Available online at www.sciencedirect.com
ScienceDirect
Computer Communications 31 (2008) 1–4
computer communications
www.elsevier.com/locate/comcom

Short Communication

On modeling multiplexed VBR videoconference traffic from H.263 video coders

P. Koutsakis *

Department of Electrical & Computer Engineering, McMaster University, ITB-A318, 1280 Main Street West, Hamilton, Ont., Canada L8S 4K1

Received 14 April 2007; accepted 30 October 2007
Available online 12 November 2007

Abstract

Due to the burstiness of video traffic, video traffic modeling is very important in order to evaluate the performance of future wired and wireless networks. In this letter, we build a highly accurate Discrete Autoregressive model to capture the behavior of multiplexed H.263 videoconference movies from VBR coders.

Keywords: Videoconference traffic; Traffic modeling; H.263; Autoregressive models

1. Introduction

Videoconference traffic is expected to be a substantial portion of the traffic carried by emerging wired and wireless networks [8,12]. For Variable Bit Rate (VBR) coded video, statistical models are needed to design networks which are able to guarantee the strict Quality of Service (QoS) requirements of video traffic, in terms of video packet delays and video packet dropping.

H.263 is a video standard that can be used for compressing the moving picture component of audio-visual services at low bit rates [1]. Although H.264 has been recently introduced as the new video coding standard, H.263 is still very widely used [3,13]. The problem of modeling H.263 video traffic has been addressed in a few papers in the literature (e.g., [15,16]). Videoconference traffic, however, has some inherent characteristics (e.g., very high autocorrelation) which differentiate the problem of its modeling from that of modeling video traffic.

In this work, we build a model which accurately captures the behavior of multiplexed H.263 videoconference traces. To the best of our knowledge, this is the first work in the literature which addresses the specific problem. Very few other papers in the literature (e.g., [17]) have considered the same problem, but all of them have actually used low-rate versions of movies, i.e., not videoconference traces, for their work.

2. The Pearson type V distribution fit

In [2], Heyman et al. analyzed three H.261 videoconference sequences and found that the marginal distributions for all the sequences could be described by a gamma (or equivalently negative binomial) distribution. In [15], the authors study one videoconference sequence and another sequence, from a football game from a H.263+ coder with a high degree of compression and note that the negative binomial distribution provides a good *approximation* to the distribution of the number of bits per video frame.

In our work, we have studied three different long sequences (“Office Cam”, “ARD Talk”, “N3 Talk”, i.e., a camera showing the events within an office and two talk shows) of low motion H.263 VBR encoded videos, from Reisslein and Fitzek [4]. 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. We have investigated the possibility of modeling the traces with a

* Tel.: +1 9055259140.
E-mail address: polk@ece.mcmaster.ca

0140-3664/$ - see front matter
doi:10.1016/j.comcom.2007.10.042

number of well-known distributions and our results have shown that the *best* fit (not a perfect fit, but a very good one) 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 Probability Density Function (PDF) of a Pearson type V distribution with parameters (α, β) is $f(x) = [x^{-(\alpha+1)} e^{-\beta/x}]/[\beta^{-\alpha}\Gamma(\alpha)]$, for all $x > 0$, and zero otherwise. The mean and variance are given by the equations: $\text{Mean} = \beta/(\alpha - 1)$, $\text{Variance} = \beta^2/[(\alpha - 1)^2(\alpha - 2)]$. The parameters α, β for each trace were computed based on the trace's mean and variance.

To test, statistically, which distribution provides a good fit for the above traces, we have used *Q–Q* plots. The *Q–Q* plot is a powerful goodness-of-fit test [2,11], 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 $Pr(X \leqslant x) = z$).

In Fig. 1, we have plotted the 0.03-, 0.06-, 0.09-,..., quantiles of the actual "office cam" trace versus the respective quantiles of the various distribution fits (all points in both axes are in bytes). The results were common for all traces under study (the respective Figures are omitted here due to space limitations), showing that the Pearson V distribution fit is the best in comparison to the gamma, 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). The Pearson V distribution provides not only the best fit among all the examined distributions, but an excellent fit in itself, as for more than 90% of the distributions of the three studied traces, the points of the *Q–Q* plots fall either completely along the 45-degree reference line or very close to it. Still, the autocorrelation coefficient of videoconference traces is always very large, i.e., videoconference traffic is highly correlated between successive frames; this high autocorrelation can obviously 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.

3. The DAR(1) model

Autoregressive models have been used in the past to model the output bit rate of VBR encoders, e.g. [5,6]). A Discrete Autoregressive model of order p, denoted as DAR(p) [7], 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. A DAR(1) process is a Markov chain with discrete state space S and a transition matrix: $\mathbf{P} = \rho\mathbf{I} + (1 - \rho)\,\mathbf{Q}$ **(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 $i, j \in S$. Autocorrelations are usually plotted for a range W of lags. The autocorrelation is calculated by the formula $\rho(W) = E[(X_i - \mu)(X_{i+w} - \mu)]/\ \sigma^2$ **(2)**, where μ is the mean and σ^2 the variance of the frame size for a specific video trace. As in [2], 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 (which is typically very high for videoconference sources). According to [10], the DAR(1) model can be used with any marginal distribution.

More specifically, in our model the rows of the $\mathbf{Q}$ matrix consist of the Pearson type V probabilities (f_0, $f_1, \ldots f_k$, 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. The lag-1 autocorrelation coefficient ρ is estimated by Equation (2) to be equal to 0.943 for the office camera trace, 0.867 for the ARD Talk trace and 0.872 for the N3 Talk trace. From the transition matrix in (1) it is evident that if the current frame has, for

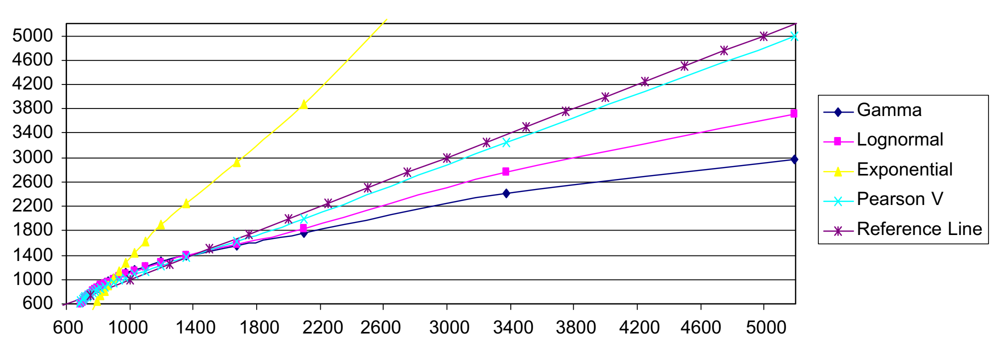

Fig. 1. *Q–Q* plot for the office camera trace.