5

On the Relevance of Long Term Correlation in MPEG-1 Video Traffic

Marco Conti, Enrico Gregori, Andreas Larsson
CNUCE, Institute of National Research Council
Via S. Maria 36 - 56126 Pisa - Italy, Phone: +39-50-593111
Fax: +39-50-589354, e-mail: man@cnuce.cnr.it

Abstract

Variable Bit Rate, *VBR*, video is expected to become increasingly important with the large scale deployment of Broadband-Integrated Services Networks, *B-ISDN*s, over the next few years. Although the modeling of VBR video sources has recently received significant attention, there is currently no widely accepted model which lends itself to mathematical analysis. Furthermore, new video compression standards, such as the *MPEG* family, are emerging. On the basis of results of a detailed statistical analysis of a long sample of a movie encoded with the MPEG-1 algorithm, an analytically tractable model is developed and analyzed in detail. The model is able to capture both the distributional and temporal characteristics of this kind of traffic. The model was validated using a two-hour long sequence generated by the MPEG coding of the movie "Star Wars". We show that our model is a flexible tool to study network issues such as bandwidth allocation and statistical multiplexing.

Keywords

VBR video, MPEG, Markov chain, simulation, ATM, statistical multiplexing

1 INTRODUCTION

Recent technological advances in fiber optics and switching systems have provided the technological basis for the development of high capacity Broadband-Integrated Services Digital Networks (B-ISDNs), which are capable of supporting transmission speeds of several hundred Mbps [1]. This enormous potential for fast and massive information transport should be able to support not only the traditional data and voice services, but also a variety of new applications, including the transport of

Work carried out in the framework of CNR coordinated projects. Andreas Larsson's research was performed while visiting the CNUCE institute in the framework of the ERASMUS project. He is now with Ericsson Hewlett-Packard Telecommunications AB.

L. Mason et al. (eds.), *Broadband Communications*

images, teleconferencing, moving video, and large volumes of interactive computer data. Asynchronous Transfer Mode (ATM) is the transfer technique for the implementation of such B-ISDNs, due to its efficiency and flexibility [1].

Variable Bit Rate (*VBR*) video is by far the most interesting and challenging application. VBR video traffic is highly variable and dependent on the coding scheme adopted and the activity of the movie. A variable bit rate encoder attempts to keep the quality of video output constant at the price of changing the bit rate. A better utilization of network resources is also obtained, since only the real amount of information has to be transferred.

While the modeling of VBR video sources has recently received significant attention [1, 2, 3, 9, 10, 11, 12, 13], there is currently no widely accepted model which lends itself to mathematical analysis. Furthermore, while most of the previous studies focus on teleconferencing traffic, the *MPEG* standards for VBR encoding of moving images are currently under development [6, 7], and there is no proof that previous models can also be used for characterizing *MPEG* traffic (e.g. movies coded with the *MPEG* algorithm). This work focuses on the analysis and characterization of the traffic generated by an MPEG-1 encoder. Specifically, Section 2 presents the characteristics of an MPEG-1 source relevant for our investigation. An MPEG-1 analytical model is presented and validated in Section 3. The relevance of the long term correlation in MPEG-1 modeling is pointed out in Section 4.

2 MPEG-1 VIDEO SOURCE

MPEG-1 is a specification for coding moving pictures, developed by the ISO Joint Motion Pictures Experts group. MPEG-1 is an interframe coder. Coders in this class exploit, in addition to intraframe coding, the temporal redundancy between adjacent frames by predicting the next frame from the current one. A key feature that distinguishes MPEG-1 from previous coding algorithms is *bidirectional temporal prediction*. For this type of prediction, some of the frames are encoded using two reference frames, one in the past and one in the future. This results in higher compression gains.

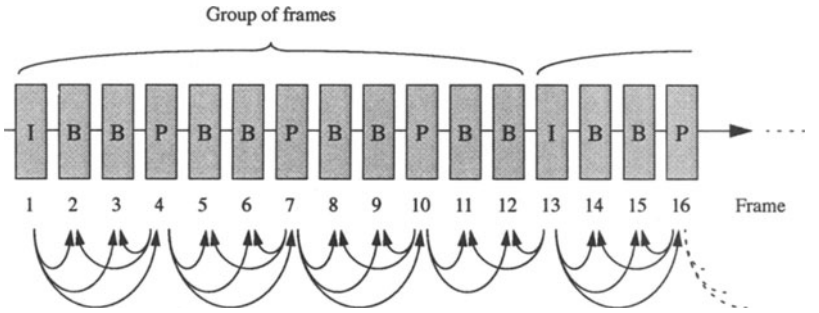

Figure 1 A sequence of MPEG-1 frames and their relationship.

As indicated above, when applying MPEG-1 to video, one of three different coding modes can be used for each frame. The terminology used for the resulting frame is related to the mode used and is as follows:

- *I-frame*: intra frame coded.

- *P-frame*: predictive coded with reference to a past picture.
- *B-frame*: bidirectional predictive coded.

I-frames provide access points for random access but only with moderate compression. Predictive coded frames are also generally used as a reference for future P-frames. Type B frames provide the highest amount of compression but require both a past and future reference prediction. In addition B-frames are never used as reference frames.

In the encoded sequence, the frames are arranged into *groups*, as shown in Figure 1. In this case a group consists of 12 frames - one I-frame, three P-frames and eight B-frames. Figure 1 also shows the relationship between the frames. We can see that I-frames are independent, P-frames are predicted, and B-frames bidirectionally predicted.Statistics of MPEG-1 coded movie

This section presents a statistical analysis of an MPEG-1 encoded movie. The source for the analysis is a bit per frame trace, released by M. Garret at Bellcore, obtained from an MPEG-1 encoder fed with approximately two hours of the movie "Star Wars". Specifically, frames are coded in groups of twelve frames as defined in Figure 1 (i.e. the frame pattern is IBBPBBPBBPBB).

Several basic statistics of the MPEG-1 coder trace are shown in Table 1. The table contains statis-

Table 1 Basic statistics for the MPEG-1 "Star Wars" movie

Measure	Original Sequence	Aggregate Sequence
Mean bandwidth, μ	15598 bits/frame	187185 bits/group
Standard deviation, σ	18165 bits/frame	72468 bits/group
Coefficient of variation, μ/σ	1.16	0.39
Peak bandwidth	185267 bits/frame	932710 bits/group
Minimum bandwidth	476 bits/frame	77754 bits/group
Peak/mean bandwidth	11.88	4.98

tics related to the following sequences

- *original sequence*: the sequence of the number of bits per frame;
- *aggregate sequence:* the sequence of the number of bits per group i.e. IBBPBBPBB-PBB.

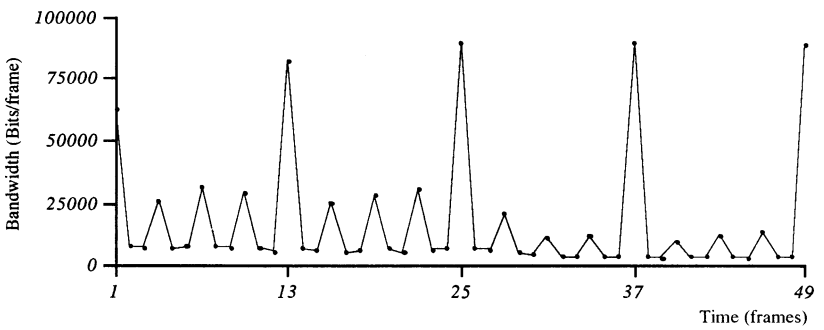

Figure 2 Part of the MPEG-1 coder trace, revealing group length and frame pattern.

As shown in Figure 2 the bandwidth varies greatly due to the different frame-types, I, P and B. An MPEG encoder can therefore be seen as a source generating three different kinds of traffic, each

according to the individual characteristics of a frame-type. In order to understand the level of inter-dependence between these three subsequences, we measure the correlation $Corr\,[x(k), y(k)]$, where $x(k)$ and $y(k)$ are the kth values in each sequence. Specifically, to use the correlation for our purposes we let k correspond to the kth group in the original sequence. Within each group, we sum the number of bits generated for each kind of frame. The correlation is then computed among the three different sequences, i.e., $I(k)$, $\sum P(k)$ and $\sum B(k)$. The resulting correlation values of the

Table 2 Correlation between frame-types.

$Corr\,[I(k), \sum P(k)]$	0.35
$Corr\,[I(k), \sum B(k)]$	-0.13
$Corr\,[\sum P(k), \sum B(k)]$	0.71

MPEG-1 coder trace are summarized in Table 2. They tell us that dependencies exist between the three sequences, and therefore they cannot be represented with three independent processes.

3 MPEG-1 MODELING

As indicated by the statistical analysis presented in the previous section, the output of an MPEG-1 encoder should be described by three partially correlated submodels where each submodel describes the output process corresponding to one frame-type. Obviously this leads to a model with a very large state space.

The model space complexity is reduced by avoiding a separate representation for the different frame-types. Specifically, this is obtained by considering a different time scale in which the time unit is the group (i.e. a sequence IBBPBBPBBPBB) and the bit rate per time unit is the sum of the amount of bandwidth generated by all the frames in a group. In this case one group is equal to 12 frames and each frame is generated every 1/24th second. The resulting sequence is hereafter named *aggregate sequence.*

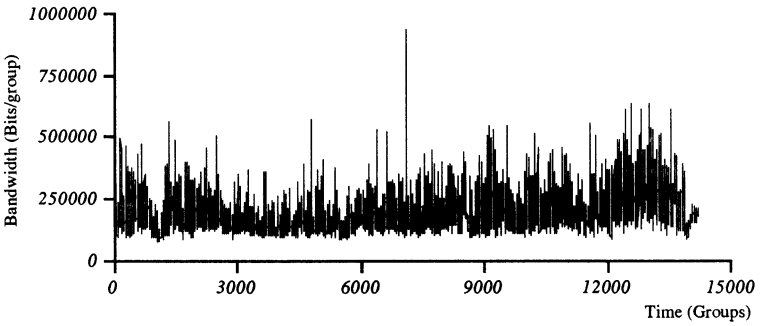

Figure 3 The aggregate sequence.

Figure 3 shows a plot of the *aggregate sequence* generated by an MPEG-1 coder with the movie Star Wars as a source. A time unit, on the x-axis, is equal to 0.5 seconds, i.e. a group interarrival.

The bit rate of consecutive frames shows that the bandwidth changes in a rapid but bursty way. However, there is also a slowly changing underlying structure. This *low frequency* underlying struc-

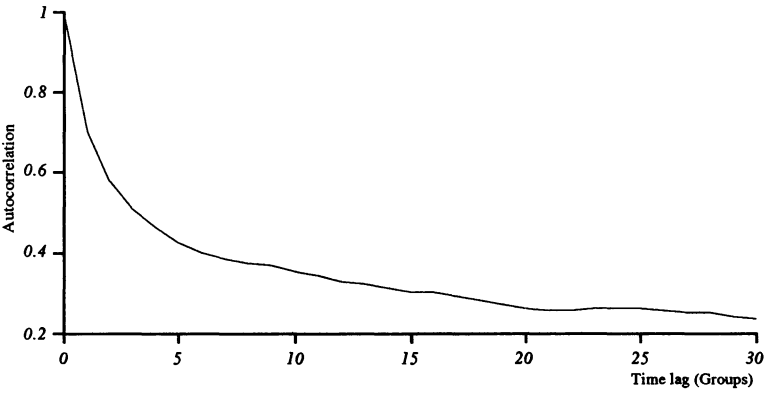

Figure 4 Short range dependencies

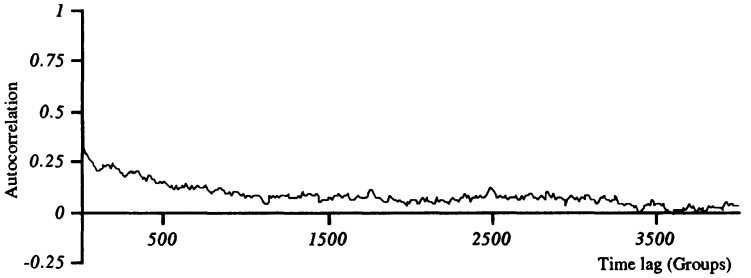

Figure 5 Long range dependencies

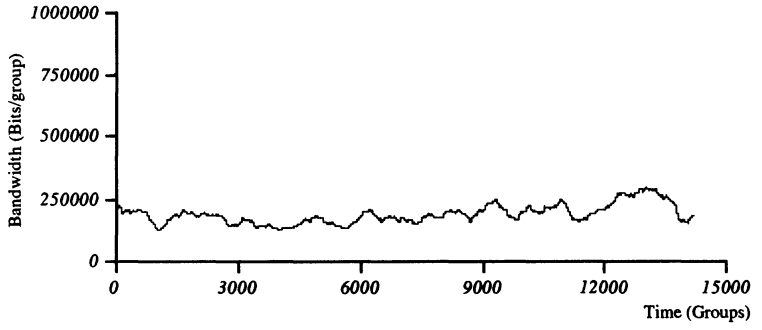

Figure 6 Low frequency component of the aggregate sequence.

ture of the sequence can be better highlighted by passing the aggregate sequence through a moving average filter of length *W*. The result with a *window size* W=300 groups (i.e. two and a half minutes), is shown in Figure 4

The autocorrelation function for the aggregate sequence is plotted in Figures 5 and 6, showing the short $(0 \le n \le 30)$ and long range $(n > 30)$ dependencies, respectively. Figure 5 shows the existence of a strong short-range dependence for time lags below approximately 30 groups, which corresponds to 15 seconds. In this range the autocorrelation function drops quickly (the autocorrelation with $n = 30$ is about 0.2). However, after this sharp initial decrease, as shown in Figure 6, it takes a very long time before the autocorrelation function drops to zero. Specifically, Figure 6 highlights the existence of a significant long-range dependence which lasts for time lags of up to 3500 groups, i.e. about 29 minutes. The tail of the autocorrelation function decreases slowly.

3.1 Description of the Model

Figures 5 and 6 show that in the aggregate sequence there are both short-range dependencies which last for around 20-30 groups (some seconds), and long-range dependencies which last for thousands of groups (some minutes). In order to capture both types of dependencies a bidimensional Markov chain $\{L_k, H_k, k \ge 0\}$ is used, in which $\{L_k | k \ge 0\}$ is used to represent the long term correlation, while $\{H_k | k \ge 0\}$ represents the short term correlation. Specifically, in our model the process $\{H_k | k \ge 0\}$ describes the bit rate per group of an MPEG-1 encoder. To avoid unnecessary complexity (in the state space $\{H_k | k \ge 0\}$) we quantize the bit rate information into a number of levels. The number of quantization levels for the process $\{H_k | k \ge 0\}$ will hereafter be denoted by N (i.e. $H_k \in \{0, 1, 2, \ldots, N-1\}$). The question of which quantization method should be used is not discussed here. For us it seemed natural to use uniform quantization. For this reason, let *max* and *min* denote the maximum and minimum bit rates observed in the aggregate sequence. The possible bit rates between *max* and *min* are quantized with a constant step size $\Delta = (max - min)/N$, resulting in the actual bit rate of the source equal to $j \cdot \Delta + min$ where j is the quantization level holding the property $0 \le j \le N-1$.

To represent the low-frequency component of an MPEG source, a modulating process $\{L_k | k \ge 0\}$ is included in the model as well $(L_k \in \{0, 1, 2, \ldots, M-1\})$.

How is L_k controlled? To this end let the random variables X_i denote the time in which the low-frequency process has stayed in level i. If the process has stayed in a level i for t-1 slots, there is the probability of leaving i at slot t equal to

$$P(X_i = t | X_i > t-1) = 1 - q_i(t) \tag{3.1}$$

and a probability of staying in i more than t equal to $q_i(t)$. Hence to control the transitions in the $\{L_k\}$ process, the variable t is added to the Markov chain state, and the transitions between the low frequency levels are managed according to (3.1).

It is still impossible to tell which distribution the "real" X_i has to be. The general approach would entail estimating the mass function from the real sequence. Since we only have a single trace, the estimated statistics may be biased. To minimize the biasing effect, in this work we assume a Geometric distribution for the duration of a low frequency state which only entails estimating a single parameter, i.e., the average time the $\{L_k\}$ spends in each state. The model is thus named *Geometric Model.*

The process we now want to model takes the form $\{L_k, H_k, k \ge 0\}$, where $L_k \in \{0, \ldots, M-1\}$ is the status of the low frequency process corresponding to the kth group, and $H_k \in \{0, \ldots, N-1\}$ is the corresponding state in the high frequency process.

The source is modelled with a Markov chain whose transition probabilities are

$$p_{ij,lm} = p^{(1)}{}_{ij,lm} = P(L_k = l, H_k = m | L_{k-1} = i, H_{k-1} = j). \tag{3.2}$$

The procedure for fitting it to a real source (i.e., to construct the transition probabilities, $p_{ij,lm}$,

of the Markov chain starting from a real source) is shown in [4].

3.2 Model Analysis

In this section we analyse the characteristics of the model to see whether it can imitate the behavior of the real source well.

To compute the statistics of our model (e.g., average, variance, peak/average ratio, autocorrelation function) we first need to compute the steady-state probabilities $\pi_{i,j} = P(L_k = i, H_k = j)$ of our Markov chain.

Several basic statistics which can be immediately calculated from the steady-state probabilities are shown in Table 3. The results were obtained with $M = 8$ and $N = 8$.[1] We see that the values coincide well.

Table 3 Comparison of some basic statistics for the real source and the model.

Measure	*Real source*	*Model*
Mean bandwidth level, μ	0.526	0.527
Standard deviation, σ	0.712	0.712
Coefficient of variation, σ/μ	1.354	1.351
Peak bandwidth level	7	7
Minimum bandwidth level	0	0
Peak/mean bandwidth level	13.307	13.283

We will now examine whether the model is successful in capturing the time correlation structure of our source. The autocorrelation function $r(n)$ of the model is thus needed:

$$r(n) = \frac{E[H_k H_{k+n}] - \mu^2}{\sigma^2} \quad . \tag{3.3}$$

Figure 7 shows a plot of $r(n)$ for the real source and four different models constructed with $M = 8$, $N = 8$ and various moving average window lengths, W. Time lags used for the calculation range from 0 to 30 groups. The plot thus compares the short-range dependence of the real source, and different parametrizations of the model.

The model constructed with W=20, has a stronger short-range dependence than the real source. It has, however, a faster decay. Even though $r(n)$ of the model is still above the real source one for a time lag equal to 30 groups, the difference is smaller than for $n = 10$.

As the value of W is increased, the autocorrelation function of the model tends to fall off at the beginning but it decreases slower. The model with $W = 40$ is a good example to emphasize this behavior. For n less than 7, the short-range dependence of the model takes on values lower than the real source. For time lags beyond this point, the plot shows that $r(n)$ of the model decay slower than for the real source. The autocorrelation functions for models constructed with $W = 60$ and $W = 80$ follow the same pattern.

We know that the long-range dependence of MPEG-1 coded VBR video is very strong. Figure 8 compares the autocorrelation function of the model and the real source for time lags of 0 up to 2000 groups. Several values of W have been used. A model constructed with $W = 100$ has a long-range dependence which is not as strong as that of the real source for n greater than 100. It also reaches zero at a time lag approximately equal to 1100, which is much earlier than in the real source.

1. The choice of the value 8 is shown in [4] to be a good compromise between accuracy of the model and complexity of the state space in the Markov chain.

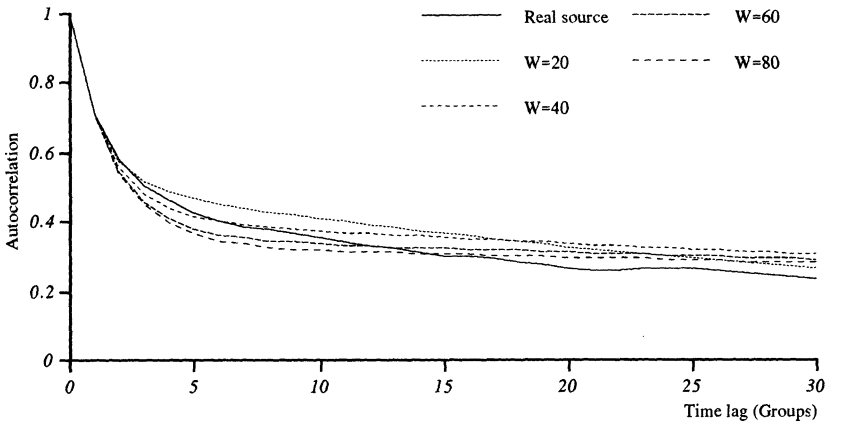

Figure 7 Comparison of the real source's and the model's short-range dependencies.

$r(n)$ for the other models plotted in Figure 8 tells us that the long-range dependence of the model tends to get stronger, and thus approaches the real source, as W is increased. For example, a model constructed with $W = 300$ matches the long-range dependence of the real source better than if it is created with $W = 200$. At the same time we know, from the previous subsection, that a higher value of W implies a weaker short-range dependence.

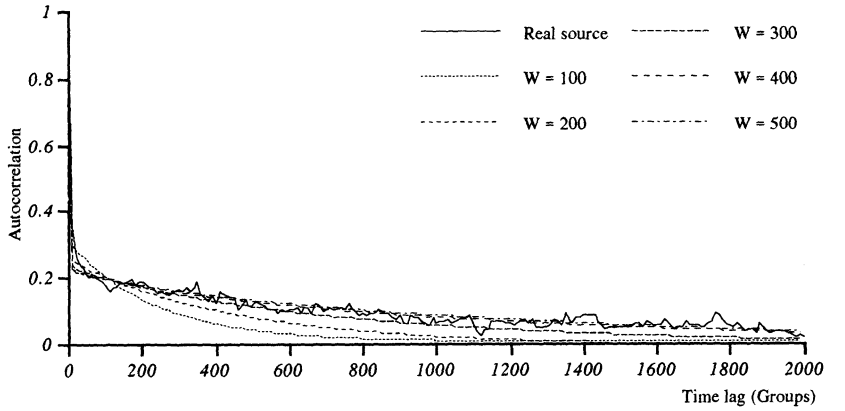

Figure 8 Comparison of the real source's and the model's long-range dependencies.

We can make the following conclusions from the model analysis:

- Distributional properties of a source, i.e. steady-state probabilities and basic statistics, are captured well by the model.
- The parameter W can be used to construct source models with different time correlation structures. The rules to follow are: stronger short- and weaker long-range depend-

encies for lower W-values. On the other hand, weaker short-range and stronger long-range dependencies are obtained for higher W-values.

Validation of the model is currently underway by exploiting the new traces which have been recently released [14]. Preliminary results of this validation process [5] show that the model can be tuned to fit different kinds of MPEG sources (e.g. movies, sports events, talk shows).

4 IMPORTANCE OF THE LONG-TERM DEPENDENCIES

At the moment ATM networks assign peak rate bandwidth to real time applications, that is, by avoiding multiplexing and thus utilizing the residual bandwidth for non real time traffic. This implies that high priority must be assigned to real time applications. Data services, on the other hand, tolerate delays and can compensate for loss by retransmission. Low priority can thus be assigned to these applications. In Section 4.1 the effect of an MPEG-1 source on the quality of service experienced by low priority traffic is investigated.

However, since the ratio peak/average for VBR video traffic may be greater than four, the peak rate approach for bandwidth allocation significantly reduces the number of VBR applications that can concurrently use the network. For this reason the peak rate allocation will be abandoned as soon as the problem of characterizing and multiplexing the traffic generated by VBR video is addressed successfully. In Section 4.2 the potential gains which can be achieved by multiplexing several MPEG video sources are investigated.

4.1 Peak rate allocation

We assume a peak rate allocation for VBR video and we consider a network with two categories of traffic: i) VBR video which has high priority and peak rate allocation; ii) data traffic which is transmitted utilizing the bandwidth reserved but not used by VBR video traffic (low priority traffic).

We consider the system which consists of a VBR video source, a data source, a priority queueing system and a transmission channel, or *server*. The server has the capacity C, to transfer a certain amount of information per time unit. Two queues, each associated with one class of traffic, are used to buffer arriving information from the sources when the server's capacity is exceeded. Items from the buffers are transmitted according to their priority. Inside each queue, transmission is first in first out (*FIFO*).

The VBR video source is described by the Geometric model. For each time slot, equal to the length of a group, the VBR video source generates traffic corresponding to one of the eight levels (between 0 and 7) of the aggregate bit rates of the MPEG-1 coded movie Star Wars. Thus for each time slot the VBR video requires between 1 and 8 *units* of bandwidth. The channel has a static allocation which implies that C is equal to 8 units/group.

Data is generated according to a Poisson process with rate λ. On average data can use a percentage p of the units not used by video. This means that the average number of data units in a group is equal to $\lambda = p(7 - \mu)$ units/group, where μ is the average bandwidth level for the video source. Hereafter, p is set to 0.9, and therefore $\lambda = 0.9 \cdot (7 - 0.53) = 5.823$ units/group ($\mu = 0.53$ is obtained from Table 1).

Three simulations with different sources representing the VBR video were run. Simulation I was a trace-driven simulation with the quantized aggregate sequence of the "Star Wars" movie. Using a real trace in a simulation has some drawbacks. First of all, using only one realisation does not make it possible to draw any conclusions on the accuracy of the simulation. The behavior of the resulting video traffic is not general either, and thus not very representative. Handling a long trace is, furthermore, impractical. We do, however, include the trace-driven simulation here to have some sort of

comparison with the other results.

In simulations II and III we used the Geometric instance model as a source, first constructed with W=20 and then with W=300. Simulative estimates are obtained with a confidence level of 90%.

We will now look at the simulation results. Specifically, we investigate the mean queue length μ_{ql} and mean queueing time μ_{qt}, in relation to the size of the low priority queue, and the time that data has to wait there, respectively. For the case where the real trace is used, i.e. simulation I, we can obtain meaningful estimates only for the first moments of queue length and queueing time. This is due to the fact that we only have one realisation of such a real trace. Table 4 compares these values obtained from simulations I, II and III. The results show that only estimates obtained for the long-range model are close to those obtained from real data.

Table 4 Data queue-length and queueing time

Measure	*Real trace*	*Model, W=20*	*Model, W=300*
μ_{ql}	21.77	11.56	23.01
μ_{qt}	29.96	15.88	31.63

In simulations II and III, on the other hand, we used a number of realisations which made accurate estimations on the queue length and queueing time. Figure 9 highlights that the long-range dependence impact on queueing time is significant. For example, as shown in Figure 9, when W=300 is used for high-priority traffic model the 99th percentile of the queueing time is about three times higher than that obtained with W=20.

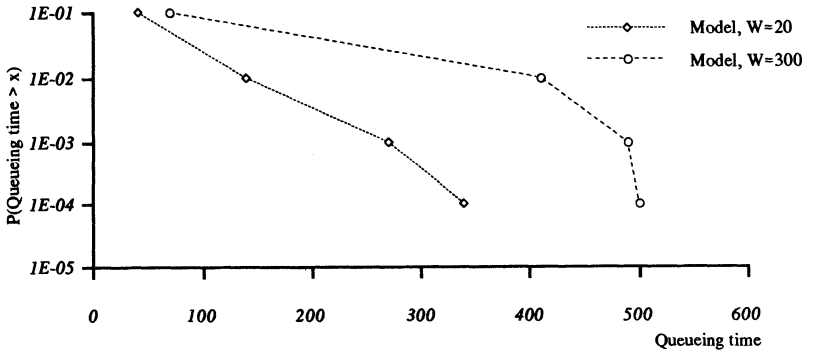

Figure 9 Data queue-length tail distribution for the two models.

The only difference between the two models studied here is their temporal behavior. Furthermore, as shown in Table 4, the model with long-range dependence was shown to be more similar to the real trace. Consequently, as a conclusion from the simulation results, one can say that by neglecting the long-range correlation one would obtain a rather optimistic estimate on queue length and queueing time.

4.2 Multiplexing of i.i.d sources

Mutliplexing of VBR video sources is complex, as these applications have low tolerance towards network congestion. Although sufficient buffer capacity may be available, excessive buffering may

not be possible, due to the resulting unacceptable delays. In this section we therefore investigate the queueing time distribution experienced by VBR video traffic as a function of the bandwidth reserved for each source. As shown before (see Table 1), the peak rate for our MPEG source corresponds to a bandwidth level equal to c=7, while the average is about bandwidth level equal to 0.53. Below we

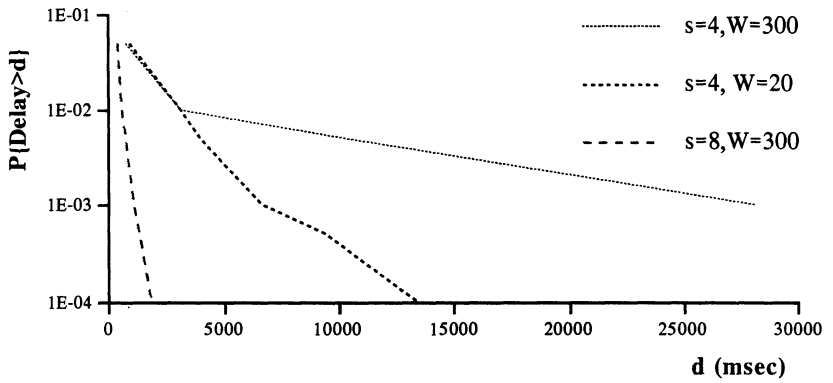

Figure 10 Tail of delay distribution, c=1.0

investigate the delay experienced by VBR video traffic by assuming that the bandwidth allocated for each source is about twice the average, i.e., c=1. The results reported in Figure 10 were obtained by studying via simulation the queueing delay distribution in a single server queueing system with a deterministic service time, FIFO, and input traffic generated by s independent and identically distributed MPEG-1 sources. Figure 10 shows that a 75% network utilization and acceptable delays can be achieved if at least eight sources are multiplexed. In addition, the figure clearly indicates that the tail estimated with W=20 is extremely underestimated in the region (1E-04,1E-02).

Acknowledgements

The authors wish to express their gratitude to the Mark Garret (Bellcore) for releasing the Star Wars trace.

References

[1] R. O. Onvural, *Asynchronous Transfer Mode Networks: Performance Issues*, Artech House, Inc., Norwood, MA, 1994.

[2] N. Ohta, *Paket Video: Modeling and signal processing*, Artech House, Inc., Norwood, MA, 1994.

[3] D.Heyman, A.Tabatabai and T. LAKshaman, "Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 2, No. 1, March. 1992, pp. 49-59.

[4] A. Larsson, "Traffic Engeneering of Broadband Networks: Modelling of MPEG 1 coded VBR Video", Master's Thesis, Linkoping University 1995, performed at CNUCE, Pisa, Italy, in the framework of the ERASMUS project.

[5] M. Conti, E. Gregori, M. Nava, "Analysis and Modeling of an MPEG-1 Video Source", Third IFIP Workshop on Performance Modelling and Evaluation of ATM Networks, 2nd-6th July, 1995, Ilkley, West Yorkshire, U.K.

[6] D. Le Gall, MPEG: A Video Compression Standard for Multimedia Applications, *Communications of the ACM*, Vol. 34, No. 4, April 1991, pp. 46-58.

[7] L. Chiariglione,"The development of an integrated audiovisual coding standard: MPEG", *IEEE Proceeding*,Vol. 83, No. 2, February 1995, pp. 151-157.

[8] M. Nomura, T. Fujii and N. Ohta, "Basic Characteristics of Variable Rate Video Coding in ATM Environment", *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 5, June 1989, pp. 752-760.

[9] G.Ramamurthy, B. Sengupta, "Modeling and Analysis of a Variable Bit Rate Video Multiplexer", Proceedings of IEEE INFOCOM, Florence 1992, pp.817-827.

[10] P. Skelly, M. Schwartz and S. Dixit, "A Histogram-Based Model for Video Traffic Behavior in an ATM Multiplexer", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 4, Aug. 1993, pp. 446-459.

[11] M. R. Frater, P. Tan and J. F. Arnold, "Variable Bit Rate Video Traffic on the Broadband ISDN: Modeling and Verification", *ITC 14*, 1994, pp. 1351-1360.

[12] M. W. Garret, *Contributions Toward Real-Time Services on Packet Switched Networks*, Ph.D. Dissertation CU/CTR/TR 340-93-20, Columbia University, New York, N.Y., May 1993.

[13] M. W. Garret,W. Willinger., "Modeling and Generation of Self-Similar VBR Video Traffic", *SIGCOMM'94*, London, Sept. 1994, pp.269-280.

[14] O.Rose, " Statistical Properties of MPEG Video Traffic and Their Impact on Traffic Modeling in ATM Systems", University of Wurzurg, Research Report No.101, February 1995.