Minimal ON/OFF Source Models for ATM Traffic

F. Bonomi\textsuperscript{a}, J. Meyer\textsuperscript{b}, S. Montagna\textsuperscript{c} and R. Paglino\textsuperscript{c}

\textsuperscript{a}AT&T Bell Laboratories, 200 Schulz Drive, Red Bank, NJ 07701, USA
\textsuperscript{b}University of Michigan, Ann Arbor, MI 48109, USA
\textsuperscript{c}Italtel, Central R&D Labs., Castelletto di Settimo Milanese, 20019 Milan, Italy

Recent studies have cast serious doubt on the robustness of the simple ON/OFF source models with geometric (or exponential, in the case of continuous time models) ON and OFF periods, so often encountered in the literature on ATM systems. In this paper we focus our attention on more refined but still computationally tractable source models, where the ON state is the mixture of $M$ geometric distributions. We analyze a discrete-time homogeneous model for a single queue with ON/OFF traffic described by such models, providing a simple approximation for the output traffic which allows a straightforward analysis of basic multistage systems. Our conclusions indicate that a a reasonable improvement over the well known 2-state Markov source model is provided by the next simplest Markov model, wherein a single state represents the OFF condition and two states represent the ON condition of the source.

1. INTRODUCTION

The analysis of queueing models for Asynchronous Transfer Mode (ATM) systems requires the identification of suitable mathematical descriptions for bursty sources. The most popular [1,2,3, and references thereof] is also the least complex, namely, a 2-state Markov model with one state representing the ON (or active) condition and one state representing the OFF (or idle) condition. The ON and OFF periods are geometrically distributed, in the case of discrete-time models, or exponentially distributed, in the case of continuous-time models. The introduction of these models demonstrated that the first moments of the ON and OFF period distributions have a considerable impact on the performance of the ATM systems. They also played a key role in the sizing and understanding of the early ATM systems.

However, a set of recent studies [3,4,5,6,7,8] has indicated the need for a re-evaluation of the robustness and validity of the 2-state bursty source model, for the following key reasons:

1. It has been shown ([5,6] among others) that the performance of ATM queueing systems with ON/OFF sources is extremely sensitive to both the first and the second moment of the distribution of the ON periods of the source activity, while it is less sensitive to the second moment of the OFF periods. These observations seem to cast serious doubt on the validity of the simple ON/OFF source models with geometric (or exponential, in the case of continuous-time models) ON and OFF periods.
2. Other studies, particularly focused on models for variable bit rate video traffic [7], have found that more sophisticated models are necessary in order to adequately capture the nature of that traffic. Also, models of interactive computer usage may require a more complex distribution of the idle periods than the geometric distribution (e.g., a bi-modal distribution might be required).

3. Furthermore, it has been shown [3,4,9,10] that the output streams from typical ATM multiplexing or switching elements cannot be adequately modeled by an ON/OFF source with geometric (or exponential) ON and OFF periods. Since these are input streams of subsequent multiplexing and switching elements within a multistage ATM network, they too should not be modeled in this manner.

In our study we concentrate on a model which is slightly more sophisticated than the popular 2-state model for ON/OFF sources with geometrically distributed periods. The model, in its general form, is a discrete-time, time-homogeneous Markov model for a single queue with \( N \) homogeneous ON/OFF traffic sources, where each source is modeled by an \((M+1)\)-state Markov model, as shown in Figure 1a. While the OFF condition is still represented by a single state and, consequently, the OFF period is geometrically distributed, the ON period becomes a mixture of \( M \) independent geometric distributions. Our somewhat unconventional analysis of such models is quite efficient, being based on the determination of the busy and idle period distributions for the considered queueing model.

The key contributions of this investigation are results concerning the basic properties of the special 3-state model \((M=2)\) given in Figure 1b. We propose this model as a reasonably general model for the description of ATM bursty traffic traffic, since it overcomes the limitations of the 2-state ON/OFF model cited above, while remaining sufficiently simple and manageable. Specifically, we discuss methods for determining values of the parameters for such a model. We also introduce a "back of the envelope" approximate method for determining the characteristics of the output traffic based on our queueing model of a multiplexing/switching stage. Finally, we discuss the application of this results to the study of a simple multiple stage ATM system.

![Figure 1a. General (M-state) source model](image)

![Figure 1b. 3-state source model](image)
The structure of this paper is the following. Section 2 describes the mathematical models and solution approaches we employ. Section 3 develops the output traffic characterization for the queueing model considered and also presents a simple approximation of key parameters describing such output traffic. Numerical results and applications are then discussed in Section 4, followed by a final section that summarizes the conclusions.

2. THE MATHEMATICAL MODEL AND ITS ANALYSIS

The discrete-time model considered presumes the system illustrated in Figure 2. It includes \( N \) homogeneous ON/OFF sources, a finite buffer of size \( K \), and an output link of capacity one (i.e., at most one cell is transmitted per unit time, or time slot). The access to the buffer follows a FCFS discipline. Each source is modeled by an \((M+1)\)-state Markov chain. State transitions in the source model states occur at the end of a time slot. State 0 describes the OFF condition, whose duration is geometrically distributed, with parameter \( t_{00} \), i.e.,

\[
D_{\text{OFF}}(n) = (1-t_{00})t_{00}^{(n-1)}, \quad n = 0,1,2,\ldots
\]

The mean value and the squared coefficient of variation are given by \( m_{\text{OFF}} = 1/(1-t_{00}) \) and \( CV_{\text{OFF}}^2 = t_{00} \), respectively. The ON condition for the source, during which cells are generated at the same rate of the output link, is represented by the remaining states 1,2,\ldots, \( M \). The distribution of the ON period is thus a mixture of geometric distributions and is given by

\[
D_{\text{ON}}(n) = \sum_{i=1}^{M} \alpha_i (1-t_{ii}) t_{ii}^{(n-1)}, \quad n = 0,1,2,\ldots
\]

Hence, the mean and squared coefficient of variation of the ON period (or burst length) are given by

\[
m_{\text{ON}} = L = \sum_{i=1}^{M} \alpha_i \frac{1}{(1-t_{ii})}.
\]

---

Figure 2. Multiplexer model

\[\text{Figure 2. Multiplexer model}\]

\[\text{D_{OFF}(n) = (1-t_{00})t_{00}^{(n-1)}, \quad n = 0,1,2,\ldots}\]

\[\text{D_{ON}(n) = \sum_{i=1}^{M} \alpha_i (1-t_{ii}) t_{ii}^{(n-1)}, \quad n = 0,1,2,\ldots}\]

\[\text{m_{ON} = L = \sum_{i=1}^{M} \alpha_i \frac{1}{(1-t_{ii})}.}\]
\[ Cv_{ON}^2 = \sum_{i=1}^{M} \alpha_i \frac{(1-2t_{i}) (1-t_{i})}{m_{ON}} - m_{ON}. \]

The choice of a source model with geometric OFF-period distribution is motivated by the results in [5] and [6], which show that the second and higher moments of this distribution do not have a very significant impact on performance (assuming ON/OFF sources). The same studies indicate, however, that the second moments of the ON period distribution do have a significant influence. By choosing a mixture of geometric distributions for the ON periods, we are able to further investigate this phenomenon. The continuous-time version of a special case (M=2) of the source model discussed here is also being studied in [11].

The behavior of the ATM queueing model just described can be studied by using a discrete-time, finite-state Markov chain. The time domain of the chain is defined as the set \( T = \{ 0, 1, 2, \ldots \} \) of non-negative integers, corresponding to beginning instants of possible output link transmissions. A system state \( \theta \in \Theta \) is defined as the ordered pair \( \theta = (i, s) \), \( 0 \leq i \leq K, \ s \in \Sigma \), where
\begin{itemize}
  \item \( i \): total number of cells in the buffer,
  \item \( s \): combined state of the \( N \) sources, represented as an \( M \)-tuple \( s = (n_1, \ldots, n_M) \), with \( 0 \leq n_i \leq N \), for all \( i \),
  \item \( \Sigma \): set of all possible (combined) source states.
\end{itemize}

Let \( \Sigma_j \subseteq \Sigma \) be the subset of source states such that exactly \( j \) sources are in the ON sources. Note that \( \Sigma_0 \) contains a single element, the \( M \)-tuple \( s_0 = (0, 0, 0, \ldots, \theta) \). Let \( A \), a matrix with elements \( A(s_i \mid s_j) \), be the transition matrix of the combined source model, where \( s_j \in \Sigma_j \) and \( s_i \in \Sigma_i \). Also, let \( \pi \), be the vector that gives the steady-state occupancy probabilities \( \pi(\theta) = \pi(i, s) \), and let \( P \), a matrix of elements \( P(\theta \mid \sigma) \), with \( \theta, \sigma \in \Theta \), be the one-step transition probability matrix of the model. The analysis of the model seeks to determine the idle and busy period distributions of the output line. The loss probability can then be obtained without extra effort. If required, the complete steady-state queue length distribution can also be determined via a simple procedure.

Beginning with the busy period, let \( B(m) \) denote its steady-state probability distribution, i.e., \( B(m) \) is the steady-state probability that, after an idle-to-busy transition, the output is busy for exactly \( m \) slots. Further, let \( F_m(\theta) \) denote the probability that, starting from state \( \theta \in \Theta - \theta_0 \) (where \( - \) denotes set difference), the system arrives at the idle state \( \theta_0 = (0, s_0) \) after exactly \( m \) time slots. Once the one-step transition probability matrix \( P \) is determined, \( F_m(\theta) \) can then be obtained using the following recursive formula [12]:

\[ F_m(\theta) = \begin{cases} P(\theta \mid \theta) & \text{for } m = 1, \\ \sum_{\theta' \in \Theta - \theta_0} F_{m-1}(\theta') P(\theta' \mid \theta) & \text{for } m = 2, 3, \ldots. \end{cases} \]

Finally, let \( C(\theta \mid \theta_0) \) be the probability that a busy period begins with a transition from \( \theta_0 \) to \( \theta \in \Theta - \theta_0 \) and let \( C_0 \) be the vector of these probabilities. Then
\[ C(\theta \mid \theta_0) = \frac{P(\theta \mid \theta_0)}{1 - P(\theta_0 \mid \theta_0)}; \]

and, combining the above, we have

\[ B(m) = \sum_{\theta \in \Theta - \theta_0} F_m(\theta) \ C(\theta \mid \theta_0), \quad m = 1, 2, \ldots \]

Observe that the above formulation of the busy period did not require knowledge of the steady-state probability distribution \( \pi \). Such ability to so obtain \( B(m) \) therefore results in a simplification of the overall analysis.

It is also useful to observe that \( B(m) \) can be alternatively expressed using matrix notation, as the product of matrices

\[ B(m) = P_0' \ P_B^m \ C_0. \]

where \( P_0 \) is the vector of elements \( P(\theta \mid \theta_0) \), prime signifies the transpose, and \( P_B \) is the submatrix of the transition probability matrix \( P \) including only the busy states \( \theta \in \Theta - \theta_0 \). Given this formulation for \( B(m) \), it is possible to apply the Perron-Frobenius theorem [13] and conclude that the asymptotic behavior of the busy period distribution \( B(m) \) exhibits a geometric decay, conditioned by the dominating, real eigenvalue of \( P_B \).

Focusing next on the idle period distribution, note that it is geometrically distributed (with mean \( 1/(1-A(s_0 \mid s_0)) \)) as the minimum of \( N \) geometrically distributed random variables with parameter \( t_{00} \). Note also that the only system state which can initiate or sustain an idle period is \( \theta_0 = (0, s_0) \). Thus the distribution of an idle period is geometric with mean

\[ m_I = \frac{1}{1-A(s_0 \mid s_0)} = \frac{1}{1-t_{00}}. \]

Accordingly, when \( B(m) \) and, consequently, its mean \( m_B \) are known, we can immediately determine the utilization of the output line \( \rho_o \) via the equation

\[ \rho_o = \frac{m_B}{m_B + m_I}; \quad (1) \]

whence the steady-state probability \( \pi(\theta_0) \) of an empty queue and all sources OFF

\[ \pi(\theta_0) = 1 - \rho_o. \]

In turn the loss probability \( P_{\text{loss}} \) can be expressed as

\[ P_{\text{loss}} = \frac{\rho - \rho_o}{\rho}, \]

where \( \rho = N r \) is the total load offered by the \( N \) input sources, each offering a load \( r \). Hence, as with \( B(m) \), we have a formulation that does not require the prior solution of the steady-state occupancy probabilities. Once \( \pi(\theta_0) = \pi(0,s_0) \) is known, the whole steady-state distribution can be recovered quickly by a simple procedure, not presented here because of space limitations.
3. OUTPUT TRAFFIC CHARACTERIZATION

In this section we present results concerning the main characteristics of the output traffic of the ATM queueing model of the previous section. It is obvious that this traffic exhibits an alternation of ON and OFF periods corresponding to the busy and idle periods, respectively, of the queueing model. As mentioned in Section 2, the idle period distribution for this model is geometric, with average length \( m_I = 1/(1 - r_{ON}) \), and thus does not require further attention. Our focus, therefore, is the distribution of the busy periods, where prior work in this regard has appeared in a number of earlier papers [3,4,9,14,15,16].

First, let us assess the general shape of the busy period distribution. Two typical examples of such a distribution are presented by the solid curves in Figure 3. Both curves are obtained letting \( N = 8 \), \( K = 500 \), and assuming an offered load of \( \rho = 0.6 \). However, the source models used to compute the two busy period distributions are different. In one case, corresponding to the label \( CV_{ON}^2 = 0.983 \), the source model is a 2-state model (\( M = 1 \), providing a geometric ON-period distribution), with mean \( L = 5 \). In the other case, corresponding to the label \( CV_{ON}^2 = 2 \), the source model is a 3-state model (\( M = 2 \)), yielding a distribution for the ON period with the same mean but a higher squared coefficient of variation. Figure 3 indicates that the busy period distribution is highly affected by the source characterization. Also, it is clear that the shape of the busy period distribution is far from the linear shape of a geometric distribution, even though it manifests a linear asymptotic behavior. In order to stress this point, in Figure 3 we show the dotted line corresponding to a geometric distribution with the same average as the busy period distribution obtained with \( M = 2 \). Note that the non-geometric nature of the busy period distribution of ATM queueing systems has been observed in [3,4,9] and seems to be independent of the characterization of the sources. Also, the deviation from the geometric behavior is less evident at queue utilization levels below 50%. A fundamental consequence of this observation is that the ON/OFF traffic seen over links which are internal to an ATM network is, in general, not well described by a 2-state Markov model, even when the external sources may be faithfully represented by such models. Geometric approximations for the busy period distribution could also be considered. However, as discussed in [4], a simple matching of mean values is inadequate for this purpose, thus necessitating a more complex approach that may be difficult to justify.

![Figure 3. Busy period distribution on inter-stage links](image1)

![Figure 4. Modelling of busy-period distribution on first stage outputs](image2)
A much more flexible and accurate approximation of the output traffic (for the system considered here) can be achieved through its representation as a 3-state model, i.e., the output (if regarded as a source) is modeled as in Figure 1b. In Figure 4 we show the actual busy period distribution, together with distribution obtained with this output model \((M = 2)\), using the approach proposed in [5] to determine the values of the parameters (namely, \(\alpha\), \(t_{11}\), and \(t_{22}\)). Note that, in all instances shown in Figure 4, the distributions based on the 3-state model are in close agreement with those given by the exact analysis of Section 2. In particular, they exhibit the bi-modal behavior of the exact solution, i.e., an early region of non-linear decrease, followed by an asymptotic linear decrease. Observe that the simple 3-state model matches well the third moment of the busy period distribution. In most cases, the asymptotic slope is likewise well matched. As we will show later, slight mismatches in this slope do not lead to relevant errors in the evaluation of the performance measures of interest.

The approach used in determining the parameter values of the 3-state approximation model consists of matching the first moment and the squared coefficient of variation of the exact distribution (e.g., the busy period distribution), that is,

\[
\alpha = \frac{1}{2} \left( 1 + \sqrt{\frac{(C_v^2 - 1)m_B + 1}{(C_v^2 + 1)m_B + 1}} \right); \quad t_{11} = \frac{(m_B - 2\alpha)}{m_B}; \quad t_{22} = \frac{(m_B - 2(1 - \alpha))}{m_B}.
\]

Note that these equations can also be used to characterize a 3-state source model, once the values of the three parameters \(m_{ON}\), \(m_{OFF}\) and \(C_v^2\) are known. For this model, it is also interesting to observe that in this model the larger of values for parameters \(t_{11}\) and \(t_{22}\) could be regarded as an estimate for the dominating eigenvalue of the transition probability matrix among states corresponding to the busy period \(P_B\). This is due to the fact that, from our observations, the exact distribution and that of the approximating 3-state model match well with regard to asymptotic slopes. Note that one of the two ON states is thus used to describe the asymptotic behavior of the busy period distribution.

In summary, to determine the parameter values of the approximate 3-state model we need either exact or approximate values for the first moments of the busy and idle period distributions, along with the value of the second moment (or, equivalently, the squared coefficient of variation) of the busy period distribution. While exact values can be computed using the analysis presented in Section 3, we have recently found a surprisingly simple and yet precise mean of approximating the values of \(m_1\), \(m_B\), and \(C_v^2\). Let the estimates of these be denoted by \(\hat{m}_1\), \(\hat{m}_B\), and \(\hat{C}_v^2\). Then, from above,

\[
\hat{m}_1 = m_1 = \frac{1}{1 - \hat{t}_{00}^{N}}.
\]  

In order to determine the remaining estimates we consider a "reference" infinite queueing model where arrivals are Poisson, but services are distributed according to the discrete distribution of the ON (burst) period of the sources in the original model. Accordingly, in this new model, the number of sources is infinite, and the service time for a customer (number of time slots) is distributed the same as an ON period for a source in the original model. For this reference model, we calculate the first moment and the squared coefficient of variation of the busy period distribution, following a procedure similar to the continuous time solution proposed in [17]. Where needed, \(z\)-transforms are used, instead of Laplace
transforms. The expressions for the first moment and for the coefficient of variation for the reference model under offered load $\rho'$ are given by

$$m = \frac{m_{ON}}{(1 - \rho')}.$$  \hspace{1cm} (3)

$$Cv^2 = \frac{(Cv_{ON}^2 + 1)}{(1 - \rho')} - 1.$$ \hspace{1cm} (4)

Using the above formulas, we then obtain the estimates

$$\hat{m}_B = \hat{m}_I - \frac{\rho}{1 - \rho},$$ \hspace{1cm} (5)

$$\rho' = \frac{\hat{m}_B - m_{ON}}{\hat{m}_B},$$ \hspace{1cm} (6)

$$Cv_{B}^2 = \frac{(Cv_{ON}^2 + 1)}{(1 - \rho')} - 1.$$ \hspace{1cm} (7)

Equations (2),(5)-(7) thus constitute our "back of the envelope" approximation of the output traffic. Regarding its derivation, note that (5) is obtained from (1), neglecting for this purpose the (small) difference between $\rho$, the offered load, and $\rho_o$, the load on the output link. Furthermore, (6) derives the "equivalent load" to be used in the reference model via (3), using the knowledge of $m_{ON}$ and the estimate $\hat{m}_B$ we just obtained in (5). Finally, (7) comes from (4) by using the "equivalent load" value. Note that, in some sense, this value captures the effect of the finite source model within the infinite reference model. Note also that the only parameters involved in (2),(5)-(7) are the input parameters of the original model. A validation of this approximation is presented in the section that follows.

4. NUMERICAL RESULTS AND APPLICATIONS

We now consider some numerical results concerning sensitivity of performance to the values of certain model parameters, based on the 3-state source model considered above.

With regard to cell loss probability, Figures 5 and 6 display its sensitivity to the squared coefficient of variation of the distribution of the ON periods for the sources. The model parameter values assumed are $N=8$ and $16$, mean burst length $L=5$, buffer size $K=500$, and offered load $\rho$ of 0.6 and 0.8. Observe that in the relevant region of loss probabilities between $10^{-6}$ and $10^{-11}$, the sensitivity to the squared coefficient of variation is striking. Indeed, assuming geometric burst distributions for sources with higher ON period variability leads to an underestimation of the loss probability by many orders of magnitude. Moreover, the influence of the number of input streams $N$ in Figure 5 is not appreciable.

Figure 6 shows the loss probability as a function of the variability of the ON periods, for two different mean values of the ON period distribution, and for two values of the buffer size.
Again, for loss probabilities higher than $10^{-6}$, the curve remains rather flat, while, below this value, loss probabilities become increasingly more sensitive to changes in the squared coefficient of variation.

This seems to indicate, as is intuitive, that an accurate modeling of an input source, including higher moments of the ON distribution, is necessary when the Quality of Service (QoS) requirement (in terms of loss performance for that source) is below $10^{-6}$. On the other hand, if higher loss probabilities can be tolerated then simpler (e.g., 2-state Markov) models might still be appropriate. One can also examine the impact of the third moment of the ON-period distribution on the cell loss probability, as shown in Figure 7. Here, the model parameter values are $N = 4, 8, 16$, $K=500$, $p=0.4$, and mean burst length $L = 15.3$. The squared coefficient of variation of the ON distribution is held constant at the value 2.23. As indicated by the curves, the loss probability is still somewhat sensitive to the third moment of the burst length, although to a lesser degree than observed above for the squared coefficient of variation.

By way of justifying the 3-state output model developed in Section 3, Table 1 compares the exact second moments of the busy period distribution for one queueing stage, assuming $N$ 2-state ON/OFF sources ($M = 1$), with those obtained with the approximate procedure.
proposed above, for various values of $\rho$, $L$, and $N$. The relative errors in the second moments are plotted as a function of $\rho$ in Figure 8. These results provide remarkably favorable testimony to the validity of this output model. In particular, we note that the approximation error is insensitive to the number of sources $N$, decreases with increasing source average burst length $L$, and increases with the offered load $\rho$.

As further evidence of the validity of the approximate output traffic characterization, we consider its application in the analysis of the 2-stage ATM multiplexing/switching configuration shown in Figure 9. Specifically, the first stage consists of four identical first stage multiplexer queues whose output streams are further multiplexed together by a second stage queue. For the sake of simplicity, all links in this model are characterized by the same capacity (one cell/slot). Note that multistage ATM systems are discussed also in [10,15,16]. In a first experiment, shown in Figure 10, we compare the buffer occupancy (queue length distribution) at the second stage as it is obtained by three different methods.

1. Simulation of the complete two stage model.

2. Analytical solution for the the idle/busy distributions of the first stage and then determination of the parameter values of a 3-state ($M=2$) input model to be used in the exact analysis of the second stage.

3. Application of the simple estimation method described by equations (2) and (5)-(7) to determine the the 3-state input models for the second stage, which is then again approximately analyzed using the same method.

These are compared, in turn, with the occupancy distribution that results if the traffic on an interstage link is represented by a 2-state model (i.e., $M = 1$ as opposed to the choice $M = 2$ in the method 2) above). The input models for the first stage are assumed to be 2-state ON/OFF models with mean burst length $L = 5$ and utilization $\rho = 0.025$. On closer examination of Figure 10, we see that the distributions given by methods 1) - 3) agree very closely. In sharp contrast, use of a 2-state model for the interstage link underestimate the queue length distribution by one or two orders of magnitude.

![Figure 9. Two-stages ATM multiplexers network](image)

![Figure 10. Buffer occupancy distribution of second stage mux](image)

Figure 11 displays a similar comparison for input sources that have mean ON period duration $L = 20$. In this case, methods 1)-3) yield results that are essentially identical,
whereas a 2-state interstage model again underestimates occupancy probabilities, particularly for larger queue lengths.

Finally, in Figure 12 we compare the busy period distributions of the second stage obtained by methods corresponding to 2) and 3) above. Again, these two methods yield results that compare very favorably. Of particular significance is how well method 3) (the "back of the envelope approach") approximates the probabilities of long busy periods when the input burstiness is high ($L = 20$).

![Figure 11. Buffer occupancy distribution of second stage mux](image1)

![Figure 12. Busy period distribution on second-stage links](image2)

5. CONCLUSIONS

The steady-state behavior of a discrete ATM queueing model was studied presuming homogeneous ON/OFF sources described by ($M + 1$)-state Markov chains. An initial goal of the investigation was to provide some insight into the sensitivity of certain relevant performance measures to the characterization of the ON-period distribution of the sources.

Our principal conclusion is that simple 2-state Markov models for ON/OFF sources may be seriously inadequate, particularly when the loss probability requirements for the queueing system are rather small (e.g., below $10^{-6}$) and the queuing utilization is relatively high (e.g., above 50%). However, the use of source models capturing higher moments of the ON distribution may not be necessary for lower queuing utilizations and higher loss requirements, but also when the source peak rate is much smaller than that of the output link.

Our study also indicates that the simplest extension of the classic ON/OFF models may provide a much more appropriate source model, namely, a 3-state model with one OFF state and two independent ON states. This model has the following advantages

1. It is only slightly more complex than the 2-state model.
2. It can represent ON-period distributions with coefficients of variation that have a wide range of values.
3. It captures the general behavior of the busy period of typical ATM queueing systems, and can provide a simple approximation of its probability distribution.

Finally, we have proposed and validated a simple "back of the envelope" approximate
characterization for the output traffic of the model considered, and discussed its application to a multistage system.

**Acknowledgements:** We sincerely appreciate the insight gained in useful discussions with Bharat Doshi of AT&T Bell Laboratories.

**APPENDIX**

The calculation of $\pi(i, s)$ is performed by a recursive scheme, described in the following. Note that $\pi(i, s)$, the steady-state distribution of an aperiodic and irreducible Markov Chain, is unique.

1. We first determine the probabilities $\pi(0, s_0)$, $\pi(1, s_0)$, noting that $\Sigma_0$ contains a single element $s_0 = (0,0,0,...,0)$, and $\pi(1, s_1)$, for all $s_1 \in \Sigma_1$.

   First, assume $\pi(0, s_0)$ known. For example, it may be determined as a result of the computation of the busy period distribution, as described in Section 2. However, since the steady-state distribution to be determined is unique, we can also guess a value for $\pi(0, s_0)$, and determine its correct value at the end, via normalization of the distribution. Then, we can determine the remaining probabilities by solving the linear system (for all $s_1 \in \Sigma_1$):

   \[
   \begin{align*}
   \pi(0, s_0) &= [\pi(0, s_0) + \pi(1, s_0)]A(s_0 \mid s_0) + \sum_{s_1 \in \Sigma_1} \pi(1, s_1)A(s_0 \mid s_1), \\
   \pi(1, s_1) &= [\pi(0, s_0) + \pi(1, s_0)]A(s_1 \mid s_0) + \sum_{\overline{s}_1 \in \Sigma_1} \pi(1, \overline{s}_1)A(s_1 \mid \overline{s}_1).
   \end{align*}
   \]

2. Next, we can determine in a similar way the probabilities $\pi(2, s_2)$, for all $s_2 \in \Sigma_2$, and then $\pi(1, s_0)$, and $\pi(2, s_1)$, for all $s_1 \in \Sigma_1$.

3. In general we can determine the probabilities $\pi(i+1, s_0)$, and $\pi(i+1, s_1)$, for all $s_1 \in \Sigma_1$, by solving the following system of linear equations:

   \[
   \begin{align*}
   \pi(i, s_0) &= \sum_{j=2}^{\min(N, i+1)} \sum_{s_j \in \Sigma_j} \pi(i+1, s_j)A(s_0 \mid s_j) + \sum_{s_1 \in \Sigma_1} \pi(i+1, s_1)A(s_0 \mid s_1), \\
   \pi(i+1, s) &= \sum_{j=2}^{\min(N, i+1)} \sum_{s_j \in \Sigma_j} \pi(i+1, s_j)A(s \mid s_j) + \sum_{s_1 \in \Sigma_1} \pi(i+1, s_1)A(s \mid s_1) + \\
   &\quad + \pi(i+1, s_0)A(s \mid s_0) \quad \text{for all } s \in \Sigma_1.
   \end{align*}
   \]

4. The remaining unknown probabilities, $\pi(K, s_j)$, for all $s_j \in \Sigma_j$, are determined by the final linear equations system (for all $s_j \in \Sigma_j$):

   \[
   \pi(K-1, s_0) = \sum_{j=0}^{\min(N, K-1)} \sum_{s_j \in \Sigma_j} \pi(K, s_j)A(s_0 \mid s_j),
   \]
\[ \pi(K, s_1) = \sum_{j=0}^{\min(N, K)} \sum_{s_j \in \Sigma_j} \pi(K, s_j) A(s_1 | s_j). \]
\[ \pi(K, s_2) = \sum_{j=0}^{\min(N, K)} \sum_{s_j \in \Sigma_j} \pi(K, s_j) A(s_2 | s_j) + \sum_{j=0}^{\min(N, K-1)} \sum_{s_j \in \Sigma_j} \pi(K-1, s_j) A(s_2 | s_j). \]
\[ \pi(K, s_N) = \sum_{j=0}^{\min(N, K-N+1)} \sum_{s_j \in \Sigma_j} \pi(K, s_j) A(s_N | s_j) + \sum_{j=0}^{\min(N, K-N+1)} \sum_{s_j \in \Sigma_j} \pi(K-N+1, s_j) A(s_N | s_j). \]

REFERENCES


Table 1
Error in the approximation of the second-order moment of the busy period distribution

<table>
<thead>
<tr>
<th></th>
<th>L=5</th>
<th>L=50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact value</td>
<td>Approx.</td>
</tr>
<tr>
<td>N=4</td>
<td>( \rho )</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>7.76E+01</td>
<td>7.89E+01</td>
</tr>
<tr>
<td>0.4</td>
<td>1.60E+02</td>
<td>1.68E+02</td>
</tr>
<tr>
<td>0.7</td>
<td>1.01E+03</td>
<td>1.13E+03</td>
</tr>
<tr>
<td>N=8</td>
<td>( \rho )</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>8.43E+01</td>
<td>8.58E+01</td>
</tr>
<tr>
<td>0.4</td>
<td>1.90E+02</td>
<td>1.99E+02</td>
</tr>
<tr>
<td>0.7</td>
<td>1.38E+03</td>
<td>1.55E+03</td>
</tr>
<tr>
<td>N=16</td>
<td>( \rho )</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>8.77E+01</td>
<td>8.95E+01</td>
</tr>
<tr>
<td>0.4</td>
<td>2.06E+02</td>
<td>2.16E+02</td>
</tr>
<tr>
<td>0.7</td>
<td>1.59E+03</td>
<td>1.79E+03</td>
</tr>
</tbody>
</table>

Table 2
Error in the third order moment of the busy period distribution using a 3-state model

<table>
<thead>
<tr>
<th></th>
<th>Exact value</th>
<th>Approx.</th>
<th>Rel. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=8, L=50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r )</td>
<td>Exact value</td>
<td>Approx.</td>
<td>Rel. Err.</td>
</tr>
<tr>
<td>0.2</td>
<td>2.33E+06</td>
<td>2.33E+06</td>
<td>3.68E-03</td>
</tr>
<tr>
<td>0.5</td>
<td>1.48E+07</td>
<td>1.39E+07</td>
<td>5.99E-02</td>
</tr>
<tr>
<td>0.8</td>
<td>2.03E+07</td>
<td>3.46E+07</td>
<td>7.07E-01</td>
</tr>
</tbody>
</table>