Simulation Modelling Practice and Theory 75 (2017) 127–145

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Generation of synthetic video traffic using time series

Christos Katris*, Sophia Daskalaki

Department of Electrical & Computer Engineering, University of Patras, Rio 26504 Greece

ARTICLE INFO

Article history:
Received 9 June 2016
Revised 5 February 2017
Accepted 13 April 2017

Keywords:
Synthetic video traffic
Time series models
FARIMA
FARIMA-GARCH
Non-normality
FARIMA-to-LN

ABSTRACT

Generating traffic has always been an important part of network simulations but has turned to an even more challenging task with modern networks. The statistical properties of the input stochastic processes traced in the networks used all along Information Era turned out to be complicated and difficult to reproduce. Taking into account successful efforts in modeling Internet traffic with FARIMA time series models, this paper attempts to extend their applicability and employ them to generate synthetic video traffic. It is known that FARIMA can model both the Short Range (SRD) and Long Range Dependence (LRD) existing in video traffic; however the traces it produces fail to describe correctly the moments (mean, standard deviation, skewness, kurtosis) of the distribution behind the data. Since an efficient traffic generator should capture both the statistical properties and queuing behavior of video traffic we experiment with models such as FARIMA with Student's *t* errors and FARIMA-GARCH with Normal and Student's *t* errors, improving somewhat the accuracy of the generated traffic. Furthermore, the paper suggests the projection of the traces generated by a FARIMA model to values of a Lognormal distribution. It is shown that such a methodology produces synthetic traces that can emulate very closely the behavior of real traces. In order to quantify closeness the generated traces are fed into a simple FIFO queuing system with finite buffers, where loss probability is calculated and compared to that experienced by the corresponding real traces. Using five different real traces, MPEG-4 or H.263, it is shown that the proposed methodology produces traffic generators that can capture satisfactorily several statistical properties of the real traffic and also its queuing behavior for a wide range of buffer sizes and service rates.

1. Introduction

Generation of synthetic traffic is an important procedure quite necessary for performance evaluation when designing a network and for network management. Its necessity has become even more pressing due to the remarkable growth of video applications and multimedia users reported during the recent years. When reliable, synthetic traffic generators can be used with simulations to resolve problems in multimedia traffic management, such as bandwidth allocation, design of protocols and admission controls [1], while testing real networks for their performance can be expensive and sometimes prohibitive. For this reason it is preferable to use mathematical models and/or simulations to evaluate the performance of a network. In general, statistical models are considered better choice than trace-driven simulations since they can shed light to traffic characteristics, they are stochastic and not static, and can describe different situations when tuning the parameters in the models.

* Corresponding author.
E-mail addresses: chriskatris@upatras.gr (C. Katris), sdask@upatras.gr (S. Daskalaki).

http://dx.doi.org/10.1016/j.simpat.2017.04.005
1569-190X/

In order to generate synthetic video traffic using time series one should consider models that are suitable for this kind of traffic, such as FARIMA (Fractional ARIMA), since these models capture properties like SRD (Short Range Dependency) and LRD (Long Range Dependency) that normally video traffic appear to carry. According to Alheraish [2] a successful video traffic model should be able to capture the statistical characteristics of a real trace (mean, standard deviation, coefficient of variation, mode and autocorrelation) and also its dynamical behavior, e.g. to expose similar queuing behavior. It should be as simple as possible, generate traffic with low computational complexity and if possible characterize a wide range of video sources. According to Casilari et al. [3] approximating the probability density function (pdf) of the underlined random process is of utmost importance for matching the queuing behavior of a real traffic trace, while SRD is of secondary importance and LRD of even less. The method proposed in this paper satisfies relatively well the above criteria and certainly improves proximity to real traffic compared to the classical FARIMA models.

The main scope of the paper is to present a new procedure that is based on time series models and can be used for generating synthetic traffic for full motion video. Several VBR source models that have been used in the past are based on Autoregressive (AR) processes and a comprehensive description can be found in [4], along with a thorough and most recent survey on models for video traffic. Here we briefly refer only to those that have been used for video traffic generation. In general the AR approach is simple and requires few parameters. For full motion video, however, which could comprise frames of different types (I, B, or P), single AR models are not suitable and usually a combination of AR processes with different residual distributions for each frame type is required. As a consequence frame-based [5,6], scene-based and nested AR models [7] have been adopted. Of course this increases complexity to modeling and formulation. Other alternative variations of AR models include the DAR (Discrete Autoregressive) model of order p [8,9], GAR (AR model with Gamma residual process), GBAR (Gamma-Beta Autoregressive model) [10] and the General AR model that also includes GACS (Gaussian Autoregressive and Chi-Square) [11,12].

The projected AR (PAR) models form yet another interesting category of models for generating video traffic. They attempt to preserve autocorrelation and at the same time to fit the frame-size histogram. Previous work on such traffic models uses different types of AR processes according to the type of video and encoding scheme [13]. They appear to capture in general the autocorrelation behavior of a video and estimation of their coefficients from empirical data is relatively simple. On the other hand, it is not possible to find a single AR model that can capture different statistical characteristics and also there is no single video model that is suitable for all video sequences and for all purposes.

Moreover, self-similar models have been used successfully for video traffic modeling and generation. FARIMA models are the most widely used, individually [14,15] and in more composite frameworks [16,17,3]. Casilari et al. [3] considered briefly projections of the FGN (Fractional Gaussian Noise) and FARIMA models to a desired distribution. Apart from the time series models there has been a long standing effort for modeling traffic using Markov and Markov modulated models, most of them quite successfully for some type of video [e.g. [18,19]]. Their level of complexity is considered much higher than the level required from the AR family of models. Lastly, wavelets have formed another group of successful models for video traffic. Their success has been reported in a number of articles [e.g. [20,21]].

The ultimate goal of our approach is to provide a general framework, which can be adapted to several circumstances and is able to generate traffic that mimics real video traces, using only a certain time unit and very little information of the empirical traffic, such as the mean and standard deviation. In the core of our approach is the dependence structure (long and short range) that exists in video traffic, and for this reason we employ FARIMA models. On the other hand, for capturing the queuing behavior of real traffic, the marginal distribution of the underlying process should be matched. For this reason, we consider a 2-stage approach where an initially fitted FARIMA model to a video trace is transformed to a data series with a pre-determined target distribution. It builds upon the PAR types of models using FARIMA modeling in conjunction with the Lognormal distribution, which requires only two parameters and its CDF is easily invertible. With this approach we can accurately capture the marginal distribution of the traffic. Although this approach is superior to other FARIMA and ARIMA alternatives in capturing queuing performance of real traffic, we also have to assess its ability to preserve the SRD and LRD property of the FARIMA model. We perform this assessment experimentally.

The classical FARIMA model is usually very successful in capturing the LRD and SRD properties of a real traffic trace. However our experimentation showed that it cannot capture its marginal distribution, when it is far from normal, neither its non-linear dependencies, when they are present. To overcome this we considered as traffic generators FARIMA models with Student's t and Normal Inverse Gaussian error distributions, and also FARIMA-GARCH models with Normal, Normal Inverse Gaussian and Student's t errors. A number of these models achieved slight improvements for some traces, but none achieved a clear and definite improvement. As an alternative we develop a two-step traffic generating procedure which combines FARIMA models with Lognormal distribution through a transformation of percentiles. Such a procedure first embodies the desired SRD and LRD characteristics with the help of a FARIMA model and then attempts to approximate the pdf of a recommended distribution, the Lognormal in this case. The method is analogous to ARTA (AR To Anything) process [22,23] which combines an AR model with a desired marginal distribution and generates a series that displays the autocorrelation of the AR model and at the same time achieves a desired marginal distribution. The method is tested with simulated traces which include different characteristics, and then applied to real traces successfully compared with the aforementioned models.

Next, we describe analytically the procedure, argue on its efficacy and set up a queuing framework for evaluating its performance. The evaluation is performed experimentally and shows that transforming a FARIMA generated trace to variates of a Lognormal distribution matches sufficiently well real video traces. Section 2 presents the time series models that will