Available online at www.sciencedirect.com
ScienceDirect
Performance Evaluation 65 (2008) 51–70
PERFORMANCE EVALUATION
An International Journal
www.elsevier.com/locate/peva

A new model for video traffic originating from multiplexed MPEG-4 videoconference streams

Aggelos Lazaris[a], Polychronis Koutsakis[b,*], Michael Paterakis[a]

[a] Department of Electronic and Computer Engineering, Technical University of Crete, Greece
[b] Department of Electrical and Computer Engineering, McMaster University, Canada

Received 15 May 2005; received in revised form 5 February 2007
Available online 17 February 2007

Abstract

Due to the burstiness of video traffic, video modeling is very important in order to evaluate the performance of future wired and wireless networks. In this paper, we first study the behavior of single MPEG-4 videoconference traces and investigate the possibility of modeling this type of traffic with well-known distributions. Our results show that the Pearson type V distribution is the best fit among all the examined distributions, for all the traces under study. However, the behavior of single videoconference traces can never be perfectly "captured" by a distribution generating independently frame sizes according to a declared mean and standard deviation, due to the high autocorrelation of videoconference; therefore none of the fitting attempts can achieve high accuracy. Still, our results on attempting to model single MPEG-4 videoconference sources provide significant insight and help to build a Discrete Autoregressive (DAR(1)) model to "capture" the behavior of *multiplexed MPEG-4 videoconference movies* from VBR coders. Based on our results and on comparisons with other existing approaches, we discuss the contribution of our proposed method to the field.

Keywords: Videoconferencing; MPEG-4 video encoding; Video traffic modeling; Discrete autoregressive model

1. Introduction

The popularity of video streaming over the Internet is continuously growing, with hundreds of new subscribers registered daily. In addition, existing and emerging wireless systems such as EGPRS, UMTS, CDMA-2000 and WLAN enable multimedia transmission and reception at any place and time at reasonable and sufficient data rates; video transmission for mobile terminals is likely to be a major application in future mobile systems and may be a key factor to their success [1].

Therefore, as traffic from video services is expected to be a substantial portion of the traffic carried by emerging wired and wireless networks, statistical source models are needed for Variable Bit Rate (VBR) coded video in order to design networks which are able to guarantee the strict Quality of Service (QoS) requirements of the video traffic. Video packet delay requirements are strict, because delays are annoying to a viewer; whenever the delay experienced

* Corresponding author. Tel.: +1 9055259140x26542; fax: +1 9055212922.
E-mail addresses: alazaris@telecom.tuc.gr (A. Lazaris), polk@ece.mcmaster.ca (P. Koutsakis), pateraki@telecom.tuc.gr (M. Paterakis).

0166-5316/$ - see front matter
doi:10.1016/j.peva.2007.02.004

by a video packet exceeds the corresponding maximum delay, the packet is dropped, and the video packet dropping requirements are equally strict. There are three areas where single video source models are useful [2]:

a. Studying what types of traffic descriptors are needed for parameter negotiation with the network at call setup, as the source model will be crucial in determining if, under the current load, an additional source with certain traffic characteristics and service requirements may be accepted or not [3–5],
b. testing rate control algorithms and
c. predicting the quality of service degradation caused by congestion on an access link.

Hence, the problem of modeling video traffic, in general, and videoconferencing, in particular, has been extensively studied in the literature. VBR video models which have been proposed in the literature include first-order autoregressive (AR) models [6], discrete AR (DAR) models [2,7], Markov renewal processes (MRP) [8], MRP transform-expand-sample (TES) [9], finite-state Markov chain [10,11], and Gamma–beta-auto-regression (GBAR) models [12,13]. The GBAR model, being an autoregressive model with Gamma-distributed marginals and geometric autocorrelation, "captures" data-rate dynamics of VBR video conferences well; however, it is not suitable for general MPEG video sources [13].

In [14], various differences in successive video frame sizes of VBR video traffic were investigated, while in [15] packet generation intervals, at various levels of video activity, were studied. In [7,16] the authors show that H.261 videoconference sequences generated by different hardware coders, using different coding algorithms, have gamma marginal distributions (this result was also employed by [17], which proposes an Autoregressive Model of order one for sequences of H. 261 encoding) and use this result to build a Discrete Autoregressive (DAR) model of order one, which works well when several sources are multiplexed.

In [18–20], different approaches are proposed for MPEG-1 traffic, based on the lognormal, Gamma, and a hybrid Gamma/lognormal distribution model, respectively. Standard MPEG encoders generate three types of video frames: I (intracoded), P (predictive) and B (bidirectionally predictive); i.e., while I frames are intra-coded, the generation of P and B frames involves, in addition to intra-coding, the use of motion prediction and interpolation techniques. More specifically, an I frame uses only transform coding and provides a point of access to the compressed video data. A P frame uses motion-compensated prediction from the most recent previous I or P frame. I frames and P frames are called anchor frames because they are used to predict other frames. As P frames use information already transmitted in previous anchor frames, their size (number of bits required for representation) can be much less than that of an Iframe. B frames are coded based on both past and future I or P frames, offering the greatest opportunity for data compression; the size of a B frame is typically about an order of magnitude smaller than that of an I frame [13]. In synopsis, I frames are, on average, the largest in size, followed by P frames and then by B frames.

An important feature of common MPEG encoders (both hardware and software) is the manner in which frame types are generated. Typical encoders use a fixed Group-of-Pictures (GOP) pattern when compressing a video sequence; the GOP pattern specifies the number and temporal order of P and B frames between two successive I frames. A GOP pattern is defined by the distance N between I frames and the distance M between P frames. In practice, the most frequent value of M is 3 (two successive B frames) while the most frequent values of N are 6, 12, and 15, depending on the required video quality and the transmission rate.

Generally, as analyzed in [19], all the video modeling studies presented above can be classified into two categories: (a) data-rate models, and (b) frame-size models.

In a data-rate model, only the rate at which data are arriving at a link is generated for performance prediction purposes. Almost all models, including AR, DAR, MRP, MRP TES and the GBAR model, fall under this category. These models achieve good and often very good results in predicting average packet-loss probability and ATM buffer overflowing probability. However, they have the shortcoming of failing to identify such details as the percentage of frames affected, as even a small rate of data loss involving I frames may affect perceptual quality of received video significantly, but the same amount of data loss in B frames would have far less impact.

In a frame-size model, sizes of individual MPEG frames are generated, and hence, data-rate information can be obtained from the frame-size information. The inherent frame-by-frame burst nature of MPEG videos is preserved in this category of models.

In this work, we focus on the problem of modeling videoconference traffic from MPEG-4 encoders, which is a relatively new and yet open issue in the relevant literature (all previously mentioned references to MPEG modeling research efforts addressed MPEG-1 and MPEG-2 modeling). The MPEG-4 standard is particularly designed for video streaming over wireless networks [21,42]. We use four different long sequences of MPEG-4 encoded videos and we show that the use of the Gamma and lognormal distributions (which are considered the best choice for modeling many types of video traffic and especially the Gamma distribution is the basis for many of the above-mentioned models of the literature), *is not* the most appropriate for MPEG-4 videoconference traffic. We show that, for modeling *single* videoconference sources, the best choice among all the examined distributions is the Pearson type V distribution. However, the high autocorrelation characteristic of videoconference traffic can never be perfectly "captured" by a distribution generating independently frame sizes according to a declared mean and standard deviation, and therefore none of the fitting attempts (including the Pearson V), as good as they might be, can achieve perfect accuracy. For this reason, we extend our work in order to build models which "capture" well the behavior of *multiplexed* MPEG-4 videoconference movies, by generating frame sizes independently for I, P and B frames (i.e., we build a frame-size model). *Our work follows the steps of the work conducted by Heyman* et al. *in* [7,16] *in order to build Discrete Autoregressive (DAR) models of order one*. The work in [7,16], as already mentioned earlier in this section, focused on video traffic originating from previous technology encoders (H.261); MPEG-4 traffic has quite different characteristics from H.261 and H.263 traffic (newer H.26x technology encoding), but also from MPEG-1 and MPEG-2 traffic. However, our results agree with those in [7,16] in that it is shown that the DAR models work well when several sources are multiplexed (which is most often the case in both wired and wireless networks). Of course, as will be shown in the description of our approach, the different nature of MPEG-4 videoconference traffic compared to H.261 traffic demands the use of three DAR models instead of one, as in [7,16].

A brief reference to the MPEG-4 standard and the most important differences of MPEG-4 encoding with H.261, H.263, MPEG-1 and MPEG-2 encoding follows.

H.261 was targeted for teleconferencing applications where motion is naturally more limited, and therefore H.261 motion vectors' accuracy is reduced in comparison to MPEG. Also, H.261 encoding does not use B frames. The coding algorithm of H.263 is similar to that used by H.261; however, with some improvements and changes to improve performance and error recovery. H.263 supports five resolutions: in addition to QCIF and CIF that were supported by H.261 there is SQCIF, 4CIF, and 16CIF. SQCIF is approximately half the resolution of QCIF. 4CIF and 16CIF are 4 and 16 times the resolution of CIF respectively. The support of 4CIF and 16CIF means the codec could then compete with other higher bit-rate video coding standards such as the MPEG standards.

MPEG-1 is basically a standard for storing and playing video on a single computer at low bit-rates. It is focused on bit-streams of about 1.5 Mbps and for storage of digital video on CDs. The focus is on compression ratio rather than picture quality. It can be considered as traditional VCR quality, with the difference that it is digital instead of analog.

MPEG-2 is a standard for digital TV. It meets the requirements for HDTV and DVD (Digital Video/Versatile Disc). The MPEG-2 project focused on extending the compression technique of MPEG-1 to cover larger pictures and higher quality at the expense of a lower compression ratio and therefore also higher bandwidth usage. MPEG-2 also provides more advanced techniques to enhance the video quality at the same bit-rate.

The MPEG group initiated the new MPEG-4 standards in 1993 with the goal of developing algorithms and tools for high efficiency coding and representation of audio and video data to meet the challenges of videoconferencing applications. The standards were initially restricted to low bit-rate applications but were subsequently expanded to include a wider range of multimedia applications and bit-rates. The most important addition to the standards was the ability to represent a scene as a set of audiovisual objects. The MPEG-4 standards differ from the MPEG-1 and MPEG-2 standards in that they are not optimized for a particular application but integrate the encoding, multiplexing, and presentation tools required to support a wide range of multimedia information and applications. In addition to providing efficient audio and video encoding, the MPEG-4 standards include such features as the ability to represent audio, video, images, graphics, text, etc. as separate objects, and the ability to multiplex and synchronize these objects to form scenes. Support is also included for error resilience over wireless links, the coding of arbitrary shaped video objects, and content-based interactivity such as the ability to randomly access and manipulate objects in a video scene [24]. In comparison to MPEG-2, an MPEG-4 encoder achieves a bit-rate reduction by a factor of two or three without affecting the subjective video quality [22,40]. However, the bit-rate variability of the encoded streams and their statistics are very different from MPEG-2 streams, in particular when low bit-rate encoding is used [26]. These differences were a motivation for our work on designing a model capable of accurately reproducing MPEG-4 videoconference traffic.

To the best of our knowledge, the subject of modeling MPEG-4 videoconference traffic has been addressed in the relevant literature only in [42,43]. However, important work has been presented on modeling MPEG-4 encoded

Table 1
Statistics for the I, P, B frames of the ARD Talk trace

Trace	Mean bit rate (Kbps)	Peak bit rate (Mbps)	Standard deviation (Kbps)
Office cam	400	2	434
Lecture room cam	210	1.5	182
N3 Talk	550	3.4	329
ARD Talk	540	3.1	346

movies in [21–23]. The different characteristics of videoconference traffic (e.g., much higher autocorrelation than movies, different ratios in the sizes of P and B frames than movies, no significant changes between scenes) were also a motivation for our work on an efficient model for this type of traffic. After the detailed presentation of our modeling approach, we discuss the results of our proposed modeling method in respect to the results presented in [21–23,42,43], as we believe that this comparison can lead to significant conclusions.

The rest of the paper is organized as follows: Section 2 introduces our study on the possibility of modeling single videoconference traces based on a distribution generating frame sizes independently, and discusses its accuracy. Section 3 contains a brief explanation of the DAR(1) model and then proceeds to discuss its implementation in order to acquire an accurate model for multiplexed MPEG-4 videoconference streams. Section 4 includes the aforementioned comparison of our method and results with the approaches and results presented in [21–23,42,43]. Finally, Section 5 presents the conclusions of our study.

2. A study on single-source MPEG-4 videoconference modeling

2.1. Frame-size histograms

As already mentioned in the introduction, we use four different long sequences of MPEG-4 encoded videos (from [25,26]) with low or moderate motion (i.e., traces with very similar characteristics to the ones of actual videoconference traffic), in order to derive a statistical model which fits well the real data. The length of the videos varies from 45 to 60 min and the data for each trace consists of a sequence of the number of cells per video frame and the type of video frame, i.e., I, P, or B. We use packets of ATM cell size throughout this work, but our modeling mechanism can be used equally well with packets of other sizes. We have investigated the possibility of modeling the four videoconference videos with quite a few well-known distributions and our results show that the best fit among these distributions is achieved for all the traces studied with the use of the Pearson type V distribution. The Pearson type V distribution (also known as the "inverted Gamma" distribution) is generally used to model the time required to perform some tasks (e.g., customer service time in a bank); other distributions which have the same general use are the exponential, Gamma, Weibull and lognormal distributions [31]. Since all these distributions have been often used for video traffic modeling in the literature, they have been chosen as fitting candidates in order to compare their modeling results in the case of MPEG-4 videoconferencing.

We have used the high quality coding version of the four traces which were used in our study. The traces are, respectively:

1. A video stream extracted and analyzed from a camera showing the events happening within an office (Video Name: "Office Cam").
2. A video stream extracted and analyzed from a camera showing a lecture (Video Name: "Lecture Room Cam").
3. A video stream extracted and analyzed from a talk-show (Video Name: "N3 Talk").
4. A video stream extracted and analyzed from another talk-show (Video Name: "ARD Talk").

The trace statistics are presented in Table 1.

New video frames arrive every 40 ms, in Quarter Common Intermediate Format (QCIF) resolution. The compression pattern used to encode all the examined video streams is IBBPBBPBBPBB, i.e., $N = 12, M = 3$, according to the definitions used in Section 1.

The frame-size histogram based on the complete VBR streams is shown, for all four sequences, to have the general shape of a Pearson type V distribution (this is shown in Fig. 1, which presents indicatively the histogram for the lecture sequence).

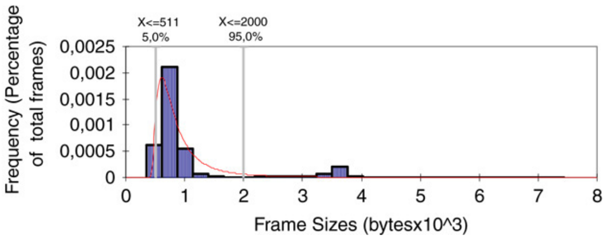

Fig. 1. Histogram for the frame size of the lecture camera trace.

Our observation for the shape of the histograms, derived here for MPEG-4 videoconference traffic, was also derived for MPEG-1 traffic in [30], for a 23 min long sequence of the movie *The Wizard of Oz*. The authors in [30] reached the conclusion that, regardless of the fact that the Pearson type V distribution may be a better fit than other distributions for the whole trace, the development of a model based on a distribution for all frames is not a good choice, as the impact of the frame type (I, P, B) would not be "captured" in such a model. In our work, we first proceeded by testing, statistically, which distribution provides a better fit for the above traces (i.e., if the Pearson V is indeed the better fit for each whole trace), and then we extended our study on each video frame type.

2.2. Statistical tests and autocorrelations

The statistical test was made with the use of Q–Q plots. The Q–Q plot is a powerful goodness-of-fit test [7,31], which graphically compares two data sets in order to determine whether the data sets come from populations with a common distribution (if they do, the points of the plot should fall approximately along a 45 deg reference line). More specifically, a Q–Q plot is a plot of the quantiles of the data versus the quantiles of the fitted distribution (a) z-quantile of X is any value x such that $P((X \leq x) = z)$.

We have plotted the 0.01-, 0.02-, 0.03-, ... quantiles of the actual trace versus the respective quantiles of the various distribution fits for the ARD Talk, N3 Talk, office camera and lecture camera traces, and the common characteristic observed in all four figures was that the Pearson V distribution fit was the best in comparison to the Gamma, Weibull, lognormal and exponential distributions.

However, as in [30], although the Pearson V was shown to be the better fit among all distributions, the degree of goodness-of-fit for the Pearson V varied significantly, and even in the cases of a good fit the fit was not highly accurate.

This was expected, as the gross differences in the number of bits required to represent I, P and B frames impose a degree of periodicity on MPEG-encoded streams, based on the cyclic GOP formats (i.e., the cyclicity leads to very high autocorrelation among subsequent video frames, therefore this behavior cannot be captured by a distribution generating frame sizes independently). Any model which purports to reflect the frame-by-frame correlations of an MPEG-encoded video stream must account for GOP cyclicity, otherwise the model could produce biased estimates of cell loss rate for a network with some given traffic policing mechanism [13,20].

Hence, we proceeded to study the frame size distribution *for each of the three different video frame types* (I, P, B), in the same way we studied the frame size distribution for the whole trace. This approach was also used:

a. In [30], where the authors reached the conclusion that, although the frame-size histogram of the MPEG-1 sequence has the general shape of a Pearson type V distribution, each of the three frame-size subsequences is best modeled with the use of a fit based on the lognormal distribution.
b. In [13], where the GOP GBAR model, proposed by the authors, attempts to "capture" overall statistical properties of I, P and B frames of MPEG movies by using three GBAR models for the generation of three random variables that have Gamma (or Gamma–Weibull)-distributed marginals and geometric autocorrelations.

Another approach, similar to the above, was proposed in [18]. This scheme uses again lognormal distributions and assumes that the change of a scene alters the average size of I frames, but not the sizes of P and B frames. However, it is shown in [8,19] that the average sizes of P and B frames can vary 20% and 30% (often more than that), respectively, in subsequent scenes; therefore the size changes are statistically significant.

As it will be shown from our results, none of the above choices of distribution fits are relevant to the case of I, P and B frames of MPEG-4 videoconference traffic.

Table 2
Statistics for the I, P, B frames of the ARD Talk trace

Frame type	Mean frame size (B)	Peak frame size (B)	Variance-frame size	Pearson type V parameters (α, β)	Autocorrelation (lag-1)
I frames	7035	13 857	3 093 441	(18, 119 608.99)	0.932357
P frames	3020	15 579	1 663 610	(7.48, 19 579.61)	0.778138
B frames	2075	10 486	642 279	(8.7, 15 974.7)	0.944969

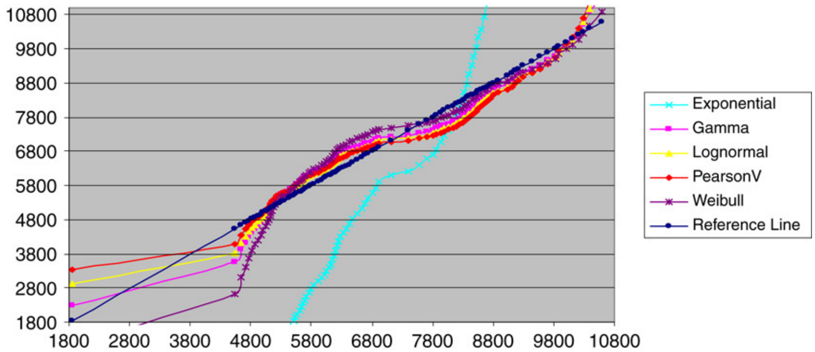

Fig. 2. Q–Q plot for the ARD Talk I frames.

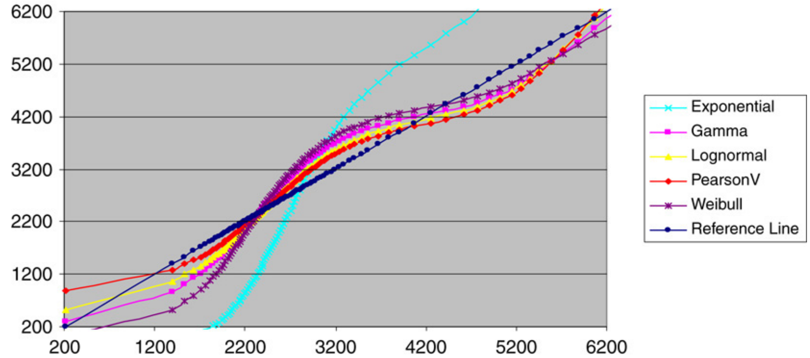

Fig. 3. Q–Q plot for the ARD Talk P frames.

The mean, peak and variance of the video frame sizes for each video frame type of the ARD Talk trace are given in Table 2, along with the respective parameters of the Pearson type V distribution and the autocorrelation coefficient of lag-1. The autocorrelation coefficient of lag-1 shows the very high degree of correlation between successive frames of the same type and will be used in the following sections of this work, in order to build a Discrete Autoregressive Model for each video frame type. The Probability Density Function (PDF) of a Pearson type V distribution with parameters (α, β) is $f(x) = x^{-(\alpha+1)}\mathrm{e}^{-\beta/x}/\beta^{-\alpha}\Gamma(\alpha)$, for all $x > 0$, and zero otherwise. The mean and variance are given by the equations: Mean $= \beta/(\alpha - 1)$, Variance $= \beta^2/[(\alpha - 1)^2(\alpha - 2)]$. The Pearson type V parameters were similarly calculated for each video frame type of the other three traces under study.

From the five distributions examined (Pearson V, exponential, Gamma, lognormal, Weibull) the Pearson V distribution once again provided the best fitting results for 11 of the 12 cases examined, i.e., for all video frame types of the office, ARD Talk and lecture camera traces, and for the I, B frame types of the N3 Talk trace. The only case in which the Pearson V distribution exhibits worse fitting results than another distribution is that of the N3 Talk P frames, where the best fitting result is derived with the use of the lognormal distribution (still, even in this case the difference in the goodness-of-fit results was very marginal). We present indicatively some of these results in Figs. 2–4 (the results omitted here are identical in nature with the ones presented in the figures).

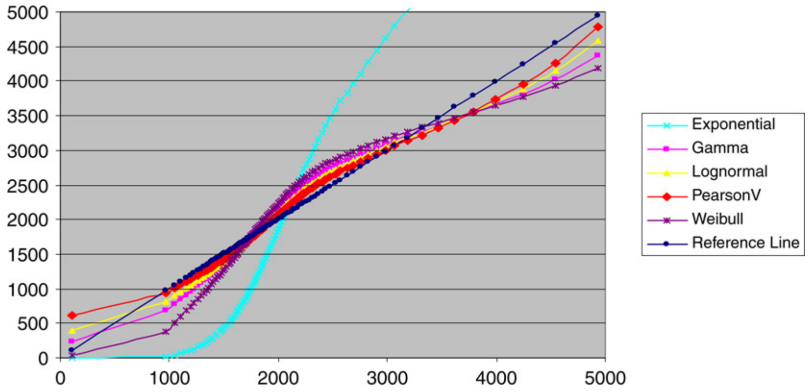

Fig. 4. Q–Q plot for the ARD Talk B frames.

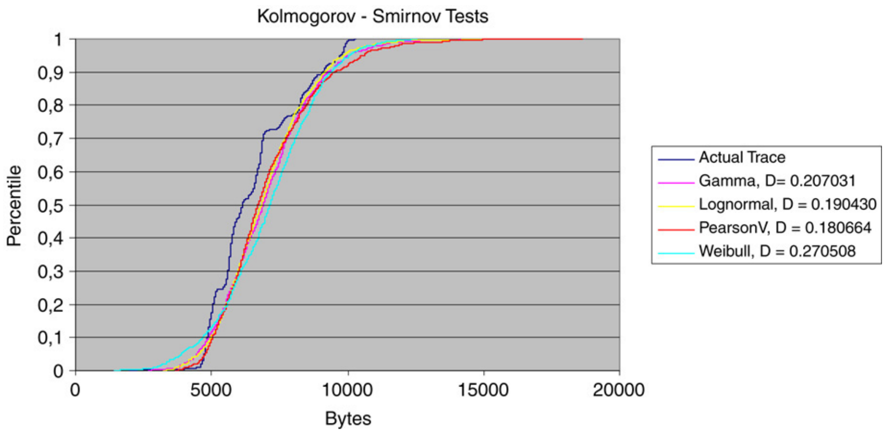

Fig. 5. KS-test (Comparison percentile plot) for the ARD Talk I frames.

In order to further verify the validity of our results, we performed Kolmogorov–Smirnov tests for all 12 fitting attempts. The Kolmogorov–Smirnov test (KS-test) tries to determine if two datasets differ significantly. The KS-test has the advantage of making no assumption about the distribution of data, i.e., it is non-parametric and distribution free. The KS-test uses the maximum vertical deviation between the two curves as its statistic D. For more information on the KS-test the interested reader is referred to [31]. As explained in [31], the use of KS-tests is a good statistical tool; however it has the drawback that KS-tests give the same weight to the difference between the actual data and the fitted distribution for all values of data, whereas many compared distributions differ primarily in their tails. The results of our KS-tests, in Figs. 5 and 6 (which are presented here indicatively), confirm our respective conclusions based on the Q–Q plots (i.e., the Pearson V distribution is the best fit). Similar results were deduced by all 12 KS-tests. The KS-test for the N3 Talk P frames also confirms the respective Q–Q plot result that, just in this case, the lognormal distribution provides a very marginally better fit ($D_{\text{lognormal}} = 0.171875$, $D_{\text{PearsonV}} = 0.173828$).

Although controversy persists regarding the prevalence of Long Range Dependence (LRD) in VBR video traffic [27,28,38], in the specific case of MPEG-encoded video, research has shown that LRD is important [18,29]. The results of our study on single MPEG-4 videoconferencing agree with this conclusion. The sizes of the frames produced by an MPEG video encoder are strongly correlated, as the correlations are a natural consequence of the recurrent GoP pattern and the similarities between the successive images which form the basic elements of a compressed video stream [22]. The autocorrelation functions for the N3 Talk and office camera traces are shown in Figs. 7 and 8 (the autocorrelation function of the ARD Talk trace is similar to that of the N3 Talk trace and the autocorrelation function of the lecture camera trace is similar to that of the office camera trace; they are not presented here in order to avoid repetitive results). Two apparent periodic components are observed from Figs. 7 and 8, one containing lags with low autocorrelation and the other lags with high autocorrelation. We observe that autocorrelation remains high even for

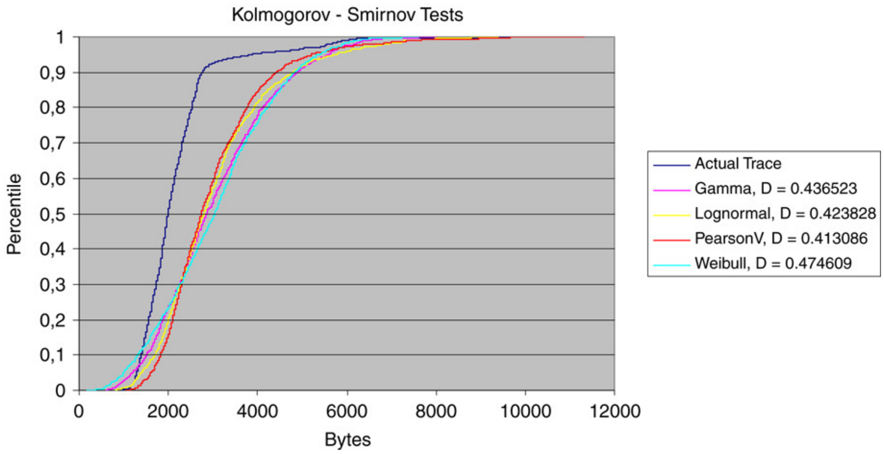

Fig. 6. KS-test (Comparison percentile plot) for the ARD Talk P frames.

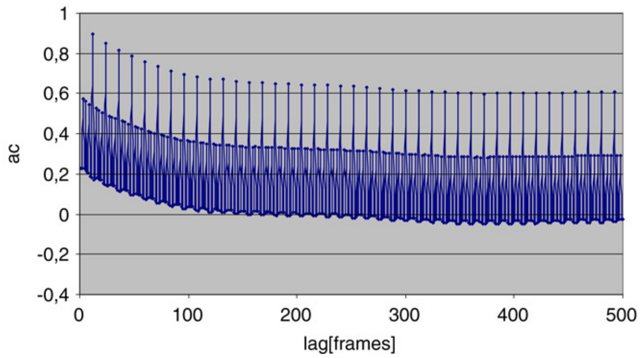

Fig. 7. Autocorrelation function of the N3 Talk trace.

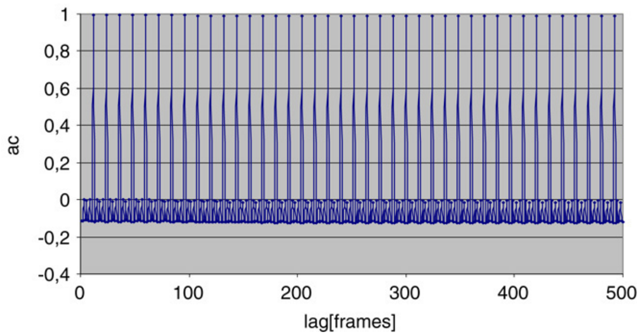

Fig. 8. Autocorrelation function of the office camera trace.

large numbers of lags and that both components decay very slowly; both these facts are a clear indication of the importance of LRD. Our results for videoconferencing traffic agree in this matter with the respective ones in [22] for video traffic: strong autocorrelation coefficients are found due to the periodic recurrence of I, B and P frames, and the autocorrelation function has a very slowly exponentially decreasing envelope.

As mentioned earlier and shown in the figures above, a general comment which stands for all types of frames of the four traces of videoconference traffic is that the autocorrelation coefficient for frames of the same type is always very large, i.e., traffic is highly correlated between successive frames. Although the fitting results when modeling each video frame type separately, with the use of the Pearson V distribution, are clearly better than the results produced by modeling the whole sequence uniformly, this high autocorrelation can never be perfectly “captured” by a distribution generating frame sizes independently, according to a declared mean and standard deviation, and therefore none of the fitting attempts (including the Pearson V), as good as they might be, can achieve perfect accuracy. However, the identification of the Pearson V distribution as the best one for fitting the video frame sizes is very useful (as it was in [7,16] for the respective identification of the Gamma and negative binomial distributions for H.261 traffic); following the steps of [7,16], we extend our work in order to build DAR models which inherently use the autocorrelation coefficient of lag-1 in their estimations and which will be shown to "capture" well the behavior of *multiplexed MPEG-4 videoconference movies*, by generating frame sizes independently for I, P and B frames with the use of the Pearson V distribution.

3. Modeling multiplexed MPEG-4 videoconference traffic

3.1. The DAR(1) model

Autoregressive models have been used in the past to model the output bit-rate of VBR encoders, e.g. [32,33].

A Discrete Autoregressive model of order p, denoted as DAR(p) [34,35], generates a stationary sequence of discrete random variables with an arbitrary probability distribution and with an autocorrelation structure similar to that of an Autoregressive model. DAR(1) is a special case of a DAR(p) process and it is defined as follows: let $\{V_n\}$ and $\{Y_n\}$ be two sequences of independent random variables. The random variable V_n can take two values, 0 and 1, with probabilities 1-ρ and ρ, respectively. The random variable Y_n has a discrete state space S and $P\{Y_n = i) = \pi(i)$. The sequence of random variables $\{X_n\}$ which is formed according to the linear model:

$$X_n = V_n X_{n-1} + (1 - V_n)Y_n$$

is a DAR(1) process.

A DAR(1) process is a Markov chain with discrete state space S and a transition matrix:

$$\mathbf{P} = \rho\mathbf{I} + (1 - \rho)\mathbf{Q} \tag{1}$$

where ρ is the autocorrelation coefficient, $\mathbf{I}$ is the identity matrix and $\mathbf{Q}$ is a matrix with $Q_{ij} = \pi(j)$ for $i, j \varepsilon S$.

Autocorrelations are usually plotted for a range W of lags. The autocorrelation can be calculated by the formula:

$$\rho(W) = E[(X_i - \mu)(X_{i+w} - \mu)]/\sigma^2, \tag{2}$$

where μ is the mean and σ^2 the variance of the frame size for a specific video trace.

3.2. Modeling results and discussion based on Q–Q plots

As in [7,16,36], where a DAR(1) model with negative binomial distribution was used to model the number of cells per frame of VBR teleconferencing video, we want to build a model based only on parameters which are either known at call set-up time or can be measured without introducing much complexity in the network. DAR(1) provides an easy and practical method to compute the transition matrix and gives us a model based only on four physically meaningful parameters, i.e., the mean, peak, variance and the lag-1 autocorrelation coefficient ρ of the offered traffic (these correlations, as already explained, are typically very high for videoconference sources). According to [37], the DAR(1) model can be used with any marginal distribution.

As shown in our work on modeling a single MPEG-4 videoconference trace, the lag-1 autocorrelation coefficient for the I, P and B frames of each trace is very high in almost all the studied cases. Therefore, we proceeded to build a DAR(1) model for each video frame type for each one of the four traces under study. More specifically, in our model the rows of the $\mathbf{Q}$ matrix consist of the Pearson type V probabilities $(f_0, f_1, \dots f_k, F_K)$, where $F_K = \Sigma_{k>K} f_k$, and K is the peak rate. Each k, for $k < K$, corresponds to possible source rates less than the peak rate of K.

From the transition matrix in (1) it is evident that if the current frame has, for example, i cells, then the next frame will have i cells with probability $\rho + (1 - \rho) * f_i$, and will have k cells, $k \neq i$, with probability $(1 - \rho) * f_k$. Therefore the number of cells per video frame stays constant from one (I, P or B) video frame to the next (I, P or B) video frame, respectively, in the realization of our model, with a probability slightly larger than ρ (for example, in the ARD Talk trace, with probability slightly larger than 93.23%, 77.81%, 94.49% for the I, P and B frames of the trace, respectively). This is evident in Figs. 9–11, where we compare the actual I, P and B video frames' sequences

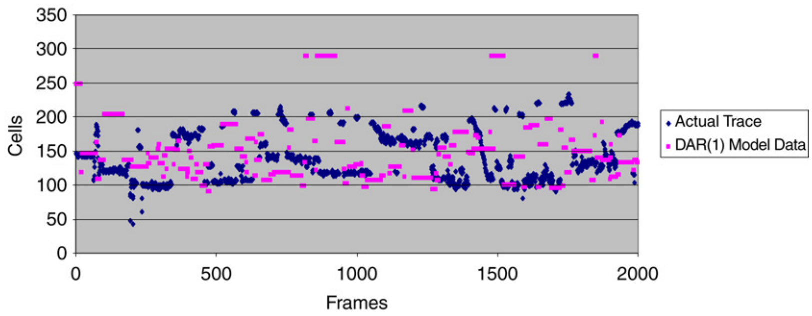

Fig. 9. Comparison for a single trace between a 2000 frame sequence of the actual I frames of the ARD Talk trace and the respective DAR(1) model in number of cells/frame (Y axis).

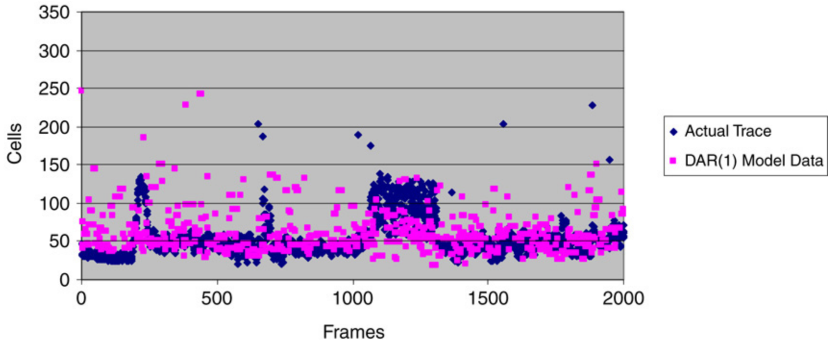

Fig. 10. Comparison for a single trace between a 2000 frame sequence of the actual P frames sequence of the ARD Talk trace and the respective DAR(1) model in number of cells/frame (Y axis).

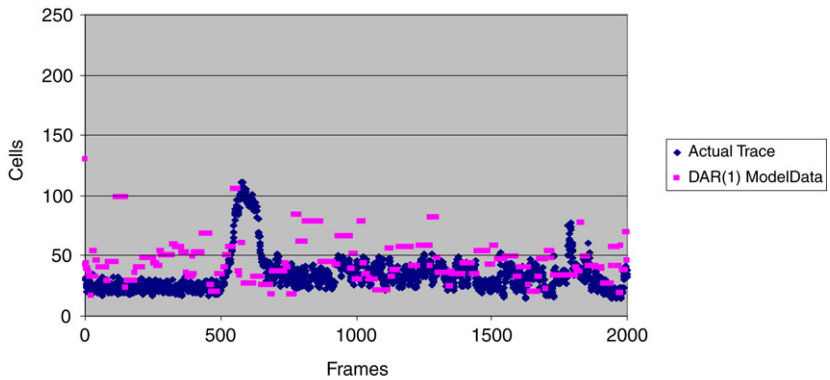

Fig. 11. Comparison for a single trace between a 2000 frame sequence of the actual B frames sequence of the ARD Talk trace and the respective DAR(1) model in number of cells/frame (Y axis).

of the ARD Talk trace and their respective DAR(1) model realizations and it is shown that the DAR(1) models' data produce a "pseudo-trace" with a periodically constant number of cells for a number of video frames. This causes a significant difference when comparing a segment of the sequence of I, P, *or* B frames of the actual ARD Talk video trace and a sequence of the same length produced by our DAR(1) model. The same vast differences also appeared when we plotted the DAR(1) models versus the actual I, P and B video frames of the actual N3 Talk, office camera and lecture camera traces for a single movie.

However, our results have shown that the differences presented above become small for all types of video frames and for all the examined traces *for a superposition of 10 or more sources*, and are almost completely smoothed out in most cases, as the number of sources increases (the authors in [7,36] have reached similar conclusions for

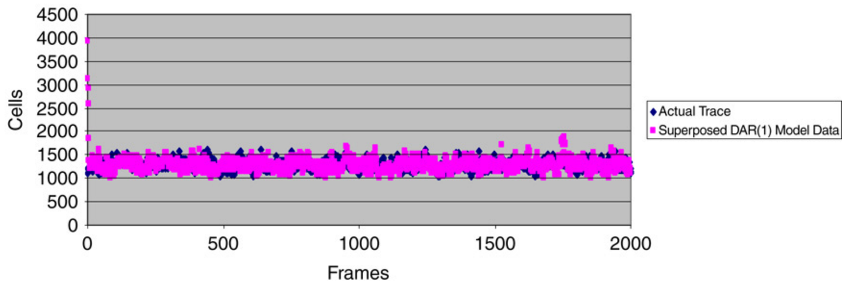

Fig. 12. Comparison for 20 superposed sources between a 2000 frame sequence of the actual P frames sequence of the ARD Talk trace and the respective DAR(1) model in numbers of cells/frame (Y axis).

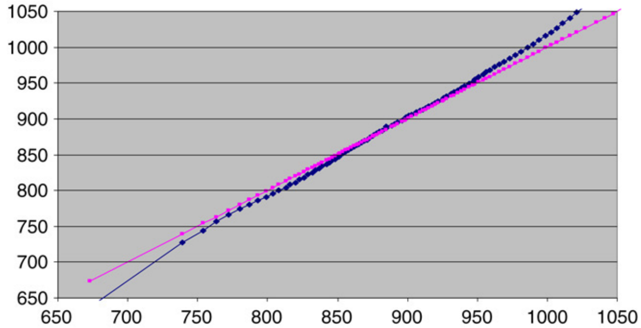

Fig. 13. Q–Q plot of the DAR(1) model versus the actual video for the B frames of the N3 Talk trace, for 20 superposed sources.

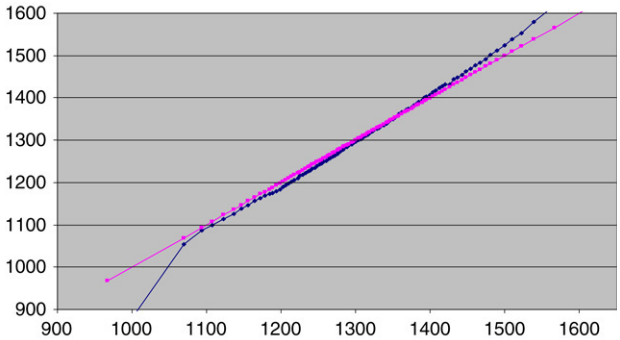

Fig. 14. Q–Q plot of the DAR(1) model versus the actual video for the P frames of the N3 Talk trace, for 20 superposed sources.

their own DAR(1) model and they present results for a superposition of 20 traces). This is clear in Fig. 12, which presents the comparison between our DAR(1) model and the actual P frames' sequences of the ARD Talk video, for *a superposition* of 20 traces[1] (the results were similar for all video frame types of all four traces). The common property of these results (derived by using a queue to model multiplexing and processing frames in a FIFO manner) is that the DAR(1) model seems to provide very accurate fitting results for P and B frames, and relatively accurate results for I frames.

However, although Fig. 12 and our respective results for the other traces suggest that the data distribution of a realization of DAR(1) "captures" well the behavior of the multiplexed actual traces, they do not suffice as a result. Therefore, we proceeded again with testing our model statistically in order to study whether the data distribution of our model's realization is a good fit for the I, P, B frames for the trace superposition.

For this reason we have used again Q–Q plots, and we present indicatively some of these results in Figs. 13–16, where we have plotted the 0.01-, 0.02-, 0.03-, …quantiles of the actual video frames' types versus the respective

[1] We have used the initial trace sequences to generate traffic for 20 sources, by using different starting points in the trace.

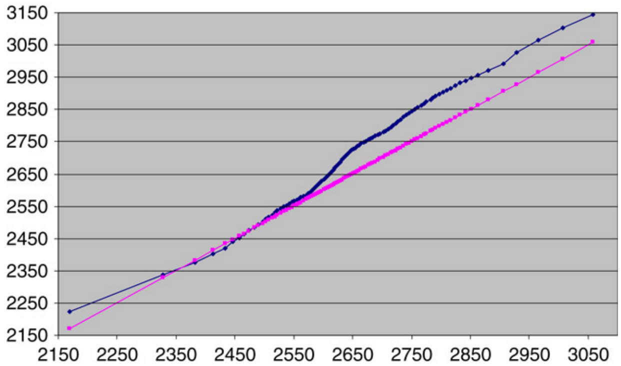

Fig. 15. Q–Q plot of the DAR(1) model versus the actual video for the I frames of the N3 Talk trace, for 20 superposed sources.

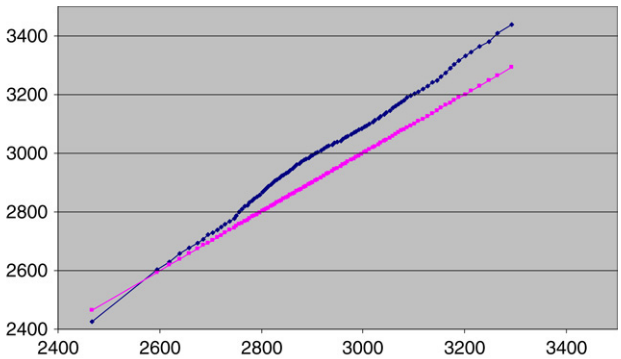

Fig. 16. Q–Q plot of the DAR(1) model versus the actual video for the I frames of the ARD Talk trace, for 20 superposed sources.

quantiles of the respective realized DAR(1) models, for a superposition of 20 traces (*the quantities in both axes for all of the Figs. 13–16 refer to the total number of packets of ATM cell size generated by the sources*). As shown in Figs. 13 and 14, which present the comparison of actual P and B frames with the data distributions of the respective DAR(1) models, the points of the Q–Q plot fall almost completely along the 45 deg reference line (or very close to it), with the exception of the first and last 3% quantiles (left- and right-hand tail), for which the DAR(1) model underestimates and overestimates, respectively, the probability of frames with a very small (large) number of cells. The very good fit shows that the superposition of the P and B frames of the actual traces can be modeled very well by a respective superposition of data produced by the DAR(1) model, as was suggested in Fig. 12.

Figs. 15 and 16 present the comparison of actual I frames with the respective DAR(1) models, for the ARD Talk and the N3 Talk traces. Again, the result suggested from the comparison between a 2000 frame sequence of the actual I frames' sequence of the ARD Talk trace and the respective DAR(1) model (discussed earlier, with our results on Fig. 12), i.e., that our method for modeling I frames of multiplexed MPEG-4 videoconference streams provides only relative accuracy, is shown to be valid with the use of the Q–Q plots. As shown in the figures, the DAR(1) model significantly overestimates, for a large area of quantiles, the traffic which is generated in the I frames of the superposed sources.

The results for all the other cases which are not presented in Figs. 13–16 are similar in nature to the ones shown in the figures. More detailed comments on these results are discussed in Section 4, which presents a comparison of our model with three other efficient MPEG-4 traffic models of the recent relevant literature.

One problem which could arise with the use of DAR(1) models is that such models take into account only short range dependence, while, as shown earlier, MPEG-4 videoconference streams show LRD. This problem is overcome by our choice of modeling I, P and B frames separately, instead of modeling the whole trace. This is shown in Table 3 (which presents the lag-1, lag-2 and lag-3 autocorrelation for various numbers of I, P, B frames' superposition), and in Figs. 17–20, which again present indicatively the autocorrelation versus the number of lags for various video frames' types of superpositions, for the traces under study.

Table 3
Lag-1, lag-2 and lag-3 autocorrelation for various superposed video frames' types

Frame type	Lag number	Autocorrelation for 10 superposed ARD Talk traces	Autocorrelation for 10 superposed lecture camera traces	Autocorrelation for 15 superposed office camera traces	Autocorrelation for 20 superposed N3 Talk traces
I frames	lag-1	0.923099	0.988417	0.977590	0.846386
	lag-2	0.851932	0.979339	0.968813	0.708864
	lag-3	0.790611	0.970938	0.959997	0.599729
P frames	lag-1	0.677880	0.810007	0.839514	0.798678
	lag-2	0.622169	0.788674	0.771694	0.644500
	lag-3	0.581068	0.925405	0.755593	0.587320
B frames	lag-1	0.924995	0.974796	0.786822	0.946025
	lag-2	0.873943	0.961141	0.764408	0.903311
	lag-3	0.823523	0.952967	0.679838	0.872255

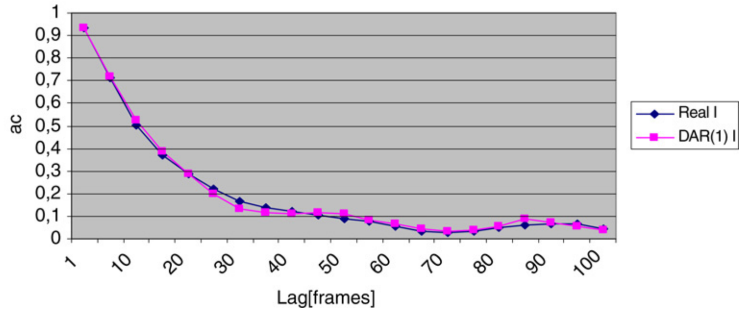

Fig. 17. Autocorrelation vs. number of lags for the *I* frames of the actual ARD Talk trace and the DAR(1) model, for 20 superposed sources.

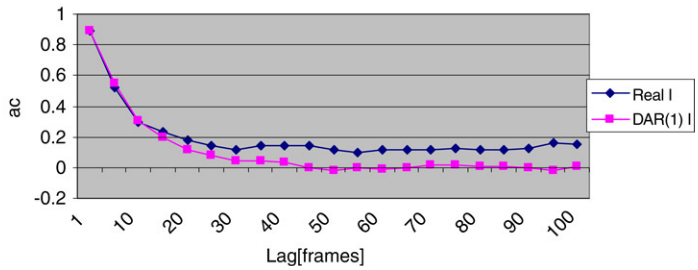

Fig. 18. Autocorrelation vs. number of lags for the *I* frames of the actual N3 Talk trace and the DAR(1) model, for 20 superposed sources.

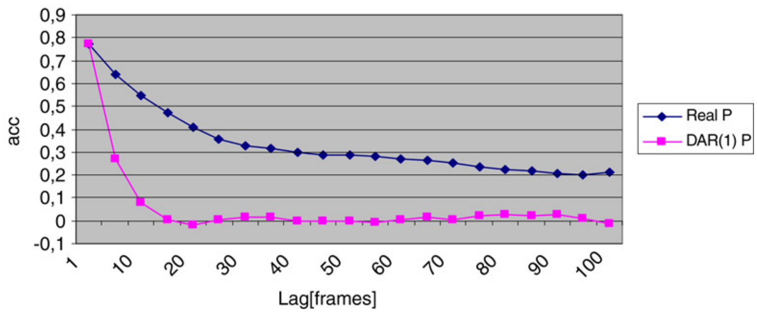

Fig. 19. Autocorrelation vs. number of lags for the *P* frames of the actual N3 Talk trace and the DAR(1) model, for 20 superposed sources.

Table 3 shows that the degree of correlation between successive frames (lag-1) of the same type, which was shown to be very high for single traces in Table 2, remains very high for various numbers of superposed sources. However

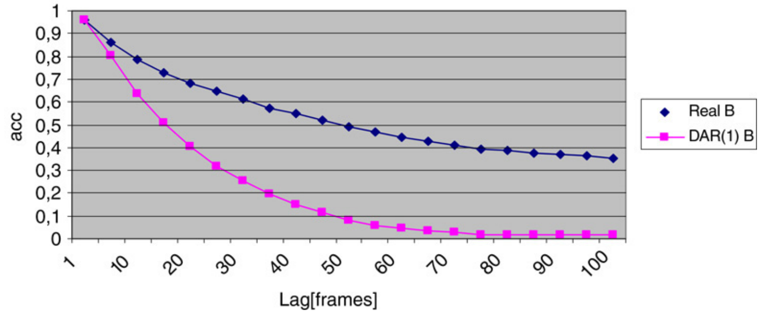

Fig. 20. Autocorrelation vs. number of lags for the B frames of the actual N3 Talk trace and the DAR(1) model, for 20 superposed sources.

it is clear from the combined results in Table 3 and in Figs. 17–20 that, even for a small number of lags the high autocorrelation of the superposition of frames decreases quickly, for all the traces (this number varies from as low as 3, as shown in many cases in Table 3, to a maximum of 10; as shown in all the figures, for lags greater than 10 the autocorrelation decreases dramatically). This decrease is in some cases "captured" well by the DAR(1) model (e.g., Figs. 17 and 18) and in some cases not (e.g., Figs. 19 and 20). However, the important conclusion in all cases is that the quick decrease of the autocorrelation causes its "capture" (good or not) to have minimal impact on the fitting quality of the DAR(1) model. Figs. 17–20 were chosen specifically among the 12 respective results for the autocorrelation versus the number of lags for a superposition of frames' types, in order to emphasize this minimal impact and to further support our choice of using a first-order model. This is clear when comparing:

a. Figs. 13 and 20, which show that, as long as the autocorrelation of the actual video frames sequence decreases quickly, the existence of a difference between this autocorrelation and the autocorrelation of the respective DAR(1) model (which decreases even more quickly) does not affect the high quality of the fit.
b. Figs. 16 and 17, which show that, even in the case of a nearly perfectly accurate fit among the autocorrelations of the actual sequence and the model, the model does not necessarily provide equally accurate fitting results.

The important conclusion from all of the above results is that if the autocorrelation of the video frames' sequence remained high for multiplexed traffic (as it does for singe-source traffic), the DAR(1) approach would not work. The reason that our approach works so well is that the autocorrelation of multiplexed traffic diminishes so quickly, that it becomes of negligible importance in terms of modeling the behavior of multiplexed traffic.

3.3. Modeling results and discussion based on a queuing performance study

Finally, we completed our statistical analysis with a queuing performance study similar to the one presented in [22], in order to acquire additional validation for the quality of our results. More specifically, we fed a discrete time queuing system (representing a downlink channel) with unlimited buffer size, for a 20 Mbps channel transmission rate. The transmission slot had a 40 ms duration (equal to the inverse of the video frame rate). We assumed, in our simulation, that up to 10 packets (this value is taken from [22]) of length equal to 48 B (payload of an ATM-sized packet), may be served during each transmission slot (i.e., we consider a TD/CDMA channel frame with a duration of 12 ms [41] and 62 slots/frame). We studied the packet waiting time and the packet loss ratio to validate the model for various load factors, given the delay constraints of real time video streams (*packets of a video frame need to be transmitted before the arrival of the next video frame, i.e., within* 40 ms, *otherwise they are dropped; an upper bound of just* 0.01% *is allowed for videoconference packet dropping* [41]). A load of 0.4, e.g., corresponds to a load of 0.4*20 Mbps = 8 Mbps. It should be noted here that in our queuing performance study we compare the results between the whole actual traces and our models, i.e., we do not compare I, P and B frames separately between the actual trace and the models, but instead we use our separate models for I, P and B frames in order to generate a "pseudo-trace" based on the actual traces' GOP pattern ($N = 12$, $M = 3$).

We have derived results for all the traces under study; we present here selectively some of them, for brevity reasons. The results for all cases not shown below are of very similar nature to the ones presented.

Figs. 21–23 present a comparison of the waiting time cdfs for superposed "Office Cam" and "N3 Talk" traces and our respective models; it is clear from all three figures that our modeling approach allows a quite accurate

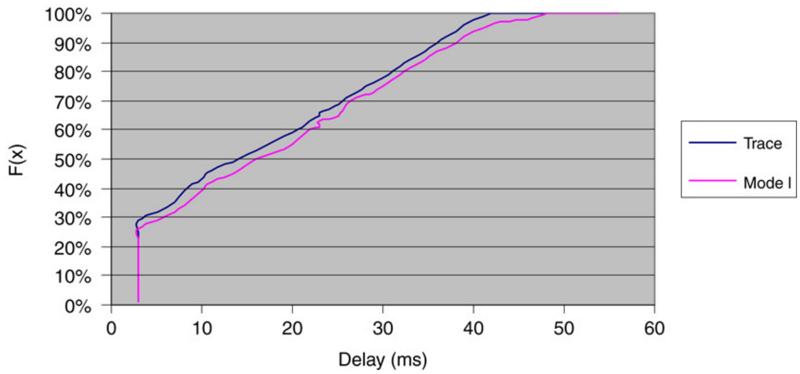

Fig. 21. Waiting time cdfs for the office camera traces and the model, for 10 superposed sources (offering a 20% average channel load). Unlimited buffer size.

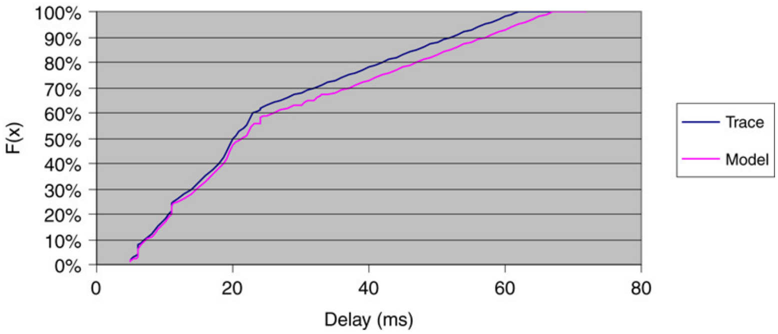

Fig. 22. Waiting time cdfs for the office camera traces and the model, for 20 superposed sources (offering a 40% average channel load). Unlimited buffer size.

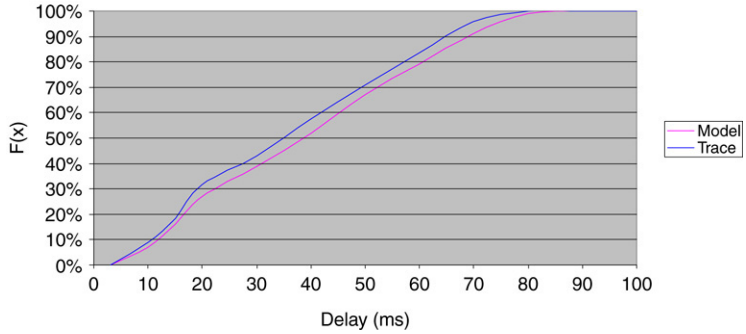

Fig. 23. Waiting time cdfs for the N3 Talk traces and the model, for 20 superposed sources (offering a 55% average channel load). Unlimited buffer size.

characterization of the waiting time experienced by the packets of the multiplexed MPEG-4 videoconference sources, for various but realistic load factors (20%–55%). This result, combined with the results presented in Figs. 13–16, verifies the validity of our modeling approach, hence showing that a first-order model is a competent candidate for modeling MPEG-4 videoconference traffic from multiplexed video sources. However, as shown in Figs. 22 and 23, the loads of 20 superposed "Office Cam" and "N3 Talk" sources, respectively, is too large for the system to handle (both for the real traces and our pseudo-traces) without an excessive delay (significantly larger than the 40 ms upper bound) being experienced by a large percentage of the video packets.

The fact that the results for the actual traces and the model do not perfectly coincide in Figs. 21 and 22 is mainly due to the fact that our modeling approach is not perfectly accurate in modeling I frames.

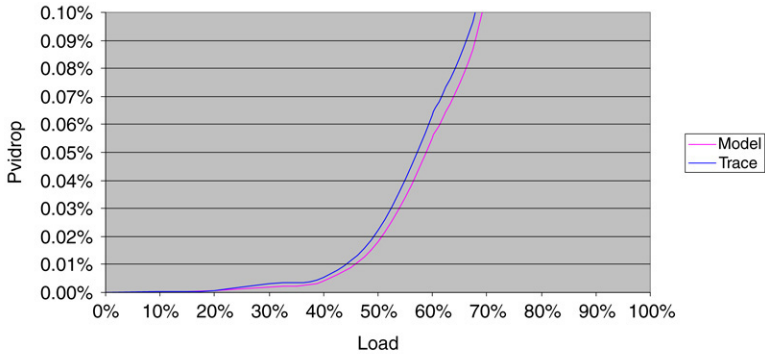

Fig. 24. Video packet dropping ratio versus offered load from superposition of N3 Talk traces. Unlimited buffer size.

The system's inadequacy to handle (in both the cases of the model and the actual trace) the offered traffic is more intense in the results presented in Fig. 23, where a larger percentage of the video packets than in Fig. 22 experiences longer delays than the 40 ms upper bound. The reason for this is not only the higher offered load, but also that the "Office Cam" trace (used in Figs. 21 and 22) has a mean rate of 400 Kbps and peak rate of 2 Mbps, whereas the "N3 Talk" trace has a mean rate of 550 Kbps and peak rate of 3.4 Mbps; therefore not only is the "N3 Talk" trace more demanding in bandwidth, but most importantly it is much burstier (peak/mean = 6.18) than the "Office Cam" trace (peak/mean = 5). This inadequacy is made clearer in Fig. 24, which presents the video packet loss ratio for various system loads offered by a superposition of "N3 Talk" traces. For both the model and the actual traces, the system is unable to satisfy the QoS requirement of maximum 0.01% video packet dropping, for loads higher than 46%.

The most important conclusion from Figs. 21–24 is, again, that all the results derived from our modeling approach are shown to be very close to the ones derived with the use of the actual traces.

4. Conceptual comparison with other MPEG-4 modeling schemes

As stated in the introduction section, to the best of our knowledge the subject of modeling MPEG-4 videoconference traffic has been addressed in the relevant literature only in [42,43]. However, in [21–23], three efficient schemes for modeling MPEG-4 video traffic were proposed and evaluated. In this section, we will conceptually compare our work with all of these studies.

In [22], the author proposes a model based on a detailed analysis of the first and second order statistics of real MPEG-4 traces, which uses a customization of the Discrete Batch Markovian Process. Although the four traces used by the author (MPEG-4, low bit rate encoding) exhibit LRD, as in the case of our traces, the fact that the author is considering video traffic instead of videoconference causes his model to be significantly different to ours. The model in [22] is based on the *identification of the start of each new scene*, which is not the case in videoconference traffic, where more or less there is basically one scene with small changes. Also, the author uses the lognormal distribution to provide a good fit for the subsequence of the mean I frames sizes of each scene, which was shown earlier to be a second-rate choice for MPEG-4 videoconference traffic; the same choice (of the lognormal distribution) is made in [22] for the approximation of the B and P frames' size distributions. Finally, the author compares the correlations of the B, P and I frames' size sequences and denotes that the I sequence correlation is always clearly superior to the B or P sequence correlation, which is often not the case in MPEG-4 videoconference traffic (e.g., in the ARD Talk trace, the correlation of the B frames' size sequence is always superior to the autocorrelation of the I frames' size sequence).

In [21,23] the authors use wavelet analysis to model the behavior of MPEG-4 video traffic. This approach is used in [21] to show the self-similar behavior of MPEG traffic, and it is implemented on the whole trace. The detailed analytical work in [23] models I, P and B frames separately, based on three MPEG-4 video traces. The use of wavelet analysis by the authors to model I-frame sizes is shown to provide excellent results; however, in comparison to our work, we find again the significant difference that I frame sizes are considered to follow a Gamma distribution. The authors proceed with modeling P frame sizes based on a time-domain model which attempts to "capture" the intra-GOP correlation between I and P frames (this correlation was first observed for MPEG-1 video traffic in [39]), and finally they model B-frames' sizes with a lognormal distribution, based on their observation that the sizes of B-frames are relatively small compared to those of P frames and that the correlations between the B-frame and I-frame sequences are much smaller than those between the P-frame and I-frame sequences. This, however, is not the case in MPEG-4 videoconference traffic, as B frames have, on average, a size equal to 70%–80% of the P frames, and, as shown from our results, their modeling with a lognormal distribution is not the best choice. Finally, although the authors' results in modeling the MPEG-4 traces are shown to be excellent, the fact that P-frames' sizes are generated based on the respective modeled I-frames' sizes of each GOP, could lead, *especially in the case of superposed traces*, to an unnecessary bad estimation of the P-frame size due to a mediocre modeling of the specific I frame (i.e., it could lead to certain "chain mistakes").

In [42,43], similarly to our work, the authors study the possibility of modeling MPEG-4 videoconference traffic (from [25,26]) with autoregressive models. Among the traces studied by the authors, three are also used in our study, namely "Lecture Room Cam", "ARD Talk" and "Office Cam" (however, for the latter trace they use the low quality coding version, whereas we use the high quality coding version for all traces); the most important difference with our work is that in [42,43] the authors attempt to model *single* traces, not multiplexed traffic, and they do it by using a third order autoregressive model. Despite the high computational complexity of the model (which is noted by the authors) in comparison to the first order model used in our work for multiplexed traffic, the authors explain that the model is still not sufficient in itself to provide a good fit for the actual trace; therefore a gamma distortion of the marginal distribution function was required to derive a good fit with their approach.

Therefore, based on the very good results of our study in modeling P- and B-frames' sizes of multiplexed MPEG-4 videoconference traffic, and the significantly lower complexity of our scheme in comparison to the above-mentioned approaches, we feel that our approach is best for this type of traffic. However, since our modeling scheme clearly overestimates I-frames' sizes for multiplexed MPEG-4 videoconference traffic, the use of wavelet modeling for the I-frames' size sequence may provide a very competent solution, and our future work will be pointed towards this direction.

5. Conclusions

Models of video traffic will prove very important in the immediate future, as networks will need to handle video traffic competently (i.e., to guarantee its strict QoS requirements despite its burstiness). Hence, in this paper we have investigated, for the first time in the literature to the best of our knowledge, the subject of modeling MPEG-4 videoconference traffic.

Initially, we have investigated the possibility of modeling single MPEG-4 videoconference traces with well-known distributions. Our results have shown that the Gamma and lognormal distributions, which are considered the best choice for modeling many types of video traffic and are used as the basis for many video models in the literature, *are not* the most appropriate choice for modeling MPEG-4 videoconference traffic. This is the first significant contribution of this paper.

We then proceeded to use the Pearson V distribution, which was shown to be the best fit among all the examined distributions for single MPEG-4 videoconference sources, for all the video traces under study. Our approach was to model separately the I, P and B frames of each trace, in order to achieve better modeling accuracy. Indeed, our results have shown that *this is a clearly better choice* than modeling the whole trace; however, the behavior of videoconference traffic can never be perfectly "captured" by a distribution generating independently frame sizes according to a declared mean and standard deviation, due to the high autocorrelation of videoconference traffic. Hence, none of the fitting attempts can achieve high accuracy. Still, important insight was gained from our results on modeling single MPEG-4 videoconference sources, as it led us to propose a new approach in order to "capture" the behavior of *multiplexed* MPEG-4 videoconference movies from VBR coders; this is the second significant contribution of this paper.

Our approach, which is based on building a Discrete Autoregressive (DAR(1)) model, has been used in the past in the relevant literature for multiplexed H.261 sources, *but it is used for the first time for MPEG-4 traffic*; we show, in our work, that the different nature of MPEG-4 videoconference traffic compared to H.261 traffic demands the use of three DAR models instead of one, as in the relevant work for H.261.

The subject of modeling multiplexed videoconference sources is especially significant, since wireless and wired networks in the near future will need to handle a continuously growing number of superposed videoconference calls. Based on our results and on comparisons of our new modeling method with other existing approaches, we have shown that our modeling approach provides very accurate results in modeling P and B video frames' sizes, but moderately accurate results in modeling I frames. We have also discussed relevant approaches for MPEG-4 video (not videoconference, which has much higher autocorrelation) modeling; we have explained our scheme's advantages in comparison to them for videoconference traffic, and we have discussed how, with a combination of our modeling approach with another approach, used for MPEG-4 video modeling, the moderate accuracy of our scheme in I-frames' size modeling may be overcome.

Acknowledgements

The authors would like to thank the Editor and the anonymous reviewers of the paper. Their valuable remarks and suggestions have significantly helped us enhance our work.

References

[1] T. Stockhammer, H. Jenkac, G. Kuhn, Streaming video over variable bit-rate wireless channels, IEEE Transactions on Multimedia 6 (2) (2004) 268–277.
[2] D.M. Lucantoni, M.F. Neuts, A.R. Reibman, Methods for performance evaluation of VBR video traffic models, IEEE/ACM Transactions on Networking 2 (1994) 176–180.
[3] N.-E. Rikli, Modeling techniques for VBR video: Feasibility and limitations, Performance Evaluation 57 (1) (2004) 57–68.
[4] E. Gelenbe, X. Mang, R. Onvural, Bandwidth allocation and call admission control in high-speed networks, IEEE Communication Magazine 35 (5) (1997) 122–129.
[5] E. Gelenbe, X. Mang, R. Onvural, Diffusion based call admission control in ATM, Performance Evaluation 27–28 (1996) 411–436.
[6] M. Nomura, T. Fuji, N. Ohta, Basic characteristics of variable rate video coding in ATM environment, IEEE Journal on Selected Areas in Communications 7 (5) (1989) 752–760.
[7] D.P. Heyman, A. Tabatabai, T.V. Lakshman, Statistical analysis and simulation study of video teleconference traffic in ATM networks, IEEE Transactions on Circuits and Systems for Video Technology 2 (1) (1992) 49–59.
[8] A.M. Dawood, M. Ghanbari, Content-based MPEG video traffic modeling, IEEE Transactions on Multimedia 1 (1) (1999) 77–87.
[9] B. Melamed, D.E. Pendarakis, Modeling full-length VBR video using Markov-renewal modulated TES models, IEEE Journal on Selected Areas in Communications 16 (5) (1998) 600–611.
[10] K. Chandra, A.R. Reibman, Modeling one- and two-layer variable bit rate video, IEEE/ACM Transactions on Networking 7 (3) (1999) 398–413.
[11] Q. Ren, H. Kobayashi, Diffusion approximation modeling for Markov modulated bursty traffic and its applications to bandwidth allocation in ATM networks, IEEE Journal on Selected Areas in Communications 16 (5) (1998) 679–691.
[12] D.P. Heyman, The GBAR source model for VBR videoconferences, IEEE/ACM Transactions on Networking 5 (4) (1997) 554–560.
[13] M. Frey, S. Ngyuyen-Quang, A gamma-based framework for modeling variable-rate video sources: The GOP GBAR model, IEEE/ACM Transactions on Networking 8 (6) (2000) 710–719.
[14] H.S. Chin, J.W. Goodge, R. Griffiths, D.J. Parish, Statistics of video signals for viewphone-type pictures, IEEE Journal on Selected Areas in Communications 7 (5) (1989) 826–832.
[15] R. Kishimoto, Y. Ogata, F. Inumaru, Generation interval distribution characteristics of packetized variable rate video coding data streams in an ATM network, IEEE Journal on Selected Areas in Communications 7 (5) (1989) 833–841.
[16] D.P. Heyman, T.V. Lakshman, A. Tabatabai, H. Heeke, Modeling teleconference traffic from VBR video coders, in: Proceedings of the IEEE International Conference on Communications (ICC) 1994, New Orleans, USA, pp. 1744–1748.
[17] S. Xu, Z. Huang, A Gamma autoregressive video model on ATM networks, IEEE Transactions on Circuits and Systems Video Technology 8 (2) (1998) 138–142.
[18] M. Krunz, S.K. Tripathi, On the characterization of VBR MPEG streams, in: Proceedings of ACM SIGMETRICS, vol. 25, June 1997.
[19] U.K. Sarkar, S. Ramakrishnan, D. Sarkar, Modeling full-length video using Markov-modulated gamma-based framework, IEEE/ACM Transactions on Networking 11 (4) (2003) 638–649.
[20] O. Rose, Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems, in: Proceedings of the 20th Annual Conference on Local Computer Networks, October 1995.
[21] D. Arifler, B.L. Evans, Modeling the self-similar behavior of packetized MPEG-4 video using wavelet-based methods, in: Proceedings of the IEEE International Conference on Image Processing, Rochester, New York, USA, 2002, pp. 848–851.
[22] S. Baey, Modeling MPEG4 video traffic based on a customization of the DBMAP, in: Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems, SPECTS'04, San Jose, California, July 2004, pp. 705– 714.
[23] M. Dai, D. Loguinov, Wavelet and time-domain modeling of multi-layer VBR video traffic, in; Proceedings of the 14th International Packet Video Workshop, Irvine, California, USA, December 2004.
[24] S. Gringeri, R. Egorov, K. Shuaib, A. Lewis, B. Basch, Robust compression and transmission of MPEG-4 video, http://www.kom.e-technik.tu-darmstadt.de/acmmm99/ep/gringeri/. Online.
[25] http://www-tkn.ee.tu-berlin.de/research/trace/trace.html. Online.
[26] F.H.P. Fitzek, M. Reisslein, MPEG-4 and H.263 video traces for network performance evaluation, IEEE Network 15 (6) (2001) 40–54.

[27] D.P. Heyman, T.V. Lakshman, What are the implications of long-range dependence for VBR-video traffic engineering, IEEE/ACM Transactions on Networking 4 (3) (1996) 301–317.

[28] B.K. Ryu, A. Elwalid, The importance of long-range dependence of VBR video traffic in ATM traffic engineering: Myths and realities, in: Proceedings of the ACM SIGCOMM, Stanford, CA, USA, 1996, pp. 3–14.

[29] J. Beran, R. Sherman, M.S. Taqqu, W. Willinger, Long-range dependence in variable bit-rate video traffic, IEEE Transactions on Communications 43 (2–4) (1995) 1566–1579.

[30] M. Krunz, H. Hughes, A traffic model for MPEG-coded VBR streams, in: Proceedings of the ACM SIGMETRICS, Ottawa, Canada, 1995, pp. 47–55.

[31] A.M. Law, W.D. Kelton, Simulation Modeling & Analysis, 2nd ed., McGraw-Hill, 1991.

[32] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, J.D. Robbins, Performance models of statistical multiplexing in packet video communications, IEEE Transactions on Communications 36 (7) (1988) 834–844.

[33] C. Shim, I. Ryoo, J. Lee, S. Lee, Modeling and call admission control algorithm of variable bit rate video in ATM networks, IEEE Journal on Selected Areas in Communications 12 (2) (1994) 332–344.

[34] A. Adas, Traffic models in broadband networks, IEEE Communications Magazine 35 (7) (1997) 82–89.

[35] P.A. Jacobs, P.A.W. Lewis, Time series generated by mixtures, Journal of Time Series Analysis 4 (1) (1983) 19–36.

[36] A. Elwalid, D.P. Heyman, T.V. Lakshman, D. Mitra, A. Weiss, Fundamental bounds and approximations for ATM multiplexers with applications to video teleconferencing, IEEE Journal on Selected Areas in Communications 13 (6) (1995) 1004–1016.

[37] T.V. Lakshman, A. Ortega, A.R. Reibman, VBR video: Trade-offs and potentials, Proceedings of the IEEE 86 (5) (1998) 952–973.

[38] K. Park, W. Willinger (Eds.), Self-Similar Network Traffic and Performance Evaluation, John Wiley & Sons, 2000.

[39] A. Lombardo, G. Morabito, G. Schembra, An accurate and treatable markov model of MPEG-video traffic, in: Proceedings of the IEEE Infocom 1998, San Francisco, CA, USA.

[40] R. Schafer, MPEG-4: A multimedia compression standard for interactive applications and services, IEE Electronics and Communications Engineering Journal 10 (6) (1998) 253–262.

[41] D.A. Dyson, Z.J. Haas, A dynamic packet reservation multiple access scheme for wireless ATM, Mobile Networks and Applications 4 (2) (1999) 87–99.

[42] A. Golaup, A.H. Aghvami, Modelling of MPEG4 traffic at GoP level using autoregressive processes, in: Proceedings of the IEEE Vehicular Technology Conference-fall, Vancouver, Canada, 2002, pp. 854–858.

[43] A. Golaup, A.H. Aghvami, A multimedia traffic modeling framework for simulation-based performance studies, Computer Networks 50 (12) (2006) 2071–2087.

Aggelos Lazaris was born in Athens, Greece, in 1982. He received his 5-year Diploma in Electronic and Computer Engineering from the Technical University of Crete, in 2006. He is currently working towards the M.Sc. degree in the same department. His research interests include video encoding and transmission; traffic modeling; multiple access control protocols design for wireless networks.

Polychronis Koutsakis was born in Hania, Greece, in 1974. He received his 5-year Diploma in Electrical Engineering in 1997 from the University of Patras, Greece and his M.Sc. and Ph.D. degrees in Electronic and Computer Engineering in 1999 and 2002, respectively, from the Technical University of Crete, Greece. He was a Visiting Lecturer at the Electronic and Computer Engineering Department of the same University for three years (2003–2006). He is currently an Assistant Professor at the Electrical and Computer Engineering Department of McMaster University, Canada.

His research interests focus on the design, modeling and performance evaluation of computer communication networks, and especially on the design and evaluation of multiple access schemes for multimedia integration over wireless networks, on call admission control and traffic policing schemes for both wireless and wired networks, on multiple access control protocols for mobile satellite networks, wireless sensor networks and powerline networks, and on traffic modeling. Dr. Koutsakis has authored more than 55 peer-reviewed papers in the above mentioned areas, has served as a Guest Editor for an issue of the ACM Mobile Computing and Communications Review, as a TPC member for conferences such as IEEE GLOBECOM 2006, IEEE GLOBECOM 2007, IEEE LCN 2006, IEEE LCN 2007 and IEEE PerCom 2006, as Session Chair for the IEEE GLOBECOM 2006 Symposium on Satellite & Space Communications and the IEEE WCNC 2007 and serves as a reviewer for most of the major journal publications focused on his research field. He is a member of the IEEE.

Michael Paterakis received his Diploma degree from the National Technical University of Athens, Greece, his M.Sc. degree from the University of Connecticut, and his Ph.D. degree from the University of Virginia, in 1984, 1986, and 1988, respectively, all in Electrical Engineering. Since 1995, he is a faculty member in the Department of Electronic and Computer Engineering (ECE) at the Technical University of Crete, Greece, where he is currently a Professor and Vice Rector of the University responsible for Academic Affairs & Personnel. During the period 2000–2005, he served a five year term as the Director of the Telecommunication Systems Institute, a national research institute operating within the framework of the Technical University of Crete. During Sept. 1999 – Aug. 2001, he served a two year term as the Chairman of the ECE Department. He was an Associate Professor in the Department of Computer and Information Sciences (CIS) at the University of Delaware, on the faculty of which he has been between 1988 and 1995.

His research interests include computer communication networks with emphasis on protocol design, modeling and performance evaluation of broadband wireline and wireless networks; queueing and applied probability theory and their application to computer communication networks and to distributed multimedia information systems. He has published over 100 papers in the abovementioned technical areas. He has served on the Technical Program Committees of major international conferences and as reviewer for almost all of the major IEEE Transactions and other international technical journals and conferences in his research areas. He also served as invited speakers co-chair for the 2004 IFIP Networking Conference, and as general co-chair for the 2005 14th IEEE LANMAN Workshop. Professor Paterakis is a senior member of the IEEE.