International Journal of Advances in Computer Science & Its Applications – IJCSIA
Volume 4: Issue 3 [ISSN 2250-3765]
Publication Date : 30 September,2014

Estimation of Hurst Exponent and Filtering Gaussian Effect on Fractional Brownian Motion

Danladi Ali and V.V. Gnatushenko

***Abstract*—In this work, four different sets of data traffic have been generated from fractional brownian motion (fBm) to estimate true values of Hurst exponent boundaries in order to determine the degree of self-similarity in terms of long range dependence (LRD). One-dimensional multilevel wavelet decomposition and filtering algorithm is applied to filter fractional Gaussian noise (fGn) in the fBm generated. Autocorrelation function (ACF) and fast Fourier transform (FFT) energy spectrum is used to validate the result of the filtering effect. The result of the filtering process revealed that fGn in the fBm is de-noised successfully as the coefficient of ACF grow above zero and energy rate in the FFT- spectrum increases tremendously**

***Keywords*—fBm, ACF, FFT, LRD Hurst exponent and self-similarity.**

I. Introduction

The introduction of the audio and video streaming, in addition to the usual web browsing have made the traffic structure of the telecommunications network complex, up to date the demand for internet service is growing in a geometric progression while the internet traffic exhibit fractal property; mathematically become self-similar and long range dependence (LRD)[8]. Presence of self-similarity in a network is associated with amplified queuing delay, packet retransmission, an overflow probability, eventually may involve packet drop. Self-similarity degrades the overall performance of the network by affecting its quality of service and capacity utilization. Quite number of models had been developed to address self-similarity issues on internet network since its discovery, but still situation prevails due to the huge or the varying nature of the input characteristics [10]. Therefore, when we want to study the self- similar nature of traffic of all kinds on internet network usually question of constructing a model of input characteristics (volume of the traffic) arise [4][11].

1. Danladi Ali. Department of Information and Technology, National Metallurgical Academy Ukraine

2. V. V. Gnatushenko. Department of Information and Technologies, National Metallurgical Academy, Ukraine.

However, to design a suitable model that can capture a picture of network behavior and to develop fast algorithms of free flow of information across a network from source to destination, self-similarity problem must be put into consideration if not the network may behave in an unexpected manner. This is a critical issue, because the fundamental aim of network monitoring and design is to deliver outstanding quality of service to the end user with little or no interference. This is an enormous task. Many researches in this field show that, internet traffic are fractal in nature; self-Similar and LRD been a packet switching or a teletraffic network [1][2][5][6][9][13][14].

In this work, we propose to investigate self-similarity boundaries in the data generated from fBm by estimating the true values of Hurst exponent and filter the content of the fGn in the fBm using a wavelet method, which is expected to be the cause of the self-similar effect. In order to estimate the true values of the Hurst exponent we used three (3) estimators; discrete second order derivative (DSOD), wavelet version of discrete second order derivative (WDSOD) and regression wavelet version (RWV). In what followed we generate data with different realization lengths from fBm, using a specific value of Hurst exponent, (H=0.8). The first realization length is generated to 100 while the subsequent lengths are generated in multiple of 10 which means that the highest length is 100,000 the idea of varying the realization lengths is to assume LRD. As we know, fBm is a Gaussian self-similar process with stationary increments. Haven estimate the true values of Hurst exponent , that tells us about the degree of self-similarity, as well as the fGn, next we will apply one-dimensional multilevel wavelet filtering to filter the effect of the fGn present in the fBm generated, when applying the one-dimensional multilevel wavelet filtering we will consider autocorrelation function (ACF), and fast Fourier transform (FFT) energy spectrum to validate our filtering results. Hurst exponent is a measure of long-term memory of time series, it relates to autocorrelation of the time series [7]. One obvious observation with ACF is that its coefficient decay slowly to zero when data in a time series is an LRD and self-similar while in FFT- energy spectrum, the energy against frequency in terms of rate will be reduces tremendously.

II. Data Traffic generation and Method of Data Analysis

As earlier mention the data generated from the fBm are then classified into four different sets assigning realization length to A =100, B=1,000, C=10,000 and D=100,000. We then estimate Hurst exponent of the data sets generated using a wavelet method as shown in Figure 1. Having known the true values of the Hurst exponent and ascertain the degree of self-similarity in each case we then applied one-dimensional multilevel wavelet to decompose and filter the fractional Gaussian noise in the generated data while observing the ACF and FFT energy spectrum.

A. *Hurst Exponent Estimation with Wavelet*

The derivation of expression (4) is adopted from [12]. Let's consider the sum of data traffic observed as given in expression (1)

$$X = x_1 + x_2 + \dots\dots + x_n \tag{1}$$

The derivation below is adopted from [12]. Let us consider X as a stationary process of the second order. The wavelet coefficient $d_x(j,k)$ satisfies the relation below

$$M\left|d_x(j,k)^2\right| = \int F(\lambda) 2^j \left|\psi(2^j\lambda)\right| d\lambda \tag{2}$$

Where $F(\lambda)$ and $\psi(\lambda)$ are the power spectrum for X and the Fourier transform for the wavelet function $\psi_0(.)$ respectively.

$$M\left|d_x(j,k)^2\right| \approx 2^{j(2H-1)} C_f C(H,\psi_0)$$

Where $C(H,\psi_0) = \int |\lambda|^{-(2H-2)} |\psi(\lambda)|^2 \, d\lambda$ is a constant that depends on H and ψ_0

If X is equal to m, then the available number of the wavelet coefficient in the octave is $m_j, m_j = 2^{1} n$ as a result

$$\mu_j = M\left[d_x(j,k)^2\right] = \frac{1}{m_j} \sum_{k=1}^{m_j} \left|d_x(j,k)\right|^2 \tag{3}$$

Normally μ_j is unbiased and it is the constant estimation for $M\left|d_x(j,k)^2\right|$, the expansion of (3) represents a possible way to estimate H of the LRD process.

$$\log_2 \mu_j = \log_2 \left(\frac{1}{m_j} \sum_{1}^{m_j} \left|d_x(j,k)^2\right| \right) \approx 2H - 1 + C \tag{4}$$

Where, $C = \log_2, C_w = \log_2 \left[C_f C(H,\psi_0)\right]$ is a constant, then the slope of the function of plot μ_j on j have a linear slope whereby the series can be analyzed while the expansion operator tells us more about the time scale or the frequency range. This situation is always better understood in a fast Fourier transforms.

B. *Autocorrelation Function*

The general form of evaluating ACF coefficient of the data observed may be evaluated by the expression (5)

$$r(k) = \frac{1}{N} \sum \frac{(x_1 - \overline{X})(x_{z+k} - \overline{X})}{\sigma^2(x)} \tag{5}$$

Where $\overline{X}$, the mean of the observed data, σ is the variance of x while kϵ(0,1,……,n). ACF of x may be related to self - similarity, when we consider the expression (6);

$$r(k) \sim k^{-\beta} L(k), \qquad k \to \infty \tag{6}$$

Where L is the function that changes slowly to infinity for all x is greater than zero, then the process may be strictly self-similar with β=2H-1, where, 0<β<1, this implies that as β approaches 1 the degree of the self-similarity increases while ACF coefficient decay to zero.

C. *Fast Fourier Transform Energy Spectrum*

The most common method to study or show the amount of fractional Gaussian noise in the signal is achieve through FFT spectrum analysis. In this work, we are going to examine the FFT energy spectrum of the fBm signal, with respect to the growth of the fBm signal in the spectrum as we filter the Gaussian effect. Our expectation is that, as the filtering level increases the signal spectrum will also increases, the evidence may be confirmed in the later part of this work. The discrete Fourier transforms (DFT); usually determined by decomposing the signal into different frequency components; this process is slow and hard to investigate alternatively, use of FFT technique overshadow the DFT process by enhancing the process and easy to notice significant changes in terms of spectrum size and speed. Let's consider (1), the DFT may be computed using expression (7), below.

$$X_k = \sum_{n=0}^{N=1} x_n e^{-i2\pi k \frac{n}{N}} \tag{7}$$

k ϵ 0, . . . , N-1. In DFT, evaluation of X_k may requires N^2 operation, while FFT, requires mathematical operation $N\log_2 N$ to improve the size and the speed of the DTF spectrum. This means that as filtering process increases the FFT speed increases by the ratio of $N^2/N\log_2 N$ which may be writing as $N/\log_2 N$, this is approximately 10 times the speed of DFT spectrum. See [16][17] for further details. Where, $e^{-\frac{2\pi i}{N}}$ denotes N^{th} primitive root of unity and N represents vector sum of the sampled sets.

III. Result and Discussion

TABLE I HURST EXPONENT BOUNDARIES ESTIMATED

	DSOD	WVDSOD	RWV
A	0.65 < H < 0.90	0.65 < H < 0.95	0.30 < H < 1.00
B	0.62 < H < 0.80	0.62 < H < 1.00	0.42 < H < 1.00
C	0.75 < H < 0.83	0.75 < H < 0.84	0.62 < H < 0.84
D	0.76 < H < 0.81	0.78 < H < 0.82	0.68 < H < 0.79

Actually, there are many methods of estimating or analyzing Hurst exponent which may include aggregate variance plot, R/S - method, periodogram, whittle, abry-vietch, variance of residuals, absolute moments estimators and so on but, our approach in this work is by using wavelet because, this method seems to have promising results and gives better confidence interval than the other methods and it has filtering ability, most of these methods have their relevant applications in different networks; most times they tend to be bias in solving general issues unlike wavelet method which always shows better result in almost all cases. Self - similarity can be evaluated by Hurst exponent values if H range between ½<H<1 [12] in any telecommunication network, then we can conclude that traffic in the network is self-similar. Figure 1, shows the true values of the Hurst exponent for all the three estimators for a set of generated from fBm data. The true values of the other sets of fBm can be obtain in the same manner as shown in Figure 1. "In some cases, it is not necessary that we have to determine the specific or exact value of Hurst exponent for a given data. Otherwise, we could decide to which boundary Hurst exponent belong to" [15] as shown in the Figure 1, below;

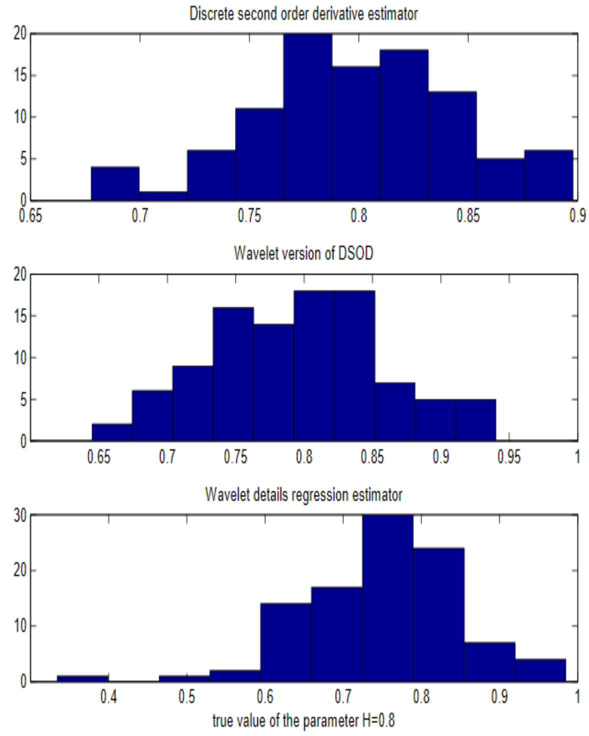


Figure 1. True values of the Hurst exponent for set A

As shown Figure 1, the discrete second order derivative and the wavelet version of the second order derivative give almost similar results with very little dispersion while the wavelet regression seems to be slightly biased and has higher dispersion. In all the estimation, the true value of the Hurst exponent is presented on X-axis while the frequency of their occurrence is represented on Y - axis. It was observed that in all the estimators. Hurst exponent increases with the increase in realization length (LRD) and does not depend much on the input characteristics. We noticed that the third method in Figure 1, slightly falls into anti-persistent behavior with the Hurst exponent value less than 0.5, which may be attributed to the short range dependence behaviors of input characteristics. Now that we have established the fact that, in almost all cases, there is the presence of the fGn attributing to the strong self – similarity in all. What next; we applied multilevel one - dimensional wavelet filtering process to filter the fGn significantly or eliminate completely the effect of self - similarity while closely observing ACF and FFT

- Energy spectrum. In the filtering process, we used db10 wavelet type to decompose the fBm. After successful de-noising, we noticed that for A level 2 decomposition is adequate while for B, C, D level 3, 4, 5 respectively which means that the decomposition level is also is directly related to the LRD. We also noticed that the decomposition level increased by one level when the realization length increases by x10. During the decomposition and the reconstruction process, we observed the ACF and FFT - Energy spectrum as shown in Figure 2 to 5 for each of set fBm generated.

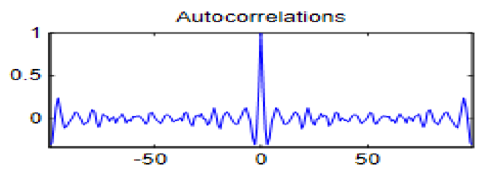
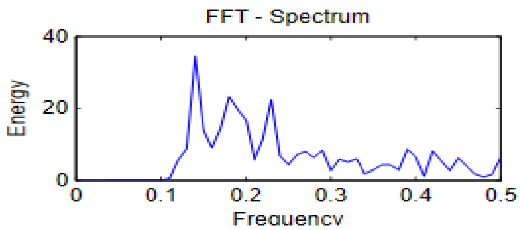


Figure 2a. ACF of A after de-noising



igure 2b. FFT-Energy spectrum of A after de-noising

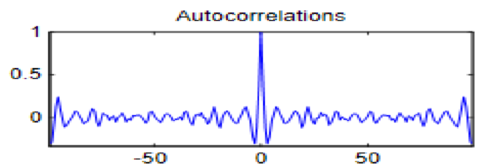


Figure 3a. ACF of B after de-noising

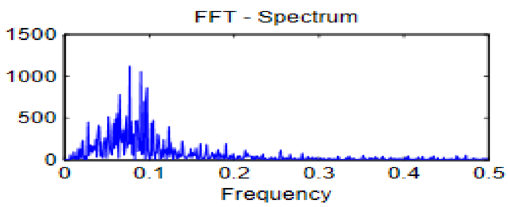


Figure 3b. FFT-Energy Spectrum of B after de-noising.

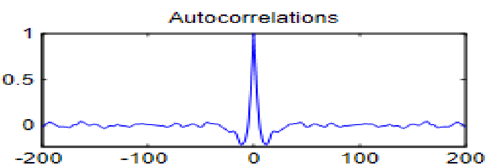


Figure 4a. ACF of C after de-noising

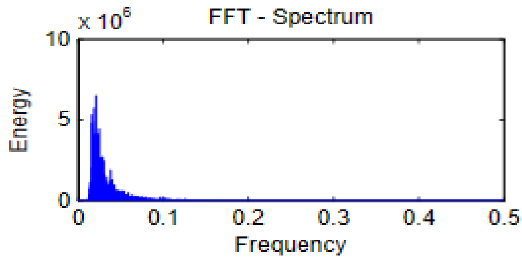


Figure 4b. FFT-Energy Spectrum of C after de-noising.

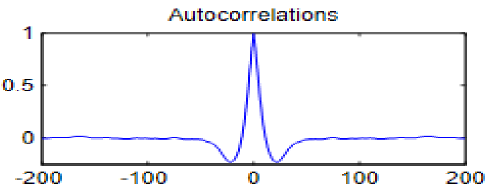


Figure 5a. ACF of D after de-noising

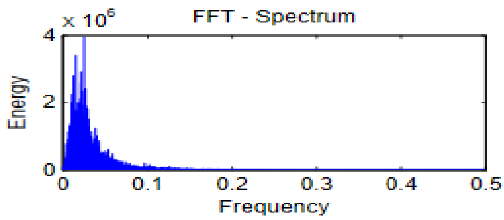


Figure 5b. FFT-Energy Spectrum of D after de-noising

From Figure 2 to 5, the ACFs for all the sets of the fBm generated after decomposition and reconstruction; clearly show that each set had been de-noised, since the property of ACF with LRD or self – similarity do not comply as its coefficient grow above zero. In FFT - Energy spectrum for all sets show tremendous data rate with high energy in all cases.

IV. Conclusion

In this study, we investigated the self - similar nature of fBm by estimating the true values of Hurst exponent in four different sets of fBm generated using wavelet method. The work revealed that the degree of self – similarity depends on LRD. It was also observed the decomposition level of each realization length depend on the LRD and wavelet is a good tool for filtering Gaussian effect. We hope this work has provided some basic information on how to estimate Hurst exponent by their boundaries using wavelet and its filtering effect. Further research work should be conducted to provide better robust way of solving self-similarity problems.

References

[1] A. Adas, "Traffic, models in broadband network" IEEE Communication Magazine 2006.

[2] P. Abry and D. Veitch, "Wavelet analysis of long - range dependence traffiC"IEEE Transaction on information theory, 1998.

[3] A. F. Behrouz, "Data communication and Computer network" Mc-Grew Hill Higher Education, 2007

[4] V. V. Gnatushenko, Y. S. Proedak and A. Danladi, "Study of effect of pareto distribution structure form on information flow characteristics by applying simulation model" Journal system and technology 1(84), pp. 140-146, 2013

[5] S. Chaoming, H. Shlomo and A. M. Hernan , "Self – Similarity of complex network"'Cond-mat, dis-nn, 2005

[6] E. M. Crovella, A. Bestavros, "Self – similarity in worldwide web traffic" In IEEE transactions on networking, vol. 5(6), pp. 835-846, December, 1997.

[7] T. Higuchin, " Approach to an irregular time series on the basis of the fractal theory: Physic vol. 30 pp. 277-283, 1984.

[8] E. H. John, "Fractal and self - similarity" Journal of Mathematics University of Indiana vol. 30, 713-747, 1981.

[9] A. Nagurney, Q. Qiang , "A measure for congested network" Epl Journal, 2007.

[10] L. Nikolai, and T. Boris, "An analysis of an ATM buffer with self- similarity (fractal) input traffic" 14th Annual IEEE conference. Psu.edu, 1995.

[11] A. Nils, M. Peter, and G. Michael, "Characteristics and Modeling of voice traffic in first responders network" 32nd IEEE conference on local computer network, 2007.

[12] I. S. Oleg, M. S. Sergey, and V. O Andrey, "Self – similar process in Telecommunications" John Wiley and Sons, Ltd, 2007.

[13] D. Beyers, " Packet radio network research, development and application " Proceeding SHAPE Conference on packet radio, Amsterdam, 2007.

[14] W. Walter, S. T. Murad , S. Robert, and V. W. Daniel, "Self – Similarity through high - variability; Statistical analysis of Ethernet LAN traffic at the source level" IEEE transactions 5(1):71-86,1997. http//link.springer.com/chapter/10.846428.048-6.14

[15] J.Jelinskis, Heuristic algorithm for robust approximate Hurst parameter estimation with Wavelet analysis and neural networks, Telecommunication and electronics. ISSN 1407-8880. 2008

[16] B. Peters. "FFT making technology fly" SIAM News, 26(3) May,1993.http//www.siam.org/siamnews/mtc/mtc593.htm

[17] R. Peters, "Fourier transform construction by vector graphics", Am Journal of Physics 60, 439-441, 1992.

Danladi Ali was born in Garkida, Gombi, Adamawa State, Nigeria. He received his B. Eng. (Electrical and Electronic) 2000 at University of Maiduguri Borno State, Nigeria and M.Eng.(Electronic and Telecommunication) 2006 at University of Benin, Nigeria. He is a Ph. D student, Department of Information Technology and System at National Metallurgical Academy of Ukraine. He is a Lecturer at Adamawa State University, Mubi. Nigeria. His area of interest is; computer networking and communications with special interest in AD-HOC Wireless Network, Self – Similarity and Fractality of traffic in computer networks.

V. V. Gnatushenko. Hold Ph. D and a Lecturer at the Department of Information Technology and System, National Metallurgical Academy of Ukraine.