

Research Article

Cross-Layer Admission Control Policy for CDMA Beamforming Systems

Wei Sheng and Steven D. Blostein

Department of Electrical and Computer Engineering, Queen's University, Walter Light Hall (19 Union Street), Kingston, Ontario, Canada K7L 3N6

Received 31 October 2006; Revised 24 June 2007; Accepted 1 August 2007

Recommended by Robert W. Heath Jr.

A novel admission control (AC) policy is proposed for the uplink of a cellular CDMA beamforming system. An approximated power control feasibility condition (PCFC), required by a cross-layer AC policy, is derived. This approximation, however, increases outage probability in the physical layer. A truncated automatic retransmission request (ARQ) scheme is then employed to mitigate the outage problem. In this paper, we investigate the joint design of an AC policy and an ARQ-based outage mitigation algorithm in a cross-layer context. This paper provides a framework for joint AC design among physical, data-link, and network layers. This enables multiple quality-of-service (QoS) requirements to be more flexibly used to optimize system performance. Numerical examples show that by appropriately choosing ARQ parameters, the proposed AC policy can achieve a significant performance gain in terms of reduced outage probability and increased system throughput, while simultaneously guaranteeing all the QoS requirements.

Copyright © 2007 W. Sheng and S. D. Blostein. 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

In a code division multiple access (CDMA) system, quality-of-service (QoS) requirements rely on interference mitigation schemes and resource management, such as power control, multiuser detection, and admission control (AC) [1–3]. Recently, the problem of ensuring QoS by integrating the design in the physical layer and the admission control (AC) in the network layer is receiving much attention. In [4, 5], an optimal semi-Markov decision process (SMDP)-based AC policy is presented based on a linear-minimum-mean-square-error (LMMSE) multiuser receiver for constant bit rate traffic and circuit-switched networks. In [6], optimal admission control schemes are proposed in CDMA networks with variable bit rate packet multimedia traffic.

The above algorithms [4–6] integrate the optimal AC policy with a multiuser receiver, and as a result, are able to optimize the power control and the AC across the physical and network layers. However, [4–6] only consider single antenna systems, which lack the tremendous performance benefits provided by multiple antenna systems [7–17]. Furthermore, [4–6] rely on an asymptotic signal-to-interference ratio (SIR) expression proposed in [18] which requires a large number of users and a large processing gain. This specific

signal model limits the application of the proposed AC policies. Motivated by these facts, in this paper, we investigate cross-layer AC design for an arbitrary-size CDMA system with multiple antennas at the base station (BS).

To derive an optimal AC policy, a feasible state space and exact power controllability are required but are hard to evaluate for the case of multiple antenna systems. This motivates an approximated power control feasibility condition (PCFC) proposed for admission control of a multiple antenna system. This approximation, however, introduces outage in the physical layer, for example, a nonzero probability that a target signal-to-interference ratio (SIR) cannot be satisfied. To reduce the outage probability in the physical layer, a truncated ARQ-based reduced-outage-probability (ROP) algorithm can be employed. Truncated ARQ is an error-control protocol which retransmits an error packet until correctly received or a maximum number of retransmissions is reached. It is well known that retransmissions can significantly improve transmission reliability, and as a result, can reduce the outage probability. Although retransmissions increase the transmission duration of a packet and thus degrade the network layer performance, this degradation can be controlled to an arbitrarily small level by appropriately choosing the parameters of a truncated ARQ scheme, such as the maximum

number of allowed retransmissions and target packet-error rate (PER).

To date, there is no research on cross-layer AC design which considers both link-layer error control schemes and multiple antennas. We remark that this paper differs from prior investigations, for example, [4–6], in the following aspects: (a) here multiple antenna systems are investigated which provide a large capacity gain, while in [4–6], only single antenna systems are discussed; (b) in this paper, a cross-layer AC policy is designed by including error-control schemes, while in [4–6], no such error control schemes are exploited; (c) prior investigations in [4–6] rely on a large system analysis which requires an infinite number of users and infinite length spreading sequences, while here, no such requirements are imposed. In summary, this paper provides a framework for joint optimization across physical, data-link, and network layers, and as a result, is capable of providing a flexible way to handle QoS requirements.

We remark that in the current third generation (3G) system, the application of more efficient methods for packet data transmission such as high-speed uplink packet access (HSUPA) has become more important [19]. In HSUPA, a threshold-based call admission control (CAC) policy is employed, which admits a user request if the load reported is below the CAC threshold. Although the CAC decision can be improved upon by taking advantage of resource allocation information [19], and it is simple to implement, it is well known that the threshold-based CAC policy cannot satisfy QoS requirements in the network layer [5]. Our proposed AC policy provides a solution to guarantee the QoS requirements in both physical and network layers.

The proposed AC policy can be derived offline and then stored in a lookup table. Whenever an arrival or departure occurs, an optimal action can be obtained by table lookup, resulting in low enough complexity for admission control at the packet level. Similar to call/connection level admission control, in a packet-switched system, a packet admission control policy decides if an incoming packet can be accepted or blocked in order to meet quality-of-service (QoS) requirements. In a packet-switched network, blocking a packet instead of blocking the whole user connection can be more spectrally efficient. In this paper, we consider the packet level AC problem.

The rest of this paper is organized as follows. In Section 2, we present the signal model. In Section 3, an approximated PCFC and ARQ-based ROP algorithm are discussed. The formulation and solution of Markov-decision-process (MDP)-based AC policies are proposed in Section 4. Section 5 summarizes the cross-layer design of ARQ parameters. Simulation results are then presented in Section 6.

We will use the following notation: $\ln x$ is the natural logarithm of x , and $*$ denotes convolution. The superscripts $(\cdot)^H$ and $(\cdot)^t$ denote hermitian and transpose, respectively; $\text{diag}(a_1, \dots, a_n)$ denotes a diagonal matrix with elements a_1, \dots, a_n , and \mathbf{I} denotes an identity matrix. For a random variable X , $E[X]$ is its expectation. The notation and definitions used in this paper are summarized in Table 1.

TABLE 1: Notation and definitions.

Notation	Definition
M	Number of antennas at the BS
K	Number of users
J	Number of classes
R_i	Data rate for packet i
p_i	Transmitted power for packet i
B	Bandwidth
G_i	Link gain for packet i
\mathbf{a}_i	Array response vector for packet i
λ_j	Arrival rate for class j
μ_j	Departure rate for class j
Ψ_j	Blocking probability constraint for class j
D_j	Connection delay constraint for class j
L_j	Maximum number of retransmissions for class j
ρ_j	Target PER for class j
$\text{PER}_{\text{overall}}^j$	Achieved overall PER for class j
PER_{in}^j	Achieved instantaneous PER for class j
γ_j	Target SIR for class j
B_j	Buffer size for class j
\mathbf{w}_i	Beamformer weight for packet i
η_0	One-sided power spectral density of additive white Gaussian noise (AWGN)

2. SIGNAL MODEL AND PROBLEM FORMULATION

2.1. Signal model at the physical layer

We consider an uplink CDMA beamforming system, in which M antennas are employed at the BS and a single antenna is employed for each packet. There are K accepted packets in the system, and a channel with slow fading is assumed.

To highlight the design across physical and upper layers considered in this paper, the effects due to multipath are neglected. However, the proposed schemes in this paper can be extended straightforwardly to the case where multipath exists, provided multipath delay profile information is available.

The received vector at the BS antenna array can be written as

$$\mathbf{x}(t) = \sum_{i=1}^K \sqrt{P_i} G_i \mathbf{a}_i s_i(t - \tau_i) + \mathbf{n}(t), \quad (1)$$

where P_i and G_i denote the transmitted power and link gain for packet i , respectively; \mathbf{a}_i is defined as the array response vector for packet i , which contains the relative phases of the received signals at each array element, and depends on the array geometry as well as the angle of arrival (AoA); $s_i(t)$ is the transmitted signal, given by $s_i(t) = \sum_n b_i(n) c_i(t - nT)$, where $b_i(n)$ is the information bit stream, and $c_i(t)$ is the spreading sequence; τ_i is the corresponding time delay and $\mathbf{n}(t)$ is the thermal noise vector at the input of antenna array.

It has been shown that the output of a matched filter sampled at the symbol interval is a sufficient statistic for the estimation of the transmitted signal [14]. The matched filter for a desired packet k is given by $c_k^H(-t)$. The output of the matched filter is sampled at $t = nT$, where T denotes symbol interval. Hence, the received signal at the output of the matched filter is given by [14]

$$\begin{aligned} \mathbf{x}_k(n) &= \mathbf{x}(t) * c_k^H(-t)|_{t=nT} \\ &= \sum_{i=1}^K \sqrt{P_i} G_i \mathbf{a}_i \int_{(n-1)T+\tau_k}^{nT+\tau_k} \sum_m b_i(m) c_i(t - mT - \tau_i) \\ &\quad \times c_k(t - nT - \tau_k) dt + \mathbf{n}_k(n), \end{aligned} \quad (2)$$

where $\mathbf{n}_k(n) = \mathbf{n}(t) * c_k^H(-t)|_{t=nT}$.

In order to reduce the interference, we employ a beamforming weighting vector \mathbf{w}_k for a desired packet k . We can write the output of the beamformer as

$$\begin{aligned} y_k(n) &= \mathbf{w}_k^H \mathbf{x}(n) \\ &= \sum_{i=1}^K \sqrt{P_i} G_i \mathbf{w}_k^H \mathbf{a}_i \int_{(n-1)T+\tau_k}^{nT+\tau_k} \sum_m b_i(m) c_i(t - mT - \tau_i) \\ &\quad \times c_k(t - nT - \tau_k) dt + \mathbf{w}_k^H \mathbf{n}_k(n). \end{aligned} \quad (3)$$

We assume the signature sequences of the interfering users appear as mutually uncorrelated noise. As shown in [14], the received signal-to-interference ratio (SIR) for a desired packet k can be written as

$$\text{SIR}_k = \frac{B}{R_i} \frac{p_k \phi_{kk}^2}{\sum_{l \neq i} p_l \phi_{il}^2 + \eta_0 B}, \quad (4)$$

where B and R_i denote the bandwidth and data rate for packet i , respectively, and the ratio B/R_i represents the processing gain; $p_i = P_i G_i^2$ denotes the received power for packet i , and η_0 denotes the one-sided power spectral density of background additive white Gaussian noise (AWGN); the parameters ϕ_{ii}^2 and ϕ_{ik}^2 are defined as

$$\phi_{ik}^2 = |\mathbf{w}_k^H \mathbf{a}_i|^2 \quad (5)$$

which capture the effects of beamforming. In the following, we consider a spatially matched filter receiver, for example, $\mathbf{w}_k = \mathbf{a}_k$.

QoS requirements in the physical layer

In a wireless communication network, we must allow for outage, defined as the probability that a target SIR, or equivalently, a target packet-error rate (PER), cannot be satisfied. The QoS requirement in the physical layer can be represented by a target outage probability.

In this paper, we rely on a relationship between a target SIR and a target PER. Although an exact relationship may not be available, we can obtain the target SIR according to

an approximate expression of PER. As discussed in [20], in a system with packet length N_p (bits), the target SIR for a desired packet i , denoted by γ_i , can be approximated by

$$\gamma_i = \frac{1}{g} [\ln a - \ln \rho_i] \quad (6)$$

for $\gamma_i \geq \gamma_0$ dB, where ρ_i denotes the overall target PER; a , g , and γ_0 are constants depending on the chosen modulation and coding scheme. In the above expression, the interference is assumed to be additive white Gaussian noise, which is reasonable in a system with enough interferers.

2.2. Signal model in data-link and network layers

We consider a single-cell CDMA system which supports J classes of packets, characterized by different target PERs ρ_j , different blocking probability requirements Ψ_j , and different connection delay requirements D_j , where $j = 1, \dots, J$. Requests for packet connections of class j are assumed to be Poisson distributed, with arrival rates λ_j , $j = 1, \dots, J$.

The admission control (AC) is performed at the BS. An AC policy is derived offline, and stored in a lookup table. When a packet is generated at the mobile station (MS), the MS sends an access request to the BS. In this request, the class of this packet is indicated. After receiving the request, the BS makes a decision, which is then sent back to the MS, on whether the incoming packet should be either accepted, queued in the buffer, or blocked. Similarly, whenever a packet departs, the BS decides whether the packet in the queue can be served (transmitted).

Once a packet is accepted, its first transmission round will be performed, and then the receiver will send back an acknowledgement (ACK) signal to the transmitter. A positive ACK indicates that the packet is correctly received while a negative ACK indicates an incorrect transmission.

If a positive ACK is received or the maximum number of retransmissions, denoted by L , is reached, the packet releases the server and departs. Otherwise, the packet will be retransmitted. Therefore, the service time of a packet can comprise at most $L + 1$ transmission rounds. Each transmission round includes the actual transmission time of the packet and the waiting time of an ACK signal (positive or negative). The duration of a transmission round for a packet in class j is assumed to have an exponential distribution with mean duration $1/\mu_j$, $j = 1, \dots, J$. However, in this paper, a sub-optimal solution is also provided for a generally distributed duration.

If the packet is not accepted by the AC policy, it will be stored in a queue buffer provided that the queue buffer is not full. Otherwise, the packet will be blocked. Each class of packets shares a common queue buffer, and B_j denotes the queue buffer size of class j .

The QoS requirements in the network layer can be represented by the target blocking probability and connection delay, denoted by Ψ_j and D_j for class j , respectively. For each class j , where $j = 1, \dots, J$, there are K_j packets physically present in the system, which have the same target packet-error-PER, blocking probability, and connection delay constraints.

We note that there are two types of buffers in the system: queue buffers and server buffers. The queue buffer accommodates queued incoming packets, while the server buffer accommodates transmitted packets in the server in case any packet in the server requires retransmission. For simplicity, we assume that the size of the server buffer is large enough such that all the packets in the server can be stored. In the following, the generic term “buffer” refers to the queue buffer.

2.3. Problem formulation

The AC policy considered in this paper is for the uplink only. However, with an appropriate physical layer model for power allocation, the methodology can be extended straightforwardly to the downlink AC problem. The uplink AC is performed at the BS, and the following information is necessary to derive an admission control policy: traffic model in the system, such as arrival and departure rate, and QoS requirements in both physical and network layers.

The overall system throughput is defined as the number of correctly received packets per second, given by

$$\text{Throughput} = \sum_{j=1}^J (1 - P_b^j)(1 - \rho_j)(1 - P_{\text{out}}^j)\lambda_j, \quad (7)$$

where P_b^j , ρ_j and P_{out}^j denote the blocking probability, target PER, and outage probability for class j packets, respectively.

In this paper, we aim to derive an optimal AC policy which incorporates the benefits provided by multiple antennas and ARQ schemes. The objective is to maximize the overall system throughput given in (7), while simultaneously guaranteeing QoS requirements in terms of outage probability, blocking probability, and connection delay.

The above optimization problem can be formulated as a Markov decision process (MDP). With a required power control feasibility condition (PCFC), combined with an ARQ-based reduced-outage-probability (ROP) algorithm, a target outage probability constraint can be satisfied. Blocking probability and connection delay requirements can be guaranteed by the constraints of this MDP.

In the following, we first derive an approximate PCFC combined with an ARQ-based reduced-outage-probability (ROP) algorithm that can guarantee the outage probability constraint. Based on these results, we then formulate the AC problem as a Markov decision process. Afterward, we discuss how to design ARQ parameters optimally in order to achieve a maximum system throughput.

3. PHYSICAL LAYER INVESTIGATION: PCFC DERIVATION AND OUTAGE REDUCTION

To investigate the physical layer performance, we must derive an approximate PCFC, which ensures a positive power solution to achieve target SIRs. Due to the approximation of the derived PCFC, we then propose an ARQ-based ROP algorithm to reduce the resulting outage probability.

3.1. PCFC

In the physical layer, the SIR requirements of packet i can be written as

$$\text{SIR}_i \geq \gamma_i \quad (8)$$

for $i = 1, \dots, K$, where SIR_i is given in (4).

Inserting the SIR expression in (4) into (8), and letting SIR_i achieve its target value, γ_i , we have the matrix form [15]

$$[\mathbf{I} - \mathbf{QF}]\mathbf{p} = \mathbf{Q}\mathbf{u}, \quad (9)$$

where \mathbf{I} is the identity matrix, $\mathbf{p} = [p_1, \dots, p_K]^t$, $\mathbf{u} = \eta_0 B [1, \dots, 1]^t$,

$$\mathbf{Q} = \text{diag} \left\{ \frac{\gamma_1 R_1/B}{1 + \gamma_1 R_1/B}, \dots, \frac{\gamma_K R_K/B}{1 + \gamma_K R_K/B} \right\},$$

$$\mathbf{F} = \begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & F_{1,K} \\ F_{2,1} & F_{2,2} & \cdots & F_{2,K} \\ \cdots & \cdots & \cdots & \cdots \\ F_{K,1} & F_{K,2} & \cdots & F_{K,K} \end{bmatrix} \quad (10)$$

in which $F_{ij} = \phi_{ij}^2 / \phi_{ii}^2$.

To ensure a positive solution for power vector \mathbf{p} , we require the following power control feasibility condition [15],

$$\rho(\mathbf{QF}) < 1, \quad (11)$$

where $\rho(\cdot)$ denotes the maximum eigenvalue.

The outage probability can be obtained as the probability that the above condition is violated. Although the state space, required by an optimal AC policy, can be formulated by evaluating the above outage probability, this evaluation relies on the number of packets as well as the distribution of AoAs for all the packets in the system, and thus results in a very high computation complexity. An approach to evaluate the above outage probability with reasonably low complexity is currently under investigation.

In this paper, we propose an alternative solution, which employs an approximated PCFC, and as a result can dramatically simplify the formulation of the state space.

Without loss of generality, we consider an arbitrary packet i in class 1, where $i = 1, \dots, K_1$. By considering specific traffic classes and letting SIR achieve its target value, the expression in (4) can be written as

$$\gamma_i = \frac{p_i \phi_{ii}^2 (B/R_1)}{\sum_{l=1, l \neq i}^{K_1} p_l \phi_{il}^2 + \sum_{l=1}^{K_2} p_l \phi_{il}^2 + \cdots + \sum_{l=1}^{K_J} p_l \phi_{il}^2 + \sigma^2}, \quad (12)$$

where $\sigma^2 \triangleq \eta_0 B$ denotes noise variance, and p_i represents received power for packet i .

It is not difficult to show that packets in the same class have the same received power. By denoting the received power in class j as p_j , where $j = 1, \dots, J$, the above

expression can be written as

$$\begin{aligned} \gamma_i &= \frac{p_1 \phi_{ii}^2(B/R_1)}{\sum_{l=1, l \neq i}^{K_1} p_l \phi_{il}^2 + \dots + \sum_{l=1}^{K_j} p_l \phi_{il}^2 + \sigma^2} \\ &= \frac{p_1 \phi_{ii}^2(B/R_1)}{p_1(K_1 - 1)\beta_1 + \sum_{j=2}^J p_j K_j \beta_j + \sigma^2}, \end{aligned} \quad (13)$$

where $\beta_1 = (1/(K_1 - 1)) \sum_{l=1, l \neq i}^{K_1} \phi_{il}^2$ and $\beta_j = (1/K_j) \sum_{l=1}^{K_j} \phi_{il}^2$, in which $j = 2, \dots, J$.

By exchanging the numerator and denominator, (13) is equivalent to

$$\frac{p_1(K_1 - 1)\beta_1 + \sum_{j=2}^J p_j K_j \beta_j + \sigma^2}{p_1(B/\gamma_1 R_1)} = \phi_{ii}^2, \quad (14)$$

where $i = 1, \dots, K_1$.

Summing the above K_1 equations, and calculating the sample average, we obtain

$$\frac{p_1(K_1 - 1)\alpha_1 + \sum_{j=2}^J K_j p_j \alpha_j + \sigma^2}{p_1(B/\gamma_1 R_1)} = \frac{1}{K_1} \sum_{i=1}^{K_1} \phi_{ii}^2, \quad (15)$$

where $\alpha_1 = (1/K_1) \sum_{i=1}^{K_1} \beta_1$ and $\alpha_j = (1/K_1) \sum_{i=1}^{K_1} \beta_j$.

When the number of packets is large enough, by the weak law of large numbers, the above $\alpha_1, \dots, \alpha_j$ can be approximated by their mean values, and (15) can be further simplified as

$$\frac{p_1(K_1 - 1)E_{11}[\phi_{\text{int}}] + \sum_{j=2}^J K_j p_j E_{1j}[\phi_{\text{int}}] + \sigma^2}{p_1(B/\gamma_1 R_1)} = E_1[\phi_{\text{des}}] \quad (16)$$

in which $E_{mn}[\phi_{\text{int}}]$ is the expected fraction of an interferer packet in class n passed by a beamforming weight vector for a desired packet in class m , where $m, n = 1, \dots, J$, while $E_j[\phi_{\text{des}}]$ is the expected fraction of a desired packet in class j passed by its beamforming weight vector, where $j = 1, \dots, J$.

The AoAs of active packets in the system are assumed to be independent and identically distributed, that are independent of a packet's specific class. Therefore, it is reasonable to assume that $E_{mn}[\phi_{\text{int}}]$ is also independent of specific classes m and n , which can be denoted by $E[\phi_{\text{int}}]$. Similarly, $E_j[\phi_{\text{des}}]$ is independent of class j , and can be denoted by $E[\phi_{\text{des}}]$. $E[\phi_{\text{des}}]$ and $E[\phi_{\text{int}}]$ represent the expected fractions of the desired packet's power and interference, respectively.

From the above discussion, (16) can be written as

$$\frac{p_1(K_1 - 1)E[\phi_{\text{int}}] + \sum_{j=2}^J K_j p_j E[\phi_{\text{int}}] + \sigma^2}{p_1(B/\gamma_1 R_1)} = E[\phi_{\text{des}}]. \quad (17)$$

By exchanging the numerator and denominator of the above equation, we have

$$\begin{aligned} p_1 \frac{B}{\gamma_1 R_1} / \left(p_1(K_1 - 1) \frac{E[\phi_{\text{int}}]}{E[\phi_{\text{des}}]} \right. \\ \left. + \sum_{j=2}^J K_j p_j \frac{E[\phi_{\text{int}}]}{E[\phi_{\text{des}}]} + \frac{\sigma^2}{E[\phi_{\text{des}}]} \right) = 1. \end{aligned} \quad (18)$$

The QoS requirement for class 1 in (18) can be extended to any class j ,

$$\begin{aligned} p_j \frac{B}{\gamma_j R_j} / \left(p_j(K_j - 1) \frac{E[\phi_{\text{int}}]}{E[\phi_{\text{des}}]} \right. \\ \left. + \sum_{m=1, m \neq j}^J K_m p_m \frac{E[\phi_{\text{int}}]}{E[\phi_{\text{des}}]} + \frac{\sigma^2}{E[\phi_{\text{des}}]} \right) = 1, \end{aligned} \quad (19)$$

where $j = 1, \dots, J$.

The power allocation solution can be obtained by solving the above J equations [21]

$$\begin{aligned} p_j = \frac{\sigma^2}{E[\phi_{\text{int}}]} / \left(\left(1 + \frac{B}{\gamma_j R_j (E[\phi_{\text{int}}]/E[\phi_{\text{des}}])} \right) \right. \\ \left. \times \left[1 - \sum_{j=1}^J \frac{K_j}{1 + (B/\gamma_j R_j (E[\phi_{\text{int}}]/E[\phi_{\text{des}}]))} \right] \right), \end{aligned} \quad (20)$$

where $j = 1, \dots, J$.

Positivity of the power solution implies the following power control feasibility condition:

$$\sum_{j=1}^J \frac{K_j}{1 + (B/\gamma_j R_j (E[\phi_{\text{int}}]/E[\phi_{\text{des}}]))} < 1. \quad (21)$$

As shown in [22], $E[\phi_{\text{int}}]$ and $E[\phi_{\text{des}}]$ can be determined numerically from (5) for a beamforming system.

We note that the above approximated power control feasibility condition is independent of the angle of arrivals, and thus can provide a less-complicated offline AC policy, which does not require estimation of the current AoA realizations of each packet. However, due to the randomness of the actual SIR, this deterministic power control feasibility condition introduces outage. In the next section, we discuss how to mitigate the outage.

3.2. ARQ-based ROP

We first define two types of PERs. The overall achieved PER, denoted by $\text{PER}_{\text{overall}}^j$, is defined as the probability that a class j packet is incorrectly received after its maximum number of ARQ retransmissions is reached, for example, an error occurs in each of the $L_j + 1$ transmission rounds, where L_j denotes the maximum number of retransmissions. The achieved instantaneous PER, denoted as $\text{PER}_{\text{in}}^j(l)$, is defined as the probability that an error occurs in a single transmission round l for a class j packet.

Under the assumption that each retransmission round is independent from the others, by using an ARQ scheme with a maximum of L_j retransmissions for class j , the achieved overall PER is constrained by [20]

$$\begin{aligned} \text{PER}_{\text{overall}}^j &= \prod_{l=1}^{L_j+1} \text{PER}_{\text{in}}^j(l), \\ &\leq \rho_j, \end{aligned} \quad (22)$$

where ρ_j denotes the target overall PER for class j .

The achieved outage probability for class j , denoted by P_{out}^j , can be written as

$$\begin{aligned} P_{\text{out}}^j &= \text{Prob}\{\text{PER}_j^{\text{overall}} > \rho_j\} \\ &= \text{Prob}\left\{\prod_{l=1}^{L_j+1} \text{PER}_{\text{in}}^j(l) > \rho_j\right\}, \end{aligned} \quad (23)$$

where $\text{Prob}\{A\}$ denotes the probability of event A . By maintaining PCFC, $\text{PER}_{\text{in}}^j(l)$ remains unchanged. Therefore, by increasing L_j , the outage probability in the above equation can be reduced.

4. AC PROBLEM FORMULATION BY INCLUDING ARQ

In the previous section, we have derived an approximated PCFC combined with an ARQ-based ROP algorithm in the physical layer. In the following, we discuss how to derive an AC policy in the network layer.

An optimal semi-Markov decision process (SMDP)-based AC policy as well as a low-complexity generalized-Markov decision process (GMDP)-based AC policy is discussed.

4.1. SMDP-based AC policy

Traditionally, the decision epoches are chosen as the time instances that a packet arrives or departs. In the system under consideration, the duration of each packet may include several transmission rounds due to ARQ retransmissions, and as a result, the time duration until next system state may not be exponentially distributed. Therefore, the SMDP formulation approach discussed in [4–6], which assumes an exponentially distributed duration, cannot be applied here.

In the following, we propose a novel formulation in which the decision epoch is chosen as the arrival and departure of each transmission round. Based on these decision epoches, the time duration until the next state remains exponentially distributed. The components of a Markov decision process, such as state space, action space, and dynamic statistics, are modified accordingly to represent the characteristics of different transmission rounds. The formulation of this SMDP as well as its LP solution are now described.

State space and action space

Class j packets are divided into $L_j + 1$ subclasses, in which the state of the i th subclass can be represented by the number of packets which are under the i th round transmission, that is, the $(i - 1)$ th retransmission, where $i = 1, \dots, L_j + 1$.

In admission problems, the discrete-value (finite) state at time t , $s(t)$, can be written as

$$\begin{aligned} s(t) &= \left[\underbrace{n_q^1(t), k^{1,1}(t), \dots, k^{1,L_j+1}(t), \dots,}_{n_q^j(t), k^{j,1}(t), \dots, k^{j,L_j+1}(t)} \right]^T, \end{aligned} \quad (24)$$

where $k^{j,i}(t)$ represents the number of active packets in class j and subclass i served in the system, and $n_q^j(t)$ denotes the number of packets in the queue buffer of class j . Since the arrival and departure of packets are random, $\{s(t), t > 0\}$ represents a finite state stochastic process [4]. From here on, we will drop the time index.

The state space S is comprised of any state vector s , in which SIR requirements can be satisfied or, equivalently, the power control feasibility condition (PCFC) holds,

$$\begin{aligned} S &= \left\{ s : n_q^j \leq B_j, j = 1, \dots, J; \right. \\ &\quad \left. \sum_{j=1}^J \frac{(\sum_{l=1}^{L_j+1} k^{j,l})}{1 + (B/\gamma_j R_j (E[\phi_{\text{int}}]/E[\phi_{\text{des}}]))} < 1 \right\}, \end{aligned} \quad (25)$$

where B_j denotes the buffer size of class j . We have mentioned that the PCFC for the case of no ARQ is used in our AC problem, no matter how many retransmissions are allowed.

At each state s , an action is chosen that determines how the admission control will perform at the next decision moment [4]. In general, an action, denoted as a , can be defined as a vector of dimension $\sum_{j=1}^J L_j + 2J$,

$$a = \left[\underbrace{a_1, d_1^1, \dots, d_1^{L_1+1}}_{\dots}, \underbrace{a_j, d_j^1, \dots, d_j^{L_j+1}} \right]^T, \quad (26)$$

where a_j denotes the action for class j if an arrival occurs, $j = 1, \dots, J$. If $a_j = 0$, the new arrival is placed in the buffer provided that the buffer is not full or is blocked if the buffer is full; if $a_j = 1$, the arrival is admitted as an active packet, and the number of servers of class j is incremented by one.

The quantity d_j^i , where $1 \leq i \leq L_j$, denotes the action for class j packet if the i th transmission round is finished, and is received correctly. If $d_j^i = 0$, where $1 \leq i \leq L_j$, $k^{j,i}$ is decremented by one, and no packets that are queued in the buffer are made active; if $d_j^i = 1$, the number of servers is maintained by admitting a packet at the buffer as an active packet.

The quantity $d_j^{L_j+1}$ denotes the action for class j packet if a connection has finished its $(L_j + 1)$ th transmission round. If $d_j^{L_j+1} = 0$, no packets that are queued in the buffer are made active, and k^{j,L_j+1} is decremented by one; if $d_j^{L_j+1} = 1$, the number of servers is maintained by admitting a packet at the buffer as an active packet.

The admissible action space for state s , denoted by A_s , can be defined as the set of all feasible actions. A feasible action ensures that after taking this action, the next transition state is still in space S [4].

State dynamics $p_{sy}(a)$ and $\tau_s(a)$

The state dynamics of an SMDP are completely specified by stating the transition probabilities of the embedded chain $p_{sy}(a)$ and the expected holding time $\tau_s(a)$: $p_{sy}(a)$ is defined as the probability that the state at the next decision epoch is

TABLE 2: Expression of transition probability p_{sy} .

y	$p_{sy}(a)$
$y = s + q^j$	$\lambda_j a_j \tau_s(a)$
$y = s + b^j$	$\lambda_j (1 - a_j) \delta(B_j - n_q^j) \tau_s(a)$
$y = s + c_i^j$	$(1 - \rho_j) [\mu_j k^{j,i} (1 - d_i^j) \tau_s(a)]$ $+ (1 - \rho_j) [\mu_j k^{j,i} d_i^j (1 - \delta(n_q^j)) \tau_s(a)]$
$y = s + r^j$	$\sum_{i=1}^{L_j+1} (1 - \rho_j) \mu_j k^{j,i} d_i^j \tau_s(a) \delta(n_q^j)$
$y = s + e_i^j$	$\rho_j \mu_j k^{j,i} \tau_s(a)$
$y = s + f^j$	$\mu_j k^{j,L_j+1} d_j^{L_j+1} \delta(n_q^j) \tau_s(a)$
$y = s + g^j$	$\mu_j k^{j,L_j+1} (1 - d_j^{L_j+1}) \tau_s(a)$ $+ \mu_j k^{j,L_j+1} d_j^{L_j+1} (1 - \delta(n_q^j)) \tau_s(a)$
Otherwise	0

y if action a is selected at the current state s , while $\tau_s(a)$ is the expected time until the next decision epoch after action a is chosen in the present state s [4].

Derivations of $\tau_s(a)$ and $p_{sy}(a)$ rely on the statistical properties of arrival and departure processes [4]. Since the arrival and departure processes are both Poisson distributed and mutually independent, it follows that the cumulative process is also Poisson, and the cumulative event rate is the sum of the rates for all constituent processes [4]. Therefore, the expected sojourn time, $\tau_s(a)$, can be obtained as the inverse of the event rate,

$$\begin{aligned} \tau_s(a)^{-1} &= \lambda_1 a_1 + \lambda_1 (1 - a_1) \delta(B_1 - n_q^1) \\ &+ \sum_{i=1}^{L_1+1} \mu_1 (k^{1,i}) + \dots + \lambda_j a_j + \lambda_j (1 - a_j) \delta(B_j - n_q^j) \\ &+ \sum_{i=1}^{L_j+1} \mu_j (k^{j,i}), \end{aligned} \quad (27)$$

where

$$\delta(z) = \begin{cases} 1 & \text{if } z > 0, \\ 0 & \text{if } z = 0. \end{cases} \quad (28)$$

To derive the transition probabilities, we employ the decomposition property of a Poisson process, which states that an event of a certain type occurs with a probability equal to the ratio between the rate of that particular type of event and the total cumulative event rate $1/(\tau_s(a))$ [4]. Transition probability $p_{sy}(a)$ is shown in Table 2, where ρ_j denotes the target packet-error rate for class j packets. The set of vectors $\{q_j, b^j, c_i^j, r^j, e_i^j, f^j, g^j\}$ represents the possible state transitions from current state s . Each vector in this set has a dimension of $\sum_{j=1}^J L_j + 2J$, and contains only zeros except for one or two positions. The nonzero positions of this set of vectors, as well as the possible state transitions represented by these vectors, are specified in Tables 3 and 4, respectively.

Policy and cost criterion

For any given state $s \in S$, an action a , which decides if the new packet at the next decision epoch will be blocked or ac-

TABLE 3: Definition of vectors in Table 2: each vector defined in this table has a dimension of $\sum_{j=1}^J L_j + 2J$, which contains only zeros except for the specified positions.

Vector	Nonzero positions
q^j	Position $2(j-1) + \sum_{t=1}^{j-1} L_t + 2$ contains a 1
b^j	Position $2(j-1) + \sum_{t=1}^{j-1} L_t + 1$ contains a 1
c_i^j	Position $2(j-1) + \sum_{t=1}^{j-1} L_t + i + 1$ contains a -1
r^j	Position $2(j-1) + \sum_{t=1}^{j-1} L_t + 1$ contains a -1
e_i^j	Position $2(j-1) + \sum_{t=1}^{j-1} L_t + i + 1$ contains a -1 and position $2(j-1) + \sum_{t=1}^{j-1} L_t + i + 2$ contains a 1
f^j	Position $2(j-1) + \sum_{t=1}^{j-1} L_t + 1$ contains a -1
g^j	Position $2(j-1) + \sum_{t=1}^{j-1} L_t + L_j + 2$ contains a -1

TABLE 4: Representation of vectors in Table 2: each defined vector represents a possible state transition from current state s .

Notation	State transition
$s + q^j$	An increase in subclass 1 of class j by 1
$s + b^j$	An increase in queue j by 1
$s + c_i^j$	A decrease in subclass i of class j by 1
$s + r^j$	A decrease in queue j by 1
$s + e_i^j$	An increase of subclass $i + 1$ by 1, and a decrease in subclass i of class j by 1
$s + f^j$	A decrease in queue j by 1
$s + g^j$	A decrease in subclass $L_j + 1$ of class j by 1

cepted, is selected according to a specified policy R . A stationary policy R is a function that maps the state space into the admissible action space.

We consider average cost criterion [4]. The cost criterion for a given policy R and initial state s_0 , which includes blocking probability as a special case, is given as follows:

$$J_R(s_0) = \lim_{T \rightarrow \infty} \frac{1}{T} E \left\{ \int_0^T c(s(t), a(t)) dt \right\}, \quad (29)$$

where $c(s(t), a(t))$ can be interpreted as the expected cost until the next decision epoch and is selected to meet the network layer performance criteria [4].

In the system under investigation, we are interested in blocking probability and connection delay constraints. If the cost criterion $J_R(s_0)$ represents blocking probability, we have $c(s, a) = (1 - a_j)(1 - \delta(B_j - n_q^j))$, and if the cost criterion $J_R(s_0)$ represents connection delay, we have $c(s, a) = n_q^j$.

An optimal policy R^* that minimizes an average cost criterion $J_R(s_0)$ for any initial state s_0 exists,

$$J_{R^*}(s_0) = \min_{R \in \mathbf{R}} J_R(s_0), \quad \forall s_0 \in S \quad (30)$$

under the weak unichain assumption [23], where \mathbf{R} is the class of admissible AC policies.

Solving the AC policy by linear programming (LP)

The optimal AC policy, which can minimize the blocking probability, can be obtained by using the decision variables $z_{sa}, s \in S, a \in A_s$.

The optimal AC policy R^* in (30) can be obtained by solving the following linear programming (LP):

$$\min_{z_{sa} \geq 0, s, a} \sum_{s \in S} \sum_{a \in A_s} \sum_{j=1}^J \eta_j (1 - a_j) (1 - \delta(B_j - n_q^j)) \tau_s(a) z_{sa} \quad (31)$$

subject to

$$\begin{aligned} \sum_{a \in A_y} z_{ya} - \sum_{s \in S} \sum_{a \in A_s} p_{sy}(a) z_{sa} &= 0, \quad y \in S, \\ \sum_{s \in S} \sum_{a \in A_s} \tau_s(a) z_{sa} &= 1, \\ \sum_{s \in S} \sum_{a \in A_s} (1 - a_j) (1 - \delta(B_j - n_q^j)) \tau_s(a) z_{sa} &\leq \Psi_j, \\ \sum_{s \in S} \sum_{a \in A_s} n_q^j \tau_s(a) z_{sa} &\leq D_j, \end{aligned} \quad (32)$$

where D_j and Ψ_j denote the connection delay and blocking probability constraints, respectively, and η_j is the coefficient representing the weighting of the cost function for a particular class j , where $j = 1, \dots, J$.

The optimal policy will be a randomized policy: the optimal action $a^* \in A_s$ for state s , where A_s is the admissible action space, is chosen probabilistically according to the probabilities $z_{sa} / \sum_{a \in A_s} z_{sa}$.

We remark that the above randomized AC policy can optimize the long-run performance. The decision variables, z_{sa} , where $s \in S$ and $a \in A_s$, act as the long-run fraction of decision epoches at which the system is in state s and action a . At each state s , there exists a set of feasible actions, and each action induces a different cost $c(s, a)$. The long-run performance can be optimized by appropriately allocating these time fractions, and the allocation leads to a randomized AC policy. When a deterministic policy is desired, a constraint regarding the decision variables z_{sa} should be imposed into the above optimization problem, in order to ensure that at each state s , there is one and only one nonzero decision variable. It is obvious that the more constraints we impose, the worse the achieved performance becomes. We choose a randomized AC policy in order to achieve long-run optimal performance.

4.2. GMDP-based AC policy

In the above, we provide an optimal SMDP formulation. The state space has dimension of $2J + \sum_{j=1}^J L_j$ for J classes of traffic. For large J and retransmission number, this leads to a computation problem of excessive size.

In order to reduce complexity, we consider the decision epoch as the time instances that a packet arrives or departs. As we discussed in the previous section, based on these decision epoches, the time interval until the next state is not exponentially distributed. Therefore, we have a generalized Markov decision process (GMDP). While an optimal solution for this GMDP problem is hard to obtain, a linear programming approach provides a suboptimal solution [5].

We remark that the formulation of a GMDP is very similar to the AC problem formulation employed in [4–6], except that the state space has been modified to include beamforming and the mean duration of a packet is modified to consider the impact of ARQ schemes.

In the formulated GMDP, decision epoches are chosen as the time instances that a packet arrives or departs. The arrival process for class j is assumed to have a Poisson distribution with arrival rate λ_j . The duration of the class j packets may have a general distribution, with mean $(1/\mu_j)(1 + \rho_j + \dots + \rho_j^{L_j})$, where μ_j denotes the departure rate for each transmission round for the class j packets.

The state space S is comprised of any state vector s , which satisfies SIR requirements,

$$S = \left\{ s = [n_q^1, k^1, \dots, n_q^J, k^J]^T : n_q^j \leq B_j, \right. \\ \left. j = 1, \dots, J; \sum_{j=1}^J \frac{k^j}{1 + (B/\gamma_j R_j (E[\phi_{\text{int}}]/E[\phi_{\text{des}}]))} < 1 \right\}, \quad (33)$$

where k^j denotes the number of active packets for class j .

At each decision epoch, an action is chosen as $a = [a_1, d_1, \dots, a_J, d_J]^T$, where a_j denotes the action for class j if an arrival occurs, $j = 1, \dots, J$ and d_j denotes the action for class j packet if a packet in this class departs. The admissible action space for state s , denoted by A_s , can be defined as the set of all feasible actions.

The state dynamics of a SMDP are completely specified by stating the transition probabilities of the embedded chain $p_{sy}(a)$ and the expected holding time $\tau_s(a)$, which are given in [4, 5].

After formulating the AC problem as a GMDP, the AC policy, which minimizes the blocking probability, can be obtained by using the decision variables $z_{sa}, s \in S, a \in A_s$ from linear programming which is presented in (31).

In a low instantaneous PER region, the suboptimal solution proposed in the above is very close to the SMDP-based AC policy. Intuitively, when the PER is very low, retransmission occurs only occasionally, and the duration of a packet would be very close to an exponential distribution. In this case, the LP approach would provide an optimal solution to the above GMDP.

We remark that unlike the SMDP-based AC policy in which the transmission round is assumed to have an exponential distribution, the GMDP-based AC policy discussed in the subsection can be applied to a system with a generally distributed transmission round.

4.3. Complexity

SMDP or GMDP-based AC policies are always calculated offline and stored in a lookup table. Whenever an arrival or departure occurs, an optimal action can be obtained by table lookup using the current system state. This facilitates the implementation of packet-level admission control.

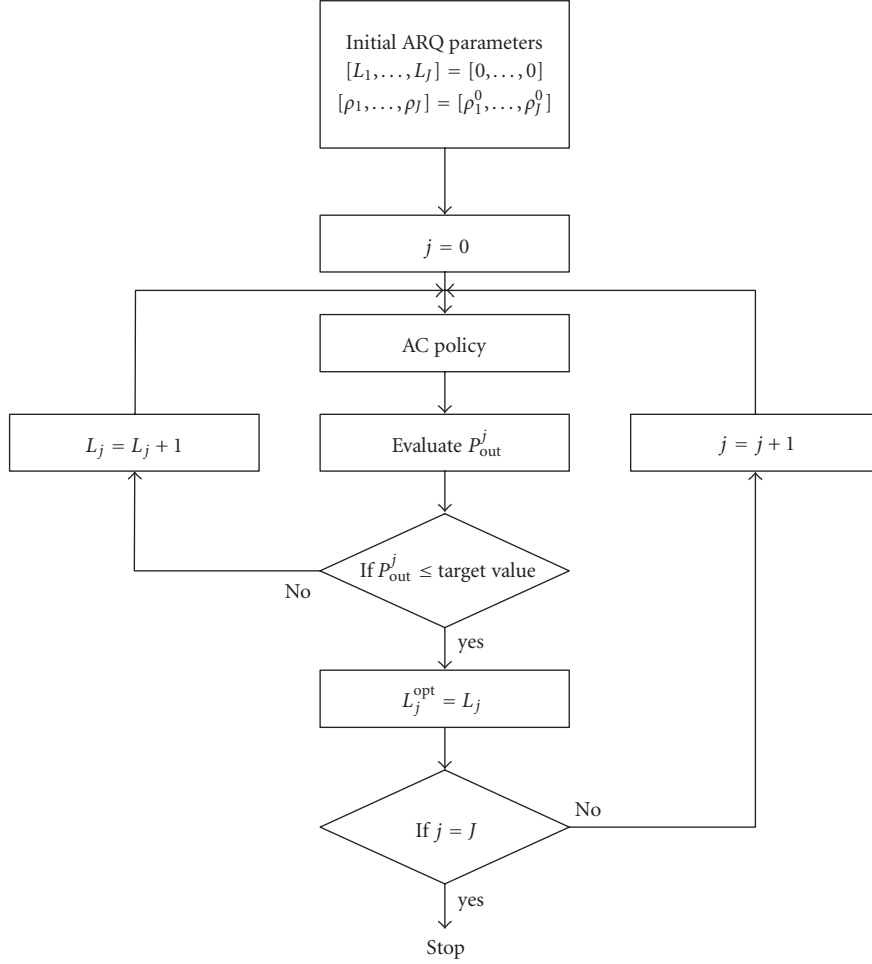


FIGURE 1: Search procedure of the optimal number of retransmissions.

Once system parameters change, an updated policy is required. However, in the system we investigate, the policy only depends on buffer sizes, long-term traffic model, and QoS requirements. These parameters are generally constant for the provision of a given profile of offered services. Therefore, an SMDP or GMDP-based policy has a very reasonable computation complexity.

5. CROSS-LAYER DESIGN OF ARQ PARAMETERS

In the previous sections, we discuss how to derive the PCFC in the physical layer and how to derive admission control in the network layer. These derivations assume that ARQ parameters such as L_j and ρ_j , where $j = 1, \dots, J$, are already known. In this section, we discuss how to choose these parameters in order to guarantee outage probability constraints and optimize overall system throughput.

The search procedures for optimal ARQ parameters, denoted as vectors $L^{\text{opt}} = [L_1^{\text{opt}}, \dots, L_J^{\text{opt}}]$ and $\rho^{\text{opt}} = [\rho_1^{\text{opt}}, \dots, \rho_J^{\text{opt}}]$, are demonstrated in Figures 1 and 2, respectively. The initial parameters are set to $[L_1, \dots, L_J] = [0, \dots, 0]$ and $[\rho_1, \dots, \rho_J] = [\rho_1^0, \dots, \rho_J^0]$, where ρ_j^0 represents the upper

bound target PER for class j , which can be specified for the system. In Figure 2, Δ_j represents the adjustment step size.

From the search procedures presented in Figures 1 and 2, it is observed that the number of allowed retransmissions L_j^{opt} , which can achieve a target outage probability, is minimized; and as a result, the network layer performance degradation can be minimized. Thus, network layer QoS requirements in terms of blocking probability and connection delay can be guaranteed by formulating the AC problem as an SMDP or GMDP.

Summing above, by choosing ARQ parameters in a cross-layer context, QoS requirements in the physical and network layers can be guaranteed, and the overall system throughput can be maximized.

6. SIMULATION RESULTS

We consider a 3-element circular antenna array, for example, $M = 3$, with a uniformly distributed angle of arrival (AoA) over $[0, 2\pi)$ [22]. Numerical values of parameters $E[\phi_{\text{des}}]$ and $E[\phi_{\text{int}}]$ in (21), derived in [22], are shown in Table 5. We remark that the proposed AC policies can be applied to any other array geometry and AoA distribution. Without loss of

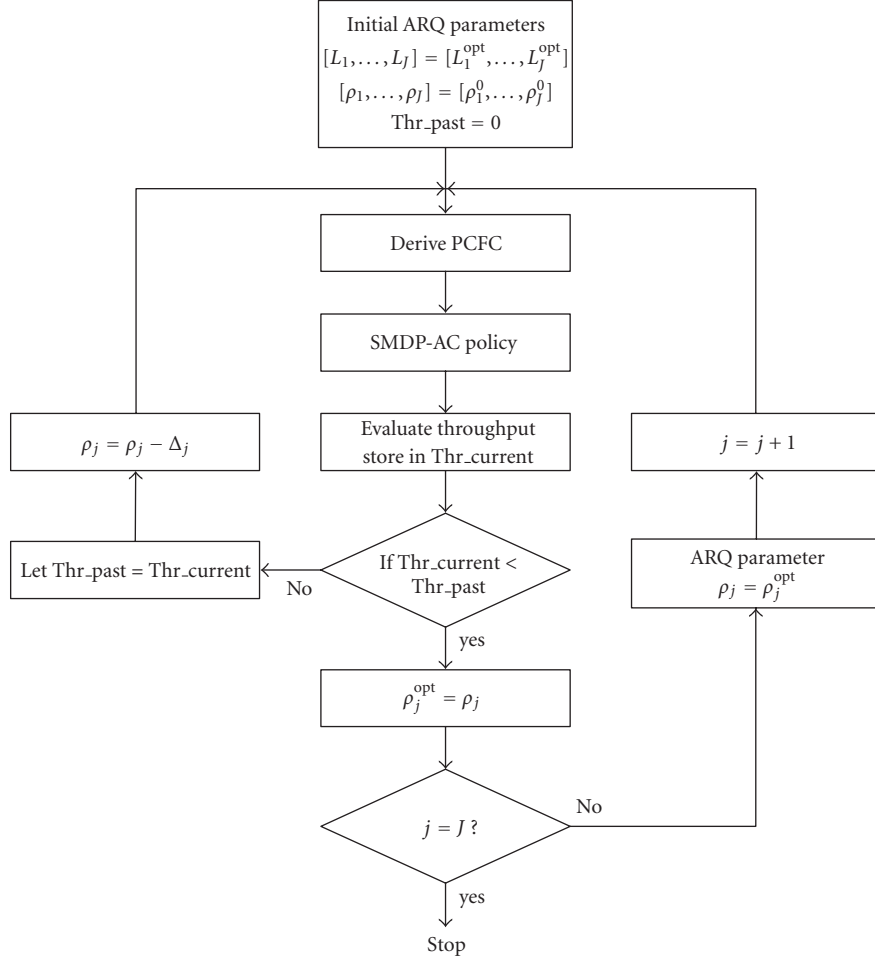


FIGURE 2: Search procedure of the optimal target PER.

TABLE 5: Numerical values of $E[\phi_{\text{des}}]$ and $E[\phi_{\text{int}}]$ in (20) and (21).

M	1	2	3	4	5	6
$E[\phi_{\text{des}}]$	1.0	1.0	1.0	1.0	1.0	1.0
$E[\phi_{\text{int}}]$	1.0	0.5463	0.3950	0.3241	0.2460	0.2058

generality, we consider a single-path channel and a two-class system with a QPSK and convolutionally coded modulation scheme with rate 1/2 and a packet length $N_p = 1080$. Under this scheme, the parameters of a , g , and γ_0 in (6) can be obtained from [20]. For simplicity, no buffer is employed in the simulation. Simulation parameters are presented in Table 6.

6.1. Performance of SMDP-based AC policies

Here, we investigate how the ARQ scheme can reduce outage probability while only slightly degrading the network layer performance.

We examine the case in which only the class 2 packets can be retransmitted once, for example, $L_1 = 0$ and $L_2 = 1$, and an optimal SMDP-based AC policy is employed. The target PER for the class 1 packets is set to 10^{-4} , while different target

TABLE 6: Simulation parameters.

B	3.84 MHz	a	90.2514
g	3.4998	γ_0	1.0942 dB
R_1	144 kbps	R_2	384 kbps
λ_1	1	λ_2	0.5
μ_1	0.25	μ_2	0.1375
Ψ_1	0.1	Ψ_2	0.2
D_1	2.25	D_2	0.5360
M	3	η_0	10^{-6}
η_1	0.5	η_2	0.5

PERs for class 2 are evaluated. We focus on the performance for the class 2 packets since only these packets are allowed retransmission. Figure 3 presents the analytical and simulated blocking probabilities as a function of ρ_2 . It is observed that the simulation results are very close to the analytical results. Figure 4 presents the outage probability and throughput for the class 2 packets. It is observed that at a reasonably low PER, the outage probability can be reduced dramatically, and overall system throughput can be significantly improved by allowing only one retransmission. Figure 5, which presents

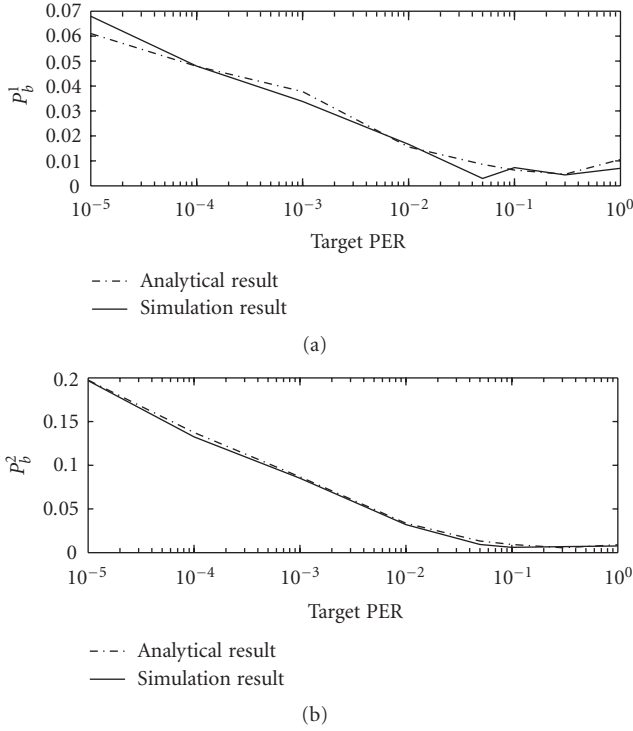


FIGURE 3: SMDP-based AC policy: analytical and simulated blocking probabilities as a function of ρ_2 in which $L_1 = 0, L_2 = 1$, and $\rho_1 = 10^{-4}$.

the network layer performance degradation by employing ARQ, shows that the degradation can be ignored in a low PER region.

6.2. Performance of GMDP-based AC policies

In the above, we discussed the performance of SMDP-based AC policies, which require high computation. To reduce complexity, a GMDP-based AC policy can be employed. The target PER for class 1 is set to 10^{-4} , while different target PER requirements for class 2 are considered.

Figure 6 shows the analytical and simulated blocking probabilities as a function of target PER for the class 2 packets. The gap between the simulated and analytical results is due to the non-exponential distribution of the packet duration.

Figure 7 demonstrates that for a small number of retransmissions, SMDP and GMDP-based AC policies have similar performance. Although performance comparison for large L_j is not presented here since an SMDP-based AC policy would involve excessive computation, it is expected that for low PER, these two AC policies would still have similar performance. For a high PER, however, the packet duration is far from exponentially distributed, and thus linear programming cannot provide an optimal solution to a GMDP and its performance would be inferior to that of SMDP. In summary, GMDP-based AC policy provides a simplified approach which is capable of achieving a near-optimal system

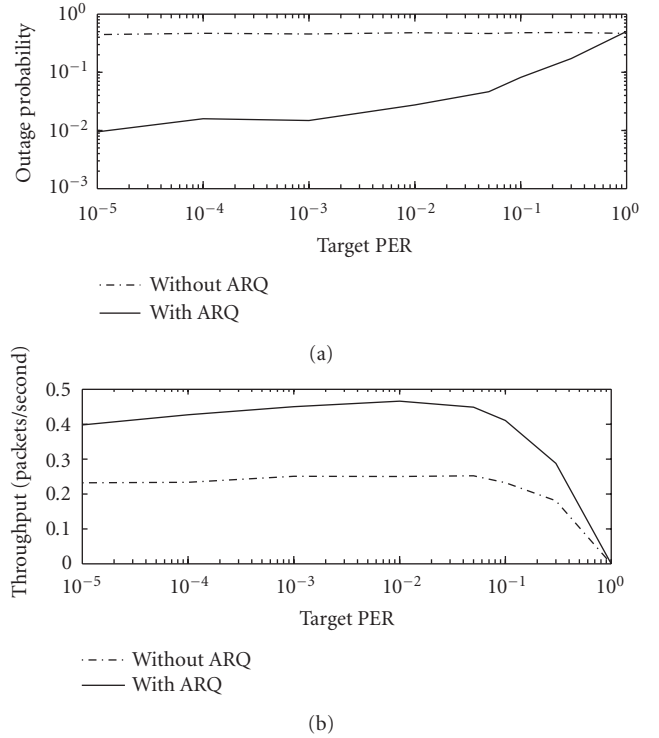


FIGURE 4: SMDP-based AC policy: outage probability and throughput for class 2 packets as a function of ρ_2 in which $L_1 = 0, L_2 = 1$, and $\rho_1 = 10^{-4}$.

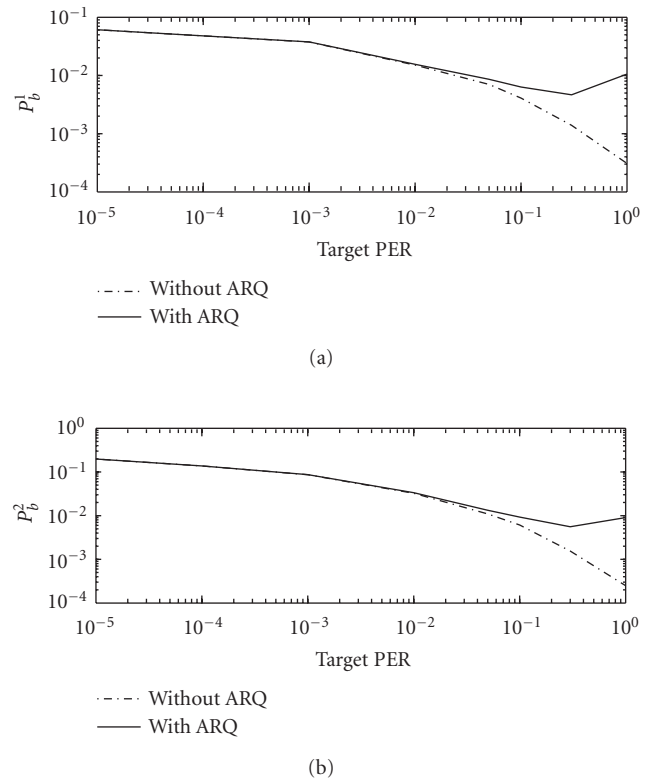


FIGURE 5: SMDP-based AC policy: blocking probability degradation as a function of ρ_2 in which $L_1 = 0, L_2 = 1$, and $\rho_1 = 10^{-4}$.

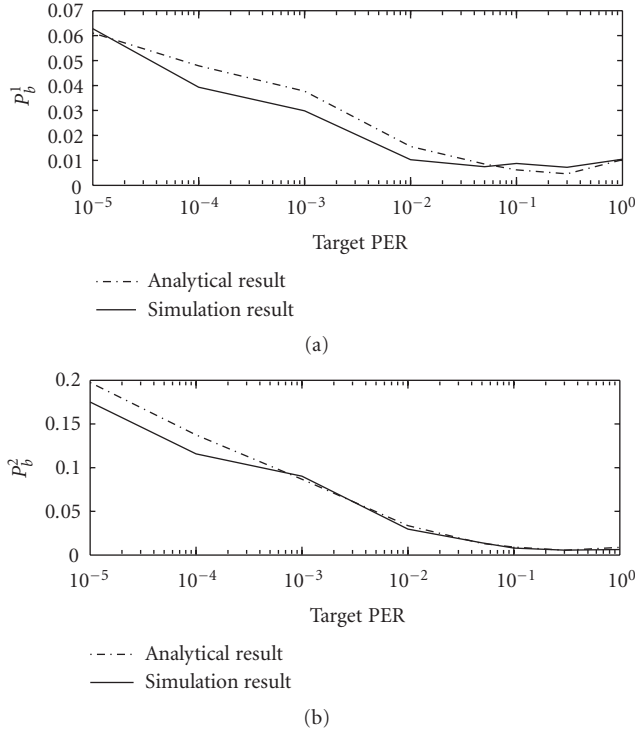


FIGURE 6: GMDP-based AC policy: analytical and simulated blocking probabilities as a function of ρ_2 in which $L_1 = 0$, $L_2 = 1$, and $\rho_1 = 10^{-4}$.

performance for a system with low PER or a small number of retransmissions.

Figures 8–10 compare the performance among different numbers of retransmissions in which $\rho_1 = \rho_2$, and $L_1 = L_2$. From here on, L_j is denoted by L in the figures. We investigate the performance for $L_j = 0, 1$, and 2, respectively. The results for large L_j can be extended straightforwardly. It is observed that in a low PER region, for example, $\rho_j \leq 0.01$, with an increased L_j , outage can dramatically be reduced, while the blocking probability is only slightly degraded. With only one retransmission allowed, the throughput can be improved by 100%. However, when L_j is increased beyond a certain level, for example, $L_j = 2$ in the system under consideration, the outage reduction and throughput improvement are not significant. Beyond this threshold, further increasing L_j may even lead to a performance degradation due to a degraded network layer performance. From Figures 8–10, we also conclude that at high PER, the proposed ARQ-based ROP algorithm is not as efficient as in low PER.

6.3. Performance of a complete-sharing-based admission control policy

For a complete-sharing (CS)-based policy, whenever a packet arrives, the power control feasibility condition in (21) is evaluated by incorporating information of this newly arrived packet. If this condition is satisfied, the incoming packet can be accepted, otherwise, the packet is stored in a buffer or blocked if the buffer is full. CS-based AC policy provides a

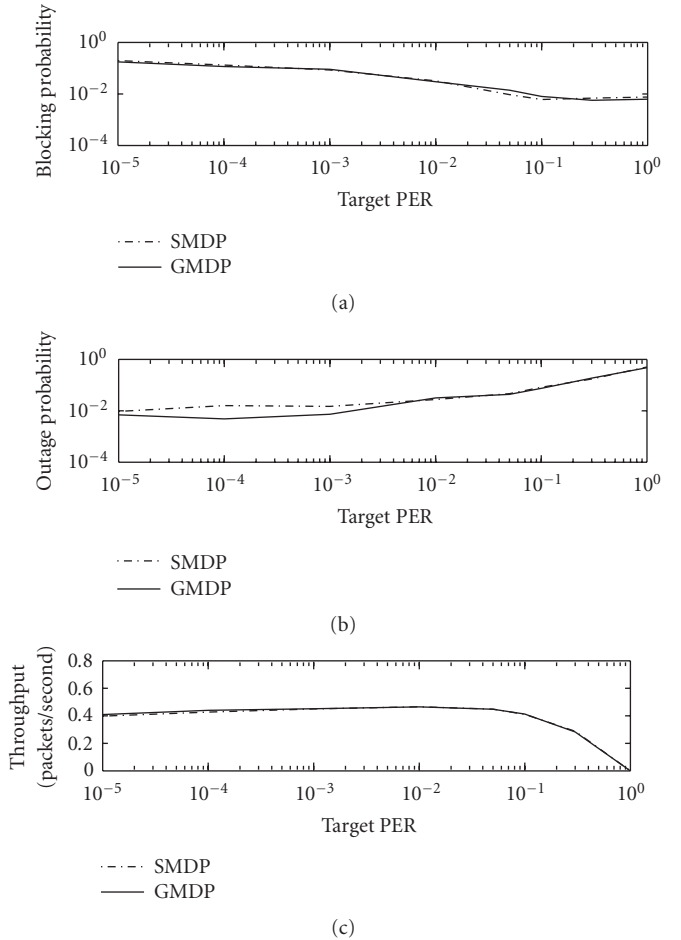


FIGURE 7: Performance comparison between SMDP and GMDP-based AC policies as a function of ρ_2 in which $L_1 = 0$, $L_2 = 1$, and $\rho_1 = 10^{-4}$.

simple admission control algorithm but ignores the QoS requirements in the network layer.

We now provide a simple example for complete-sharing (CS)-based AC policy. For comparison purposes, the simulation results for a GMDP-based AC policy is also presented. In this example, both classes of packets are allowed to retransmit twice, for example, $L_1 = L_2 = 2$.

We note that in a system with relaxed blocking probability constraints, even a CS-based AC policy can satisfy all the QoS requirements. To illustrate the shortcoming of a CS-based AC policy, we now restrict the blocking probability constraint for class 2 to 0.05 without loss of generality, and all the other parameters in Table 6 remain unchanged.

The results for a GMDP-based AC policy and a CS-based AC policy are shown in Table 7, in which P_b^j denotes the blocking probability for class j packets, where $j = 1, 2$ and P_b denotes the overall blocking probability. It is observed that for a CS-based AC policy, the blocking probability constraint cannot be guaranteed. For example, when the buffer size is $[0, 3]$, the blocking probability for class 1 packets is 0.1185, which exceeds its constraint 0.1. When the buffer size

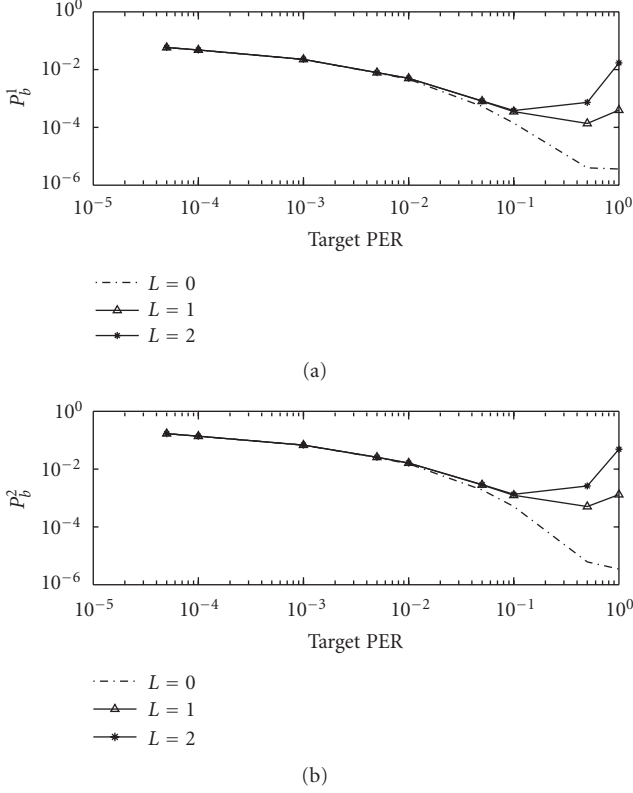


FIGURE 8: GMDP-based AC policies: blocking probability as a function of target PER in which $\rho_1 = \rho_2$ and $L_1 = L_2$.

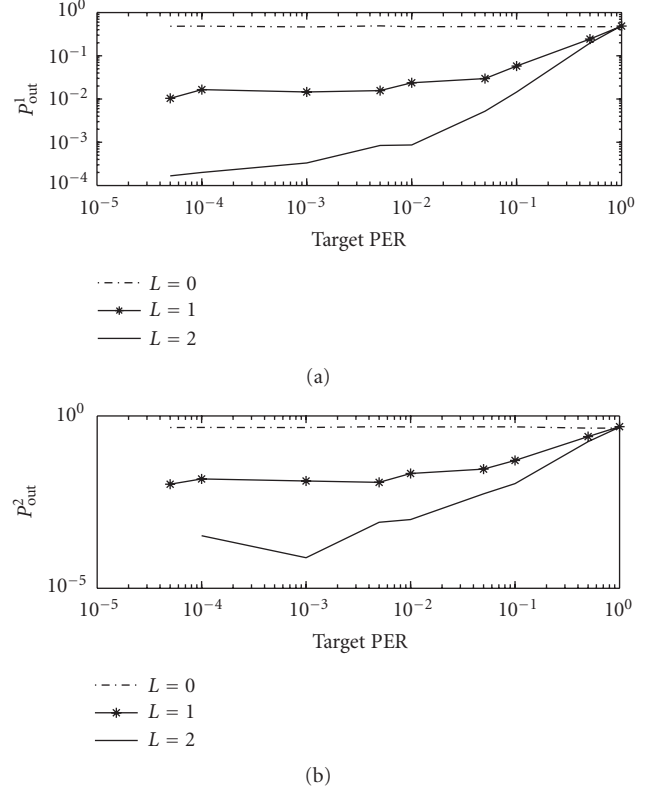


FIGURE 9: GMDP-based AC policies: outage probability as a function of target PER in which $\rho_1 = \rho_2$ and $L_1 = L_2$.

TABLE 7: Comparison between CS-based and GMDP-based AC policies, in which $L_1 = L_2 = 2, \rho_1 = \rho_2 = 10^{-4}$. The blocking probability constraint is set to $[0.1, 0.05]$, and the connection delay constraint is $[2.25, 0.5360]$.

$[B_1, B_2]$	GMDP: P_b^1	GMDP: P_b^2	GMDP: P_b	CS: P_b^1	CS: P_b^2	CS: P_b
$[0, 2]$	0.0714	0.0359	0.0537	0.0978	0.0390	0.0684
$[0, 3]$	0.0764	0.0280	0.0522	0.1185	0.0171	0.0678
$[1, 2]$	0.0434	0.0412	0.0423	0.0505	0.040	0.0452
$[3, 2]$	0.0179	0.0379	0.0279	0.0210	0.0569	0.0389

is $[3, 2]$, the blocking probability for class 2 packets is 0.0569, which exceeds its blocking probability constraint 0.05. However, for the same buffer sizes, GMDP-based AC policy can always guarantee blocking probability constraints for both classes.

6.4. Choosing ARQ parameters

As discussed in Section 5, ARQ parameters, such as L_j and ρ_j , should be chosen appropriately in order to achieve maximum throughput while simultaneously satisfying the QoS requirements in the physical and network layers.

We now provide a simple example to illustrate how to obtain optimal ARQ parameters by using the algorithm proposed in Section 5. The initial target PERs $\rho_j^0 = 0.05$, where

$j = 1, 2$, are given by the system which represents the upper bound of the target PER.

Using the algorithm presented in Section 5, the optimal ARQ parameters are derived as $L_1^{\text{opt}} = 1, L_2^{\text{opt}} = 1, \rho_1^{\text{opt}} = 0.005$, and $\rho_2^{\text{opt}} = 0.005$, respectively, for outage probability constraint $[0.01, 0.01]$ and blocking probability constraints $[0.1, 0.2]$. If the blocking probability constraint remains unchanged, and the outage probability constraint is reduced to $[10^{-3}, 10^{-3}]$, the optimal ARQ parameters can be derived as $L_1^{\text{opt}} = 2, L_2^{\text{opt}} = 2, \rho_1^{\text{opt}} = 0.01$, and $\rho_2^{\text{opt}} = 0.01$, respectively.

6.5. Sensitivity of the proposed algorithm to traffic load

In this subsection, we study the sensitivity of the proposed AC policy to different traffic loads. Traffic load can be represented by the packet occupancy ratio, defined as $[\lambda_1/\mu_1, \lambda_2/\mu_2]$. The following traffic loads are investigated: $[(1, 1/2); (2, 1(1/2)); (3, 2(1/2)); (4, 3(1/2)); (5, 4(1/2))]$.

Let λ and μ denote the overall arrival rate and the average departure rate, respectively, which can be expressed as $\lambda = \lambda_1 + \lambda_2$ and $\mu = (\lambda_1/(\lambda_1 + \lambda_2))\mu_1 + (\lambda_2/(\lambda_1 + \lambda_2))\mu_2$. The overall traffic load is represented by λ/μ . In the following examples, the target PER is assumed to be 10^{-3} for both classes, and a GMDP-based AC policy is employed, which would achieve a very similar performance to an optimal SMDP-based AC policy due to the low target PER under investigation.

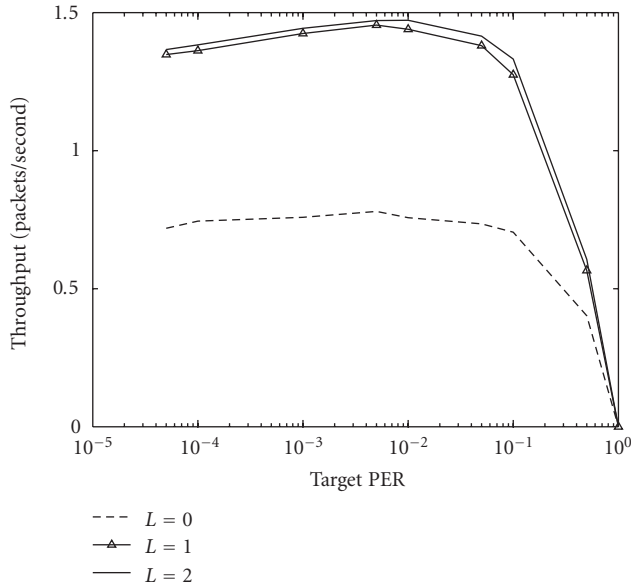


FIGURE 10: GMDP-based AC polices: throughput as a function of target PER in which $\rho_1 = \rho_2$ and $L_1 = L_2$.

Figure 11 presents the average blocking probability, outage probability and throughput as a function of overall traffic load. With an increased traffic load, there will be an increased interfering power and thus the performance is degraded. We remark that for all the traffic loads investigated, the proposed ARQ-based ROP algorithm is able to reduce the outage probability significantly at the cost of a slightly degraded network layer performance. Therefore, the proposed ARQ-based ROP algorithm can be applied to a wide variety of traffic conditions.

7. CONCLUSIONS

This paper provides a novel framework which exchanges information among physical, data-link, and network layers, and as a result provides a flexible way to handle the QoS requirements as well as the overall system throughput. In this paper, we propose a cross-layer AC policy combined with an ARQ-based ROP algorithm for a CDMA beamforming system. Both optimal and suboptimal admission control policies are investigated. We conclude that in a low PER region, for example, less than 10^{-2} , the proposed AC policies are capable of achieving significant performance gain while simultaneously satisfying all QoS requirements. Numerical examples show that the throughput can be improved by 100% by employing only one retransmission. Although ARQ schemes may degrade network layer performance, this degradation can be adequately controlled by appropriately choosing ARQ parameters. Furthermore, the proposed AC policy and ARQ-based ROP algorithm can be applied to any traffic load.

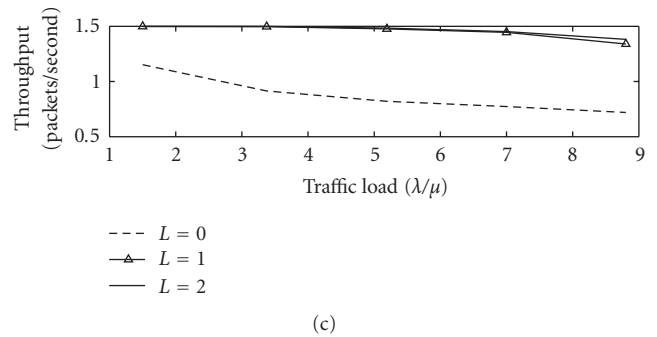
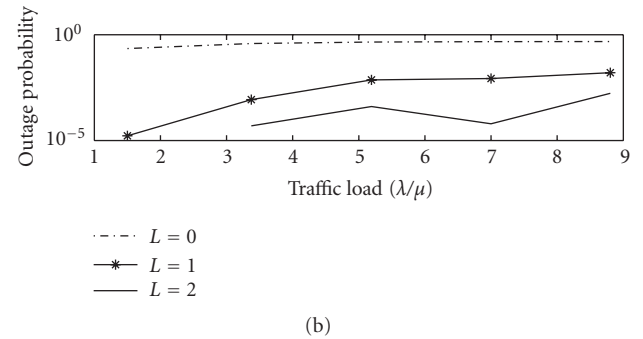
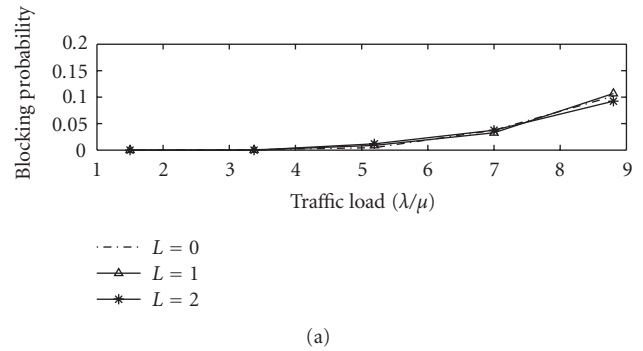


FIGURE 11: Blocking probability, outage probability, and throughput as a function of overall traffic load in which $\rho_1 = \rho_2 = 10^{-3}$.

ACKNOWLEDGMENT

The support of the Natural Sciences and Engineering Research Council of Canada, discovery Grant 41731, is gratefully acknowledged.

REFERENCES

- [1] M. Andersin, Z. Rosberg, and J. Zander, "Soft and safe admission control in power-controlled mobile systems," *IEEE/ACM Transactions on Networking*, vol. 5, no. 2, pp. 255–265, 1997.
- [2] Y. Bao and A. S. Sethi, "Performance-driven adaptive admission control for multimedia applications," in *Proceedings of IEEE International Conference on Communications (ICC '99)*, vol. 1, pp. 199–203, Vancouver, BC, Canada, June 1999.
- [3] 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.
- [4] C. Comaniciu and H. V. Poor, "Jointly optimal power and admission control for delay sensitive traffic in CDMA networks

- with LMMSE receivers,” *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2031–2042, 2003.
- [5] 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.
- [6] F. Yu, V. Krishnamurthy, and V. C. M. Leung, “Cross-layer optimal connection admission control for variable bit rate multimedia traffic in packet wireless CDMA networks,” *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 542–555, 2006.
- [7] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [8] A. Yener, R. D. Yates, and S. Ulukus, “Combined multiuser detection and beamforming for CDMA systems: filter structures,” *IEEE Transactions on Vehicular Technology*, vol. 51, no. 5, pp. 1087–1095, 2002.
- [9] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [10] B. Hassibi and B. M. Hochwald, “High-rate codes that are linear in space and time,” *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, 2002.
- [11] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communication: performance criterion and code construction,” *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [12] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [13] S. D. Blostein and H. Leib, “Multiple antenna systems: their role and impact in future wireless access,” *IEEE Communications Magazine*, vol. 41, no. 7, pp. 94–101, 2003.
- [14] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, “Joint optimal power control and beamforming in wireless network-using antenna arrays,” *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1313–1324, 1998.
- [15] G. Song and K. Gong, “Performance comparison of optimum beamforming and spatially matched filter in power-controlled CDMA systems,” in *Proceedings of IEEE International Conference on Communications (ICC ’02)*, vol. 1, pp. 455–459, New York, NY, USA, April-May 2002.
- [16] A. M. Wyglinski and S. D. Blostein, “On uplink CDMA cell capacity: mutual coupling and scattering effects on beamforming,” *IEEE Transactions on Vehicular Technology*, vol. 52, no. 2, pp. 289–304, 2003.
- [17] K. I. Pedersen and P. E. Mogensen, “Directional power-based admission control for WCDMA systems using beamforming antenna array systems,” *IEEE Transactions on Vehicular Technology*, vol. 51, no. 6, pp. 1294–1303, 2002.
- [18] J. Evans and D. N. C. Tse, “Large system performance of linear multiuser receivers in multipath fading channels,” *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2059–2078, 2000.
- [19] S. Brueck, E. Jugl, H.-J. Ketttschau, M. Link, J. Mueckenheim, and A. Zaporozhets, “Radio resource management in HSDPA and HSUPA,” *Bell Labs Technical Journal*, vol. 11, no. 4, pp. 151–167, 2007.
- [20] Q. Liu, S. Zhou, and G. B. Giannakis, “Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1746–1755, 2004.
- [21] A. Sampath, P. S. 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.
- [22] A. Wyglinski, “Performance of CDMA systems using digital beamforming with mutual coupling and scattering effects,” M.S. thesis, Queen’s University, Kingston, Ontario, Canada, 2000.
- [23] H. C. Tijms, *Stochastic Modelling and Analysis: A Computational Approach*, John Wiley & Sons, Chichester, UK, 1986.