A Model for the Performance Analysis of Voice/Data ATM Multiplexers

Shiouming Stanley Wang and John A. Silvester

CENG Technical Report 92-07

Department of Electrical Engineering - Systems
University of Southern California
Los Angeles, California 90089-2562
(213)740-4579

A Model for the Performance Analysis of Voice/Data ATM Multiplexers

Shiouming Stanley Wang and John A. Silvester

Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, California 90089

Abstract

In this paper, cell arrivals from integrated voice and data sources are modeled as a two-state *Markov Modulated Poisson Process* with batch arrivals ($MMPP^{[X]}$). A voice/data ATM multiplexer is modeled as an $MMPP^{[X]}/D/1$ queue. Inaccuracy and high computation overhead of existing similar approaches are overcome by introducing a new set of parameters for the MMPP. Simulation is used to verify the accuracy of the approximation. Comparisons to a recently proposed method by other authors are included in the numerical results.

1 Introduction

Numerous studies that characterize packetized voice source, e.g., [1], [3], [4], [7], [8], [10], [13], [20], [21] and [22], have been made. Among these studies, Brady's *ON-OFF process* [4] has been widely adopted as the model for the arrival process corresponding to a single voice source. In the ON-OFF process model, the voice source alternates between exponentially distributed (or geometrically distributed if we use a discrete time scale) *ON* periods and exponentially distributed *OFF* periods. The *ON* periods correspond to talkspurts while the *OFF* periods represent silence durations. Packets are created with a constant interarrival time during the *ON* periods and no packets are generated during the OFF periods. The transition rates (or *transition probabilities* for the discrete time case), from ON-to-OFF (α) and from OFF-to-ON (β), are determined by the expected length of the talkspurts and the expected length of the silence periods (see Fig. 1).

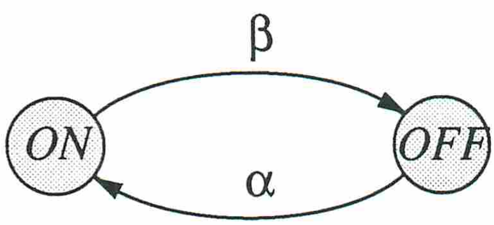

Fig. 1. An ON-OFF process.

The superposition of N such ON-OFF processes forms a *finite-state birth-death process* with the states representing the number of voice sources in talkspurt (in the ON-state) as shown in Fig. 2 [21].

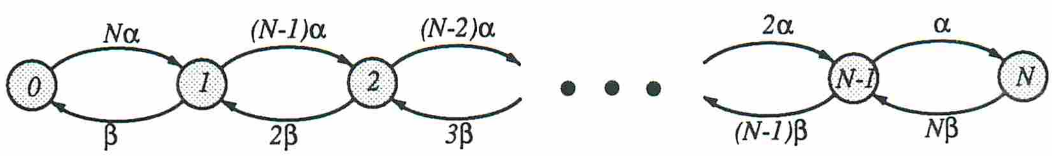

Fig. 2. The birth-death process for the number of voice sources in talkspurt.

The difficulty in generating useful analytical results from a queueing system based on this the birth-death process to determine the arrivals has forced researchers to look for approximations. Two main approaches have been proposed. The first approach uses a fluid flow approximation [5], [14], [21], which somewhat overlooks the randomness (or, in other words, the short-term variation) of the arrivals. The second approach carefully matches the parameters of the process to a simpler one, e.g., the *two-state Markov Modulated Poisson Process* (MMPP) used in [1], [7] and [14], and the *renewal process* used in [14]. The later approach suffers from its inaccuracy and high computational overhead.

In this paper, cells from superposed voice sources are approximated by means of a two-state MMPP as in [1] but the procedure for parameter matching is different. Data packets are assumed to from a Poisson process, i.e., data cells form a Poisson process with batch arrivals. The arrivals from the integrated data and voice sources are then modeled as a two-state MMPP with batch arrivals (denoted $MMPP^{[X]}$ in this paper). The new matching procedure results in better accuracy and much lower computational overhead then existing techniques.

The paper is organized as follows. First, the approximation technique is discussed in section 2. Numerical results and some discussions are presented in section 3. Finally, some conclusions are drawn in section 4.

2 The Approximation Technique

In this section, we present the model using $MMPP^{[X]}$ as an approximation for integrated voice and data traffic. Some similar models are discussed first.

2.1 Using two-state MMPP as an approximation

Heffes and Lucantoni, in their work [7], use the following two-state MMPP statistical characteristics to match with that of the superimposed process:

i. the expected arrival rate;

ii. the variance-to-mean ratio of the number of arrivals in some time interval;

iii. the long-term variance-to-mean ratio of the number of arrivals;

iv. the third moment of the number of arrivals in some time interval.

Upon observing the strong role of the overload† period in determining the performance of the multiplexer, Nagarajan *et al.*, in their matching procedure [14], replace the last step in [7] by "*the variance of the number of arrivals in some time interval giving that the system is in overload state,*" and display an improvement on predicting packet loss in a finite-buffered system. In [1], Baiocchi *et al.* attack the problem from a different angle. They use a theorem proved in [15] (theorem 2.3.1, p. 62) and present an "asymptotic matching procedure" which leads to more accurate results than those of [7] and [21] (see the reference for details).

MMPP's have been used in the research mentioned above to approximate aggregation of several ON-OFF processes. We extend this model to specifically include both voice and data sources in the same model. We then propose a new way to match parameters that improves the accuracy of the approximation for the performance prediction of an ATM multiplexer.

† Defined to be when the number of active sources exceeds the capacity of the system.

2.2 The new approach

Consider N independent voice sources modeled by ON-OFF processes with the following parameters:

i. Γ, the constant arrival rate in the ON-state;

ii. α, the transition rate from the OFF-state to the ON-state;

iii. β, the transition rate from the ON-state to the OFF-state.

Let C denote the capacity of the multiplexer; $M = \lfloor C/\Gamma \rfloor$ be the maximum number of active voice sources which the multiplexer can support; π_k defined by:

$$\pi_k = \binom{N}{k} \left(\frac{\alpha}{\alpha+\beta}\right)^k \left(\frac{\beta}{\alpha+\beta}\right)^{N-k}, \quad 0 \le k \le N \tag{1}$$

be the steady state probability that k out of N voice source are in the ON-state; and γ_1, λ_1, γ_2, λ_2 be the four parameters for the two-state MMPP (see appendix for details). The states of the arrival process are divided into two disjoint subsets: those states in overload situation, $\{M+1, M+2, \ldots, N\}$, and the others, $\{0, 1, 2, \ldots, M\}$. These two sets of states are mapped into the state-II (the overload-state) and the state-I (the underload-state) of a two-state MMPP respectively (see Fig. 3).

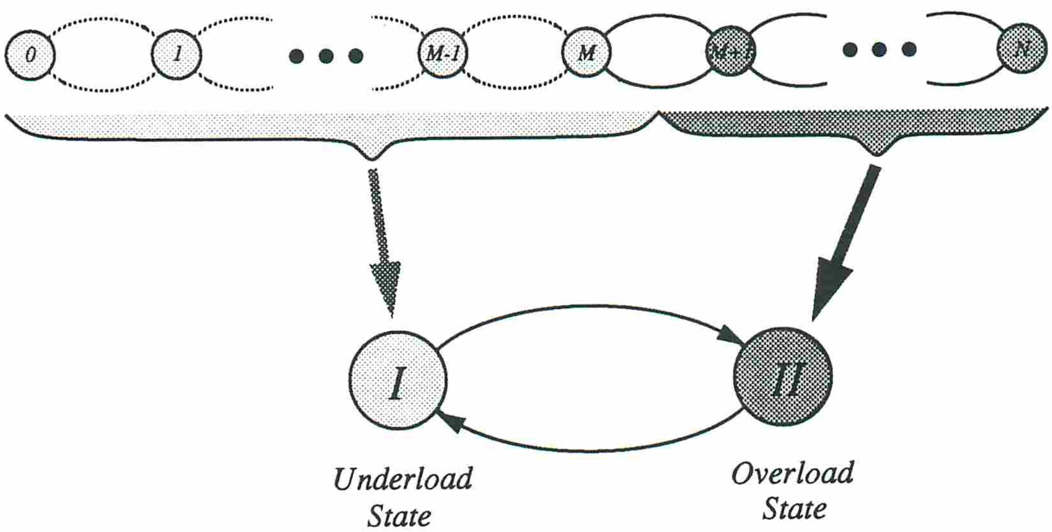

Fig. 3. State mapping between the arrival process and the two-state MMPP for integrated voice and data traffic.

We then have the following recurrence relation for T_i, the expected time until the process visits state i-1 for the first time starting from state i:

$$T_i = \frac{1}{(N-i)\alpha + i\beta} + \frac{(N-i)\alpha}{(N-i)\alpha + i\beta}(T_{i+1} + T_i), \quad 1 \le i \le N-1 \tag{2}$$

with the boundary condition, $T_N = 1/(N\beta)$. Note that, in (2), the first term on the right-hand side gives the expected sojourn time in state i. The first part of the second term on the right-hand side specifies the probability that the process will make a transition to state $i+1$ given that it is currently in state i. It can be easily proved by successive substitution, for example, that T_{N-i} satisfies the following equation:

$$T_{N-i} = \sum_{k=0}^{i} \frac{\frac{i!}{k!}\alpha^{i-k}}{\frac{(N-k)!}{(N-i-1)!}\beta^{i+1-k}}$$

Thus, T_{N-i} has the following explicit closed-form expression:

$$T_{N-i} = \frac{1}{N\beta\binom{N-1}{i}} \sum_{k=0}^{i} \binom{N}{k} \left(\frac{\alpha}{\beta}\right)^{i-k}, \quad 0 \le i \le N-1 \tag{3}$$

Using (3), we can find the expected time until the system load drops below the system capacity, T_{M+1}, once the load exceeds the system limit, M. We then propose the following approach to match the parameters (Note that, *ii-iv* are basically the same as those used in [1]):

i. $$\gamma_2 = \frac{1}{T_{M+1}} = \left(\frac{1}{N\beta\binom{N-1}{N-M-1}} \sum_{k=0}^{N-M-1} \binom{N}{k} \left(\frac{\alpha}{\beta}\right)^{N-M-1-k}\right)^{-1}$$

ii. $$\lambda_1 = \Gamma \sum_{k=0}^{M} k\left(\frac{\pi_k}{\Pi_u}\right), \text{ where } \Pi_u = \sum_{k=0}^{M} \pi_k$$

iii. $$\lambda_2 = \Gamma \sum_{k=M+1}^{N} k\left(\frac{\pi_k}{\Pi_o}\right), \text{ where } \Pi_o = \sum_{k=M+1}^{N} \pi_k$$

iv. $$\gamma_1 = \gamma_2 \frac{\Gamma\phi - \lambda_1}{\lambda_2 - \Gamma\phi}, \text{ where } \phi = N\left(\frac{\alpha}{\alpha+\beta}\right) \text{ is the expected number of active calls.}$$

As one might expect, the arrival stream from the superposed process is burstier than that from the corresponding MMPP especially when the MMPP is in the overload-state, which implies a relatively larger number of voice sources in talkspurt. This can be visualized by observing the "smoothing" effect introduced by the Poisson processes, since when the MMPP is in any particular state the arrival process is just an ordinary Poisson process. Otherwise, each active source would emit cells to the system at its peak rate. Meanwhile, an under-estimation on the burstiness of the arrival process of a queueing system will convert to an under-estimation on the average system time. Since on the average, a longer queue is expected by each new arrival for a burstier arrival process due to a shorter interarrival time within the same burst. Hence, the approximation is expected to perform worse and worse as the system load increases, i.e, the number of active sources increases. The same problem will also exist for the asymptotic matching procedure proposed in [1].

2.3 Improve the accuracy of the approximation

In order to improve the accuracy of our model, let us observe that if we over-estimate T_{M+1}, we may, in fact, improve the model's accuracy. This is because the arrival rate in the overload-state is higher than that in the underload-state (see step *ii* and *iii* of the matching procedure). Also, increasing the average overload time causes a decrease on the average underload time (see step *iv* of the matching procedure). Therefore, overestimating the average overload time, T_{M+1}, will increase the burstiness of the MMPP. Note also that (for the same number of voice calls) the number of states which are mapped into the overload-state of the MMPP increases, i.e., M decreases, as the channel capacity decreases. This corresponds to the MMPP staying a longer time in the overload-state on the average. And, the longer the average overload time is, the severer the under-estimation on the burstiness is. Thus, we need a higher over-estimation for T_{M+1} to offset the increasingly severe under-estimation of the burstiness.

One possible way to over-estimate T_{M+1} is to replace (3) by its upper-bound. After several trials, we found that the following upper-bound for (3) leads to very good results:

$$\frac{i+1}{N\beta\binom{N-1}{i}} \max_{0 \le k \le i} \binom{N}{k} \left(\frac{\alpha}{\beta}\right)^{i-k} \tag{4}$$

By applying *Stirling's formula* [6] to (4), we have:

$$\frac{i+1}{N\beta\binom{N-1}{i}} \max_{0\le k\le i} O(\sqrt{N})\left(\frac{N^N}{k^k(N-k)^{N-k}}\right)\left(\frac{\alpha}{\beta}\right)^{i-k}$$

$$= \frac{i+1}{N\beta\binom{N-1}{i}} \max_{0\le k\le i} O(\sqrt{N})e^{N\log N-(N-k)\log(N-k)-k\log k+(i-k)\log(\frac{\alpha}{\beta})}$$

$$= \frac{i+1}{N\beta\binom{N-1}{i}} \max_{0\le k\le i} O(\sqrt{N})e^{N(-\frac{N-k}{N}\log\frac{N-k}{N}-\frac{k}{N}\log\frac{k}{N})+(i-k)\log(\frac{\alpha}{\beta})} \tag{5}$$

The problem now reduces to determining which value of k provides the largest exponent in (5). So, let us replace k/N by x and rewrite the exponent of (5) as:

$$N\left[(x-1)\log(1-x)-x\log x+\left(\frac{i}{N}-x\right)\log\left(\frac{\alpha}{\beta}\right)\right] \tag{6}$$

After taking the derivative of (6) and equating to zero, we get:

$$\log(1-x)-\log x-\log\left(\frac{\alpha}{\beta}\right) = 0 \tag{7}$$

From (7), it can be easily verified that (6) is maximized at $k = \frac{N\beta}{\alpha+\beta}$. Thus, we conclude that if

$$i < \frac{N\beta}{\alpha+\beta} \tag{8}$$

$k = i$ should be chosen to find the upper-bound for (3); otherwise, $k = \left\lfloor\frac{N\beta}{\alpha+\beta}\right\rfloor$ should be used.

Theorem 1. For any stable system, i.e., utilization $\rho = (N\alpha\Gamma)/[C(\alpha+\beta)] < 1$, $N-M-1 < N[\beta/(\alpha+\beta)]$ always holds.

Proof. Assume the theorem is not true, i.e., for some α, β, and N:

$$N-M-1 \geq \frac{N\beta}{\alpha+\beta}$$

$$\Rightarrow M+1 \leq N - \frac{N\beta}{\alpha+\beta}$$

$$\Rightarrow \frac{C}{\Gamma} < M+1 \leq N - \left(1-\frac{\alpha}{\alpha+\beta}\right)N$$

$$\Rightarrow \frac{N\alpha\Gamma}{(\alpha+\beta)C} > \frac{\alpha+\beta}{\alpha}\left[1-\left(1-\frac{\alpha}{\alpha+\beta}\right)\right]$$

$$\Rightarrow \rho > 1$$

which contradicts the assumption. Therefore, the theorem is always true. □

Theorem 1 implies that for the systems of interest, (8) always holds for $i = N-M-1$. The upper-bound that we found for $T_{N-(N-M-1)}$, denoted $\tilde{T}_{M+1}$, is then given by:

$$\tilde{T}_{M+1} = \frac{N-M}{N\beta\binom{N-1}{N-M-1}}\binom{N}{N-M-1} = \frac{N-M}{(M+1)\beta} \tag{9}$$

Note that, in the extreme case, i.e., when $M+1 = N$, $\tilde{T}_N$ is equal to T_N. Furthermore, for the same number of voice calls, as the value of M+1 decreases, i.e., the channel capacity decreases, the value of $\tilde{T}_{M+1} - T_{M+1}$ increases. This is precisely how we want $\tilde{T}_{M+1}$ to behave, as pointed out before.

2.4 Including data traffic to the model

We can now refine our matching procedure using the result of (9) as the following:

i. $\gamma_2 = \dfrac{1}{\tilde{T}_{M+1}} = \dfrac{(M+1)\beta}{N-M}$;

ii-iv. (*as before*).

We now have a two-state MMPP representing the arrival process from a number of voice sources. To include data sources into the model, let us assume that data sources generate data packets according to a Poisson process with rate μ. The packet length has an arbitrary probability mass function P_k, the probability that a data packet has a size of k cells. If we combine the data sources with the MMPP, we have a two-state MMPP with batch arrivals. The transition rates, γ_1 and γ_2, are determined by the matching procedure. The arrival rates are $\lambda_1 + \mu$ and $\lambda_2 + \mu$ for the underload state and the overload state respectively. The probability that a batch has a size of k cells is:

$$P_k \frac{\mu}{\mu + \lambda_1} + \delta_{k-1} \frac{\lambda_1}{\mu + \lambda_1}, \quad k = 1, 2, 3, \ldots \tag{10}$$

if the process is in the state-I; and:

$$P_k \frac{\mu}{\mu + \lambda_2} + \delta_{k-1} \frac{\lambda_2}{\mu + \lambda_2}, \quad k = 1, 2, 3, \ldots \tag{11}$$

if the process is in the state-II, where:

$$\delta_{k-1} = \begin{cases} 1, & \text{if} \quad k = 1 \\ 0, & \text{otherwise} \end{cases}$$

3 Numerical Results and Discussions

In this section we present some numerical results followed by some discussions.

3.1 Numerical results

We model a voice/data ATM multiplexer as an $\mathrm{MMPP}^{[X]}/D/1$ queue and the procedure described in appendix are applied to get the expected system time. Each voice call is characterized by $\Gamma = 1/6$ cells/msec (assume 64 Kbps PCM coding with speech activity detector and standard 48-octet payload size), $\alpha = 1.538$ and $\beta = 2.778$ (according to the conclusions drawn by [3]). Aggregated data traffic is assumed to have an arrival rate of $20N_d/3$ packets per second and average packet size of 5 cells per packet, where N_d is the number of data calls. Fig. 4 and Fig. 5 assume a fixed number of voice and data calls (20 voice calls plus 20 data calls for Fig. 4 and 200 voice calls plus 200 data calls for Fig. 5), i.e., fixed system load, and plot average system time versus channel utilization. Fig. 6 shows the relationship between average system time and channel capacity with the system utilization kept unchanged at 0.8 and a fixed ratio of voice traffic to data traffic. In these figures, MMPP-1 is the model suggested by [1] extended to include data sources and MMPP-2 is the model presented in this study.

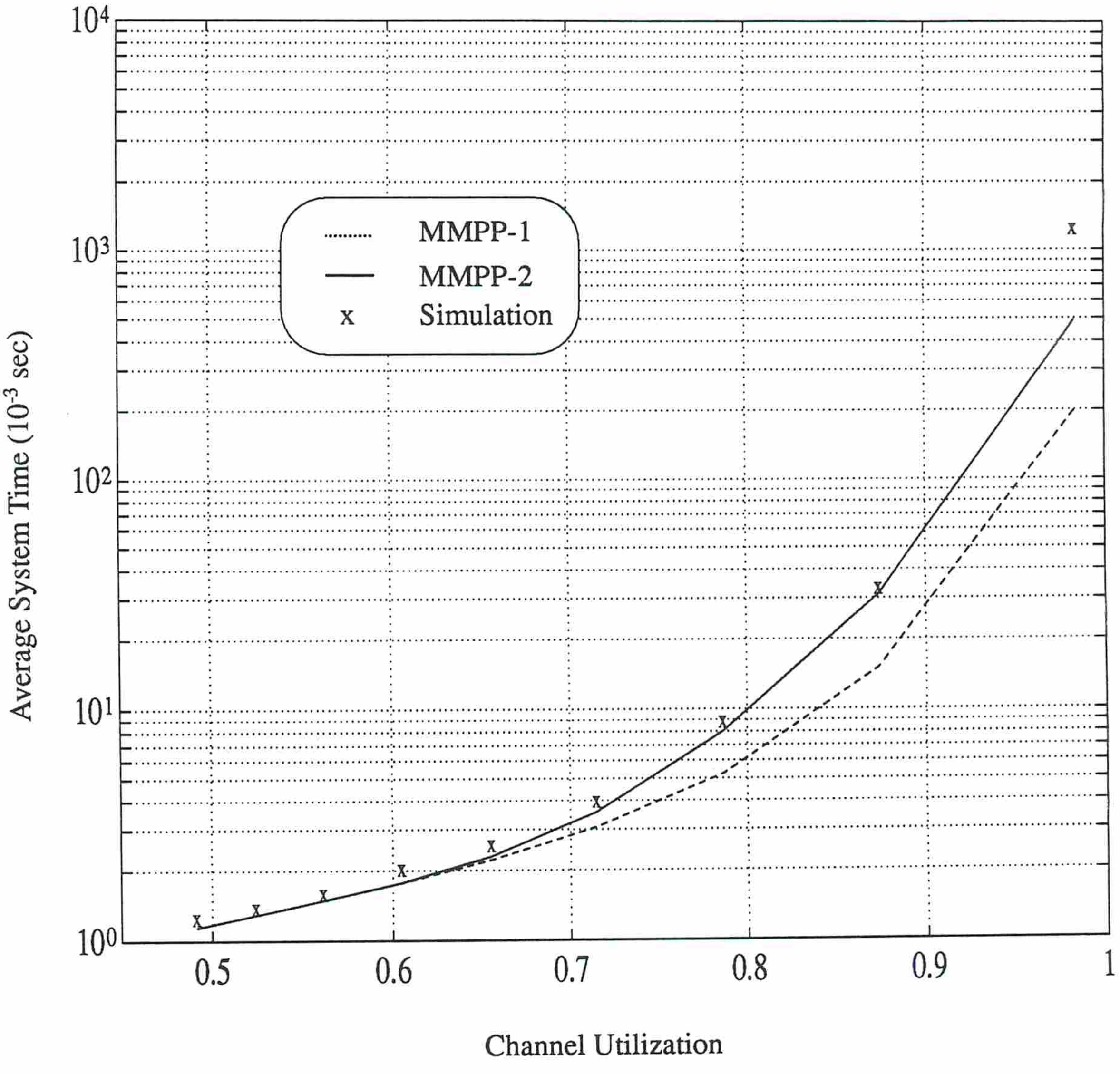

Fig. 4. The expected system time versus channel utilization for 20 voice calls and 20 data calls.

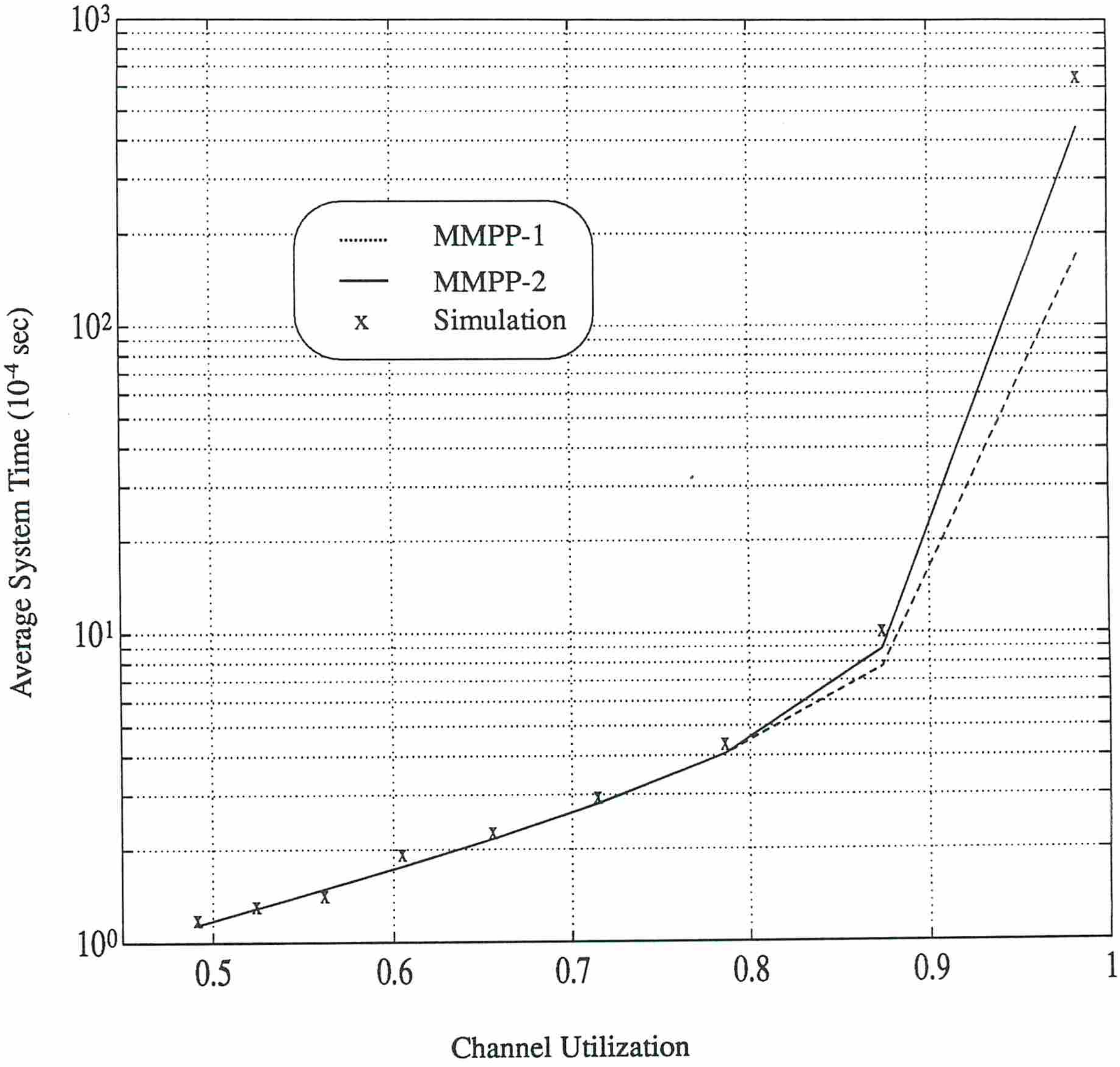

Fig. 5. The expected system time versus channel utilization for 200 voice calls and 200 data calls.

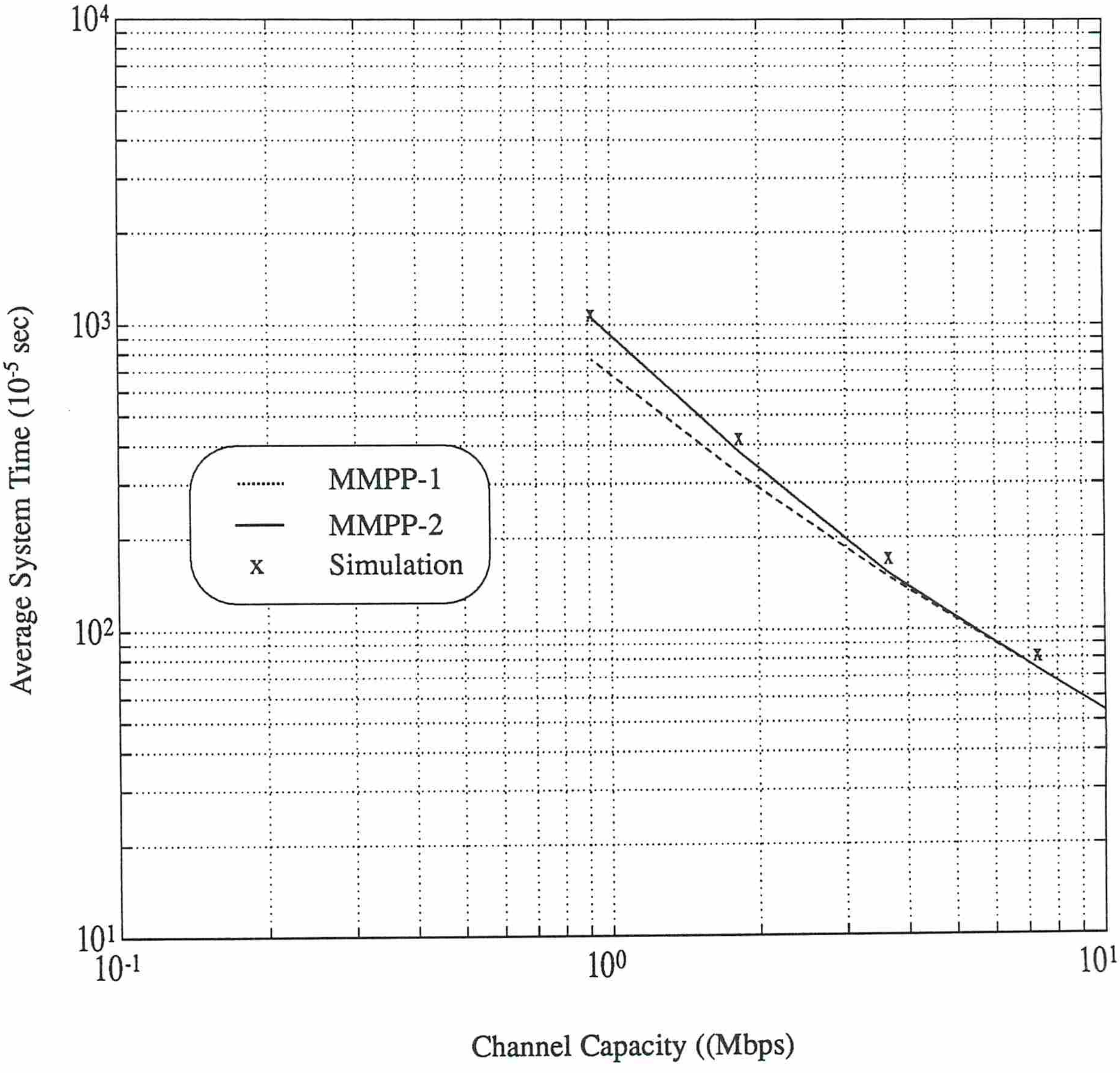

Fig. 6. The expected system time versus channel capacity for $\rho = 0.8$.

3.2 Discussions

Several advantages of our model can be identified. *First, a lower computation overhead is involved.* In [1], solving an (M+1)-dimensional eigen-system is required in order to find the parameters for the MMPP, which can be a significant effort for a large M. Whereas in our model, the computation overhead is negligible. *Second, a real-time traffic control algorithm can be developed.* Note that if the system is in an overload situation, the queue size is expected to build-up until the system returns to an underload status. Hence $\tilde{T}_{M+1}$ can be used as an indicator for the seriousness of the overload, once the system is in the overload-state. Using our results $\tilde{T}_{M+1}$ can be computed very quickly for any given number of voice and data calls and used as a criterion for traffic control. *Third and more importantly, it provides better performance predictions.* As the results indicated by Fig. 4 through Fig. 6, our model constantly out-performs the one proposed in [1] and agrees very well with simulation.

4 Conclusions

In this paper, we have studied the performance of an ATM multiplexer loaded with voice and data traffic. The actual arrival process has been approximated by an two-state Markov modulated Poisson process with batch arrivals; and the multiplexer has been modeled as an $\mathrm{MMPP}^{[X]}/D/1$ queue.

The modeling technique developed here leads to very accurate results and has very low computational overhead; hence it is useful for real-time traffic control.

More studies are necessary to draw conclusions on corresponding loss model (limited buffer space). Some studies on the higher moments of the measurements may be of use in order to understand the impact of highly variable arrival process.

Acknowledgement

The authors wish to express their appreciations to Dr. David M. Lucantoni for his valuable suggestions and assistance on the analytical results for two-state MMPP with batch arrivals.

Appendix

This appendix summarizes some analytical results on two-state MMPP and two-state $\mathrm{MMPP}^{[X]}/D/1$ queue. We begin with the definition of MMPP followed by a procedure to calculate the expected waiting time. A detailed analysis and more general cases can be found in [11] and [12].

An MMPP, a special case of the *Neuts' versatile Markovian process* (the *N-process*†), is a *doubly stochastic Poisson process* where the arrival rate is determined by the state (called the *phase*) of the underlying finite-state Markov chain. For example, a two-state MMPP has the following four parameters:

i. the transition rate from phase one to phase two, γ_1;

ii. the transition rate from phase two to phase one, γ_2;

iii. the arrival rate at phase one, λ_1;

iv. the arrival rate at phase two, λ_2.

In performance modeling, an MMPP is of interest due to its analytical tractability and versatility. Because detail analysis for an MMPP can be worked out, see later in this appendix, and the arrivals created by an MMPP are highly correlated.

Some results related to Markov modulated queueing systems are available in the literature. Knessl and Mathkowsky study the stationary distribution of the unfinished work for a two-state MMPP/G/1 queue [9]. Prabhu and Zhu complete a detailed analysis, including the waiting time, the idle time and the busy period, of an MMPP/G/1 queue [17]. Recently, Zhu, in his work [23], includes new results for an MMPP/M/1 queue with bulk arrivals. In [19], Ramaswami details the analysis of an N/G/1 queue, a more general case of MMPP/G/1 queue, including the arrival process, the stationary queue length at departures, and the waiting time. In the rest of this appendix, we summarize the results for a single server queueing system in which the arrival process is a two-state MMPP with batch arrivals and the service times are constant. We denote such a system as a two-state $\mathrm{MMPP}^{[X]}$/D/1 queueing system.

We define the following notation for a two-state $\mathrm{MMPP}^{[X]}$:

- π_i the steady state probability of the system being in phase i;
- λ_i the arrival rate for the system at phase i;
- γ_i the transition rate for the system at phase i;
- $g_i(k)$ *Prob*{bulk size is k given that the system is at phase i};

Throughout this appendix, $\tilde{\mathcal{F}}(z)$ represents the *Z-transform* of a discrete probability distribution function f_n, e.g., $\tilde{G}_j(z)$ means the Z-transform of $g_j(k)$. We use boldfaced uppercase characters as the notation for matrixes and the corresponding lowercase characters with subscripts being the notation for the elements. A character representing a vector will be arrow-headed, like $\vec{\pi} = (\pi_1, \pi_2)$.

Let $S(t) = u(t - \bar{t})$ be the service time distribution where $u(\cdot)$ is the unit step function and $\bar{t}$ is

† A detail analysis for N-process, which is called *Batch Markov Arrival Process* later by Lucantoni in [11], can be found in [16]. [11] and [19] detail the N/G/1 queueing system; and [2] extends the results to finite-buffered case.

the constant service time. The following is a procedure to calculate the expected waiting time, w, for a two-state MMPP$^{[X]}$/D/1 queueing system:

i. $$D_k = \begin{bmatrix} \lambda_1 g_1(k) & 0 \\ 0 & \lambda_2 g_2(k) \end{bmatrix} \quad \forall\ k \geq 1 \quad \text{with} \quad D_0 = \begin{bmatrix} -(\lambda_1+\gamma_1) & \gamma_1 \\ \gamma_2 & -(\lambda_2+\gamma_2) \end{bmatrix}, \quad \text{and}$$

$$D = \sum_{k=0}^{\infty} D_k = \begin{bmatrix} -\gamma_1 & \gamma_1 \\ \gamma_2 & -\gamma_2 \end{bmatrix};$$

ii. $$\vec{d} = \sum_{k=1}^{\infty} k D_k \vec{e}, \text{ where } \vec{e} = \begin{pmatrix} 1 \\ 1 \end{pmatrix};$$

iii. $$\tilde{\mathcal{D}}(G) = \sum_{k=0}^{\infty} D_k G^k \quad \text{and} \quad G = \int_0^{\infty} e^{\tilde{\mathcal{D}}(G)x} dx = \sum_{k=0}^{\infty} \left(\int_0^{\infty} e^{-\theta x} \frac{(\theta x)^k}{k!} dS(x) \right) \left(I + \frac{\tilde{\mathcal{D}}(G)}{\theta} \right)^k,$$

where $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\theta = max_i(-D(G)_{ii})$ (note that: starting with $G = \vec{e}\vec{\pi}$, the algorithm suggested by [11] can be used to solve G);

iv. compute $\vec{g}$ using $\vec{g}G = \vec{g}$ and $\vec{y} = (1-\rho)\vec{g}$ where ρ is the system utilization;

v. $V_1 = -\vec{t}\tilde{\mathcal{D}}'(1)$ and $V_2 = \vec{t}^2 (\tilde{\mathcal{D}}'(1) + \tilde{\mathcal{D}}''(1))$;

vi. $\vec{v}_1 = V_1 \vec{e}$ and $\vec{v}_2 = V_2 \vec{e}$;

vii. $$w_v = \frac{1}{2(1-\rho)} [2\rho + 2(\vec{y} - \vec{\pi}V_1)(\vec{e}\vec{\pi} + D)^{-1} \vec{v}_1 + \vec{\pi}\vec{v}_2];$$

viii. $$\vec{w} = -w_v \vec{\pi} - \vec{\pi} + (\vec{y} - \vec{\pi}V_1)(\vec{e}\vec{\pi} + D)^{-1};$$

ix. $$w = -(\vec{\pi}\vec{d})^{-1} \vec{w}\vec{d} + \frac{\vec{t}\vec{\pi}\tilde{\mathcal{D}}''(1)\vec{e}}{2\vec{\pi}\vec{d}}.$$

References

[1] A. Baiocchi et al., "Loss Performance Analysis of an ATM Multiplexer Loaded with High-Speed On-Off Processes," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 388-393, Apr. 1991.

[2] C. Blondia, "The N/G/1 Finite Capacity Queue," *Stochastic Models*, vol. 5, no. 2, pp. 273-294, 1989.

[3] P. T. Brady, "A Statistical Analysis of On-Off pattern in 16 Conversations," *Bell Systems Technical Journal*, pp. 73-91, Jan. 1968.

[4] P. T. Brady, "A Model for Generating On-Off Speech Patterns in Two-Way Conversation," *Bell Systems Technical Journal*, pp. 2445-2472, Sep. 1969.

[5] J. N. Daigle and J. D. Langford, "Models for Analysis of Packet Voice Communications Systems," *IEEE J. Select. Areas Commun.*, vol. 4, pp. 847-855, Sep. 1986.

[6] R. L. Graham, *Concrete Mathematics: A Foundation for Computer Science*, Menlo Park, CA: Addison-Wesley, 1989.

[7] H. Heffes and D. M. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance," *IEEE J. Select. Areas Commun.*, vol. 4, pp. 856-868, Sep. 1986.

[8] I. Ide, "Superposition of Interrupted Poisson Processes and Its Application to Packetized Voice Multiplexers," in *Proc. ITC '88*, pp. 3.1B2.

[9] C. Knessl and B. J. Matkowsky, "A Markov Modulated M/G/1 Queue I: Stationary Distribution," *Queueing Systems*, vol. 1, pp. 355-374, 1987.

[10] S.-Q. Li, "A General Solution Technique for Discrete Queueing Analysis of Multimedia Traffic on ATM," *IEEE Trans. Commun.*, vol. 39, pp. 1115-1132, Jul. 1991.

[11] D. Lucantoni, "New Results on the Single Server Queue with a Batch Markov Arrival Process," *Stochastic Models*, vol. 7, no. 1, pp. 1-46, 1991.

[12] D.Lucantoni and M. F. Neuts, *unpublished works.*

[13] B. Mukherjee, "Integrated Voice-Data Communications Over High-Speed Fiber Optic Networks," *Computer*, pp. 49-58, Feb. 1991.

[14] R. Nagarajan et al., "Approximation Techniques for Computing Packet Loss in Finite-Buffered Voice Multiplexers," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 368-377, Apr. 1991.

[15] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, Baltimore, MD: Johns Hopkins Univ. Press, 1981.

[16] M. F. Neuts, "A Versatile Markovian Point Process," *J. Applied Prob.*, vol. 16, pp.764-779, Dec. 1979.

[17] N. U. Prabhu and Y. Zhu, "Markov Modulated Queueing Systems," *Queueing Systems*, vol. 5, pp. 215-246, 1989.

[18] G. Ramamurthy and B. Sengupta, "Delay Analysis of a Packet Voice Multiplexer by the $\Sigma D_i/D/1$ Queue," *IEEE Trans. Commun.*, vol. 39, pp. 1107-1114, Jul. 1991.

[19] V. Ramaswami, "The N/G/1 Queue and Its Detailed Analysis," *Adv. Appl. Prob.*, vol. 12, pp. 222-261, Mar. 1980.

[20] K. Sriram and W. Whitt, "Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data," *IEEE J. Select. Areas Commun.*, vol. 4, pp. 833-846, Sep. 1986.

[21] R. C. F. Tucker, "Accurate Method for Analysis of a Packet-Speech Multiplexer with Limited Delay," *IEEE Trans. Commun.*, vol. 36, pp. 479-483, Apr. 1988.

[22] C. Yuan and J. A. Silvester, "Queueing Analysis of Delay Constrained Voice Traffic in a Packet Switching System," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 729-738, Jun. 1989.

[23] Y. Zhu, "A Markov Modulated M/M/1 Queue with Group Arrival," *Queueing Systems*, vol. 8, pp. 255-264, Apr. 1991.