Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2007, Article ID 21808, 13 pages
doi:10.1155/2007/21808

Research Article

Performance of a Two-Level Call Admission Control Scheme for DS-CDMA Wireless Networks

Abraham O. Fapojuwo and Yinggan Huang

Department of Electrical and Computer Engineering, The University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

Received 8 May 2007; Accepted 23 August 2007

Recommended by Sudip Misra

We propose a two-level call admission control (CAC) scheme for direct sequence code division multiple access (DS-CDMA) wireless networks supporting multimedia traffic and evaluate its performance. The first-level admission control assigns higher priority to real-time calls (also referred to as class 0 calls) in gaining access to the system resources. The second level admits nonreal-time calls (or class 1 calls) based on the resources remaining after meeting the resource needs for real-time calls. However, to ensure some minimum level of performance for nonreal-time calls, the scheme reserves some resources for such calls. The proposed two-level CAC scheme utilizes the delay-tolerant characteristic of non-real-time calls by incorporating a queue to temporarily store those that cannot be assigned resources at the time of initial access. We analyze and evaluate the call blocking, outage probability, throughput, and average queuing delay performance of the proposed two-level CAC scheme using Markov chain theory. The analytic results are validated by simulation results. The numerical results show that the proposed two-level CAC scheme provides better performance than the single-level CAC scheme. Based on these results, it is concluded that the proposed two-level CAC scheme serves as a good solution for supporting multimedia applications in DS-CDMA wireless communication systems.

Copyright © 2007 A. O. Fapojuwo and Y. Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Recent years have witnessed a great amount of activity on developing the next-generation wireless networks that are expected to provide a wide range of services, such as voice, data, video, and web traffic at very high data rates. Since the radio spectrum is a very scarce resource, call admission control (CAC) is becoming one of the most important elements of radio resource management. The direct sequence code division multiple access (DS-CDMA) technique is widely used in the second- and third-generation mobile communication systems. The problem of CAC in DS-CDMA multimedia wireless networks is very challenging due to the different quality of service (QoS) requirements of the traffic classes, traffic asymmetry between uplink and downlink and, for a given traffic class, different treatments between handoff and new calls. A signal-to-interference ratio- (SIR-) based CAC scheme is proposed in [1]. In [2], downlink admission control based on the output power level from base stations in a CDMA system is studied. The traffic asymmetry between uplink and downlink of multimedia communication is researched in [3, 4]. In [5–9], multilayer medium access control schemes for wireless multimedia services are proposed. In [4, 10–13], CAC algorithms for serving multiple traffic classes requiring different QoS and multiple transmission rates are considered. In [14], a CAC scheme under imperfect power control is studied. Seeking solution to the CAC problem in wireless networks continues to be of active research interest in academia and industry; the works in [15–18] are examples of recently published works in the literature. Jeon and Jeong [10] proposed a CAC scheme based on SIR measurements for DS-CDMA cellular systems supporting mobile multimedia services. Under the CAC scheme studied in [10], a call is admitted only when the SIR requirements of both the existing and the new calls are guaranteed. This CAC scheme takes into account the different QoS requirements of multiple traffic classes, assigns the total available bandwidth to the uplink and downlink asymmetrically, and guarantees the priority of handoff call requests over the new call requests within a service class. However, the CAC scheme described in [10] only focuses on the admission control at a single level and does not contain any mechanism that takes advantage of the delay-tolerant characteristic of nonreal-time calls. The two-level CAC scheme proposed in this paper is similar to the single-level CAC scheme presented in [10] by using the SIR as the metric for call admission, assigns priority to real-time calls over nonreal-time calls, and accounts for traffic and resource asymmetry in the uplink and downlink. Our work differs from that of [10] in two respects. First, the two-level CAC scheme proposed in this paper accounts for the provisioned physical resources (e.g., channel elements at the base station) in DS-CDMA wireless networks and incorporates queuing of nonreal-time calls (to take advantage of their delay-tolerant characteristic) during physical resource shortage. Second, in the SIR calculation, the shadowing effect is taken into account in addition to the distance-dependent path loss that was only considered in the analysis presented in [10]. Our work is similar to Singh et al.'s work [13] by queuing nonreal-time calls, but it differs from that in [13] by performing CAC analysis for both the uplink and downlink directions, and considering system-level design parameters (e.g., different admission control thresholds for new and handoff real-time and nonreal-time calls, reservation bandwidth for nonreal-time calls, and traffic and resource asymmetry in downlink and uplink) that are of practical interest in actual wireless network deployment and provisioning.

The main contribution of this paper is the proposal of a two-level call admission control scheme for DS-CDMA wireless networks and the evaluation of its performance. The proposed scheme is discussed in the context of two classes of services: real-time and nonreal-time calls. At the first level, real-time calls are always given a higher priority over nonreal-time calls in gaining access to system resources. In addition, within the real-time service class, the handoff calls are given higher priority over new calls. At the second level, the nonreal-time calls are scheduled to transmit on a first-come, first-served basis at the beginning of each CDMA frame (slot) according to the available capacity (i.e., residual capacity) obtained after subtracting the real-time resource requirements from the total resource available. Further, a variable parameter, called reservation capacity, is introduced to guarantee some minimum level of performance for nonreal-time calls. To make use of delay-tolerant characteristic of nonreal-time traffic, a finite queue is used to temporarily store the nonreal-time calls that cannot begin at the time of initiation, due to lack of resources. The proposed two-level CAC scheme is flexible, allowing the total available bandwidth in a cell to be distributed unequally in the uplink and downlink directions to account for traffic asymmetry.

The remainder of this paper is organized as follows. Section 2 presents the proposed two-level call admission control scheme. Section 3 contains performance analysis of the proposed two-level CAC scheme; the analysis outputs are the performance metrics of call blocking and outage probabilities, average throughput for both real-time and nonreal-time calls, and average waiting time of nonreal-time calls in the queue. Section 4 presents numerical results and discussion. Finally, Section 5 concludes the paper.

2. PROPOSED TWO-LEVEL CALL ADMISSION CONTROL SCHEME

Consider a DS-CDMA cellular system offering multimedia services each with different QoS requirements. As it is well known, performance of CDMA-based cellular systems is interference-limited. Hence, in this paper, the SIR is used as the metric for call admission. Specifically, the received energy-per-bit to interference spectral density ratio (E_b/I_0) for a call must be higher than a desired threshold to achieve and maintain the required service quality. To this end, a call request (new or handoff) is admitted only when the received E_b/I_0 for the call and those of all the other active calls (in progress) are above the E_b/I_0 threshold value required for acceptable communication. Without loss of generality, in this paper, multimedia services are classified into real-time and nonreal-time categories. Due to their different QoS requirements, the two classes are given two different treatments in call admission control, resulting in the two-level call admission control (CAC). The proposed two-level CAC relies on two ideas. First, real-time calls are given a higher priority over nonreal-time calls in accessing the system resources (i.e., bandwidth). Second, instead of blocking a nonreal-time call whose resource request cannot be met at time of initiation, a finite queue is introduced to temporarily store the nonreal-time call. The proposed two-level CAC scheme is described quantitatively as follows.

2.1. *Level 1 call admission control for admission of real-time calls*

The call admission control is based on the noise rise condition [5, 19, 20]. Denoting the system bandwidth by W, the noise rise condition can be expressed as

$$L = \sum_{i=1}^{N_0} \Gamma_{0,i} R_{0,i} + \sum_{i=1}^{N_1} \Gamma_{1,i} R_{1,i} \leq W(1-\eta), \tag{1}$$

where L is the aggregate system load, N_0 and N_1 denote the number of real-time (referred to as class 0, henceforth) and nonreal-time (referred to as class 1, henceforth) users supported, respectively, $R_{0,i}$ and $R_{1,i}$ are the transmission rates for the ith class 0 and ith class 1 calls, respectively, $\Gamma_{0,i}$ and $\Gamma_{1,i}$ are the target energy-per-bit to interference spectral density ratio for the ith class 0 and ith class 1 calls, respectively, η is the noise rise coefficient defined as the ratio of the background noise power spectral density to the total (intracell + intercell + background noise) received power density, and $(1-\eta)$ is the loading factor threshold. To guarantee some minimum performance for class 1 calls, some amount of system bandwidth W, denoted by W_{res}, is reserved for class 1 calls. The problem is to determine the number of class 0 calls that can be supported by the remaining bandwidth. Consider first the uplink direction and assume that $\Gamma_{0,i}^u = \Gamma_0^u$, $R_{0,i}^u = R_0^u$ for all class 0 calls, and $\Gamma_{1,i}^u = \Gamma_1^u$, $R_{1,i}^u = R_1^u$ for all class 1 calls. Using (1), we have that N_0^u, the maximum number of class 0 calls supported in the uplink direction when bandwidth W^u_{res} is reserved for class 1 calls, is given by

$$N_0^u = \left\lfloor \frac{W^u(1-\eta^u) - W^u_{\text{res}}}{\alpha_0^u \Gamma_0^u R_0^u} \right\rfloor, \tag{2}$$

where W^u is the total available bandwidth in the uplink and α_0^u represents the uplink activity factor of class 0 calls. The superscript "u" denotes uplink direction and the other notations in (2) are as defined previously. The corresponding expression for N_0^d, the maximum number of class 0 calls supported in the downlink direction, is calculated by

$$N_0^d = \left\lfloor \frac{w^d(1-\eta^d) - (1-\rho)w^d_{\text{res}}}{\alpha_0^d \Gamma_0^d (1-\rho) R_0^d} \right\rfloor, \tag{3}$$

where ρ is the average orthogonality factor for the cell due to multipath and the superscript "d" denotes downlink direction. The overall number of class 0 calls supported in either the downlink or the uplink direction is

$$N_0 = \min(N_0^u, N_0^d). \tag{4}$$

The resource (bandwidth) required to support the N_0 real-time calls is therefore reserved. However, the allocation of resource to class 0 call requests (i.e., both new and handoff calls) is based on SIR call admission criteria described as follows. A class 0 new call request is admitted to the system (i.e., allocated resources) if

$$E_0^x \geq \Gamma_0^x, \tag{5a}$$
$$E_{k,0}^x \geq \Phi_{k,0}^x, \tag{5b}$$

where E_0^x is the received E_b/I_0 in the x direction, $x \in \{\text{uplink}, \text{downlink}\} \equiv \{u, d\}$ for the class 0 new call request, $E_{k,0}^x$ is the received E_b/I_0 in the x direction for an active class k call, $k \in \{0, 1\}$ given that the class 0 new call request is admitted, and $\Phi_{k,0}^x$ is the E_b/I_0 threshold in the x direction that an active class k call uses to control the admission of a class 0 new call request. As done in [10], $\Phi_{k,0}^x = \beta_0^n \Gamma_k^x$, where β_0^n (> 1) is the multiplicative factor that controls the admission of class 0 new call requests. Note that the inequality (5b) is checked for all class k calls that are in progress when the class 0 new call request is made. Similarly, a class 0 handoff call request is admitted to the system if

$$E_0^x \geq \Gamma_0^x, \tag{6a}$$
$$E_{k,0}^x \geq \Omega_{k,0}^x, \tag{6b}$$

where $\Omega_{k,0}^x$ is the E_b/I_0 threshold in the x direction that an active class k call uses to control the admission of a class 0 handoff call request. Also let $\Omega_{k,0}^x = \beta_0^h \Gamma_k^x$ [10], where β_0^h (> 1) is the multiplicative factor that controls the admission of class 0 handoff call requests. The inequality (6b) is checked for all class k calls that are in progress when the class 0 handoff call request is made. From the foregoing observation, a class 0 new call (or class 0 handoff call) is admitted if the inequalities (5a) and (5b) (or inequalities (6a) and (6b)) are satisfied.

2.2. *Level 2 call admission control for admission of nonreal-time calls*

The resource manager assigns resources to service nonreal-time calls based on the residual resources after those supporting real-time calls have been allocated. Using the noise rise equation, the number of nonreal-time calls (i.e., class 1 calls) supported in the uplink and downlink directions is given by

$$\begin{aligned} N_1^u &= \max\left(\left\lfloor \frac{W^u(1-\eta^u) - \alpha_0^u N_0^u \Gamma_0^u R_0^u}{\alpha_1^u \Gamma_1^u R_1^u} \right\rfloor, \left\lfloor \frac{W^u_{\text{res}}}{\alpha_1^u \Gamma_1^u R_1^u} \right\rfloor\right), \\ N_1^d &= \max\left(\left\lfloor \frac{W^d(1-\eta^d) - \alpha_0^d N_0^d \Gamma_0^d (1-\rho) R_0^d}{\alpha_1^d \Gamma_1^d (1-\rho) R_1^d} \right\rfloor, \right. \\ &\qquad \left. \left\lfloor \frac{W^d_{\text{res}}}{\alpha_1^d \Gamma_1^d (1-\rho) R_1^d} \right\rfloor\right). \end{aligned} \tag{7}$$

The overall number of nonreal-time calls supported in either downlink or uplink direction is

$$N_1 = \min(N_1^u, N_1^d). \tag{8}$$

Equation (8) implies that the remaining resources, after subtracting the resources for supporting the N_0 class 0 calls, can handle a maximum of N_1 class 1 calls. The allocation of the remaining resources to the class 1 new and handoff call requests is governed by SIR-based admission criteria. Specifically, a class 1 new call is admitted if

$$\begin{aligned} E_1^x &\geq \Gamma_1^x, \\ E_{k,1}^x &\geq \Phi_{k,1}^x, \end{aligned} \tag{9}$$

where E_1^x, $E_{k,1}^x$, and $\Phi_{k,1}^x$ have similar definitions as the terms in (5a) and (5b), but are now defined with respect to class 1 calls and $\Phi_{k,1}^x = \beta_1^n \Gamma_k^x, \beta_1^n > 1$. Similarly, a class 1 handoff call is admitted if

$$\begin{aligned} E_1^x &\geq \Gamma_1^x, \\ E_{k,1}^x &\geq \Omega_{k,1}^x, \end{aligned} \tag{10}$$

where $\Omega_{k,1}^x = \beta_1^h \Gamma_k^x, \beta_1^h > 1$. If sufficient physical resources are available to support the requested data rate for the class 1 call, the call is made active. However, it is possible for the inequalities (9) for new calls (or (10) for handoff calls) to be met, but the call cannot be made active due to physical resource (e.g., channel elements) shortage. Instead of blocking such class 1 calls, we take advantage of their delay-tolerant characteristic and temporarily store such calls in a queue, to be served at a later time. Note that the queued class 1 calls, even though admitted into the system, do not generate interference to the other active calls since they are not yet allocated resources. At the beginning of every CDMA frame (slot), an attempt is made to serve (i.e., allocate physical resources) to the queued nonreal-time calls, on a first-come, fist-served (FCFS) basis. If the requested physical resources by the head-of-queue call are now available and the inequalities (9) for new call (or (10) for handoff call) are still met, then the scheduler allocates resources to the head-of-queue call and the call then becomes active. A queued class 1 call is removed from the queue once it is allocated resources.

3. PERFORMANCE ANALYSIS

The objective of the analysis is to study the performance of a two-level call admission control that assigns priority of resource access to class 0 calls, reserves bandwidth for class 1 calls and temporarily queue class 1 calls that cannot be served. For analysis, we consider a multicellular DS-CDMA system. We assume that all the cells are homogeneous and in statistical equilibrium. Hence, we perform the analysis with respect to a test user (located in one arbitrary cell) whose transmission is affected by both intracell interference from other users in the same cell as the test user, and intercell interference from the surrounding cells.

3.1. Analysis assumptions

The assumptions made in analysis are listed as follows.

A1. Class 0 and class 1 new calls arrive according to independent Poisson processes with mean call rate Λ_0 and Λ_1 per time unit, respectively. Total mean call arrival rate $\Lambda = \Lambda_0 + \Lambda_1$; the decomposition of Λ into Λ_0 and Λ_1 depends on the assumed call mix in the calculations.

A2. Class 0 and class 1 handoff calls arrive according to independent Poisson processes with mean call rate λ_0 and λ_1 per time unit, respectively.

A3. Duration of a class 0 (class 1) call is exponentially distributed with mean $1/\mu_0$ $(1/\mu_1)$.

A4. Dwell time of a user engaged in a class 0 (class 1) call in a cell is exponentially distributed with mean $1/v_0$ $(1/v_1)$.

A5. Connection time for a class 0 call (or class 1 call) alternates between active and dormant states. The length of active period for a class 0 (or class 1) call is exponentially distributed with mean $1/\zeta_0^x$ $(1/\zeta_1^x), x \in \{u, d\}$, where u and d denote uplink and downlink, respectively. Similarly, the length of dormant period for a class 0 (class 1) call is exponentially distributed with mean $1/\omega_0^x$ $(1/\omega_1^x), x \in \{u, d\}$.

A6. Class 1 calls that cannot be allocated resources at the resource request instant are temporarily stored in a queue of size Q_{pq}.

A7. Patience time for a class 1 call waiting in the queue is exponentially distributed with mean $1/\mu_{pq}$. For class 1 applications such as web page or file download, patience time is interpreted as the maximum time to download a page or a file. Note that the concept of patience time, as used in this paper, is not for the purpose of resource allocation, but instead it serves to prevent too long waiting time for the class 1 calls that are served. A queued class 1 call that has not been served when its patience time expires, therefore, reneges from the queue without receiving service.

A8. The received signal or interference power (in dB unit) is normally distributed with mean determined by the distance-dependent path loss model and standard deviation σ dB.

3.2. System model

Based on assumptions A1–A4, A6, and A7, the system is modeled by a continuous-time, discrete-state, two-dimensional Markov chain whose state transition diagram is shown in Figure 1. Denote the system state by $s = (n_0, n_{\text{tot}})$, where n_0 is the number of class 0 calls in a cell and n_{tot} is the total number of class 1 calls in a cell of which n_1 are active and $n_q = (n_{\text{tot}} - n_1)$ are waiting in the queue. Denote the state space of n_0 by S, where $S = \{0, 1, 2, \ldots, N_0\}$. Within the feasible state space S, any state transition is caused by one of the following events: (1) arrival of a class 0 new call, (2) arrival of a class 0 handoff call from a neighboring cell, (3) departure (i.e., handoff) of an active class 0 call to a neighboring cell, and (4) successful completion of a class 0 call. Similarly, denote the state space of n_{tot} by Ψ, where $\Psi = \{0, 1, 2, \ldots, N_1 + Q_{pq}\}$. Recall that $n_{\text{tot}} = n_1 + n_q$ so that the state space of n_1 and n_q are, respectively, $n_1 \in \{0, 1, \ldots, N_1\}$ and $n_q \in \{0, 1, \ldots, Q_{pq}\}$. Within the feasible state space Ψ, any state transition is caused by one of the following events: (1) arrival of a class 1 new call, (2) arrival of a class 1 handoff call from a neighboring cell, (3) departure (handoff) of a class 1 call to a neighboring cell, and (4) completion of a class 1 call. For event 4 causing state transition in Ψ, note that completion refers to successful or unsuccessful completion where the latter is due to the departure of a class 1 call from the queue when its patience time expires (assumption A7). In Figure 1, the label $Y_0(\cdot, \cdot)$ denotes the transition rate from state n_0 to state (n_0+1), caused by the arrival of a class 0 (new or handoff) call. Similarly, $Y_1(\cdot, \cdot)$ represents the transition rate from state n_{tot} to state $(n_{\text{tot}} + 1)$, caused by the arrival of a class 1 call. Also in Figure 1, the label $X_0(\cdot, \cdot)$ denotes the transition rate from state (n_0+1) to state n_0, caused by the departure of an active class 0 call to a neighboring cell or successful completion of an active class 0 call. For the active class 1 calls, transition from state $(n_{\text{tot}} + 1)$ to state n_{tot} occurs at rate $X_1(\cdot, \cdot)$. The mathematical expressions for the transition rates are

$$\begin{aligned} Y_i(n_0, n_{\text{tot}}) &= q_i^n(n_0, n_{\text{tot}}) + q_i^h(n_0, n_{\text{tot}}), \quad i \in \{0, 1\}, \\ x_i(n_0, n_{\text{tot}}) &= q_i^c(n_0, n_{\text{tot}}) + q_i^b(n_0, n_{\text{tot}}), \quad i \in \{0, 1\}, \end{aligned} \tag{11}$$

where, for class 0 calls, $q_0^n(n_0, n_{\text{tot}})$, $(q_0^h(n_0, n_{\text{tot}}))$ are the transition rates from state n_0 to state $(n_0 + 1)$ due to the arrival of a new (handoff) call, and $q_0^b(n_0, n_{\text{tot}})$, $(q_0^c(n_0, n_{\text{tot}}))$ are the transition rates from state $(n_0 + 1)$ to state n_0 caused by the departure of an active class 0 call to a neighboring cell (successful completion of an active class 0 call). For class 1 calls, $q_1^n(n_0, n_{\text{tot}})$, $(q_1^h(n_0, n_{\text{tot}}))$ are the transition rates from state n_{tot} to state $(n_{\text{tot}} + 1)$ due to the arrival of a new (handoff) call, and $q_1^b(n_0, n_{\text{tot}})$, $(q_1^c(n_0, n_{\text{tot}}))$ are the transition rates from state $(n_{\text{tot}} + 1)$ to state n_{tot} caused by the departure of an active class 1 call to a neighboring cell (successful completion of an active class 1 call). Note that, for class 1 calls, $X_1(\cdot, \cdot)$ only accounts for successful completion, and the unsuccessful completion occurs at rate $q_1^r(\cdot, \cdot)$ (see Figure 1). It now remains to determine the expressions for $q_i^n(n_0, n_{\text{tot}}), q_i^h(n_0, n_{\text{tot}}), q_i^b(n_0, n_{\text{tot}}), q_i^c(n_0, n_{\text{tot}}), i \in \{0, 1\}$, and $q_1^r(n_0, n_{\text{tot}})$ which are presented in what follows.

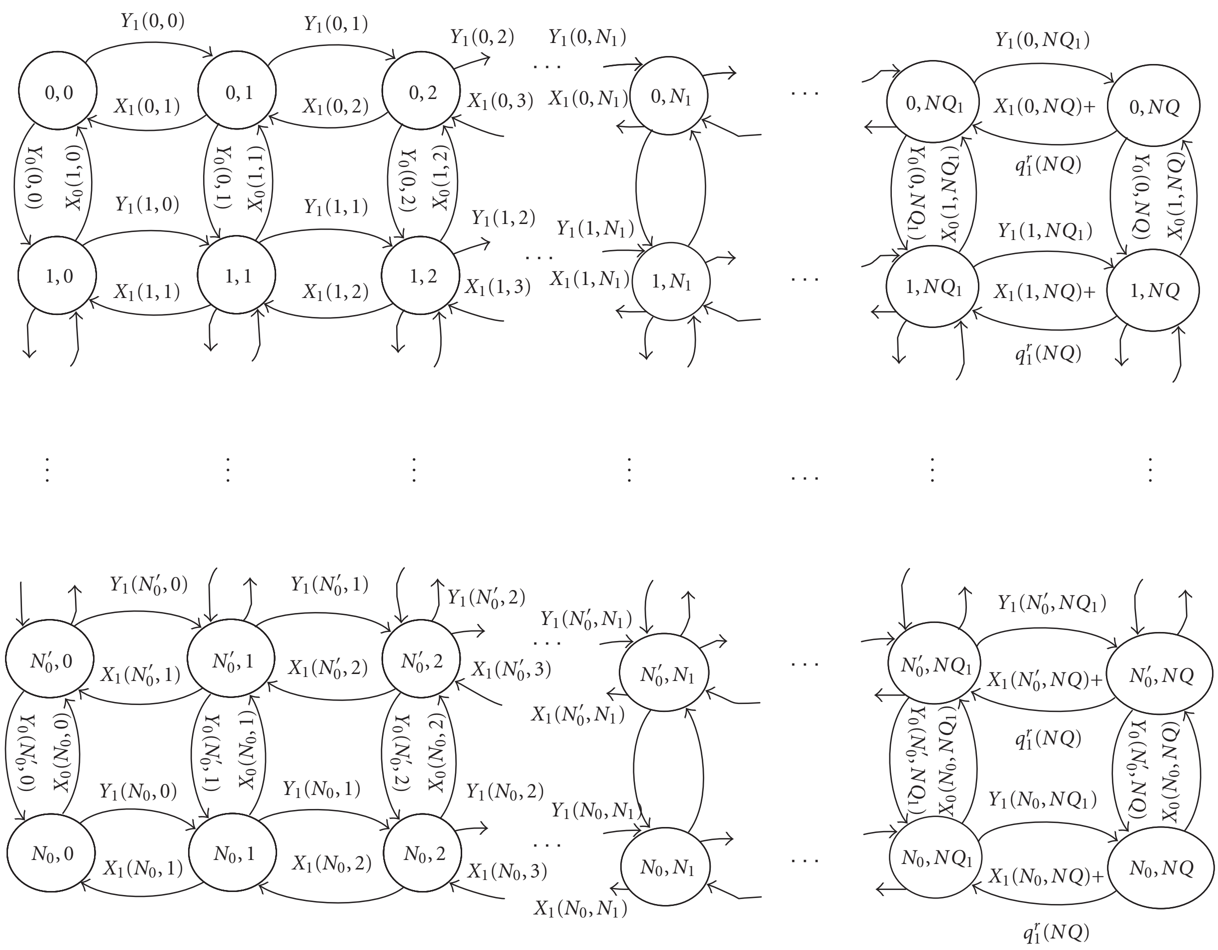

Figure 1: Markov state transition diagram for the system ($NQ_1 : N_1 + Q_{pq} - 1$; $NQ : N_1 + Q_{pq}$; $N_0' : N_0 - 1$).

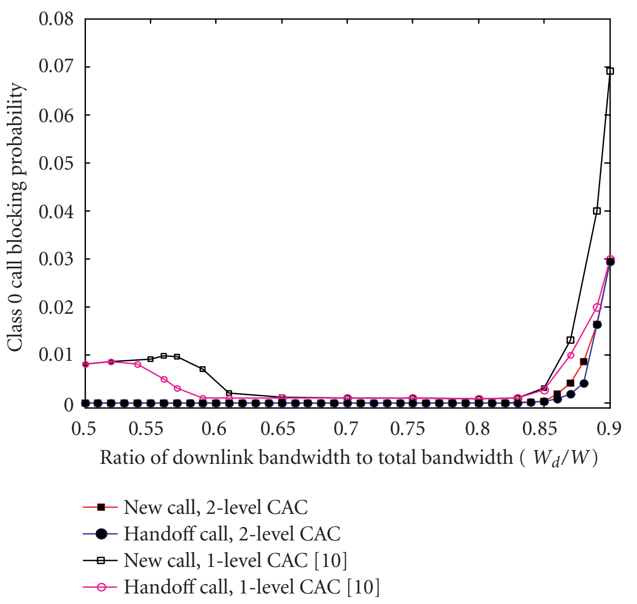

Figure 2: Class 0 call blocking probabilities versus $W_d/W(\Lambda = 0.1, W_{\text{res}}/W = 0.1)$.

3.2.1. *Expression for $q_0^n(n_0, n_{\text{tot}})$*

Let $A_0^n(n_0, n_{\text{tot}})$ be the probability that the BS, in state (n_0, n_{tot}), admits a class 0 new call. Using $A_0^n(n_0, n_{\text{tot}})$ along with assumption A1 gives the expression for $q_0^n(n_0, n_{\text{tot}})$ as

$$q_0^n(n_0, n_{\text{tot}}) = A_0^n(n_0, n_{\text{tot}})\Lambda_0, \tag{12}$$

where $A_0^n(n_0, n_{\text{tot}})$ is determined as follows. Recall from Section 2 that a class 0 new call is admitted if the admission criteria of all the active class $k, k \in \{0, 1\}$, calls in both the uplink and downlink directions (5b) are satisfied. Consequently, we must estimate the mean received E_b/I_0 for both uplink and downlink channels of all active calls. Suppose that when the system is in state $\mathbf{s} = (n_0, n_{\text{tot}})$, there exists an activity state $\phi_s = (u_0, d_0; u_1, d_1)$ that explicitly describes the actual number of active class 0 and class 1 calls in each of the uplink (u_0 and u_1) and downlink (d_0 and d_1) directions. Let $\mathbf{Q}(\mathbf{s})$ be the state space of all feasible activity states of system state $\mathbf{s}$, where $\mathbf{Q}(\mathbf{s}) = \{\phi_s : 0 \leq u_0, d_0 \leq n_0; 0 \leq u_1, d_1 \leq n_1\}$. Note that $0 \leq u_0, d_0 \leq n_0$ and $0 \leq u_1, d_1 \leq n_1$ because not all the n_0 and n_1 calls are active at the same time. Note also that the upper limit for d_1 is n_1 (not n_{tot}) because none of the n_q class 1 calls in the queue is active. Define $\Pr\{\phi_s\}$ as the probability that the activity state is ϕ_s when the system is in state s. Since the downlink and uplink transmissions are independent,

$$\Pr\{\phi_s\} = \left[\binom{n_0}{u_0}(\alpha_0^u)^{u_0}(1-\alpha_0^u)^{n_0-u_0}\binom{n_0}{d_0}(\alpha_0^d)^{d_0}(1-\alpha_0^d)^{n_0-d_0}\right] * \left[\binom{n_1}{u_1}(\alpha_1^u)^{u_1}(1-\alpha_1^u)^{n_1-u_1}\binom{n_1}{d_1}(\alpha_1^d)^{d_1}(1-\alpha_1^d)^{n_1-d_1}\right], \quad (13)$$

where α_j^u and α_j^d are the uplink and downlink activity factors for call class j, $j = \{0, 1\}$. The activity factors are defined by $\alpha_j^x = \omega_j^x/(\omega_j^x + \zeta_j^x)$, $x \in \{u, d\}$. Define $\Pr\{G_0^n(n_0, n_1, \Phi_{k,0}^x)|\phi_s\}$ as the conditional probability that an active class k call allows for the admission of a class 0 new call when the system is in activity state ϕ_s. As such, the notation $\{G_0^n(n_0, n_1, \Phi_{k,0}^x)|\phi_s\}$ denotes the event that $E_{\text{rcv}}^x(\phi_s)$, the received E_b/I_0 when the system is in activity state ϕ_s, exceeds $\Phi_{k,0}^x$. The analysis in [10] computes the admission probability using indicator variables whose values (1 or 0) are based on whether or not $E_{k,0}^x(\phi_s)$, the *average* received E_b/I_0, exceeds the call admission threshold for acceptable communication. In this paper, we compute the admission probability by modeling $E_{\text{rcv}}^x(\phi_s)$, the received E_b/I_0 when the system is in activity state ϕ_s, as a random variable (which follows from assumption A8), where the variation in the received signal and interference power is due to shadowing. The effect of shadowing was not considered in the analysis presented in [10]. Hence, the conditional call admission probability is calculated by

$$\Pr\left\{G_0^n(n_0, n_1, \Phi_{k,0}^x)|\phi_s\right\} = 1 - \Pr\left\{E_{\text{rcv}}^x(\phi_s) < \Phi_{k,0}^x\right\}. \quad (14)$$

From the assumption of log-normal shadowing, $E_{\text{rcv}}^x(\phi_s)$ (in dB unit) is a normal random variable with mean $E_{k,0}^x(\phi_s)$ and standard deviation $\sigma(\phi_s)$, both expressed in dB. Hence,

$$\Pr\{E_{\text{rcv}}^x(\phi_s) < \Phi_{k,0}^x\} = \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{E_{k,0}^x(\phi_s) - \Phi_{k,0}^x}{\sqrt{2}\sigma(\phi_s)}\right)\right], \quad (15)$$

where $\operatorname{erf}(\cdot)$ is the error function. By unconditioning (14) on ϕ_s, making use of (15) and considering all possibilities, we have

$$A_0^n(n_0, n_{\text{tot}}) = \sum_{\phi_s \in \mathbf{Q}(\mathbf{s})} \prod_{k=0}^{1} \prod_{x \in \{u,d\}} \left[0.5 - 0.5 \operatorname{erf}\left(\frac{\Phi_{k,0}^x - E_{k,0}^x(\phi_s)}{\sqrt{2}\sigma(\phi_s)}\right)\right] \Pr\{\phi_s\}, \quad (16)$$

where $\Pr\{\phi_s\}$ is given by (13). Note that in (15) and (16), $E_{k,0}^x(\phi_s)$, the estimated mean E_b/I_0 value in the x direction, now depends on the activity state ϕ_s and is given by

$$E_{k,0}^u(\phi_s) = \left[\frac{1}{M_k^u(\phi_s)} + \frac{R_0^u \Gamma_0^u}{W^u \Gamma_k^u}\right]^{-1}, \quad k = 0, 1,$$
$$E_{k,0}^d(\phi_s) = \left[\frac{1}{M_k^d(\phi_s)} + (1-\rho)\frac{R_0^d \Gamma_0^d}{W^d \Gamma_k^d}\right]^{-1}, \quad k = 0, 1, \quad (17)$$

where $M_k^u(\phi_s)$ and $M_k^d(\phi_s)$ denote the average E_b/I_0 when the system is in activity state ϕ_s in the uplink and downlink, respectively, and they are calculated by

$$M_k^u(\phi_s) = \frac{W^u \Gamma_k^u}{\sum_{j=0}^{1} u_j R_j^u \Gamma_j^u - R_k^u \Gamma_k^u + \xi^u \sum_{j=0}^{1} n_j \alpha_j^u R_j^u \Gamma_j^u}, \quad k = 0, 1, \quad (18)$$

$$M_k^d(\phi_s) = \frac{(1-z) W^d \Gamma_k^d}{(1-\rho)\sum_{j=0}^{1} d_j R_j^d \Gamma_j^d - (1-\rho)(1-z) R_k^d \Gamma_k^d + \xi^d \sum_{j=0}^{1} n_j \alpha_j^d R_j^d \Gamma_j^d}, \quad k = 0, 1. \quad (19)$$

In (19), z is the proportion of the total BS transmission power spent on overhead channels, n_j is the number of class j calls in progress in a cell, and $\xi^u(\xi^d)$ is the ratio of average uplink (downlink) interference from other cells to average uplink (downlink) interference from own cell.

3.2.2. *Expression for* $q_1^n(n_0, n_{\text{tot}})$

Applying the same approach described above, the expression for $q_1^n(n_0, n_{\text{tot}})$ can be written as

$$q_1^n(n_0, n_{\text{tot}}) = A_1^n(n_0, n_{\text{tot}})\Lambda_1, \quad (20)$$

where

$$A_1^n(n_0, n_{\text{tot}}) = \sum_{\phi_s \in \mathbf{Q}(\mathbf{s})} \prod_{k=0}^{1} \prod_{x \in \{u,d\}} \left[0.5 - 0.5 \operatorname{erf}\left(\frac{\Phi_{k,1}^x - E_{k,1}^x(\phi_s)}{\sqrt{2}\sigma(\phi_s)}\right)\right] \Pr\{\phi_s\}. \quad (21)$$

3.2.3. *Expression for* $q_0^h(n_0, n_{\text{tot}})$

Let $A_0^h(n_0, n_{\text{tot}})$ be the probability that the BS, in state (n_0, n_{tot}), admits a class 0 handoff call. Using $A_0^h(n_0, n_{\text{tot}})$ along with assumption A2 gives the expression for $q_0^h(n_0, n_{\text{tot}})$ as

$$q_0^h(n_0, n_{\text{tot}}) = A_0^h(n_0, n_{\text{tot}})\lambda_0. \quad (22)$$

Derivation of the expression for $A_0^h(n_0, n_{\text{tot}})$ follows the same approach as for $A_0^n(n_0, n_{\text{tot}})$, but now using $\Omega_{k,0}^x$. Hence,

$$A_0^h(n_0, n_{\text{tot}}) = \sum_{\phi_s \in \mathbf{Q}(\mathbf{s})} \prod_{k=0}^{1} \prod_{x \in \{u,d\}} \left[0.5 - 0.5 \operatorname{erf}\left(\frac{\Omega_{k,0}^x - E_{k,0}^x(\phi_s)}{\sqrt{2}\sigma(\phi_s)}\right)\right] \Pr\{\phi_s\}. \quad (23)$$

3.2.4. Expression for $q_1^h(n_0, n_{\text{tot}})$

The transition rate $q_1^h(n_0, n_{\text{tot}})$ is given by

$$q_1^h(n_0, n_{\text{tot}}) = A_1^h(n_{\text{tot}})\lambda_1, \tag{24}$$

where

$$A_1^h(n_0, n_{\text{tot}}) = \sum_{\phi_s \in \mathbf{Q}(\mathbf{s})} \prod_{k=0}^{1} \prod_{x \in \{u,d\}} \left[0.5 - 0.5\,\text{erf}\left(\frac{\Omega_{k,1}^x - E_{k,1}^x(\phi_s)}{\sqrt{2}\sigma(\phi_s)}\right)\right]\Pr\{\phi_s\}. \tag{25}$$

3.2.5. Expression for $q_0^c(n_0 + 1, n_{\text{tot}} + 1)$

The parameter $q_0^c(n_0+1, n_{tot}+1)$ defines system transitioning from state $(n_0 + 1)$ to state n_0 when an active class 0 call is successfully completed. By assumption A3, the holding time of a class 0 call is exponentially distributed with mean $1/\mu_0$. Hence,

$$q_0^c(n_0 + 1, n_{\text{tot}} + 1) = (n_0 + 1)\mu_0. \tag{26}$$

3.2.6. Expression for $q_1^c(n_0 + 1, n_{\text{tot}} + 1)$

Similarly as above, $q_1^c(n_0 + 1, n_{\text{tot}} + 1)$ describes state transitioning from state $(n_{\text{tot}} + 1)$ to state n_{tot} due to successful completion of an active class 1 call. Using assumption A3,

$$q_1^c(n_0 + 1, n_{\text{tot}} + 1) = (n_1 + 1)\mu_1. \tag{27}$$

Recall that only n_1 of the n_{tot} class 1 calls in the system are active, each completing at rate μ_1.

3.2.7. Expression for $q_0^b(n_0 + 1, n_{\text{tot}} + 1)$

When an MS engaged in a class 0 call while the system is in state $(n_0 + 1, n_{\text{tot}} + 1)$ moves to a neighboring cell, the call is handed off to the cell for continuity of conversation. In this case, the dwell time in the cell of interest is less than the call duration. By assumption A4, the cell dwell time of a class 0 call is exponentially distributed with mean $1/\upsilon_0$. Hence, the mobility induced handoff rate for class 0 calls is given by

$$q_0^b(n_0 + 1, n_{\text{tot}} + 1) = (n_0 + 1)\upsilon_0. \tag{28}$$

3.2.8. Expression for $q_1^b(n_0 + 1, n_{\text{tot}} + 1)$

Similarly as above, $q_1^c(n_0 + 1, n_{\text{tot}} + 1)$ describes state transitioning from state $(n_{\text{tot}} + 1)$ to state n_{tot} due to handoff of an active class 1 call to a neighboring cell. Using assumption A4,

$$q_1^b(n_0 + 1, n_{\text{tot}} + 1) = (n_1 + 1)\upsilon_1. \tag{29}$$

3.2.9. Expression for $q_1^r(n_0 + 1, n_{\text{tot}} + 1)$

A class 1 call that is temporarily stored in the queue departs once its patience time expires. By assumption A7, the patience time of a class 1 call waiting in the queue is exponentially distributed with mean $1/\mu_{pq}$. Hence,

$$q_1^r(n_0 + 1, n_{\text{tot}} + 1) = ((n_{\text{tot}} + 1) - N_1)\mu_{pq}. \tag{30}$$

3.3. Steady-state equations

Having determined the expressions for the state transition rates in Figure 1, we now can write the steady-state balance equations. Let $p(n_0, n_{\text{tot}})$ denote the steady-state probability that the system is in state (n_0, n_{tot}). Using the rate equality principle [21], we write the following balance equations for all the possible values of $n_0 \in S$ and $n_{\text{tot}} \in \Psi$:

$$
\begin{aligned}
& n_0 = 0, n_{\text{tot}} = 0: \\
& p(0,0)[Y_0(0,0)+Y_1(0,0)] = p(1,0)X_0(1,0)+p(0,1)X_1(0,1), \\
& n_0 = 0, 1 \le n_{\text{tot}} \le N_1 - 1: \\
& p(0,n_{\text{tot}})[Y_0(0,n_{\text{tot}}) + Y_1(0,n_{\text{tot}}) + X_1(0,n_{\text{tot}})] \\
& \quad = p(1,n_{\text{tot}})X_0(1,n_{\text{tot}}) + p(0,n_{\text{tot}}+1)X_1(0,n_{\text{tot}}+1) \\
& \qquad + p(0,n_{\text{tot}}-1)Y_1(0,n_{\text{tot}}-1), \\
& n_0 = 0,\ n_{\text{tot}} = N_1: \\
& p(0,N_1)[Y_0(0,N_1) + Y_1(0,N_1) + X_1(0,N_1)] \\
& \quad = p(1,N_1)X_0(1,N_1) + p(0,N_1-1)Y_1(0,N_1-1) \\
& \qquad + p(0,N_1+1)[X_1(0,N_1) + q_1^r(N_1+1)], \\
& n_0 = 0,\ N_1 < n_{\text{tot}} < N_1 + Q_{pq}: \\
& p(0,n_{\text{tot}})[Y_0(0,n_{\text{tot}}) + Y_1(0,n_{\text{tot}}) + X_1(0,n_{\text{tot}})] \\
& \quad = p(1,n_{\text{tot}})X_0(1,n_{\text{tot}}) + p(0,n_{\text{tot}}-1)Y_1(0,n_{\text{tot}}-1) \\
& \qquad + p(0,n_{\text{tot}}+1)[X_1(0,N_1) + q_1^r(n_{\text{tot}}+1)], \\
& n_0 = 0,\ n_{\text{tot}} = N_1 + Q_{pq}: \\
& p(0,N_1+Q_{pq})[Y_0(0,N_1+Q_{pq})+X_1(0,N_1)+q_1^r(N_1+Q_{pq})] \\
& \quad = p(1,N_1+Q_{pq})X_0(1,N_1+Q_{pq}) \\
& \qquad + p(0,N_1+Q_{pq}-1)Y_1(0,N_1+Q_{pq}-1),
\end{aligned} \tag{31}
$$

$$
\begin{aligned}
& 1 \le n_0 \le N_0 - 1,\ n_{\text{tot}} = 0: \\
& p(n_0,0)[Y_0(n_0,0) + Y_1(n_0,0) + X_0(n_0,0)] \\
& \quad = p(n_0+1,0)X_0(n_0+1,0) \\
& \qquad + p(n_0-1,0)Y_0(n_0-1,0) + p(n_0,1)X_1(n_0,1), \\
& 1 \le n_0 \le N_0 - 1, 1 = n_{\text{tot}} \le N_1 - 1: \\
& p(n_0,n_{\text{tot}})[Y_0(n_0,n_{\text{tot}}) + Y_1(n_0,n_{\text{tot}}) \\
& \qquad\qquad + X_0(n_0,n_{\text{tot}}) + X_1(n_0,n_{\text{tot}})] \\
& \quad = p(n_0+1,n_{\text{tot}})X_0(n_0+1,n_{\text{tot}}) + p(n_0-1,n_{\text{tot}})Y_0(n_0-1,n_{\text{tot}}) \\
& \qquad + p(n_0,n_{\text{tot}}+1)X_1(n_0,n_{\text{tot}}+1) + p(n_0,n_{\text{tot}}-1)Y_1(n_0,n_{\text{tot}}-1), \\
& 1 \le n_0 \le N_0 - 1, n_{\text{tot}} = N_1: \\
& p(n_0,N_1)[Y_0(n_0,N_1)+Y_1(n_0,N_1)+X_0(n_0,N_1)+X_1(n_0,N_1)] \\
& \quad = p(n_0+1,N_1)X_0(n_0+1,N_1)+p(n_0-1,N_1)Y_0(n_0-1,N_1) \\
& \qquad + p(n_0,N_1-1)Y_1(n_0,N_1-1) \\
& \qquad + p(n_0,N_1+1)[X_1(n_0,N_1) + q_1^r(N_1+1)], \\
& 1 \le n_0 \le N_0 - 1, N_1 < n_{\text{tot}} < N_1 + Q_{pq}: \\
& p(n_0,n_{\text{tot}})[Y_0(n_0,n_{\text{tot}}) \\
& \qquad + Y_1(n_0,n_{\text{tot}}) + X_0(n_0,n_{\text{tot}}) + X_1(n_0,N_1) + q_1^r(n_{\text{tot}})]
\end{aligned}
$$

$$
\begin{aligned}
&= p(n_0+1,n_{\text{tot}})X_0(n_0+1,n_{\text{tot}})\\
&\quad + p(n_0-1,n_{\text{tot}})Y_0(n_0-1,n_{\text{tot}})\\
&\quad + p(n_0,n_{\text{tot}}-1)Y_1(n_0,n_{\text{tot}}-1)\\
&\quad + p(n_0,n_{\text{tot}}+1)\left[X_1(n_0,N_1)+q_1^r(n_{\text{tot}}+1)\right],
\end{aligned}
\tag{32}
$$

$$
\begin{aligned}
&n_0 = N_0, n_{\text{tot}} = 0:\\
&p(N_0,0)\left[Y_1(N_0,0)+X_0(N_0,0)\right]\\
&\quad = p(N_0-1,0)Y_0(N_0-1,0)+p(N_0,1)X_1(N_0,1),\\
&n_0 = N_0, 1 \le n_{\text{tot}} \le N_1-1:\\
&p(N_0,n_{\text{tot}})\left[Y_1(N_0,n_{\text{tot}})+X_0(N_0,n_{\text{tot}})+X_1(N_0,n_{\text{tot}})\right]\\
&\quad = p(N_0-1,n_{\text{tot}})Y_0(N_0-1,n_{\text{tot}})\\
&\quad + p(N_0,n_{\text{tot}}+1)X_1(N_0,n_{\text{tot}}+1)\\
&\qquad + p(N_0,n_{\text{tot}}-1)Y_1(N_0,n_{\text{tot}}-1),\\
&n_0 = N_0,\ n_{\text{tot}} = N_1:\\
&p(N_0,N_1)\left[Y_1(N_0,N_1)+X_0(N_0,N_1)+X_1(N_0,N_1)\right]\\
&\quad = p(N_0-1,N_1)Y_0(N_0-1,N_1)\\
&\qquad + p(N_0,N_1+1)\left[X_1(N_0,N_1)+q_1^r(N_1+1)\right]\\
&\qquad + p(N_0,N_1-1)Y_1(N_0,N_1-1),\\
&n_0 = N_0,\ N_1 < n_{\text{tot}} < N_1+Q_{pq}:\\
&p(N_0,n_{\text{tot}})\left[Y_1(N_0,n_{\text{tot}})+X_0(N_0,n_{\text{tot}})+X_1(N_0,N_1)+q_1^r(n_{\text{tot}})\right]\\
&\quad = p(N_0-1,n_{\text{tot}})Y_0(N_0-1,n_{\text{tot}})\\
&\qquad + p(N_0,n_{\text{tot}}+1)\left[X_1(N_0,N_1)+q_1^r(n_{\text{tot}}+1)\right]\\
&\qquad + p(N_0,n_{\text{tot}}-1)Y_1(N_0,n_{\text{tot}}-1),\\
&n_0 = N_0,\ n_{\text{tot}} = N_1+Q_{pq}:\\
&p(N_0,N_1+Q_{pq})\left[X_0(N_0,N_1+Q_{pq})+X_1(N_0,N_1)+q_1^r(N_1+Q_{pq})\right]\\
&\quad = p(N_0-1,N_1+Q_{pq})Y_0(N_0-1,N_1+Q_{pq})\\
&\qquad + p(N_0,N_1+Q_{pq}-1)Y_1(N_0,N_1+Q_{pq}-1).
\end{aligned}
\tag{33}
$$

The balance equations (31)–(33) along with the normalization condition $\sum_{n_0\in S}\sum_{n_{\text{tot}}\in\Psi} p(n_0,n_{\text{tot}}) = 1$ are solved to obtain the steady-state probabilities $p(n_0,n_{\text{tot}})$ for all $n_0 \in S$ and $n_{\text{tot}} \in \Psi$. Let $\pi(n_0)$ and $\tau(n_{\text{tot}})$, respectively, denote the marginal steady-state probabilities for the number of class 0 and class 1 calls in the system; these marginal probabilities are calculated by

$$
\begin{aligned}
\pi(n_0) &= \sum_{n_{\text{tot}}\in\psi} p(n_0,n_{\text{tot}}),\\
\tau(n_{\text{tot}}) &= \sum_{n_0\in S} p(n_0,n_{\text{tot}}).
\end{aligned}
\tag{34}
$$

3.4. Performance measures

The performance measures for the proposed two-level call admission control scheme are presented in this section.

3.4.1. Blocking probability for new calls

The blocking of a class 0 or class 1 new call is due to two factors: blocking due to insufficient E_b/I_0 or blocking due to insufficient channel resources. The expressions for blocking probability of a new call are given by

$$
\begin{aligned}
P_{nb}^{(0)} &= \sum_{n_0\in S^-}\left[1-A_0^n(n_0)\right]\pi(n_0)+\pi(N_0),\\
P_{nb}^{(1)} &= \sum_{n_{\text{tot}}\in\Psi^-}\left[1-A_1^n(n_{\text{tot}})\right]\tau(n_{\text{tot}})+\tau(N_1+Q_{pq}),
\end{aligned}
\tag{35}
$$

where $S^- = \{0,1,2,\ldots,N_0-1\}$, $\Psi^- = \{0,1,2,\ldots,N_1+Q_{pq}-1\}$, and the superscripts 0 and 1 represent class 0 and class 1 calls, respectively.

3.4.2. Blocking probability for Handoff calls

The blocking or forced termination of a class 0 or class 1 handoff call is also due to two factors: blocking due to insufficient E_b/I_0 or blocking due to insufficient channel resources. The expressions for blocking probability of a handoff call are given by

$$
\begin{aligned}
P_{hb}^{(0)} &= \sum_{n_0\in S^-}\left[1-A_0^h(n_0)\right]\pi(n_0)+\pi(N_0),\\
P_{hb}^{(1)} &= \sum_{n_{\text{tot}}\in\Psi^-}\left[1-A_1^h(n_{\text{tot}})\right]\tau(n_{\text{tot}})+\tau(N_1+Q_{pq}).
\end{aligned}
\tag{36}
$$

3.4.3. Outage probability for class 0 calls

A real-time call is in outage if the received E_b/I_0 of the call falls below the required threshold for acceptable communication. Since the received E_b/I_0 is measured at both the MS and BS, both the downlink and uplink must be tested for outage. Let θ_0^x denote the outage probability for real-time calls in the x direction, where $x \in \{u,d\}$.θ_0^x is calculated by the formula

$$
\theta_0^x = \sum_{n_0\in S}\pi(n_0)\sum_{\phi_s\in\mathbf{Q(s)}}\left[0.5-0.5\,\text{erf}\left(\frac{M_0^x(\phi_s)-\Gamma_0^x}{\sqrt{2}\sigma(\phi_s)}\right)\right]\Pr\{\phi_s\}.
\tag{37}
$$

3.4.4. Outage probability for class 1 calls

Similarly, the expression for the outage probability for class 1 calls in the x direction ($x \in \{u,d\}$) is given by

$$
\begin{aligned}
&\theta_1^x\\
&= \sum_{n_{\text{tot}}\in\Psi'}\tau(n_{n\,\text{tot}})\sum_{\phi_s\in\mathbf{Q(s)}}\left[0.5-0.5\,\text{erf}\left(\frac{M_1^x(\phi_s)-\Gamma_1^x}{\sqrt{2}\sigma(\phi_s)}\right)\right]\Pr\{\phi_s\},
\end{aligned}
\tag{38}
$$

where the set $\Psi\prime = \{0,1,2,3,\ldots,N_1\}$ spans only the possible number of class 1 calls that can be in progress and does not include those waiting in the queue.

3.4.5. Throughput for class 0 calls

Throughput is defined as the allocated data rate under the condition that the received E_b/I_0 exceeds the required E_b/I_0.

Let Z_0^x denote the throughput for class 0 calls in the x direction. The expression for Z_0^x is given by

$$Z_0^x = \sum_{n_0 \in S} \pi(n_0) \sum_{\phi_s \in \mathbf{Q}(\mathbf{s})} \left[0.5 - 0.5\,\mathrm{erf}\left(\frac{\Gamma_0^x - M_0^x(\phi_s)}{\sqrt{2}\sigma(\phi_s)}\right)\right] \Pr\{\phi_s\}(n_0 R_0^x). \tag{39}$$

The total system throughput due to the active class 0 calls, Z_0, is the sum of the throughput for the uplink and downlink, that is, $Z_0 = Z_0^u + Z_0^d$.

3.4.6. Throughput for class 1 calls

Similarly, let Z_1^x denote the throughput for class 1 calls in the x direction. Then,

$$Z_1^x = \sum_{n_{\text{tot}} \in \Psi'} \tau(n_{\text{tot}}) \sum_{\phi_s \in \mathbf{Q}(\mathbf{s})} \left[0.5 - 0.5\mathrm{erf}\left(\frac{\Gamma_1^x - M_1^x(\phi_s)}{\sqrt{2}\sigma(\phi_s)}\right)\right] \Pr\{\phi_s\}(n_{\text{tot}} R_1^x). \tag{40}$$

The total system throughput due to class 1 calls, Z_1, is the sum of the throughput for the uplink and downlink, that is, $Z_1 = Z_1^u + Z_1^d$.

3.4.7. Average queuing delay for class 1 calls

Using Little's law [22], the average waiting time of class 1 calls in the queue, $E[W]$, is calculated using the formula

$$E[W] = \frac{\sum_{n_{\text{tot}}=N_1}^{N_1+Q_{pq}} (n_{\text{tot}} - N_1)\tau(n_{\text{tot}})}{(1 - P_{nb}^{(1)})\Lambda_1 + (1 - P_{hb}^{(1)})\lambda_1}. \tag{41}$$

4. PERFORMANCE RESULTS AND DISCUSSION

One goal of the performance results is to compare the performance of the proposed two-level CAC scheme with that of a previously proposed single-level CAC scheme [10]. Another objective is to conduct sensitivity analysis to study the effect of downlink bandwidth ratio and bandwidth reservation ratio parameters on the system performance metrics of call blocking, outage probability, average throughput achieved for both class 0 (real-time) and class 1 (nonreal-time) calls, and average waiting time in the queue of class 1 calls. The purpose of the sensitivity analysis is to provide guidance in the selection of system parameter values to achieve optimal system performance.

4.1. Assumed input parameter values

Values of system parameters which are specific to the proposed two-level CAC scheme are selected as follows: noise rise coefficient $\eta = 0.1$, standard deviation of log-normal shadowing $\sigma = 8$ dB, maximum queue size $Q_{pq} = 20$ class 1 calls, and the average patience time for calls stored in the queue is 100 seconds. Values of the remaining system parameters are chosen similarly as those used in [10]. For both class 0 and class 1 calls, the assumed values for data rates, activity factors, call mix, mean call duration, and mean cell dwell time are summarized in Table 1. The required nominal E_b/I_0 thresholds for quality communication for the two traffic classes in the uplink and downlink directions are selected as $\Gamma_0^u = \Gamma_1^u = \Gamma_0^d = \Gamma_1^d = 4$ dB. The call admission control parameters are set as $\beta_0^h = 1.05, \beta_0^n = 1.1$, and $\beta_1^h = \beta_1^n = 1.2$. The ratio of other cell interference to own cell interference in the uplink (ξ^u) and downlink (ξ^d) is chosen as 0.5. The proportion of the total base-station power spent on overhead channels, $z = 0.3$, and the downlink average orthogonality factor in a cell, ρ, is set at 0.5. Unless otherwise stated, nominal values of downlink bandwidth ratio, bandwidth reservation ratio, and total new call arrival rate are chosen as $W_d/W = 0.8$, $W_{\text{res}}/W = 0.1$, and $\Lambda = 0.1$ calls per second, respectively. Finally, due to the interdependence between the mean handoff rate (λ_0 and λ_1) and the state probabilities $p(n_0, n_{\text{tot}})$, values of the mean handoff rates are not specified explicitly, but instead they are computed iteratively.

4.2. Performance results

Figure 2 presents class 0 new and handoff call blocking probability. As the value of downlink bandwidth ratio (W_d/W) is varied from 0.5 to 0.9, this range is selected to account for the higher traffic flow in the downlink compared to uplink. Other parameter values are set to their nominal values, as stated earlier. Note the very good agreement between the simulation results (represented by symbols) and the analysis results (depicted by lines) for the proposed 2-level CAC scheme. It is observed from Figure 2 that the 2-level CAC scheme exhibits a lower blocking probability than the 1-level scheme. The lower blocking performance is due to the fact that class 0 calls are given higher priority in accessing the bandwidth resources and the balance is used by class 1 calls. Notice also that, for the 2-level CAC, the downlink bandwidth ratio has no effect on the class 0 call blocking probability when the value of W_d/W lies in the range of 0.5 to 0.85. Beyond $W_d/W = 0.85$, the blocking level increases sharply. One implication of the observed call blocking behavior for the 2-level CAC is the flexibility in selecting the bandwidth ratio to meet a specified downlink traffic level without a negative impact on class 0 call blocking level. The above observation and explanation also apply to the call blocking performance for class 1 calls, which is shown in Figure 3. A comparison of the call blocking performance of class 0 and class 1 calls in Figures 2 and 3 shows that, for the 2-level CAC scheme and values of W_d/W in the range of $[0.5, 0.85]$, the class 1 call blocking is the same as class 0 call blocking. However, for the single-level CAC, class 1 call blocking is much higher than class 0 call blocking. Class 1 call blocking is identical to class 0 call blocking for the 2-level CAC scheme because of the flexibility of queuing class 1 calls that cannot be assigned resources at initial access request instant. Queuing of class 1 calls therefore translates to a reduction in their blocking level caused by resource shortage. The sensitivity of call blocking level to new call arrival rate is presented in Figure 4, for both the 1-level and 2-level CAC schemes.

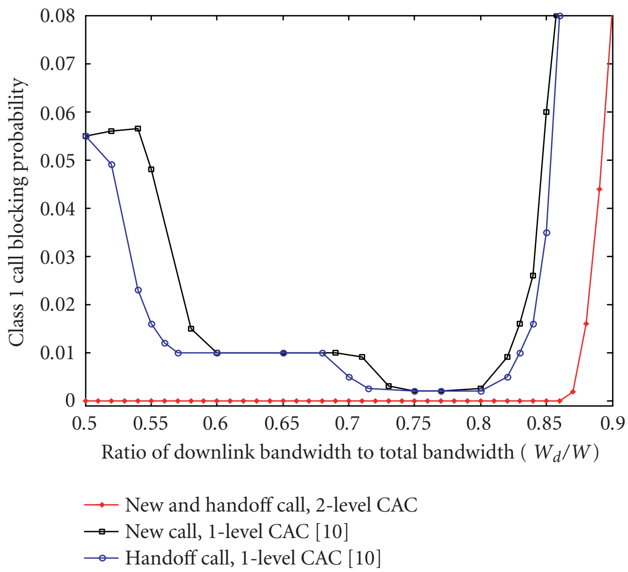

Figure 3: Class 1 call blocking probabilities versus $W_d/W(\Lambda = 0.1, W_{\text{res}}/W = 0.1)$.

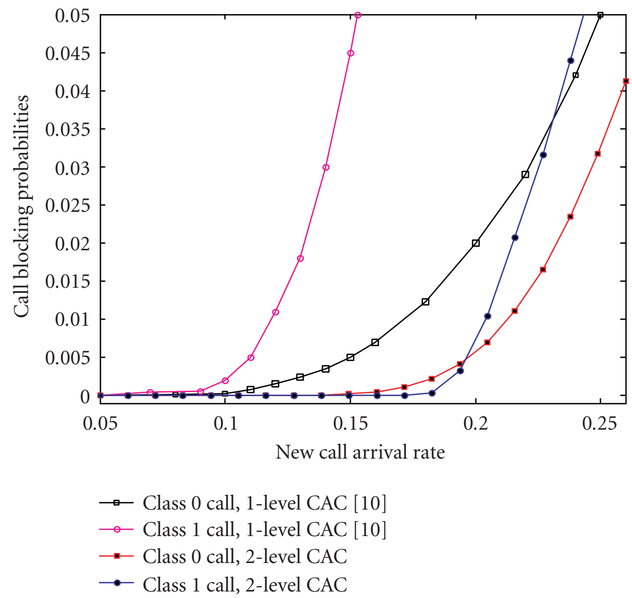

Figure 4: Call blocking probabilities versus aggregate new call arrival rate.

Table 1: Traffic model parameter values [10].

	Class 0 call		Class 1 call	
Link	Uplink	Downlink	Uplink	Downlink
Data rate, R_i	16 kbps	16 kbps	64 kbps	384 kbps
Activity factor, α_i	0.5	0.5	0.00285	0.015
Call mix	90%		10%	
Mean call duration, $1/\mu_i$	120 seconds		3000 seconds	
Mean cell dwell time, $1/\nu_i$	300 seconds		1200 seconds	

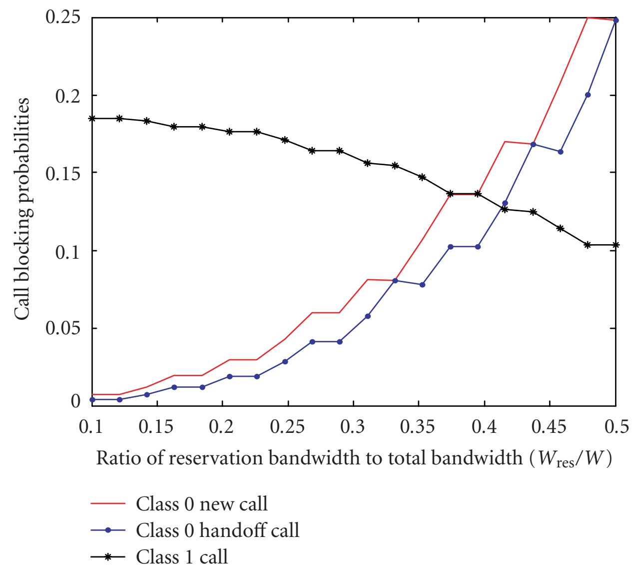

Figure 5: Call blocking probabilities versus reservation bandwidth ratio.

Figure 4 is useful for determining the maximum call arrival rate that can be supported at a desired call blocking performance objective. For example, at a grade of service objective of 1% call blocking, the proposed 2-level CAC scheme can support maximum call arrival rates of 0.22 and 0.21 calls/sec for class 0 and class 1 calls, respectively. These maximum call arrival rates represent 30% and 75% capacity gain over the corresponding numbers achieved with the 1-level CAC scheme. Figure 5 shows the effect of increasing the reservation bandwidth ratio on the call blocking performance for class 0 and class 1 calls. It is observed from Figure 5 that class 1 call blocking probability is reduced as the reservation bandwidth ratio increases. The penalty though is the concomitant increase in the blocking probabilities for class 0 new and handoff calls. Figure 5 is useful for determining the proper value of reservation bandwidth ratio that simultaneously satisfies the call blocking performance objectives for class 0 and class 1 calls.

Figure 6 presents the uplink and downlink outage probabilities of class 0 and class 1 calls at different values of W_d/W for both the 1-level and 2-level CAC schemes. For either CAC scheme, the downlink outage probability decreases as W_d/W increases. An opposite trend is observed for uplink outage probability. The preceding statements imply that higher bandwidth improves the outage probability. This is so because the received E_b/I_0 is directly proportional to the bandwidth so that a larger bandwidth ensures that the received E_b/I_0 is large enough to always exceed the required threshold, thereby preventing an outage. It is also interesting to find that, at $W_d/W < 0.53$, the downlink outage probability obtained with the 2-level CAC scheme is higher than the corresponding result for the 1-level CAC scheme. Beyond W_d/W of 0.53, the outage performance for the 2-level CAC is better than the 1-level CAC scheme. Note that the performance improvement is not due to the queuing of class 1 calls because such calls are not actually active and do not generate

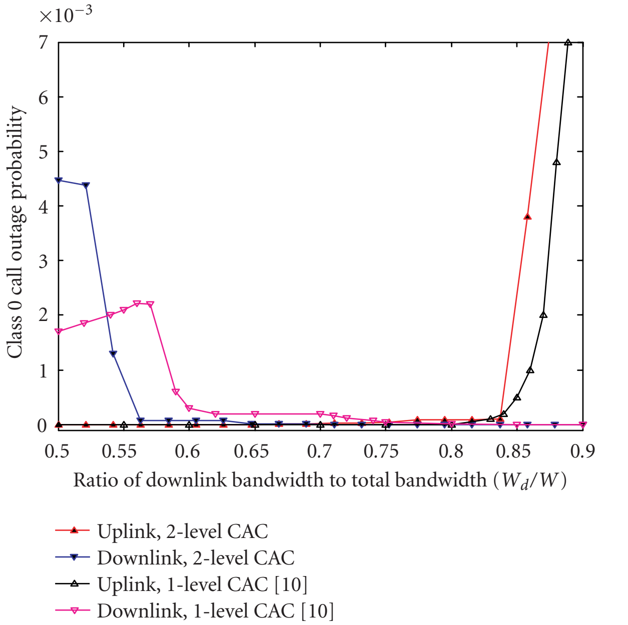

Figure 6: Outage probabilities versus downlink bandwidth ratio, W_d/W.

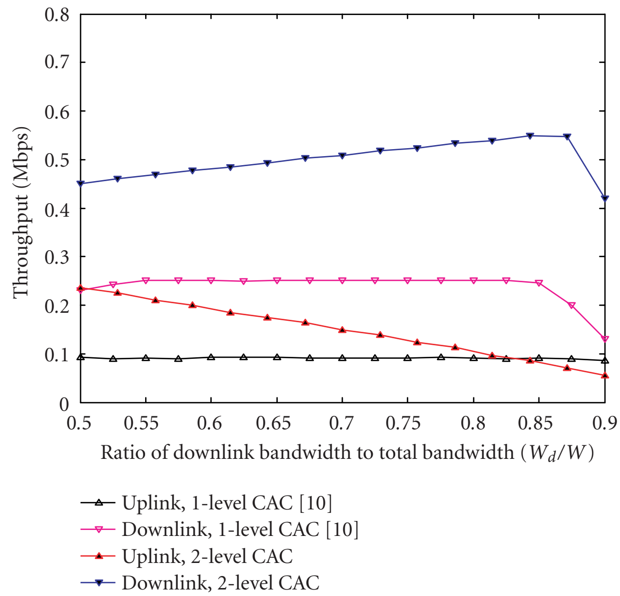

Figure 8: Average throughput versus downlink bandwidth ratio, W_d/W.

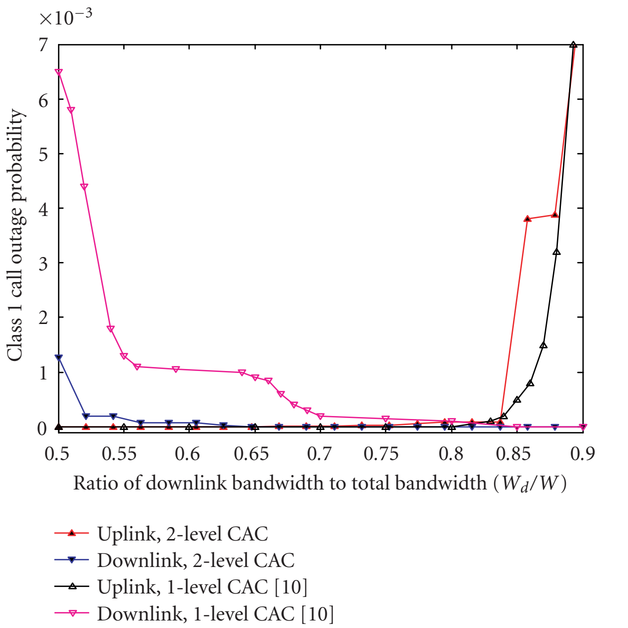

Figure 7: Outage probabilities versus downlink bandwidth ratio, W_d/W.

interference to the other existing calls. The performance improvement is therefore explained by the higher priority assigned to class 0 calls in gaining access to system resources prior to class 1 calls. It is also interesting to find that the 2-level CAC scheme exhibits a higher downlink outage probability when $W_d/W \geq 0.84$. The preceding observation for uplink and downlink outage probabilities suggests that the 2-level CAC scheme is very sensitive to low bandwidth (i.e., $W_u/W \leq 0.16$ for uplink and $W_d/W \leq 0.55$ for downlink) resulting in poor performance. Figure 7 presents the corresponding outage probability results for class 1 calls. A comparison of Figure 6 with Figure 7 shows that while for the 1-level CAC scheme, class 1 call downlink outage probability is higher than the corresponding results for class 0 calls, the reverse is the case for the 2-level CAC scheme. The uplink outage probabilities for class 0 and class 1 calls are similar for either CAC scheme. Figures 6 and 7 are useful for determining the outage performance targets that correspond to a specified call blocking objective. For example, the downlink bandwidth ratio for a 1% call blocking objective is found to be 0.87 from Figures 2 and 3. Using Figures 6 and 7 and assuming the 2-level CAC scheme, the downlink and uplink outage objectives are 0% and 0.7%, respectively, for class 0 calls and 0% and 0.4% for class 1 calls. Clearly, the outage probabilities are much less than call blocking, as desired.

Figure 8 compares the uplink and downlink throughput obtained using the proposed 2-level CAC scheme with that achieved by the 1-level CAC scheme at different values of downlink bandwidth ratio. Two observations are evident from Figure 8. First, the downlink throughput achieved with the 2-level CAC scheme is roughly double that obtained with the 1-level CAC scheme. The improvement is due to the queuing of class 1 calls in the 2-level CAC scheme. The 2-level CAC scheme fully makes use of delay-tolerant characteristic of class 1 calls to balance the traffic flow between class 0 and class 1 calls. Hence, the system resources are used more efficiently as manifested by the improved throughput levels. The second observation from Figure 9 is that while the downlink throughput begins to decrease at W_d/W of 85% for the 1-level CAC scheme, the degradation in throughput for the 2-level CAC scheme begins at about 87% downlink bandwidth ratio demonstrating better

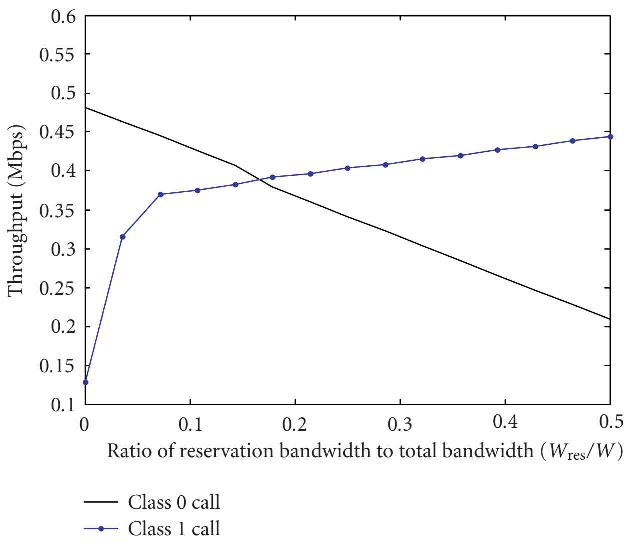

FIGURE 9: Throughput versus reservation bandwidth ratio, W_{res}/W.

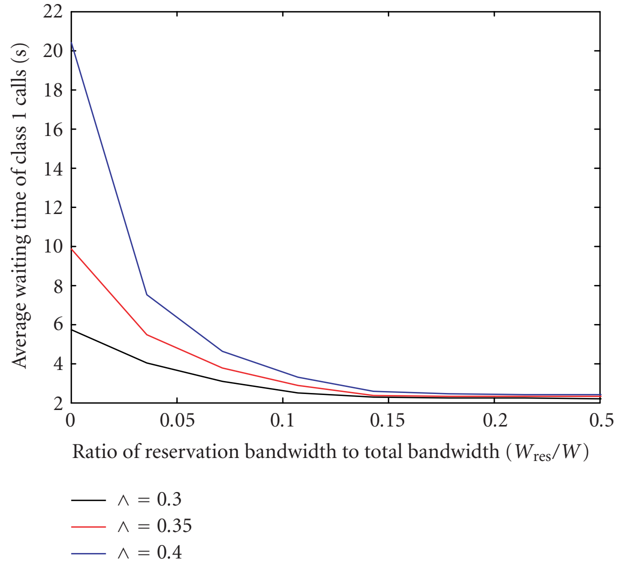

FIGURE 10: Average waiting time of class 1 calls versus reservation bandwidth ratio, W_{res}/W.

robustness for the 2-level CAC scheme. In case of the uplink throughput level, the 1-level CAC scheme is more robust than the 2-level CAC scheme whose throughput level decreases with W_d/W. Figure 9 presents the sensitivity of class 0 and class 1 call throughput levels to reservation bandwidth ratio. Clearly, the class 1 call throughput increases as the reservation bandwidth increases but with a reduction in the class 0 call throughput levels, as expected. Figures 5 and 9 present the tradeoff between increasing the reservation bandwidth ratio to meet a minimum performance objective for class 1 calls and the degradation in the class 0 call blocking and throughput performance. It is concluded from Figures 5 and 9 that reserving more bandwidth for class 1 calls translates to lower call blocking and higher throughput but at the expense of higher call blocking and lower throughput for class 0 calls.

Figure 10 presents the average waiting time of class 1 calls assuming the 2-level CAC scheme. The results are plotted against reservation bandwidth ratio and parameterized by aggregate new call arrival rate, Λ. For a given value of Λ, the average waiting time decreases very rapidly as the bandwidth reservation ratio is increased. As an example, at an aggregate new call arrival rate of 0.35 calls/sec, the average waiting time decreases by 66% when the reservation bandwidth ratio is increased from 0.05 to 0.1. Note from Figure 10 that, for a given Λ, the average waiting time can be reduced to a small value (i.e., approximately 2 seconds) by an appropriate choice of reservation bandwidth ratio. It is found that the reservation bandwidth ratio to make the average waiting time equal to zero is higher at large values of aggregate new call arrival rate. Note, however, that the reservation bandwidth ratio cannot be increased arbitrarily because of its negative impact on the throughput and blocking level for class 0 calls, as found earlier from Figures 5 and 9. It is also observed that, at a given value of reservation bandwidth, the average waiting time for class 1 calls increases with the call arrival rate, as expected. For example, at a reservation bandwidth ratio of 10%, the average waiting times of a class 1 call (e.g., file download) are 3 seconds, 2 seconds and 1 second for $\Lambda = 0.4$, 0.35, and 0.3 calls/sec, respectively.

5. CONCLUSION

In this paper, we propose a two-level call admission control scheme for wireless DS-CDMA networks carrying multimedia traffic. The scheme fully utilizes the traffic characteristics of wireless multimedia communication; it assigns higher priority to the real-time traffic class (i.e., class 0 calls) in gaining access to the system resources and implements queuing of nonreal-time calls (i.e., class 1 calls) that cannot be allocated resources at initial request instant. To ensure that nonreal-time calls are not starved of resources due to the higher priority given to real-time calls, the scheme also incorporates some reserved capacity for nonreal-time calls. For each traffic class, the scheme manages the uplink and downlink resources separately. Further, the scheme also manages the resources to new and handoff calls within each traffic class. Performance of the proposed two-level CAC scheme is analyzed using Markov chain theory to derive system performance metrics of call blocking, outage probability, average throughput, and average waiting time of nonreal-time calls in the queue. The numerical results obtained from analysis show that the proposed two-level CAC scheme exhibits a lower call blocking, lower outage probability, and a higher throughput than the corresponding results obtained using a single-level call admission control. For example, our results show that the two-level call admission control can achieve up to 75% capacity gain over the single-level CAC scheme. It is found that the average waiting time introduced by the queuing of nonreal-time calls can be reduced by appropriate selection of the reservation bandwidth ratio but this value must be chosen carefully so as not to seriously degrade the performance of real-time calls. Based on the results, it is concluded that the proposed two-level CAC scheme would serve as a viable alternative for managing the resources of a DS-CDMA wireless communication system.

ACKNOWLEDGMENT

This research is supported in part by a grant from Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES

[1] Z. Liu and M. El Zarki, "SIR-based call admission control for DS-CDMA cellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 4, pp. 638–644, 1994.

[2] J. Lee and Y. Han, "Downlink admission control for multimedia services in WCDMA," in *Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '02)*, vol. 5, pp. 2234–2238, Lisbon, Portugal, September 2002.

[3] W. S. Jeon and D. G. Jeong, "Call admission control for mobile multimedia communications with traffic asymmetric between uplink and downlink," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 1, pp. 59–66, 2001.

[4] M. Casoni, G. Immovilli, and M. L. Merani, "Admission control in T/CDMA systems supporting voice and data applications," *IEEE Transactions on Wireless Communications*, vol. 1, no. 3, pp. 540–548, 2002.

[5] C. Comaniciu, N. B. Mandayam, D. Famolari, and P. Agrawal, "Wireless access to the World Wide Web in an integrated CDMA system," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 472–483, 2003.

[6] P. R. Larijani, R. H. Hafez, and I. Lambadaris, "Two level access control strategy for multimedia CDMA," in *Proceedings of IEEE International Conference on Communications (ICC '98)*, vol. 1, pp. 487–492, Atlanta, Ga, USA, June 1998.

[7] C. Comaniciu, N. B. Mandayam, D. Famolari, and P. Agrawal, "QoS guarantees for third generation (3G) CDMA systems via admission and flow control," in *Proceedings of the 52nd Vehicular Technology Conference (VTC '00)*, vol. 1, pp. 249–256, Boston, Mass, USA, September 2000.

[8] T.-K. Liu and J. Silvester, "Joint admission/congestion control for wireless CDMA systems supporting integrated services," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 6, pp. 845–857, 1998.

[9] M. Soroushnejad and E. Geraniotis, "Multi-access strategies for an integrated voice/data CDMA packet radio network," *IEEE Transactions on Communications*, vol. 43, no. 2–4, pp. 934–945, 1995.

[10] W. S. Jeon and D. G. Jeong, "Call admission control for CDMA mobile communications systems supporting multimedia services," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 649–659, 2002.

[11] A. Sampath and J. M. Holtzman, "Access control of data in integrated voice/data CDMA systems: benefits and trade-off," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 8, pp. 1511–1526, 1987.

[12] D. Ayyagari and A. Ephremides, "Optimal admission control in cellular DS-CDMA systems with multimedia traffic," *IEEE Transactions on Wireless Communications*, vol. 2, no. 1, pp. 195–202, 2003.

[13] S. Singh, V. Krishnamurthy, and H. V. Poor, "Integrated voice/data call admission control for wireless DS-CDMA systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1483–1495, 2002.

[14] A. Sampath, P. Kumar, and J. M. Holtzman, "Power control and resource management for a multimedia CDMA wireless system," in *Proceedings of the 6th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '95)*, vol. 1, pp. 21–25, Toronto, Canada, September 1995.

[15] S. Aissa, J. Kuri, and P. Mermelstein, "Call admission on the uplink and downlink of a CDMA system based on total received and transmitted powers," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2407–2416, 2004.

[16] G. S. Paschos, I. D. Politis, and S. A. Kotsopoulos, "A quality of service negotiation-based admission control scheme for WCDMA mobile wireless multiclass services," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1875–1886, 2005.

[17] X. Yang, G. Feng, and D. S. C. Kheong, "Call admission control for multiservice wireless networks with bandwidth asymmetry between uplink and downlink," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 1, pp. 360–368, 2006.

[18] W. Li and X. Chao, "Call admission control for an adaptive heterogeneous multimedia mobile network," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 515–525, 2007.

[19] A. M. Viterbi and A. J. Viterbi, "Erlang capacity of a power controlled CDMA system," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 6, pp. 892–900, 1993.

[20] A. Sampath, N. B. Mandayam, and J. M. Holtzman, "Erlang capacity of a power controlled integrated voice and data CDMA system," in *Proceedings of the 47th IEEE Vehicular Technology Conference (VTC '97)*, vol. 3, pp. 1557–1561, Phoenix, Ariz, USA, May 1997.

[21] S. M. Ross, *Introduction to Probability Models*, Academic Press, London, UK, 1985.

[22] L. Kleinrock, *Queueing Systems. Volume 1: Theory*, John Wiley & Sons, New York, NY, USA, 1975.