MODELING PACKET DELAY IN MULTIPLEXED VIDEO TRAFFIC

Charles Thompson, Kavitha Chandra * ,Sudha Mulpur ** and Jimmie Davis ***

Center for Advanced Computation and Telecommunications University of Massachusetts, Lowell, MA 01854

In this paper a Markov chain based characterization of a single video source is used to model multiplexed video traffic and the resulting packet delay. The multiplexed video stream of statistically identical sources is represented using a reduced order approximation of the superposed Markov generator matrix. It is shown that inadequate spectral content in the single source generator matrix is manifested as an underestimate of the packet delay probabilities for the multiplexed stream. A new method for simplifying the generator matrix for the multiplexed video stream that amply models its spectral content is given. The simulated and calculated results using the aforementioned model are shown be in good agreement.

Keywords: Fluid buffer, VBR, Multiplexer

1. Introduction

Variable bit rate (VBR) video traffic sources can be successfully modeled by finite state, discrete time Markov chains [1,2,3]. The state structure, in conjunction with a rate process attributed to each state, allows modeling of the short-time correlation in traffic generated by a video-coder. The impact of temporal correlation in the output rate of a VBR video source on the queue response has been examined in [15] and it has been shown that macro-level correlations can be modeled by Markov chain based models. Long-range dependence seen in VBR video has also been examined and the queueing results have been compared to those obtained using a Markov chain based discrete autoregressive model [16]. It was concluded that for moderate buffer sizes, the short-range correlations obtained

* Partially supported by NSF CAREER grant NCR-9734585
** Currently at Tellabs, Willington, MA 01887
*** Supported by Lucent Technologies Cooperative Research Fellowship Program

using Markov chain models are sufficient to estimate the buffer characteristics.

Traffic analysis [2,3] of H.261 and MPEG-2 coders has shown that typically 15 to 20 states are required to faithfully model the queue behavior of moderate to high activity video sources. Single source models form the basis for determining the queueing delay and loss distributions for multiplexed video sources. The distribution of the queueing delay is considered to be the sum of two components: one resulting from packet-level (cell-level) and the other from burst-level queueing [4]. In packet-level (cell-level) queues, the arrival packet stream is modeled as a periodic process with random displacements in the starting epoch [5,6]. Burst-level queues have been analyzed using the fluid buffer approximation [7]. The approximation requires that the rate process be prescribed and this is done using a generator matrix and a rate vector to model the packet traffic. The multiplexing of N identical VBR sources each having K states and K associated rates yields a source model for the aggregate traffic which has $O(K^N)$ states. Hence, the number of states required exponentially increases with the number of sources. To counter this problem, reduced order traffic models have been used as approximations. Multi-state Markov sources obtained by the superposition of multiple two state on-off sources have been used to model multiplexed sources [8-10]. In these models, M multiplexed two state on-off sources require $O(M)$ states to represent the aggregate process. For moderate to high activity video sources [2], the aforementioned approach can be shown to underestimate the packet delay. In [11], uniform sampling of the bit rate histogram of the video source has been used to develop a Markov modulated rate process model. The sampling rate and scalability of the model with increasing numbers of sources remain open issues. The asymptotic behavior of the tail probability of the queue length for large buffers can be effectively used when a small loss or delay probability is required. In such a case, the largest negative eigenvalue of the product of the inverse of drift matrix and generator matrix yields the exponential decay rate of the delay distribution. The link capacity associated with this eigenvalue is called the effective bandwidth [12]. However, in order to assess the effectiveness of multiplexing and the impact of coding algorithms, the buffer behavior in the low to medium occupancy regime must be considered as well. Recently, Choudhury *et. al* [13] have determined higher order approximations that include the asymptotic constant that multiplies the exponential term of the delay distribution.

In this work, a model for the traffic generated by multiplexed video sources and the resulting packet delay distribution for a fluid buffer is given. A Markov chain based traffic model for a single VBR video source derived from measured data is used. Using this source model, the characteristics of the multiplexed traffic stream is obtained. It is shown that the degree of sampling of the rate histogram influences spectral content of the generator matrix. Undersampling the rate process is manifested as an underestimate of the packet delay probability. This feature becomes problematic as one increases the number of multiplexed sources. Therefore, in order to obtain a scalable model for the aggregate source traffic, a minimum number of states in the single source model should be retained. Methods for reducing the state space dimension of the aggregate source model need to be identified. We consider the minimal Markov chain representation of the video source that can be justified by measurement. This approach is implemented using an invariance constraint on the eigenvalue distribution of the transition matrix with increasing state dimension. The delay distribution for multiplexed VBR sources is computed using a fluid approximation of the buffer. Next, when aggregated into packets, the bit stream of the encoder yields the mean packet rate and a fundamental packet rate. By virtue of this feature, N multiplexed sources each emitting packets at M harmonics of the fundamental packet rate, require a maximum of $O(NM)$ states for the analysis of the delays. This simplification results in a substantial reduction in the complexity of the model.

2. VBR Source Model

The VBR video model used in this work is derived from a teleconferencing video scene encoded using the H.261 recommendation. In this scheme, the encoder first generates a single intraframe (I-frame) at the beginning of the video sequence. Subsequent frames are encoded as predictive frames (P-frames). The frames are generated at $1/30^{th}$ of second intervals and carry a variable number of bits per frame. Successive P-frames carry differential information with respect to a previous scene change. This state of affairs causes the number of bits in successive P-frame to P-frame transitions to be correlated. The P-frame to P-frame transitions are modeled by a Markov chain and the output rate distribution for each of the constituent states is taken to be a Gaussian distribution.

The parameters for the video source model were obtained from an analysis of approximately 5 minutes of the H.261 coded video sequence. To segment the rate distribution, a K-means cluster detection algorithm is used on the two dimensional space formed by the rates $\hat{r}(n)$ and $\hat{r}(n+1)$ where the index n is the frame index and $\hat{r}$ is given in bits per frame. The number of cluster centers retained is equal to the number of states. The temporal dependence in the bit rate is modeled by a K-state Markov chain [2,3]. The number of states required is a function of the number of scene types in the video sequence. For each state, the number of bits per frame is a random variable that is drawn from a rate probability distribution. As a result, a continuous range of the frame bit rates is identified with each scene type.

Correlation effects afforded by the generator matrix play a dominant role in the analyses of burst-level queueing. Therefore, the magnitude of the transition probabilities must be identified from the data. The ratio of the number of transitions from state i to state j and the total number of transitions out of state i is used to estimate the transition probabilities in the generator matrix. This is done once the cluster centers are determined. The probability of the video source being in a given state is expressed by the vector $\hat{\mathbf{p}}$. The i^{th} element $\hat{p}_i(n)$ of the vector $\hat{\mathbf{p}}$ corresponds to the probability that the source is in the i^{th} state at frame n. The transition probability between the states is determined by the Markov chain with transition probability matrix $\hat{Q}$ and

$$\hat{\mathbf{p}}(n) = \hat{\mathbf{p}}(n-1)\hat{Q} \tag{1}$$

To retain the burst-level traffic characteristics, an appropriate number of the states as identified by the number of cluster centers must be retained. The number of clusters chosen will govern the ability to capture the dynamics of the short-time correlation in the video source traffic. This correlation effects the queue behavior and is captured by accurately modeling the state sojourn times. Therefore, the order of the generator must be chosen such that these features are captured. When using too few states, the tail probabilities of the rate histogram will not be captured, thereby yielding an underestimate of the packet delay or loss probability. This situation has been observed by Haβlinger[14] in modeling VBR sources using semi-Markov models.

In the discrete-time Markov chain model, the temporal correlation in the rate time series is captured by the eigenvalues of $\hat{Q}$ which have magnitudes between zero and one. Since the correlation depends on the eigenvalues geometrically, the model order K must be increased until all subsequent increases of the model order yield eigenvalues that fall below a prescribed threshold value. In this

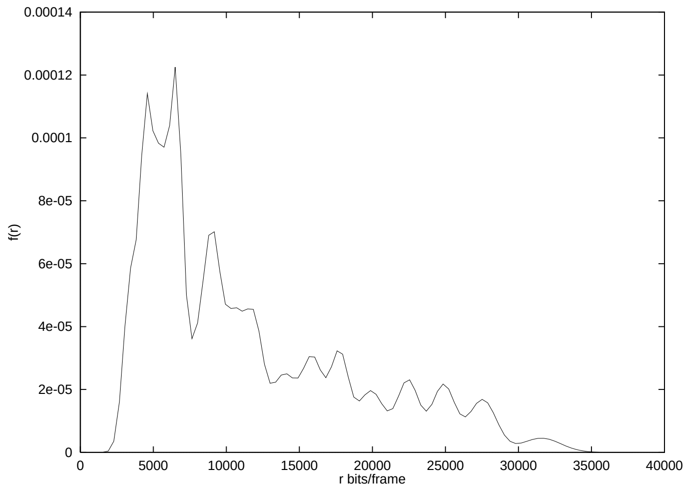

Figure 1. Rate pdf $f(r)$

way, the dominant eigenvalues of the source are retained in the model. Chandra $et. al$[3] have shown that for the video conferencing data used in this model, a 16 state model is sufficient in this regard. Adding states yields eigenvalues having a magnitude of less than 10^{-1}.

The elements of the rate vector $\hat{r}_i$ in bits per frame and the variance in the rate $\hat{\sigma}_i^2$ are given in Table 1. The corresponding state index is given by the integral value i. In such case, the steady state rate *pdf* is given by

$$f(\hat{r}) = \sum_{i=1}^{K} \bar{p}_i \, \frac{e^{-(\hat{r}-\hat{r}_i)^2/(2\sigma_i^2)}}{\sigma_i\sqrt{2\pi}} \tag{2}$$

The coefficient $\bar{p}_i$ is the steady state probability of state i obtained from the normalized right eigenvector corresponding to the unity eigenvalue of $\hat{Q}^T$. The expected value for the rate is equal to the first moment of $f(\hat{r})$. The rate *pdf* for the 16 state model derived from the teleconferencing date set is given in Figure 1. In this case, the mean rate is $11,586$ bits per frame. At 30 frames per second, the bit rate is approximately equal to $0.35 Mbits/sec$. When the number of states are reduced, the weaker peaks occurring at the tail of the *pdf* are aggregated.

Table 1
Mean and variance of the rate distribution for each state

state i	rate (bps) $\hat{r}_i$	std. deviation $\hat{\sigma}_i$
1	3447.41	494.06
2	4507.80	427.26
3	5524.21	504.81
4	6566.66	406.32
5	7838.01	1066.35
6	9048.14	564.76
7	10528.33	679.95
8	11897.90	613.91
9	13934.40	906.08
10	15931.20	708.13
11	17735.60	631.07
12	19902.00	924.07
13	22449.10	757.52
14	24870.30	799.91
15	27550.30	938.45
16	31564.90	1329.96

3. Multiplexed sources and fluid buffers

In this section, an outline the procedure for developing a model for multiplexed video traffic is presented. The single source model will serve as the basis for the multiplexed video model. First, the discrete-time model presented in the previous section will be recast in terms of its continous-time analog. Let $\hat{\mathbf{p}}(n)$ be the samples of its continuous-time analog $\mathbf{p}(\mathbf{t})$. The samples occur at equal intervals of time, ΔT. The time interval ΔT is equal to the frame repetition time which is equal to 1/30 of a second. The forward difference $\hat{\mathbf{p}}(n) - \hat{\mathbf{p}}(n-1)$ is approximated by $\Delta T \partial \mathbf{p}/\partial T$, where the variable T denotes time. Using the non-dimensionalization $t = T/\Delta T$ the probability of the source being in one of K-states at non-dimensional time t is modeled by the continuous time Markov chain

$$\frac{\partial \mathbf{p}}{\partial t} = \mathbf{p}Q \tag{3}$$

where $Q = \hat{Q} - I$ is the generator and the component p_i of the row vector $\mathbf{p}$ represents the probability of being in the i^{th} state at time t. The mean rate of the number of packets generated in a video frame is given by the rate vector $\mathbf{r}$. The elements of the vector are obtained by scaling the elements of $\hat{\mathbf{r}}$ by a factor $12,000$ bits (1500 bytes) per packet. The component r_i of $\mathbf{r}$ is taken to be the output rate when the source is in the i^{th} state.

The multiplexing of two mutually independent and identical sources yields two dimensional state probabilities which can be modeled and expressed in form similar to that given in Eqn. (3). The mapping of the set of two-dimensional states to a one-dimensional state vector is accomplished by replacing the generator and the state probability vector by the Kronecker sum $Q \bigoplus Q$ and product $\mathbf{p} \bigotimes \mathbf{p}$, respectively. The rate matrix R, initially comprised of the elements of $\mathbf{r}$ on its diagonal, is replaced by the Kronecker sum $R \bigoplus R$. The Kronecker sum operation increases the dimension of the generator matrix to K^2. In the case of N sources, the generator and rate matrices are obtained by taking an N-fold Kronecker sum of Q and R, respectively. The dimension of each of these matrices is K^N. The objective of this work is to develop a model that reduces the state space dimension while maintaining scalability. Furthermore, the model should adequately estimate the complementary delay distribution.

Scalable computation of the generator is possible under the assumption that the first moment of the rate probability distribution of N multiplexed sources is known. For N identical independent sources, the mean rate is equal to $N\bar{r}$ where $\bar{r}$ is the mean rate of a single source. The approach is to quantize the packet rate into integral multiples of a fundamental rate r_o. In the case of a VBR source having K-states, the packet rate per frame is approximated as

$$\mathbf{r} = \{r_1, r_2, \cdots, r_K\} \approx r_o \{m_1, m_2, \cdots, M\} \tag{4}$$

Here, $r_{j+1} \geq r_j$ and m_1, m_2, $\cdots$, M are integer harmonics of the fundamental. The fundamental rate r_o is determined using $r_o = r_1/m_1$. Starting at $m_1 = 1$, the value of m_1 is increased until the first moment of the resulting rate distribution is within a prescribed tolerance of the exact result. The value for the first multiplier m_1 is limited to range from 1 to K. The remaining multipliers are evaluated as $m_i = r_i/r_o$. Therefore, each component of the packet rate vector is a harmonic of r_o. The diagonal rate matrix R is constructed by using the elements of $\mathbf{r}$ as the diagonal elements.

The matrix R may contain columns that have identical rate elements. The approach is to prescribe a single rate for each state. A reduced dimension generator matrix is obtained by summing the columns and weighting the rows that are associated with an equivalent rate. The number of unique rates associated with the superposition of N sources each defined by M quantized rates is given by the length of the N-fold convolution of vectors each of dimension M. This yields a system model of $O(NM)$ states.

For the case where two rates r_k and r_l have equal rate values, the state reduction process is outlined as follows. These rates correspond to the state represented by the k and l columns of the K-state generator matrix. For this purpose, let $k < l$. The steady state probability for these two states are $\bar{p_k}$ and $\bar{p_l}$. The reduction process is given in Eqns (5) and (6). The reduced transition matrix Q'_{ij} is obtained by summing the coefficents in the k and l columns and placing the result in column k. In addition, the k and l rows are also combined. Rows k and l are weighted by the fractions $\bar{p}_k/(\bar{p}_k + \bar{p}_l)$ and $\bar{p}_l/(\bar{p}_k + \bar{p}_l)$, respectively and summed. The resultant sum replaces row k.

$$Q'_{ij} = \begin{cases} Q_{ij} & for\ i \neq k, j \neq k \\ Q_{ik} + Q_{il} & for\ i \neq k, j = k \\ \dfrac{\bar{p}_k Q_{kj} + \bar{p}_l Q_{lj}}{\bar{p}_k + \bar{p}_l} & for\ i = k, j \neq k \\ \dfrac{\bar{p}_k (Q_{kk} + Q_{kl}) + \bar{p}_l (Q_{lk} + Q_{ll})}{\bar{p}_k + \bar{p}_l} & for\ i = k, j = k \end{cases} \tag{5}$$

$$R'_{ij} = \begin{cases} r_i & for\ i = j \\ 0 & otherwise \end{cases} \tag{6}$$

Column and row l are eliminated from the matrix. The remaining columns where $K \geq j > l$ are shifted to the left by one column. As a result, Q' has $K - 1$ columns. The remaining rows where $K \geq i > l$ are shifted up by one row. Hence, Q' becomes a matrix having dimension $K - 1$. The rate matrix R' is simply comprised of the diagonal elements of the rate vector with the r_l element removed. The two source model is obtained by replacing the matrices R and Q by R' and Q', respectively.

The number of sources can be succesively doubled by recursively applying the aformentioned method. The index of the first harmonic is chosen using the following process: For a single source, the fundamental $r_o^{(1)}$ is chosen to yield the minimum error in the mean value of the rate. The index m_1 ranges from 1 to $K^{(1)}$ where $K^{(1)}$ is the dimension of the one source generator matrix. When N is a power of two the fundamental is obtained recusively using Eqn (7).

$$r_o^{(N)} = \begin{cases} r_o^{(N/2)} & for\ \dfrac{r_1^{(N)}}{2r_o^{(N/2)}} \geq K^{(N)} \\ \\ 2r_o^{(N/2)} & for\ \dfrac{r_1^{(N)}}{2r_o^{(N/2)}} < K^{(N)} \end{cases} \tag{7}$$

The variables N and K denote the number of sources and the dimension of the generator, respectively. The superscript denotes the variable that is evaluated at that iteration of the source count.

The model for the multiplexed VBR sources will be evaluated using a fluid buffer model of a queue. The aggregated traffic stream enters a queue with infinite waiting room. Packets in the buffer are serviced on a first-in first-out basis at a constant rate of c packets per frame. In the steady state, burst-level delay can be modeled using the fluid buffer approximation. The variable x is taken to be the continuous analog of the buffer occupancy in packets. Following [7], the complementary probability of the buffer having more than x packets is $Pr(Delay > x) = 1 - \mathbf{p}\mathbf{p}^{\mathbf{T}}|\mathbf{p}|^{-\mathbf{1}}$ where $\mathbf{p}$ satisfies

$$\frac{\partial \mathbf{p}}{\partial x}\left[R - Ic\right] = \mathbf{p}\mathbf{Q} \tag{8}$$

The drift matrix $R-Ic$ is a diagonal matrix comprised of the elements $r_i - c$. The solution of Eqn (8) follows that of an eigenvalue problem. For an infinite buffer, subject to consideration that the solution is bounded at x equal to infinity, the leading coefficients of the exponentially growing modes are set equal to zero. The amplitude of the remaining modes is determined by equating $p_i(x = 0)$ equal to zero for cases where $r_i - c$ is greater than zero.

4. Results

The fluid buffer analysis is carried out for a H.261 encoded video source described by Markov transition and rate parameters. The estimation of these parameters from data is discussed in Section 2. Each state is then associated

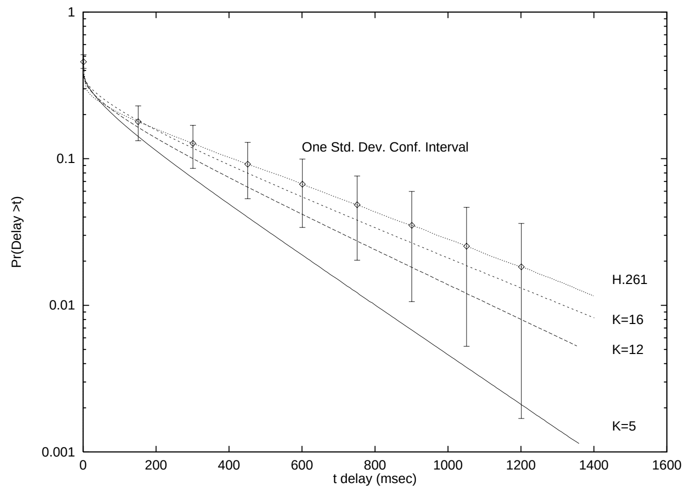

Figure 2. Complementary delay distribution for 4 multiplexed H.261 sources. $\rho = 0.8$

with an average arrival rate. The selection of the optimal number of states is based on statistical matching of the rate histogram and an invariance criteria on the dominant eigenvalue distribution of the transition matrix with increasing value of K. A 16-state Markov chain satisfies these constraints for the considered source. The source average and peak rates are approximately 0.35 and 1.0 $Mbits$ per second at 30 frames per second. The packet size is 1500 bytes.

A fluid buffer model is used to examine how the choice of the number of states in the single source model influences the queueing characteristics of the multiplexed stream. Using a 5 minute duration video sequence from a single source, multiple sources are constructed by concatenating segments of 300 consecutive frames from the original source with randomly selected starting points. The delay distributions are obtained for many such ensembles and averaged to yield the representative queueing behavior of the multiplexed stream. For each value of the queue length, the mean and standard deviation of the buffer occupancy is obtained. A K-state Markov chain model is derived for a single source, for $K = 5, 12$ and 16 states. Figure 2 depicts the delay distributions obtained from the fluid-buffer model for these values of K. Also shown is the ensemble averaged queue distribution obtained from the H.261 video data along with one standard deviation interval values. The results shown are for multiplexing $N = 4$ sources. It can be seen that inadequate modeling of the spectral content in the single source as in the $K = 5$ case, significantly underestimates the packet-delay

Table 2
Dimension of the simplified generator matrix

Number of Sources, N	Order of the Reduced Generator Matrix	Relative error
1	16	
2	112	1.26×10^{-4}
4	423	-1.36×10^{-3}
8	556	9.32×10^{-4}
16	652	7.30×10^{-4}
32	661	5.30×10^{-4}

probabilities experienced by the video traffic. As K is increased the asymptotic decay rate obtained from the fluid buffer model approaches that obtained using actual source data. The difference in the decay rates is magnified as N is increased, if the magnitude of K is insufficient to completely capture the spectral content of the single source generator matrix. It can be seen from the figure that 16 states can be used to adequately estimate the delay distribution.

The 16-state Markov chain model is used to model each of the multiplexed video sources. The generator matrix for the aggregate sources was simplified by combining the states resulting from the Kronecker sum with those which have common quantized rates values. To do so, a fundamental rate was chosen using the method described in Section 3. The relative error in the mean rate using the reduced dimension generator was evaluated as $(\bar{r}^{(N)} - N\bar{r})/(N\bar{r})$. The relative error in the mean rate versus the number of sources multiplexed is given in Table 2. It is evident that the number of states required to match the mean rate of N multiplexed sources is far less than the 16^N states that would be required if no approximations were imposed on Kronecker sum of the generator matrices. The rate probability density function approaches a Gaussian density function as the number of sources is increased.

The aforementioned reduced state generator matrices were used to compute the complementary delay distribution in a first-in first-out fluid buffer. The complementary delay distributions $Pr(Delay > x)$ for $N = 2$, 4, 8, 16 and 32 multiplexed sources are depicted in Figure 3. In each case, the utilization ρ is equal to 0.8. The service rate $c = \bar{r}N/\rho$. The results of the models are compared to those obtained by simulation. The simulation results are bounded by the confidence intervals shown in the figure. The extreme values denoted in

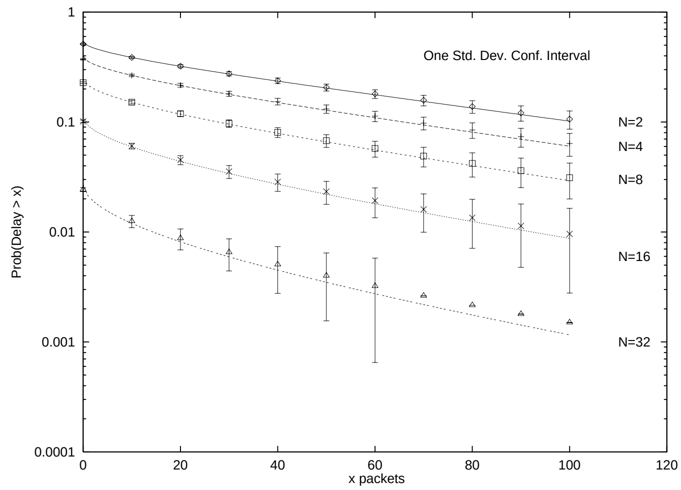

Figure 3. Complementary delay distribution for 2,4,8,16 and 32 H.261 sources , $\rho = 0.8$

the errorbars correspond to the location of one standard deviation value and the center point represents the mean value. In the simulation, each source emits 10^6 frames of data per realization. Each point is the result of an ensemble average of 60 realizations. The simulation and model results are in good agreement for $N \leq 16$ sources. Departure between these results occur at $N = 32$ sources. This difference is the result of incomplete modeling of the subdominant states at $N = 16$.

MPEG-2 two layer encoded video sources from entertainment applications are also used to verify the model. The entertainment video corresponds to a clip from "The Blues Brothers" movie. In MPEG-2, an I-frame is generated only at scene changes. Therefore, between I-frames, an arbitrary number of P-frames can occur depending on the video source characteristics such as the scene duration. The frame rate for MPEG-2 data is $1/24^{th}$ of a second. MPEG-2 data is modeled by a 20-state Markov chain. The fluid buffer analysis is carried out for the MPEG-2 multiplexed video data. The results for the complementary delay distribution for the multiplexing of $N = 2$ and $N = 4$ sources are given in Figure 4, in which the model results are compared with those of the simulation. The solid lines represent the results from the simulation and the dotted lines represent the results from the model. For a utilization of 0.8, it can be seen that the model is in good agreement with the simulation results.

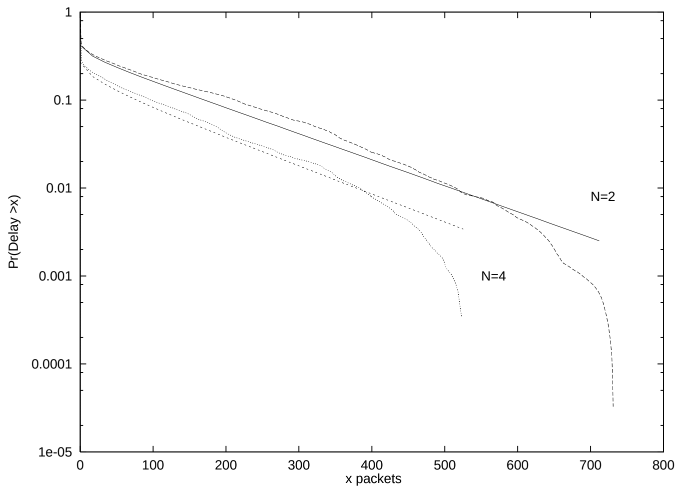

Figure 4. Complementary delay distribution for 2 and 4 MPEG-2 sources. $\rho = 0.8$

5. Summary

In this work, a fluid buffer model based analysis of the delay distribution for VBR sources was presented. A method for reducing the state space dimension of generator was given. The reduction process involved the quantization of the rates and the fundamental rate was chosen to yield the best match to the mean and variance of the rate. States having equal rates were aggregated thereby reducing the number of states in the generator matrix. The resulting model allowed for scalable analysis as the number of sources was increased. These approaches have been used to study the effect of model order on traffic descriptors. By comparison with the actual video data, it is shown that the queueing probabilities of multiplexed sources are significantly underestimated by approximating a single source with a reduced number of states. It was shown that the state reduction procedure yields reasonable results for modest numbers of sources if the single source models are sufficiently accurate.

References

[1] F. Yegenoglu, B. Jabbari and Y. Zhang, Motion-Classified Autoregressive Modeling of Variable Bit Rate Video, *IEEE Trans. Ckts. and Sys. for Video Tech.*, **3** (1993) 42-53.

[2] K. Chandra and A. R. Reibman, Modeling two-layer SNR scalable MPEG-2 video traffic, *Proc. 7th Intl Workshop on Packet Video*, Brisbane, Australia, 1996, pp. 7-12.

[3] K. Chandra and A. Reibman, Modeling one- and two-layer variable bit rate video, *IEEE/ACM Trans. on Networking*, **7** (1999) 398-413.

[4] J. W. Roberts, editor, Performance Evaluation and Design of Multiservice Networks, *COST 224 Final Rep.*, Commission of the European Communities, Brussels, 1992, Chapter 2.

[5] I. Norros, J.W. Roberts, A. Simonian, J.T. Virtamo, The superposition of variable bit rate sources in an ATM multiplexer, *IEEE J. Sel. Areas Comm.*, **9** (1991) 378-387.

[6] [6] J. W. Roberts and J.T. Virtamo, The superposition of periodic cell arrival streams in an ATM multiplexer, *IEEE Trans. Comm.*, **39** (1991) 298-303.

[7] D. Anick, D. Mitra and M.M. Sondhi, Stochastic theory of a data-handling system with multiple sources, *Bell Syst. Tech. J.*, **61** (1982) 1871-1894.

[8] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson and J.D. Robbins, Performance models of statistical multiplexing in packet video communications, sources, *IEEE Trans. on Comm.*, **36** (1988) 834-844.

[9] P. Sen, B. Maglaris, N. Rikli, and D. Anastassiou, Models for packet switching of variable-bit-rate sources *IEEE J. Sel. Areas Comm.*, **7** (1989) 865-869.

[10] S.-Q. Li ,and J. W. Mark, Traffic characterization for integrated services networks, *IEEE Trans. Comm.*, **38** (1990) 1231-1242.

[11] Skelly, M. Schwartz, and S. Dixit, A histogram-based model for video traffic behavior in an ATM multiplexer, *IEEE/ACM Trans. Networks*, **1** (1993) 446-459.

[12] A.I. Elwalid and D. Mitra, Effective bandwidth of general Markovian traffic sources and admission control of high speed networks, *Proc. IEEE INFOCOM 1993*, **1** (1993) 3a.2.1-3a.2.10.

[13] G. Choudhury, D. M. Lucantoni, and W. Whitt, Squeezing the most out of ATM, *IEEE Trans. Comm.*, **44** (1996) 203-217.

[14] G. Haβlinger, Semi-Markovian modelling and performance analysis of variable rate traffic in ATM networks, *Telecommunications Systems*, **7**, (1997) 281-298.

[15] S.-Q Li and C. Hwang, Queue response of input correlation functions: Discrete spectral analysis, *IEEE/ACM Trans. on Networking*, **1** (1993) 522-533.

[16] D. Heyman and T.V. Lakshman, What are the implications of long-range dependence for VBR-Video traffic engineering?, *IEEE/ACM Trans. on Networking*, **4** (1996) 301-317.