Efficient On-Line Generation of the Correlation Structure of F-ARIMA Processes

Maria-Estrella Sousa-Vieira, Andrés Suárez-González, José-Carlos López-Ardao, and Cándido López-García

Department of Telematics Engineering, University of Vigo, Spain
estela@det.uvigo.es

Abstract. Several traffic measurement studies have shown the presence of persistent correlations in modern networks. The use of stochastic processes able to capture this kind of correlations, as self-similar processes, has opened new research fields in network performance analysis, mainly in simulation studies, where the efficient synthetic generation of samples is one of the main topics. Although F-ARIMA processes are very flexible to capture both short- and long-range correlations in a parsimonious way, only off-line methods for synthesizing traces are efficient enough to be of practical use. In order to overcome this disadvantage, in this paper we propose a M/G/∞-based efficient and on-line generator of the correlation structure of F-ARIMA processes.

Keywords: F-ARIMA processes, M/G/∞ process, Correlation, Synthetic efficient on-line generation.

1 Introduction

Several traffic measurement results have convincingly shown the existence of persistent correlations in the traffic of modern networks [12,21,3,28,5,8,19]. These experimental findings stimulated the opening of a new branch in the stochastic modeling of traffic, since the impact of the correlation on the performance metrics may be drastic [22,26,23,11]. The use of classes of stochastic processes for traffic modeling purposes, displaying forms of correlation as diverse as possible by making use of few parameters, as self-similar processes, is important. Usually, real traces are of limited length and lack the necessary diversity required to perform simulation studies.

Some of these processes are Fractional Gaussian Noise [25] (FGN), Fractional AutoRegressive Integrated Moving Average [14,16] (F-ARIMA) and M/G/∞ [6].

Unlike the FGN process, whose correlation structure, determined by a single parameter, is too rigid, F-ARIMA processes are very flexible to capture both short- and long-range correlations using few parameters. In fact, these processes have been widely used for modeling traffic sources [12,4,1]. Nevertheless, only off-line methods for synthesizing F-ARIMA traces are efficient enough to be of practical use.

K. Al-Begain, D. Fiems, and G. Horváth (Eds.): ASMTA 2009, LNCS 5513, pp. 131–143, 2009.

The M/G/∞ process is a stationary version of the occupancy process of an M/G/∞ queueing model. In addition to its theoretical simplicity, it can be used to model different traffic sources, because it is flexible enough to exhibit both Short-Range Dependence (SRD) and Long-Range Dependence (LRD). Moreover, queueing analytical studies are sometimes feasible [10,35,30,27], but when they are not, it has important advantages in simulation studies [20,29], such as the possibility of on-line generation and the lower computational cost (exact methods for the generation of a trace of length n require only $O(n)$ computations).

In this paper, we propose an efficient and on-line generator of M/G/∞ processes for matching the correlation structure of F-ARIMA processes.

The remainder of the paper is organized as follows. In Section 2 we review the main concepts related to SRD, LRD and self-similarity. In Section 3 we describe F-ARIMA processes and the most important methods for synthesizing traces that have been proposed. The M/G/∞ process is described in Section 4. We also remind briefly the method that we have presented in [31,32] to improve the efficiency of the generator when the distribution of the service time of the M/G/∞ system has subexponential decay. In Section 5 we explain the main concepts related to the Whittle estimator that we are going to use in order to compare FGN and F-ARIMA processes for VBR video traffic modeling purposes. In Section 6 we present the M/G/∞-based generator of the correlation structure of F-ARIMA processes, and we evaluate its efficiency and the quality of the samples, and in Section 7 we apply it to the modeling of the correlation structure of VBR video traffic, and we use a method based on the prediction error of the Whittle estimator to choose the best among several processes. Finally, in Section 8 we summarize the conclusions.

2 SRD, LRD and Self-similarity

It is said that a process exhibits SRD when its autocorrelation function is summable, i.e., $\sum_{k=1}^{\infty} \mathrm{r}_k < \infty$, like in those processes whose autocorrelation function decays exponentially:

$$\exists \alpha \in (0,1) \,\Big|\, \lim_{k\to\infty} \frac{\mathrm{r}_k}{\alpha^k} = \mathsf{c}_\mathrm{r} \in (0,\infty) \,.$$

Its spectral density is bounded at the origin.

Conversely, it is said that a process exhibits LRD [7] when its autocorrelation function is not summable, i.e., $\sum_{k=1}^{\infty} \mathrm{r}_k = \infty$, like in those processes whose autocorrelation function decays hyperbolically:

$$\exists \beta \in (0,1) \,\Big|\, \lim_{k\to\infty} \frac{\mathrm{r}_k}{k^{-\beta}} = \mathsf{c}_\mathrm{r} \in (0,\infty) \,. \quad (1)$$

Its spectral density has a singularity at the origin.

Let $X = \{X_n; n = 1, 2, \dots\}$ be a stationary stochastic process with finite variance and let $X^{(m)}$ be the corresponding aggregated process, with aggregation