal Variable-Spreading-Factor Codes in W-CDMA," *IEEE J. Selected Areas Comm.,* vol. 18, no. 8, pp. 1429-1440, Aug. 2000.

[23] A.K. Parekh and R.G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case," *IEEE/ACM Trans. Networking,* vol. 1, no. 3, pp. 344-357, June 1993.

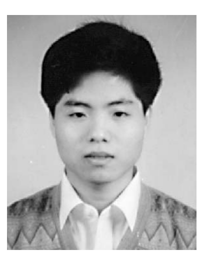
[24] I.F. Akyildiz, D.A. Levine, and I. Joe, "A Slotted CDMA Protocol with BER Scheduling for Wireless Multimedia Networks," *IEEE/ACM Trans. Networking,* vol. 7, no. 2, pp. 1-13, Apr. 1999.

[25] S.J. Golestani, "A Self-Clocked Fair Queueing for Broadband Applications," *Proc. INFOCOM'94 Conf.,* pp. 636-646, 1994.

[26] T.-H. Lee and J.T. Wang, "Admission Control for Variable Spreading Gain CDMA Wireless Packet Networks," *IEEE Trans. Vehicular Technology,* vol. 49, no. 2, pp. 565-575, Mar. 2000.

[27] D.J. Goodman and S.X. Wei, "Efficiency of Packet Reservation Multiple Access," *IEEE Trans. Vehicular Technology,* vol. 37, pp. 302-312, May 1991.

[28] M. Decina and T. Toniatti, "Bandwidth Allocation and Selective Discarding for a Variable Bit Rate Video and Bursty Data Calls in ATM Networks," *Int'l J. Digital Analog Comm. Systems,* vol. 5, no. 2, pp. 85-96, Apr.-June 1992.



Xudong Wang received the BE degree in electric engineering and the PhD degree in automation from Shanghai Jiao Tong University, Shanghai, China, in 1992 and 1997, respectively. He also received the PhD degree in electrical and computer engineering from the Georgia Institute of Technology in 2003. Since 1998, he has been with the Broadband and Wireless Networking (BWN) Lab at the Georgia Institute of Technology. His research interests include computer networking, bioinspired schemes for networking, cross-layer optimization, software radios, and communication protocols for next generation wireless networks, wireless sensor networks, wireless LANs, and ultra wide-band (UWB) networks. He is a member of the IEEE.