Modeling and Call Admission Control for Wireless Videoconference Traffic

Master Thesis

By

Aggelos D. Lazaris

Submitted to the Department of Electronic and Computer Engineering in partial fulfillment of the requirements for the Master Degree.

Technical University of Crete

Advisor: Professor Michael Paterakis

Co-advisor: Assistant Professor Polychronis Koutsakis (McMaster University, Canada)

Committee Member: Professor Vasileios Digalakis

Committee Member: Professor Nikolaos Sidiropoulos

September 2008

Ευχαριστίες

Στο σημείο αυτό, θα ήθελα να εκφράσω θερμές ευχαριστίες στον επιβλέποντα Καθηγητή κ. Μιχάλη Πατεράκη για το πολύ ενδιαφέρον θέμα που μου ανέθεσε καθώς και για την υποστήριξη που μου παρείχε κατά την διάρκεια της εκπόνησης της παρούσας εργασίας. Επίσης ευχαριστώ θερμά τον συνεπιβλέποντα Επίκουρο Καθηγητή κ. Πολυχρόνη Κουτσάκη για την πολύτιμη βοήθεια που μου παρείχε μέχρι την ολοκλήρωση της εργασίας αυτής, την υποστήριξη καθώς και την άριστη συνεργασία του. Θα ήθελα να ευχαριστήσω επίσης τον Καθηγητή κ. Νικόλαο Σιδηρόπουλο για την ανάγνωση της παρούσας εργασίας και τις εύστοχες παρατηρήσεις του, καθώς και για τις πολύτιμες γνώσεις που μας μετέδωσε μέσα από τα μαθήματα «Θεωρία Εκτίμησης και Ανίχνευσης» και «Κυρτή Βελτιστοποίηση». Επίσης ευχαριστώ πολύ τον Καθηγητή κ. Βασίλειο Διγαλάκη για την ανάγνωση της παρούσας εργασίας και τις εύστοχες παρατηρήσεις του. Τέλος ευχαριστώ την οικογένειά μου, για την υποστήριξη που μου παρείχε, την υπομονή και την κατανόηση που επέδειξε κατά την διάρκεια της φοίτησης μου στο Πολυτεχνείο Κρήτης.

Abstract

The burstiness of video traffic makes accurate video modeling a very important tool for network engineers in order to evaluate the performance of future wired and wireless networks. In the first part of this work, we introduce a new traffic model for medium quality MPEG-4 videoconference traffic. We then proceed to use this model, as well as a previously developed model for high quality MPEG-4 traffic, in the implementation of a new Call Admission Control scheme which makes decisions on the acceptance/rejection of a new video call not only based on the predicted capacity that users will consume, but also on the possible revenue gained for the provider when degrading current users in order to accommodate new ones. Our scheme is shown, via an extensive simulation study, to provide excellent Quality of Service (QoS) to wireless videoconference users. Our work focuses on next generation cellular networks, but with the proper modifications its concept can be used in other types of wireless networks as well. In the second part of this work, we propose and test a new model for traffic originating from multiplexed H.264 videoconference streams. We first investigate the behavior of single H.264 videoconference sources, by trying to model it with well-known distributions. Our results provide significant insight and help to build a Discrete Autoregressive (DAR(1)) model which is shown to be able to capture the behavior of multiplexed H.264 videoconference movies from VBR coders, combining simplicity with high accuracy.

Contents

List of Figures

List of Tables

Chapter 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 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][4][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][7], discrete AR (DAR) models [2][8], Markov renewal processes (MRP) [9], MRP transform-expand-sample (TES) [10], finite-state Markov chain [11][12], discrete-time Semi-Markov Processes (SMP) [13], wavelets [14], multifractal and fractal methods [15], and Gamma-beta-auto-regression (GBAR) models [16][17]. 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 [16]. Also, in [18][19], different approaches are proposed for MPEG-1 traffic, based on the log-normal, 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 I frame. 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 [16]. 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 our recent work [6], we focused on the problem of modeling high quality videoconference traffic from MPEG-4 encoders. The MPEG-4 standard is particularly designed for video streaming over wireless networks [20][21]. We used four different long sequences of MPEG-4 encoded videos and we showed that the use of the Gamma and lognormal distributions (which was 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 showed 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 extended 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 followed the steps of the work conducted by Heyman et al. in [8][22] in order to build Discrete Autoregressive (DAR) models of order one. The work in [8][22], 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 agreed with those in [8][22] in that we found that the DAR models work well when several sources are multiplexed (which is most often the case in both wired and wireless networks). In [6] we showed that the different nature of MPEG-4 videoconference traffic, compared to H.261 traffic, demanded the use of three DAR models instead of one, as in [8][22].

A brief reference to the MPEG-4 and H.264 standards and their above-mentioned important differences 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 bitrate 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 [23]. 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 [24][25]. 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].

H.264 is the latest video coding standard of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG). It has recently become the most widely accepted video coding standard since the deployment of MPEG2 at the dawn of digital television, and it may soon overtake MPEG2 in common use. It covers all common video applications ranging from mobile services and videoconferencing to IPTV, HDTV, and HD video storage [27]. In the process, a standard was created that improved coding efficiency by a factor of at least about two (on average) over MPEG-2 – the most widely used video coding standard today – while keeping the cost within an acceptable range. In July, 2004, a new amendment was added to this standard, called the Fidelity Range Extensions (FRExt, Amendment 1), which demonstrates even further coding efficiency against MPEG-2, potentially by as much as

3:1 for some key applications [28].

These differences were the motivation for our work on designing a model capable of accurately simulating the behavior of multiplexed MPEG-4 and H.264 videoconference traffic.

The rest of this work is organized as follows. In Chapter 2 we propose a new revenue-based call admission control mechanism based on accurate medium and high quality MPEG-4 traffic prediction. In Chapter 3, we present a new video traffic model for H.264 videoconference streams. Finally, Chapter 4 concludes this work.

Chapter 2

Revenue-Based Call Admission Control for MPEG-4 Wireless Videoconference Traffic

Next generation wireless technologies will need to incorporate new sets of traffic control procedures in order to cope with the challenges related to supporting both the existing and the ever-increasing new multimedia services. Emerging wireless networks, in particular, will need to accommodate significant loads of traffic related to real-time video services, and especially videoconference traffic [29]. The QoS requirements of video users are particularly strict. The reason is that video packet transmission delays and the subsequent video packet dropping when the delay exceeds an upper bound result in the viewer's annoyance. Call Admission Control (CAC) is a strategy used to limit the number of call connections into the network in order to reduce network congestion, therefore enabling the system to provide the desired QoS to newly incoming as well as existing calls. In wireless cellular networks, in particular, the traffic conditions in the cells can change very quickly due to user mobility; also, when mobile users change their point of attachment (handoff), the end-to-end path may be changed while they still expect to receive the same QoS. An efficient CAC mechanism should be able to cope with this strict user requirement.

In recent work, in [30], such an efficient CAC scheme for high quality MPEG-4 videoconference traffic was proposed. Like many CAC schemes in the literature, the scheme in [30] adopted the idea of a probabilistic service, as described in [31]. This type of service does not provide for the worst-case scenario, but instead guarantees a bound on the rate of lost/delayed packets based on statistical characterization of the

traffic. However, [30] did not adopt the standard method of implementation for this service type which is the use of an "equivalent bandwidth" estimation, larger than the average rate but less than the peak rate of the sources. Although widely used, "equivalent bandwidth"-based schemes are known to significantly overestimate the sources' actual bandwidth requirements and therefore to provide quite conservative CAC schemes, which fail to use efficiently all the available bandwidth [32][33]. Instead, [30] proposed the use of our recent modeling approach in [6] for traffic originating from MPEG-4 videoconference sources, in order to design a new CAC scheme for wireless cellular networks which uses the traffic parameters (peak, mean, standard deviation) which the video source either declares at call setup or has agreed on in its contract with the wireless provider, in order to precompute a large number of traffic scenarios for its decision-making.

In the present work we introduce a new model for medium quality MPEG-4 videoconference traces. We then proceed to take advantage of the accuracy of our modeling approach both for medium quality, in this work, and high quality (in [6]) MPEG-4 videoconference traffic, in order to propose a new CAC scheme for wireless cellular networks which makes decisions not only based on the system's ability to accommodate newly arriving users in terms of capacity, but also on the profit that can be made by the provider if existing users are degraded in order for new video calls to be accepted.

2.1 Medium Quality MPEG-4 Videoconference Traffic Model

We have studied the medium quality version of three different long sequences of MPEG-4 encoded videos from [26]. The difference between the medium and the high

quality encoding of the movies lies in the quantization parameters used in each case, in [26]. The three traces ("Office Cam", "Lecture Room Cam", "Boulevard Bio") are movies with low motion. We have investigated the possibility of modeling the traces with a number of well-known distributions (gamma, lognormal, log-logistic, exponential, geometric, Weibull, Pearson V). Our results (derived with the use of Q-Q plots [34], Kolmogorov-Smirnov (KS) tests [34] and Kullback-Leibler (KL) tests [35]) have shown that, similarly to our work in [6] on modeling high quality MPEG-4 videoconference traffic, the best fit among these distributions for modeling a single movie is achieved for all traces examined with the use of the Pearson type V distribution (also known as the inverted gamma distribution). The Pearson type V 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 [34]. Since all of these distributions have been often used for video traffic modeling in the literature, they have been included in this work as fitting candidates, in order to compare their modeling results in the case of MPEG-4 videoconferencing.

The data for each trace consists of a sequence of the number of cells per video frame. The length of the videos varies from 45 to 60 minutes. Table 1 presents the trace statistics for each trace, in the Medium Quality (MQ) columns. Table 1 also contains the trace statistics for the high quality version of the same movies, as both versions will be needed in our work on an efficient CAC scheme. The sets of parameters of the traces comprise the "modes" adopted by videoconference users in our study. This will be further explained in Section 2.1.

Movie	Mean (Mbps)		Peak (Mbps)		Standard Deviation (Mbps)		Revenue Weight	
	HQ	MQ	HQ	MQ	HQ	MQ	HQ	MQ
Office	0.4	0.11	2	1	0.434	0.253	6	2
Lecture	0.21	0.058	1.5	0.69	0.182	0.094	4	1
Boulevard Bio	0.65	0.19	2.6	1.3	0.368	0.197	8	3

Table 1. Statistics for the High and Medium Quality versions of the video traces.

The Probability Density Function (PDF) of a Pearson type V distribution with parameters (α, β) is

$$f(x) = \frac{x^{-(a+1)} e^{\frac{-\beta}{x}}}{\beta^{-a}\Gamma(\alpha)} \tag{1}$$

for all x>0, and zero otherwise. The mean and variance are given by the equations:

$$Mean = \frac{\beta}{\alpha - 1} \tag{2}$$

$$Variance = \frac{\beta^2}{(\alpha-1)^2(a-2)} \tag{3}$$

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 quite good fit, the fit was not perfectly 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. 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 [16]. 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. The Pearson V distribution once again provided the best

fitting results for all types of video frames' sequences, and the modeling results were much improved in comparison with those of modeling the trace as a whole. We present, indicatively, the results from our KS-tests for the I, P and B frames of the lecture trace in Figures 1-3. The KS-test tries to determine if two data sets differ significantly and has the advantage that it makes no assumption about the distribution of data, i.e., it is non-parametric and distribution-free. It uses the maximum vertical deviation between the two curves as its statistic D. 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; for this reason we confirmed our modeling results with Q-Q plots and KL-tests. The Q-Q plot is a powerful goodness-of-fit test, 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-degree 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 ($(X \leq x) = z$). The Kullback-Leibler divergence test (KL-test) is a measure of the difference between two probability distributions.

The results of our KS-tests, in Figures 1-3 show that the Pearson V distribution is the best fit, as it has the smallest maximum vertical deviation from all the distributions. Similar results were deducted by all our statistical tests.

The goal of our work in this Section is to build a model which, based on the good but not perfect fit of the Pearson V distribution for modeling a single movie, will accurately capture the behavior of multiplexed medium quality MPEG-4 videoconference movies from VBR coders.

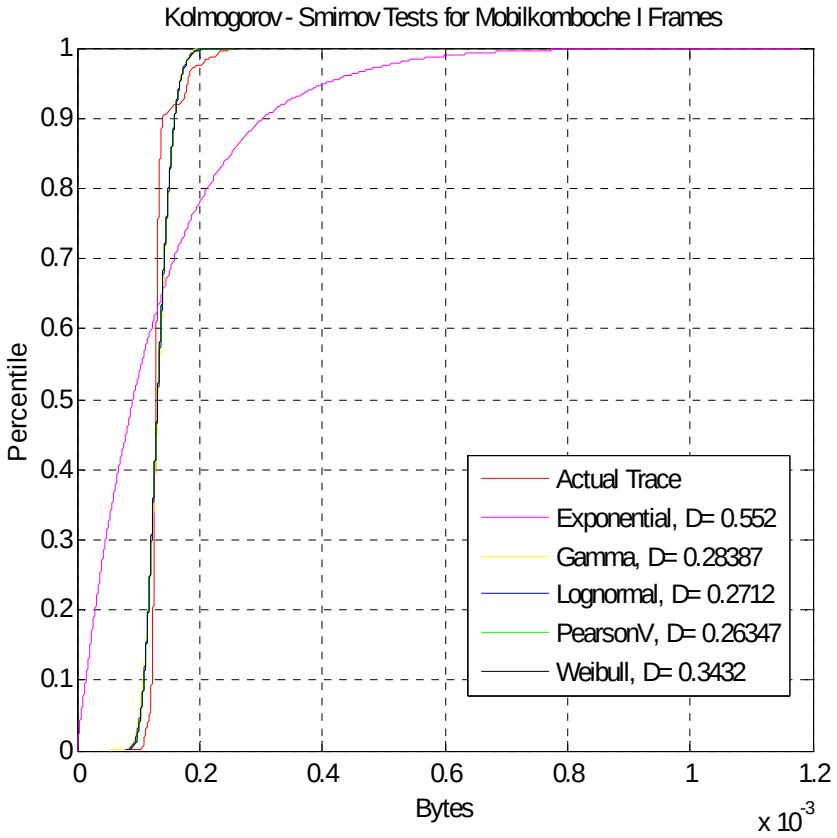

Figure 1. K-S test for the Lecture movie I frames.

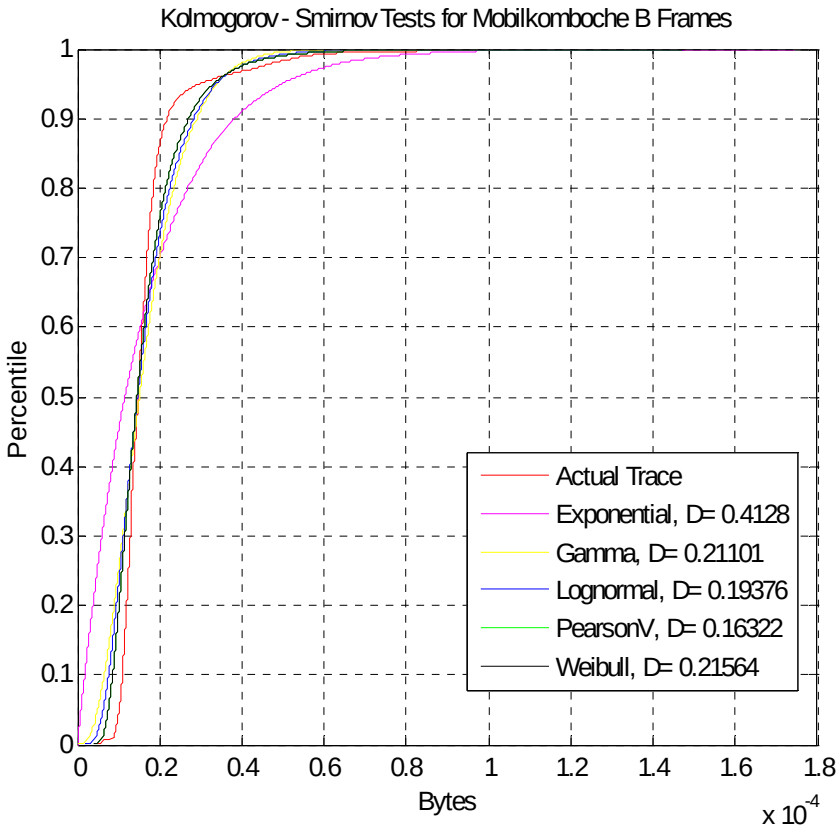

Figure 2. K-S test for the Lecture movie B frames.

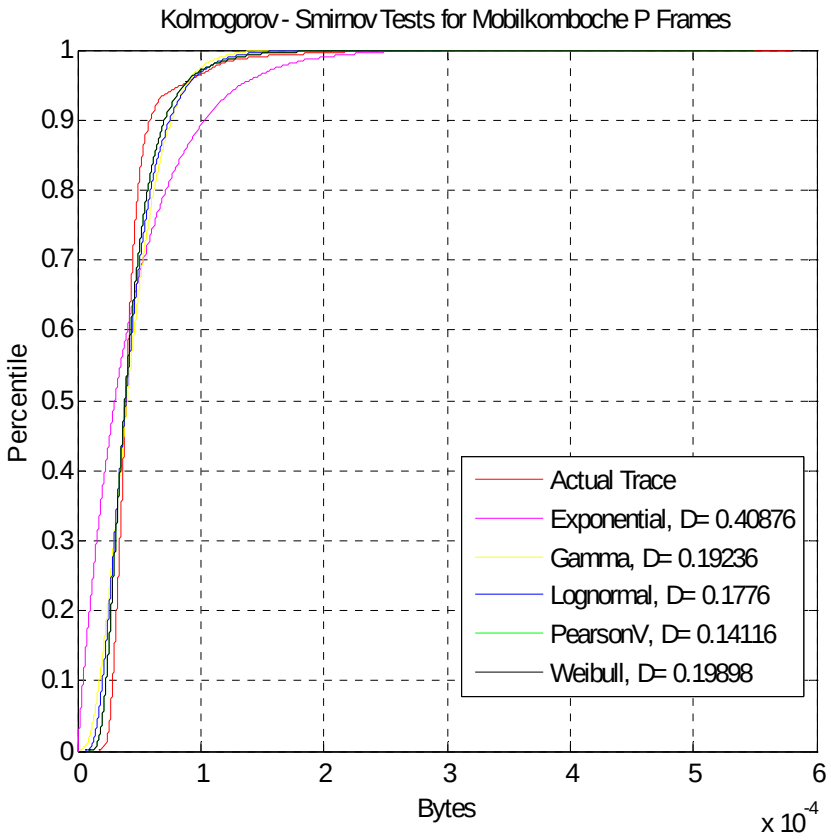

Figure 3. K-S test for the Lecture movie P frames.

2.1.1 DAR Model

A Discrete Autoregressive model of order p, denoted as DAR(p) [36], 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

$$P\{Y_n = i\} = \pi(i) \tag{4}$$

The sequence of random variables $\{X_n\}$ which is formed according to the linear model:

$$X_n = V_n X_{n-1} + (1 - V_n) Y_n \tag{5}$$

is a DAR(1) process.

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

$$\boldsymbol{P} = \rho \boldsymbol{I} + (1-\rho)\, \boldsymbol{Q} \tag{6}$$

where ρ is the autocorrelation coefficient, $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{Q}$ is a matrix with

$$Q_{ij} = \pi(j) \text{ for } i, j \in S. \tag{7}$$

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

$$\rho(W) = \frac{\mathrm{E}\left[\left(X_i - \mu\right)\left(X_{i+w} - \mu\right)\right]}{\sigma^2} \tag{8}$$

where μ is the mean and σ^2 the variance of the frame size for a specific video trace. In our model the rows of the Q matrix consist of the Pearson type V probabilities ($f_0, f_1, \ldots f_k, F_K$), where $F_K = \sum_{k>K} f_k$, and K is the peak rate. Each k, for $k<K$, corresponds to possible source rates less than the peak rate of K.

Figure 4 presents the Markov chain corresponding to the DAR(1) model.

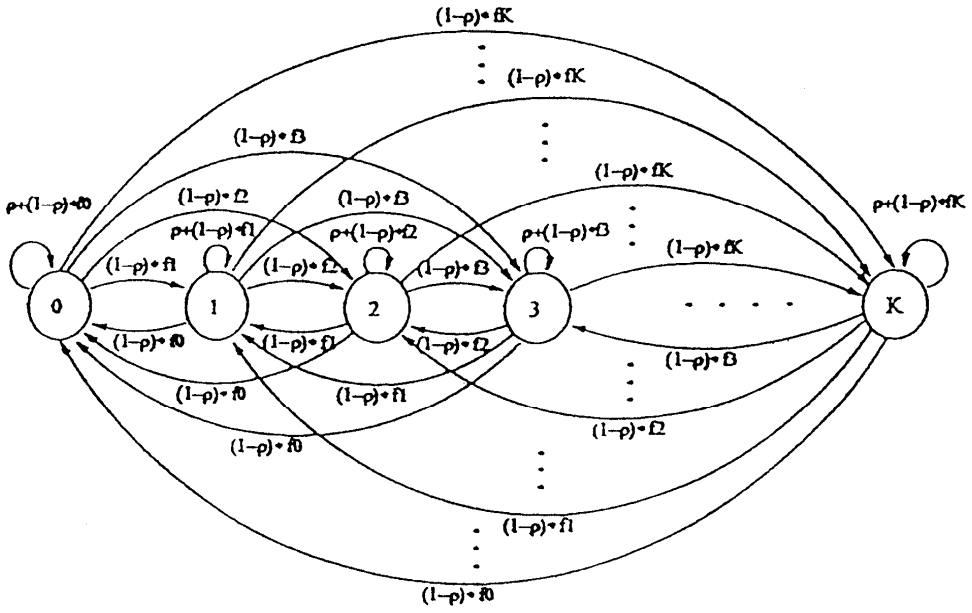

Figure 4. Markov Model for video traffic.

So, we build for each video frame type a model based only on the above parameters which are either known at call set-up time or can be measured without introducing much complexity in the network. In short, 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 (which is typically very high for videoconference sources). We proceeded with testing our models statistically (with the same methods used for single traces) in order to study whether it produces a good fit for the trace superposition. The accurate fits in our results have shown that the superposition of the actual traces can be modeled well by a respective superposition of data produced by our modeling approach.

In Figure 5, we have plotted the 0.01-, 0.02-, 0.03-,… quantiles of the actual office camera trace versus the respective quantiles of the DAR(1) model for the superposition of 30 traces' P frames. As shown in the Figure, the points of the Q-Q plot fall either very close or completely along the 45-degree reference line (which would correspond to a perfect match of the actual trace quantiles), with the exception of the first and last 3% quantile (left and right-hand tail), for which the DAR(1) model greatly overestimates the probability of frames with a very large number of cells. The very good fit (shown from all our results, which are similar in nature to those in [6]) shows that the superposition of the actual traces can be accurately modeled by a respective superposition of data produced by the DAR(1) model.

The accuracy of our modeling approach will be used in our proposal for a revenue-based CAC scheme for next generation wireless cellular networks.

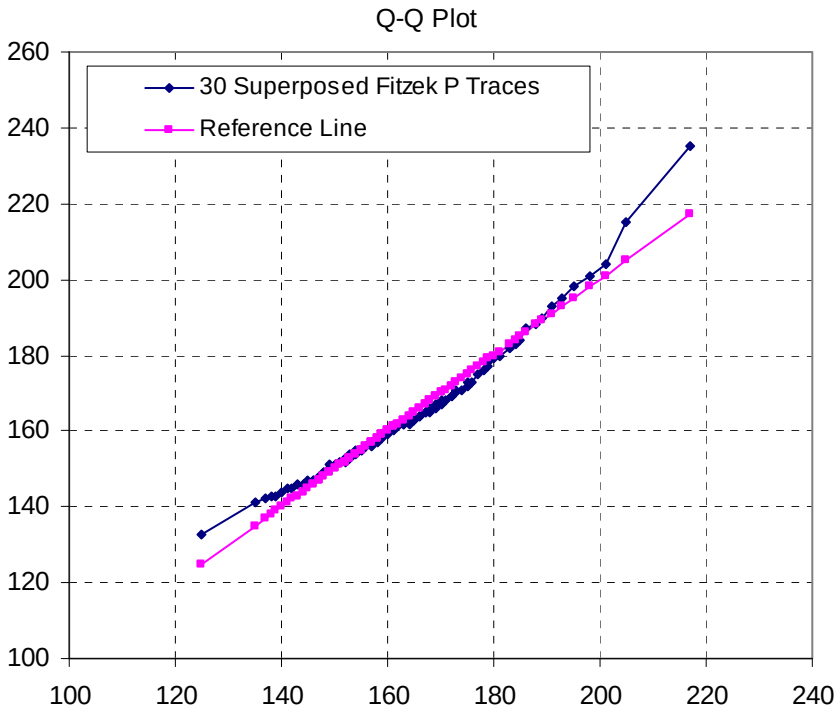

Figure 5. Q-Q plot of DAR(1) model versus the actual office camera trace for the P frames of 30 superposed sources. Values in both axes are in packets.

2.2 Channel Error Model

Errors in the wireless channel due to noise typically occur over relatively short bursts and are highly correlated in successive slots, but uncorrelated over long time windows. In our simulation study we adopt a channel error model similar to the widely studied Gilbert-Elliot [37] model, where the channel switches between a "good state" and a "bad state", but with the modification that the good state is not always error-free, neither the bad state always with errors. In our model, we consider a case where the error probability in the bad state is in the order of 104 times larger than the error probability in the good state (this ratio is taken from one set of parameters in [38]).

The parameters of our model are: Pr (good) =0.99994, Pr (good-bad) =0.000006, Pr (bad-good) =0.1, Error probability (good state) =$0.4*10^{-4}$, Error probability (bad state) = 0.7. The total probability of a transmission error occurring is equal to: Pr (good)*Error

prob (good) + Pr (bad)*Error prob (bad) = $8.2*10^{-5}$. That is, the total probability of a transmission error is only slightly smaller than the maximum acceptable video packet dropping probability of 10^{-4} [39] which we consider in this work for real-time videoconference traffic, making the need for very efficient call admission control imperative.

2.3 The Revenue-Based Call Admission Control Mechanism

The reason for which the scheme in [30] excelled in comparison to the "equivalent bandwidth" approach was the use of the accurate MPEG-4 videoconference traffic model from our work in [6], in order to precompute various traffic scenarios and combine that knowledge with online simulation, in order to be able to make accurate decisions on the acceptance or rejection of a new call. The precomputation, along with the online simulation, was made in [30] based on the traffic parameters declared by the video sources at call setup. These parameters are used for the "identification" of the source as a user adopting a specific "mode". In order to explain what a "mode" is, we first note that a logical assumption for next generation wireless networks is that videoconference users will be allowed to adopt one of a few specific "modes", each corresponding to a set of traffic parameters. Therefore, we used in our work each user's declared set of parameters in order to examine the respective precomputed traffic scenario, based on our MPEG-4 model for a source with such a set of parameters. This approach is especially plausible for wireless videoconference traffic, as the number of variations between source bandwidth requirements is naturally restricted by the type of application (a much larger pool of "modes" would have to be used in the case of regular video traffic). In [40] it was shown that the approach used in [30] works equally well for

H.263 videoconference traffic and excels, again, in comparison to the equivalent bandwidth approach.

Similarly to the work in [30], we denote in the present work as "modes" for the MPEG-4 videoconference users the sets of traffic parameters presented in Table 1. Hence, we use six "modes" for MPEG-4 videoconference traffic, one high and one medium quality "mode" for each one of the three traces. High-paying users adopt the high quality (HQ) modes, due to the increased bandwidth these modes offer.

The CAC scheme proposed in [30] focused only on the system's ability to accommodate a newly arriving user in terms of the total channel capacity which is needed for all terminals after the inclusion of the new user. In the case when the channel load, with the admission of a new call, was precomputed (or computed online) to be higher than the channel information rate, users which had agreed in their contracts on the possibility of degradation were gradually degraded up to the point where the new call could be admitted. One parameter not included in [30] was that in a real-life scenario, the decision of admitting or rejecting a new call in the network will be made by the provider not only based on the capacity needed to accommodate the call, but also on the revenue that the admission of the new call will provide. That is, if the admission of a new call (and the subsequent increase in bandwidth utilization) can only be made with the degradation of a higher-paying customer who enjoys higher QoS, the CAC module should compute whether this is a profitable decision.

For this reason, in our new CAC scheme we not only adopt the idea of [30] for precomputation and online computation of various traffic scenarios and implement it for MPEG-4 traffic, but we also assign "revenue weights" to each one of the six MPEG-4 "modes", thereby differentiating them into different service classes. These weights are shown in Table 1, and are assigned in an ad-hoc manner here, without loss of generality,

based on the traffic parameters of each “mode”. Therefore, users adopting the high quality “Boulevard Bio” MPEG-4 videoconference mode are the ones demanding the highest QoS and paying respectively for it, followed by users adopting the high quality “office” mode and the high quality “lecture” mode; users adopting the MQ versions of the traces are the low-paying users. Users choose one of the six “modes” with a probability when they enter the system (in Section 5 we will discuss our results when altering this probability). The assignment of the weights is not linear, since there are significant differences among the traffic parameters of each “mode”. We consider that 50% of the users of the high quality “office” and high quality “lecture” modes can accept degradation to a lower quality mode. By “lower quality mode” we refer to the MQ mode of the same movie. Users who have adopted the high quality “Boulevard Bio” mode are considered to be the highest paying users, therefore we assume that only a small percentage of them (20%) accept degradation. The choice of the percentages of users who accept degradation is again used here indicatively, and a change in these percentages would not alter the nature of our results.

Our CAC scheme uses the traffic models presented in Section 2 for high quality and in [6] for medium quality MPEG-4 traffic, in order to precompute a number of traffic scenarios. Naturally, not all traffic scenarios can be precomputed, due to the very large number of all possible traffic loads; still, with the use of an adequate number of precomputed scenarios and our accurate video model, when a non-precomputed traffic load occurs in the system an online simulation can be conducted relatively quickly by our system in order to compute the “deviation” between the bandwidth needed currently and the “closest” (in terms of the synthesis of modes) precomputed traffic scenario. This new traffic scenario will then be added into the CAC scheme’s database of precomputed scenarios.

As already explained, our scheme does not make its decision based only on the maximization of system capacity, as in [30] and in most CAC schemes in the literature, but also on the maximization of provider revenue. Therefore, the current revenue R is computed as:

$$R = \sum_i N_i W_i \tag{9}$$

where N_i is the total number of video users of "mode" i, and W_i is the revenue from each user of "mode" i (shown in the last columns of Table 1). Then our proposed CAC algorithm proceeds with the following steps, at the arrival of a new user request.

1. Identify the User Service Class ("Mode").
2. Check precomputed scenarios or compute online the Total Bandwidth (TB) needed for all video calls, with the inclusion of the new call
3. ***if*** TB ≤ System Global Rate ***then***
4. Admit Call, Compute new revenue R*
5. ***else***
6. ***if*** new user accepts degradation ***then***
7. Degrade New User to its MQ mode
8. Check precomputed scenarios or compute online the TB needed for all video calls, with the inclusion of the new degraded call
9. ***if*** TB ≤ System Global Rate ***then***
10. Admit New Degraded Call, Compute new revenue R*
11. ***else***
12. Check precomputed scenarios or compute online to find the smallest number Y of users of lesser priority "modes" or equal priority "modes" with the new user, who would need to be degraded for TB ≤ System Global Rate. Procedure starts with users of lesser priority mode.
13. ***if*** Y>0 ***then***
14. ***if*** new user is handoff ***then***
15. Degrade All Y users
16. Admit New Degraded Call
17. ***else***
18. Compute the revenue R**, which can be acquired for this Y and with the acceptance of the new user.
19. ***if*** R**>R ***then***
20. Degrade All Y users
21. Admit New Degraded Call
22. ***else***
23. Reject New Call
24. ***end if***
25. ***end if***
26. ***else*** (Y= 0, which means that step 12 cannot be satisfied)
27. Reject New Call
28. ***end if***
29. ***end if***
30. ***else*** (new user does not accept degradation)
31. Repeat (steps 12-28)
32. ***end if***
33. ***end if***

The logic of the CAC algorithm is that, when a new video user arrives (either from within the picocell or from handoff), the system first checks whether it can be accommodated in terms of the total bandwidth which will be needed when the user is multiplexed with the existing users in the system. If this is not possible, the algorithm attempts to degrade the user, if the user accepts degradation. The rationale behind this

decision is that the arrival of a new user should cause the minimum possible number of degradations, and hence irritation, to users who are already in the system, therefore it is preferable that the new user is accepted with degradation. One point which needs to be stressed here is that in most of the relevant works in the literature (including [30][40]), it is commonly accepted that handoff calls have absolute priority in obtaining an equal amount of channel bandwidth as the one they were occupying in their previous picocell location, i.e., handoff calls are not expected to endure any quality degradation, as this would lead to user dissatisfaction. We take a different approach in this work. It is indeed crucial for a handoff user to not experience call dropping when moving from one picocell to the next, as this would lead to significant user irritation (call dropping is much more irritating than the blocking of the call of a new user who attempts to transmit). However, if the mobile user experiences, during handoff, a degradation for which he has agreed in his contract, this should not be a cause for user irritation and therefore is allowed in our algorithm. If after degradation (of either a new or a handoff videoconference call) the acceptance of the call is still not possible, the CAC scheme checks all possibilities of degrading users of the same or lesser priority of the new call in order to accommodate it. If such a possibility exists and the call comes from handoff, it is accepted. If, however, it is a new call originating from within the picocell, it will be accommodated only if its acceptance will lead to higher revenue; otherwise, even if the total bandwidth that will be used with the acceptance of the new call is larger than the bandwidth previously used, there is no reason to degrade a significant number of users (and cause them even a slight irritation) if the provider will receive no extra revenue. In the case that the new call does not accept any degradation, the attempt is still made to degrade lesser or equal priority users who are already in the system, and a new call from within the picocell is again accepted only if it leads to higher revenue.

Regarding the applicability of our scheme as a more general proposal for wireless cellular networks, it needs to be stressed that it is actually not necessary to adopt the same approach (i.e., of basing the CAC scheme on traffic models) for all types of flows in the wireless network. If an accurate model exists for video traffic, which is the most bursty type of traffic in the network, this would be enough for the provision of a revenue-based CAC scheme which will be much less conservative than the equivalent bandwidth probabilistic service approach. The remaining types of flows (e.g., voice and data flows), which are much less greedy in terms of their bandwidth requirements, could be admitted based simply on their declared mean rate, or with any other of the many efficient approaches proposed in the literature.

2.4 Results and Discussion

Fourth generation mobile data transmission rates are planned to be up to 20 Mbps, therefore in this work we study a channel of this rate. The maximum allowed transmission delay for the video packets of a Video Frame (VF) is equal to the time before the arrival of the next VF, with packets being dropped when the deadline is reached (the interframe period in MPEG-4 encoded movies is 40 ms).

Our scheme is evaluated in 12 different scenarios versus the actual traffic generated by the real video traces, under handoff loads ranging from 5% to 15% of the total traffic (a handoff call can belong to any of the "modes" with equal probability). The first scenario is one in which the six "modes" in our study are used with a 16.67% probability (i.e., a user which "wakes up" chooses one of the six "modes" with equal probability). When studying this scenario, we found that the maximum number of users that the system could accommodate, without violating the strict QoS requirement of 0.01% maximum video packet dropping, was 51. As it will be shown in Table 3, this

corresponds to less than 75% utilization of the channel capacity. The burstiness of video traffic is responsible for the system's inability to accommodate more sources without violating their QoS requirements. In each one of the other traffic scenarios studied (scenarios 2-12), we have considered various combinations of the cases where each one of the six modes is selected by users with one of the probabilities: 10%, 20%, 30%, 40% and the total number of users present in the system is 51 (we have chosen to keep the maximum number of traces equal to the maximum that can be achieved when all modes have equal probability). The percentages shown in Table 2 refer to the initial "mode" with which a user enters the network; this mode may change due to degradation, based on the user's contract.

Each simulation point is the result of an average of 10 independent runs, each simulating one hour of network operation. Table 3 presents in its first column the bandwidth which is actually needed by the video traces, in its second column the estimated bandwidth that the traces will need based on our DAR(1) modeling approach, and in its third column the bandwidth that is utilized with the use of our CAC scheme. Two significant conclusions can be drawn from the Table. The first is that the estimation provided by our mechanism yields an overestimation of the actual bandwidth requirements of the superposed sources (the reasons for this have to do with an overestimation of the I frames size and are explained in our modeling work in [6]); still, this overestimation is small and ranges in all simulated scenarios from a minimum of 3.3% to a maximum of 7.92%. The average overestimation provided by our scheme is 5.04% over all the studied scenarios, which is acceptable, especially given that a small overestimation of the actual bandwidth requirements of video traffic is usually preferable, in order for the system to cope with the bursty nature of video users. The second conclusion has to do with the efficiency of our CAC scheme. By comparing the

three respective columns for 5%, 10% or 15% it is clear that the actual bandwidth utilized when our revenue-based CAC scheme is enforced is smaller than the bandwidth needed by the real traces, and hence also smaller than the estimated bandwidth with the use of the DAR(1) model. The reason is that some of the MQ videoconference calls are rejected from the system in order to achieve higher revenue for the provider; hence, with the use of the DAR(1) modeling approach we reserve slightly more bandwidth than actually needed, and then, with the use of the CAC scheme we hinder a number of MQ users from accessing the system in order to keep HQ users continuously content with the service they are receiving (i.e., they seldom need to be degraded).

In all the studied scenarios, the maximum system throughput is achieved in the case of 15% handoff traffic. This is expected, as in this case a larger number of video traces are accommodated by the system, by the gradual degradation of HQ users to medium video quality. The maximum throughput achieved is 76.55%, in Scenario 3.

The high precision in our scheme's predictions can show even better results with the use of a slightly larger pool of videoconference "modes" from which traffic scenarios will be precomputed. The use of a slightly larger number of modes will guarantee the existence of a larger variety of parameter sets, so that an incoming call's traffic parameters will always be well-matched with those of one of the modes in the pool.

We further investigate our mechanism's performance in the results presented in Figure 6, where we present in the y axis both the estimation provided by the DAR model and the actual bandwidth usage from our revenue-based CAC scheme; we also indicatively present the estimation provided by the use of the equivalent bandwidth approach from [41]. All the above are presented versus the normalized real system utilization; this indicates the actual traffic load generated by the traces, normalized to the channel capacity, e.g., a traffic load equal to 40% represents 40% of the 20 Mbps

uplink capacity, i.e., 8 Mbps system throughput (these loads have been created with different combinations of probabilities for the six "modes" under study, and the results presented are the average over all the combinations used). As shown in the Figure, the equivalent bandwidth estimation significantly overestimates the actual traffic load in all cases (this agrees with the results in [30][40]). Additionally, the points of the plot of the DAR model estimation fall very close to the 45-degree reference line, showing that our estimation is in all cases very accurate. Finally, and most importantly, the Figure shows once again that the use of our CAC scheme leads to a slight underallocation to the videoconference users, in comparison to their offered load. The reason is that some of the MQ videoconference calls are rejected from the system in order to achieve higher revenue for the provider. As the offered load (normalized over the 20 Mbps channel capacity) increases, this under-allocation increases as well, in order not to allow low-paying users to fill the channel capacity at the cost of degrading high-paying ones.

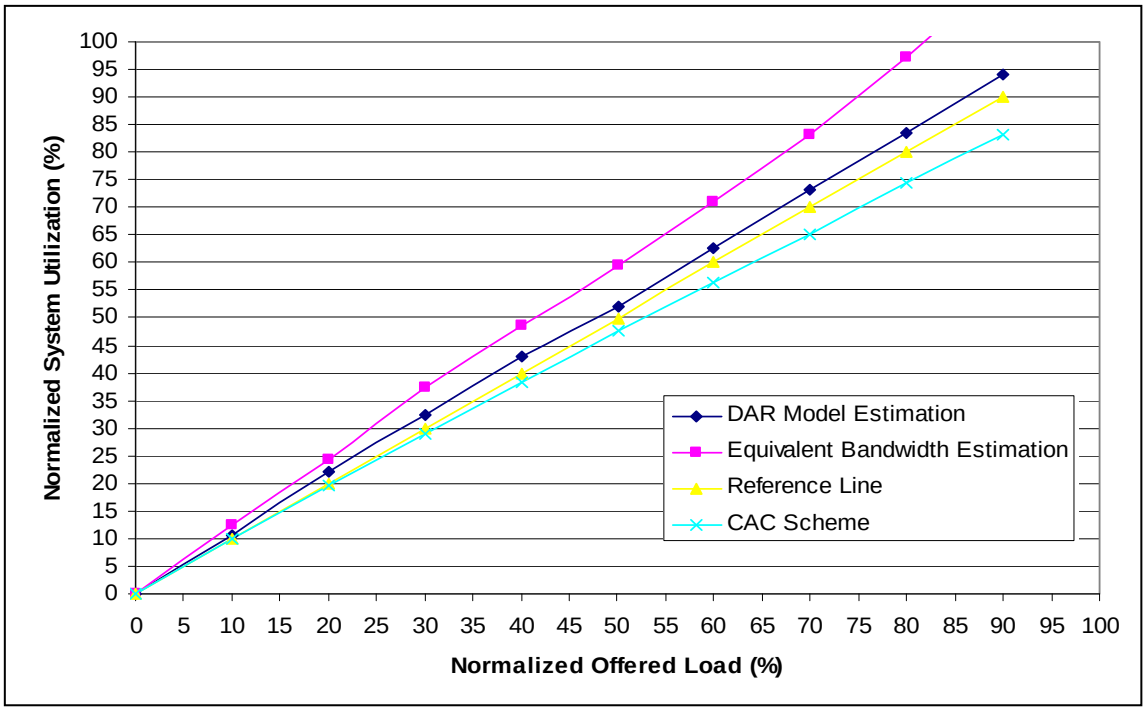

Figure 6. Capacity Utilization with the CAC scheme, with 10% handoff traffic.

Scenario	Office		Lecture		Boulevard Bio	
	HQ	MQ	HQ	MQ	HQ	MQ
1	16.67	16.67	16.67	16.67	16.67	16.67
2	10	30	20	10	10	20
3	20	10	20	10	20	20
4	30	10	20	20	10	10
5	10	20	10	10	10	40
6	40	20	10	10	10	10
7	20	10	30	20	10	10
8	10	30	10	10	20	20
9	10	10	10	20	20	30
10	10	10	40	10	20	10
11	20	20	20	20	10	10
12	10	30	10	20	20	10

Table 2. Traffic Scenarios – Percentages of HQ and MQ "modes".

Scenario	Real Traces - Bandwidth (Mbps) under various handoff loads (%)			DAR Model - Bandwidth (Mbps) under various handoff loads (%)			Actual Bandwidth Used with the CAC Scheme (Mbps) under various handoff loads (%)		
	5%	10%	15%	5%	10%	15%	5%	10%	15%
1	14.55	14.81	14.98	15.26	15.84	16.27	14.27	14.42	14.52
2	11.48	11.72	12.14	12.05	12.52	12.96	11.24	11.35	11.70
3	15.44	15.85	16.19	16.07	16.45	16.81	14.75	15.03	15.31
4	14.23	14.51	15.20	14.96	15.21	15.82	13.67	13.88	14.28
5	11.87	12.23	12.63	12.28	12.93	13.44	11.73	11.99	12.34
6	14.93	15.19	15.61	15.47	15.95	16.47	14.12	14.39	14.65
7	13.07	13.28	13.49	13.72	13.97	14.64	12.42	12.63	12.85
8	13.78	14.01	14.42	14.25	14.78	15.19	13.59	13.78	14.20
9	13.83	14.12	14.64	14.34	14.96	15.35	13.61	13.93	14.36
10	14.71	15.21	15.62	15.29	15.94	16.52	14.03	14.48	14.81
11	12.30	12.62	13.17	12.76	13.39	13.91	11.92	12.16	12.64
12	13.06	13.57	14.07	13.67	14.25	14.88	12.85	13.21	13.72

Table 3. Estimations of the required bandwidth and bandwidth utilization with the CAC scheme.

Chapter 3

Modeling Video Traffic Originating from H.264 Videoconference Streams

3.1 Single-Source H.264 Traffic Modeling

3.1.1 Frame-size histograms

We have studied two different long sequences of H.264 VBR encoded videos in eighteen formats, from the publicly available Video Trace Library of [42]. The selected videos are of 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 two traces are:

a. A demo from the Sony Digital Video Camera

b. An excerpt of NBC News

The length of the videos is 10 and 30 minutes, respectively. 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. Without loss of generality, we use 48-byte packets throughout this work, but our modeling mechanism can be used equally well with packets of other sizes. Table 4 presents the statistics for each trace. The interframe period is 33.3 ms.

We have investigated again, as in the case of medium quality MPEG-4 videoconference traffic, the possibility of modeling the eighteen traces with quite a few well-known distributions and our results show once again that the best fit among these distributions is achieved for all the traces studied with the use of the Pearson type V distribution.

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. Figure 7 presents indicatively the histogram for the NBC News ([CIF, G16, B7, F28]) sequence.

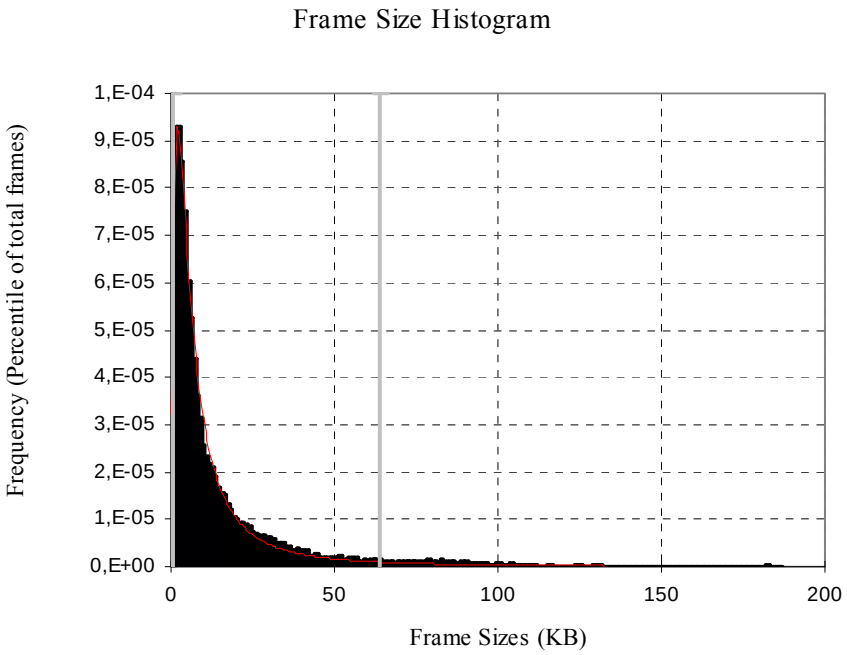

Figure 7. Frame size histogram for the NBC News trace with parameters: [CIF, G16, B7, F28].

Video Name	**[Resolution, G, B, F]**	**Mean (bits)**	**Peak (bits)**	**Std. Deviation (bits)**
NBC News	[CIF, 16, 1, 28]	15816	181096	21705,24
NBC News	[CIF, 16, 1, 48]	1197	28032	2219,26
NBC News	[CIF, 16, 3, 28]	14632	182520	21631,47
NBC News	[CIF, 16, 3, 48]	1084	28216	2237,75
NBC News	[CIF, 16, 7, 28]	15081	186872	21613,23
NBC News	[CIF, 16, 7, 48]	1054	29768	2275,84
NBC News	[CIF, 16, 15, 28]	16624	192272	21365,03
NBC News	[CIF, 16, 15, 48]	1059	31840	2290,61
Sony Demo	[CIF, 16, 1, 28]	14067	221664	27439,89
Sony Demo	[CIF, 16, 1, 48]	954	23096	2166,51
Sony Demo	[CIF, 16, 3, 28]	12801	222888	27752,93
Sony Demo	[CIF, 16, 3, 48]	887	23232	2203,77
Sony Demo	[CIF, 16, 7, 28]	13129	227680	28053,90
Sony Demo	[CIF, 16, 7, 48]	898	25480	2289,92
Sony Demo	[CIF, 16, 15*, 28]	14861	233296	28338,22
Sony Demo	[CIF, 16, 15*, 48]	933	28224	2412,26
Sony Demo	[HD, 12, 2, 48]	22513	398544	51815,57
Sony Demo	[HD, 12, 2, 38]	7618	143408	18103,28

G: GoP Size, B: Number of B Frames, F: Quantization Parameters

*When B=15 and G=16 there are no P frames in the trace sequence

Table 4. Trace Statistics

3.1.2 Statistical Tests and Autocorrelations

Our statistical tests were made with the use of Q-Q plots [34][8], Kolmogorov-Smirnov [34] tests and Kullback-Leibler divergence tests [35].

The Pearson V distribution fit was shown to be the best in comparison to the gamma, weibull, lognormal and exponential distributions, which are presented here (comparisons were also made with the negative binomial and Pareto distributions, which were also worse fits than the Pearson V). However, as already mentioned, although the Pearson V was shown to be the better fit among all distributions, the fit is not perfectly 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 H.264-encoded streams, based on the cyclic GoP formats (therefore, this case is different than the case of H.263 traffic we studied in [43], where the number of I frames was so small in each trace that the trace could be modeled as a whole).

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 in [16][44].

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 [9][19] that the average sizes of P and B frames can vary by 20% and 30% (often more than that), respectively, in subsequent scenes, therefore the size changes are statistically significant.

The mean, peak and variance of the video frame sizes for each video frame type (I, P and B) of each movie were taken again from [42] and the Pearson type V parameters are calculated based on the equations (2) and (3), for the mean and variance of Pearson V,

respectively (the parameters for the other fitting distributions are similarly obtained based on their respective formulas).

The autocorrelation coefficient of lag-1 was also calculated for all types of video frames of the eighteen movies, as it shows the very high degree of correlation between successive frames of the same type. The autocorrelation coefficient of lag-1 will be used in the following Sections of this work, in order to build a Discrete Autoregressive Model for each video frame type.

From the five distributions examined (Pearson V, exponential, gamma, lognormal, weibull) the Pearson V distribution once again provided the best fitting results for the 54 cases (18 movies, 3 types of frames per movie) studied.

In order to further verify the validity of our results, we performed Kolmogorov-Smirnov and Kullback-Leibler tests for all the 54 fitting attempts. The results of our tests confirm our respective conclusions based on the Q-Q plots (i.e., the Pearson V distribution is the best fit). Figure 8 presents indicative results from the KS-test. Regarding the KL-test, the results for the {I, P, B} frames of the Sony Demo ([CIF, G16, B3, F48]) trace are respectively, for the Pearson V distribution {0.364, 0.721, 0.432}, for the Lognormal distribution {0.378, 0.864, 0.479}, for the Gamma distribution {0.387, 1.027, 0.543} and for the Weibull distribution {0.453, 1.024, 0.533}.

Although controversy persists regarding the prevalence of Long Range Dependence (LRD) in VBR video traffic ([45][46][47]), in the specific case of H.264-encoded video, we have found that LRD is important. The autocorrelation function for the NBC News ([CIF, G16, B7, F28]) trace is shown in Figure 9 (the respective Figures for the other three traces are similar). Three apparent periodic components are observed, one containing lags with low autocorrelation, one with medium autocorrelation and the

other lags with high autocorrelation. We observe that autocorrelation remains high even for large numbers of lags and that both components decay very slowly; both these facts are a clear indication of the importance of LRD. The existence of strong autocorrelation coefficients is due to the periodic recurrence of I, B and P 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, the high autocorrelation shown in the Figure above 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, these results lead us to extend our work in order to build a DAR model, which inherently uses the autocorrelation coefficient of lag-1 in its estimation. The model will be shown to accurately capture the behavior of multiplexed H.264 videoconference movies, by generating frame sizes independently for I, P and B frames.

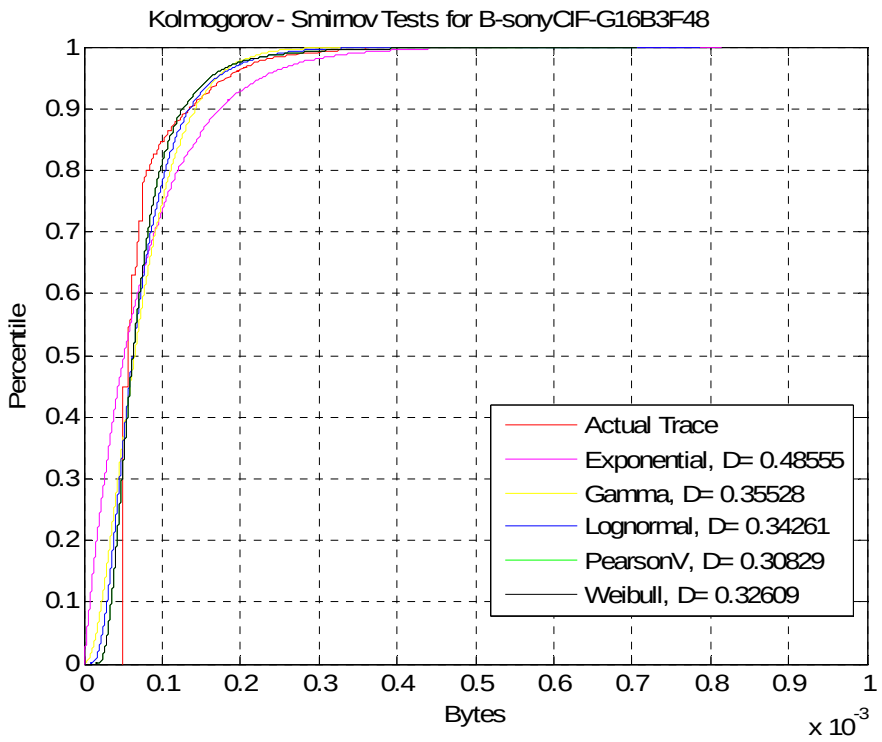

Figure 8. KS-test (Comparison Percentile Plot) for the Sony Demo B frames ([CIF, G16, B3, F48]).

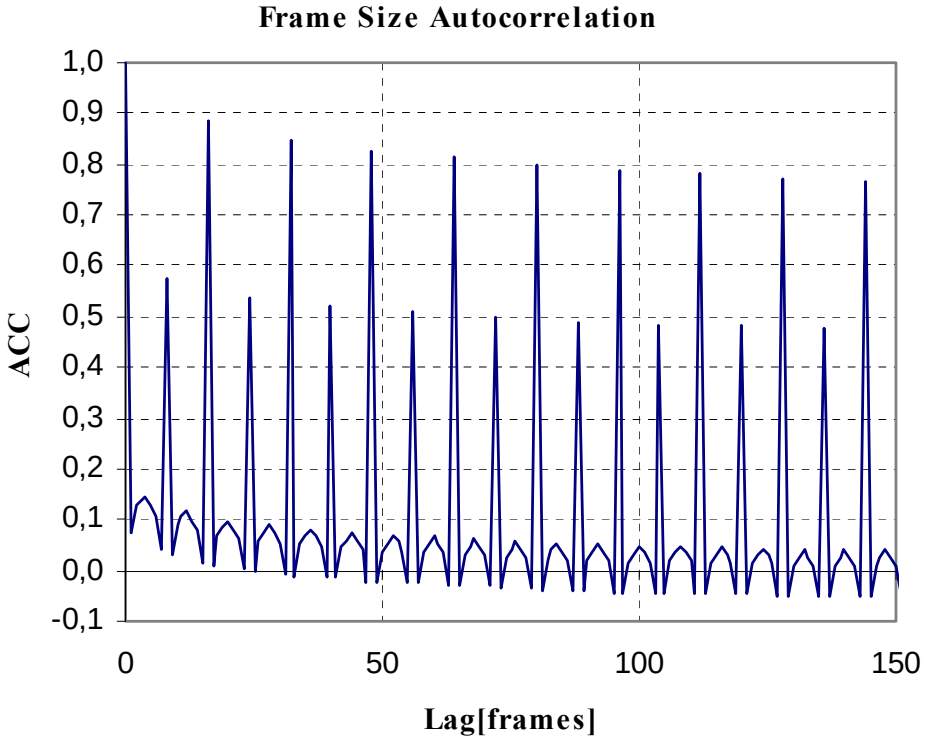

Figure 9. Autocorrelation Coefficients of the NBC News trace ([CIF, G16, B7, F28]).

3.2 The DAR (1) Model – Results and discussion

As in [8], 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. As already mentioned in Section 2.1.1, 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). The DAR(1) model can be used with any marginal distribution [48].

As already explained, the lag-1 autocorrelation coefficient for the I, P and B frames of each trace is very high in all the studied cases. Therefore, we proceeded to build a DAR(1) model for each video frame type for each one of the eighteen traces under study. More specifically, in our model the rows of the Q matrix consist of the Pearson type V probabilities ($f_0, f_1, \ldots f_k, F_K$), where $F_K = \sum_{k>K} f_k$, 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 (6) 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 our model with a probability slightly larger than ρ. This is evident in Figure 10, where we compare the actual B frames sequence of the NBC News ([CIF, G16, B15, F28]) trace and their respective DAR(1) model and it is shown that the DAR(1) model's 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 NBC News 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 other traces under study.

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 5 or more sources, and are almost completely smoothed out in most cases, as the number of sources increases (the authors in [8] have reached similar conclusions for their own DAR(1) model and they present results for a superposition of 20 traces). This is clear in Figures 11-13, which present the comparison between our DAR(1) model and the actual I, P, B frames' sequences of the NBC News ([CIF, G16, B1, F28)] , for a superposition of 30 traces (the results were perfectly similar for all video frame types of the other three traces; we have used the initial trace sequences to generate traffic for 30 sources, by using different starting points in the trace). The common property of all 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 for I frames.

However, although Figures 11-13 suggest that the DAR(1) model captures very 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 it produces 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 Figures 14-15, where we have plotted the 0.01-, 0.02-, 0.03-,… quantiles of the actual P and I video frames' types of the NBC News trace versus the respective quantiles of the

respective DAR(1) models, for a superposition of 30 traces.

As shown in Figure 14, which presents the comparison of actual P frames with the respective DAR(1) models for the NBC News ([CIF, G16, B3, F48]) trace, the points of the Q-Q plot fall almost completely along the 45-degree reference line, with the exception of the first and last 3% quantiles (left- and right-hand tail), for which the DAR(1) model underestimates the probability of frames with a very small and very large, respectively, number of cells. The very good fit shows that the superposition of the P frames of the actual traces can be modeled very well by a respective superposition of data produced by the DAR(1) model (similar results were derived for the superposition of B frames), as it was suggested in Figures 12, 13. Figure 15 presents the comparison of actual I frames with the respective DAR(1) model, for the NBC News ([CIF, G16, B7, F48]) trace. Again, the result suggested from Figure 11, i.e., that our method for modeling I frames of multiplexed H.264 videoconference streams provides only relative accuracy, is shown to be valid with the use of the Q-Q plots. The results for all the other cases which are not presented in Figures 14-15 are similar in nature to the ones shown in the Figures.

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, H.264 videoconference streams show LRD. This problem is overcome by our choice of modeling I, P and B frames separately. This is shown in Figure 16. It is clear from the Figure that, even for a small number of lags, (e.g., larger than 10) the autocorrelation of the superposition of frames decreases quickly, for all the traces. Therefore, although in some cases the DAR(1) model exhibits a slower decrease than that of the actual traces' video frames sequence, this has minimal impact on the fitting quality of the DAR(1) model. This result further supports our choice of using a first-order model.

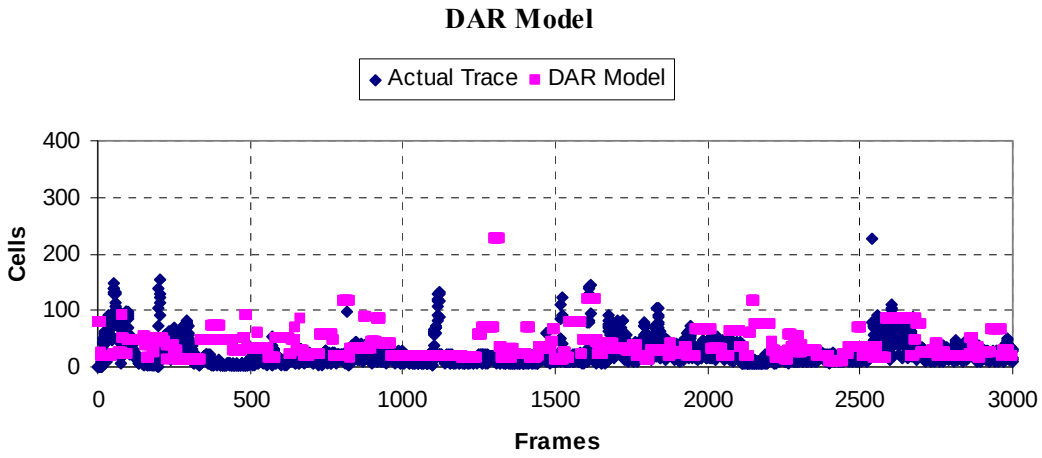

Figure 10. Comparison for a single trace between a 10000 frame sequence of the actual B frames sequence of the NBC News ([CIF, G16, B15, F28]) trace and the respective DAR(1) model in number of cells/frame (Y-axis).

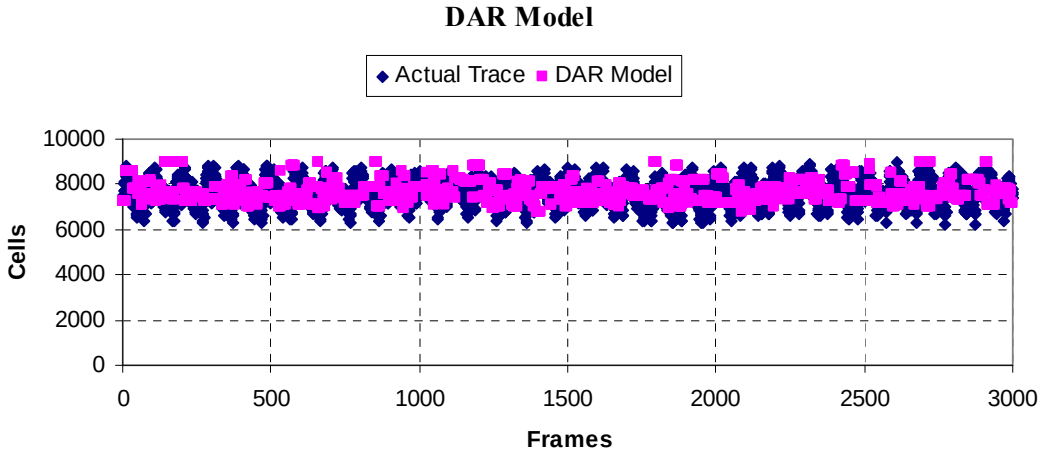

Figure 11. Comparison for 30 superposed sources between a 3000 I frame sequence of the actual NBC News ([CIF, G16, B1, F28)] trace and the respective DAR(1) model in number of cells/frame (Y-axis).

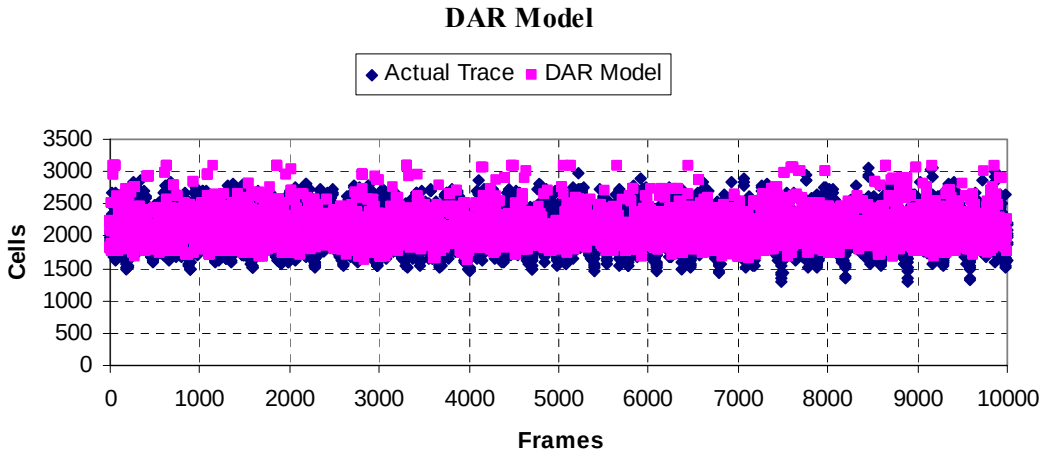

Figure 12. Comparison for 30 superposed sources between a 10000 P frame sequence of the actual NBC News ([CIF, G16, B1, F28)] trace and the respective DAR(1) model in number of cells/frame (Y-axis).

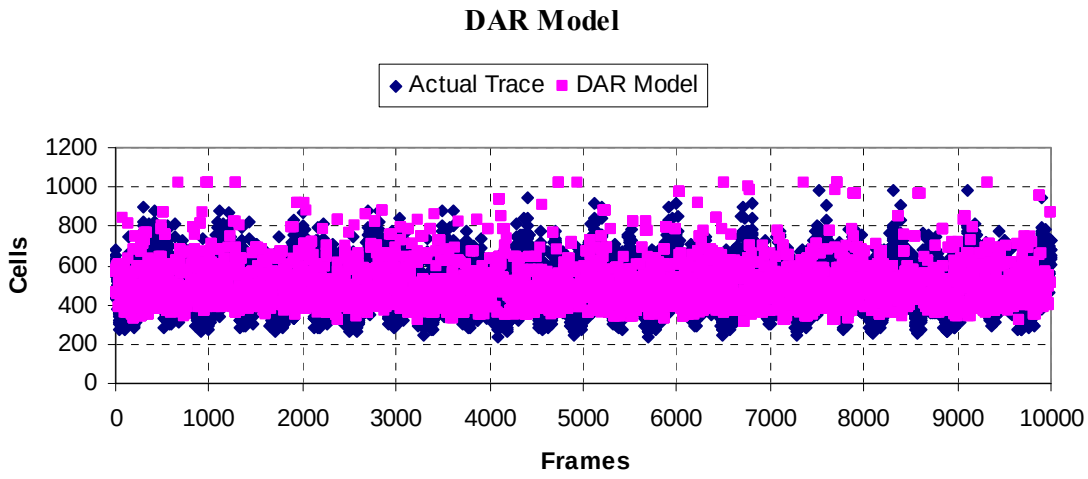

Figure 13. Comparison for 30 superposed sources between a 10000 B frame sequence of the actual NBC News ([CIF, G16, B1, F28)] trace and the respective DAR(1) model in number of cells/frame (Y-axis).

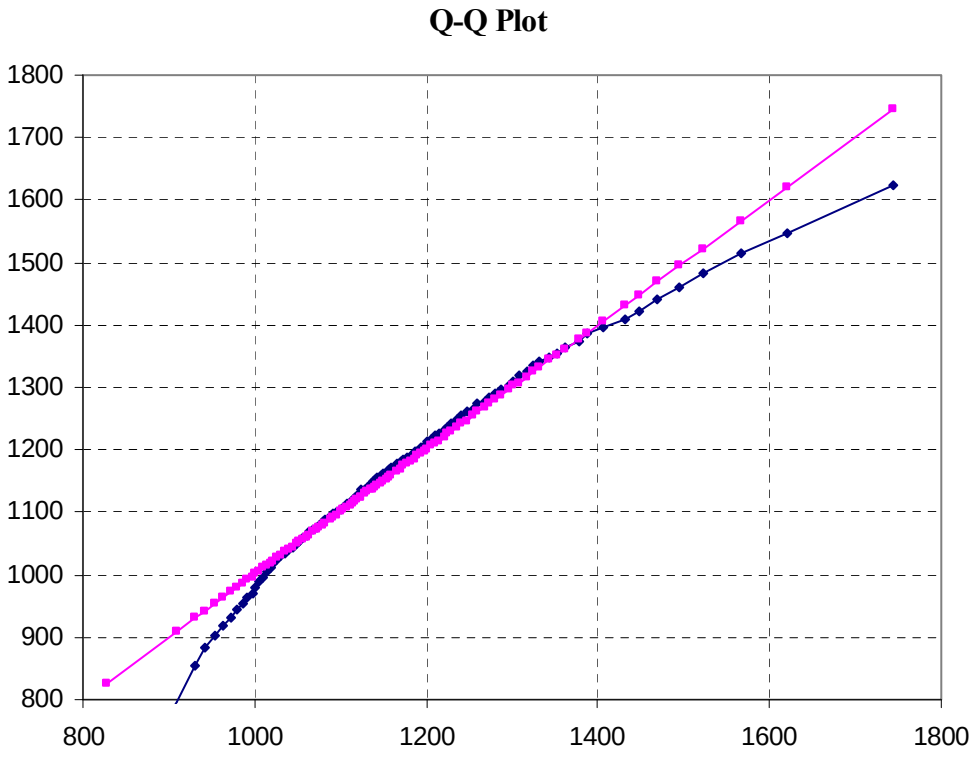

Figure 14. Q-Q plot of the DAR(1) model versus the actual video for the P frames of NBC News ([CIF, G16, B3, F48]), for 30 superposed sources.

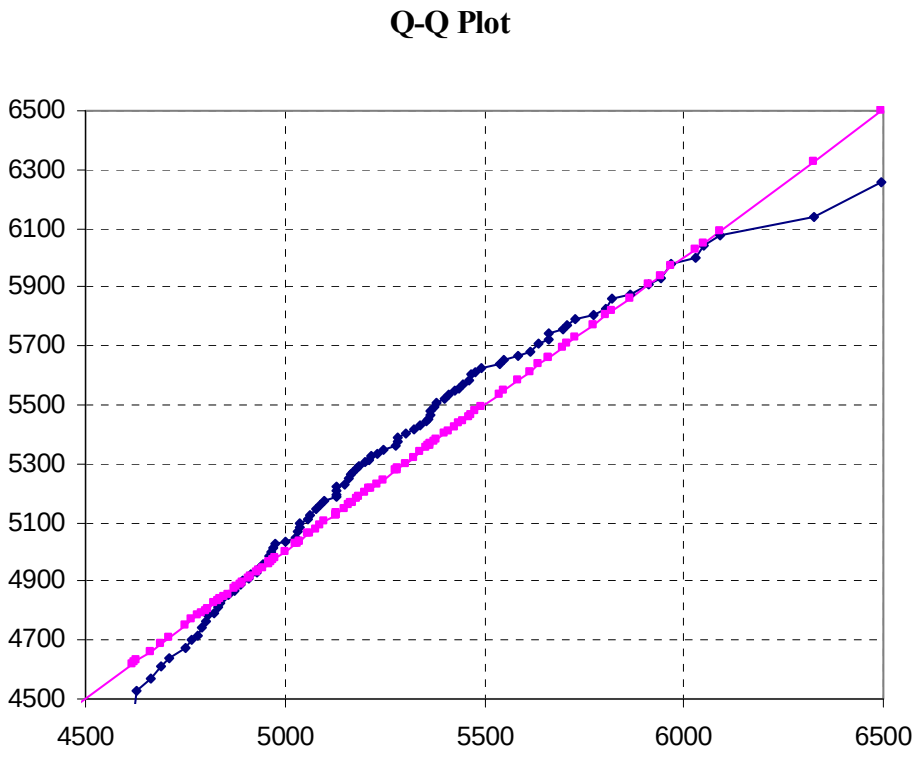

Figure 15. Q-Q plot of the DAR(1) model versus the actual video for the I frames of NBC News ([CIF, G16, B7, F48]), for 30 superposed sources.

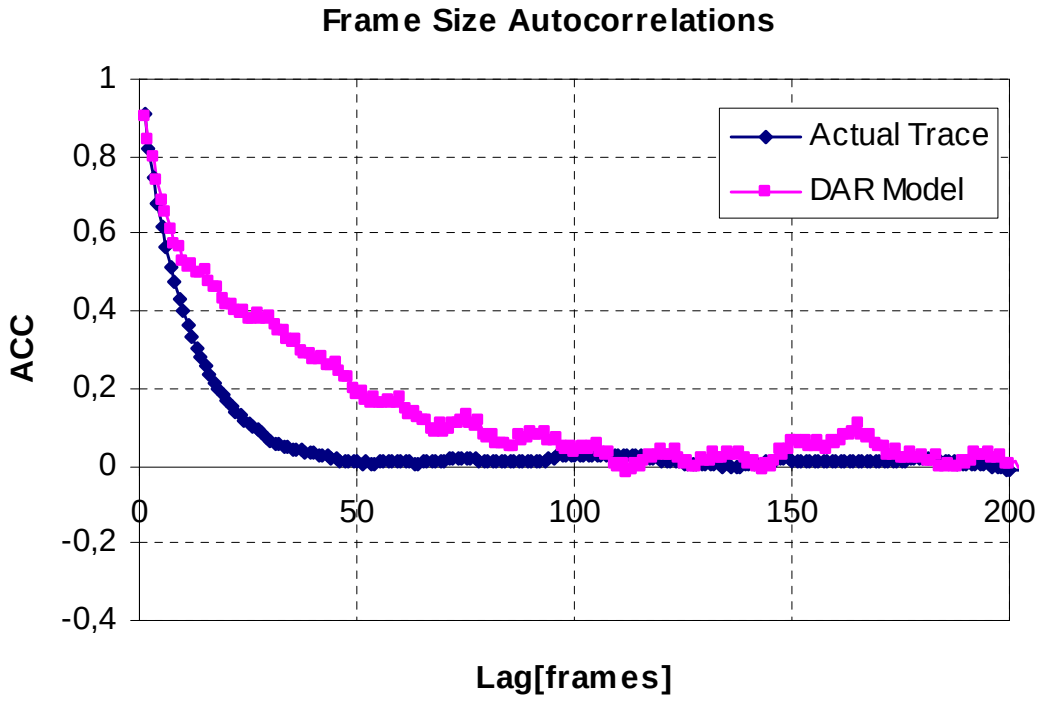

Figure 16. Autocorrelation vs. number of lags for the I frames of the actual NBC News ([CIF, G16, B15, F28]) trace and the DAR(1) model, for 30 superposed sources.

Chapter 4

Conclusions

In the first part of this work we have proposed a highly efficient new Call Admission Control mechanism for MPEG-4 videoconference traffic transmission over next generation cellular networks. The novelty of the scheme lies in the utilization of precomputed traffic scenarios for decision-making on the acceptance or rejection of a new MPEG-4 videoconference call. The precomputation is based on the traffic parameters declared by the video source at call setup; these parameters are used for the "identification" of the source as a user adopting a specific "mode" from the pool of "modes" which have provided the basis for the precomputation of our traffic scenarios. Then, this precomputation is used in combination with an algorithm which maximizes the provider revenue, in order to make decisions on the acceptance or rejection of a new videoconference call.

In the second part of this work, we have proposed and tested a new model for traffic originating from multiplexed VBR H.264 videoconferencing sources. Our work is one of the first in the relevant literature, to the best of our knowledge, to focus especially on H.264 videoconference traffic; the significance of the specific problem, in comparison to H.264 video sources in general, lies in the fact that videoconference traffic exhibits much higher autocorrelation.

Initially, we have investigated the possibility of modeling single H.264 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 H.264 videoconference

traffic. We then proceeded to use the Pearson type V distribution, which was shown to provide the relatively best fit for single H.264 videoconference sources, in order to build a Discrete Autoregressive model (separately for I, B, and P frames) for multiplexed H.264 sources. Our DAR model was shown to combine high accuracy with simplicity. Based on the very accurate results of our study in modeling P- and B-frames' sizes of multiplexed H.264 videoconference traces, and the low complexity of our first-order model, we believe that our approach is very promising for modeling this type of traffic. However, since our modeling scheme shows relative accuracy in modeling I -frames' sizes, the use of wavelet modeling for the I -frames' size sequence (similarly to the highly accurate but computationally heavy wavelet-based approach in [26]) may provide a very competent solution, and our future work will be pointed towards this direction, as well as towards using our H.264 modeling scheme for a Call Admission Control mechanism which will take into consideration not only the total bandwidth required by requesting users and the possible provider revenue, but will also make decisions of call acceptance/rejection based on user irritation.

Bibliography

[1] T. Stockhammer, H. Jenkac and G. Kuhn, "Streaming Video over Variable Bit-Rate Wireless Channels", IEEE Transactions on Multimedia, Vol. 6, No. 2, 2004, pp. 268-277.

[2] D. M. Lucantoni, M. F. Neuts and A. R. Reibman, "Methods for Performance Evaluation of VBR Video Traffic Models," IEEE/ACM Trans. Networking, Vol. 2, 1994, pp. 176-180.

[3] N-E. Rikli, "Modeling Techniques for VBR Video: Feasibility and Limitations", Performance Evaluation, Vol. 57, No. 1, 2004, pp. 57-68.

[4] E. Gelenbe, X. Mang and R. Onvural, "Bandwidth Allocation and Call Admission Control in High-Speed Networks", IEEE Communication Magazine, Vol. 35, No. 5, 1997, pp. 122-129.

[5] E. Gelenbe, X. Mang and R. Onvural, "Diffusion Based Call Admission Control in ATM", Performance Evaluation, Vol. 27-28, 1996, pp. 411-436.

[6] A. Lazaris, P. Koutsakis and M. Paterakis, "A New Model for Video Traffic Originating from Multiplexed MPEG-4 Videoconference Streams", Performance Evaluation, Vol. 65, No. 1, 2008, pp. 51-70.

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

[8] D. P. Heyman, A. Tabatabai and T. V. Lakshman, "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, 1992, pp. 49-59.

[9] A. M. Dawood and M. Ghanbari, "Content-based MPEG Video Traffic Modeling", IEEE Transactions on Multimedia, Vol. 1, No. 1, 1999, pp. 77-87.

[10] B. Melamed and D. E. Pendarakis, "Modeling Full-Length VBR Video Using Markov-Renewal Modulated TES Models", IEEE Journal on Selected Areas in Communications, Vol. 16, No. 5, 1998, pp. 600-611.

[11] K. Chandra and A. R. Reibman, "Modeling One- and Two-Layer Variable Bit Rate Video", IEEE/ACM Transactions on Networking, Vol. 7, No. 3, 1999, pp. 398-413.

[12] Q. Ren and 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, Vol. 16, No. 5, 1998, pp. 679-691.

[13] S. Kempken and W. Luther, "Modeling of H.264 High Definition Video Traffic Using Discrete-Time Semi-Markov Processes", in Proceedings of ITC 2007, Lecture Notes in Computer Science (LNCS) Vol. 4516, pp. 42–53, 2007.

[14] M. Dai and D. Loguinov, “Analysis and Modeling of MPEG-4 and H.264 Multi-Layer Video Traffic”, in Proceedings of the IEEE Infocom 2005.

[15] I. Reljin, A. Samcovic and B. Reljin, “H.264/AVC Video Compressed Traces: Multifractal and Fractal Analysis”, EURASIP Journal on Applied Signal Processing, Vol. 2006, Issue 1, 2006, pp. 1-12.

[16] M. Frey, S. Ngyuyen-Quang, “A Gamma-Based Framework for Modeling Variable-Rate Video Sources: The GOP GBAR Model”, IEEE/ACM Transactions on Networking, Vol. 8, No. 6, 2000, pp. 710-719.

[17] D. P. Heyman, “The GBAR Source Model for VBR Videoconferences”, IEEE/ACM Transactions on Networking (1997), Vol. 5, No. 4, 1997, pp. 554-560.

[18] M. Krunz and 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 and D. Sarkar, “Modeling Full-Length Video Using Markov-Modulated Gamma-Based Framework,” IEEE/ACM Trans. on Networking, Vol. 11, No.4, 2003, pp. 638-649.

[20] D. Arifler and 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.

[21] A. Golaup and A. H. Aghvami, “Modelling of MPEG4 Traffic at GoP Level Using Autoregressive Processes”, in Proceedings of the IEEE Vehicular Technology Conference-fall, 2002, Vancouver, Canada, pp. 854-858.

[22] D. P. Heyman, A. Tabatabai and T. V. Lakshman, “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, 1992, pp. 49-59.

[23] S. Gringeri, R. Egorov, K. Shuaib, A. Lewis and B. Basch, “Robust Compression and Transmission of MPEG-4 Video”, [Online] http://www.kom.e-technik.tu-darmstadt.de/acmmm99/ep/gringeri/

[24] 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.

[25] R. Schafer, “MPEG-4: A Multimedia Compression Standard for Interactive Applications and Services”, IEE Electronics and Communications Engineering Journal, Vol. 10, No. 6, 1998, pp. 253-262.

[26] F. H. P. Fitzek, M. Reisslein, “MPEG-4 and H.263 video traces for network performance evaluation”, IEEE Network, Vol. 15, No. 6, 2001, pp. 40-54.

[27] D. Marpe, T. Wiegand and G. Sullivan, “The H.264/MPEG4 Advanced Video Coding Standard and its Applications”, IEEE Communications Magazine, vol. 44, no. 8, Aug. 2006, pp. 134–143.

[28] G. J. Sullivan, P. Topiwala and A. Luthra, “The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions”, in Proc. of the SPIE Conference on Applications of Digital Image Processing XXVII, Special Session on Advances in the New Emerging Standard:H.264/AVC, Denver, USA, August 2004.

[29] M. Etoh and T. Yoshimura, “Advances in Wireless Video Delivery”, Proceedings of the IEEE, Vol. 93, (1), 2005, pp. 111-122.

[30] P. Koutsakis, “Efficient Call Admission Control for MPEG-4 Wireless Videoconference Traffic”, in Proc. of the IEEE ICCCN 2007, Honolulu, USA.

[31] H. Zhang and E. W. Knightly, “Providing end-to-end statistical performance guarantee with bounding interval dependent stochastic models”, in Proc. of the ACM SIGMETRICS, 1994, Nashville, USA.

[32] R. Guerin, H. Ahmadi and M. Naghsineh, “Equivalent capacity and its application to bandwidth allocation in high-speed networks”, IEEE Journal on Selected Areas in Communications, Vol. 9, No. 7, 1991, pp. 968-981.

[33] J. Q.-J. Chak and W. Zhuang, “Capacity Analysis for Connection Admission Control in Indoor Multimedia CDMA Wireless Communications”, Wireless Personal Communications, Vol. 12, 2000, pp. 269-282.

[34] A. M. Law and W. D. Kelton, “Simulation Modeling & Analysis”, 2nd Ed., McGraw Hill Inc., 1991.

[35] K. P. Burnham and D. R. Anderson, “Model Selection and Multi-Model Inference”, Springer-Verlag, New York, 2002.

[36] P. A. Jacobs and P. A. W. Lewis, “Time series generated by mixtures”, Journal of Time Series Analysis, Vol. 4, No. 1, pp. 19-36, 1983.

[37] E. N. Gilbert, “Capacity of a burst-noise channel”, Bell Systems Technical Journal, Vol.39, 1960, pp. 1253-1265.

[38] J. G. Kim and M. Krunz, “Effective Bandwidth in Wireless ATM networks”, in Proc. of the 4th ACM/IEEE MOBICOM, Dallas, USA, 1998, pp. 233-241.

[39] D. A. Dyson and Z. J. Haas, “A Dynamic Packet Reservation Multiple Access Scheme for Wireless ATM”, Mobile Networks and Applications (MONET) Journal, Vol. 4, No. 2, pp. 87-99, 1999.

[40] S. Chatziperis, P. Koutsakis and M. Paterakis, “A New Call Admission Control Mechanism for Multimedia Traffic over Next Generation Wireless Cellular Networks”, IEEE Transactions on Mobile Computing, Vol. 7, No. 1, 2008, pp. 95-112.

[41] W. Verbiest, L. Pinoo and B. Voeten, “The impact of the ATM concept on video coding”, IEEE Journal on Selected Areas in Communications, Vol. SAC-6, 1988, pp. 1623-1632.

[42] [Online] http://trace.eas.asu.edu/hd/index.html

[43] P. Koutsakis, “A New Model for Multiplexed VBR H.263 Videoconference Traffic”, in Proceedings of the 49th IEEE GLOBECOM 2006, San Francisco, USA.

[44] M. Krunz and H. Hughes, “A Traffic Model for MPEG-coded VBR Streams”, in Proceedings of the ACM SIGMETRICS 1995, pp. 47-55.

[45] K. Park and W. Willinger (editors), “Self-Similar Network Traffic and Performance Evaluation”, John Wiley & Sons, Inc., 2000.

[46] D. P. Heyman and T. V. Lakshman, “What are the Implications of Long-Range Dependence for VBR-Video Traffic Engineering”, IEEE/ACM Transactions on Networking, Vol. 4, No.3, 1996, pp. 301-317.

[47] B. K. Ryu and 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 1996, pp. 3-14.

[48] T. V. Lakshman, A. Ortega and A. R. Reibman, “VBR Video: Trade-offs and potentials”, Proceedings of the IEEE, Vol. 86, No. 5, 1998, pp. 952-973.