Effective Bandwidth in Wireless ATM Networks*

Jeong Geun Kim and Marwan Krunz
Department of Electrical and Computer Engineering
University of Arizona
Tucson, AZ 85721
{jkkim, krunz}@ece.arizona.edu

Abstract

Wireless ATM aims at extending ATM services to the wireless environment. In contrast to wireline ATM, which is primarily based on reliable fiber optic, wireless ATM will have to cope with an unreliable radio channel. This poses a host of technical challenges related to the provisioning of quality of service (QoS). Key issues include incorporating the characteristics of the wireless channel in the provisioning of cell-level QoS, and improving the perceived quality by using error control mechanisms. In this study, we propose a model for analyzing the cell loss statistics due to buffer overflow in a wireless ATM environment. This model incorporates the service disruption caused by the unreliable radio channel as well as the impact of error control schemes, e.g., automatic repeat request (ARQ) and forward error correction (FEC) mechanisms. The main theme of this study is to investigate the cell loss behavior of a wireless ATM link as a function of the assigned bandwidth and error control schemes. Using fluid analysis, we present an approximate expression for the cell loss rate (CLR). This expression is used to derive a closed-form expression for the wireless effective bandwidth. It is also used to investigate the optimal FEC code rate that guarantees a given CLR while maximizing the utilization of the wireless bandwidth. The validity of our analytical results are tested by contrasting them to simulation results. Our observations indicate that the proposed expression of CLR is quite accurate over a range of moderate bit error rates.

1 Introduction

Wireless asynchronous transfer mode (ATM) is emerging as a potential transport solution for broadband wireless networks [1, 2, 20]. It is being designed to be comparable with wireline ATM in terms of service capabilities, so that it will offer constant-bit-rate (CBR), variable-bit-rate (VBR), available-bit-rate (ABR), and unspecified-bit-rate (UBR) services, and will support a wide range of multimedia applications [20]. To offer such services, wireless ATM must be able to support the quality-of-service (QoS) requirements that are associated with various ATM services. Due to the scarcity of bandwidth and the high bit-error rates over the radio channel, the provisioning of QoS over a wireless medium is a challenging issue. In particular, the severe bandwidth and reliability mismatch between wireline and wireless links poses several technical issues.

ATM traffic control and resource allocation strategies are being designed with the assumption that the underlying physical media are highly reliable. Since this assumption does not apply to wireless links, applying the same resource allocation strategies of wireline ATM to the wireless environment could result in poor performance [6]. For instance, in transporting TCP traffic over wireless ATM, TCP makes the implicit assumption that lost packets are caused by congestion, although such losses might, in fact, be caused by cell discarding due to channel errors [5, 6] (errored cells are discarded by the ATM Adaptation Layer (AAL5), which sits under TCP). Consequently, TCP time outs and invokes its congestion control mechanism, which unnecessarily reduces the throughput of the wireless channel.

Other potential problems in wireless ATM are related to Header Error Control (HEC) and Cell Delineation (CD) process. ATM cells over radio links will be affected by burst errors, which will be uncorrectable by HEC. This will in turn cause cell discarding and poor cell delineation performance.

To address the above problems, the quality of a wireless link must be improved by using error control schemes. Two classes of error control are commonly used in wireless communications: automatic repeat request (ARQ) and forward error correction (FEC) [6]. In general, ARQ is used to deliver data requiring higher reliability, whereas FEC is suitable for delay-sensitive traffic [6]. Recent studies in [6] and [16] suggest that a hybrid ARQ/FEC might be more appropriate for wireless ATM. This argument stems from the fact that wireless ATM is designed to support various services with different QoS parameters. For instance, the ABR connections with relaxed time constrains can use ARQ. In contrast, CBR and VBR connections requiring low delay, jitter, and minimal cell loss will need a combination of FEC and ARQ with time-constrained retransmission [3].

In this paper, we consider a general framework for analyzing the cell-level performance over a wireless ATM link. A block diagram of this framework is shown in Fig. 1. In our framework, a traffic source (a mobile terminal or a basestation) generates a stream of ATM cells, which are first fed into a finite-size buffer. Each cell undergoes cyclic redundancy check (CRC) coding, followed by FEC coding. The cell is then transmitted over the wireless channel. At the receiver

* This research was supported in part by NSF CAREER Award ANI-9733143.

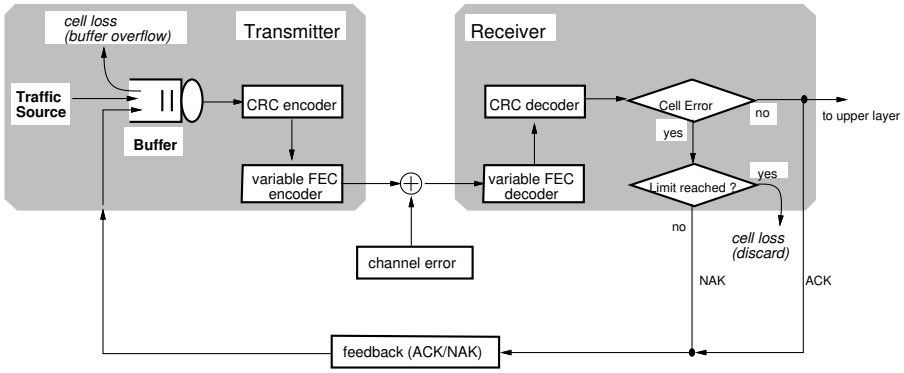

Figure 1: Framework for analyzing the performance of a wireless ATM link.

side, cells are passed through FEC and CRC decoders, in this order. If the decoded cell is correct, it is passed to the upper layer. Otherwise, the receiver will request retransmission unless the cell exceeds the limit on retransmission. The limit on the number of retransmissions is needed for delay-sensitive traffic, e.g., voice and video. Depending on the implementation, a cell that reaches the retransmission limit may be dropped or passed to the upper layer (typically, the the errored cell is dropped if its header is corrupted). Note that this model is a cell-level abstraction of the interactions between a transmitter and a receiver at data link level.

The above model has three control parameters: service rate (or assigned bandwidth), code rate of FEC, and limit on the number of retransmission. These parameters can be negotiated and tuned during the connection setup phase depending on QoS requirements, so that the network can deliver the promised QoS while maximizing the utilization of the wireless bandwidth. From the network point of view, the selection of these parameters is very crucial and requires thorough understanding of their impact on the cell-level performance. The main theme of this study is to investigate the cell loss behavior of a wireless ATM link as a function of the assigned bandwidth and error control schemes.

In order to guarantee the requested QoS while maintaining an acceptable level of bandwidth utilization, ATM employs the concept of effective bandwidth in connection admission control (CAC) and service scheduling [8, 12]. The effective bandwidth is *generally* defined as the minimum amount of bandwidth needed to provide a specific QoS given the traffic parameters of a connection and the buffer size at the multiplexer. ATM can provide a required level of QoS by guaranteeing the effective bandwidth associated with that connection. Significant research has been done on the notion of effective bandwidth over wireline networks [7, 8, 11, 12, 15]. Anick et al. provided a framework for analyzing buffer statistics under multiplexed on-off fluid sources [4]. Guérin et al. proposed an approximate expression for the effective bandwidth of both individual and multiplexed connections, arguing that this approximation is necessary in real-time network traffic control applications [12]. Elwalid and Mitra presented a way to obtain the effective bandwidth for general Markovian traffic sources [8]. Elwalid et al. proposed approximation to the cell loss rate for multiplexed Markovian sources using a hybrid Chernoff Bound/Dominant Eigenvalue approache [9]. The research on effective bandwidth has generally been addressed in the context of high-speed (wired) ATM networks. Recently, Mohammadi et al. extended the concept of effective bandwidth to wireless ATM networks [18]. However, only the cell losses due to channel error were considered, ignoring the role of error control schemes. Zukerman et al. addressed the problem of bandwidth and FEC code optimization in wireless ATM networks [22]. Although they identified some of the problems that we address in this paper, their simplified model might be inappropriate to investigate the QoS issues in wireless ATM networks.

As shown in Fig. 1, cell losses can be caused by buffer overflow at the transmitter side or cell discarding at the receiver side. In this work, we will focus on the cell loss due to buffer overflow, which implies no limit is imposed on the number of retransmissions. Cell losses due to buffer overflow are aggravated by the retransmission process, consequently leading to a reduction in the actual service rate. Therefore, extra bandwidth must be assigned in order to compensate for the reduced service rate. The extra bandwidth can be assigned by increasing the service rate, redundancy in FEC, or both. Increasing the service rate may reduce the CLR although the quality of the wireless channel stays unchanged. In contrast, increasing the FEC code rate will improve the quality of wireless channel at the expense of extra bandwidth, which may in turn reduce the effective service rate. Thus, controlling the code rate of FEC has a subtle effect on CLR performance. In general, adaptive coding strategy yields the optimal throughput for a channel with variable error statistics [21]. In this study, we address the problem of finding an optimal bandwidth and code rate which can satisfy QoS parameters specified in terms of cell loss while maximizing the utilization of bandwidth. We call this bandwidth with code rate *wireless effective bandwidth*, which is equivalent, in concept, to the effective bandwidth in wireline ATM networks.

Quantification of wireless effective bandwidth requires an analytic expression for the CLR. In our works, each traffic source is modeled as an on-off fluid process, with on and off periods being exponentially distributed. We use Gilbert-Elliott's channel model to capture the correlated behavior of channel errors [10, 13]. Using fluid-flow analysis, we provide an approximate value for the wireless effective bandwidth with code rate. We evaluate the goodness of this approximation using simulations.

The rest of the paper is organized as follows. In Section 2, we present the analytical model for wireless effective bandwidth. An approximate solution for wireless effective bandwidth is discussed in Section 3. Numerical results are reported in Section 4, followed by concluding remarks in Section 5.

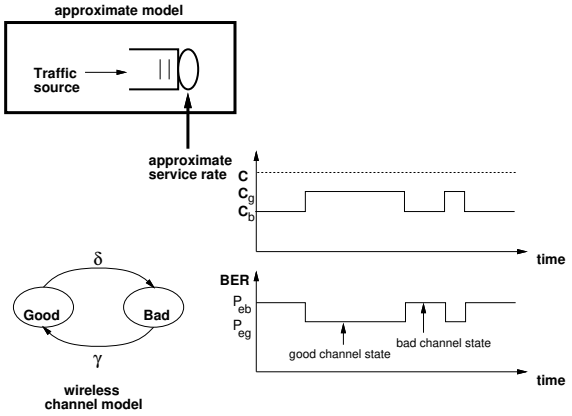

Figure 2: Wireless channel model and approximate service rate model.

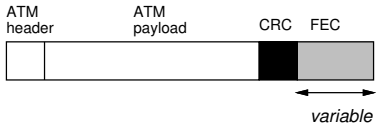

Figure 3: Cell format in a wireless ATM with hybrid ARQ/FEC error control.

2 Analytical Model

In this section, we present the analytical model for a single ATM connection transported over a wireless link. This model will be used to derive the cell loss performance of a wireless ATM link and the corresponding wireless effective bandwidth. In this model, the traffic source is characterized by an on-off fluid model, with peak rate r and with exponentially distributed on and off periods with means $1/\alpha$ and $1/\beta$, respectively. The wireless channel is modeled using Gilbert-Elliott's model, which is a two-state Markov model [10, 13] (see Fig. 2). Gilbert-Elliott's model is commonly used to investigate the performance of wireless links with correlated channel errors. According to this model, the wireless channel is modeled by two alternating states: *Good* and *Bad*. The bit error rates (BER) during the Good and Bad states are given by P_{eg} and P_{eb}, respectively, where $P_{eg} \ll P_{eb}$. The durations of the Good and Bad states are exponentially distributed with means $1/\delta$ and $1/\gamma$, respectively.

We assume that a hybrid ARQ/FEC mechanism is used for error control. Different hybrid ARQ/FEC mechanisms have been proposed in the wireless literature [6, 14], which cannot all be captured in a single model. In the underlying study, we assume a hybrid ARQ/FEC mechanism that results in the cell format shown in Fig. 3. We focus on the optimal use of bandwidth given QoS parameters, in particular, cell loss rate. Thus, we ignore all overheads related to the medium access control (MAC) protocol, including synchronization bits and CRC code. We assume that CRC can detect all error patterns. Consequently, a FEC code C is characterized by $C(n, k, \tau)$, where n is the size of a code block in bits, k is the number of parity bits, and τ is the number of correctable bits. The code rate $e(\tau)$ of a code C is defined as

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

Assuming that a code can correct up to τ bits and that bit errors in a given channel state are independent, the probability that a cell contains a non-correctable error is given by:

$$P_c(p_b, \tau) = \sum_{j=\tau+1}^{n(\tau)} \binom{n(\tau)}{j} p_b^j (1-p_b)^{n(\tau)-j} \quad (1)$$

where p_b is the bit error probability. To account for the FEC overhead, we obtain the actual service rate c_e observed by a buffer at the data link layer:

$$c_e = c \cdot e(\tau) \quad (2)$$

where c is the bandwidth assigned to the ATM connection.

Incorporating the exact dynamics of ARQ and FEC in the underlying model leads to intractable results. In particular, the exact ARQ dynamics results in a queuing system with a complicated service process (for a stop-and-wait ARQ, the service process is a Markov-modulated Bernoulli process). Therefore, some approximations are needed. To obtain analytically tractable results, we assume that the cell departure process follows a fluid process with a service rate that is modulated by the channel state (see Fig. 2). Thus, this analytic model involves two deterministic service rates: c_g during Good states and c_b during Bad states. In addition, we ignore the feedback delay associated with sending an acknowledgment from the receiver (access node) to the sender (mobile terminal). It is also assumed that the mobile terminal attempts to retransmit a cell until this cell is successfully received (i.e., number of bit errors is less than τ). In a realistic stop-and-wait ARQ scenario, the total time needed to successfully transmit a cell 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 $1 - P_{c,b}$, respectively, where $P_{c,g}$ and $P_{c,b}$ are the cell error probabilities in Good and Bad states, respectively, given by (1):

$$\begin{aligned} c_g &= c \cdot e(\tau) \cdot (1 - P_{c,g}) \\ &= c \cdot e(\tau) \cdot (1 - P_c(P_{eg}, \tau)) \quad (3) \\ c_b &= c \cdot e(\tau) \cdot (1 - P_{c,b}) \\ &= c \cdot e(\tau) \cdot (1 - P_c(P_{eb}, \tau)). \quad (4) \end{aligned}$$

It can be seen that the problem of determining the wireless effective bandwidth reduces to obtaining a service rate c that satisfies the required CLR. The system is Markovian with a state transition diagram that is shown in Fig. 4. Let S denote the state space. Thus,

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

where 0 and 1 denote the on and off states of the traffic source, respectively, and g and b denote the Good and Bad channel states, respectively.

Following a standard fluid approach (see [4], for example), the evolution of the buffer content can be described by the following differential equation:

$$\frac{d\boldsymbol{\Pi}(x)}{dx}\boldsymbol{D} = \boldsymbol{\Pi}(x)\boldsymbol{M} \quad (5)$$

where $\boldsymbol{D} \triangleq \text{diag}[-c_g, -c_b, r-c_g, r-c_b]$, $\boldsymbol{\Pi}(x) \triangleq [\ \Pi_{0,g}(x) \quad \Pi_{0,b}(x) \quad \Pi$ $\Pi_s(x) \triangleq \Pr\{\text{buffer content} \leq x, \text{and the system } s \in S\}$, and

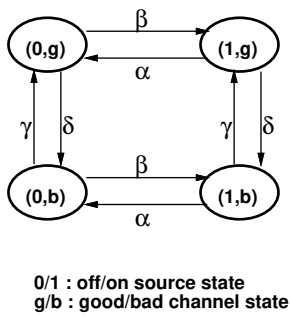

Figure 4: State transition diagram.

$\boldsymbol{M}$ 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 to (5) corresponds to the solution of the eigenvalue/eigenvector problem:

$$z\boldsymbol{\phi}\boldsymbol{D} = \boldsymbol{\phi}\boldsymbol{M} \tag{6}$$

which is generally given by

$$\boldsymbol{\Pi}(x) = \sum_{z_i \le 0} a_i \exp(z_i x)\phi_i \tag{7}$$

where a_i's are constant coefficients and the pairs $(z_i, \boldsymbol{\phi}_i)$, $i = 1, 2, \cdots$, are the eigenvalues and the right eigenvectors of the matrix $\boldsymbol{M}\boldsymbol{D}^{-1}$ [4, 17]. Let $\boldsymbol{w}$ denote the stationary probability distribution 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:

$$\boldsymbol{w} = \frac{1}{(\alpha+\beta)(\delta+\gamma)} \begin{bmatrix} \alpha\gamma & \alpha\delta & \beta\gamma & \beta\delta \end{bmatrix}. \tag{8}$$

The mean drift is given by:

$$\boldsymbol{w}\boldsymbol{D}\boldsymbol{1} = \frac{\beta}{\alpha+\beta} r - \frac{c_g\gamma + c_b\delta}{\delta+\gamma}. \tag{9}$$

For a stable system, $\boldsymbol{w}\boldsymbol{D}\boldsymbol{1} < 0$. In this paper, we consider a stable system with an infinite buffer.

In order to solve (6), we follow the general approach in [17]. Using this approach, 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

$$\boldsymbol{M}_1 = \begin{bmatrix} -\beta & \beta \\ \alpha & -\alpha \end{bmatrix} \text{ and } \boldsymbol{M}_2 = \begin{bmatrix} -\delta & \delta \\ \gamma & -\gamma \end{bmatrix} \tag{10}$$

with

$$\boldsymbol{M} = \boldsymbol{M}_1 \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{M}_2 \tag{11}$$

and $\boldsymbol{A} \otimes \boldsymbol{B}$ is the Kronecker product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. For the drift matrix $\boldsymbol{D}$, we have

$$\boldsymbol{D} = r\boldsymbol{E}_r \otimes \boldsymbol{I} - \boldsymbol{I} \otimes \boldsymbol{E}_c \tag{12}$$

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

$$\boldsymbol{E}_r = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \boldsymbol{E}_c = \begin{bmatrix} c_g & 0 \\ 0 & c_b \end{bmatrix}. \tag{13}$$

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

$$\boldsymbol{\phi} = \boldsymbol{\phi}_1 \otimes \boldsymbol{\phi}_2 \tag{14}$$

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 (12), (11), (14) into (6), we end up with

$$\begin{aligned} &\boldsymbol{\phi}_1 \otimes (\boldsymbol{\phi}_2\boldsymbol{M}_2 + z\boldsymbol{\phi}_2\boldsymbol{E}_c) \\ &\quad = (zr\boldsymbol{\phi}_1\boldsymbol{E}_r - \boldsymbol{\phi}_1\boldsymbol{M}_1) \otimes \boldsymbol{\phi}_2. \end{aligned} \tag{15}$$

Hence, for (6) and (14) to be valid, it is sufficient that there exists some real number v such that both sides of the above equation equal $zv\boldsymbol{\phi}_1 \otimes \boldsymbol{\phi}_2$:

$$z\boldsymbol{\phi}_1[r\boldsymbol{E}_r - v\boldsymbol{I}] = \boldsymbol{\phi}_1\boldsymbol{M}_1 \tag{16}$$

$$z\boldsymbol{\phi}_2[v\boldsymbol{I} - \boldsymbol{E}_c] = \boldsymbol{\phi}_2\boldsymbol{M}_2 \tag{17}$$

which implies that

$$\begin{aligned} \det[z\boldsymbol{I} - \boldsymbol{M}_1[r\boldsymbol{E}_r - v\boldsymbol{I}]^{-1}] &= 0 \\ \det[z\boldsymbol{I} - \boldsymbol{M}_2[v\boldsymbol{I} - \boldsymbol{E}_c]^{-1}] &= 0 \end{aligned} \tag{18}$$

where $\det[\boldsymbol{A}]$ is the determinant of $\boldsymbol{A}$. The first equation in (18) reduces to the quadratic polynomial

$$(zv)^2 - (zr + \alpha + \beta)zv + \beta zr = 0. \tag{19}$$

Thus,

$$zv = \frac{zr + \alpha + \beta}{2} \pm \frac{\sqrt{Q_1(z)}}{2} \tag{20}$$

where $Q_1(z) = (zr + \alpha - \beta)^2 + 4\alpha\beta$. For the second equation in (18), we have

$$\begin{aligned} &(zv)^2 - (z(c_g + c_b) - \delta - \gamma)zv \\ &\quad + z(zc_gc_b - \gamma c_g - \delta c_b) = 0. \end{aligned} \tag{21}$$

Thus,

$$zv = \frac{z(c_g + c_b) - \delta - \gamma}{2} \pm \frac{\sqrt{Q_2(z)}}{2} \tag{22}$$

where $Q_2(z) = (z(c_g - c_b) - (\delta - \gamma))^2 + 4\delta\gamma$.

Equating (20) and (22) 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 to eliminate the square roots:

$$((z(r-c_g-c_b)+\alpha+\beta+\delta+\gamma)^2 - Q_1(z)-Q_2(z))^2 - 4Q_1(z)Q_2(z) = 0. \quad (23)$$

Rearranging (23), we obtain a polynomial equation of order three, whose closed-form solution is available in the literature [19]. For brevity, we omit the algebraically tedious, but conceptually straightforward, details.

After obtaining all eigenvalues from (23), we obtain v using (20) and (22). Let v_1, v_2, v_3, v_4 denote the corresponding v values. Once z and v are determined, the eigenvectors can be obtained from (16) and (17):

$$\boldsymbol{\phi}_1 = [\ \alpha \quad -zv+\beta\ ] \quad (24)$$

$$\boldsymbol{\phi}_2 = [\ \gamma \quad zv-zc_g+\delta\ ]. \quad (25)$$

Thus,

$$\boldsymbol{\phi}_1 \otimes \boldsymbol{\phi}_2 = \begin{bmatrix} \alpha\gamma \\ \alpha(zv-zc_g+\delta) \\ (-zv+\beta)\gamma \\ (-zv+\beta)(zv-zc_g+\delta) \end{bmatrix}^T. \quad (26)$$

Finally, the stationary buffer content distribution $\boldsymbol{\Pi}(x)$ is given by

$$\boldsymbol{\Pi}(x) = a_d \boldsymbol{w} + \sum_{z_i<0} a_i \exp(z_i x)\boldsymbol{\phi}_i \quad (27)$$

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

For an infinite buffer system, we have the following boundary conditions [17]:

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

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

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -\begin{bmatrix} \phi_{1g}(z_1) & \phi_{1g}(z_2) \\ \phi_{1b}(z_1) & \phi_{1b}(z_2) \end{bmatrix}^{-1} \begin{bmatrix} w_{1g} \\ w_{1b} \end{bmatrix} \quad (28)$$

where $\phi_s(z)$ is the eigenvector element of state s by (26).

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

$$a_1 = -w_{1b}/\phi_{1b}(z_1). \quad (29)$$

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

$$G(x) = 1 - \mathbf{1}\boldsymbol{\Pi}(x). \quad (30)$$

In Section 4, we provide some numerical results related to the CLR. In addition, we compare the above analytic result with simulation results to validate the correctness of our analysis.

3 Wireless Effective Bandwidth

The analytic expression for the CLR that was 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 (CLR, buffer size, channel BERs, etc.). In general, it is not possible to obtain a closed-form expression for the effective bandwidth, which would be needed for real-time traffic control [12]. Even for a single source (i.e., no multiplexing), the closed-form solution is not available without approximation [7, 12].

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

Consider (6). In this case, the drift matrix $\boldsymbol{D}$ can be written as

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

where

$$\boldsymbol{B}_r = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and} \quad (31)$$

$$\boldsymbol{B}_c = \begin{bmatrix} \eta_g & 0 & 0 & 0 \\ 0 & \eta_b & 0 & 0 \\ 0 & 0 & \eta_g & 0 \\ 0 & 0 & 0 & \eta_b \end{bmatrix} \quad (32)$$

$\eta_g = 1 - P_c(P_{eg}, \tau)$, and $\eta_b = 1 - P_c(P_{eb}, \tau)$. Substituting $\boldsymbol{D}$ into (6), we obtain the following relation:

$$z\boldsymbol{\phi}(r\boldsymbol{B}_r - ce(\tau)\boldsymbol{B}_c) = \boldsymbol{\phi}\boldsymbol{M}. \quad (33)$$

Let $c = g(z)$. Then,

$$zr\boldsymbol{\phi}\boldsymbol{B}_r - zg(z)e(\tau)\boldsymbol{\phi}\boldsymbol{B}_c = \boldsymbol{\phi}\boldsymbol{M}. \quad (34)$$

Rearranging the previous equation, we obtain

$$g(z)e(\tau)\boldsymbol{\phi} = \boldsymbol{\phi}\left(-\frac{1}{z}\boldsymbol{M} + r\boldsymbol{B}_r\right)\boldsymbol{B}_c^{-1}. \quad (35)$$

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 (35).

From (35) 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

$$\begin{aligned} C_1 &= (\eta_g+\eta_b)r - ((\alpha+\beta+2\gamma)\eta_g + (\alpha+\beta+2\delta)\eta_b)\xi \\ C_2 &= \sqrt{(r-(\alpha-\beta)\xi)^2 + 4\alpha\beta\xi^2} \\ C_3 &= \eta_b^2((r-(\alpha+\delta)\xi)^2 + ((\beta+\delta)^2+2\alpha\beta)\xi^2) \\ &\quad -2\eta_g\eta_b(r^2-(2\alpha+\delta+\gamma)r\xi \end{aligned}$$

$$
\begin{aligned}
& +((\alpha+\beta)(\alpha+\beta+\delta+\gamma)-2\delta\gamma)\xi^2) \\
& +\eta_g^2((r-(\alpha+\gamma)\xi)^2+((\beta+\gamma)^2+2\alpha\beta)\xi^2) \\
C_4 &= (\eta_g-\eta_b)(\eta_g(r-(\alpha+\beta+2\gamma)\xi) \\
& -\eta_b(r-(\alpha+\beta+2\delta)\xi)) \\
\xi &= -B/\log p,\ B \text{ is buffer size and } p \text{ is cell loss rate.}
\end{aligned}
$$

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

$$
g(z) = \frac{C_1+(\eta_g+\eta_b)C_2+\sqrt{2(C_3+C_2C_4)}}{4\eta_g\eta_b e(\tau)}. \quad (36)
$$

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

4 Numerical Results

In this section, we provide numerical results for the CLR due to buffer overflow at a mobile terminal and for the approximate value of wireless effective bandwidth. In order to validate the accuracy of our analytical results, we compare them to those obtained by simulations.

In our simulations, an on-off traffic source is assumed in which the on periods are exponentially distributed with mean 0.02304 second and the off periods are exponentially distributed with mean 0.2304 second. The ARQ retransmission process is simulated in a realistic manner, in contrast to our analysis in which the ARQ retransmission process is approximated by a fluid process with deterministic service rates that are modulated by the channel state.

In the simulation study, we vary the buffer size and the BER during the Bad state (P_{eb}), and we fix the BER during the Good state at 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 [13]. The parameters used in both the simulations and the analytic model are summarized in Table 1. All the plots for the simulation results are obtained with 95% confidence intervals. Our simulations are based on the following assumptions:

- Conditioned on the channel state, bit errors are independent.
- Transitions between Good and Bad states do not occur in the middle of a transmitted cell.
- A cell is retransmitted if one or more bits are in error (i.e., 100% error detection).
- Propagation delay is not considered.

4.1 Case 1: ARQ

In this section, we consider the performance when only ARQ is used for error control (i.e., no FEC). Fig. 5 shows the CLR as a function of the buffer size for different values of P_{eb} and for a service rate of $c = 400$ cells/sec. As expected, the CLR increases as the buffer size decreases and as P_{eb} increases. In Fig. 5, it is observed that the analytic results slightly overestimate the actual CLR. In particular, the case of $P_{eb} = 10^{-2}$ shows higher deviation than the other cases, which can be explained by examining the variance of the BER. The mean, variance, and the squared coefficient of variation of the cell service time (T) in Bad state (which is governed by a geometric process) are given by:

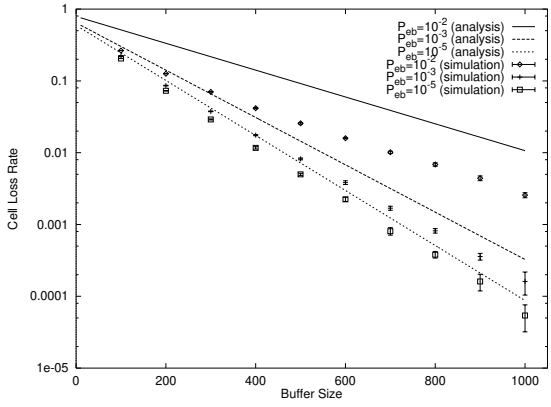

Figure 5: Cell loss rate versus buffer size ($c = 400$ cells/sec).

$$
\begin{aligned}
E[T] &= \frac{1}{c(1-P_{c,b})} = \frac{1}{c((1-P_{eb})^{53\times 8})} \\
Var[T] &= \frac{P_{c,b}}{c^2(1-P_{c,b})^2} = \frac{1-(1-P_{eb})^{53\times 8}}{c^2((1-P_{eb})^{53\times 8})^2} \\
c_T^2 &= \frac{Var[T]}{E^2[T]}.
\end{aligned}
$$

P_{eb}	$P_{c,b}$	$E[T]$	$Var[T]$	c_T^2
10^{-2}	0.9789	0.1188	0.0137	0.97
10^{-3}	0.3189	0.0036	4.29×10^{-6}	0.33
10^{-4}	0.0376	0.00259	2.54×10^{-7}	0.038
10^{-5}	0.0038	0.0025	2.41×10^{-8}	0.0039

Table 2: Mean, variance, the squared coefficient of variation of actual cell service time (T) for Bad state ($c = 400$ cells/sec).

The values for $E[T]$, $Var[T]$, and c_T^2 at $c = 400$ are listed in Table 2. As P_{eb} decreases, the variance decreases significantly. This reduced variance leads to a more deterministic cell departure process, thereby justifying the goodness of the deterministic approximation over the range of P_{eb} from 10^{-3} to 10^{-5}. Furthermore, c_T^2 drops rapidly and approaches zero as P_{eb} goes to zero. Since $c_T^2 = 0$ for the deterministic service time, this also justifies our deterministic (fluid) approximation of the average service time.

Fig. 6 and 7 depict the CLR for $c = 500$ and 550 cells/sec, respectively (BER in Bad state $= 10^{-2}, 10^{-3}$, and 10^{-5}). In Fig. 6 it is observed that for $P_{eb} = 10^{-3}$ and $P_{eb} = 10^{-5}$, the simulation and analytical results are quite close to each other. In addition, the case of $P_{eb} = 10^{-2}$ shows less difference when compared with the previous plot ($c = 400$).

Fig. 8, 9, and 10 depict the wireless effective bandwidth computed based on the results in Section 3. The curves of wireless effective bandwidth are plotted for $P_{eb} = 10^{-3}$ and $P_{eb} = 10^{-5}$, along with the results obtained from CLR based on (27). It is observed from these figures that the approximation for the wireless effective bandwidth based on one (dominant) eigenvalue overestimates the actual bandwidth requirement as predicted by (27) (which uses all eigenvalues for both BER cases). It is well known that the effective bandwidth approximation based on one eigenvalue

Parameter	Symbol	Value
peak rate	r	1 Mbps (or 2604.1667 cells/sec)
buffer size	B	$100 - 1000$ cells
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}	$10^{-2} - 10^{-5}$

Table 1: Parameters used in numerical results.

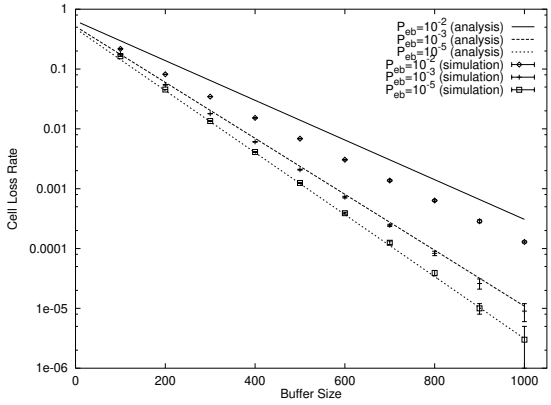

Figure 6: Cell loss rate versus buffer size ($c = 500$ cells/sec).

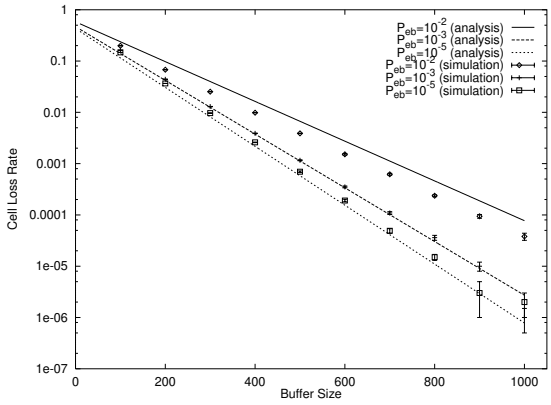

Figure 7: Cell loss rate versus buffer size ($c = 550$ cells/sec).

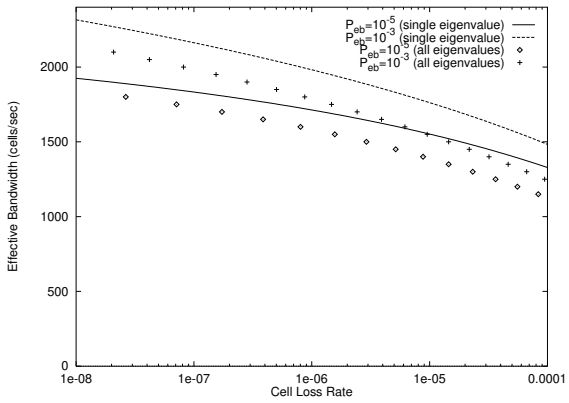

Figure 8: Effective bandwidth versus CLR ($B = 300$).

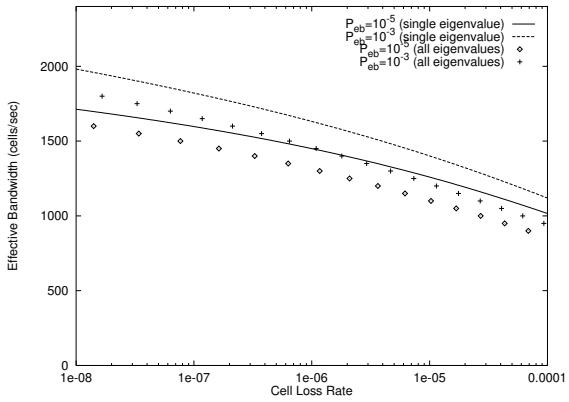

Figure 9: Effective bandwidth versus CLR ($B = 400$).

could overestimate the required bandwidth for the target CLR [7].

When $P_{eb} = 10^{-5}$ and the buffer size is 300 cells, the wireless effective bandwidth is between 1300 and 1900 cells/sec for a CLR ranging from 10^{-4} to 10^{-8}. The mean and peak rate of a traffic source is 263.74 cells/sec and 2604.1667 cells/sec, respectively. For low BERs, the wireless effective bandwidth expectedly lies between mean and peak rates (see Table 1). Interestingly, it is observed that the wireless effective bandwidth in the case of higher bit error rate is closer to the peak rate than to the mean rate. When $P_{eb} = 10^{-3}$ and the buffer size is 300 cells, the wireless effective bandwidth falls between 1500 and 2300 (which is much higher than the mean rate) for a CLR in the range 10^{-4} to 10^{-8}. Thus, we can see that the severe channel condition significantly affects wireless cell loss performance. Also, it may

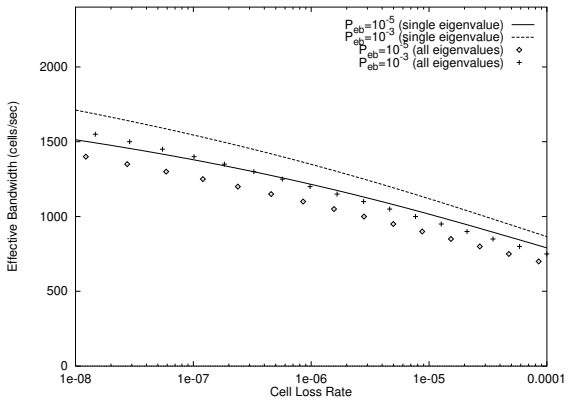

Figure 10: Effective bandwidth versus CLR ($B = 500$).

be difficult to achieve reasonable quality of service without allocating additional bandwidth capacity.

4.2 Case 2: ARQ and FEC

In this section, we investigate the performance when hybrid ARQ/FEC is used for error control. For FEC we adopt Bose-Chaudhuri-Hocquenghem (BCH) code. Since we treat both the ATM header and the payload as information bits, the total bits add up to 424 bits. Table 3 shows the size, the code rate, and the number of correctable bits using BCH code.

(n, k)	τ	$e(\tau)$
$(442, 424)$	2	0.959
$(451, 424)$	3	0.940
$(460, 424)$	4	0.921
$(469, 424)$	5	0.904
$(478, 424)$	6	0.887

Table 3: The size of BCH code versus the number of correctable bits (τ).

Fig. 11, 12, and 13 show the plots of wireless effective bandwidth versus the number of correctable bits over the target CLR of $10^{-4}, 10^{-5}, 10^{-6}$. These figures show that more bandwidth is required as the target CLR increases. In particular, each curve clearly indicates that there exists an optimal code rate for a given BER. For instance, to provide a target CLR 10^{-6} at $P_{eb} = 0.001$, we can pick up 1262.4 (cells/sec) as a service rate c and 1 as the number of correctable bits (τ). Since the mean number of bit error for $P_{eb} = 0.001$ is 0.423, the excessive capability of correction (τ), i.e., more than 1, leads to waste of bandwidth. In addition, it is observed that the optimal code rate increases as P_{eb} increases. This is a strong evidence that there exists optimal code rate for each BER and thus adaptive coding strategy is required for a wireless channel with time-varying characteristics. Unlike Fig. 11, the next two figures show that an excessive amount of bandwidth is required when only ARQ is used. This result indicates that the proper use of FEC is necessary for maximizing the utilization of the wireless bandwidth.

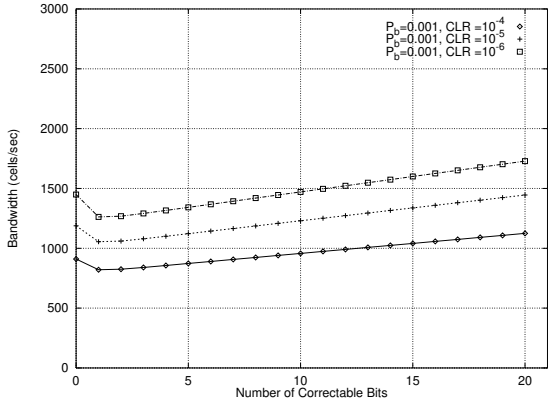

Figure 11: Effective bandwidth versus τ ($P_{eb} = 0.001$).

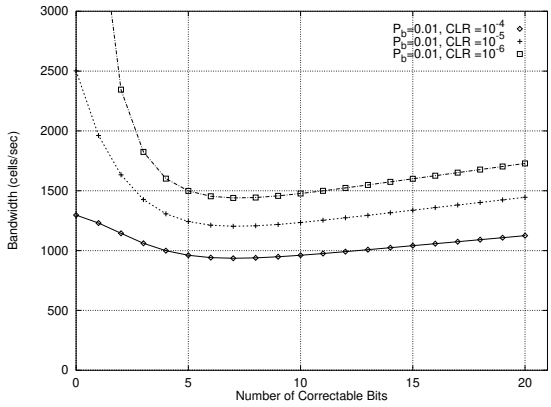

Figure 12: Effective bandwidth versus τ ($P_{eb} = 0.01$).

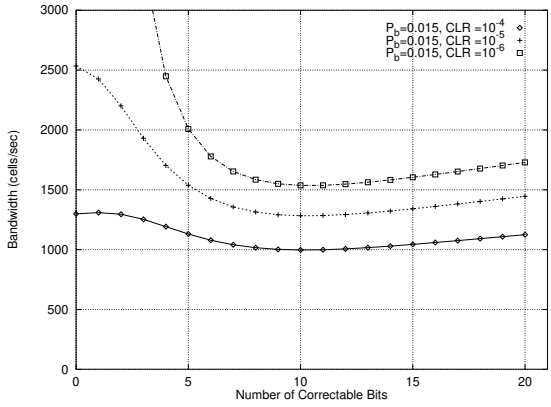

Figure 13: Effective bandwidth versus τ ($P_{eb} = 0.015$).

5 Conclusions

In this study, we investigated the cell loss behavior of a wireless ATM link. For a single connection, we derived the exact expression for CLR. This expression was then used to derive an approximate expression for the wireless effective bandwidth. We validated the accuracy of our analytic results by contrasting them with those obtained by simulations.

Our results indicate that the proposed expression of CLR in the case of a single connection is quite accurate over a range of moderate bit error rate. We also found that in the case of severe BER, wireless effective bandwidth is much greater than the average source rate, and is closer to the peak rate. In particular, we provided the expression which can be used to find the effective bandwidth and optimal code rate. We expect that wireless effective bandwidth can be applied to CAC process in wireless ATM and can also be used as a basis for service scheduling.

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] 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, 1996.

[4] 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.

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

[6] 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.

[7] 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.

[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, Jun. 1993.

[9] 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.

[10] 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.

[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] 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, Sep. 1991.

[13] 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, Sep. 1994.

[14] I. Joe and I. F. Akyildiz, "An adaptive hybrid ARQ scheme with concatenated FEC codes for wireless ATM," *Proc. MobiCom '97*, pp. 131–138, Sep. 1997.

[15] F. P. Kelly, "Effective bandwidths at multi-type queues," *Queueing Syst.*, vol. 9, pp. 5–15, 1991.

[16] 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.

[17] 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.

[18] 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 *Proc. IEEE ICC*, vol. 2, pp. 1019–1024, 1997.

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

[20] 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.

[21] 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.

[22] M. Zukerman, P. L. Hiew, and M. Gitlits, "FEC code rate and bandwidth optimization in WATM networks," *Multiaccess, Mobility, and Teletraffic: Advances in Wireless Networks*, Edited by D. Everitt and M. Rumsewicz, pp. 207–220, Kluwer Academic Publishers, Boston, 1998.