Expert Systems With Applications 62 (2016) 373–384

Contents lists available at ScienceDirect

Expert Systems With Applications

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

A multifractal wavelet model for the generation of long-range dependency traffic traces with adjustable parameters

Miguel Tuberquia-David, Fernando Vela-Vargas, Hans López-Chávez, Cesar Hernández*

Faculty of Engineering and Technology, Universidad Distrital Francisco José de Caldas, Bogotá, D.C., Colombia

ARTICLE INFO

Article history:
Received 20 January 2016
Revised 8 April 2016
Accepted 3 May 2016
Available online 20 May 2016

Keywords:
Multifractal spectrum
Hurst parameter
Multifractal wavelet model
Binomial cascade
Network traffic

ABSTRACT

The available multifractal traffic finite-length time series to implement performance test of the management, control and admission algorithms, and level of service about M/M/1 models for WAN/LAN communication systems are very few and their recollection through current mechanisms is very slow due to the amount of data that must be obtained. Hence, it is necessary to develop a tool which synthesizes traces with multifractal features and allows the stochastic parameters configuration as its average, Hurst parameter and, multifractal spectrum width. This article describes the development of a proposed algorithm to generate multifractal traffic finite-length time series with a Hurst parameter and the multifractal spectrum width, sampling and adjustable, called MultiFractal Hurst Spectrum Width (MFHSW). The MFHSW algorithm is based on the MultiFractal Hurst model (MFH) and on the Multifractal Wavelet Model (MWM), to construct the time series through a binomial multiplicative cascade. The main contribution of the MFHSW algorithm is to allow adjusting both the Hurst parameter and the multifractal spectrum width, the aforementioned is achieved by appropriately modifying the beta distributions that conform the binomial cascade. Consequently, the impact developed by the algorithm to the trace generation with multifractal features will be the improvement in the simulation and data network modeling.

The MFHW algorithm behaves as an expert system when inferring to distribution of the beta coefficients present in the scales that make part of the binomial cascade starting from the stochastic parameters configured by the user, and obtaining the corresponding time series through an inference engine. To validate the algorithm effectiveness, a trace with the Hurst parameter sampling and the multifractal spectrum width similar to the presented in a network traffic time series are synthetized. The MFHSW happens to be a promising tool for the modeling of time series applicable to diverse fields as the traffic engineering, finances, biomedical signals, among other real traces with multifractal features.

1. Introduction

The communication networks have been studied for more than two decades. It has been shown that the inter arrival times of the users' demand arrivals and the demand itself on the network presents a correlation that persists through diverse time scales. It has been shown that such events can be properly categorized by using self-similar models (Alzate, 2001; Riedi & Vehel, 1997; Taqqu, Teverovsky, & Willinger, 1997). Outstanding studies as the ones conducted in the Bellcore laboratories in the 80′s have set a standard for the modeling and study of the traffic in communication networks with the purpose of predicting, controlling and generating a better service (Leland & Wilson, 1991). It has been shown that the use of the discrete wavelet transform as a synthesis and estimation tool for the fractal traffic analysis has been computationally effective (Abry, Flandrin, Taqqu, & Veitch, 2000; Alzate Monroy, 2002). Additionally, it proved that the time series of input packets of the users' demand in a network presents a behavior of stochastic self-similarity, main characteristic of multifractal signals (Chen, Cai, & Li, 1997; Contreras, Ospina, & Alzate, 2006; Leland, Taqqu, Willinger, & Wilson, 1994). The multifractal nature of network traffic was equally validated in the different time scales in the network (Riedi & Vehel, 1997; Wang & Qiu, 2005; Yu, Song, Fu, & Song, 2013) just like in other areas such as: financial time series (Resta, 2004; Thompson & Wilson, 2014), climate modeling (Das & Ghosh, 2015), biomedical signals (Lopes & Betrouni, 2009; Yu, Qi, & Introduction, 2012) among others.

However, such modeling must capture diverse characteristics of the multifractal behavior such as the long-range dependence (LRD), the variability in the scales and the fluctuations between small

* Corresponding author.
E-mail addresses: lmtuberquiad@correo.udistrital.edu.co (M. Tuberquia-David), fjvelav@correo.udistrital.edu.co (F. Vela-Vargas), hilopezc@udistrital.edu.co (H. López-Chávez), cahernandezs@udistrital.edu.co (C. Hernández).

http://dx.doi.org/10.1016/j.eswa.2016.05.010
0957-4174/

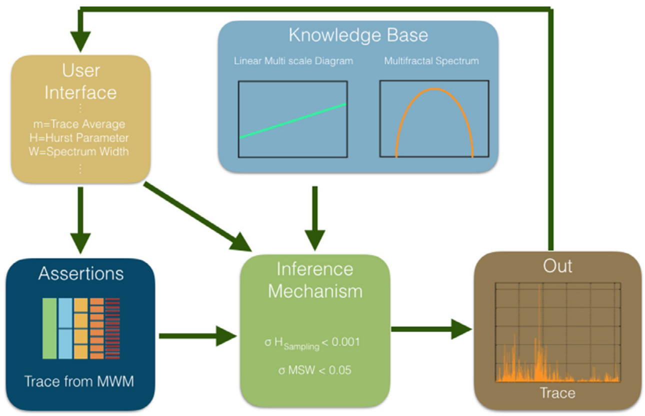

Fig. 1. MFHSW algorithm from the point of view of a system expert.

and large magnitudes (burstiness) (Ge, Fan, Zhu, Deng, & Wang, 2014; Ihlen, 2012), just as the data generation strictly positive. This last characteristic causes this task not to be modeled just with a process of Fractional Gaussian Noise (fGn) (Fan & Li, 2015), because this process normally generates distributed data with an average equal to zero (although it models appropriately the power spectrum and the LRD). In the search of an integral modeling, the Wavelet Multifractal Model (MWM) (Riedi, Crouse, Ribeiro, & Baraniuk, 1999) proposed in 1999, combines the multifractal features and the wavelet transform for the modeling of this type of traces. MWM involves the higher order statistic due to the multiplicative construction and it has given rise to the development of other models as the MultiFractal Hurst algorithm (MFH) (Contreras et al., 2006); The MFH algorithm is the basis of the current development, although it permits the configuration of the average and the Hurst parameter sampling, it does not adjust the multifractal spectrum width to a determined size by the user, which is the main difference in the current paper.

Accordingly, this paper describes the proposal of the MultiFractal Hurst Spectrum Width (MFHSW) model which allows adjusting the multifractal spectrum width (MSW) of a synthetic time series, also the Hurst parameter. The introduction of the multifractal spectrum to the MFH algorithm enables a more precise analysis of the time series, since it uses all the moments of q order. The proposed algorithm adjusts the Hurst parameter sampling and reallocates the MSW adjusting the beta distributions that conform the binomial cascade. Such binomial cascade is supported by MWM and at the moment of modifying its multiplicative coefficients it can adjust the MSW giving rise to the MFHSW algorithm. As input packets for the algorithm, the average of the trace is entered, the expected Hurst parameter, the MSW, the scales j1 and j2 that represent the octaves to be analyzed to calculate the Log-scale Diagram (LD), that allows determining the estimated H sampling and finally, the type of Wavelet to use.

Traditionally, the multifractal analysis is based on the trace parameterization. By contrast, the approach of this paper is the multifractal synthesis, this is, to generate multifractal traces that offer some expected characteristics by the users. In such a way, that is possible to specify the Hurst parameter, as the MSW, to obtain the corresponding trace.

The MFHSW algorithm is associated with an expert system (Miller, 1986) as is described in Fig. 1. Initially, the user interface captures the Hurst and the MSW parameters required by the user. The basis of rules is implemented through the LD and the multifractal spectrum (MS), because they are in charge of evaluating the Hurst and MSW parameters. The function of the inference engine is determining if the generated trace meets the predetermined Hurst and MSW parameter, and that their standard deviations are lower than 0.005 and 0.01, respectively. Finally, the working memory is associated with the generated trace without the verification of the basis of the rules and the inference engine.

The rest of the article is structured as follows. In Section 2 the most relevant works that approach the synthesize traces problem with multifractal characteristics are show. In Section 3 the mathematical resources for the estimation of the Hurst parameter sampling and the MSW are described; the formalism of the discrete wavelet is presented to calculate the detail coefficients and thus calculate LD. In Section 4 is displayed how the scale coefficients proposed by MWM were modified with the purpose of establishing a relationship between the Hurst parameter and MSW. In Section 5 the MFHSW algorithm is developed and validated through the comparison between a trace generated by the proposed algorithm and one as reference. Finally, in Section 6 the conclusions are drawn.

2. Related work

The first steps in generating traffic communications networks were presented at Intel Corporation by Kant (1999). Under his direction, the generation of request times between consecutive arrivals in a Web server was investigated; this generation looked for asymptotically self-similar processes having multifractal characteristics in small time scales. To find a solution, a Markov process M/G/oo was used to allow the introduction of multifractal behaviors in small/medium scales of time without affecting asymptotic self-similarity in the trace. Using this information, Kant modified the M/G/oo process to scale in time and queue behavior in the aggregated traffic and studied irregularities at small scales of time due to the construction of the cascade (Kant, 1999). Although both the queue and the scale behavior traditionally depended on the second order of the cascade, Kant showed that the long timescales of mass distribution had greater influence compared to the number of stages in the cascade. Based on this, he established a cascade construction method to properly select the scale and queue properties using one or two stages of the mass redistribution derived from the real networks dynamics (Kant, 1999).

Later in 2004, at the IBM research laboratory in Tokyo, Shimizu (2004) proposed a method for generating a network traffic pattern that could be used as a stress test on a web page. In the first step of the method, Shimizu generated a time series of positive root using a wavelet α-stable distribution to represent network traffic characteristics with both multifractibility and heavy queue. In the second step, the time series is restructured by Volterra filtering the high and low frequency components to finally re-synthesize the extracted components and obtain a designated traffic pattern by changing the components intensities. Shimizu simulated this method to generate traffic patterns containing long range dependence and an appropriate distribution of short-term (Shimizu, 2004).

An updated model was presented in 2010 by Zhao and Zhang (2010) at China National Digital Switching System Center department of computing and science, using the latency of the network frames, the Hurst exponent, and the types of network traffic. To predict their relationship, Zhao not only used the multifractal wavelet model based on multi–patterns, but also modified the Hurst parameter to generate multi-patterns of network traffic that include stochastic traffic and self-similar traffic at various distributions. All of this under a precise control of the frames latency to ensure the desired characteristics for the generated trace (Zhao & Zhang, 2010).

In 2012, Chávez and Monroy (2012) created an algorithm that generated traces of given length with configurable average sampling rate as well as Hurst parameter using the Multifractal