Bandwidth Allocation in Wireless Networks with Guaranteed Packet Loss Performance

Jeong Geun Kim, *Student Member, IEEE*, and Marwan M. Krunz, *Member, IEEE*

Abstract— **Providing quality of service (QoS) guarantees over wireless packet networks poses a host of technical challenges that are not present in wireline networks. One of the key issue is how to account for the characteristics of the time-varying wireless channel and for the impact of link-layer error control in the provisioning of packet-level QoS. In this paper, we accommodate both aspects in analyzing the packet loss performance over a wireless link. We consider the cases of a single and multiplexed traffic streams. The link capacity fluctuates according to a fluid version of Gilbert-Elliot channel model. Traffic sources are modeled as on-off fluid processes. For the single-stream case, we derive the exact packet loss rate (PLR) due to buffer overflow at the sender side of the wireless link. We also obtain a closed-form approximation for the corresponding wireless effective bandwidth. In the case of multiplexed streams, we obtain a good approximation for the PLR using the Chernoff-Dominant Eigenvalue (CDE) approach. Our analysis is then used to study the optimal FEC code rate that guarantees a given PLR while minimizing the allocated bandwidth. Numerical results and simulations are used to verify the adequacy of our analysis and to study the impact of error control on the allocation of bandwidth for guaranteed packet loss performance.**

Index Terms— **Wireless networks, QoS, effective bandwidth, fluid analysis**

I. Introduction

CURRENT trends in wireless networks indicate a desire to provide a flexible broadband wireless infrastructure that can support emerging multimedia services along with traditional data services [1], [2], [3]. In such a multi-service wireless environment, quality-of-service (QoS) guarantees are critical for real-time voice and video. The provisioning of these guarantees over wireless links is a challenging problem whose difficulty stems from the need to explicitly consider the harsh radio-channel transmission characteristics and the underlying link-layer error control mechanisms. This difficulty is further compounded by host mobility and its impact on the sustained bandwidth capacity.

Traffic control and resource allocation strategies in wireline packet networks have been designed with the assumption that the underlying physical media are highly reliable. Since this assumption does not hold for wireless links, such strategies may not be efficient in the wireless environment [4]. For instance, in transporting TCP traffic over wireless links, TCP makes the implicit assumption that packet losses are caused by congestion, although such losses may in fact be caused by packet discarding due to channel errors [5], [4]. Eventually, TCP times out and invokes its congestion control mechanism, which unnecessarily reduces the throughput of the wireless channel.

The transport performance over the radio channel can be improved by using link-layer error control, namely, automatic repeat request (ARQ) and/or forward error correction (FEC) [4]. In general, ARQ is used to deliver data requiring high reliability, whereas FEC is more suitable for delay-sensitive traffic [4]. Recent studies suggest that hybrid ARQ/FEC schemes might be more appropriate for a broadband wireless network [4], [6], particularly when the transported traffic streams exhibit diverse characteristics and QoS requirements. For instance, data connections with relaxed time constraints can use ARQ. In contrast, packet voice and video connections that require low delay jitter and minimal packet loss may best be supported by a combination of FEC and ARQ with time-constrained retransmission [7].

In order to provide QoS guarantees while achieving an acceptable level of bandwidth utilization, integrated networks often employ the concept of effective bandwidth in call admission control (CAC) and service scheduling [8], [9]. Significant research has been done on the notion of effective bandwidth over wireline networks [10], [8], [11], [9], [12]. Guérin et al. proposed an approximate expression for the effective bandwidth of both individual and multiplexed connections, arguing that this approximation is necessary for real-time network traffic control [9]. Elwalid and Mitra studied the effective bandwidth for general Markovian traffic sources [8]. Elwalid et al. proposed an approximation for the packet loss rate (PLR) at a statistical multiplexer using a hybrid Chernoff-Dominant Eigenvalue (CDE) approach [13]. The research on effective bandwidth has generally been addressed in the context of high-speed (wired) asynchronous transfer mode (ATM) networks. Recently, Mohammadi et al. extended the concept of effective bandwidth to wireless ATM networks [14]. However, the impact of error control was not considered. Bandwidth allocation and FEC code optimization in wireless ATM networks were discussed in [15] using a simplified framework that did not involve traffic models and queuing analysis (traffic sources were characterized by their mean rates). The provisioning of a guaranteed PLR was not investigated. In that study, the authors observed the existence of a tradeoff between FEC and the number of retransmitted packets. Chaskar et al. [16] studied the performance of TCP over a wireless link. Their analysis was used to investigate the "effective link capacity," defined as the maximum arrival rate of TCP packets for which buffer overflow is less than $1/W_{\rm bd}^2$; $W_{\rm bd}$ being the bandwidth-delay product. In [19] the authors presented a framework for providing QoS guarantees over

This work was supported in part by the National Science Foundation under grants ANI-9733143 and CCR-9979310. An abridged version of this paper appeared in the Proceedings of INFOCOM '99.
The authors are with the Department of Electrical and Computer Engineering, The University of Arizona, Tucson, Arizona 85721 (e-mail: jkkim@ece.arizona.edu; krunz@ece.arizona.edu).

a Rayleigh faded wireless downlink assuming FEC-based link shaping.

In this work, we investigate the packet loss performance due to buffer overflow at the transmitter side of a wireless link. We consider two scenarios (Figure 1). In the first scenario, a single stream is transported over the wireless link. This stream represents, for example, the traffic from a mobile terminal (MT) to an access point (AP) in a cellular system. In Figure 1, the terrestrial link between the MT and an AP is used to transport the traffic stream of Connection A. In the second scenario, several streams are multiplexed onto the same wireless link, which can be, for example, an intermediate satellite or microwave link. For both scenarios, we assume that the transmitter side of the wireless link maintains a finite-capacity packet buffer, which may occasionally overflow.

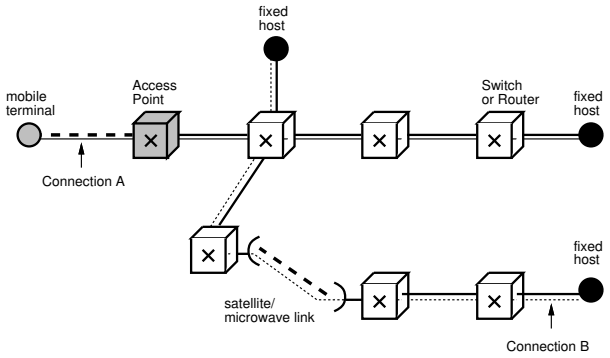


Fig. 1. Two scenarios for transporting traffic over a wireless link.

Packet losses are aggravated by the ARQ retransmission process, leading to a reduction in the effective service rate. Therefore, extra bandwidth must be assigned to compensate for such reduction. This is done by increasing the service rate and/or enhancing the error correction capability. Increasing the service rate reduces the PLR although the quality of the wireless channel stays unchanged. In contrast, increasing the effectiveness of FEC improves the quality of the wireless channel at the expense of extra bandwidth, which may in turn reduce the effective service rate. Thus, the FEC code rate has a subtle effect on the PLR. In general, adaptive coding techniques are expected to yield the optimal throughput for a channel with variable error statistics [17].

In this study, we address the problem of finding an optimal bandwidth and FEC code rate for a guaranteed PLR. To compute this bandwidth, which we refer to as the *wireless effective bandwidth*, we first evaluate the packet loss performance at the transmitter, taking into account the channel behavior and the underlying error control schemes (ARQ and/or FEC). For this purpose, we model each traffic source by an on-off fluid process and use a fluid version of Gilbert-Elliott (GE) model to capture the behavior of the wireless channel. Using fluid-flow analysis, we compute the PLR due to buffer overflow for a single and multiplexed streams. In the single-stream case, we also provide a closed-form approximation of the wireless effective bandwidth. Our results are used to study the impact of FEC on the effective bandwidth for a guaranteed PLR.

The rest of the paper is organized as follows. In Section II we describe the general framework of our study. The packet loss performance for a single stream is studied in Section III. The corresponding wireless effective bandwidth is analyzed in Section IV. The PLR for multiplexed streams is studied in Section V. Numerical results and simulations are reported in Section VI, followed by concluding remarks in Section VII.

II. Wireless Link Model

To analyze the PLR over a wireless link, we consider the framework shown in Figure 2. In this framework, traffic streams from one or more connections are fed into a finite-size FIFO buffer. The buffer is drained at a constant rate c (in packets/second), which corresponds to the capacity of an error-free wireless channel. The constant rate is produced, for example, by assigning slots periodically in a Time Division Multiple Access (TDMA) system or by using a weighted fair queueing algorithm [18]. A combination of ARQ and FEC is used to improve the transport performance of the wireless link. Several hybrid ARQ/FEC mechanisms have been proposed in the wireless literature [4], [20], which cannot all be captured in one model. In our study, we consider a particular combination in which cyclic redundancy check (CRC) is first applied to a packet, followed by FEC (i.e., the input to the FEC coder consists of the original packet plus its CRC code). We assume a very strong CRC code, with close to 100% error detection capability. In contrast, only a subset of packet errors can be corrected by FEC. Note that the purpose of FEC here is to reduce the number of retransmissions, and consequently the per-stream allocated bandwidth (as demonstrated in this paper). Another scenario, which is not considered here, involves a CRC code with partial error detection capability. We focus on the tuning of the FEC coder, treating the original packet and its CRC code as payload from the standpoint of the FEC coder. For simplicity, we ignore the overhead of the medium access control (MAC) layer.

At the receiver side, an arriving packet is passed through a FEC decoder followed by a CRC decoder. If one or more bit errors are detected by the CRC decoder, then a negative acknowledgement (NAK) is sent back to the sender. The sender must then retransmit the errored packet.

The above model has two control parameters: the service rate (or assigned bandwidth) and the FEC code rate. These parameters can be adjusted during connection setup to satisfy certain QoS requirements. From the network point of view, the selection of these parameters is very crucial and requires thorough understanding of their impact on the packet-level performance.

III. Packet Loss Performance for a Single Stream

In this section, we analyze the packet loss performance for a single traffic stream transported over a wireless link. The stream is characterized by an on-off fluid model, with

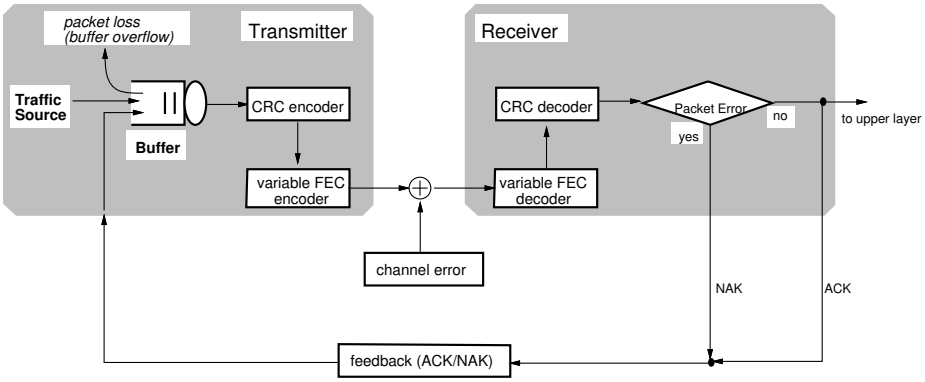


Fig. 2. Framework for analyzing the performance over a wireless link.

peak rate r and with exponentially distributed on and off periods with means $1/\alpha$ and $1/\beta$, respectively. At the bit level, the channel is represented by the two-state Markovian GE model, which is often used in performance studies of wireless links [21], [22]. As illustrated in Figure 3, in the GE model the channel alternates between *Good* and *Bad* states, with corresponding bit error rates (BER) P_{eg} and P_{eb} ($P_{eg} \ll P_{eb}$). The durations of the Good and Bad states are exponentially distributed with means $1/\delta$ and $1/\gamma$, respectively.

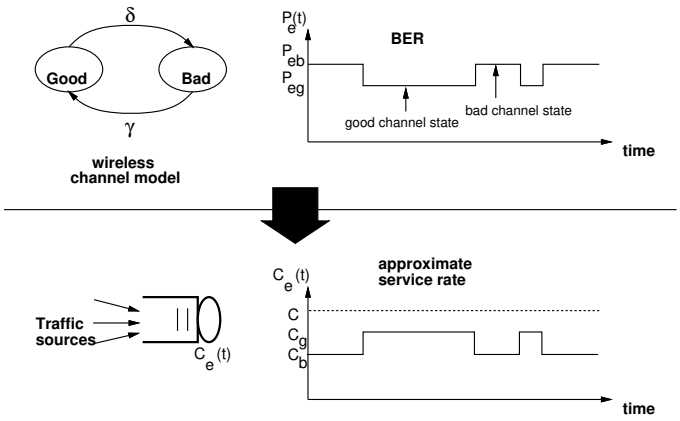


Fig. 3. Wireless channel model and corresponding service rate model.

The FEC capability of the underlying hybrid ARQ/FEC mechanism is characterized by three parameters: the number of bits in a code word (n), the number of payload bits (k), and the maximum number of correctable bits in a code block (τ). Note that n includes the k payload bits. The FEC code rate $e(\tau)$ is defined as

$$e(\tau) = \frac{k}{n(\tau)}.$$

Let τ be the maximum number of erroneous bits in a code block that can be corrected by FEC. Assuming that bit errors during a given channel state are independent, the probability that a received packet contains a non-correctable error when the channel is in the Good state is given by:

$$P_{c,g} = \sum_{j=\tau+1}^{n(\tau)} \binom{n(\tau)}{j} P_{eg}^{j} (1 - P_{eg})^{n(\tau)-j} \qquad (1)$$

A similar expression applies to $P_{c,b}$; the probability of a non-correctable packet error during Bad channel periods.

Incorporating the exact effects of ARQ and FEC in the underlying queuing model leads to intractable results. In particular, the actual pattern of retransmissions gives rise to a complicated service process; even in the simple case of stop-and-wait ARQ, the service process is a Markov-modulated Bernoulli process (the service time of a packet refers to the time it takes to *successfully* send that packet to the receiver, which includes all retransmission attempts). Therefore, approximations are needed. We assume that packet departures follow a fluid process whose service rate is modulated by the channel state (see Fig. 3). This approximation implies that there are two deterministic service rates: c_g during Good states and c_b during Bad states. For tractability purposes, our simplified model does not explicitly capture some aspects of wireless communications (e.g., interleaving). However, it can be argued [23] that from the network layer viewpoint, the packet transmission process is adequately approximated by the binary (success/failure) process, as done in our model. We assume that the feedback message from the receiver arrives back at the sender before the next transmission slot. Furthermore, we assume that the sender attempts to retransmit a packet until this packet is successfully received (i.e., number of bit errors is less than τ). In this scenario, the total time needed to successfully transmit a packet has a conditional geometric distribution (conditioned on the channel state). In this case, c_g and c_b correspond to the mean values of geometric distributions with parameters $1 - P_{c,g}$ and $1 - P_{c,b}$:

$$c_i = c \cdot e(\tau) \cdot (1 - P_{c,i}), \quad \text{for } i \in \{g, b\}$$

where c is the bandwidth assigned to the connection. The factor $e(\tau)$ in the above equation accounts for the FEC overhead, which reduces the effective service rate observed at the output of the network buffer. It can be seen that the problem of determining the wireless effective bandwidth

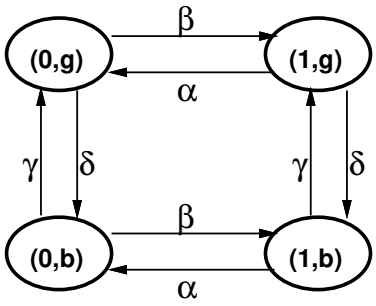


Fig. 4. State transition diagram.

reduces to obtaining a service rate c that satisfies a required PLR.

The underlying queueing system is controlled (modulated) by a four-state Markov chain with state space

$$S = \{(0,g),(1,g),(0,b),(1,b)\}$$

where 0 and 1 denote the on and off states of the traffic source, respectively, and g and b denote Good and Bad channel states, respectively. Figure 4 depicts the state transition diagram of the Markov chain.

Following a standard fluid approach (see [24], for example), the evolution of the buffer content can be described by the following differential equation:

$$\frac{d\boldsymbol{\Pi}(x)}{dx}\boldsymbol{D} = \boldsymbol{\Pi}(x)\boldsymbol{M} \tag{2}$$

where

$$\begin{aligned}
\boldsymbol{D} &\triangleq \text{diag}[-c_g, -c_b, r-c_g, r-c_b], \\
\Pi_s(x) &\triangleq \Pr\{\text{buffer content} \le x \text{ and system is in state } s\}, \\
\boldsymbol{\Pi}(x) &\triangleq [\ \Pi_{0,g}(x) \quad \Pi_{0,b}(x) \quad \Pi_{1,g}(x) \quad \Pi_{1,b}(x)\ ],
\end{aligned}$$

and $\boldsymbol{M}$ is the generator matrix of the underlying Markov chain:

$$\boldsymbol{M} = \begin{bmatrix} -(\beta+\delta) & \delta & \beta & 0 \\ \gamma & -(\beta+\gamma) & 0 & \beta \\ \alpha & 0 & -(\alpha+\delta) & \delta \\ 0 & \alpha & \gamma & -(\alpha+\gamma) \end{bmatrix}.$$

Throughout the paper, matrices and vectors are boldfaced.

The solution of (2) corresponds to the solution of the eigenvalue/eigenvector problem:

$$z\boldsymbol{\phi}\boldsymbol{D} = \boldsymbol{\phi}\boldsymbol{M} \tag{3}$$

which is generally given by

$$\boldsymbol{\Pi}(x) = \sum_{z_i \le 0} a_i \exp(z_i x)\phi_i$$

where a_i's are constant coefficients and the pairs $(z_i, \boldsymbol{\phi}_i), i = 1, 2, \cdots$, are the eigenvalues and right eigenvectors of the matrix $\boldsymbol{M}\boldsymbol{D}^{-1}$ [24], [25]. Let $\boldsymbol{w}$ denote the stationary probability vector of the Markov chain; $\boldsymbol{w}$ satisfies $\boldsymbol{w}\boldsymbol{M} = 0$ and $\boldsymbol{w}\boldsymbol{1} = 1$, where $\boldsymbol{1}$ is a column vector of ones. Then $\boldsymbol{w}$ is given by:

$$\boldsymbol{w} = \frac{1}{(\alpha+\beta)(\delta+\gamma)}[\ \alpha\gamma \quad \alpha\delta \quad \beta\gamma \quad \beta\delta\ ].$$

The mean drift is given by:

$$\boldsymbol{w}\boldsymbol{D}\boldsymbol{1} = \frac{\beta}{\alpha+\beta}r - \frac{c_g\gamma + c_b\delta}{\delta+\gamma}.$$

For a stable system, we have $\boldsymbol{w}\boldsymbol{D}\boldsymbol{1} < 0$.

In order to solve (3), we follow the approach used in [25]. The four-state Markov process is decomposed into two processes; one describes the on-off source and the other describes the state of the channel. These two processes are parameterized by the two generator matrices $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$, where

$$\boldsymbol{M}_1 = \begin{bmatrix} -\beta & \beta \\ \alpha & -\alpha \end{bmatrix} \text{ and } \boldsymbol{M}_2 = \begin{bmatrix} -\delta & \delta \\ \gamma & -\gamma \end{bmatrix} \tag{4}$$

with

$$\boldsymbol{M} = \boldsymbol{M}_1 \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{M}_2 \tag{5}$$

where $\otimes$ is the Kronecker product operator. For the drift matrix $\boldsymbol{D}$, we have

$$\boldsymbol{D} = r\boldsymbol{E}_r \otimes \boldsymbol{I} - \boldsymbol{I} \otimes \boldsymbol{E}_c \tag{6}$$

where $\boldsymbol{E}_r$ and $\boldsymbol{E}_c$ are given by

$$\boldsymbol{E}_r = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \boldsymbol{E}_c = \begin{bmatrix} c_g & 0 \\ 0 & c_b \end{bmatrix}. \tag{7}$$

Consider the following decomposition for the eigenvector of $\boldsymbol{M}\boldsymbol{D}^{-1}$:

$$\boldsymbol{\phi} = \boldsymbol{\phi}_1 \otimes \boldsymbol{\phi}_2 \tag{8}$$

where $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ are two dimensional vectors.

On substituting the expressions for $\boldsymbol{D}$, $\boldsymbol{M}$, and $\boldsymbol{\phi}$ from (6), (5), (8) into (3), we end up with

$$\boldsymbol{\phi}_1 \otimes (\boldsymbol{\phi}_2\boldsymbol{M}_2 + z\boldsymbol{\phi}_2\boldsymbol{E}_c) = (zr\boldsymbol{\phi}_1\boldsymbol{E}_r - \boldsymbol{\phi}_1\boldsymbol{M}_1) \otimes \boldsymbol{\phi}_2.$$

Hence, for (3) and (8) to be valid, it is sufficient that there exists some real number v such that both sides of the above equation equal $zv\boldsymbol{\phi}_1 \otimes \boldsymbol{\phi}_2$:

$$z\boldsymbol{\phi}_1[r\boldsymbol{E}_r - v\boldsymbol{I}] = \boldsymbol{\phi}_1\boldsymbol{M}_1 \tag{9}$$

$$z\boldsymbol{\phi}_2[v\boldsymbol{I} - \boldsymbol{E}_c] = \boldsymbol{\phi}_2\boldsymbol{M}_2 \tag{10}$$

which implies that

$$\begin{aligned}
\det[z\boldsymbol{I} - \boldsymbol{M}_1[r\boldsymbol{E}_r - v\boldsymbol{I}]^{-1}] &= 0 \\
\det[z\boldsymbol{I} - \boldsymbol{M}_2[v\boldsymbol{I} - \boldsymbol{E}_c]^{-1}] &= 0.
\end{aligned} \tag{11}$$

The first equation in (11) reduces to the quadratic polynomial

$$(zv)^2 - (zr + \alpha + \beta)zv + \beta zr = 0.$$

Thus,

$$zv = \frac{zr + \alpha + \beta}{2} \pm \frac{\sqrt{Q_1(z)}}{2} \tag{12}$$

where

$$Q_1(z) = (zr + \alpha - \beta)^2 + 4\alpha\beta. \tag{13}$$

For the second equation in (11), we have

$$(zv)^2 - (z(c_g + c_b) - \delta - \gamma)zv + z(zc_gc_b - \gamma c_g - \delta c_b) = 0.$$

Thus,

$$zv = \frac{z(c_g + c_b) - \delta - \gamma}{2} \pm \frac{\sqrt{Q_2(z)}}{2} \tag{14}$$

where

$$Q_2(z) = (z(c_g - c_b) - (\delta - \gamma))^2 + 4\delta\gamma. \tag{15}$$

Equating (12) and (14) leads to the following four equations that provide four distinct eigenvalues:

$$\begin{aligned}
&z(r - c_g - c_b) \\
&\quad + \alpha + \beta + \delta + \gamma + \sqrt{Q_1(z)} + \sqrt{Q_2(z)} = 0 \\
&z(r - c_g - c_b) \\
&\quad + \alpha + \beta + \delta + \gamma + \sqrt{Q_1(z)} - \sqrt{Q_2(z)} = 0 \\
&z(r - c_g - c_b) \\
&\quad + \alpha + \beta + \delta + \gamma - \sqrt{Q_1(z)} + \sqrt{Q_2(z)} = 0 \\
&z(r - c_g - c_b) \\
&\quad + \alpha + \beta + \delta + \gamma - \sqrt{Q_1(z)} - \sqrt{Q_2(z)} = 0
\end{aligned}$$

From the above set of equations, we can establish the following equation by eliminating the square roots:

$$((z(r - c_g - c_b) + \alpha + \beta + \delta + \gamma)^2 - Q_1(z) - Q_2(z))^2 - 4Q_1(z)Q_2(z) = 0. \tag{16}$$

Rearranging (16), we obtain a polynomial of order three, whose real roots correspond to the three nonzero eigenvalues.

After obtaining the eigenvalues, the corresponding values for v can be obtained using (12) and (14). Once z and v are determined, the corresponding eigenvectors can be obtained from (9) and (10):

$$\boldsymbol{\phi}_1 = [\ \alpha \quad -zv + \beta\ ] \tag{17}$$

$$\boldsymbol{\phi}_2 = [\ \gamma \quad zv - zc_g + \delta\ ]. \tag{18}$$

Thus,

$$\begin{aligned}
\boldsymbol{\phi} &= \boldsymbol{\phi}_1 \otimes \boldsymbol{\phi}_2 \\
&= \begin{bmatrix} \alpha\gamma \\ \alpha(zv - zc_g + \delta) \\ (-zv + \beta)\gamma \\ (-zv + \beta)(zv - zc_g + \delta) \end{bmatrix}^T
\end{aligned} \tag{19}$$

where $[\boldsymbol{A}]^T$ is the transpose of the matrix $\boldsymbol{A}$. Finally, the stationary buffer content distribution $\boldsymbol{\Pi}(x)$ is given by

$$\boldsymbol{\Pi}(x) = a_d\boldsymbol{w} + \sum_{z_i<0} a_i \exp(z_i x)\boldsymbol{\phi}_i \tag{20}$$

where a_d is the coefficient associated with the zero eigenvalue. For a stable system, the coefficients a_i's associated with positive eigenvalues are set to zero.

For an infinite buffer system, we have the following boundary conditions [25]:

$$\begin{aligned}
\Pi_s(0) = 0, &\quad \text{for } z_i < 0, \text{with } r > c_g, r > c_b \\
a_d = 1, &\quad (\text{since } \boldsymbol{\Pi}(\infty) = \boldsymbol{w}).
\end{aligned}$$

If $r > c_g$, we have two negative eigenvalues, z_1 and z_2. In this case, the coefficients are given by

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -\begin{bmatrix} \phi_{1g}(z_1) & \phi_{1g}(z_2) \\ \phi_{1b}(z_1) & \phi_{1b}(z_2) \end{bmatrix}^{-1} \begin{bmatrix} w_{1g} \\ w_{1b} \end{bmatrix}$$

where $\phi_s(z)$ is the eigenvector element of state s by (19).

If $r < c_g$, there is only one negative eigenvalue z_1. In this case, the coefficient a_1 is simply given by:

$$a_1 = -w_{1b}/\phi_{1b}(z_1).$$

After obtaining the eigenvalues, the eigenvectors, and the coefficients, we can construct the stationary buffer content distribution $\boldsymbol{\Pi}(x)$. Consequently, the PLR due to buffer overflow $G(x)$ is given by:

$$G(x) = 1 - \mathbf{1}\boldsymbol{\Pi}(x). \tag{21}$$

IV. Wireless Effective Bandwidth

The expression for the PLR obtained in the previous section can, in principle, be used to compute the effective bandwidth. However, this requires expressing the service rate c as a function of other variables (PLR, buffer size, channel BERs, the number of correctable bits, etc.). In general, it is not possible to obtain an *exact* closed-form expression for the effective bandwidth[1]; even for a single source (i.e., no multiplexing), the closed-form solution cannot be obtained without approximation [10], [9].

In this section, we derive an approximate expression for the effective bandwidth following the approach used in [8]. In [8] the authors consider the service rate c to be a variable parameter and the eigenvalues to be functions of c, i.e., $z =$

[1] In this paper, we use the term 'effective bandwidth' in a general sense to refer to the minimum service capacity that is needed to achieve a given PLR. We use the term 'effective bandwidth approximation' to refer to the more specific definition of Elwalid et al. [8].

$f(c)$. Since the problem is to obtain c for a given z, c can be expressed as the inverse function, i.e., $c = f^{-1}(z) \stackrel{\triangle}{=} g(z)$. The key point in the analysis is that this inversion problem reduces to another eigenvalue problem (see [8] for details).

Consider (3). In this case, the drift matrix $\boldsymbol{D}$ can be written as

$$\boldsymbol{D} = r\boldsymbol{B}_r - ce(\tau)\boldsymbol{B}_c$$

where

$$\boldsymbol{B}_r = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{B}_c = \begin{bmatrix} \eta_g & 0 & 0 & 0 \\ 0 & \eta_b & 0 & 0 \\ 0 & 0 & \eta_g & 0 \\ 0 & 0 & 0 & \eta_b \end{bmatrix}$$

$\eta_g = 1 - P_{c,g}$, and $\eta_b = 1 - P_{c,b}$. Substituting $\boldsymbol{D}$ into (3), we obtain the following relation:

$$z\boldsymbol{\phi}(r\boldsymbol{B}_r - ce(\tau)\boldsymbol{B}_c) \quad = \quad \boldsymbol{\phi}\boldsymbol{M}.$$

Let $c \stackrel{\triangle}{=} g(z)$. Then,

$$zr\boldsymbol{\phi}\boldsymbol{B}_r - zg(z)e(\tau)\boldsymbol{\phi}\boldsymbol{B}_c \quad = \quad \boldsymbol{\phi}\boldsymbol{M}.$$

Rearranging the previous equation, we obtain

$$g(z)e(\tau)\boldsymbol{\phi} \quad = \quad \boldsymbol{\phi}\left(-\frac{1}{z}\boldsymbol{M} + r\boldsymbol{B}_r\right)\boldsymbol{B}_c^{-1}. \qquad (22)$$

Note that the problem of obtaining $g(z)$ translates into another eigenvalue problem. According to [8], the effective bandwidth is approximated by the maximal eigenvalue $g(z)$ satisfying (22).

From (22) and after some tedious algebraic manipulations, we obtain the following four eigenvalues:

$$g(z)e(\tau) = \begin{cases} \frac{C_1 - (\eta_g+\eta_b)C_2 \pm \sqrt{2(C_3 - C_2C_4)}}{4\eta_g\eta_b} \\ \frac{C_1 + (\eta_g+\eta_b)C_2 \pm \sqrt{2(C_3 + C_2C_4)}}{4\eta_g\eta_b} \end{cases}$$

where

$$\begin{aligned} C_1 &= (\eta_g+\eta_b)r - ((\alpha+\beta+2\gamma)\eta_g + (\alpha+\beta+2\delta)\eta_b)\xi \\ C_2 &= \sqrt{(r-(\alpha-\beta)\xi)^2 + 4\alpha\beta\xi^2} \\ C_3 &= \eta_b^2((r-(\alpha+\delta)\xi)^2 + ((\beta+\delta)^2 + 2\alpha\beta)\xi^2) \\ &\quad -2\eta_g\eta_b(r^2 - (2\alpha+\delta+\gamma)r\xi \\ &\quad +((\alpha+\beta)(\alpha+\beta+\delta+\gamma) - 2\delta\gamma)\xi^2) \\ &\quad +\eta_g^2((r-(\alpha+\gamma)\xi)^2 + ((\beta+\gamma)^2 + 2\alpha\beta)\xi^2) \\ C_4 &= (\eta_g-\eta_b)(\eta_g(r-(\alpha+\beta+2\gamma)\xi) \\ &\quad -\eta_b(r-(\alpha+\beta+2\delta)\xi)) \\ \xi &= -B/\log p,\ B \text{ is buffer size and } p \text{ is PLR.} \end{aligned}$$

Thus, the maximal eigenvalue, corresponding to the wireless effective bandwidth approximation, is given by:

$$g(z) \quad = \quad \frac{C_1 + (\eta_g+\eta_b)C_2 + \sqrt{2(C_3 + C_2C_4)}}{4\eta_g\eta_b e(\tau)}. \qquad (23)$$

Note that this is an asymptotic result, i.e., the buffer size is very large and the PLR is very small.

V. Packet Loss Performance for Multiplexed Streams

As indicated in Figure 1, it is also possible to multiplex several connections onto the same wireless link. In this section, we extend the previous packet loss analysis to the case of multiplexed streams.

The wireless link is modeled as a multiplexer with a randomly varying service rate. This model is similar to the producer-consumer model investigated in [25]. In [25] the traffic generation (production) and the packet delivery (consumption) processes are coupled by a buffer, which enables their decomposition and consequently facilitates the analysis. In the case of a single connection, the traffic generation and delivery processes are specified by (9) and (10), respectively. For K multiplexed sources, the traffic generation part is governed by the following equation [25]:

$$z\boldsymbol{\phi}_s[\boldsymbol{\Lambda}_s - v\boldsymbol{I}] = \boldsymbol{\phi}_s\boldsymbol{M}_s \qquad (24)$$

where $\boldsymbol{M}_s$ and $\boldsymbol{\Lambda}_s$ are the infinitesimal generator and rate matrices for the aggregate traffic, respectively. These matrices can be written as

$$\begin{aligned} \boldsymbol{M}_s &= \boldsymbol{M}_1 \oplus \boldsymbol{M}_2 \oplus \cdots \oplus \boldsymbol{M}_K \\ \boldsymbol{\Lambda}_s &= \boldsymbol{\Lambda}_1 \oplus \boldsymbol{\Lambda}_2 \oplus \cdots \oplus \boldsymbol{\Lambda}_K \end{aligned}$$

where $\boldsymbol{M}_i$ and $\boldsymbol{\Lambda}_i$ are the infinitesimal generator and rate matrices of the ith source, $i = 1, 2, \cdots, K$. In (24), z and $\boldsymbol{\phi}_s$ are the eigenvalue and eigenvector of the matrix $\boldsymbol{M}_s[\boldsymbol{\Lambda}_s - v\boldsymbol{I}]^{-1}$, respectively. The operator $\oplus$ denotes the Kronecker sum.

As for the packet delivery process, it is governed by

$$z\boldsymbol{\phi}_r[v\boldsymbol{I} - \boldsymbol{E}_c] \quad = \quad \boldsymbol{\phi}_r\boldsymbol{M}_r \qquad (25)$$

where $\boldsymbol{M}_r$ is the generator matrix for the service process (the consumption part), and z and $\boldsymbol{\phi}_r$ are the corresponding eigenvalue and eigenvector. For the two-state GE channel model, $\boldsymbol{M}_r$ and $\boldsymbol{E}_c$ correspond to $\boldsymbol{M}_2$ in (4) and $\boldsymbol{E}_c$ in (7), respectively.

Exact analysis of the producer-consumer system described by (24) and (25) was provided in [25]. However, the buffer consumption rates in [25] were taken to be integer multiples of some unit rate. In the appendix, we provide the exact solution for the PLR based on Mitra's results, with adaptation to our wireless link model. Note that in our model, there are two consumption rates (c_b and c_g). But c_g is not necessarily an integer multiple of c_b.

While an exact solution for the PLR is feasible, the computational complexity associated with this solution is rather high. A simpler yet sufficiently accurate approximation can be obtained based on the Chernoff-dominant eigenvalue (CDE) approach. This approach was previously used by Elwalid et al. [13] to approximate the PLR at a constant-rate ATM multiplexer fed by general Markovian sources. In this approximation, the PLR at the multiplexer with buffer size x is approximated by

$$G(x) \approx Le^{zx} \qquad (26)$$

where z is the dominant eigenvalue and L is the corresponding coefficient. A procedure was presented for computing z and L for a multiplexer with a constant service rate c. A similar procedure will be used in this paper to approximate the PLR at a wireless link.

A. Calculation of the Dominant Eigenvalue

According to the results in [8], the dominant eigenvalue of (24) is the unique solution of the equation:

$$\sum_{i=1}^{K} g_i(z) = v \tag{27}$$

where $g_i(z) = \text{MRE}[\mathbf{\Lambda}_i - \boldsymbol{M}_i/z]$ and $\text{MRE}[\boldsymbol{A}]$ denotes the maximal real eigenvalue of the matrix $\boldsymbol{A}$. Note that v is a variable coupling (24) and (25), which differs from [8] in that the constant rate c is replaced by the variable v.

Consider the packet delivery part corresponding to (25). From (14) v is given by

$$v = \frac{z(c_g + c_b) - \delta - \gamma \pm \sqrt{Q_2(z)}}{2z}. \tag{28}$$

We can obtain the dominant eigenvalue by equating the left-hand side of (27) and the right-hand side of (28). Since the dominant eigenvalue is the unique solution of (27) and z is a decreasing function in v, the smaller value of v in (28), i.e., the term containing $+\sqrt{Q_2(z)}$, must be equated with (27) to obtain the dominant eigenvalue:

$$\sum_{i=1}^{K} g_i(z) = \frac{z(c_g + c_b) - \delta - \gamma + \sqrt{Q_2(z)}}{2z}.$$

Thus, the following proposition can be established.

Proposition V.1: The dominant eigenvalue in the wireless multiplexer model represented by (24) and (25) is the value of z satisfying the following equation:

$$\sum_{i=1}^{K} g_i(z) = \frac{z(c_g + c_b) - \delta - \gamma + \sqrt{Q_2(z)}}{2z}. \tag{29}$$

For a single on-off source, $g(z)$ is given by:

$$g(z) = \frac{(rz + \alpha + \beta) - \sqrt{(rz + \alpha - \beta)^2 + 4\alpha\beta}}{2z}$$

which is the same as (12).

B. Approximation of the Coefficient L

Previous studies have shown that the coefficient L in (26) plays a critical role in obtaining an accurate estimate of the PLR [10], [13]. In this section, we provide an approximation of L using a Chernoff-bound approach, in line of Elwalid et al.'s work [13] on analyzing the PLR at a multiplexer with a constant drain rate.

Following the discussion in [13] and the references therein, one can approximate the constant L by $G(0)$, implying that L is *approximately* the packet loss probability at a bufferless multiplexer:

$$G(0) \approx L. \tag{30}$$

In a bufferless multiplexer, packet losses occur when the input rate exceeds the service rate. Let χ_i denotes the arrival rate of source i at steady state. The total traffic generation rate from K sources is $\chi = \sum_{i=1}^{K} \chi_i$. For a multiplexer with a fixed service rate R, L is estimated by

$$L \approx P[\chi \geq R].$$

Using Chernoff bound, Elwalid et al. [13] showed that as $R \to \infty$ with $K/R = o(1)$, $P[\chi \geq R]$ can be obtained as follows:

$$P[\chi \geq R] = \frac{\exp(-F(s^*))}{s^*\sigma(s^*)\sqrt{2\pi}}\,[1 + o(1)] \tag{31}$$

where

$$\begin{aligned} F(s) &\triangleq sR - \sum_{i=1}^{K} \log N_i(s) \\ N_i(s) &\triangleq E[\exp(s\chi_i)] \\ \sigma^2(s) &= \frac{\partial^2 \log E[\exp(s\chi)]}{\partial s^2} \\ &= \sum_{i=1}^{K} \left[\frac{N_i''(s)}{N_i(s)} - \left(\frac{N_i'(s)}{N_i(s)} \right)^2 \right] \end{aligned}$$

and $F(s^*) = \sup_{s \geq 0} F(s)$.

In our case, R is not fixed, but can take one of two values (c_g and c_b). Losses will occur when χ exceeds c_g during the Good periods and c_b during the Bad periods. Thus, L in (26) is given by:

$$L \approx G(0) = P[\chi \geq c_g]w_g + P[\chi \geq c_b]w_b \tag{32}$$

where w_g and w_b are the steady-state probabilities that the channel is in Good and Bad states, respectively; $w_g = \gamma/(\delta + \gamma)$, $w_b = \delta/(\delta + \gamma)$. The probabilities $P[\chi \geq c_g]$ and $P[\chi \geq c_b]$ are obtained by substituting c_g and c_b for R in (31). For K homogeneous on-off sources with a constant service rate R, we have

$$\begin{aligned} N_i(s) &= w_1 e^{sr} + w_0 \\ F(s) &= sR - K \log(w_1 e^{sr} + w_0) \\ \sigma^2(s) &= \frac{K w_1 w_0 r^2 e^{sr}}{(w_1 e^{sr} + w_0)^2} \end{aligned}$$

where r is the source peak rate, w_1 and w_0 are the steady-state probabilities that a source is in on and off states, respectively; $w_1 = \beta/(\alpha + \beta)$ and $w_0 = \alpha/(\alpha + \beta)$. For $R < Kr$, $F(s)$ is a concave function with s^* given by

$$s^* = \frac{1}{r} \log \frac{Rw_0}{(Kr - R)w_1}. \tag{33}$$

Replacing R in the above equation by c_g and c_b, one at a time, we obtain corresponding values s_g^* and s_b^*, which are then used in (32) to evaluate $P[\chi \geq c_g]$ and $P[\chi \geq c_b]$. In summary, for K homogeneous on-off sources that are

multiplexed onto the wireless link, L is given by

$$\begin{aligned} L &= P[\chi \geq c_g]w_g + P[\chi \geq c_b]w_b \\ &= \frac{e^{-s_g^* c_g}(w_1 e^{s_g^* r} + w_0)^{K+1} w_g}{s_g^* r\sqrt{2\pi K w_1 w_0 e^{s_g^* r}}} \\ &\quad + \frac{e^{-s_b^* c_b}(w_1 e^{s_b^* r} + w_0)^{K+1} w_b}{s_b^* r\sqrt{2\pi K w_1 w_0 e^{s_b^* r}}} \end{aligned} \tag{34}$$

VI. Numerical Results and Simulations

In this section, we verify the adequacy of our analytical approximations by contrasting them against more realistic simulations. Similar to the analysis, the simulation results are obtained using on-off traffic sources with exponentially distributed on and off periods. However, the length of an on period in the simulations is truncated to obtain an integer number of packets. The ARQ retransmission process is simulated in a more realistic manner, whereby a packet is transmitted repeatedly until it is received with no errors. The probability of a packet error is computed from (1) for both channel states. Conditioned on the state of the channel, the number of retransmissions of a given packet in the simulation setup follows a geometric distribution. Transitions between Good and Bad states are assumed to occur only at the beginning of a packet transmission slot. It is assumed that the propagation delay is relatively small, so that the ACK/NAK message for a packet is received at the sender before the next transmission slot. Finally, we use a finite buffer size in the simulations, in contrast to the infinite-buffer assumption in the analysis.

In our experiments, we vary the buffer size and the BER during the Bad state (P_{eb}), and we fix the BER during the Good state at $P_{eg} = 10^{-6}$. We set the mean of the off period to ten times that of the on period. In addition, we take the parameters related to the wireless channel from [22]. Without loss of generality, we adopt Bose-Chaudhuri-Hocquenghem (BCH) code [26] for FEC . The input to the FEC encoder consists of 53-byte packets (i.e., $k = 424$ bits). The error correction capability used in the numerical examples ranges from $\tau = 0$ ($\epsilon(\tau) = 1$) to $\tau = 20$ ($\epsilon(\tau) = 0.7$). All simulation results are reported with 95% confidence intervals. Table I summarizes the values of the various parameters in the simulations and numerical examples.

Figure 5 shows the PLR as a function of the buffer size when $P_{eb} = 0.01$, $\tau = 0, 4$, and $c = 500$ packets/sec. It is observed that the analytic results slightly overestimate the simulated values. The gap is negligible over the whole range of buffer sizes. As expected, the PLR increases as the buffer size decreases and as τ decreases.

Figure 6 depicts the PLR versus the buffer size using $c = 900$ packets/sec ($P_{eb} = 10^{-2}, \tau = 8, 20$). As c increases, the analytic results begin to slightly underestimate the PLR. This trend is shown clearly in Figures 7 and 8, which depict the PLR as a function of the service rate c using two different buffer sizes and two FEC code rates. The inaccuracy of the approximate analytical results is acceptable even for higher τ (e.g., $\tau = 20$), given that PLRs are often contrasted in the orders of magnitude by which they differ.

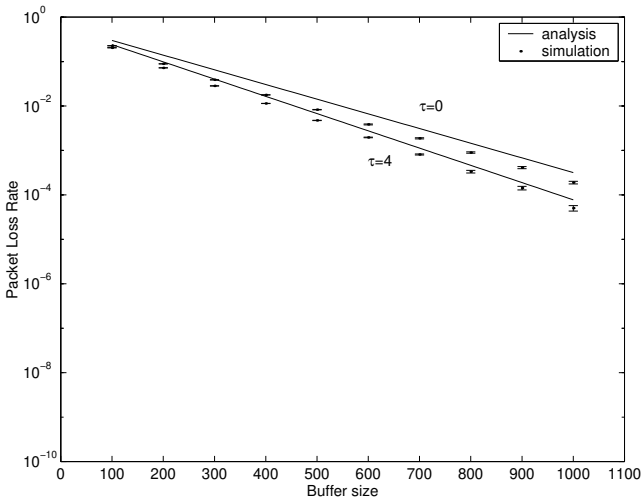


Fig. 5. PLR versus buffer size for a single stream ($c = 500$, $P_{eb} = 10^{-2}$).

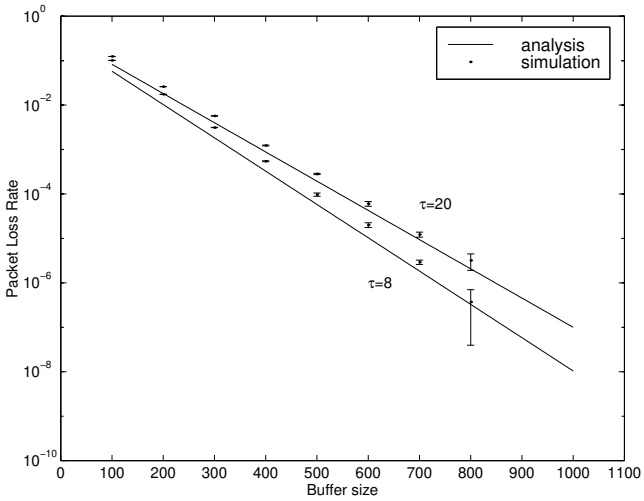


Fig. 6. PLR versus buffer size for a single stream ($c = 900$ packets/sec, $P_{eb} = 10^{-2}$).

Figure 9 depicts the exact and approximate wireless effective bandwidth versus the PLR. Only ARQ ($\tau = 0$) is used in obtaining these figures; the case of hybrid ARQ/FEC will be discussed later. The exact value is obtained using all eigenvalues, while the approximate one is obtained based on the analysis in Section IV. It can be observed that the effective bandwidth approximation is slightly pessimistic, particularly at small PLRs. In the case of multiplexed streams (not shown), we have observed that this approximation becomes quite conservative, in line of similar findings for wireline multiplexers (see [10], for example). For this reason, the use of this approximation is to be restricted to the single-stream case.

In Figure 9 when $P_{eb} = 10^{-4}$, the wireless effective bandwidth (based on single eigenvalue) ranges from 1340 to 1850 packets/sec for a PLR in the range 10^{-4} to 10^{-7}. Recall

TABLE I

PARAMETER VALUES USED IN THE SIMULATIONS AND NUMERICAL RESULTS.

Parameter	Symbol	Value
peak rate	r	1 Mbps (or 2604.1667 packets/sec)
buffer size	B	$100 - 1000$ packets
mean on period	$1/\alpha$	0.02304 sec
mean off period	$1/\beta$	0.2304 sec
mean Good channel period	$1/\delta$	0.1 sec
mean Bad channel period	$1/\gamma$	0.0333 sec
BER in Good channel state	P_{eg}	10^{-6}
BER in Bad channel state	P_{eb}	$10^{-2} - 10^{-5}$

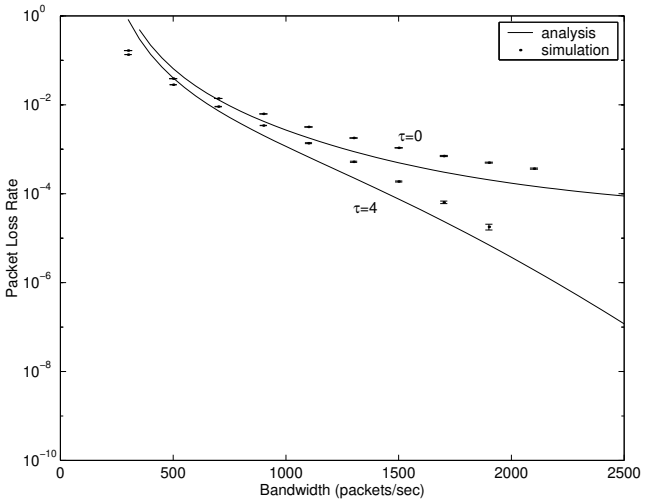


Fig. 7. PLR versus c for a single stream ($B = 300$, $P_{eb} = 10^{-2}$).

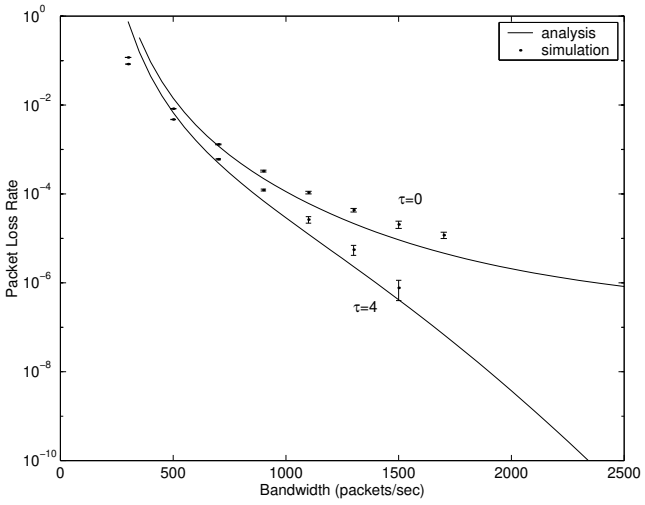


Fig. 8. PLR versus c for a single stream ($B = 500$, $P_{eb} = 10^{-2}$).

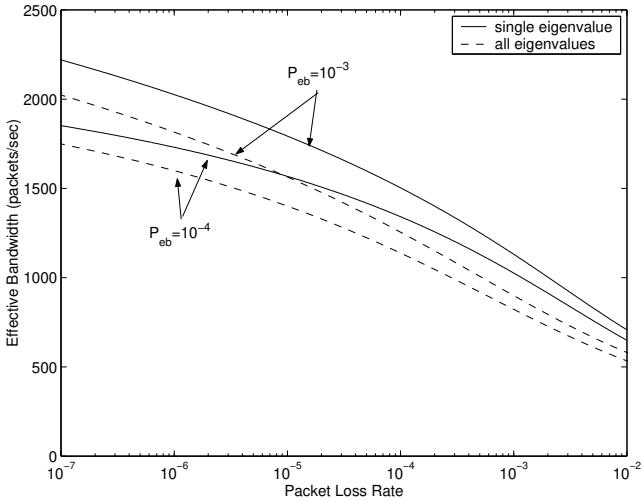


Fig. 9. Effective bandwidth versus PLR for one stream ($B = 300$, $\tau = 0$).

that the mean and peak rates of a traffic source are 263.74 and 2604.1667 packets/sec, respectively. As P_{eb} increases, the wireless effective bandwidth gets closer to the source peak rate. For example, when $P_{eb} = 10^{-3}$ the wireless effective bandwidth ranges from 1501 to 2216 packets/sec for a PLR in the range 10^{-4} to 10^{-7}. This trend is further illustrated in Figure 10, which depicts the wireless effective bandwidth as a function of P_{eb} for three target PLRs using $B = 800$ packets. Only ARQ is considered. Interestingly, the wireless effective bandwidth increases by a noticeable amount when P_{eb} goes from 10^{-3} to 10^{-2}. This is due to the fact that increasing P_{eb} from 10^{-3} to 10^{-2} has a relatively significant effect on the packet error probability. For example, when P_{eb} goes from 10^{-5} to 10^{-4}, $P_{c,b}$ increases from 0.004231 to 0.04151, resulting in a slight decrease in c_b (through the term $1 - P_{c,b}$). In contrast, when P_{eb} goes from 10^{-3} to 10^{-2}, $P_{c,b}$ increases from 0.3457 to 0.9858.

Figures 11-13 show the approximate wireless effective bandwidth versus the number of correctable bits (τ) for different target PLRs and different BERs. The buffer size is fixed at 800 packets. These figures clearly indicate the existence of an "optimal" FEC code rate at which the allocated bandwidth is minimized for a given target PLR. For instance, to guarantee a PLR of 10^{-6} when $P_{eb} = 0.001$,

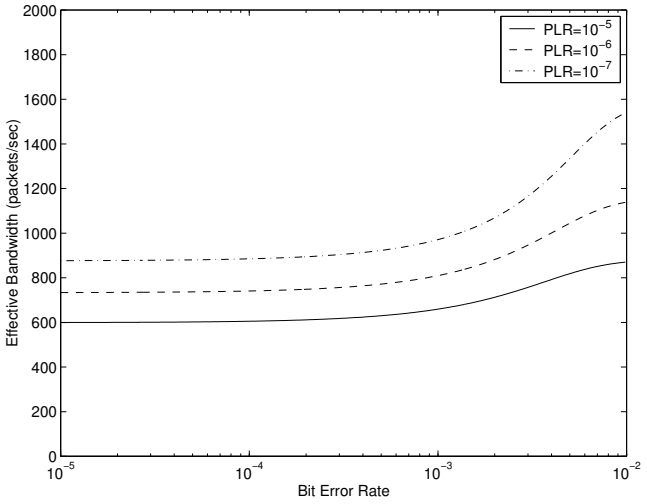


Fig. 10. Effective bandwidth approximation versus P_{eb} for a single stream ($B = 800$, $\tau = 0$).

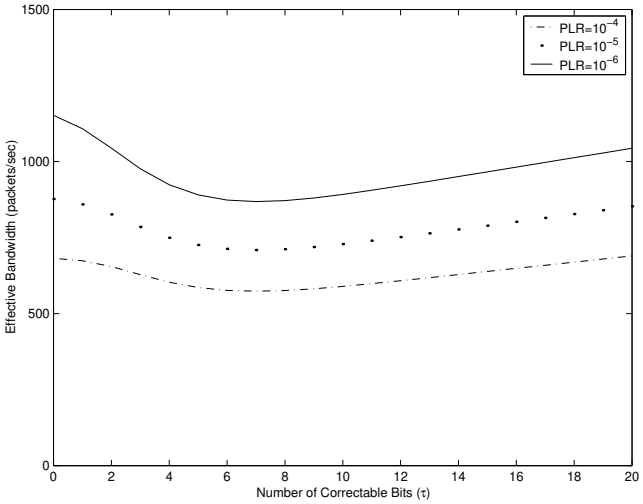


Fig. 12. Effective bandwidth versus τ for a single stream ($P_{eb} = 0.01$).

the minimum required bandwidth is $c = 761.9$ packets/sec, which is achieved when $\tau = 1$ (FEC corrects only single-bit errors). At $P_{eb} = 0.001$ the mean number of bit errors in a packet is $0.423 < 1$, so the advantage of correcting multiple-bit errors is overshadowed by the extra FEC bandwidth overhead. We have noticed that the optimal code rate is always achieved at a nonzero τ, implying that with appropriate tuning of the FEC code rate (e.g., by manipulating the puncturing rate), it is always beneficial from resource allocation standpoint to employ a hybrid ARQ/FEC scheme to guarantee a given PLR. It can be observed from Figures 11-13 that the optimal code rate increases with P_{eb}. For $P_{eb} = 0.01$ and $P_{eb} = 0.015$, the optimal values of τ are given by 7 and 10, respectively. The dependence of the optimal code rate on the channel BERs suggests the need for adaptive FEC to continuously track an optimal code rate for a channel with time-varying characteristics.

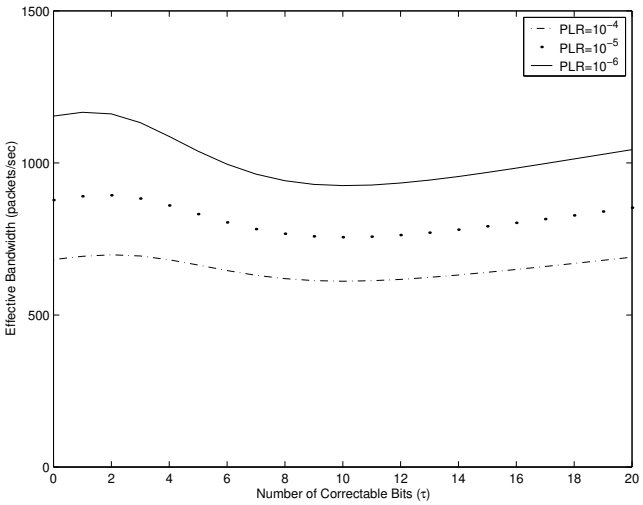


Fig. 13. Effective bandwidth versus τ for a single stream ($P_{eb} = 0.015$).

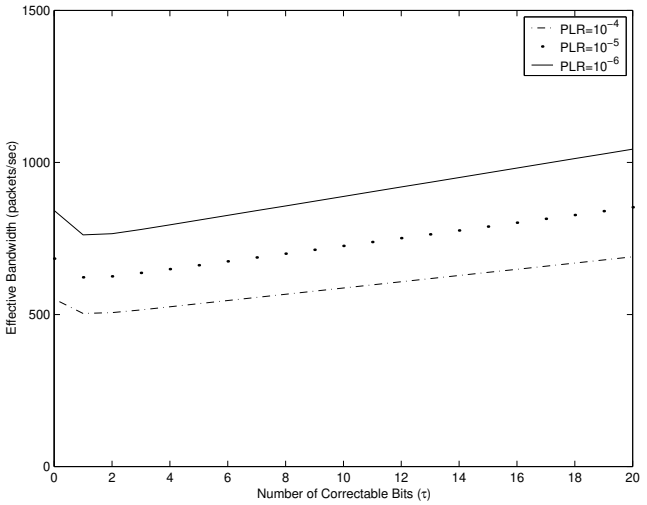


Fig. 11. Effective bandwidth versus τ for a single stream ($P_{eb} = 0.001$).

Next, we examine the PLR for multiplexed streams based on the results in Section V. We consider homogeneous on-off sources having the same target PLR. For brevity, we report the results for ARQ only (i.e., $\tau = 0$). We set $P_{eb} = 10^{-3}$ and $c = 10$ Mbps, and we vary the number of multiplexed streams. To test the goodness of our PLR analysis, which was obtained using the CDE approximation, we compare it to the exact analysis given in the appendix. We also present the results based on two other approximations: the asymptotic and Binomial approximations. Similar to the CDE approach, both approximations are based on the dominant eigenvalue. However, they differ from it in the estimation of L; the coefficient associated with the dominant eigenvalue. In the asymptotic approximation, the exact value of L, denoted by L_a, is used, which is obtained from the exact analysis in the appendix. The Binomial approximation estimates L by L_b, which is com-

puted from the binomial distribution:

$$L_b \approx G(0) = P[\chi > c] = \sum_{i=j}^{K} \left(\begin{array}{c} K \\ i \end{array} \right) w_1^i w_0^{K-i}$$

where $j \stackrel{\triangle}{=} \lceil c/r \rceil$.

Figures 14, 15, and 16 depict the PLR for 40, 50, and 60 multiplexed connections, respectively, which correspond to traffic loads of 43.6%, 54.5%, and 65.4%. It is observed that the asymptotic approximation is quite accurate compared with the exact result. Thus, it can be argued that one eigenvalue is good enough to compute the PLR for multiplexed streams. As shown in these figures, the difference between the exact results and both the CDE and Binomial approximations tends to decrease as K increases. Both approximations provide upper bounds on the exact PLR, with the Binomial approximation being the tighter of the two. Since in our case only homogeneous on-off sources are considered, the binomial approximation is rather easy to compute. However, for general heterogeneous sources, this will incur substantial numerical complexity.

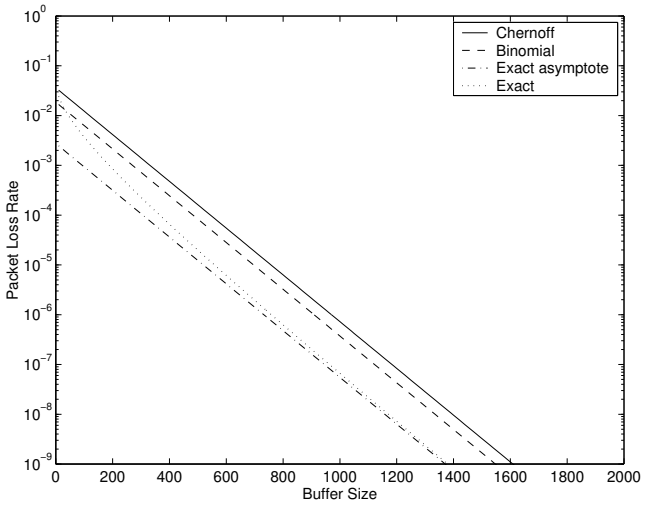


Fig. 14. PLR versus buffer size (40 multiplexed streams).

The three approximations for the PLR are compared in Table II. Under light load ($K = 10, 20, 30$), both the CDE and Binomial approximations deviate largely from the ideal asymptote. Thus, a better approximation for L is still needed when K is relatively small. As K increases, both approximations become more accurate in estimating the actual value of L.

VII. Conclusions

In this paper, we investigated the packet loss performance due to buffer overflow at the transmitter side of a wireless link. Traffic burstiness was captured through on/off fluid models. The fluctuations of the wireless channel were appropriately captured using a fluid version of the Gilbert-Elliot model. Error control schemes (ARQ and FEC) were incorporated. The packet loss performance was analyzed for a single and for multiplexed streams. In the

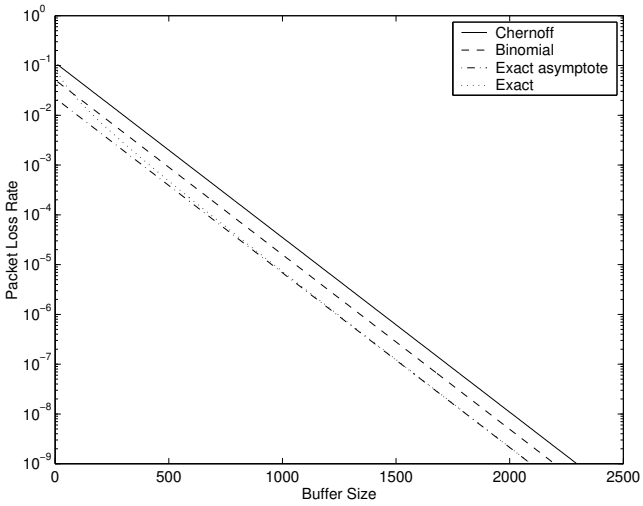


Fig. 15. PLR versus buffer size (50 multiplexed streams).

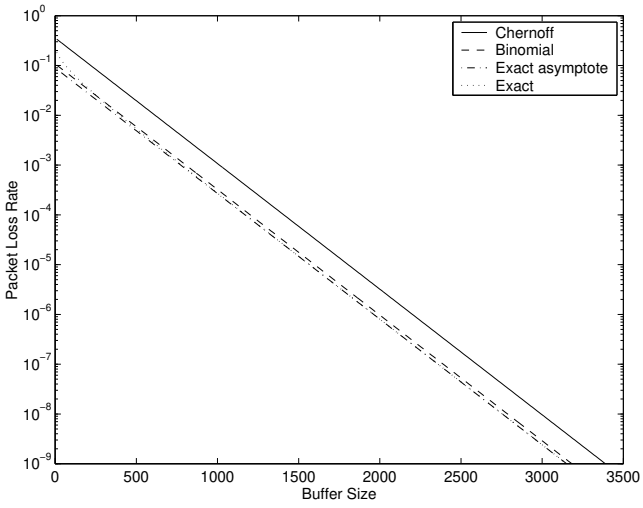


Fig. 16. PLR versus buffer size (60 multiplexed streams).

case of a single stream, the exact PLR was obtained, which only requires the solution of a polynomial of order three. The solution was then used to obtain an approximation for the wireless effective bandwidth, which can be used as a valuable tool in resource allocation and admission control in wireless networks. In the case of multiplexed streams, an approximate expression for the PLR was obtained. By contrasting the analytical results with more realistic simulations, we observed that in the case of a single stream, the analytical expressions for the PLR and (approximate) wireless effective bandwidth are acceptable over a wide range of bit error rates. The approximate PLR for multiplexed streams is adequate under medium and high loads. A better approximation is needed when the load is light. We found that with proper tuning of the FEC code rate, hybrid ARQ/FEC is always more beneficial in terms of reducing the per-stream bandwidth requirement than ARQ alone. The selection of the FEC code rate is crucial to achieving the optimal use of the wireless bandwidth. Our focus in this paper was on the packet loss rate as the primary QoS

TABLE II
THE COEFFICIENTS L_c, L_b, L_a VERSUS THE NUMBER OF MULTIPLEXED CONNECTIONS K.

K	L_c (CDE)	L_b (Binomial)	L_a (Asymptotic)	L_c/L_a	L_b/L_a
10	4.763×10^{-6}	1.2×10^{-6}	1.151×10^{-8}	413.75	104.31
20	7.667×10^{-4}	3.441×10^{-4}	8.254×10^{-7}	928.92	416.88
30	8.078×10^{-3}	4.169×10^{-3}	1.061×10^{-4}	76.13	39.29
40	0.0361371	0.0185953	0.002747	13.15	6.77
50	0.113547	0.0514114	0.02214	5.13	2.32
60	0.358695	0.107971	0.08951	4.0	1.2

measure. In the future, we plan to investigate the delay performance over the wireless link with an imposed limit on the number of packet retransmissions.

APPENDIX
EXACT PACKET LOSS PERFORMANCE FOR MULTIPLEXED ON-OFF SOURCES

Consider K homogeneous on-off sources that are multiplexed over a wireless channel. The number of active sources follows the typical birth-death process with state space $\{i = 0, 1, \cdots, K\}$. The infinitesimal generator $\boldsymbol{M}_1$ for this traffic generation process is given by (35). In [25], Mitra provided a procedure for obtaining the eigenvalues and eigenvectors of the system in (35). Due to the decomposability of this model, as discussed in Section V, the analysis of the producer part in [25] is directly applicable. For the packet delivery part, we use the previous time-varying service rate model which is based on a two-state Gilbert-Elliott's channel model. Unlike the analysis of the packet delivery part in [25], which has a discrete range of service rates, our model has a continuous range that is the function of the BER of a wireless channel and the FEC capability.

By equating (4.7ii) in [25] and (14) in terms of zv, we obtain

$$\frac{K}{2}(rz + \alpha + \beta) + \left(k_1 - \frac{K}{2}\right)\sqrt{Q_1(z)} = \frac{z(c_g + c_b) - \delta - \gamma}{2} - \left(k_2 - \frac{1}{2}\right)\sqrt{Q_2(z)}. \quad (36)$$

Arranging the previous equation, one can define the following equation

$$f(z, k_1, k_2) \triangleq \left(k_1 - \frac{K}{2}\right)\sqrt{Q_1(z)} + \left(k_2 - \frac{1}{2}\right)\sqrt{Q_2(z)} + \frac{K}{2}(rz + \alpha + \beta) - \frac{z(c_g + c_b) - \delta - \gamma}{2} \quad (37)$$

In order to facilitate the calculation of the eigenvalues, we define $P(z, k_1, k_2)$. The equation $P(z, k_1, k_2) = 0$ is obtained from $f(z, k_1, k_2) = 0$ by rearranging and squaring terms to eliminate the square roots.

$$P(z, k_1, k_2) = \left(\left(k_1 - \frac{K}{2}\right)^2 Q_1(z) + \left(k_2 - \frac{n}{2}\right)^2 Q_2(z) - L^2(z)\right)^2 - 4\left(k_1 - \frac{K}{2}\right)^2\left(k_2 - \frac{n}{2}\right)^2 Q_1(z)Q_2(z) \quad (38)$$

where

$$L(z) \triangleq \frac{K}{2}(rz + \alpha + \beta) - \frac{z(c_g + c_b) - \delta - \gamma}{2}. \quad (39)$$

Because of the duplicate roots, we only need to consider $0 \le k_1 \le K/2$ and $0 \le k_2 \le 1$. The eigenvalues are the solutions of the polynomials $P(z, k_1, k_2) = 0, 0 \le k_1 \le K/2$ and $0 \le k_2 \le 1$, as proved in [25].

So far, we have discussed how to obtain the eigenvalues in (9) and (10) for the case of K on-off sources. The eigenvector $\boldsymbol{\phi}_1$ can be obtained by using Equation (5.7) in [25]. The eigenvectors $\boldsymbol{\phi}_2$ for the packet delivery part (10) are given in (18). Thus, the resulting eigenvectors $\boldsymbol{\phi}$ are obtained by the following equation:

$$\boldsymbol{\phi} = \boldsymbol{\phi}_1 \otimes \boldsymbol{\phi}_2.$$

Finally, the stationary queue length distribution $\boldsymbol{\Pi}(x)$ is given by:

$$\boldsymbol{\Pi}(x) = a_d \boldsymbol{w} + \sum_{z_i < 0} a_i \exp(z_i x)\boldsymbol{\phi}_i \quad (40)$$

where $\boldsymbol{w} = \boldsymbol{w}_1 \otimes \boldsymbol{w}_2$, $\boldsymbol{w}_1$ is given in (2.18) in [25], and $\boldsymbol{w}_2$ is given by

$$\boldsymbol{w}_2 = \begin{bmatrix} \frac{\gamma}{\delta+\gamma} & \frac{\delta}{\delta+\gamma} \end{bmatrix}.$$

In order to calculate the coefficients a_i, we use the boundary conditions:

$$\Pi_s(0) = 0, \quad \text{for } z_i < 0, \text{with } \chi_i > c_g \text{ or } \chi_i > c_b$$
$$a_d = 1, \quad (\text{since } \boldsymbol{\Pi}(\infty) = \boldsymbol{w})$$

where χ_i denotes the total source rate associated with the eigenvalue z_i. Finally, we obtain the packet loss rate from (21).

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their constructive comments.

$$\boldsymbol{M}_1 = \begin{bmatrix} -K\beta & K\beta & & & \\ \alpha & -(\alpha+(K-1)\beta) & (K-1)\beta & & \\ & \vdots & & \vdots & \\ & & & (K-1)\alpha & -((K-1)\alpha+\beta) \quad \beta \\ & & & & K\alpha \quad -K\alpha \end{bmatrix}. \tag{35}$$

References

[1] A. S. Acampora and M. Naghshineh, "An architecture and methodology for mobile–executed handoff in cellular ATM networks," *IEEE J. Select. Areas Commun.*, vol. 12, no. 8, pp. 1365–1375, Oct. 1994.

[2] A. Acampora, "Wireless ATM: a perspective on issues and prospects," *IEEE Pers. Commun.*, vol. 3, no. 4, pp. 8–17, Aug. 1996.

[3] D. Raychaudhuri and N. D. Wilson, "ATM-based transport architecture for multiservices wireless personal communication networks," *IEEE J. Select. Areas Commun.*, vol. 12, no. 8, pp. 1401–1414, Oct. 1994.

[4] J. B. Cain and D. N. McGregor, "A recommended error control architecture for ATM networks with wireless links," *IEEE J. Select. Areas Commun.*, vol. 15, no. 1, pp. 16–28, Jan. 1997.

[5] H. Balakrishnan, V. N. Padmanabhan, S. Seshan, and R. H. Katz, "A comparison of mechanisms for improving TCP performance over wireless links," *IEEE/ACM Trans. Networking*, vol. 5, no. 6, pp. 756–769, Dec. 1997.

[6] H. Liu and M. El Zarki, "Performance of H.263 video transmission over wireless channels using hybrid ARQ," *IEEE J. Select. Areas Commun.*, vol. 15, no. 9, pp. 1775–1786, Dec. 1997.

[7] E. Ayanoglu, K. Y. Eng, and M. J. Karol, "Wireless ATM: limits, challenges, and protocols," *IEEE Pers. Commun.*, vol. 3, no. 4, pp. 18–34, Aug. 1996.

[8] A. I. Elwalid and D. Mitra, "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks," *IEEE/ACM Trans. Networking*, vol. 1, no. 3, pp. 329–343, June 1993.

[9] R. Guérin, H. Ahmadi, and M. Naghshineh, "Equivalent capacity and its application to bandwidth allocation in high–speed networks," *IEEE J. Select. Areas Commun.*, vol. 9, no. 7, pp. 968–981, Sept. 1991.

[10] G. L. Choudhury, D. M. Lucatoni, and W. Whitt, "Squeezing the most out of ATM," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 203–217, Feb. 1996.

[11] R. J. Gibbens and P. J. Hunt, "Effective bandwidths for the multi-type UAS channel," *Queueing Syst.*, vol. 9, pp. 17–28, 1991.

[12] F. P. Kelly, "Effective bandwidths at multi-type queues," *Queueing Syst.*, vol. 5, pp. 5–15, 1991.

[13] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra, and A. Weiss, "Fundamental bounds and approximations for ATM multiplexers with applications to video teleconferencing," *IEEE J. Select. Areas Commun.*, vol. 13, no. 6, pp. 1004–1016, Aug. 1995.

[14] A. Mohammadi, S. KuMar, and D. Klynmyshyn, "Characterization of effective bandwidth as a metric of quality of service for wired and wireless ATM networks," in *ICC*. IEEE, 1997, vol. 2, pp. 1019–1024.

[15] M. Zukerman, P. L. Hiew, and M. Gitlits, "FEC code rate and bandwidth optimization in WATM networks," in *Multi-access, Mobility, and Teletraffic: Advances in Wireless Networks*, D Everitt and M. Rumsewicz, Eds., pp. 207–220. Kluwer, Boston, 1998.

[16] H. M. Chaskar, T. V. Lakshman, and U. Madhow, "TCP over wireless with link level error control: analysis and design methodology," *IEEE/ACM Trans. Networking*, vol. 7, no. 5, pp. 605–615, Oct. 1999.

[17] M. Rice and S. B. Wicker, "Adaptive error control for slowly varying channels," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 917–925, Feb./Mar./Apr. 1994.

[18] A. Parekh and R. Gallager, "A generalized processing sharing approach to flow control in integrated services networks: The single node case," *IEEE/ACM Trans. Networking*, vol. 1, no. 3, pp. 344–357, June 1993.

[19] H. M. Chaskar and U. Madhow, "Traffic multiplexing and service guarantees on a Rayleigh faded wireless downlink," in *ICUPC '98*, Oct. 1998, vol. 1, pp. 463–467.

[20] I. Joe, "An adaptive hybrid ARQ scheme with concatenated FEC codes for wireless ATM," in *MobiCom '97*, Sept. 1997, pp. 131–138.

[21] R. Fantacci, "Queueing analysis of the selective repeat automatic repeat request protocol wireless packet networks," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 258–264, May 1996.

[22] N. Guo and S. D. Morgera, "Frequency-hopped ARQ for wireless network data services," *IEEE J. Select. Areas Commun.*, vol. 12, no. 8, pp. 1324–1336, Sept. 1994.

[23] M. Zorzi, R. R. Rao, and L. B. Milstein, "Error statistics in data transmission over fading channels," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1468–1477, Nov. 1998.

[24] D. Anick, D. Mitra, and M. M. Sondhi, "Stochastic theory of a data–handling system with multiple sources," *Bell Syst. Tech. J.*, vol. 61, pp. 1871–1894, 1982.

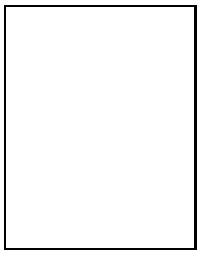
[25] D. Mitra, "Stochastic theory of a fluid model of producers and consumers coupled by a buffer," *Adv. Appl. Prob.*, vol. 20, pp. 646–676, 1988.

[26] S. Lin and Jr. D. J. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1983.

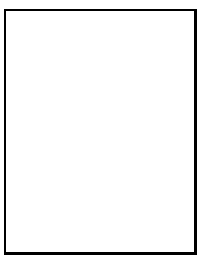
[27] A. I. Elwalid and D. Mitra, "Statistical multiplexing with loss priorities in rate-based congestion control of high-speed networks," *IEEE Trans. Commun.*, vol. 42, no. 11, pp. 2989–3002, Nov. 1994.

[28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P Flannery, *Numerical Recipes in C*, Cambridge University Press, 2 edition, 1992.

[29] T. Sato, M. Kawabe, T. Kato, and A. Fukasawa, "Throughput analysis method for hybrid ARQ schemes over burst error channels," *IEEE Trans. Veh. Technol.*, vol. 42, no. 1, pp. 110–117, Feb. 1993.



Jeong Geun Kim received his B.S. and M.S. degrees, both in electrical engineering, from Yonsei University, Seoul, Korea, in 1990 and 1992, respectively. He is currently working toward a Ph.D. degree at the University of Arizona. His research interests are in broadband wireless networks, quality of service, and performance evaluation of communication networks.



Marwan M. Krunz received the B.S. in Electrical Engineering from Jordan University, Jordan, in 1990, and the M.S. and Ph.D. degrees in Electrical Engineering from Michigan State University in 1992 and 1995, respectively. From 1995 to 1997, he was a postdoctoral research associate with the Department of Computer Science and the Institute for Advanced Computer Studies at the University of Maryland, College Park. In January 1997, he joined the Department of Electrical and Computer Engineering at the University of Arizona, where he is currently an

Assistant Professor. His research interests are in teletraffic modeling, resource allocation in wireless networks, and QoS-based routing.

Dr. Krunz received the National Science Foundation CAREER Award in 1998. He is a Technical Editor for the IEEE Communications Interactive Magazine, and has served and continue to serve on the executive and program committees of several technical conferences.