ISSN 1407-8880
TELECOMMUNICATIONS AND ELECTRONICS
TELEKOMUNIKĀCIJAS UN ELEKTRONIKA
2008-8880

Heuristic algorithm for Robust Approximate Hurst Parameter Estimation with Wavelet Analysis and Neural Networks

J.Jelinskis [1]
[1]Riga Technical University, Faculty of Electronics and Telecommunications
Riga, Āzenes str. 12, E-Mail: jans.jelinskis@etf.rtu.lv

Abstract – Fast and robust Hurst parameter estimation of traffic data traces tops the bill of nowadays problems of the field of traffic engineering. Almost every existent approach fits up the goal of as far as possible precise H parameter estimation; however this option is not as necessary as approximate estimation of boundaries of H parameter if traffic demonstrates long range dependence. Often this is the satisfactory condition for defining adequate traffic engineering operations. In this paper we verify a possibility of heuristic H parameter estimation algorithm which is based on wavelet transform of fractional Brownian motion synthesized data traces, and forthcoming operating with neural network learning capabilities. In this paper algorithm is described. Experimental data are depicted and future research subjects are pointed.

Keywords: traffic Engineering; Hurst parameter estimation; wavelet analysis; neural networks.

I. INTRODUCTION.

In 1994, Leland and Willinger et al. [1] reported that network traffic exhibited selfsimilarity – a fractal concept in data analysis. Fractal traffic analysis and modeling have been ever since a very popular research topic in the field of network and traffic engineering.

Today, the Internet is growing exponentially, with traffic statistics that mathematically exhibit fractal characteristics: self-similarity and long-range dependence. With these properties, data traffic shows high peak-to-average bandwidth ratios and causes data networks be inefficient [2].

For example, the self-similar traffic degrades network performance [3] such as buffer overflow probability. As a result, the network utilization decreases remarkably due to retransmitting the overflowed packets.

These problems make it difficult to predict, quantify, and control data traffic, in contrast to the traditional Poisson distributed traffic in classical telephone networks.

Especially in the recent years, many new bandwidth consuming applications with multimedia content have become increasingly popular, which causes network traffic to be even more fractal in nature. Multimedia content requires network providers to improve network efficiency and to improve the quality of services to customers. How to make networks more efficient while still maintaining a good quality of service becomes the key issue in the network-engineering domain.

Self-similar traffic modeling is the other topical problem of representing our understanding of dynamic demands by stochastic processes. Accurate traffic models are necessary for service providers to maintain quality of service properly. Many traffic models have been developed based on traffic measurement data, so there is a problem of finding an appropriate way of fast and robust traffic model identification methods, which allows to determine, with the certain accurateness all the vital characteristics of measured data traces.

Below we briefly discuss the definition of self-similarity and long-range dependence and their connection. This issue is discussed in very detailed manner in [4].

Suppose $X(t)$ is a second order stationary stochastic process and fx and γx its spectrum and autocorrelation function respectively. The process $X(t)$ is said to be Long Range Dependent (LRD) if either, for some constant c_f,

$$f_x(\upsilon) \sim c_f|\upsilon|^{-\alpha} \text{ as } |\upsilon| \to 0, \text{ where } \alpha \in (0,1), \quad (1)$$

Or if, for a different constant c_γ,

$$\gamma x(k) \sim c_\gamma|k|^{-(1-\alpha)} \text{ as } |k| \to \infty, \text{ where } \alpha \in (0,1). \quad (2)$$

The process is called long-range dependant because $\gamma x(k)$ goes to zero so slow, as $k \to \infty$ that $\sum_k \gamma x(k) = \infty$. A process $Y(t)$ is said to be self-similar with so called self-similarity Hurst parameter H if and only if

$$c^{-H}Y(ct) = Y(t) \text{ for all } c > 0 \quad (3)$$

There is a close connection between LRD and self-similar process. [8] The increment of any finite variance self-similar process is long-range dependent, as long as $\frac{1}{2} < H < 1$, with H and α related through H=(α+1)/2.

One common self-similar process is the fractional Brownian motion (fBm) [5], whose increments define the so called fractional Gaussian noise (fGn) which is Long range dependent.
traces, widely used in a great variety of applications, especially in traffic modeling in broadband network
Estimation is done based on previously trained wavelet neural network learning capabilities.

The relevance of this approach derives from the ability of capturing the inner nature of the data set even if the values are much approximated.

II. PROBLEM STATEMENT AND SCOPE.

Because of the widely accepted long-range dependent self-similar properties of network traffic, Hurst parameter estimation provides a natural approach to studying such models. It seems to be logical that H parameter estimation leads to the reliable traffic control technique, which can be used for traffic engineering purposes.

Many approaches for estimating the Hurst parameter have been proposed [6, 7]. Among various approaches, the wavelet method has attracted the interest owing to its robustness to non-stationary and decorrelation property. The wavelet transform with their natural scale invariance and low computational cost is really suitable for analyzing of LRD process [8].

There are also several Hurst parameter estimators, such as R/S plot, variance-time plot, periodogram and wavelet-based.

Park et al. [7] thoroughly compared three different Hurst parameter estimators by using simulated, synthetic and real Internet traffic data sets. It reveals a number of important challenges which one faces when estimating the long-range dependence parameter in Internet data traffic traces. Stoev et al. [6] explored some of these challenges in more detail by using the wavelet spectrum method.

Our approach is based on predefined fBm trace wavelet analysis and forthcoming neural network training/operating with wavelet decomposition coefficients as a training patterns, thus our approach is bodily novel.

In many cases there is no need to know precisely what H parameter value corresponds to existing traffic trace. For most traffic engineering purposes, it is quite enough to decide to which boundaries of H parameter this specific measured trace belongs to be able to take essential actions based on foreknown rules.

In fact, the fractal behavior of an important class of self-similar traffic models is unambiguously described by a single parameter - *H* [9], which gives us reliable approach of determining of the belonging of the trace to the certain level of self-similarity.

In this research we assume that we are interested in H parameter values in within 0.5<H<1, and we partition this range into subsequent subclasses. In this case we introduce 5 classes which we name accordingly:

$$
\begin{aligned}
&A=\{0{,}5<H<0{,}6\};\\
&B=\{0{,}6<H<0{,}7\};\\
&C=\{0{,}7<H<0{,}8\}; \qquad (4)\\
&D=\{0{,}8<H<0{,}9\};\\
&E=\{0{,}9<H<1\};
\end{aligned}
$$

Based on this separation all the next methodology is being done, which is described in the next sections.

In this paper we verify a possibility of fast and robust approximate H parameter estimation approach based on Wavelet filtering potential and Neural Netowork learning capabilities. We analyze a wavelet based method for the estimation of the Hurst parameter boundaries, of synthetically generated self-similar

II. METHODOLOGY AND RESULTS.

During this research we generated a large amount of data using the algorithm for fractional Brownian motion traces with predefined H parameter traces [10]. We assume that our interest is raised by long range dependant processes, where an H parameter value varies $\frac{1}{2} < H < 1$.

As a next step we applied multilevel discrete one-dimensional wavelet decomposition at the 3 level using a specific Db10 wavelet. As a result vectors of decomposition coefficients were obtained.

Coefficient vectors were divided, depending on the separation accepted above (4) and the separated decomposition coefficient vectors were assigned the class IDs, which was intended to use as a training target data for wavelet neural network.

As a next step the backpropagation neural network (NN) was built with the predefined number of hidden neurons.

Trained NN is capable to operate with real coefficients of decomposed data traces with the determined operation error rate (Fig.1 and Fig. 5). We assume that data traces are normalized before wavelet analysis, which can have a notable impact on NN operating error, if analyzed data is out of boundaries of NN training data.

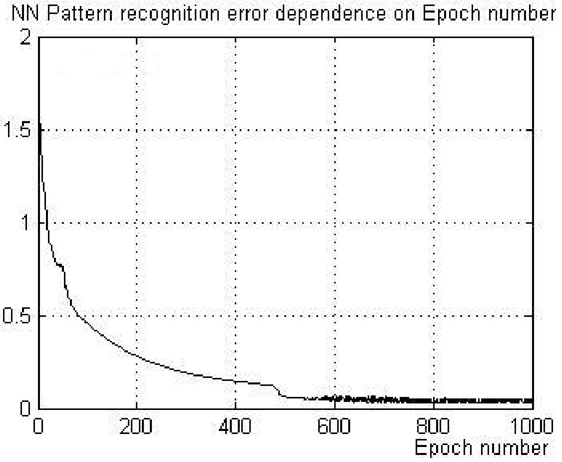

Fig. 1. Mean squared NN operating error dependence on epoch number. Training patterns decomposed form the originally generated fBm data trace. Number of hidden neurons – 20.

III. ROBUSTNESS OF THE APPROACH.

Our interest in wavelets began with the idea that the wavelet transform could be used to filter out noisy data in time series for operating with trained neural network. It is especially important when dealing with considerable outliers, because it could be the reason for big neural network operating error, when analyzed data is out of the training data boundaries.

The next important issue is a possibility to operate with compressed data, which contains considerably less information about the original data trace. This is crucial when data collecting is performed with ultrafast multi Gb/s data channels and comparatively infrequent sampling gives us much approximated scene of analyzed data (Fig. 3)

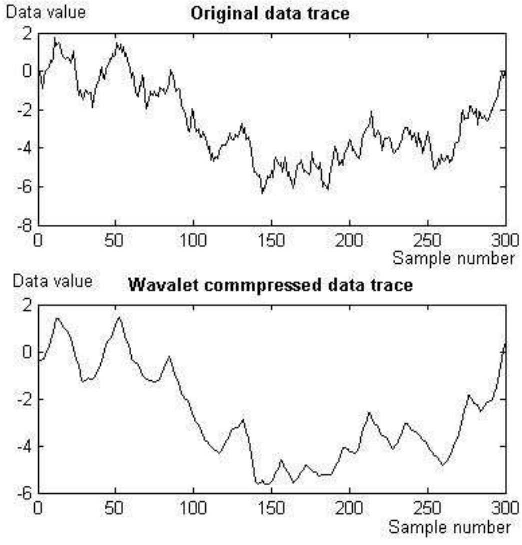

Fig. 2. Originaly generated fBm data trace with H=0.8, and wavelet compressed instance.

We have used three H estimators operating with original and compressed data traces.

The first one, due to Istas and Lang [11] is based on the discrete second-order derivative. The second one is a wavelet-based adaptation and has similar properties. The third one, proposed by Flandrin[12] estimates H using the slope of the log-log plot of the detail variance versus the level.

As it is seen from results (Fig.3 and Fig.4), when dealing with compressed data, the first two estimators shows completely incorrect estimation and the wavelet approach proposed by Flandrin[12] is also incorrect but preserves the prior variance in the outcome.

The author proposed algorithm operates with the compressed data with the similar operating error, but in this case we have to apply much more training to achieve the appropriate operating error level (Fig.5)

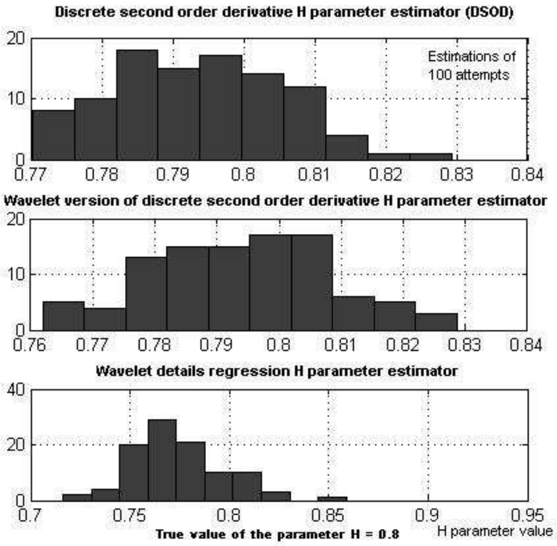

Fig. 3. Histogram of 100 approaches of H parameter estimation from original fBm trace using 3 different estimation techniques. The true H parameter value is 0.8.

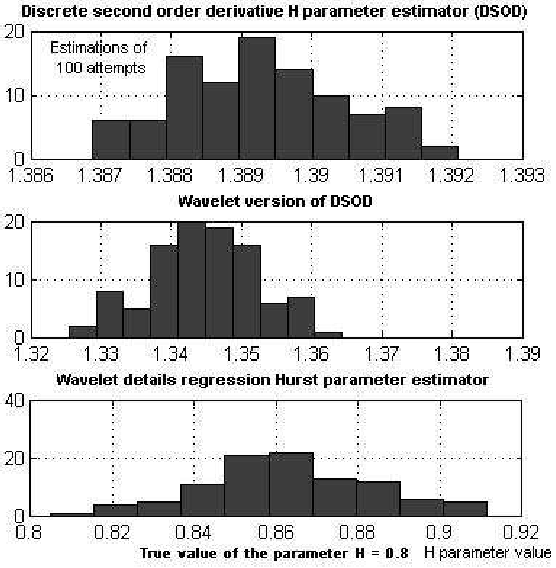

Fig. 4. Histogram of 100 approaches of H parameter estimation from Wavelet compressed fBm trace using 3 different estimation techniques. The true H parameter value is 0.8.

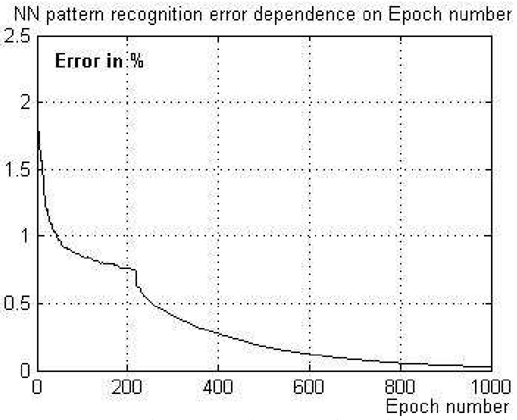

Fig. 5. Mean squared NN operating error dependence on epoch number. Training patterns decomposed form the compressed data. Number of hidden neurons – 20.

The idea behind noise filtering is to remove the noise while leaving the important detail. While this is fine as noise reduction tool, would it be so effective when meeting spacious outlier shift in data series? (Fig.6).

This question deals with concept of robustness of proposed H estimation algorithm and it is a subject of analyzed data trace size and correspondent outliers. Optimal ratio of outliers/data trace size is not a subject of this research; however results for synthesized outliers within fBm data traces show promising results. By our finding results even significant outliers do not make critical changes in wavelet decomposition scene (Fig.7), which results in correct operation of trained neural network.

To this issue there was not devoted plenty attention in this research and it is mentioned to be the subject of future research.

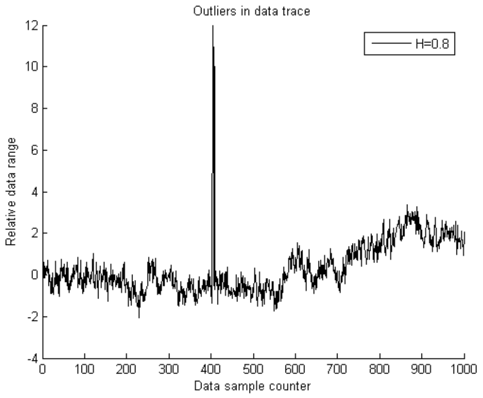

Fig. 6. Possible outlier shift in analyzed data series

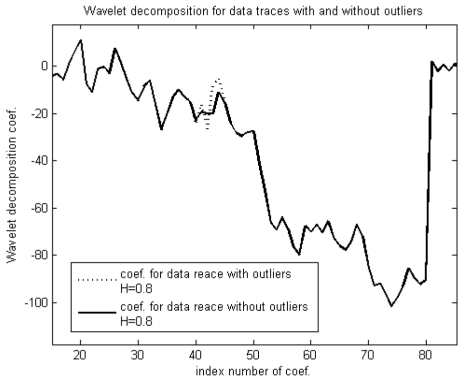

Fig. 7. Wavelet decomposition coefficients for fBm data trace with and without outliers

IV. CONCLUSIONS AND FUTURE RESEARCH.

In this paper we verify a possibility of heuristic robust and approximate wavelet analysis based algorithm for H parameter estimation. Fractional Brownian motion synthesized data traces serves as a template, for forthcoming wavelet decomposition and operating with neural network learning capabilities, where decomposed real data traces are analyzed with the means of trained neural network.

Robustness of this approach is based on wavelet analysis capability of filtering noise out of the data as well as considerable indifference for outliers in the analyzed data. Significantly compressed data traces are also suitable for operating with proposed algorithm as after wavelet decomposition they offer applicable training patterns for NN robust operating.

As a result we have testes proposed approach, using fBm synthesized data traces. Acquired results are based on assumed variables for this algorithm, in other words we have used only subjectively assumed values for wavelet selection for data trace decomposition, decomposition level and NN architecture. These options are modifiable, and search for optimal selection of above listed criteria is can have a considerable impact on a potency of proposed algorithm.

Alternate design of wavelet for data trace decomposition, as well as wavelet decomposition level and selection for NN architecture are not subject of this research and are mentioned as a ground for the future activities.

REFERENCES.

[1] **Will E. Leland, Walter Willinger, Murad S. Taqqu, Daniel V. Wilson**, On the Self-Similar Nature of Ethernet Traffic, ACM SIGCOMM Computer Communication Review, volume 25, issue 1(1995), p. 202-213.

[2] **Alexessander da Silva Couto Alves,** Internet Traffic Engineering, An Artificial Intelligence Approach, Departamento de Ciencia de Computadores Faculdade de Ciencias da Universidade do Porto (2004), PhD thesis.

[3] **Z.Sahinoglu and S.Tekinay,** On Multimedia Networks: Self-Similar Traffic and Network Performance, IEEE Commun. Mag., Vol.37 (1999), No.1, p.48-52.

[4] **Taqqu M,** Fractional Brownian motion and long-range dependence, Theory and Applications of Long-Range Dependance, 2003, p. 5 – 38.

[5] **AlShafee M., Elagha H,** On the Self-Similar Nature of ATM Network Traffic, 2007 Informing Science and Information Technology Conference Ljubljana, Slovenia, 2007, Volume 4, p. 641 – 649.

[6] **Stoev S., Taqqu's M.S., Park C., Marron J.S,** On the wavelet spectrum diagnostic for Hurst parameter estimation in the analysis of Internet traffic, Computer Networks, Volume 48, Issue 3, 2005, p. 423-445

[7] **Park J. Park C,** Robust estimation of the Hurst parameter and selection of an onset scaling //Under revision, 2007, Web available:http://www.maths.lancs.ac.uk/~parkj1/paper/robusth_Dec07.pdf

[8] **Guanghui He, Yuan Gao I., Jennifer H Hou, Kihong Park,** A Case for Exploiting Self-Similarity of Network Traffic in TCP Congestion Control, The International Journal of Computer and Telecommunications Networking Volume 45, Issue 6, 2004, p. 743 – 766.

[9] **Bin Shi.** Multiscale Statistical Analysis of Self-Similar Processes with Applications in Geophysics and Health Informatics, Georgia Institute of Technology,2005. PhD Thesis.

[10] **Dieker T.** Simulation of fractional Brownian motion**,** University of Twente, Department of Mathematical Sciences, The Netherlands, 2002. PhD Thesis.

[11]**. Istas, J.; G. Lang**, Quadratic variations and estimation of the local Hölder index of a Gaussian process, Ann. Inst. Poincaré, 33,1994, p. 407-436.

[12]**. Flandrin, P,** Wavelet analysis and synthesis of fractional Brownian motion, IEEE Trans.on Inf. Vol.38,1992 p. 910-917