Cross-Layer Optimal Connection Admission Control for Variable Bit Rate Multimedia Traffic in Packet Wireless CDMA Networks

Fei Yu, *Member, IEEE*, Vikram Krishnamurthy, *Fellow, IEEE*, and Victor C. M. Leung, *Fellow, IEEE*

***Abstract*—Next generation wireless code division multiple access (CDMA) networks are required to support packet multimedia traffic. This paper addresses the connection admission control problem for multiservice packet traffic modeled as Markov modulated Poisson process (MMPP) with the quality of service (QoS) requirements on both physical layer signal-to-interference ratio (SIR) and network layer blocking probability. Optimal linear-programming-based algorithms are presented that take into account of SIR outage probability constraints. By exploiting the MMPP traffic models and introducing a small SIR outage probability, the proposed algorithms can dramatically improve the network utilization. In addition, we propose two reduced complexity algorithms that require less computation and can have satisfactory approximation to the optimal solutions. Numerical examples illustrating the performance of the proposed schemes are presented.**

***Index Terms*—Admission control, cross layer, code division multiple access (CDMA), multimedia, Markov decision process, nearly completely decomposable Markov chain.**

I. Introduction

In recent years, there has been significant growth in the use of wireless mobile communications. Code division multiple access (CDMA) is among the most promising multiplexing technologies for next generation wireless mobile networks. With the growing demand for bandwidth-intensive multimedia application (e.g., video) in CDMA networks, quality of service (QoS) provisioning is becoming increasingly important. An efficient connection admission control (CAC) scheme is crucial to guarantee the QoS and to maximize the network utilization simultaneously [1]–[7].

There are different QoS metrics at different layers of wireless CDMA networks. At the physical layer, the QoS requirements are usually characterized by signal-to-interference ratio (SIR) or bit error probability (BEP). At the network layer, QoS metrics are blocking probabilities of connections. With notable exceptions discussed below, most previous work addresses the QoS at only one layer, e.g., [2], [3] at physical layer and [1] at network layer. However, the interplay between physical layer and network layer plays an important role in designing wireless networks [8]. Therefore, the scheme optimized for only one layer (physical layer or network layer) may result in unsatisfactory performance for the overall system. Two recent papers study the admission control problem by considering physical layer QoS and network QoS jointly [4], [5]. Optimal admission policies are designed to maximize the network utilization (minimize blocking probabilities) with constraints on SIR and blocking probabilities. An optimal power control policy is also given in [5] under this cross-layer design framework.

Although [4], [5] guarantee both physical layer QoS and network QoS by means of cross-layer optimization, only *constant bit rate traffic* and *circuit-switched networks* are considered. In particular, it is assumed in [4], [5] that the radio resource consumed by a connection is not changed and the SIR constraint should be guaranteed at all time instants during the lifetime of the connection. However, in order to provide integrated services and utilize radio resource more efficiently, next generation wireless mobile networks are required to support packet traffic such as multimedia and Internet Protocol (IP) data [9], where a connection may change the radio resource consumption and the QoS requirements during its lifetime. To the best of our knowledge, design of optimal admission control schemes that consider cross-layer issues with variable bit rate packet traffic has not been addressed in previous work. For example, [6], [7] consider packet traffic in the design of admission control schemes, however, the physical layer is not considered in [6] and packet traffic is not considered in the optimality framework in [7]. Cross-layer designs are proposed in [10] for the medium access control (MAC) scheme, which is different from the admission control problem considered here.

In this paper, we propose optimal connection admission control schemes in CDMA networks with variable bit rate packet multimedia traffic by considering both the physical and network layers. Specifically, packet traffic is modeled as Markov-modulated Poisson process (MMPP) [11], which has been widely used in modeling various types of multimedia traffic such as voice [12], Motion Picture Experts Group (MPEG) video [13] and self-similar traffic [14], [15]. We present four connection admission control schemes:

1) In the first proposed scheme, which we call CDMA MMPP admission control I (CMAC-I), both physical layer QoS and network QoS are guaranteed at all time instants and no statistical multiplexing is considered. We show that guaranteeing QoS at all time instants results in low network utilization for packet traffic.

Manuscript received February 9, 2004; revised December 26, 2004. This work was supported by the Canadian Natural Science and Engineering Research Council. This paper was presented in part at the Globecom'04, December 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ananthram Swami.

The authors are with the Department of Electrical and Computer Engineering, University of British Columbia, 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4 (e-mail: feiy@ece.ubc.ca; vikramk@ece.ubc.ca; vleung@ece.ubc.ca).

Digital Object Identifier 10.1109/TSP.2005.861785

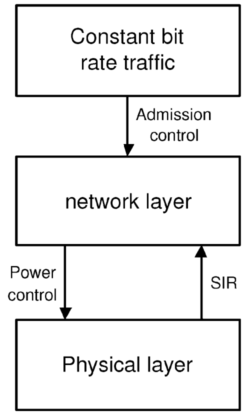

Fig. 1. Cross-layer optimization for constant bit rate traffic in circuit-switched CDMA networks.

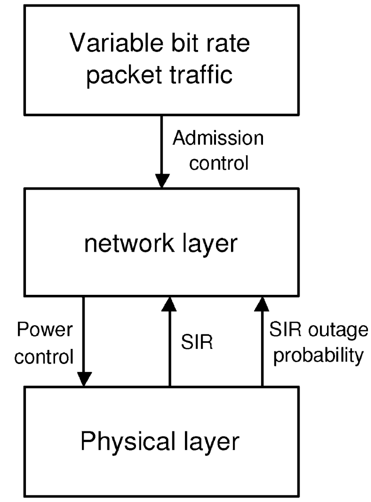

Fig. 2. Cross-layer optimization for variable bit rate packet traffic in packet CDMA networks considered in this paper.

2) In the second scheme, CMAC-II, we introduce an alternative physical layer QoS metric, namely SIR outage probability P^{out}. Instead of guaranteeing the SIR at all the time, we can guarantee P^{out}. With a small P^{out}, the blocking probability can be decreased, and hence the network utilization can be increased significantly.

Both the above schemes provide optimal solutions to the connection admission control problem with and without SIR outage. However, the computation complexity is extensive in these two schemes due to the curse of dimensionality inherent in solving any semi-Markov decision process (SMDP). It is very difficult, if not impossible, to get feasible solutions in real networks due to the problem of large dimensionality. In order to tackle this large dimensionality, we apply the notion of *nearly completely decomposability* to the CAC problem and propose two reduced CAC algorithms we call CMAC-III and CMAC-IV.

3) In CMAC-III, a conservative approach to the constraint of P^{out} is used, where the SIR outage probability can be guaranteed at all time instants.
4) CMAC-IV considers an aggressive approach to the constraint of P^{out} to increase the network utilization. In this scheme, the SIR outage probability constraint may be violated at some time instants, but the long term SIR outage probability can be guaranteed.

Figs. 1 and 2 illustrate the conceptual design difference between the approach proposed in [5] and those in this paper. We study the cross-layer optimization problem in packet-switched networks with variable bit rate packet traffic. By exploiting the MMPP packet traffic models and introducing a small SIR outage probability, the proposed algorithms can dramatically improve the network utilization. This is similar to the philosophy of CDMA physical layer where nonorthogonal codes introduce multiaccess interference (MAI) but increase network capacity. Moreover, we apply the notion of nearly completely decomposability to reduce the computation complexity of the problem. We show the effectiveness of the proposed schemes by numerical examples using voice and *Star Wars* video traffic.

The rest of this paper is organized as follows. Section II describes the traffic and physical layer models. The admission control scheme without SIR outage is presented in Section III. Section IV discusses the design of admission control with SIR outage. SectionV illustrates the nearly complete decomposability of the admission control problem. The admission control schemes with reduced complexity are presented in Section VI. Section VII illustrates the performance of the proposed schemes by numerical examples. Finally, we conclude this paper in Section VIII.

II. Traffic Model and CDMA Physical Layer Model

In this section, we introduce the MMPP and use it to model the packet traffic in packet-switched CDMA networks. In order to study the interplay between the physical layer and network layer, we derive the asymptotic system capacity and the minimum transmit power control solution for CDMA networks with linear minimum mean square error (lmmse) receivers.

A. MMPP Traffic Models for Multimedia Traffic

One of the main differences between this paper and the previous approaches [4], [5] is that we exploit the packet traffic models, which play a significant role in the design and engineering of packet networks. In this paper, we use the MMPP traffic model, which is analytically tractable and has been used extensively in the representation and study of a variety of traffic [16]. We fist introduce the general MMPP model. Then, we give some examples of using MMPP to model voice, video and self-similar traffic.

1) General MMPP Model: Assume that there are $j, j = 1, 2, \ldots, J$, classes of statistically independent traffic in the network. Class j traffic has M_j states with the process, while in any state $m_{ji}, 1 \leq m_{ji} \leq M_j$, behaving as a Poisson process with a state-dependent rate parameter R_{ji}. Transitions between states are governed by an underlying continuous-time Markov chain (or more generally a Markov renewal process). Fig. 3 shows an example of MMPP model, where θ_{in} is the rate of transition between state i and n.

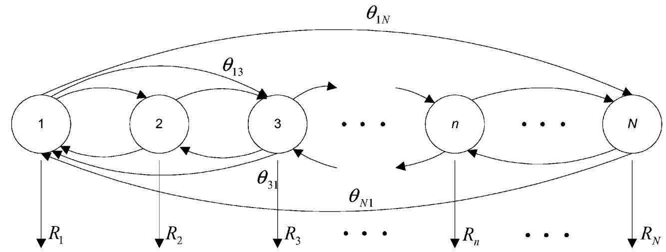

Fig. 3. MMPP traffic model.

The row vector $\pi_j = [\pi_{j1}, \pi_{j2}, \ldots, \pi_{jM_j}]$ representing the M_j stationary distributions satisfies

$$\pi_j \Theta_j = 0. \tag{1}$$

Θ_j is the infinitesimal generating matrix for the Markov chain governing the class j MMPP traffic. Θ_j is given by

$$\Theta_j = \begin{bmatrix} \theta_{11}^j & \theta_{12}^j & \cdots & \theta_{1M_j}^j \\ \theta_{21}^j & \theta_{22}^j & \cdots & \theta_{2M_j}^j \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{M_j1}^j & \theta_{M_j2}^j & \cdots & \theta_{M_jM_j}^j \end{bmatrix} \tag{2}$$

where the diagonal rate parameter θ_{ii}^j is given by $\theta_{ii}^j = -\sum_{n=1, n\neq i}^{M_j} \theta_{\text{in}}^j$. The rate entering state m_{ji} is $\alpha_{ji} = \sum_{n=1}^{M_j} \theta_{ni}^j$ and the rate leaving state m_{ji} is $\beta_{ji} = \sum_{n=1}^{M_j} \theta_{\text{in}}^j$. Class j MMPP connections arrive according to Poisson process with rate λ_j. The durations of class j connections are exponentially distributed with mean μ_j. When an MMPP connection arrives, the initial state is m_{ji} with probability r_{ji}.

2) MMPP Models for Voice, Video and Self-Similar Traffic: It is well known [12] that human speech consists of an alternating sequence of active, or talk spurt, intervals, typically averaging 0.4–1.2 s in length, followed by silence intervals averaging 0.6–1.8 s in length. The durations of talk spurt and silence can be assumed to be exponentially distributed. This gives rise an MMPP with two states, with α representing the rate of transition from the salience to talk spurt and β the rate in the reverse direction. The source generates packets while in talk spurt with rate R. This voice model is widely used in the literature [6], [12], [17].

For a video stream, a natural model is to quantize the stream and then approximate the video signal by its quantized version [16]. it is found that eight quantization levels suffice to provide an accurate representation for a number of broadcast-quality (NTSC) video strips, as well as MPEG video sequences [13]. Therefore, a given video source can be represented by eight levels of traffic, changing from frame to frame. The packets for a given frame, and hence quantized level, are transmitted as a Poisson stream. Then we have a eight-state MMPP model for a video stream. There are a variety of ways to determine its parameters. One can calculate the video histogram and the autocovariance function for an actual sequence and then use these measured quantities to determine the transition probabilities. Alternately, one can measure the transition probabilities from an actual sequence and use these probabilities to calculate the steady-state probabilities and the resultant autocovariance function.

Recent measurement studies of packet traffic in local area networks (LANs) and the Internet show that the traffic in these networks appears to be *self-similar* [18], [19]. It looks the same regardless of time scales over a long range interval. Self-similar traffic is characterized by the property that correlation does not decay exponentially over time. Superpositions of two-state MMPPs can be used to model the self-similar traffic [14], [15]. Assume that $Q_1, Q_2, \ldots, Q_H$ are the infinitesimal generating matrices of underlying MMPPs. We can construct an MMPP with self-similar behavior over several time scales by superposing them to make a new MMPP with infinitesimal generating matrix $Q = Q_1 \oplus Q_2 \oplus \cdots \oplus Q_H$, where $\oplus$ is the Kronecker's sum. The parameters of MMPP are determined so as to fit the covariance function of the MMPP to that of second-order self-similar processes over several time scales [14], [15].

In this paper, we use the general form of the MMPP model, so that the proposed approaches in this study are general enough to be applicable to a variety of traffic in real networks.

B. Fading Channel and Linear Multiuser Detector Physical Layer Model

Physical layer aspects such as CDMA interference suppression algorithms and channel fading characteristics play important roles in designing CAC schemes. Consider a synchronous CDMA system with spreading gain N and K users. Each user generates MMPP traffic in the system. An important physical layer performance measure of class j users is the signal-to-interference ratio SIR_j, which should be kept above the target value ω_j. The signature sequences of all users are independent and randomly chosen. A class j connection in state m is assigned $R_{jm}(t)$ signature sequences and transmits at $R_{jm}(t)$ times the basic rate (obtained using the highest spreading gain N). We assume that the receiver estimates the channels of all users. The estimation is based on training data over the training period. Due to multipath fading, each user appears as L resolvable paths or components at the receiver. The path l of user k is characterized by its estimated average channel gain $\bar{h}_{kl}$ and its estimation error variance ξ_k^2. The lmmse detectors are used at the receiver to recover the transmitted information. In a large system (both N and K are large) with background noise σ^2, the SIR for the lmmse receiver of the a user (say, the first one) can be expressed approximately as [20]

$$\text{SIR}_1 = \frac{P_1 \sum_{l=1}^{L} |\bar{h}_{1l}|^2 \eta}{1 + P_1 \xi_1^2 \eta}, \tag{3}$$

where P_1 is the attenuated transmitted power from user 1, η is the unique fixed point in $(0, \infty)$ that satisfies

$$\eta = \left[\sigma^2 + \frac{1}{N} \sum_{k=2}^{K} \left((L-1) I\left(\xi_k^2, \eta\right) + I\left(\sum_{l=1}^{L} P_1 \left(|\bar{h}_{1l}|^2 + \xi_1^2 \right), \eta \right) \right) \right] \tag{4}$$

and

$$I(\nu, \eta) = \frac{\nu}{1 + \nu\eta}. \tag{5}$$

Assume that all users in the same class have the same average channel gain $|\bar{h}_j|^2 = \sum_{l=1}^{L} |\bar{h}_{jl}|^2, j = 1, 2, \ldots, J$. Authors in [5] show that a minimum received power solution exists such that all users in the system meet their target SIRs if and only if

$$\omega_j < \frac{|\bar{h}_j|^2}{\xi_j^2} \tag{6}$$

and

$$\frac{1}{N} \sum_{j=1}^{J} \sum_{m=1}^{M_j} n_{jm} R_{jm} \Upsilon_j < 1, \tag{7}$$

where n_{jm} is the number of class j users in state m and

$$\Upsilon_j = (L-1)\omega_j \frac{\xi_j^2}{|\bar{h}_j|^2} + \frac{\omega_j \left(1 + \frac{\xi_j^2}{|\bar{h}_j|^2}\right)}{1 + \omega_j}. \tag{8}$$

The minimum received power solution for a class 1 user is shown in (9) at the bottom of the page.

Remark 1: The above physical layer model can be extended to a multicell system. Consider the multicell model suggested by Wyner in [21], where the received signal at each site is the sum of the signals received from intracell users, plus a factor $\vartheta, 0 \leq \vartheta \leq 1$, times the sum of the signals of users in the adjacent cells, as received at their cell sites. Nonadjacent cell users are assumed to produce no interference. The lmmse receivers at each site treat all intercell interference as noise. Assume that there are K_a users in the adjacent cells. Let P_{k_a} be the received power from user k_a at her/his serving site. The intercell interference power is $\sum_{k=1}^{K_a} \vartheta^2 P_k$. Therefore, the total noise power in the local cell is $\sigma^2 + \sum_{k=1}^{K_a} \vartheta^2 P_k$, where σ^2 is the background noise. The local cell can obtain the information about K_a and P_{k_a} in the adjacent cells by exchanging signaling messages or by local estimation functions. For simplicity, we only consider the single-cell model in this paper.

Admission control in this paper is designed independently of access control [6]. The resulting admission control algorithms are concerned with a user's channel and power parameters at admission request time. These parameters may change after admission due to time variation in the channel characteristics. It is the role of access control to dynamically ensure that the QoS requirements are satisfied. Moreover, access control is responsible for scheduling packets and guarantee QoS requirements at packet level, such as packet delay and packet loss rates, which are not addressed in this paper. Decoupling of admission control and access control is a practically feasible and widely used methodology [6].

III. Constrained Semi-Markov Decision Process Formulation of CAC in Packet CDMA Networks

In this section, the admission control problem in packet CDMA networks is formulated as an average cost semi-Markov decision process (SMDP). We present a linear-programming-based solution to the admission control problem that maximizes network utilization subject to both SIR and blocking probability constraints. We call this scheme CDMA MMPP admission control I (CMAC-I). In order to utilize the linear programming algorithm to obtain the optimal solution, it is necessary to identify the state space, decision epochs, actions, state dynamics, cost, and constraints.

A. State Space

Define row vector $k(t) = [k_1(t), k_2(t), \ldots, k_J(t)] \in \mathbb{Z}_+^J$, where $k_j(t)$ denotes the number of class j connections in the system. Define row vector $n(t) = [n_{12}(t), n_{13}(t), \ldots, n_{1M_1}(t), n_{22}(t), n_{23}(t), \ldots, n_{2M_2}(t), \ldots, n_{J2}(t), \quad n_{J3}(t), \ldots, n_{JM_J}(t)] \in \mathbb{Z}_+^M$, where $n_{jm}(t)$ denotes the number of class j connections in state m and $M = \sum_{j=1}^{J}(M_j - 1)$. The state vector of the system at decision epoch t is given by

$$x(t) = [k(t), n(t)]. \tag{10}$$

Note that $n_{11}(t), n_{21}(t), \ldots, n_{J1}(t)$ are not included in the state vector, because these values can be obtained from $k(t)$ and $n(t)$ described earlier. For example, $n_{11}(t) = k_1(t) - \sum_{m=2}^{M_1} n_{1m}$. The state space X comprises of any state vector, such that SIR constraints can be met at all the time. The SIR expression in (3) is derived under the assumption that K/N is a constant. Although the ratio K/N in our case is changing due to user arrivals and departures, it is a constant when the system does not change states. Consequently, we can use the SIR expression in (3) at each state to restrict the state space. In addition, the number of class j connections in state m should not be greater than the total number of class j connections. Therefore, the state space X can be defined as

$$X = \left\{ x = [k, n] \in \mathbb{Z}_+^{J+M} : \frac{1}{N} \sum_{j=1}^{J} k_j R_{jM_j} \Upsilon_j < 1; \right.$$
$$\left. k_j \geq n_{jm}, m = 1, 2, \ldots, M_j, j = 1, 2, \ldots, J \right\}. \tag{11}$$

Since the arrivals and departures of connections, as well as the state changes of MMPP traffic are random, $\{x(t)\}_{t \in \mathbb{R}_+}$ is a finite-state stochastic process.

Remark 2: Because we consider packet traffic in this paper, vector $n(t)$ is included in the state vector $x(t)$. This does not happen in previous approaches [4], [5], where only constant bit rate traffic and circuit-switched CDMA networks are studied. Note that we can also use buffers to queue the connection arrivals when the radio resources are not available. In this case, another vector denoting the number of connections in the queues should be included in the state vector, as is done in [4], [5]. To highlight the packet traffic considered in this paper, we do not

$$P_1 = \frac{\omega_1 \sigma^2}{|\bar{h}_1|^2 \left(1 - \omega_1 \frac{\xi_1^2}{|\bar{h}_1|^2}\right) \left(1 - \frac{1}{N} \sum_{j=1}^{J} \sum_{m=1}^{M_j} n_{jm} R_{jm} \Upsilon_j\right)}. \tag{9}$$

consider the queueing of connection arrivals. However, the proposed scheme in this paper can be extended straightforwardly to the systems with connection arrival queues.

B. Decision Epochs and Actions

When a user requests a connection, the network will make an admission decision. Therefore, the natural decision epochs for our model are the connection arrival points. However, each time when a connection arrival or departure occurs, the state of the system changes. Also, the state of the system also changes when an MMPP connection changes states. Therefore, similar to [4], [5], we choose the decision epochs to be the set of all connection arrival and departure instances, as well as the instances when an MMPP traffic changes states.

At each decision epoch $t_k, k = 0, 1, 2, \ldots$, the network makes a decision for each possible user arrival that may occur in the time interval $(t_k, t_{k+1}]$. In each state, there is at most two actions to choose from (accept or reject) for each class of traffic. Action a at decision epoch t is defined as

$$a(t) = [a_1(t), a_2(t), \ldots, a_J(t)] \tag{12}$$

where $a_j(t)$ denotes the action for class j connections. If $a_j(t) = 1$, admit a class j connection. If $a_j(t) = 0$, the connection is rejected. The action space is a set of all possible actions, which can be defined as

$$A = \{a : a \in \{0,1\}^J, j = 1, 2, \ldots, J\}. \tag{13}$$

For a given state $x \in X$, a selected action should not result in a transition to a state that is not in X. In addition, action $(0, 0, \ldots, 0)$ should not be a possible action in state $(0, 0, 0, \ldots, 0)$. Otherwise, new connections are never admitted into the network and the system cannot evolve. The action space A_x of a given state $x \in X$ is defined as

$$\begin{aligned} A_x = \big\{ & a \in A : a_j = 0 \\ & \text{if } \left[\left(k + e_j^u\right), \left(n + e_{jm}^s\right)\right] \notin X, \\ & j = 1, 2, \ldots, J, m = 2, 3, \ldots, M_j \\ & \text{and } a \neq (0, 0, \ldots, 0) \text{ if } x = (0, 0, 0, \ldots, 0) \big\} \end{aligned} \tag{14}$$

where $e_j^u \in \{0,1\}^J$ denotes a row vector containing only zeros except for the jth component, which is 1, and $e_{jm}^s \in \{0,1\}^M$ denotes a row vector containing only zeroes expect for the $(\sum_{i=1}^{j-1} \sum_{n=1}^{M_i} 1 + m)$th component, which is 1. $k + e_j^u$ corresponds to an increase of the number of class j users by 1. $n + e_{jm}^s$ corresponds to an increase of the number of class j users in state m by 1.

C. State Dynamics

The state dynamics of packet CDMA networks can be characterized by the state transition probabilities of the embedded chain $p_{xy}(a)$ and the expected sojourn time $\tau_x(a)$ for each state-action pair. Since connection arrival, departure and MMPP of each connection are mutually independent Poisson process, the cumulative process is also Poisson. The resulting process consists of a connection arrival process with rate $\sum_{j=1}^J \lambda_j$, if a class j connection can be admitted (i.e., $a_j = 1$), a connection departure process with rate $\sum_{j=1}^J \mu_j k_j$, an MMPP-state-changing process with rate $\sum_{j=1}^J \sum_{m=1}^{M_j} \alpha_{jm}(k_j - n_{jm}) + \sum_{j=1}^J \sum_{m=1}^{M_j} \beta_{jm} n_{jm}$. The cumulative event rate is the sum of the rates of all constituent processes and the expected sojourn time is the inverse of the event rate

$$\begin{aligned} \tau_x(a) = \Bigg[& \sum_{j=1}^J \lambda_j a_j + \sum_{j=1}^J \mu_j k_j \\ & + \sum_{j=1}^J \sum_{m=1}^{M_j} \alpha_{jm}(k_j - n_{jm}) + \sum_{j=1}^J \sum_{m=1}^{M_j} \beta_{jm} n_{jm} \Bigg]^{-1}. \end{aligned} \tag{15}$$

The state transition probabilities of the embedded chain is shown in (16) at the bottom of the page.

D. Policy, Performance Criterion, and Cost Function

For each given state $x \in X$, an action $a \in A_x$ is chosen according to a policy $u_x \in \mathcal{U}$, where $\mathcal{U}$ is a set of admissible policies defined as

$$\mathcal{U} = \{u: X \to A \,|\, u_x \in A_x, \forall x \in X\}. \tag{17}$$

The average cost criterion is considered as the performance criterion in this paper. For any policy $u \in \mathcal{U}$ and an initial state x_0, the average cost is defined as

$$J_u(x_0) = \lim_{T \to \infty} \frac{1}{T} \mathbf{E} \left\{ \int_0^T c(x(t), a(t))\, dt \right\} \tag{18}$$

where $\mathbf{E}$ is the expectation and $c(x(t), a(t))$ is the expected cost until the next decision epoch when $a(t)$ is selected at state $x(t)$. The aim is to find an optimal policy u^* that minimizes $J_u(x_0)$ for any initial state x_0. We assume that the embedded chain considered in this paper is a unichain, which is a common assumption in the CAC context [4], [5]. With the unichain assumption, an optimal policy exists and can be obtained by solving the linear program associated with the SMDP.

$$p_{xy}(a) = \begin{cases} \lambda_j a_j r_{jm} \tau_x(a), & \text{if } y = x + e_j^u + e_{jm}^s, \quad m = 2, 3, \ldots, M_j, \quad j = 1, 2, \ldots, J \\ \lambda_j a_j r_{j1} \tau_x(a), & \text{if } y = x + e_j^u, \quad j = 1, 2, \ldots, J \\ \mu_j n_{jm} \tau_x(a), & \text{if } y = x - e_j^u - e_{jm}^s, \quad m = 2, 3, \ldots, M_j, \quad j = 1, 2, \ldots, J \\ \mu_j n_{j1} \tau_x(a), & \text{if } y = x - e_j^u, \quad j = 1, 2, \ldots, J \\ \alpha_{jm}(k_j - n_{jm}) \tau_x(a), & \text{if } y = x + e_{jm}^s, \quad m = 2, 3, \ldots, M_j, \quad j = 1, 2, \ldots, J \\ \beta_{jm} n_{jm} \tau_x(a), & \text{if } y = x - e_{jm}^s, \quad m = 2, 3, \ldots, M_j, \quad j = 1, 2, \ldots, J \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

Based on the action a taken in a state x, a cost $c(x,a)$ occurs to the network. Authors in [4] show that the blocking probability can be expressed as an average cost criterion in the CAC setting. Minimizing blocking probability is equivalent to maximizing network utilization. Since the blocking probability expression in this paper is similar to that in [4], we define the cost as

$$c(x,a) = \sum_{j=1}^{J} w_j(1-a_j) \tag{19}$$

where $w_j \in \mathbb{R}_+$ is the weight associated with class j.

E. Constraints

In the current problem formulation, SIR constraints of all connections in the system can be guaranteed at all time instants by restricting the state space in (11). In addition, it is desirable for the network operator to put constraints on blocking probabilities of certain classes of traffic. Therefore, we need to formulate connection blocking probability constraints in our model. Since we have derived the expected sojourn time $\tau_x(a)$ for a given state-action pair, the blocking probability for class j can be defined as the fraction of time the system is in a set of states $X_j^b \subset X$ and the chosen action is in a set of actions $A_{x_j^b} \subset A$, where $x_j^b \in X_j^b$ and $A_{x_j^b} = \{a \in A : a_j = 0\}$,

$$\begin{aligned} P_j^b &= \lim_{T\to\infty} \frac{1}{T}\mathbf{E}\left\{\int_0^T (1-a_j(t))\tau_{x(t)}(a(t))\,dt\right\} \\ &= \frac{\sum_{x\in X_j^b}\sum_{a\in A_{x_j^b}} \tau_x(a)}{\sum_{x\in X}\sum_{a\in A}\tau_x(a)}. \end{aligned} \tag{20}$$

The constraints related to the blocking probability can be expressed as

$$P_j^b \le \gamma_j, \quad j=1,2,\ldots,J. \tag{21}$$

The blocking probability constraints can be easily addressed in the linear programming formulation by defining a cost function related to these constraints

$$c_j^b(x,a) = (1-a_j), \quad j=1,2,\ldots,J. \tag{22}$$

F. Linear Programming Solution to the SMDP

Due to the constraints in the above SMDP formulation, it is natural to use the linear programming methodology to compute the optimal policy.

The optimal policy u^* of the SMDP is obtained by solving the following linear program.

$$\min_{z_{xa}\ge 0, x\in X, a\in A_x} \sum_{x\in X}\sum_{a\in A_x}\sum_{j=1}^{J} w_j(1-a_j)\tau_x(a) z_{xa}$$

subject to

$$\begin{aligned} &\sum_{a\in A_y} z_{ya} - \sum_{x\in X}\sum_{a\in A_x} p_{xy}(a) z_{xa} = 0, \quad y\in X, \\ &\sum_{x\in X}\sum_{a\in A_x} z_{xa}\tau_x(a) = 1, \\ &\sum_{x\in X}\sum_{a\in A_x} (1-a_j) z_{xa}\tau_x(a) \le \gamma_j, \quad j=1,2,\ldots,J. \end{aligned} \tag{23}$$

The decision variables are $z_{xa}, x \in X, a \in A_x$. The term $z_{xa}\tau_x(a)$ can be interpreted as the steady-state probability of the system being in state x and a is chosen. The first constraint is a balance equation and the second constraint can guarantee that the sum of the steady-state probabilities to be one. The network layer blocking probability constraints are expressed in the third one. Since sample path constraints are included in (23), the optimal policy obtained will be a randomized policy: The optimal action $a^* \in A_x$ for state x is chosen probabilistically according to the probabilities $z_{xa}/\sum_{a\in A_x} z_{xa}$.

IV. Guaranteed SIR Outage Probability for Packet-Switched CDMA Networks

The formulation described earlier can guarantee the SIR of all connections in the system at all time instants, which is desirable from mobile users' point of view. However, as we will show in Example 1 below, this formulation will result in low network utilization, especially when the MMPP traffic is bursty in data services. Therefore, we propose another formulation in this section to address the CAC problem, which we call CMAC-II. The main idea is that instead of guaranteeing the SIR at all time instants, we guarantee the SIR outage probability. This formulation is motivated by the design of packet-switched wireline networks. It is well known [16] that allocating a connection with its peak bandwidth guarantees no packet loss, but results in the lowest network utilization and no multiplexing gain. Therefore, most bandwidth allocation and admission control schemes in wireline networks allow small packet loss probability to increase the network utilization [16]. Similarly, since most applications in wireless networks can tolerate small probability of SIR outage, we introduce the SIR outage probability in wireless CDMA networks. Similar to the blocking probability, the SIR outage probability can be given as the fraction of time the system is in a set of states $X^{\text{out}} \subset X$ and the chosen action is in a set of actions $A_{x^{\text{out}}} \subset A$ such that for a given state $x^{\text{out}} \in X^{\text{out}}$, an action $a \in A_{x^{\text{out}}}$ results in a transition into a state that the SIR is outage, i.e., $(1/N)\sum_{j=1}^{J}\sum_{m=1}^{M_j}(n_{jm}+a_j)R_{jm}\Upsilon_j > 1$

$$\begin{aligned} P^{\text{out}} &= \lim_{T\to\infty}\frac{1}{T}\mathbf{E}\left\{\int_0^T \delta\left(\frac{1}{N}\sum_{j=1}^{J}\sum_{m=1}^{M_j}(n_{jm}(t) \right.\right. \\ &\qquad \left.\left. + a_j(t))R_{jm}(t)\Upsilon_j - 1\right)\tau_{x(t)}(a(t))\,dt\right\} \\ &= \frac{\sum_{x\in X^{\text{out}}}\sum_{a\in A_{x^{\text{out}}}}\tau_x(a)}{\sum_{x\in X}\sum_{a\in A}\tau_x(a)} \end{aligned} \tag{24}$$

where

$$\delta(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \ge 0. \end{cases} \tag{25}$$

The constraints related to the SIR outage probability can be expressed as

$$P^{\text{out}} \le \zeta. \tag{26}$$

In order to introduce the SIR outage probability in CMAC-II, we need to redefine the state space, constrains and the linear programming formulation. The actions, state dynamics, performance criterion and cost function are the same as those in Section III.

The state space in the CMAC-II is defined as

$$X = \left\{ x = [k, n] \in \mathbb{Z}_+^{J+M} : \sum_{j=1}^{J} k_j \le G, n_{jm} \le k_j, \right. \\ \left. m = 2, 3, \ldots, M_j, j = 1, 2, \ldots, J \right\}, \quad (27)$$

where G is a fixed large positive integer used to restrict the state space to be finite. The cost function related to the constraint of SIR outage probability is

$$c^{\text{out}}(x, a) = \delta \left(\frac{1}{N} \sum_{j=1}^{J} \sum_{m=1}^{M_j} (n_{jm} + a_j) R_{jm} \Upsilon_j - 1 \right). \quad (28)$$

The linear program associated with CMAC-II becomes

$$\min_{z_{xa} \ge 0, x \in X, a \in A_x} \sum_{x \in X} \sum_{a \in A_x} \sum_{j=1}^{J} w_j (1 - a_j) \tau_x(a) z_{xa}$$

subject to

$$\begin{aligned} &\sum_{a \in A_y} z_{ya} - \sum_{x \in X} \sum_{a \in A_x} p_{xy}(a) z_{xa} = 0, \quad y \in X \\ &\sum_{x \in X} \sum_{a \in A_x} z_{xa} \tau_x(a) = 1, \\ &\sum_{x \in X} \sum_{a \in A_x} (1 - a_j) z_{xa} \tau_x(a) \le \gamma_j, \quad j = 1, 2, \ldots, J \\ &\sum_{x \in X} \sum_{a \in A_x} \delta \left(\frac{1}{N} \sum_{j=1}^{J} \sum_{m=1}^{M_j} (n_{jm} + a_j) R_{jm} \Upsilon_j - 1 \right) \\ &\quad \times z_{xa} \tau_x(a) \le \zeta. \end{aligned} \quad (29)$$

The fourth constraint in (29) requires that the SIR outage probability should be below the target value.

By introducing a small SIR outage probability, we can increase the network utilization significantly. We show this by the following example.

Example 1: For simplicity, we assume there is one class of two-state ON/OFF MMPP traffic in a CDMA network with spreading gain $N = 128$. The transmission rate when the MMPP traffic in the ON state is R_{on} corresponding to an equivalent spreading gain 32. No packet is transmitted in the OFF state. The rate from OFF to ON is $\alpha = 0.03$ and the rate from ON to OFF is $\beta = 0.1$. The stationary probability that the connection is in ON state is $p_{\text{on}} = \alpha/(\alpha + \beta)$. The target SIR is $\omega = 10$. The channel parameters are chosen as follows: $L = 1$ (flat fading), $|\bar{h}|^2 = 1$, $\xi^2 = 0.05$. In this system, if we guarantee the SIR at all time instants, it is not difficult to calculate from (7) that the maximum number of connections that can be admitted is 33. If we allow the SIR outage in the system but guarantee the SIR outage probability below 0.001, the maximum number of connections that can be admitted is 89, which is more than two times of the number of admissible connections without SIR outage. The calculation of this number is shown as follows. Assume that there are A connections in the system. The probability that there are B connections in the ON state is

$$\phi(A, B) = \binom{A}{B} (p_{\text{on}})^B (1 - p_{\text{on}})^{A-B}.$$

The SIR outage probability is

$$\psi(A) = \sum_{n=0}^{A} \phi(A, n) \delta(n R_{\text{on}} \Upsilon - 1)$$

where Υ is defined in (8). Given a SIR outage probability constraint, we can calculate the maximum number of admissible connections. Next, we compare the maximum achievable utilizations with and without SIR outage. Define the network utilization as the ratio between the carried traffic in the network during unit time and the maximum network capacity. Fig. 4 shows the maximum achievable utilization with three values of SIR outage probability, 0, 0.01, and 0.001. Fig. 5 shows the maximum achievable utilization versus SIR outage probability allowed in the system. We can see that the utilization can be increased significantly by introducing the SIR outage probability, especially when the MMPP traffic is bursty (p_{on} is small).

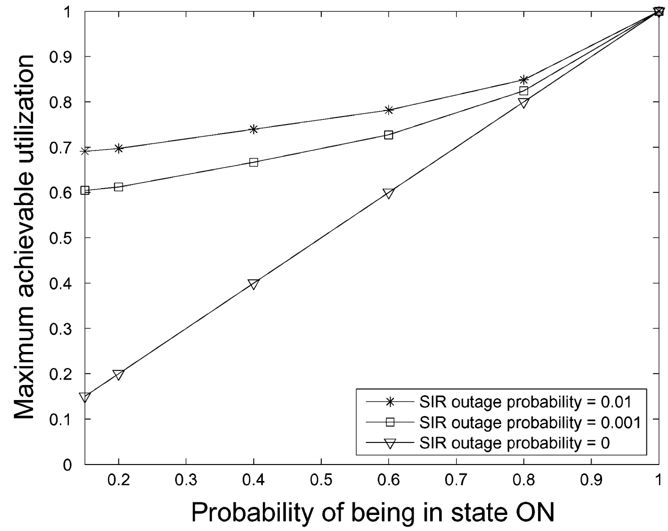

Fig. 4. Maximum achievable utilization with and without SIR outage.

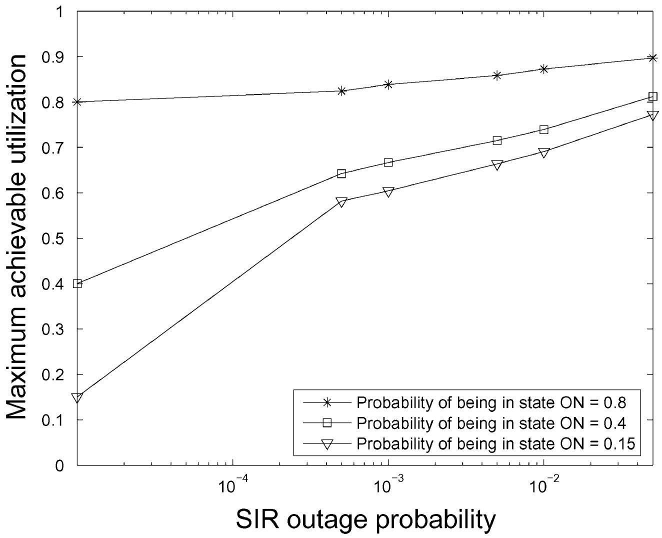

Fig. 5. Maximum achievable utilization versus SIR outage probability.

V. Nearly Complete Decomposable (NCD) SMDP

The SMDPs introduced in CMAC-I and CMAC-II are $(J+\sum_{j=1}^{J}\sum_{m=1}^{M_j} m)$ dimensional for a J class system. For large J, this leads to computational problems of excessive size. The notion of NCD [22] Markov chains can be used to reduce this to a J-dimensional problem. Intuitively, when making a connection admission decision, the number of connections of each class in progress is important, but the number of connections of each class in the different states is not, because these quantities oscillate too rapidly. Mathematically, the embedded Markov chain of the SMDP in packet-switched CDMA networks has two time-scale structure—it spends most of its time jumping in the k component, and only rarely jumps in the n component of the state descriptor (10). The main idea here is to show that the SMDP that models a packet-switched network is NCD.

Let S denote the total number of states in the embedded Markov chain of the SMDP. Mathematically, the transition probability matrix of the embedded Markov chain has the following NCD structure

$$Q = B + \epsilon C \tag{30}$$

where B has a block diagonal structure.

$$B = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ 0 & B_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{HH} \end{bmatrix} \tag{31}$$

where $B_{ii} \in \mathbb{R}^{s_i \times s_i}$, $\forall i$, $\sum_i s_i = S, \epsilon > 0$ is the *maximum degree of coupling*, and $C \in \mathbb{R}^{S\times S}$. $B_{ii}, \forall i$ are also infinitesimal generators. Denote the state partitions as $\mathcal{X}_1 = (x_1, x_2, \ldots, x_{s_1}), \mathcal{X}_2 = (x_{s_1+1}, x_{s_1+2}, \ldots, x_{s_1+s_2})$ and so on. The "super-states" $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_H$ $(H < S)$ will be called as *macro-states*. In addition, the elements of k in the state vector are identical within each macro-state. The above structure (30) of Q implies that the embedded Markov chain has a two time-scale structure—it spends most of its time jumping between states in a particular macro-state, and only rarely jumps between macro-states.

A process of aggregation can be applied to the S-state nearly completely decomposable Markov chain to approximate it with a H-state Markov chain. The approximation accuracy can be stated as follows.

Theorem 1: Consider an S-state NCD Markov chain with H macro-states and transition probability matrix $Q \in \mathbb{R}^{S\times S}, Q = B + \epsilon C$, as defined in (30). Let $Q^* \in \mathbb{R}^{H\times H}$ and $\pi_{Q^*} = [\pi_{Q^*1}, \pi_{Q^*2}, \ldots, \pi_{Q^*H}] \in \mathbb{R}^H$ denote the state transition probability matrix and the steady-state probability vector of the aggregated Markov chain with H states, respectively. $\pi_{Q^*h}, h = 1, 2, \ldots, H$, is $O(\epsilon)$ approximation of the stationary probability of the NCD Markov chain being in macro-state h.

For a Proof of Theorem 1, see [23, Chapters I and II].

We show by an example of this NCD structure in the problem considered in this paper.

Example 2: For simplicity again, we assume that there is one class of MMPP traffic with arrival rate λ and service rate μ. The MMPP has two states, x_1 and x_2. The initial state is x_1 with probability 1. The rate from x_1 to x_2 is β and the rate from x_2 to x_1 is α. In addition, we assume that at most 2 connections are allowed in the system. As shown before, the state of this system can be defined as a doublet (k, n), where k denotes the number of connections in the system and n denotes the number connections in state x_2 among k connections. The resultant two-dimensional state space, with transitions between state superimposed, appears in Fig. 6. The infinitesimal generator of the embedded Markov chain is shown in the equation at the bottom of the page. This matrix can be written as

$$Q = B + \epsilon C$$

where B has a block diagonal structure

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\alpha & \alpha & 0 & 0 & 0 \\ 0 & \beta & -\beta & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\alpha & 2\alpha & 0 \\ 0 & 0 & 0 & \beta & -\beta-\alpha & \alpha \\ 0 & 0 & 0 & 0 & 2\beta & -2\beta \end{bmatrix} = \begin{bmatrix} B_{11} & 0 & \cdots \\ 0 & B_{22} & 0 \\ \cdots & 0 & B_{33} \end{bmatrix}$$

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 \\ \mu & -\alpha-\mu-\lambda & \alpha & \lambda & 0 & 0 \\ \mu & \beta & -\beta-\lambda & 0 & \lambda & 0 \\ 0 & 2\mu & 0 & -2\mu-2\alpha-\lambda & 2\alpha & 0 \\ 0 & \mu & 2\mu & \beta & -2\mu-\beta-\alpha-\lambda & \alpha \\ 0 & 0 & 2\mu & 0 & 2\beta & -2\beta-2\mu \end{bmatrix}.$$

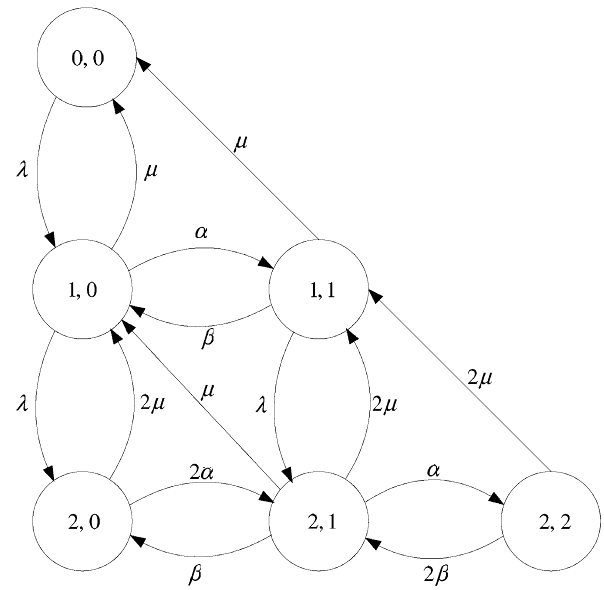

Fig. 6. Embedded Markov chain in Example 2.

where (see the equation at the bottom of the page), where $\Gamma = (\lambda + 2\mu)(\lambda + 3\mu)$.

We can see that $B_{11} \in \mathbb{R}^{1\times 1}$, $B_{22} \in \mathbb{R}^{2\times 2}$ and $B_{33} \in \mathbb{R}^{3\times 3}$ are also infinitesimal generators. Denote the sate partitions as macro-state $\mathcal{X}_1 = \{(0,0)\}, \mathcal{X}_2 = \{(1,0),(1,1)\}$, $\mathcal{X}_3 = \{(2,0),(2,1),(2,2)\}$. Note that the number of connections (k) in the system does not change within a macro-state, but the number of connections in the x_2 state (n) does. In addition, because $1/\lambda$ and $1/\mu$ are the mean connection interarrival and service times, which are usually large values compared to $1/\alpha$ and $1/\beta$, the coupling parameter $\epsilon = \lambda\mu(\lambda+\mu)(\lambda+2\mu)(\lambda+3\mu)$ is small. For example, assume that $\alpha = \beta = 0.4$ and $\mu = 0.004$ (these values are reasonable for voice traffic), if $\lambda = 0.2$ (this is a quite large arrival rate that results in about 60% blocking probability, as shown in numerical examples), the couple parameter ϵ is about 7.2×10^{-6}, which is relatively small compared to α and β.

Example 2 can be generalized to multiple classes of MMPP traffic with multiple states in each class. Specifically, we can partition the state space (11), (27) into macro-states. The k component of the state descriptor is the same within a given macro-state but different from one macro-state to another. The n component is different within a given macro-state. Moreover, because the cost function defined in (19) is related to the k component, the number of connections that can be admitted to the system, the cost incurred by an action a is the same within a macro-state.

VI. Reduced Complexity Constrained SMDP Formulation of CAC

Since the embedded Markov chain of the SMDP formulated in CMAC-I and CMAC-II for packet-switched CDMA networks has NCD structure, in this section, we formulate the CAC problem as a reduced complexity constrained SMDP by exploiting this NCD structure.

A. States, Decision Epochs, Actions and State Dynamics

In the reduced complexity problem, only the number of connections in each class appears in the state vector of the system, which is given by

$$x(t) = k(t) = [k_1(t), k_2(t), \ldots, k_J(t)]. \tag{32}$$

As in the original problem, we choose the decision epochs to be the set of instances when the state of the system changes, i.e., the connection arrival and departure instances. The actions that can be chosen is the same as before, accept and reject for each class of traffic. The state dynamics of the reduced complexity problem becomes much simpler than those of the original problem. The resulting process consists of only an arrival process with rate $\sum_{j=1}^{J} \lambda_j$, if a class j connection be admitted and a departure process with rate $\sum_{j=1}^{j} \mu_j k_j$. The sojourn time is

$$\tau_x(a) = \left[\sum_{j=1}^{J} \lambda_j a_j + \sum_{j=1}^{J} \mu_j k_j\right]^{-1}. \tag{33}$$

The state transition probabilities of the embedded chain is

$$p_{xy}(a) = \begin{cases} \lambda_j a_j \tau_x(a), & \text{if } y = x + e_j^u, \quad j = 1,2,\ldots,J \\ \mu_j k_j \tau_x(a) & \text{if } y = x - e_j^u, \quad j = 1,2,\ldots,J \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

$$B_{11} = [0], \quad B_{22} = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}, \quad B_{33} = \begin{bmatrix} -2\alpha & 2\alpha & 0 \\ \beta & -\beta-\alpha & \alpha \\ 0 & 2\beta & -2\beta \end{bmatrix}, \quad \epsilon = \lambda\mu(\lambda+\mu)(\lambda+2\mu)(\lambda+3\mu)$$

and

$$C = \begin{bmatrix} \frac{-1}{\mu(\lambda+\mu)\Gamma} & \frac{1}{\mu(\lambda+\mu)\Gamma} & 0 & 0 & 0 & 0 \\ \frac{1}{\lambda(\lambda+\mu)\Gamma} & \frac{-1}{\lambda\mu\Gamma} & 0 & \frac{1}{\mu(\lambda+\mu)\Gamma} & 0 & 0 \\ \frac{1}{\lambda(\lambda+\mu)\Gamma} & 0 & \frac{-1}{\mu(\lambda+\mu)\Gamma} & 0 & \frac{1}{\mu(\lambda+\mu)\Gamma} & 0 \\ 0 & \frac{2}{\lambda(\lambda+\mu)\Gamma} & 0 & \frac{-1}{\lambda\mu(\lambda+\mu)(\lambda+3\mu)} & 0 & 0 \\ 0 & \frac{1}{\lambda(\lambda+\mu)\Gamma} & \frac{2}{\lambda(\lambda+\mu)\Gamma} & 0 & \frac{-1}{\lambda\mu(\lambda+\mu)(\lambda+2\mu)} & 0 \\ 0 & 0 & \frac{2}{\lambda(\lambda+\mu)\Gamma} & 0 & 0 & \frac{-2}{\lambda(\lambda+\mu)\Gamma} \end{bmatrix}$$

In the reduced complexity problem, the connection blocking probability constraints can be considered the same way as before.

The above formulation in the reduced complexity problem is similar to those in circuit-switched systems [4], [5]. However, SIR outage is not considered in circuit-switched systems, which will result in low network utilization for packet traffic. By exploiting the MMPP packet traffic models, we propose two approaches to the constraint of SIR outage probability, conservative approach (CMAC-III) and aggressive approach (CMAC-IV), which are described later.

B. The Conservative Approach to the Constraint of the SIR Outage Probability

In the conservative approach, which we call CMAC-III, the SIR outage probability (24) is guaranteed in each state and the system will never go into any state for any period of time where the SIR outage probability will be violated. Before defining the state space of the reduced complexity problem in this approach, we need to derive the SIR outage probability in a given state.

The stationary state probabilities in the original problem with state vector $[k, n]$ is

$$\phi(k,n) = \prod_{j=1}^{J}\prod_{m=1}^{M_j} \binom{k_j - \sum_{l=1}^{m-1} n_{jl}}{n_{jm}} (\pi_{jm})^{n_{jm}} \tag{35}$$

where π_{jm} is the stationary state probabilities of class j MMPP being in state m, which is defined in (1) of Section II. The SIR outage probability in the reduced complexity problem for a given state $x = k$ is shown in (36) at the bottom of the page. The state space of the reduced complexity problem is restricted such that the SIR outage probability of any given state is not violated.

$$X = \{x = [k_1, k_2, \ldots, k_J] \in \mathbb{Z}_+^J : \psi(k) \le \zeta\}. \tag{37}$$

The reduced complexity policy can be obtained from the solution of the following linear program.

$$\min_{z_{xa} \ge 0, x\in X, a \in A_x} \sum_{x\in X}\sum_{a\in A_x}\sum_{j=1}^{J} w_j(1-a_j)\tau_x(a) z_{xa}$$

subject to

$$\begin{aligned}
&\sum_{a\in A_y} z_{ya} - \sum_{x\in X}\sum_{a\in A_x} p_{xy}(a) z_{xa} = 0, \quad y \in X\\
&\sum_{x\in X}\sum_{a\in A_x} z_{xa}\tau_x(a) = 1\\
&\sum_{x\in X}\sum_{a\in A_x} (1-a_j) z_{xa}\tau_x(a) \le \gamma_j, \quad j = 1,2,\ldots,J.
\end{aligned} \tag{38}$$

C. The Aggressive Approach to the Constraint of the SIR Outage Probability

In the aggressive approach, which we call CMAC-IV, the SIR outage probability is satisfied over long periods of time, but can be violated in some states for short periods. Similar to the definition (24) in CMAC-II, the SIR outage probability in CMAC-IV is given as

$$\begin{aligned}
P_{\text{reduced}}^{\text{out}} &= \lim_{T\to\infty} \frac{1}{T}\mathbf{E}\left\{\int_0^T \psi(x)\tau_{x(t)}(a(t))\,dt\right\}\\
&= \sum_{x\in X}\sum_{a\in A} \psi(x)\tau_x(a).
\end{aligned} \tag{39}$$

The state space of the reduced complexity problem is defined as

$$X = \left\{x = [k_1, k_2, \ldots, k_J] \in \mathbb{Z}_+^J : \sum_{j=0}^{J} k_j \le G\right\} \tag{40}$$

where G is a large positive integer to restrict the state space to be finite. The linear program for obtaining the reduced complexity policy becomes

$$\min_{z_{xa} \ge 0, x\in X, a \in A_x} \sum_{x\in X}\sum_{a\in A_x}\sum_{j=1}^{J} w_j(1-a_j)\tau_x(a) z_{xa}$$

subject to

$$\begin{aligned}
&\sum_{a\in A_y} z_{ya} - \sum_{x\in X}\sum_{a\in A_x} p_{xy}(a) z_{xa} = 0, \quad y \in X\\
&\sum_{x\in X}\sum_{a\in A_x} z_{xa}\tau_x(a) = 1\\
&\sum_{x\in X}\sum_{a\in A_x} (1-a_j) z_{xa}\tau_x(a) \le \gamma_j, \quad j = 1,2,\ldots,J\\
&\sum_{x\in X}\sum_{a\in A_x} \psi(x,a) z_{xa}\tau_x(a) \le \zeta.
\end{aligned} \tag{41}$$

VII. NUMERICAL EXAMPLES—VOICE AND *STAR WARS* VIDEO TRAFFIC

In this section, we illustrate the performance of the proposed admission control schemes by numerical examples. Two classes of voice traffic and one class of video traffic are considered with arrival rates λ_1, λ_2 and λ_3, respectively. The service rates are μ_1, μ_2 and μ_3. Each voice traffic is modeled as an MMPP that has two states with the transmission rates in one state being zero. The transmission rates in another state for both voice classes are equal and correspond to an equivalent spreading gain $N = 16$, which can be interpreted as multiple code transmission using a higher spreading gain (say, $N = 128$) to ensure the accuracy of the asymptotic approximation of CDMA systems with lmmse

$$\psi(x) = \psi(k) = \sum_{n_{11}=0}^{k_1}\sum_{n_{12}=0}^{k_1}\cdots\sum_{n_{1M_1}=0}^{k_1}\sum_{n_{21}=0}^{k_2}\sum_{n_{22}=0}^{k_2}\cdots\sum_{n_{2M_2}=0}^{k_2}\cdots\sum_{n_{JM_J}=0}^{k_J} \phi(k,n)\delta\left(\frac{1}{N}\sum_{j=1}^{J}\sum_{m=1}^{M_j} n_{jm}R_{jm}\Upsilon_j - 1\right). \tag{36}$$

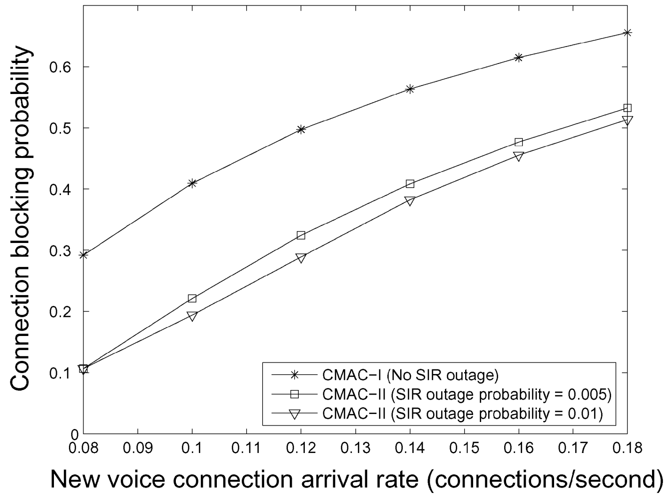

Fig. 7. Voice connection blocking probabilities in CMAC-I and CMAC-II.

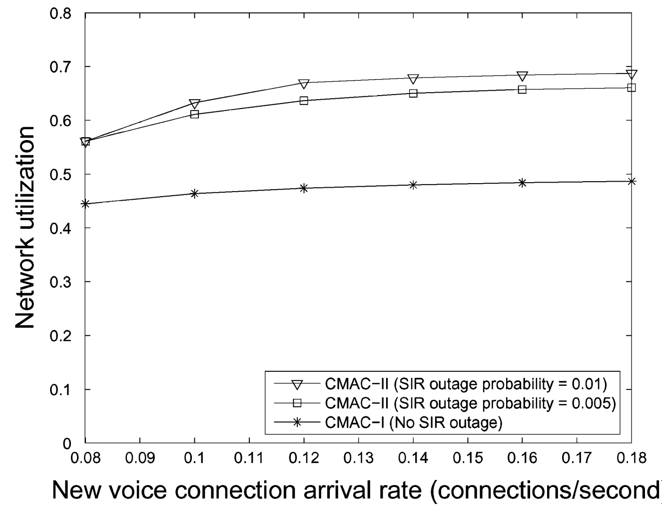

Fig. 8. Network utilizations in CMAC-I and CMAC-II.

receivers in (3). The rate from x_{j1} to x_{j2} is α_j and the rate from x_{j2} to x_{j1} is $\beta_j, j = 1, 2$, for the voice traffic. An MMPP with eight states is used to model the video traffic. We choose the infinitesimal generating matrix in [13, Table III] for a full motion coded movie *Star Wars* using discrete cosine transform (DCT) and Huffman coding techniques [24], which is a benchmark used in several papers studying video traffic. [See (42) at the bottom of the page.]

The transmission rate R_{31} in the first state of the video traffic corresponds to an equivalent spreading gain $N = 32$. The transmission rates in other states are: $R_{3m} = mR_{31}, m = 2, 3, \ldots, 8$. The target SIRs for three classes are equal, $\omega_1 = \omega_2 = \omega_3 = 10$. The channel parameters are chosen as follows: $L = 1$ (flat fading), $|\bar{h}_1|^2 = |\bar{h}_2|^2 = |\bar{h}_3|^2 = 1$, $\xi_1^2 = 0.02$, $\xi_2^2 = \xi_3^2 = 0.05$.

A. Effects of SIR Outage Probability

This subsection considers the effects of introducing SIR outage probability on both connection blocking probability and network utilization. We compare the connection blocking probabilities and the network utilizations in CMAC-I and those in CMAC-II. Since SIR outage is not allowed in both CMAC-I and the schemes in [4] and [5], they have the same performance. Due to the large state space with video traffic, only voice traffic is considered in CMAC-I and CMAC-II. Figs. 7 and 8 show the blocking probabilities and the network utilizations, respectively. In this example, the following parameters are chosen in (23) and (29): $\lambda_1 = \lambda_2$, $\mu_1 = 0.004$, $\mu_2 = 0.005$. $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0.4$. Two SIR constraints are considered for CMAC-II, 0.005 and 0.01. No blocking probabilities constraints are applied here. We can see that with a small SIR outage probability (say, 0.005) allowed in the system, the connection blocking probability can be reduced and the network utilization can be increased significantly. This is because the peak transmission rate is allocated in CMAC-I if a connection is admitted to guarantee the SIR for all connections at all the time. In CMAC-II, multiplexing gain can be achieved by introducing the SIR outage probability. It can also be observed that the larger the SIR outage is allowed, the smaller the blocking probability and the larger the network utilization can be obtained in CMAC-II.

B. NCD Structure and Randomized Policy

Here, we show the NCD structure of the CAC problem in packet-switched CDMA networks and the randomized policy due to the constraints. Fig. 9 shows an admission policy for class one traffic obtained in CMAC-II. The constraint of the connection blocking probability is 0.05. If the admission decision is 1, a connection can be admitted. A connection is rejected if the admission decision is 0. The value between 1 and 0 indicates a randomized policy. From this figure, we can see the policy is a randomized policy when the system is in state $[10, 5]$. We can also observe the dependence of the admission policy on the k

$$\Theta_3 = \begin{bmatrix} -4.737 & 4.105 & 0 & 0.316 & 0 & 0.316 & 0 & 0 \\ 0.643 & -1.714 & 0.771 & 0.043 & 0.129 & 0.043 & 0.086 & 0 \\ 0 & 1.408 & -2.254 & 0.845 & 0 & 0 & 0 & 0 \\ 0 & 0.082 & 0.74 & -1.562 & 0.616 & 0.082 & 0 & 0.041 \\ 0 & 0 & 0.111 & 0.741 & -1.333 & 0.407 & 0.074 & 0 \\ 0 & 0 & 0 & 0.095 & 1.518 & -3.13 & 1.423 & 0.095 \\ 0 & 0 & 0 & 0 & 0.137 & 2.606 & -2.743 & 0 \\ 0 & 0 & 0 & 0 & 0.189 & 0 & 0.189 & -0.378 \end{bmatrix}. \quad (42)$$

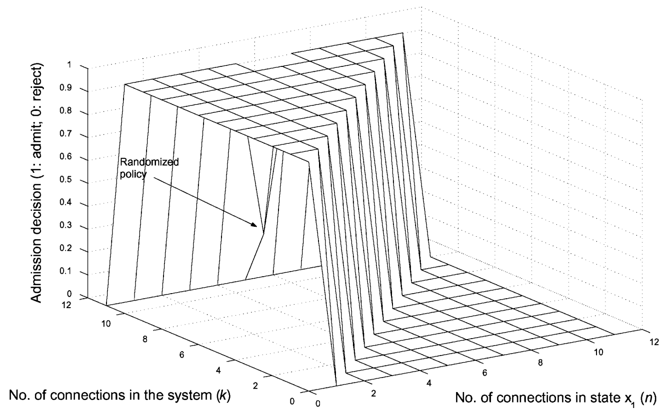

Fig. 9. Dependence of admission policy on the number of connections in the system (n) and the number of connections in different states (k).

component (the number of connections in) and the n component (the number of connections in different states). It is observed that the k component is much important in making the admission decision than the n component. For example, when there are less than 10 connections in the system $(k < 10)$, all new connection arrivals can be admitted no matter how many connections are in the state ON (the n value). This illustrates the nearly complete decomposability structure of the CAC problem in packet-switched CDMA networks.

C. The Performance of Nearly Complete Decomposability Approximation

We compare the results from the original scheme CMAC-II to those from CMAC-IV using nearly complete decomposability approximation. Table I illustrates the parameters in the experiment and the numerical results. SIR outage constraint is 0.005 and λ_1, μ_1 are fixed in this experiment. The blocking probabilities are obtained from different values of α_1 and β_1. We observe that the NCD solution can have good approximation to the optimal solution. Table I also indicates that the NCD solution performs better when the MMPP changes states faster (α and β are large). This is because when MMPP changes states faster, the maximum degree of coupling ϵ in (30) is smaller, and the process of aggregation yields better approximation.

Table II shows the comparison of the computation complexity of the proposed schemes. In particular, we are interested in the CPU time required in solving the linear program. We run the algorithms on a PC with a 1.6-GHz Pentium 4 CPU with 256 M memory. Both CMAC-III and CMAC-IV need much less CPU time than CMAC-I and CMAC-II. The NCD solutions reduce the computation complexity significantly.

D. Comparisons of the Conservative Approach With the Aggressive Approach to the SIR Outage Probability

We compare the conservative approach to the SIR outage probability of CMAC-III with the aggressive approach of CMAC-IV. The video traffic is used in this example.

Figs. 10 and 11 show the video blocking probabilities and the network utilizations, respectively, in CMAC-I, CMAC-III and CMAC-IV. The SIR outage probability constraint is 0.01 in CMAC-III and CMAC-IV. We observe that CMAC-III has

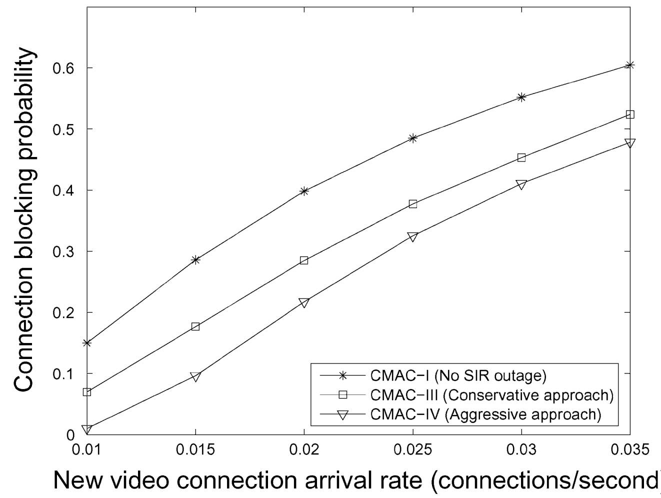

Fig. 10. Video connection blocking probabilities in CMAC-I, CMAC-III, and CMAC-IV.

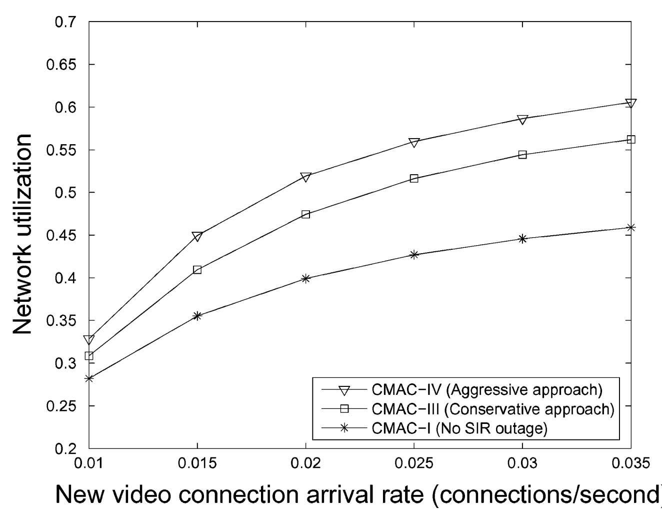

Fig. 11. Network utilizations in CMAC-I, CMAC-III, and CMAC-IV.

higher blocking probability and lower network utilization than CMAC-IV. This is because CMAC-III can guarantee the SIR outage probability in any state and is conservative in admitting connections. In contrast, in CMAC-IV, the SIR outage probability can be violated in some states, but is satisfied over long periods of time, which can increase the network utilization. It is also observed that both CMAC-III and CMAC-IV have higher network utilization than CMAC-I by introducing the SIR outage probability.

VIII. CONCLUSION

We have presented optimal connection admission control schemes in packet CDMA networks. MMPP is used to model the packet traffic. An optimal solution to this problem was given. To achieve the multiplexing gain of packet traffic, we introduced the SIR outage probability constraint and proposed the second scheme that can guarantee the SIR outage probability. By introducing a small SIR outage probability, the network utilization can be increased significantly. We also showed that the embedded Markov chain of the SMDP that models packet

TABLE I
CONNECTION BLOCKING PROBABILITY IN THE OPTIMAL SOLUTION CMAC-II AND THE NCD SOLUTION CMAC-IV

λ_1	μ_1	α_1	β_1	Blocking prob. in CMAC-II	Blocking prob. in CMAC-IV
0.1	0.004	0.02	0.02	0.2201	0.2392
0.1	0.004	0.04	0.04	0.2213	0.2392
0.1	0.004	0.4	0.4	0.2306	0.2392
0.1	0.004	4	4	0.2381	0.2392

TABLE II
CPU TIME REQUIRED TO OBTAIN THE ADMISSION CONTROL POLICY

Admission control scheme	CPU time (seconds)
CMAC-I	784.3
CMAC-II	785.6
CMAC-III	3.6
CMAC-IV	3.8

CDMA networks has nearly complete decomposability structure. Then, we proposed two other solutions that can reduce the computation complexity and yield satisfactory approximation to the optimal solutions.

The proposed approaches are based on the MMPP traffic model. It is interesting to study other traffic models in the cross-layer optimization framework. In addition, we did not address the access control issue in this paper, which is also of interest in future work.

REFERENCES

[1] W. Yang and E. Geraniotis, "Admission policies for integrated voice and data traffic in CDMA packet radio networks," *IEEE J. Sel. Areas Commun.*, vol. 12, pp. 654–664, May 1994.

[2] M. Soroushnejad and E. Geraniotis, "Multi-access strategies for integrated voice/data CDMA packet radio networks," *IEEE Trans. Commun.*, vol. 43, pp. 934–945, Feb. 1995.

[3] T. Liu and J. A. Silvester, "Joint admission/congestion control for wireless CDMA systems supporting integrated services," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 845–857, Aug. 1998.

[4] S. Singh, V. Krishnamurthy, and H. V. Poor, "Integrated voice/data call admission control for wireless DS-CDMA systems with fading," *IEEE Trans. Signal Process.*, vol. 50, pp. 1483–1495, June 2002.

[5] C. Comaniciu and H. V. Poor, "Jointly optimal power and admission control for delay sensitive traffic in CDMA networks with LMMSE receivers," *IEEE Trans. Signal Process.*, vol. 51, pp. 2031–2042, Aug. 2003.

[6] C. Comanciu, N. Mandayamn, D. Famolari, and P. Agrawal, "Wireless access to the world wide web in an integrated CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 472–483, May 2003.

[7] R. M. Rao, C. Comaniciu, T. V. Lakshman, and H. V. Poor, "An overview of cac principles in DS-CDMA networks—Call admission control in wireless multimedia networks," *IEEE Signal Process. Mag.*, vol. 21, pp. 51–58, Sep. 2004.

[8] A. Swami and L. Tong, "Signal processing for networking: An integrated approach," *IEEE Signal Process. Mag.*, vol. 21, pp. 18–19, Sep. 2004.

[9] A. Jamalipour and P. Lorenz, "Merging IP and wireless networks," *IEEE Wireless Commun.*, vol. 10, pp. 6–7, Oct. 2003.

[10] A. Maharshi, L. Tong, and A. Swami, "Cross-layer designs of multichannel reservation MAC under Rayleigh fading," *IEEE Trans. Signal Process.*, vol. 51, pp. 2054–2067, Aug. 2003.

[11] W. Fischer and K. Meier-Hellstern, "The Markov-modulated Poisson process (MMPP) cookbook," *Perform. Eval.*, vol. 18, no. 2, pp. 149–171, 1993.

[12] J. N. Daigle and J. D. Langford, "Models for analysis of packet-voice communication systems," *IEEE J. Sel. Areas Commun.*, vol. 4, pp. 847–855, Sep. 1986.

[13] P. Skelly, M. Schwartz, and S. Dixit, "A histogram-based model for video traffic behavior in an ATM multiplexer," *IEEE/ACM Trans. Netw.*, vol. 1, pp. 446–459, Aug. 1993.

[14] A. T. Anderson and B. F. Nielsen, "A Markovian approach for modeling packet traffic with long-range dependence," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 719–732, Jun. 1998.

[15] T. Yoshihara, S. Kasahara, and Y. Takahashi, "Practical time-scale fitting of self-similar traffic with Markov-modulated Poisson process," *Kluwer Telecommunication Systems*, vol. 17, pp. 185–211, Jun. 2001.

[16] M. Schwartz, *Broadband Integrated Networks*: Prentice Hall, 1996.

[17] A. Sampath and J. M. Holtzman, "Access control of data in integrated voice/data CDMA systems: Benefits and tradeoffs," *IEEE J. Sel. Areas Commun.*, vol. 15, pp. 1511–1526, Oct. 1997.

[18] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Trans. Netw.*, vol. 2, pp. 1–15, Feb. 1994.

[19] M. E. Crovella and A. Bestavros, "Self-similarity in world wide web traffic," *IEEE/ACM Trans. Netw.*, vol. 5, pp. 835–846, Dec. 1997.

[20] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2059–2078, Sep. 2000.

[21] A. Wyner, "Shannontheoretic approach to a gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, pp. 1713–1727, Nov. 1994.

[22] C. D. Meyer, "Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems," *SIAM Rev.*, vol. 31, pp. 240–272, 1989.

[23] P. J. Courtois, *Decomposability, Queueing and Computer Systems Applications*. New York, NY: Academic, 1977.

[24] M. Garrett and M. Vetterli, "Congestion control strategies for packet video," in *Proc. 4th Int. Workshop on Packet Video*, Kyoto, Japan, Mar. 1991.

Fei Yu (M'04) received the M.S. degree in computer engineering from Beijing University of Posts and Telecommunications, P.R. China, in 1998, and the Ph.D. degree in electrical engineering from the University of British Columbia (UBC), Canada, in 2003.

From 1998 to 1999, he was a system engineer with China Telecom, P.R. China, working on the planning, design, and performance analysis of national SS7 and GSM networks. From 2002 to 2004, he was a research and development engineer with Ericsson Mobile Platforms, Sweden, where he worked on dual-mode UMTS/GPRS handsets. He is currently a Postdoctoral Research Fellow at UBC. His research interests are quality of service, cross-layer design, and mobility management in wireless networks.

Vikram Krishnamurthy (F'05) was born in 1966. He received the B.S. degree from the University of Auckland, New Zealand, in 1988, and the Ph.D. degree from the Australian National University, Canberra, in 1992.

He is currently a professor and Canada Research Chair with the Department of Electrical Engineering, University of British Columbia, Vancouver. Prior to 2002, he was a chaired professor with the Department of Electrical and Electronic Engineering, University of Melbourne, Australia, where he also served as Deputy Head of department. His current research interests include ion channels and biological nanotubes, networked sensor scheduling and control, statistical signal processing, nonlinear filtering, and wireless telecommunications.

Dr. Krishnamurthy is currently an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS AEROSPACE AND ELECTRONIC SYSTEMS, and *Systems and Control Letters*. He is a Guest Editor of a Special Issue of the IEEE TRANSACTIONS ON NANOBIOSCIENCE, March 2005 on bionanotubes. He currently serves on the Signal Processing Theory and Methods (SPTM) Technical Committee of the IEEE Signal Processing Society and the International Federation of Automatic Control (IFAC) Technical Committee on Modeling, Identification and Signal Processing. He has served on the Technical Program Committee of several conferences in signal processing, telecommunications, and control.

Victor C. M. Leung (F'03) received the B.A.Sc. (Hons.) degree in electrical engineering from the University of British Columbia (UBC) in 1977, and was awarded the APEBC Gold Medal as the Head of the graduating class in the Faculty of Applied Science. He was awarded a Natural Sciences and Engineering Research Council Postgraduate Scholarship from UBC, where he received the Ph.D. degree in electrical engineering in 1981.

From 1981 to 1987, he was a Senior Member of Technical Staff with Microtel Pacific Research Ltd. (later renamed MPR Teltech Ltd.), specializing in the planning, design, and analysis of satellite communication systems. He also held a part-time position as a Visiting Assistant Professor with Simon Fraser University in 1986 and 1987. In 1988, he was a Lecturer with the Department of Electronics, Chinese University of Hong Kong. He joined the Department of Electrical Engineering at UBC in 1989, where he is currently a Professor, Associate Head of Graduate Affairs, holder of the TELUS Mobility Industrial Research Chair in Advanced Telecommunications Engineering, and a member of the Institute for Computing, Information and Cognitive Systems. His research interests are in the areas of architectural and protocol design and performance analysis for computer and telecommunication networks, with applications in satellite, mobile, personal communications, and high-speed networks.

Dr. Leung is a voting member of ACM. He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He has served on the Technical Program Committees of numerous conferences, and serves as the Technical Program Vice-Chair of IEEE WCNC 2005.