MPEG VIDEO TRAFFIC MODELS: SEQUENTIALLY MODULATED SELF-SIMILAR PROCESSES

Hai Liu, Nirwan Ansari, and Yun Q. Shi
New Jersey Center for Wireless Telecommunications
Department of Electrical and Computer Engineering
New Jersey Institute of Technology
University Heights, Newark, NJ 07102, USA

Abstract [1] New traffic models are called for to facilitate the design of effective admission and flow control algorithms, and network performance evaluation in accommodating video traffic. We propose a new approach, sequentially modulated self-similar processes (SMSSPs), to model MPEG coded video traffic. SMSSPs are shown to be able to capture both short range dependency (SRD) and long range dependency (LRD) of the video traffic accurately. Traffic data are decomposed according to the MPEG data structure, into several parts, each modeled as a self-similar process. These processes are then modulated sequentially in a manner similar to how the frames are grouped into the GOP (Group of Pictures) pattern.

Keywords: MPEG, modulated self-similar processes, video traffic modeling, long range dependency, short range dependency.

1. INTRODUCTION

The trend to transmit video over network, especially over ATM, is emerging. Traffic models are important to network design, performance evaluation, bandwidth allocation, and bit-rate control. It was, however, observed that traditional models fall short in describing video traffic because video traffic is strongly autocorrelated and bursty [1]. To accurately model video traffic, autocorrelations among data should be taken into consideration. A considerable amount of effort on video modeling has been reported. These models are used

[1] This work was done in part while N. Ansari was on leave at the Department of Information Engineering, Chinese University of Hong Kong, Hong Kong, and Y.Q. Shi was on leave at the Information Engineering Div., Nanyang Technological University, Singapore.

to capture two statistical factors: marginal distribution (first-order statistics) and autocorrelation function (second-order statistics) of traffic arrival times. The importance of long range dependency is among the most arguable issues in video modeling. Some of the results support the view that LRD has drastic impact on queuing performance [2],[3],[4],[5], while other results support the view that LRD has little impact on queuing performance because of the fact that the buffer capacity is limited in practice [6].

While the importance of long range dependency is arguable, the impact of short term autocorrelation in traffic processes on queuing performance with a finite buffer can be very drastic (see [7] and references in it). Simulation results show that the network queuing performance with strong and weak autocorrelation traffic may be quite different. Thus, a model should capture not only the first-order statistics, but also the second-order statistics. SRD models can capture short-term autocorrelation, but fail to capture long-term dependency. LRD models, on the other hand, can capture long-term dependency, but underestimate the short term dependency.

Markov-Renewal-Modulated TES (transform expand sample) models were used to model JPEG encoded motion pictures. One of the drawbacks of TES is that the ACF of a TES process for lags beyond one cannot be derived analytically. It can only be obtained by searching in the parameter space, and thus good results can hardly be guaranteed [8].

The $M/G/\infty$ input process model is a compromise between LRD and SRD models [8]. Simulation results were found to be better than those of a self-similar process when the switch buffer is relatively small. Better results than those of DAR(1) [9] model were found when the buffer size is large. The results were derived for JPEG and the I frames of MPEG sequences. As will be shown below, ACF of MPEG sequences is quite different from that of JPEG sequences and that of I sequences. In our opinion, it is almost impossible to accurately capture the ACF of MPEG compressed data by a simple function such as the exponential function, and thus this method fails to capture the second-order statistics of MPEG sequences.

In [10], I, P, and B frames were modeled separately. I frames were described by three parts: scene length, average I frame size during a scene, and variations from average frame size during the scene. P and B frames were modeled as $i.i.d$ processes. As we will show, I, P, and B frames are all LRD processes. Another characteristic of B frames is that its ACF exhibits a repeated pattern (see Fig. 1, 2, and 3) (The repeated pattern appears in B frames only!). The fact that B frames occupy a very large part of the whole sequence and B frames are also rather large (actually in *star wars*, the largest frame is a B frame) suggests that the impact of B frames cannot be ignored, and it is inappropriate to model P and B frames as $i.i.d$ processes.

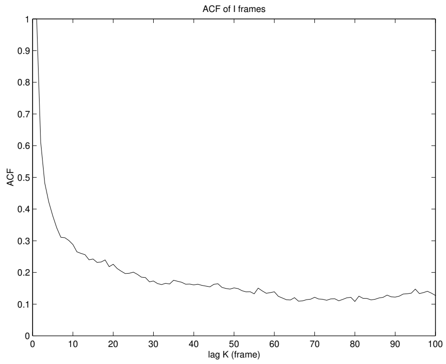


Figure 1 ACF of star wars' I frames.

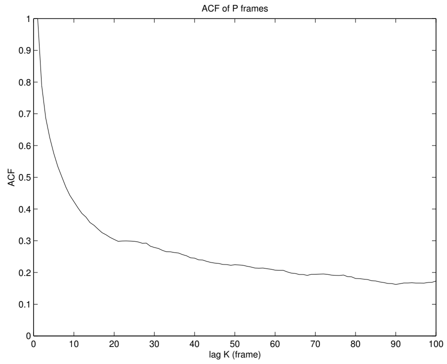


Figure 2 ACF of star wars' P frames.

In this paper, we propose to model MPEG compressed video sequence by sequentially modulated self-similar processes, in which the original sequences are decomposed into several parts that can be modeled by self-similar processes. It has been found that video traffic possesses self-similarity, and thus it is natural to model video traffic by self-similar processes. Self-similar processes have very simple ACF forms, and therefore, are easier to analyze than other kinds of processes. The rest of the paper is organized as follows. Section 2 describes empirical data and ACF. Concepts of SRD, LRD and self-similar processes are presented in Section 3. Modeling of decomposed data and the whole data set along with simulation results are discussed and presented in Section 4.

2. EMPIRICAL DATA AND ACF

The empirical data used here was MPEG coded data of *Star Wars*[2]. The source contains materials ranging from low complexity/motion scenes to those with high and very high actions. The data file consists of 174,136 frames, each having a different frame size (bytes per frame). The movie length is approximately 2 hours at 24 frames per second. The original video was captured as 408 lines by 508 pixels, and then interpolated to 240×352 (Luminance - Y), and 120×176 (Chrominance - U and V). Every frame was partitioned into blocks of 8×8 pixels. These data blocks were transformed using DCT. After DCT transformation, coefficients were quantized and Huffman coded. Run length coding was further used to reduce bit rate. Motion estimation techniques were used to compress data volume. The frames were organized as follows: IBBPBBPBBPBB IBBPBB . . . , i.e., 12 frames in a Group of Pictures

[2]The MPEG coded data were the courtesy of M.W. Garrett of Bellcore and M. Vetterli of UC Berkeley.

(GOP). I frames are those which use intra frame coding method (without motion estimation), P frames are those which use inter frame coding technique (with motion estimation), and B frames can be predicted using forward and backward prediction.

The ACF of the MPEG coded *Star War* is shown in Fig 4. It fluctuates around three envelopes, reflecting the fact that, after the use of motion estimation and forward/backward prediction techniques, the dependency between frames is reduced. This characteristic should be taken into consideration in modeling MPEG coded video sequences. We propose to decompose the sequence into I, P, B_1, $B_2, \cdots$, B_8 according to the GOP pattern, and model every part by a different self-similar process.

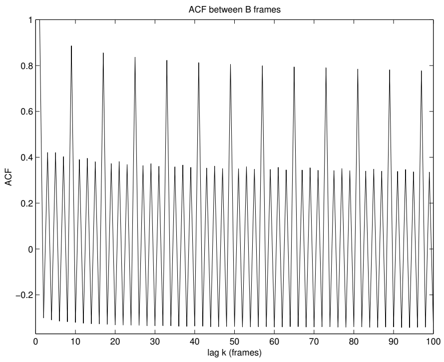


Figure 3 ACF of star wars' B frames.

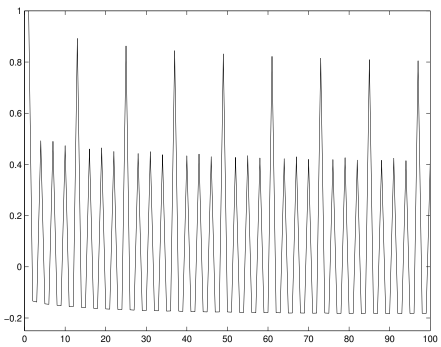


Figure 4 ACF of the MPEG video.

3. SRD, LRD, AND SELF SIMILARITY

Consider a stationary process $X = \{X_n : n = 1, 2, \dots\}$ with mean μ and variance σ^2. The autocorrelation function and the variance of X are denoted as:

$$r(k) = \frac{E[(X_n - \mu)(X_{n+k} - \mu)]}{\sigma^2} \tag{1}$$

and

$$\sigma^2 = E[(X_n - \mu)^2]. \tag{2}$$

X is said to be SRD if $\sum_{k=0}^{\infty} r(k)$ is finite; otherwise, the process is said to be LRD [11].

Let X defined above have the following autocorrelation function:

$$r(k) \rightsquigarrow k^{-\beta} L(k), k \to \infty \tag{3}$$

where $0 < \beta < 1$, and L is a slowly varying function as $k \to \infty$, i.e., $\lim_{t\to\infty} L(tx)/L(t) = 1$ for all $x > 0$. Consider the aggregated process

$$X^{(m)} = \{X_t^{(m)}\} = \{X_1^{(m)}, X_2^{(m)}, \dots\},$$

where

$$X_t^{(m)} = \frac{1}{m}(X_{tm-m+1} + \cdots + X_{tm}), t \in Q, m \in Q, \tag{4}$$

and Q is a positive integer set. X is said to be exactly second-order self-similar [11] if

$$var X^{(m)} = \sigma^2 m^{-\beta} \tag{5}$$

and

$$r^{(m)}(k) = r(k) \tag{6}$$

for all $m \in \{1, 2, 3, \cdots\}$ and $k \in \{0, 1, 2, \cdots\}$. Here $r^{(m)}(k)$ is the autocorrelation function of $X^{(m)}$. In fact, Eq. (5) is sufficient to define a self-similar process since Eq. (3) and (6) can be derived from Eq. (5) [11].

Since empirical video traffic exhibits self-similarity and long range dependency, it is intuitive to use self-similar processes to model video traffic. It is one of the most often used processes to capture LRD of video traffic. Often times, a self-similar process is simply referred to as a LRD process.

Hurst parameter $H = 1 - \beta/2 (0 < \beta < 1)$ is used to measure the similarity of a process. It is the only parameter needed to describe a second-order self-similar process. For a process with self-similarity, $1/2 < H < 1$.

4. MODELING MPEG TRAFFIC

In order to model MPEG coded data, we decompose the MPEG traffic into 10 sub-sequences $X_I, X_P, X_{B_1}, X_{B_2}, \cdots$, and X_{B_8}. X_I consists of all I frames, X_P consists of all P frames, the first B frames in all GOPs constitute X_{B_1} , the second B frames in all GOPs constitute X_{B_2}, and so on. We have used $k^{-\beta}$, $e^{-\beta k}$, and $e^{-\beta\sqrt{k}}$, corresponding to the ACFs of a self-similar process, a Markov process, and an $M/G/\infty$ input process, respectively, to approximate ACFs of these processes. For illustrative purposes, some of these approximations are shown in Fig. 5 to 12. The sums of squares of errors obtained by the three kinds of methods are tabulated in Table 1.1. It is quite obvious that self-similar processes are better choices. We therefore use self-similar processes to model these data.

Using the least squares method, $\beta = 0.4663$, 0.3546, 0.4468, 0.4779, 0.4294, 0.4656, 0.4380, 0.4682, 0.4465, and 0.4606 are derived for X_I, X_P,

Table 1.1 Least square errors obtained by self-similar process, markov and $M/G/\infty$ method

	I	P	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
LRD	0.46	0.35	0.45	0.48	0.43	0.47	0.44	0.47	0.45	0.46
$M/G/\infty$	5.13	12.3	5.40	4.80	6.23	5.19	5.83	5.05	5.03	5.34
Markov	7.98	21.0	9.45	7.30	11.1	8.12	10.2	7.95	7.22	8.75

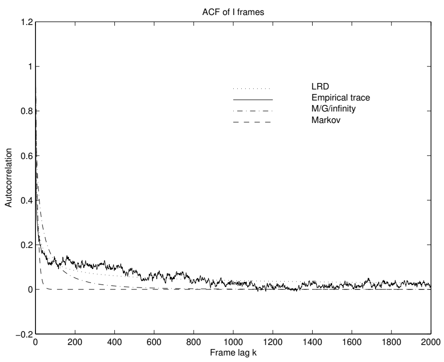


Figure 5 Approximation for ACF of I frames by : LRD, M/G/∞, and Markov processes.

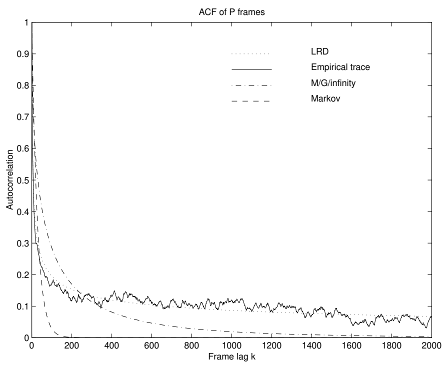


Figure 6 Approximation for ACF of P frames by : LRD, M/G/∞, and Markov processes.

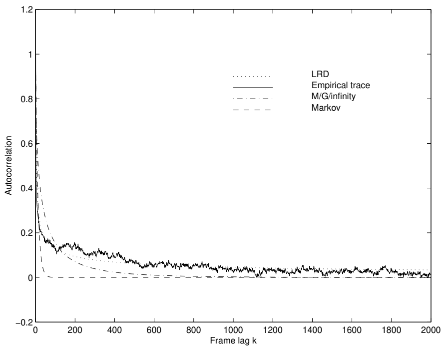


Figure 7 Approximation for ACF of B_1 frames by : LRD, M/G/∞, and Markov processes.

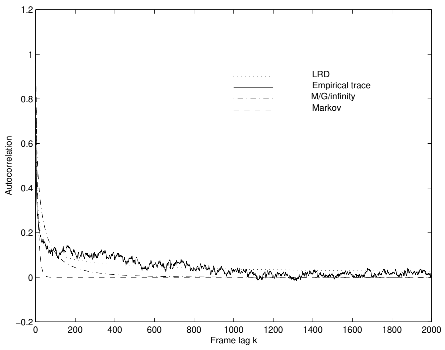
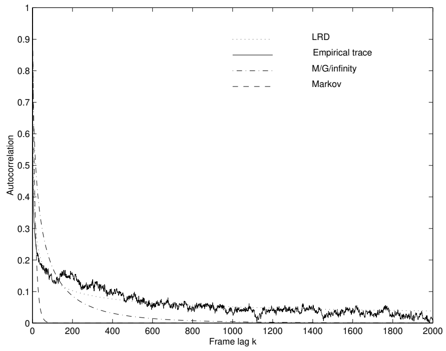


Figure 8 Approximation for ACF of B_2 frames by : LRD, M/G/∞, and Markov processes.

X_{B_1}, $X_{B_2}, \cdots$, and X_{B_8}, respectively. The corresponding Hurst parameters



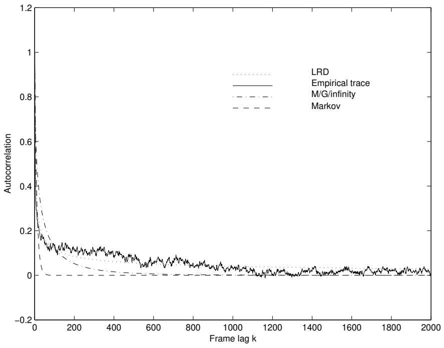
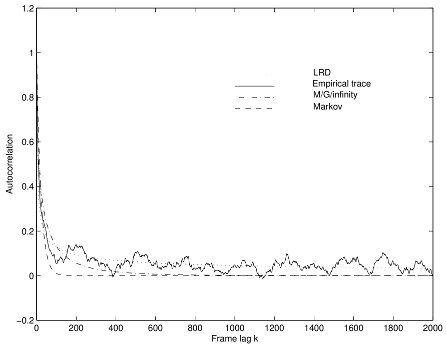


Figure 9 Approximation for ACF of B_3 frames by : LRD, M/G/∞, and Markov processes.

Figure 10 Approximation for ACF of B_4 frames by : LRD, M/G/∞, and Markov processes.



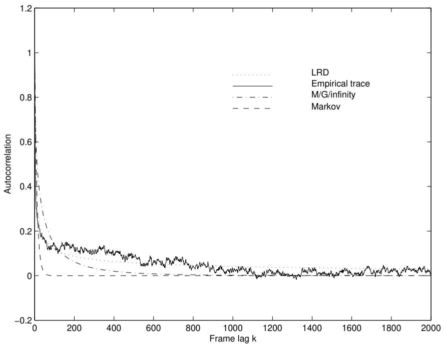


Figure 11 Approximation for ACF of B_7 frames by : LRD, M/G/∞, and Markov processes

Figure 12 Approximation for ACF of B_8 frames by : LRD, M/G/∞, and Markov processes

for these processes are $H = 0.7668$, 0.8227, 0.7766, 0.7610, 0.7853, 0.7672, 0.7810, 0.7659, 0.7768, 0.7697, respectively.

To model marginal distributions of these processes, we use Beta distributions which have the following form of probability density function:

$$f(x;\gamma,\eta,\mu_0,\mu_1) = \begin{cases} \frac{1}{\mu_1-\mu_0}\frac{\Gamma(\gamma+\eta)}{\Gamma(\gamma)\Gamma(\eta)}(\frac{x-\mu_0}{\mu_1-\mu_0})^{\gamma-1}(1-\frac{x-\mu_0}{\mu_1-\mu_0})^{\eta-1} & \\ \qquad \mu_0 \le x \le \mu_1, 0<\gamma, 0<\eta & \\ 0 \qquad \text{otherwise,} & \end{cases} \tag{7}$$

where γ and η are shape parameters, and $[\mu_0, \mu_1]$ is the domain where the distribution is defined. They can be estimated by the following formulae [12]:

$$\hat{\eta} = \frac{1 - \bar{x}}{s^2}[\bar{x}(1 - \bar{x}) - s^2] \tag{8}$$

$$\hat{\gamma} = \frac{\bar{x}\hat{\eta}}{1 - \bar{x}} \tag{9}$$

where

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{10}$$

$$s^2 = \frac{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}{N(N-1)}, \tag{11}$$

and N is the number of data in the data set. Using these formulae, $\hat{\eta} = 1.5237$, 1.5699, 1.4172, 1.3016, 1.6858, 1.6329, 1.7276, 1.4218, 4.0585, 1.5402, and $\hat{\gamma} = 12.7263$, 11.1939, 8.1089, 8.1604, 11.8499, 13.9278, 12.2180, 8.6536, 10.4233, 11.1768 are derived for X_I, X_P, X_{B_1}, X_{B_2}, $\cdots$, and X_{B_8}, respectively.

By combining X_I, X_P, X_{B_1}, X_{B_2},$\cdots$, and X_{B_8} in a manner similar to the GOP pattern, a model for MPEG coded traffic is obtained. This model can be used to generate traffic data.

Fig. 13 shows a trace of the empirical video traffic, and the trace generated by our model is shown in Fig. 14. Note the similarity between these two figures. Since traffic is random, the appropriateness of a traffic model should be judged by its statistical properties rather than the mere similarity between these two figures. This can be demonstrated by the ACF of the generated traffic shown in Fig. 15 (compare to Fig. 4), and the ACF of B frames shown in Fig. 16 (compare to Fig. 3), implying that the proposed model can capture both the LRD and SRD of B frames[3]. The ACF of P and I frames can also be captured very well (not shown here owing to the limited space.)

5. CONCLUSIONS

We have proposed a new traffic model, sequentially modulated self-similar processes, to model MPEG compressed video sequence. The model can match

[3]That is, the model can match the ACF of B frames for both large and small lag, k. Owing to the limited space, the ACF of B frames for large k is not shown.

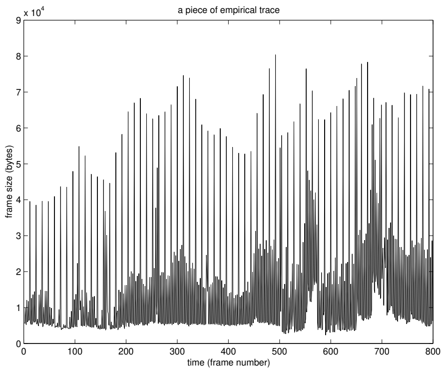


Figure 13 A trace of the empirical traffic data.

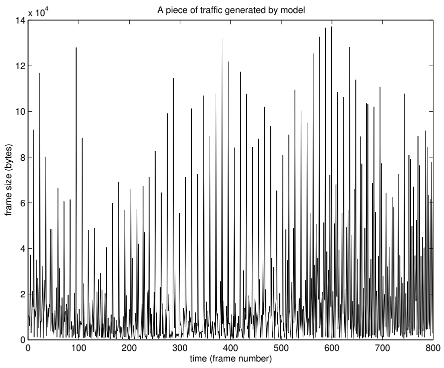


Figure 14 Traffic data generated by our model

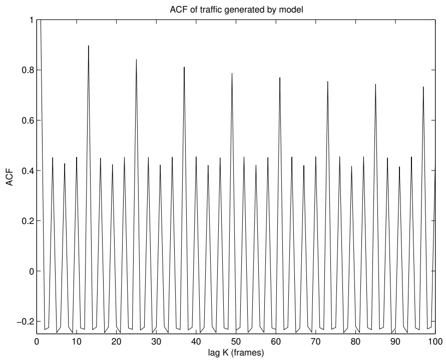


Figure 15 ACF of traffic data generated by our model

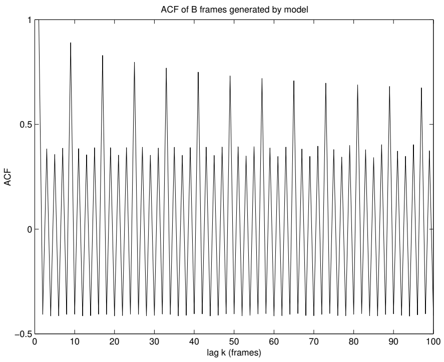


Figure 16 ACF of B frames generated by our model

the ACFs of the P, I, and B frame sequences very well. It can capture both the LRD and SRD. This model will play an important role in future network design and performance evaluation.

6. ACKNOWLEDGEMENT

The authors would like to further acknowledge M.W. Garrett for explaining their results.

References

[1] V. Paxson and S. Floyd, "Wide-area traffic: the failure of Poisson modeling," *Proc. SIGCOMM'94*, pp. 257–268, London, U.K., Aug. 1994.

[2] R. Addie, M. Zukerman and T. Neame, "Performance of a single server queue with self-similar input," *Proc. IEEE ICC'95*, volume 3, pp. 461–465, Seattle, WA, 1995.

[3] N. G. Duffield and N. O'Connell, "Large deviations and overflow probabilities for the general single server queue, with applications," DIAS-STP-93-30, Dublin Institute for Advanced Studies, 1993.

[4] C. Huang, M. Devetsikiotis, I. Lambadaris, and A. R. Kaye, "Fast simulation for self-similar traffic in ATM networks," *Proc. IEEE ICC' 95*, pp. 11–22, Seattle, June 1995.

[5] I. Norros, "A storage model with self-similar input," *Queueing Systems*, vol. 14, pp. 387–396, 1994.

[6] B. K. Ryu and A. Elwalid, "The importance of long-range dependence of VBR video traffic in ATM traffic engineering: myths and realities," *Proc. ACM Sigcomm'96*, pp. 3–14, Stanford University, CA, Aug. 1996.

[7] B. Melamed and D. E. Pendarakis, "Modeling full-length VBR video using markov-renewal-modulated TES models," *IEEE/ACM Trans. Networking*, vol. 5, pp. 600–612, Feb. 1998.

[8] M. Krunz and A. M. Makowski, "Modeling video traffic using $M/G/\infty$ input processes: a compromise between Markovian and LRD models," *IEEE J. Selected Areas in Comm.*, vol. 16, pp. 733–749, June 1998.

[9] D. P. Heyman and T. V. Lakshman, "Source models for VBR broadcast-video traffic," *IEEE/ACM Trans. Networking*, vol. 4, pp. 40–48, Feb. 1996.

[10] M. Krunz and S. K. Tripathi, "On the characterization of VBR MPEG streams," *Proc. SIGMETRICS'97*, pp. 192–202, Cambridge, MA, June 1997.

[11] B. Tsybakov and N. Georganas, "On self-similar traffic in ATM queues: definitions, overflow probability bound, and cell delay distribution," *IEEE/ACM Trans. Networking*, vol. 5, pp. 397–408, 1997.

[12] G. J. Hahn and S. S. Shapiro. *Statistical Models in Engineering*. John Wiley & Sons Inc., New York, 1967.