Asymptotic Analyses of a Statistical Multiplexor with Heterogeneous ATM Sources

Hyong Woo Lee and Jon Wei Mark

Abstract

Two asymptotic analyses of the queue length distribution at a statistical multiplexor supporting heterogeneous exponential on-off sources are considered. The first analysis is performed by approximating the cell generation rate as a multi-dimensional Ornstein-Uhlenbeck process and then applying the Benes queueing formula. In the second analysis, we start with a system of linear equations derived from the exact expressions of the dominant eigenvalue of the matrix governing the queue length distribution. Assuming that there are a large number of sources, we obtain asymptotic approximations to the dominant eigenvalue. Based on the analyses, we define a traffic descriptor to include the mean and the variance of the cell generation rate and a burstiness measure. A simple expression for the quality of service (QoS) in cell loss rate is derived in terms of the traffic descriptor parameters and the multiplexor parameters (output link capacity and buffer size). This result is then used to quantify the factors determining the required capacity of a call taking the statistical multiplexing gain into consideration. As an application of the analyses, we can use the required capacity calculation for simple yet effective connection admission control (CAC) algorithms.

I. Introduction

The Asynchronous Transfer Mode(ATM) has been recognized as a promising transfer mode of realizing the broadband ISDN. An ATM network is a mesh connection of switches and multiplexors in which switching and multiplexing are cell-oriented. Moreover, ATM offers connection-oriented service through the establishment of virtual channels and virtual paths. One physical link can support many virtual channels. It follows that an output port of an output- buffered switch must handle multiplexed transmission of a number of virtual flows. Thus, statistical multiplexing takes place at interior nodes as well as at user-network interfaces.

It is expected that an ATM network will carry a variety of traffic types including voice, data and video. Unlike data traffic in a conventional computer network, interactive voice and video calls in a packet switched environment generate cells with a high degree of time correlation [1-4]. In order to provide the quality of service (QoS) assurance to the calls having various cell generation processes, we need an analytical model for predicting the queue length distribution. Traditional queueing analysis using a Poisson or batch Poisson process to approximate the cell arrivals fails to capture the effect of the time correlation of the cell generation [5].

In this paper we assume that the cell generation process of a call can be represented by a single exponential on-off source [1], by a number of identical exponential on-off sources [3] or by two or more types of exponential on-off sources [4, 11]. Observing that the cell transmission time is often very small compared to the length of an on period, many researchers have successfully used the fluid flow approximation to predict the queue length distributions of a statistical multiplexor [4, 6, 7]. Although the fluid flow approximation simplifies the analysis, the exact analysis becomes computationally intractable, as the number of types of on-off sources increases and the number of on-off sources in each type increases.

As the ATM link capacity increases, it is likely that the number of source types and the number of source of each type increase. In this paper, we obtain simple asymptotic approximations for the queue length distribution assuming that the number of on-off sources in each type is large. We show that this approximation gives an asymptotic upper bound on the exact queue length distribution and can be used to conservatively estimate the cell loss rate. As a by-product of the analyses, we obtain a simple relationship among QoS in cell loss rate, the traffic descriptor (mean, variance of cell generation rate and a measure of burstiness) of a source and the required capacity, the portion of capacity used by the source.

The paper is organized as follows. In Section 2, we review

Manuscript received March 24, 1997; accepted May 23, 1997.

H. W. Lee is with Department of Electronics and Information Engineering Korea Univeristy.

J. W. Mark is with Department of Electrical and Computer Engineering University of Waterloo, Canada.

known results on the queue length distribution analysis using the fluid flow model. We then provide further motivation and assumptions used throughout this paper. In Sections 3 and 4, we present two asymptotic analyses of the queue length distribution. We then discuss the relationship among the traffic descriptor, the required channel capacity and the cell loss rate performance in Section 5. In Section 6, we comment on a possible method of obtaining the traffic descriptor parameters using an estimate of the autocorrleation function of the cell generation rate. The paper is concluded in Section 7.

II. Known Results and Assumptions

Under the assumption that 1) a cell can be described as an on-off source cells are generated at a constant rate during an on period and no cells are generated during an off period, 2) the length of on and off periods are exponentially distributed and 3) all the calls are independent and statistically identical (referred to as homogeneous), Anick et al. [8] have provided a simple algorithm for determining the queue length distribution of a statistical multiplexor. Kosten [9] has reported efforts to extend the analysis of Anick et al. for non-homogeneous sources (the sources may have non-identical mean on and off periods and cell generation rates) and the sources with non-exponential on periods. He offers an approximate solution for the cell loss rate as $\beta_0 e^{-z_0 B}$ where B is the multiplexor buffer size and z_0 is the dominant eigenvalue (largest negative eigenvalue) of the matrix governing a system of linear equations. He has proposed to obtain the coefficient, β_0, using simulations. Stern and Elwalid [10] have discussed the solutions to the multiplexor queueing analysis where the source generate cells at the rate modulated by a Markovian source state. Assuming that the modulating process is reversible, they have provided a computationally efficient algorithm for obtaining the eigenvalues and the bounds on the coefficients needed to assess the cell loss rate at the statistical multiplexor. Although their algorithm considerably simplifies the computation of the eigenvalues and the bounds of the coefficients, the computation effort is still significant when the number of types of the sources and/or the number of states of each source type is large.

In [11], the authors reported an effort to represent an arbitrary time correlated cell generation process using two types of exponential on-off sources. It was demonstrated that an arbitray call or a group of calls can be closely approximated by an aggregate cell generation process of two types of exponential on-off sources as far as the effect on the queue length distributions is concerned. It, therefore, will be useful to have a solution for the multiplexor queueing problem where there are K types of exponential on-off sources. By specializing the analysis of Stern and Elwalid [10], the authors [12] obtained a simple approximation to the upper bound of the coefficients for such a problem. Even though the approximation simplifies the queueing analysis, the computation can be still time consuming especially when the number of types, K, is large. Moreover, it is very difficult to gain much insight into the system behaviour using such an analysis.

We discuss analyses of the queue length distribution of statistical multiplexing under the following assumptions:

1. There are K types of exponential on-off sources. Type k sources are described using four parameters $(\lambda_k, \mu_k, \gamma_k, N_k)$, where N_k is the number of type k sources, γ_k is the cell generation rate while in an on state, and λ_k^{-1} and μ_k^{-1} are, respectively, the mean duration of the on and off states.
2. All the source are independent of each other.
3. The multiplexing buffer is B cells in length and the output link capacity is c cells/sec.
4. We will concentrate on a node at which up to B cells be stored at its output buffer. Cells arriving to a full buffer are lost. The cell loss rate, P_e, is usually very low, 10^{-9}-10^{-11}, and can be approximated by

$$P_e \approx G(B),$$

where $G(x) \triangleq \lim_{t\to\infty} Pr[Q(t)>x]$, and $Q(t)$ is the number of cells in the buffer at time t assuming an infinite buffer size.

It is noted that a cell may be characterized by a combination of an arbitrary number of (elemental) sources.

As the system grows, i.e., as c increases, the number of sources that can be accommodated in each type, $N_k (k=1,\ldots,K)$, is likely to increase. Under this scenario, even a simple approximation given in [11] tends to be computationally time consuming. In this paper, we discuss a number of asymptotic approximations to the function $G(x)$ under two limiting conditions:

1. $x \to \infty$. This implies that an asymptote of form $\beta_0 e^{-z_0 x}$ is a good approximation to $G(x)$.
2. $N_k \to \infty$ while keeping the aggregate mean, variance and the time constant of the type k source constant at Λ_k, Σ_k^2 and η_k $(k=1,\ldots,K)$, where

$$\begin{aligned} \Lambda_k &= N_k \gamma_k p_k, \\ \Sigma_k^2 &= N_k \gamma_k^2 p_k (1-p_k), \\ \eta_k &= \lambda_k + \mu_k, \\ p_k &= \lambda_k / \eta_k \quad (\textit{activity factor}) \end{aligned} \tag{1}$$

This approximation, while not preserving the original process of cell generation, gives us several advantages.

1. The number of parameters needed to describe a particular type of source is reduced to three, namely $(\Lambda_k, \Sigma_k^2, \eta_k, \infty)$, from four in $(\lambda_k, \mu_k, \gamma_k, N_k)$ or equivalently $(\Lambda_k, \Sigma_k^2, \eta_k, N_k)$.
2. The reduced parameter set is easier to measure than the

original set. This allows a possibly easier policing by the network service provider after the user-network interface (UNI) unit.

3. The collection of type k source represented by the three parameter set($\Lambda_k, \Sigma_k^2, \eta_k, \infty$) behaves in the most pessimistic manner among the sources described by ($\Lambda_k, \Sigma_k^2, \eta_k, N_k$). As will be seen later, the dominant eigenvalue increase (decreases in absolute value) as N_k increasse among the group of sources which have the same values of Λ_k, Σ_k^2 and η_k. Therefore, the cell loss probability predicted using the approximate model tends to be larger than actual cell loss probability.
4. The use of the reduced parameter set gives a useful insight into how the parameters influence in determining the capacity requirement for a given type of sources satisfying the prescribed cell loss criterion.

For a statistical multiplexor with a single type of on-off sources ($K = 1$), Simonian and Virtamo [13] gave an asymptotic solution as $N_1 \to \infty$ and $x \to \infty$ by first approximating the aggregate cell generation rate with an Ornstein-Uhlenbeck(O-U) process and then applying the Benes queueing fromula [14]. Recently, Kobayashi and Ren [15] generalized the problem for $K \geq 1$ by first approximating the aggregate cell generation rate with a multi-dimensional O-U process and then solving a set of partial differential equations. These provide very simple asymptotes. However, as shown in Section 2, they tend to significantly underestimate the dominant eigenvalues especially when the channel utilization is low. Therefore, instead of conservative estimates of the cell loss rates, they are likely to give optimistic results.

We provide a more accurate asymptotic approximation to the queue length distribution in a closed form in the following two sections. In Section 3, we first generalize the analysis of Simonian and Virtamo [13] for the case of $K = 2$. This provides an alternative derivation of the dominant eigenvalue obtained by Kobayashi and Ren [15] for $K = 2$. It also allows us to obtain a better approximation to the coefficient corresponding to the dominant eigenvalue than that used in [15]. We then demonstrate, by numerical examples, that these approximations can be too optimistic when the channel utilization is moderate to low. In Section 4, we provide an alternative approach to obtain an approximation to the dominant eigenvalue by taking the limit $N_k \to \infty$ at a later stage of the analysis than the analysis of Section 3. This solution allows us to have a more accurate approximation while keeping the level of computational complexity to that in the previous analysis. The solution obtained in Section 3 is shown to be a special case of the more general solution obtained in this section.

Ⅲ. Asymptotic Analysis I

In this section, we first summarize the asymptotic approximation to $G(x)$ for $K = 2$ given in Appendix A. The results are essentially a generalization of the solution obtained by Simonian and Virtamo in [13]. The key here is to represent the aggregate cell generation rate as a two dimensional O-U process, use the Benes queueing formula, and expand the resulting integral to obtain an asymptotic expression for $G(x)$. For $K = 2$, we have

$$G(x) \sim \widetilde{\beta}_0 e^{-\widetilde{z}_0 x}, \quad (2)$$

where

$$\tilde{z}_o = \frac{c-\Lambda}{\sum_{k=1}^{K} \Sigma_k/\eta_k},$$

$$\tilde{\beta}_o = \frac{\Sigma}{\sqrt{2}\pi(c-\Lambda)} e^{-\tilde{z}_o^2 \delta},$$

$$\Lambda = \sum_{k=1}^{K} \Lambda_k,$$

$$\Sigma = \sqrt{\sum_{k=1}^{K} \Sigma_k^2},$$

$$\delta = \sum_{k=1}^{K} (\Sigma_k/\eta_k)^2.$$

It is noted that eqn. (2) reduces to the asymptotic approximation obtained by Simonian and Virtamo [13] when K = 1. Although the derivation in Appendix A is only for K = 2, it is conjectured that the above result can be generalized to an arbirary integer $K (\geq 1)$. In fact, it was shown by Kobayashi and Ren [15] recently that the dominant eigenvalue of their solution is identical to $\tilde{z}_0$ for any $K \geq 1$. It is also noted here that $\tilde{z}_0$ can be obtained as a first order approximation of the equation solved in Section 3. The utility of the present analysis is, therefore, not in the derivation of $\tilde{z}_0$ but in the derivation of a good approximation of the coefficient corresponding to the dominant eigenvalue, $\tilde{\beta}_0$.

In Fig. 1, the $\tilde{z}_0$'s (dotted lines) are compared with the exact values of the dominant eigenvalue (solid lines) for identical sources ($K = 1$, we omit all the subscripts denoting source type). In this figure, we keep the mean, variance of the aggregated cell arrival rate and the time constant (η) of the cell arrival rate at 10, 10 and 2, respectively, and vary the number of sources from 20 to 500. Also included in the same figure are the asymptotic values of the dominant eigenvalues (dashed lines) as $N \to \infty$ (This is derived in Section 3 and shown here for comparison.). It is shown that $\tilde{z}_0$ is a good approximation to the exact dominant eigenvalue when the channel utilization, $\rho (\triangleq \Lambda/c)$, is large. However, as ρ decreases, $\tilde{z}_0$ severely underestimates the dominant eigenvalue. Therefore, the use of $\tilde{z}_0$ in calculating the cell loss rate is likely to give an optimistic result.

In Fig. 2, we plot the coefficients corresponding to the donminant eigenvalues vs N using the same parameters as in Fig. 1. Exact coefficients are shown by solid lines and $\tilde{\beta}_0$'s are shown by dashed lines. In the same figure we show the coefficient used by Kobayashi and Ren [15] in dotted lines. They have approximated the coefficient using the theory of large deviation as

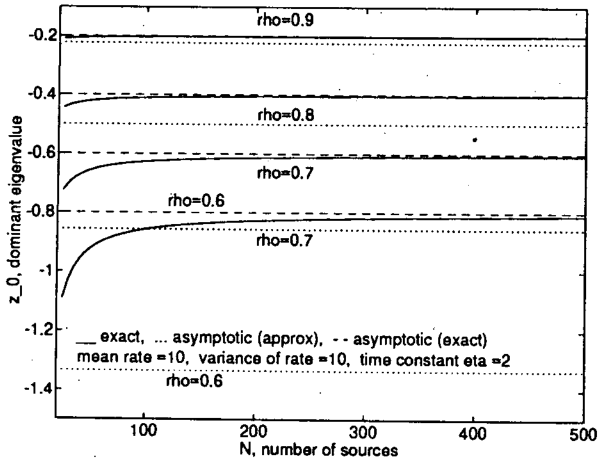


Fig. 1. Dominant eigenvalue vs number of sources(K = 1).

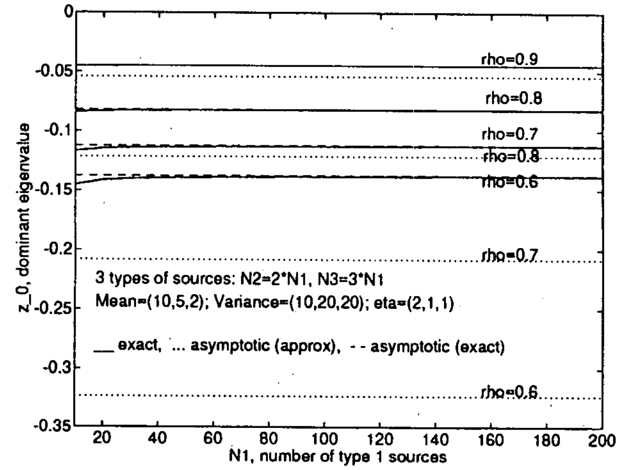


Fig. 3. Dominant eigenvalue vs number of sources(K = 3).

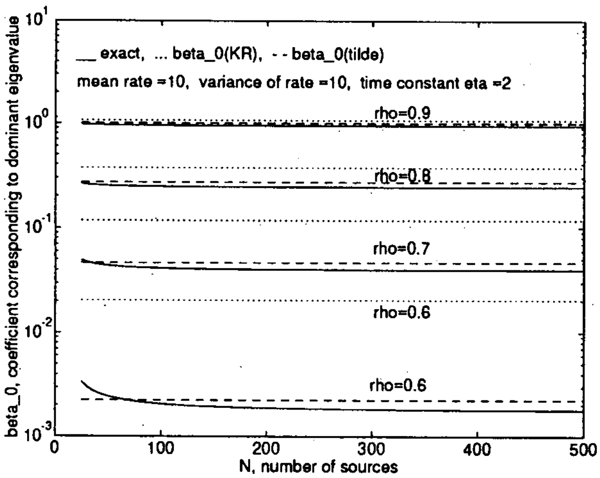


Fig. 2. Coefficient of the dominant eigenvalue vs number of sources(K = 1).

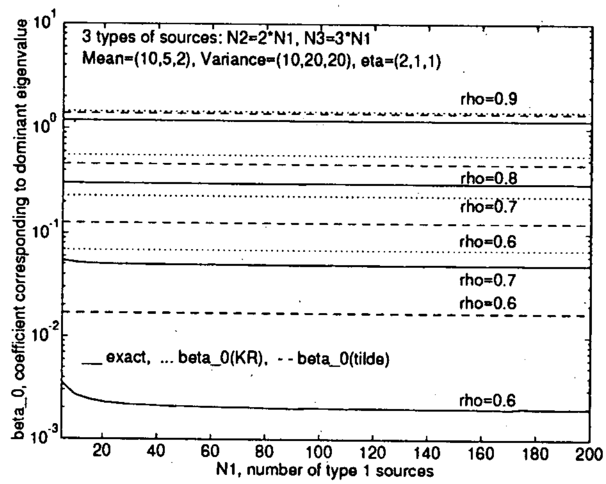


Fig. 4. Coefficient of the dominant eigenvalue vs number of sources(K = 3).

$$\beta_0^{(KR)} \approx \frac{1}{\sqrt{2\pi\theta^2\Sigma^2}}\, e^{-\Sigma^2\theta^2/2},$$

where $\theta \triangleq (c-\Lambda)/\Sigma^2$. It is seen that $\tilde{\beta}_0$ gives a very good approximation for a wide range of parameter values. It is also shown in the figure that $\beta_0^{(KR)}$ closely approximates β_0 for $\rho > 0.7$. However, as utilization decreases $\beta_0^{(KR)}$ tends to overestimate β_0 significantly. Although we do not have a theoretically exact limit on β_0 as N approaches infinity, it seems that convergence of β_0 is much more rapid than that of z_0. We also observe that the higher the channel utilization the faster convergence takes place.

In Figs. 3 and 4, respectively, we compare the exact and asymptotic approximations of dominant eigenvalue and the corresponding coefficients for $K=3$, $N_2 = 2N_1$, $N_3 = 3N_1$. We increase the number of sources in each type by the same factor with 3-parameter representations of type 1, 2 and 3 given by (10, 10, 2), (5, 20, 1) and (1, 20, 1), respectively. Similar observations as in Figs. 1 and 2 are made.

IV. Asymptotic Analysis II

It was shown in Section 2 that the asymptotic approximation obtained by first approximating the cell generation rate with an O-U process, while giving a very simple cell loss rate expression, can be too optimistic especially when the channel utilization is low. In this section we will derive an alternative asymptotic approximation which gives a pessimistic (conservative) estimate of the cell loss rate.

First we consider a system with only one type of sources (say type k). Suppose that N_k type k sources share a statistical multiplexor whose output link capacity is c_k. Assuming that the buffer size is infinite, the steady state probability that the number of cells is greater than x is asymptotically (as $x \to \infty$) given by $\beta_0 e^{-z_0 x}$, where the dominant eigenvalue z_0 is [9]

$$z_0 = \frac{\lambda_k}{c_k/N_k} - \frac{\mu_k}{\gamma_k - c_k/N_k}. \tag{3}$$

Expressing z_0 in terms of $(\Lambda_k, \Sigma_k^2, \eta_k, N_k)$ and $\rho_k(\triangleq \Lambda_k/c_k)$ we obtain

$$z_0 = -\frac{\Lambda_k \eta_k \rho_k (1-\rho_k)}{\rho_k \Sigma_k^2 - (1-\rho_k)\Lambda_k^2/N_k} \tag{4}$$

It can be seen that for $p_k < \rho_k < 1$, z_0 increases (decreases in absolute value) as we increase N_k while keeping Λ_k, Σ_k^2 and η_k constant. Define

$$z_0^* \triangleq \lim_{N_k \to \infty} z_0(N_k) = -\frac{\Lambda_k \eta_k}{\Sigma_k^2}(1-\rho_k). \tag{5}$$

Among the source whose mean and variance are respectively given by Λ_k and Σ_k^2 and time correlation is given by η_k, z_0^* is the upper bound of the dominant eigenvalue when the channel capacity is Λ_k/ρ_k. It is also noted by examining eqn. (4), z_0^* is a good approximation to z_0 for $\rho_k \approx 1$. The dashed lines represent z_0^* in Figures 1 and 3 and approach the solid lines representing exact values of z_0 as N_k and/or ρ_k increase.

Now consider the case of the heterogeneous inputs. The dominant eigenvalue can be obtained by solving

$$z_0 = -\frac{\Lambda_k \eta_k \rho_k (1-\rho_k)}{\rho_K \Sigma_k^2 - (1-\rho_k)\Lambda_k^2/N_k}, \ (k=1,\ldots,K), \tag{6}$$

with $\sum_{k=1}^{K} c_k = c$ [9, 10]. This effectively is saying that we can conceptually partition the system into K parallel subsystems: the k^{th} subsystem has an output channel of capacity c_k and has type k sources as its inputs. The dominant eigenvalue of the original system is obtained by partitioning the available capacity c into K regions, $(c_1, c_2, \ldots, c_K)$, in such a way that the dominant eigenvalues of the subsystems are the same, which in turn are identical to the dominant eigenvalue of the original system. Because the above equation does not lend itself to a simple solution, we seek a dominant eigenvalue for the case when the number of sources in each type of source approaches infinity. Therefore, we have a simpler problem.

$$z_0^* = -\frac{\Lambda_k \eta_k}{\Sigma_k^2}(1-\rho_k), \ (k=1,\ldots,K), \tag{7}$$

$$\rho_k = \Lambda_k / c_k, \tag{8}$$

$$c = \sum_{k=1}^{K} c_k. \tag{9}$$

Define $\triangle_k \triangleq \frac{\Sigma_k^2}{\Lambda_k \eta_k}$, then we have

$$\rho_k = 1 + z_0^* \triangle_k, \ (k=1,\ldots,K), \tag{10}$$

$$c_k = \frac{\Lambda_k}{1 + z_0^* \triangle_k}, \ (k=1,\ldots,K). \tag{11}$$

For $p_k < \rho_k < 1$, we have $0 < |z_0^* \triangle_k| = 1 - \rho_k < 1 - p_k < 1$. Therefore, eqn. (9) can be rewritten as

$$\sum_{k=1}^{K} \Lambda_k \sum_{j=1}^{\infty} [1 + (-1)^j (z_0^* \triangle_k)^j] = c \tag{12}$$

Because $z_0^* = -|z_0^*| \Lambda_k > 0$ and $\triangle_k \geq 0$, eqn. (12) can be written as

$$\sum_{k=1}^{K} \Lambda_k \sum_{j=0}^{\infty} (|z_0^*| \triangle_k)^j = c. \tag{13}$$

We denote $z_0^{*(J)}$ as the J^{th} order approximation of z_0^*, which satisfies the truncated equation

$$\sum_{j=0}^{J} S_j (|z_0^{*(J)}|)^j = c, \tag{14}$$

where $S_j \triangleq \sum_{k=1}^{K} \Lambda_k \triangle_k^j (j = 0, 1, \ldots)$. If there is at least one type of source whose variance of cell arrival rate Σ_k^2 is non-zero, $S_j > 0$ for all $j \geq 0$. $S_j = 0$ $(j > 0)$ corresponds to the system with only constant rate sources. For the remainder of this paper we assume that there is at least one type of sources whose cell generation rate is not constant.

Since $S_j > 0$, $(j \geq 0)$, the following facts may be easily verified.

Fact 1. $|z_0^{*(J)}|$ is a monotonically decreasing function of J.

Fact 2. There exists a unique positive solution for $|z_0^{*(J)}|$ of eqn. (14). Other solutions, if there is any, are either negative or conjugate complex pairs.

First order approximation, $z_0^{*(1)}$:

$$z_0^{*(1)} = -\frac{c - S_o}{S_1 - 1} = \frac{c - \sum_{k=1}^{K} \Lambda_k}{\sum_{k=1}^{K} \Lambda_k \triangle_k} \tag{15}$$

It is noted that $z_0^{*(1)}$ is identical to the dominant eigenvalue $\tilde{z}_0$ obtained in Section 2 for $K = 2$ and also obtained by Kobayashi and Ren for $K \geq 1$ [15]. This approximation is only good when the ρ_k's are close to 1. (See Figures 1 and 3. The dotted line corresponds to the first order approximation.)

Second order approximation, $z_0^{*(2)}$: This solution of the second order equation is

$$z_0^{*(2)} = -|z_0^{*(2)}| = \frac{S_1 - \sqrt{S_1^2 + 4S_2(c - S_0)}}{2S_2} \tag{16}$$

Third order approximation, $z_0^{*(3)}$: The solution of the second order equation is

$$z_0^{*(3)} = -|z_0^{*(3)}| = \frac{S_2/3 - (T_1 + T_2)}{S_3} \tag{17}$$

where

$$T_1 = \sqrt[3]{-\frac{a_1}{2}+\sqrt{T_o}},$$

$$T_2 = -\sqrt[3]{\frac{a_1}{2}+\sqrt{T_o}},$$

$$T_0 = \frac{a_1^2}{4}+\frac{a_0^3}{27},$$

$$a_0 = \frac{3S_3S_1-S_2^2}{3},$$

$$a_1 = \frac{2S_2^3-9S_3S_2S_1+27S_3^2(S_0-c)}{27}$$

In Appendix B, it is shown that $|z_0^{*(3)}|$ in eqn. (17) is the unique positive solution of the truncated third order equation.

In Fig. 5, the J^{th} order approximation of the dominant eigenvalue vs J is plotted for a single type of source with $\Lambda = 10$, $\Sigma^2 = 10$ and $\eta = 2$. It is seen that the approximation converges to the exact value of z_0^* rapidly as the order J increases. Even for a relatively low utilization ($\rho = 0.6$), the second or third order approximation seems to be accurate enough. Furthermore, since z_0^* is an upper bound of z_0, we may safely use these lower order approximations ($z_0^{*(2)}$ or $z_0^{*(3)}$) to predict the cell loss rate in a fairly conservative manner. We will examine these later in this section.

Before presenting numerical examples to illustrate the use of these approximations, we need an approximation to the coefficient of the corresponding dominant eigenvalue. Attempts were made to obtain an asymptotic expression for β_0 using the previous analysis on the upper bound of β_0 [12]. However, a useful, yet simple, solution has not been obtained. We, therefore, use $\tilde{\beta}_0$ obtained in Section 2 to approximate β_0. In Fig. 6, we compare the exact upper bound (solid line), exact asymptote[1] ($\beta_0 e^{-z_0 x}$, dashed line) and approximate asymptote ($\tilde{\beta}_0 e^{-z_0^{*(J)}}$: $J = 1$ star(*), $J = 2$ dotted line, $J = 3$ dot-dash(-.-)) of $G(x)$ for two types of source whose

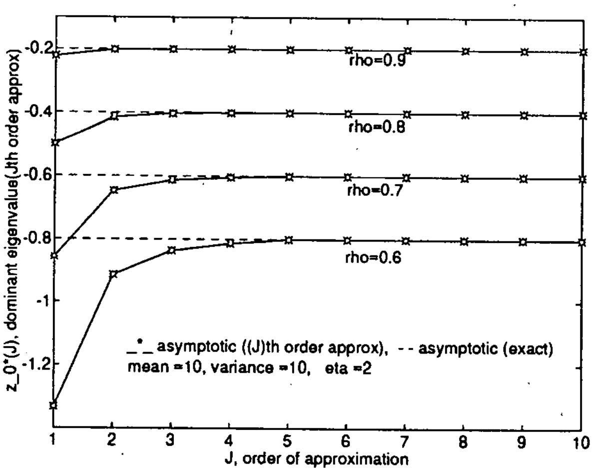


Fig. 5. Dominant eigenvalue vs the order of approximation(J).

1) Even if we refer to this as an exact asymptote, we use an approximate value of β_0 [12]. The dominant eigenvalue, however, is exact.

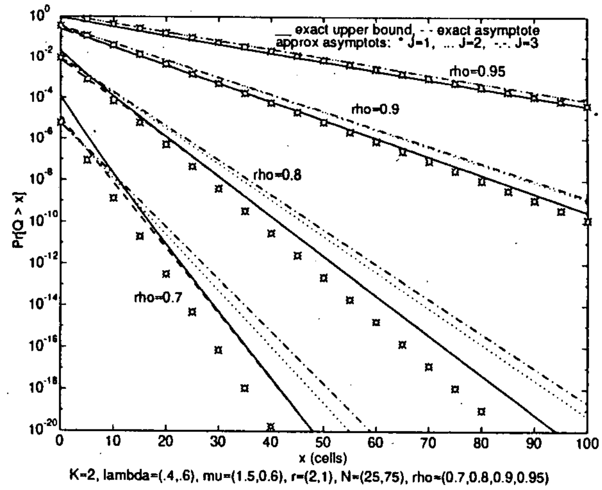


Fig. 6. Comparison of $G(x)$ with various asymptotes($K = 2$).

parameters are shown in the figure. For a large range of utilization, the approximation gives an upper bound on the complementary queue length distribution. In this example, the first order approximation seems to be reasonably accurate for $\rho \geq 0.9$. However, the cell loss rate predicted using this approximation can be lower than actual cell loss rate for $\rho < 0.9$. The difference between the second order approximation and the third order approximation seems to be negligible. The second and the third approximations consistently give an upper bound on G(x) except for a very small value of x; therefore, it seems that we can use them to conservatively estimate the cell loss rate.

V. Required Channel Capacity, Performance Measure and Traffic Descriptors

Although c_k, the portion of channel capacity assigned conceptually to type k traffic (see Section 3), is not exclusively used by the type k traffic sources, it gives a rough measure of resource utilized by type k sources collectively. Examining $c_k = \frac{\Lambda_k}{1-|z_0^*|\triangle_k}$ more carefully, we observe the following:

1. Since $0 \leq |z_0^*|\triangle_k < 1 - p_k$, we have $\Lambda_k \leq c_k < \Lambda_k / p_k$. Note that Λ_k / p_k is the aggregate peak cell generation rate of type k source.
2. $\triangle_k (= \Sigma_k^2 / (\Lambda_k \eta_k))$ is a measure of dispersion of cell generation rate which combines the effects of mean, variance and the time correlation of the cell generation rate. The larger Σ_k^2 / Λ_k is, the larger c_k is needed. Also, the larger the time corrleation of rate generation (the larger mean burst length (η_k^{-1})) is, the larger c_k is needed.
3. c_k is also a function of $|z_0^*|$ which depends on the maximum tolerable cell loss rate, P_e, and the buffer capacity, B, and the system multiplexing gain reflected on β_0 through

$$|z_0^*| = -\frac{1}{B}\log(\frac{P_e}{\beta_0}).$$

Althogh β_0 depends on type k traffic, unlike $\triangle_k$, its dependence is indircet.

Consider the first order approximation of c_k (from eqn. (11)):

$$c_k \approx \Lambda_k(1+|z_0^*|\triangle_k).$$

The extra capacity over the mean cell arrival rate Λ_k for type k traffic sources to absorb traffic dispersion in order to guarantee a certain level of cell loss performance criterion is given by

$$\triangle c_k \triangleq c_k - \Lambda_k \approx \frac{1}{B}\log(\frac{\beta_0}{P_e})\cdot(\frac{\Sigma_k^2}{\eta_k}).$$

Note that $\triangle c_k$ is approximately proportional to the rate variance (Σ_k^2) and the mean burst length ($1/\eta_k$), and inversely proportional to the buffer size (B). (Recell that we ae considering a liminting case $N_k \to \infty$.) Also note that one can reduce $\triangle c_k$ by increasing the maximum tolerable cell loss rate, P_e, and/or increasing the multiplexing gain which is accompanied by a decrease in β_0.

If the traffic of a call is smoothed by reducing its peak rate by $1/2$ thus reducing Σ_k^2 by $1/2$ then we can decrease the required excess capacity, $\triangle c_k$, roughly by $1/2$ while keeping B and P_e the same. If the contribution of the traffic to the system is not the major one, the assumption that the change in β_0 is not so significant is reasonable.

The discussions in this section so far seem to suggest that two parameters Λ_k and $\triangle_k$ are sufficient (at least asymptotically) to describe the queueing behaviour of a statistical multiplexor. However, in order to determine β_0 or even its approximation $\tilde{\beta}_0$, we need more detailed description of the call.

VI. Measurement of Traffic Descriptor Values

It is necessary for one to determine the values of mean, variance and the burst length of the cell arrivals of a call. In this section, we consider a call process consisting of a number of identical and independent exponential on-off sources [4, 6]. Then, it is possible to obtain all the necessary values of the traffic descriptor parameters of a call ($\Lambda_l, \Sigma_l^2, \eta_l$) by using the estimate of the autocorrelation fuction of the cell generation rate. Let

$$R(\tau) \triangleq E[R_l(t)R_l(t+\tau)],$$

where $R_l(t)$ is the cell generation rate of call l. Then we have relations

$$R(0) = \Lambda_l^2 + \Sigma_l^2 \quad (18)$$

$$\lim_{\tau\to\infty} R(\tau) = \Lambda_l^2 \quad (19)$$

$$\int_0^\infty [R(\tau) - \Lambda_l^2]d\tau = \frac{\Sigma_l^2}{\eta_l} \quad (20)$$

Thus we are able to estimate the values of the traffic descriptor parameters using the autocorrelation function $R(\tau)$.

Based on the discussions in [11] where, as far as the queue length distribution is concerned, an arbitrary call can be approximated by a set of two types of exponential on-off sources, it would be useful to obtain the values of two sets of traffic descriptor parameters(types l and m) constituting a call from the estimate of the autocorrelation function of the cell generation process of the call. Of course, one can use the method used in [11] in which the number of moments of steady state cell arrival rates and the indices of dispersion of the number of cells arriving in an interval are matched to obtain the two-type representation. However, the method is potentially time consuming to use. Attempts to obtain the parameter values using the autocorrelation function of the cell arrival process have not been successful. It is possible to obtain Σ_l^2, $\Sigma_m^2, \eta_l, \eta_m$ and $\Lambda_l + \Lambda_m$ by fitting the autocorrelation function with the theoretical curve. However, a separation of the mean cell arrival rate between types l and m has not been achieved. Since we are looking for a fairly gross approximation, we may be satisfied with the single type representation for which we have a simple method of determining the necessary values of the traffic descriptor parameters.

A question may be raised about how to deal with the more realistic case where the constituent sources are not exponential on-off sources. Rather the elemental constituent source is general on-off source. To that end we assume that eqn. (20) is also valid for an arbitrary source. (Note that eqns (18) and (19) are valid for an arbitrary on-off source.) This allows one to have a simple mapping of an arbitrary on-ff source to an equivalent exponential on-off source. Once the autocorrelation function of the source is obtained, from eqns (18) and (19), one obtains Λ_l and Σ_l^2. Using eqn. (20), one then obtains the equivalent value of η_l, $\eta_l^{(eq)}$.

For example, since the autocorrelation function of a periodic on-off source is also periodic, $\eta_l^{(eq)}$ of an equivalent on-off source corresponding to the source must be ∞[2]. Therefore, $\triangle_l = 0$ and $c_l = \Lambda_l$ for a periodic on-off source. This is reasonable because our analysis is based on that the queue lenght is large ($x \to \infty$) and when the queue length is sufficiently large the variation of the cell arrivals due to periodically alternating on and off states of cell generations can be absorbed by the buffer alone without providing excess link capacity.

Consider a system of N homogeneous on-off source in which the length of an off period is exponentially distributed and that of an on period is either Erlang-m or hyper-exponentially distributed.

2) In this case the $\lim_{\tau\to\infty} R(\tau)$ is defined as the average of the autocorrelation function for one period.

The mean on and off periods are μ^{-1} and λ^{-1}, respectively. Therefore, we have $\mu_0 = m\mu$ for the Erlang-m distribution and $\mu = [\sum_{i=1}^{m}(\sigma_i/\mu_i)]^{-1}$ for the hyper-exponential distribution (refer to Fig. A.1). In Appendix C, we obtained the integral in eqn. (20) as

$$\int_0^\infty [R(\tau)-\Lambda^2]d\tau = \begin{cases} (\frac{m+1}{2m}) \left[\frac{\lambda\mu}{(\lambda+\mu)^3}\right], \text{Erlang} \\ (\sum_{i=1}^{m}\frac{\sigma_i}{\mu_i^2}) \left[\frac{\lambda\mu}{(\lambda+\mu)^3}\right], \text{hyper-exponential} \end{cases} \quad (21)$$

Here, we generalize the time constant η to denote the inverse of the mean cycle length (the sum of mean of on and off periods). Then, from eqn. (20), the equivalent time constant, $\eta^{(eq)}$ is given by

$$\eta^{(eq)} = \begin{cases} \left[\frac{2m}{m+1}\right]\cdot\eta, \text{Erlang} \\ \left[\sum_{i=1}^{m}\frac{\sigma_i}{\mu_i^2}\right]^{-1}\cdot\eta, \text{hyper-exponential} \end{cases}$$

By taking the limit, $m\to\infty$, of the Erlang-m distribution we fine that $\eta^{(eq)} = 2\eta$ for the sources with constant on periods.

The squared coefficient of variation of an on period, C_{ON}^2, is given by

$$C_{ON}^2 = \begin{cases} \frac{1}{m}(\geq 1), & \text{Erlang} \\ \mu^2\sum_{i=1}^{m}\frac{\sigma_i}{\mu_i^2}(\leq 1), & \text{hyper-exponential} \end{cases}$$

It is possible to realize various values of C_{ON}^2 using Erlang-m and hyper-exponential distributions for an on period. In Fig. 7, we let the mean and the variance of the aggregate cell generation rate and η are 1, 1 and 2, respectively. In order to reduce the number of parameters of the hyper-exponential sources, we let $m = 2$ and

$$\mu_1 = 2\sigma_1\mu,$$

$$\mu_2 = 2\sigma_2\mu.$$

We then have $\eta^{(eq)} = 4\sigma_1\sigma_2\eta$. Comparisons between the exact and approximate results are made for various channel utilizations and coefficient of variations of an on period. The exact results are obtained using the formula given by Kosten [9] with $N = 10$. The approximate results are obtained using eqn. (5) with η_k replaced by $\eta^{(eq)}$. The figure indicates that the proposed approach gives a resonable estimate of the dominant eigenvalue even for the non-exponential on-off sources when the channel utilization is not less than 0.75. As the utilization decreases *and* the coefficient of variation decreases, the approximation deviates from the exact result significantly. The difference is most noticeable for the sources with constant on periods ($C_{ON}^2 = 0$). We know that there are infinite number of eigenvalues for a system with sources of constant on periods; the dominant eigenvalue alone may be insufficient to characterize the queueing behaviour in that case. The actual behaviour is more likely to be the result of many eigenvalues near the dominant one. As a result the approximation may in fact be reflecting the aggregated effect of the many eigenvalues near the dominant one more accurately than it appears in the figure.

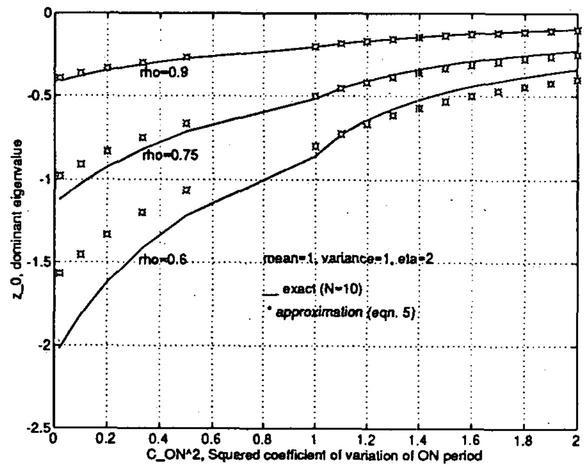


Fig. 7. Exact and approximate values of dominant eigenvalues vs C_{On}^2.

VII. Conclusions

In this paper, we have analyzed the queue length distributions of a statistical multiplexor for heterogeneous exponenetial on-off sources. Under the condition that the queue length is large ($x\to\infty$) and the number of sources of each type is large ($N_k\to\infty$), a simple, yet reasonable, approximation for the queue length distribution can be obtained. Besides giving us a conservative estimate of the cell loss rate, this simplification allows us to observe the relationship among the required capacity, maximum cell loss rate, buffer size, the mean and the variance of cell arrival rate, and the time correlation of the cell arrival rate.

We have provided a brief discussion of how one may measure the values of the traffic descriptors from the autocorrelation function of the cell arrival process. A possible extension of the proposed analysis for arbitrary (not necessarily exponential) on-off sources is also discussed.

Numerical examples indicate that the proposed approach offers a simple characterization of the ATM traffic which captures the most important features of a time correlated cell arrival process. Although, the analytic results using the proposed method deviate from the exact results, the fact that they tend to be conservative and that they provide a useful and explicit relationship among the network parametrs (maximum cell loss rate, buffer size, multiplexing gain), the user traffic parameters (mean, variance of the cell generation rate and time correlation) and the required capacity suggests the usefulness of such a simple characterization of the ATM traffic and the corresponding asymptotic queueing analysis.

Appendices

A. Derivation of β_0

We have the following problem: There are two types of sources. At time t, sources of type k generate cells at the rate of $R_k(t)$ which satisfies

$$dR_k(t) = \eta_k(\Lambda_k - R_k(t))dt + \sqrt{2\eta_k}\Sigma_k dW_k(t),$$

where η_k, Σ_k^2 and p_k are given by eqn. (1) in the text, and $W_k(t)$ is a white Gaussian process with mean zero and unit variance. It is known that as N_k approaches infinity, the aggregated cell arrival rate from sources of type k can be described by the above equation [13]. The stationary distribution of $R_k(t)$ is known to be Gaussian with mean Λ_k and variance Σ_k^2. It is also known that total number of cells arriving in $[0, \tau]$, $K(\tau)$, is also Gaussian with mean $M(\tau, l_1, l_2)$ and variance $v(\tau)$ given that $R_k(0) = l_k$ ($k = 1, 2$), where

$$\begin{aligned} M(\tau, l_1, l_2) &= \sum_{k=1}^{2}[\Lambda_k\tau + (l_k - \Lambda_k)h(\eta_k\tau)/\eta_k], \\ v(\tau) &= \sum_{k=1}^{2}(\Sigma_k/\eta_k)^2 g(\eta_k\tau), \\ h(\theta) &= 1 - e^{-\theta}, \\ g(\theta) &= 2\theta - 3 + 4e^{-\theta} - e^{-2\theta}. \end{aligned}$$

Our starting point is a generalization of equation (2.10) of [13], that is

$$G(x) \le \int_0^\infty d\tau \int\int_{l_1+l_2\le c} q_\tau(x, l_1, l_2)(c - l_1 - l_2)d\Psi(l_1)d\Psi(l_2), x>0 \quad \textbf{(A1)}$$

where $q_\tau(x, l_1, l_2) = Pr[K(\tau) - c\tau = x \mid R_1(0) = l_1, R_2(0) = l_2]$ and $\Psi_k(l_k)$ is the stationary distribution of $R_k(t)$.

Owing to the fact that $K(\tau)$ and $R_k(t)$ are Gaussian, we can write the left hand side of eqn. (A1), denoted by $S(x)$, as

$$S(x) = \int_0^\infty d\tau \int_- \int^c dl_1\psi_1(l_1)\int_{-\infty}^{c-l_1}(c - l_1 - l_2)q_\tau(x, l_1, l_2)\psi_2(l_2)dl_2$$

where

$$\begin{aligned} q_\tau(x, l_1, l_2) &= \frac{1}{\sqrt{2\pi v(\tau)}}\exp\left[-\frac{[x + c\tau - M(\tau, l_1, l_2)]^2}{2v(\tau)}\right], \\ \psi_k(l_k) &= \frac{1}{\Sigma_k\sqrt{2\pi}}\exp\left[-\frac{(l_k - \Lambda_k)^2}{2\Sigma_k^2}\right], \quad k = 1, 2. \end{aligned}$$

Let $\xi_k = (l_k - \Lambda_k) / \Sigma_k$, $\gamma_k = (c - \Lambda_k) / \Sigma_k$, ($k = 1, 2$), and $\Lambda = \Sigma_{i=1}^2 \Lambda_i$. Then,

$$S(x) = \int_0^\infty \frac{d\tau}{\sqrt{2\pi v(\tau)}} \int_\infty^{\gamma_1} \frac{d\xi_1}{\sqrt{2\pi}}\exp\left(-\frac{\xi_1^2}{2}\right)I_1(\tau, \xi_1), \qquad \textbf{(A2)}$$

where

$$\begin{aligned} I_1(\tau, \xi_1) &= \exp\left[-\frac{(x + c\tau - \Lambda\tau - \xi_1\Sigma_1 h_1)^2}{2(\Sigma_2^2 h_2^2 + v(\tau)}\right] \\ &\cdot\left\{\frac{\Sigma_2}{a}(\gamma_2 + b)N[a(\gamma_2 + b)] + \frac{\Sigma_2}{a^2\sqrt{2\pi}}\exp\left[-\frac{a^2(\gamma_2 + b)^2}{2}\right]\right\} \end{aligned}$$

$$\begin{aligned} a &= \frac{\sqrt{\Sigma_2^2 h_2^2 + v(\tau)}}{v(\tau)}, \\ b &= -\frac{(x + c\tau - \Lambda\tau - \xi_1\Sigma_1 h_1)\Sigma_2 h_2}{\Sigma_2 h_2^2 + v(\tau)}, \\ h_k &= h(\eta_k\tau)/\eta_k, \quad (k = 1, 2), \\ N(x) &= \int_\infty^x \frac{1}{\sqrt{2\pi}}e^{-\gamma^{2/2}}d\tau. \end{aligned}$$

Since $N(\cdot) \le 1$, we have an upper bound of $I_1(\tau, \xi_1)$, $\tilde{I}_1$, by replacing $N[a(\gamma_2 + b)]$ by 1 in I_1. We then have

$$\begin{aligned} S(x) &\le \int_0^\infty \frac{d\tau}{\sqrt{2\pi v(\tau)}}\int_\infty^{\gamma_1}\frac{d\xi_1}{\sqrt{2\pi}}\exp\left(-\frac{\xi_1^2}{2}\right)\tilde{I}_1(\tau, \xi_1) \\ &= \int_0^\infty \frac{d\tau}{\sqrt{2\pi v(\tau)}} I_2(\tau), \end{aligned}$$

where

$$I_2(\tau) = \int_\infty^{\gamma_1}\frac{d\xi_1}{\sqrt{2\pi}}\tilde{I}_1(\tau, \xi_1).$$

After some algebraic manipulations, we obtain, $I_2 = I_2^0 + I_2^1$, where

$$I_2^0(\tau) = \frac{A\Sigma_2}{a\sqrt{2\pi}}\exp\left\{-\frac{(x + c\tau - \Lambda\tau)^2}{2[\Sigma_1^2 h_1^2 + \Sigma_2^2 h_2^2 + v(\tau)]}\right\}$$

$$I_2^1(\tau) = \frac{\Sigma_2}{a^2 u_2\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(u_0 - u_1^2/u_2^2)\right\}N[u_2(\gamma_1 - u_1/u_2^2)],$$

$$\begin{aligned} A &= \frac{\sqrt{2\pi}}{v_0}(-w_1 v_1 + w_0)N[v_0(\gamma_1 - v_1)] + \frac{w_1}{v_0^2}\exp\left[-\frac{v_0(\gamma_1 - v_1)^2}{2}\right], \\ w_1 &= \frac{\Sigma_1}{\Sigma_2} - \frac{\Sigma_1\Sigma_2 h_1 h_2}{\Sigma_2^2 h_2^2 + v(\tau)} \\ w_0 &= \frac{c - \Lambda}{\Sigma_2} - \frac{(x + c\tau - \Lambda\tau)\Sigma_2 h_2}{\Sigma_2^2 h_2^2 + v(\tau)}, \\ v_1 &= \frac{(x + c\tau - \Lambda\tau)\Sigma_1 h_1}{\Sigma_1^2 h_1^2 + \Sigma_2^2 h_2^2 + v(\tau)} \\ v_0 &= \sqrt{\frac{\Sigma_1^2 h_1^2 + \Sigma_2^2 h_2^2 + v(\tau)}{\Sigma_2^2 h_2^2 + v(\tau)}}, \\ u_2 &= \sqrt{1 + a^2 w_1^2 + \frac{\Sigma_1^2 h_1^2}{\Sigma_2^2 h_2^2 + v(\tau)}}, \end{aligned}$$

$$u_1 = w_1 w_0 a^2 + \frac{(x+c\tau-\Lambda\tau)\Sigma_1 h_1}{\Sigma_2^2 h_2^2 + v(\tau)},$$

$$u_0 = w_0^2 a^2 + \frac{(x+c\tau-\Lambda\tau)^2}{\Sigma_2^2 h_2^2 + v(\tau)}$$

Define

$$J_i(x) = \int_0^\infty \frac{d\tau}{\sqrt{2\pi v(\tau)}} I_2^i(\tau), \quad i=1,2$$

Then, using Laplace's method [16] one can show that as $x \to \infty$

$$J_0(x) \sim \exp\left[-\frac{F(\tau)}{2}\right]\frac{\Sigma_1}{\sqrt{2\pi(c-\Lambda)}}\exp\left\{-\frac{1}{2}[v_0(\gamma_1-v_1)]^2\right\}$$

$$J_1(x) \sim \exp\left[-\frac{H(\tau)}{2}\right]\frac{\Sigma_1}{\sqrt{2\pi(c-\Lambda)}} \times\sqrt{\frac{\Sigma_2}{\Sigma_1+\Sigma_2}}N[u_2(\gamma_1-\frac{u_1}{u_2^2})]$$

with τ evaluated at $\tau^* = \frac{x}{c-\Lambda}+2(\frac{\delta_1+\delta_2}{\delta_1\eta_1+\delta_2\eta_2}), \delta_k=(\Sigma_k/\eta_k)^2 \cdot (k=1,2)$ and

$$F(\tau) = \frac{(x+c\tau-\Lambda\tau)^2}{\Sigma_1^2 h_1^2 + \Sigma_2^2 h_2^2 + v(\tau)}$$

$$H(\tau) = \frac{(x+c\tau-\Lambda\tau)^2}{\Sigma_2^2 h_2^2 + v(\tau)}$$

Noting that $H(\tau^*) \geq F(\tau^*)$, one has

$$S(x) \lesssim \exp\left[-\frac{F(\tau^*)}{2}\right]\frac{\Sigma_1+\Sigma_2(\sqrt{\Sigma_2/(\Sigma_1+\Sigma_2)})}{\sqrt{2\pi(c-\Lambda)}}. \quad \textbf{(A3)}$$

Here the symbol $\lesssim$ denotes an asymptotic upper bound. As $x \to \infty$, the left hand side of eqn. (A3) in turn approaches $\beta'_0 e^{-\tilde{z}_0 x}$, where

$$\tilde{z}_0 = \frac{c_\Lambda}{\Sigma_1^2/\eta_1 + \Sigma_2^2/\eta_2},$$

$$\beta'_0 = \frac{\Sigma_1+\Sigma_2(\sqrt{\Sigma_2/(\Sigma_1+\Sigma_2)})}{\sqrt{2\pi(c-\Lambda)}} e^{-\tilde{z}_0^2(\delta_1+\delta_2)}$$

Since the role of type 1 and 2 can be interchanged, another formula for β'_0 can be obtained by interchanging the subscripts 1 and 2 in the above expression. For simplicity we will *approximate* minimum of $\Sigma_1+\Sigma_2(\sqrt{\Sigma_2/(\Sigma_1+\Sigma_2)})$ and $\Sigma_2+\Sigma_1(\sqrt{\Sigma_1/(\Sigma_1+\Sigma_2)})$ by $\sqrt{(\Sigma_1^2+\Sigma_2^2)}$. This not only simplifies the resulting expression, but also makes the analysis of Simonian and Virtamo [13] to be a special case of our analysis. With this approximation, we finally have

$$G(x) \lesssim \beta_0 e^{-\tilde{z}_0 x}, \quad \textbf{(A4)}$$

where

$$\tilde{z}_0 = \frac{c-\Lambda}{\Sigma_1^2/\eta_1 + \Sigma_2^2/\eta_2},$$

$$\beta_0 = \frac{\Sigma}{\sqrt{2\pi(c-\Lambda)}} e^{-\tilde{z}_0^2 \delta},$$

$$\Lambda = \sum_{k=1}^{2} \Lambda_k,$$

$$\Sigma = \sqrt{\sum_{k=1}^{2} \Sigma_k^2},$$

$$\delta = \sum_{k=1}^{2} (\Sigma_k/\eta_k)^2.$$

B. Proof of eqn. (17)

There are three solutions. We need to show that of the three solutions eqn. (17) gives the desired positive solution. If $T_0 >$ eqn. (13) has only one real root [17]. Owing to fact 2 of Section 3, this root must be positive. Therefore, we only need to show that $T_0 > 0$ and $\sqrt{T_0} > \alpha_1/2$ to prove that the soultion given in eqn. (17) is indeed positive. If at least one of $\triangle_k$ is non-zero, then $S_j > 0$, $(j \geq 0)$; therefore,

$$\begin{aligned} a_0 &= 3S_3S_1 - S_2^2 \\ &> S_3S_1 - S_2^2 \\ &= \sum_{k=1}^{K}\sum_{l>k}^{K}\Lambda_k\Lambda_l\triangle_k\triangle_l(\triangle_l^2 - \triangle_k\triangle_l) \\ &= \sum_{k=1}^{K}\sum_{l>k}^{K}\Lambda_k\Lambda_l\triangle_k\triangle_l(\triangle_l^2 - 2\triangle_k\triangle_l + \triangle_k^2) \\ &= \sum_{k=1}^{K}\sum_{l>k}^{K}\Lambda_k\Lambda_l\triangle_k\triangle_l(\triangle_l - \triangle_k)^2 \\ &\geq 0. \end{aligned}$$

This proves that $T_0 > 0$. Also $\sqrt{T_0} > a_1/2$, therefore, both T_1 and T_2 are real.

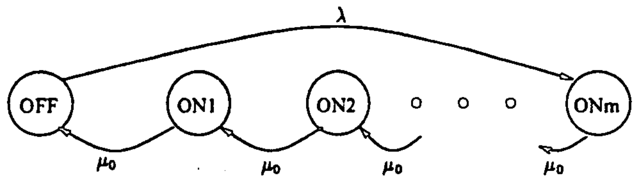
C. Derivation of eqn. (21)

A source is in one of the (m + 1) states (see Fig. A1). Let $S(t)$ and $R(t)$ denote the state and the cell generation rate of the source, respectively at t.

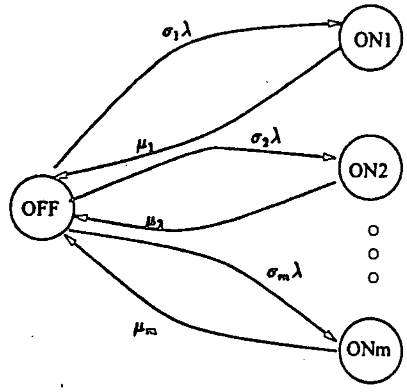
$$R(t) = \begin{cases} 0, & \text{if } S(t) = 0 \ (\textit{OFF state}) \\ 1, & \text{if } S(t) = i \ (i=1,\dots,m) \ (\textit{ONi state}) \end{cases}$$

The random process $S(t)$ is a continuous time Markov process whose infinitestimal generator matrix is denoted by Q[3]. Define

3) In this Appendix, Q, S and F are used differently than in Appendix A.



(a) Erlang-m distribution



(b) Hyper-exponential distribution

Fig. A1. Source with Erlang-m or hyper-exponential on periods.

$$\pi \triangleq (\pi_o, \pi_1, \ldots, \pi_m),$$
$$\pi_i \triangleq \lim_{t\to\infty} Pr[S(t)=i],\ (i=0,1,\ldots m),$$
$$Y(t) \triangleq \{y_{ij}(t)\},$$
$$Y^*(s) \triangleq \{Y^*_{ij}(s)\},$$
$$y_{ij}(t) \triangleq Pr[S(t)=j \mid S(0)=i],$$
$$Y^*_{ij}(s) \triangleq \mathcal{L}[y_{ij}(t)],\ (i,j=0,1,\ldots,m).$$

The equilibrium distribution of $S(t)$ is obtained by

$$\pi Q = 0, \qquad \pi e = 1,$$

where e is a column vector of ones. Using the initial condition that

$$y_{ij}(0) = \delta_{ij},$$

and the forward Chapman-Kolmogorov equation

$$\frac{dY(t)}{dt} = Y(t)Q,$$

we find $Y^*(s)$ as

$$Y^*(s) = (sI-Q)^{-1}, \tag{A5}$$

where I is an identity matrix.

The autocorrelation function of the cell generation rate of the source, $R(\tau)$, is given by

$$R(\tau) = Pr[R(t)=1]\cdot Pr[R(t+\tau)=1 \mid R(t)=1]$$
$$= \sum_{i=1}^{m} \pi_i \cdot (1 - Pr[S(t+\tau)=0 \mid S(t)=i]).$$

But

$$Pr[S(t+\tau)=0 \mid S(t)=i] = \int_0^{\tau} f_i(\tau_0) y_{oo}(\tau-\tau_0)d\tau_0,$$

where $f_i(\tau)$ is the probability density function of the remaining on period until source reaches an OFF state starting from an ONi state. The Laplace transform of $R(\tau)$, $\Phi(s)$, is given by

$$\Phi(s) = \sum_{i=1}^{m} \pi_i \left\{ \frac{1}{s} - F^*_i(s) Y^*_{00}(s) \right\},$$

where $F^*_i(s) \triangleq \mathcal{L}[f_i(\tau)]$. Using the final value theorem of the Laplace transform, we have

$$\int_0^{\infty} [R(\tau)-\Lambda^2]d\tau = \lim_{s\to 0} s\cdot \mathcal{L}\left\{\int_o^{\infty}[R(\tau)-(1-\pi_o)^2]d\tau\right\}$$
$$= \lim_{s\to 0}\left[\Phi(s) - \frac{(1-\pi_0)^2}{s}\right]. \tag{A6}$$

Case 1 : On period is Erlang-m distrbuted (Fig. A1-(a)) :

The on period is Erlang-m distrubed: there are m stages in an on period; each stage of the on period is exponentially distrubuted with mean μ_0^{-1}. The infinitesimal generator is given by

$$Q = \begin{pmatrix} -\lambda & 0 & 0 & \ldots & \lambda \\ \mu_0 & -\mu_0 & 0 & \ldots & 0 \\ 0 & \mu_0 & -\mu_0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & 0 & \ldots & -\mu_0 \end{pmatrix},$$

where $\mu_0 = m\mu$. The equilibrium distribution of the source state is

$$\pi_o = \frac{\mu}{\lambda+\mu},$$
$$\pi_i = \frac{\lambda/\mu}{\lambda+\mu},\ (i=1,\ldots,m).$$

The remaining on period while the source is in an ONi state is Erlang-i distributed. Therefore, we have

$$F^*_i(s) = \left(\frac{\mu_0}{s+\mu_0}\right)^i.$$

From eqn. (A5), we obtain,

$$Y^*_{00}(s) = \frac{(s+\mu_0)^m}{(s+\lambda)(s+\mu_0)^m \lambda\mu_0^m}.$$

Finally, after some algebriaic manipulations, using eqn. (A6), we obtain

$$\int_0^{\infty}[R(\tau)-\Lambda^2]d\tau = \left(\frac{m+1}{2m}\right)\left[\frac{\lambda\mu}{(\lambda+\mu)^3}\right]$$

Case 2 : On period is hyper-exponentially distributed(Fig. A1- (b)):

The on period is hyper-exponentially distributed: when the source ends an idle period, ONi state is chosen with probability σ_i. The length of the ONi state is exponentially distributed with mean μ_0^{-1}. We have

$$Q = \begin{pmatrix} -\lambda & \sigma_1\lambda & \sigma_2\lambda & \cdots & \sigma_m\lambda \\ \mu_1 & -\mu_1 & 0 & \cdots & 0 \\ \mu_2 & 0 & -\mu_2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \mu_m & 0 & 0 & \cdots & -\mu_m \end{pmatrix},$$

$$\pi_0 = \frac{1/\lambda}{1/\lambda + \sum_{i=1}^{m}(\sigma_i/\mu_i)},$$

$$\pi_i = \frac{\sigma_i/\mu_i}{1/\lambda + \sum_{i=1}^{m}(\sigma_i/\mu_i)}, \quad (i = 1, \ldots, m),$$

$$F_i^*(s) = \frac{\mu_i}{s+\mu_i},$$

$$Y_\infty^*(s) = \left[(s+\lambda) - \lambda\sum_{i=1}^{m}\frac{\sigma_i\mu_i}{s+\mu_i}\right]^{-1}.$$

We, therefore, have,

$$\int_0^\infty [R(\tau) - \Lambda^2]d\tau = \frac{\sum_{i=1}^{m}(\sigma_i/\mu_i^2)}{\lambda^2[1/\lambda + \sum_{i=1}^{m}(\sigma_i/\mu_i)]^3}.$$

since $\mu = [\sum_{i=1}^{m}(\sigma_i/\mu_i)]^{-1}$, we finally have

$$\int_0^\infty [R(\tau) - \Lambda^2]d\tau = \left(\sum_{i=1}^{m}\frac{\sigma_i}{\mu_i^2}\right)\left[\frac{\lambda\mu}{(\lambda+\mu)^3}\right].$$

Acknowledgement

This research has been supported in part by a grant from the Korea Research Foundation and in part by a grant from ITRC of Canada.

References

[1] M. Schwartz *Broadband Integrated Networks*, Prentice-Hall, 1996.

[2] D. P. Heyman, A. Tabatabai and T. V. Lakshman, "Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 2, No. 1, pp. 49-58, March 1992.

[3] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson and J. D. Robbins, "Performance Models of Statistical Multiplexing in Packet Video Communications," *IEEE Trans. on Commun.*, Vol. 36, No. 7, pp. 834-844, July 1988.

[4] P. Sen, B. Maglaris, N.-E. Rikli and D. Anastassiou, "Models for Packet Switching of Variable-Bit-Rate Video Sources," *IEEE Journal on Selected Areas in Commun.*, Vol. 7, No. 5, pp. 865-869, June 1989.

[5] B. Steyaert and H. Bruneel, "An Effective Algorithm to Calculate the Distribution of the Buffer Contents and the Packet Delay in Multiplexer with Bursty Sources," *Globecom*, pp. 417-475, 1991.

[6] J. N. Daigle and J. D. Langford, "Models for Analysis of Packet Voice Communications Systems," *IEEE Journal on Selected Areas in Commun.*, Vol. 4, No. 6, pp. 847-855, September 1986.

[7] A. Baiocchi, N. B. Melazzi, M. Listanti, A. Roveri and R. Winkler, "Loss Performance Analysis of an ATM Multiplexer Loaded with High-Speed ON-OFF Sources," *IEEE Journal on Selected Areas in Commun.*, Vol. 9, No. 3, pp. 388-393, April 1991.

[8] D. Anick, D. Mitra and M. M. Sondhi, "Stochastic Theory of a Data-Handling System with Multiple Sources," *The Bell Ssystem Tech.*, Vol. 61, No. 8, pp. 1871-1894, October 1982.

[9] L. Kosten, "Liquid Models for a Type of Information Buffer Problems," *Delft Progress Report*, No. 11, pp. 71-86, 1986.

[10] T. E. Stern and A. Elwalid, "Analysis of Separable Markov-Modulated Rate Models for Information-Handling Systems," *Adv. Appl. Prob.*, No. 23, pp. 105-139, 1991.

[11] H. W. Lee and J. W. Mark, "ATM Network Traffic Characterization Using Two Types of On-Off Sources," *IEEE INFOCOM*, March 1993.

[12] H. W. Lee and J. W. Mark, "Capacity Allocation In Statistical Multiplexing of ATM Sources," *IEEE/ACM Trans. on Networking*, Vol. 3, No. 2, pp. 126-138, April 1995.

[13] A. Simonian and J. Virtamo, "Transient and Stationary Distribution for Fluid Queue and Input Processes with a Density," *SIAM J. Appl. Math*, Vol. 51, No. 6, pp. 1732-1739, Dec. 1991.

[14] A. A. Borovkov, *Stochastic Process in Queueing Theory, Springer-Verlag*, Berlin, New York, 1976.

[15] H. Kobayashi and Q. Ren, "A Diffusion Approximation Analysis of an ATM Statistical Multiplexer with Multiple Types of Traffic, Part I: Equillibrium State Solutions," *ICC*, 1993.

[16] N. G. De Bruijn, *Asymptotic Methods in Analysis, Dover Publications*, New York, 1981.

[17] W. H. Beyer, ed. *CRC Standard Mathematical Tables, 28th edition*, Boca Raton, Florida, 1987.

Hyong Woo Lee received B.A.Sc, Electrical Engineering, University of British Columbia: Ph.D., Electrical Engineering, University of Waterloo. 1983-1988, Assistant Professor at University of British Columbia. 1988-1991, Assistant Professor at Carleton University. 1992-1995, Assistant Professor at University of Waterloo. Since 1995, Associate Professor at Korea University.

Jon Wei Mark received B.A.Sc, M.A.Sc, Electrical Engineering, University of Toronto: Ph.D., Electrical Engineering McMaster University. Since 1972, Professor at University of Waterloo.