Approximation of Internet Traffic using Robust Wavelet Neural Networks

J. Jeļinsiks, G. Lauks
Institute of Telecommunication, Faculty of Electronics and Telecommunication, Riga Technical University
Āzeres st. 12, LV-51368 Riga, Latvia, phone: +371 7089201; e-mails: lauks@rsf.rtu.lv, jans.jelinskis@etf.rtu.lv

Introduction

Telecommunications networks have migrated from circuit based telephony services to packet based broadband network services. Merging with computer networks, they are being integrated with non-real-time data services on classical Internet integrated multimedia services, including real time voice, video and services on the new generation Internet. Thus, the concepts and requirements of traffic engineering have also been changed significantly [1]. There is a need of handling with traffic, that is completely different nowadays and traffic measurements/generation problems have to have new approaches and methods.

One of the most significant findings of traffic measurement studies over the last decade has been the observed self-similarity in packet network traffic [2].

Self-similar processes are emerging as a powerful mathematical representation of a great variety of physical phenomena. Self-similarity has been discovered, analyzed and exploited in many frameworks, especially in the field of traffic modeling in broadband networks [3].

Recently there is revealed the relationship between the self similar property of network traffic and the heavy tail distribution of duration time of Internet connections, which makes self-similar processes a popular tool for modeling Internet traffic flows. Below we briefly discuss their definition and connection. This issue is discussed in very detailed manner in [7].

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

\[
f_X(k) \sim c_f |k|^{-\alpha} \quad \text{as} \quad |k| \to 0, \quad \alpha \in (0,1),
\]

(1)

Or if, for a different constant \( c_f \),

\[
\gamma_X(k) \sim c_f |k|^{1-\alpha} \quad \text{as} \quad |k| \to \infty, \quad \alpha \in (0,1).
\]

(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

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

(3)

There is 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 \( \alpha \) related through \( H=(\alpha+1)/2 \). One common self-similar process is the fractional Brownian motion (fBm) [9], whose increments define the so called fractional Gaussian noise (fGn) which is Long range dependent.

In this paper we verify a possibility of traffic approximation approach based on synthetically generated fractional Brownian motion traces with predefined \( H \) parameter. We analyze a wavelet based method for the estimation of the Hurst parameter boundaries, of synthetically generated self-similar traces, widely used in a great variety of applications, especially belonging to traffic modeling in broadband networks. Estimation is made 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 without introducing artificial and cumbersome models.

Traffic classification problem

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 classification tool, which can be used for traffic engineering purposes.

Many approaches for estimating the Hurst parameter have been proposed [5, 12, 13]. Among various approaches, the wavelet method has attracted the interest owing to its robustness to non-stationary and decorrelation property.
Park et al. [13] 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. [12] explored some of these challenges in more detail by using the wavelet spectrum method. While the wavelet method is reliable in practice and quite robust with respect to smooth polynomial trends in the data, it can mislead the practitioner. For example, a traffic trace with a number of deterministic shifts in the mean rate results in a steep wavelet spectrum which leads to overestimating the Hurst parameter. However we will not come back to this issue in this paper, because we assume this is not the ground of this research, as our approach is based on predefined fBm trace wavelet analysis and forthcoming neural network training/operating with wavelet decomposition coefficients, 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, when dealing with traffic classification problem it is quite enough to decide to which class of traffic 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 processes, fBm in our case, is unambiguously described by a single parameter - H [4], which gives us reliable tool for classification condition selection.

In this research we assume that traffic classification is defined by H parameter values in within 0.5<H<1, which can be considered as long range dependent process partitioning into initiated subclasses. In this case we introduce 5 classes which we name accordingly:

\[ A = \{0.5 < H < 0.6\} \]
\[ B = \{0.6 < H < 0.7\} \]
\[ C = \{0.7 < H < 0.8\} \]
\[ D = \{0.8 < H < 0.9\} \]
\[ E = \{0.9 < H < 1\} \]

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

**Methodology and results**

To this end we generated a large amount of data using the algorithm for fractional Brownian motion generation with predefined H parameter traces [11]. In this research 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 one-dimensional wavelet analysis using a specific wavelet. Db10 wavelet is chosen for the analysis here, and decomposition at 4 level was used. As a result vectors of decomposition coefficients were obtained (Fig. 1).

Coefficient vectors were clustered with Fuzzy c-means (FCM) clustering algorithm [14] and every trace was given an identifier – class ID, which was intended to use as a training target data for wavelet neural network. Obtained clusters are depicted below on Fig. 2.

![Fig. 1. Wavelet decomposition structure for five fBm data traces with predefined H parameter values](image1)

![Fig. 2. Fuzzy clustering result for fBm traces with predefined H parameter values within 0.5<H<1](image2)

**Robustness of the approach**

As a next step, to build a wavelet basis neural network (NN), the sigmoid function is replaced with the wavelet in NN [10]. The decomposed component sequences (Fig. 1) are employed to train the NN.

Trained NN is capable to operate with real coefficients of decomposed data traces with the determined operation error rate. 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.

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, especially it is 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 [6].

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. 3).

![Fig. 3. Possible outlier shift in analyzed data series](image)

This question deals with concept of robustness of proposed traffic classification 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. 4), which results in correct operation of trained neural network.

![Fig. 4. Wavelet decomposition coefficients for fBm data trace with and without outliers](image)

One notable problem which is respecting concept of robustness deals with cluster distribution when varying wavelet decomposition level of analyzed data traces (Fig. 5).

![Fig. 5. Fuzzy clustering result for fBm traces with predefined H parameter values within 0.5<H<1, and wavelet decomposition level 2](image)

Such a wide cluster expansion can negatively influence quality of training data for used NN, and a result increase of NN operating error can be expected. To this issue there was not devoted plenty attention in this research and it is mentioned to be the subject of future research.

**Conclusion and future research.**

In this paper we verify a possibility of possible robust wavelet analysis based approach for traffic approximation by introducing predefined classes within data traces that exhibit long range dependent manner, and are characterized by single H parameter. Fractional Brownian motion synthesized data traces serves as a template, for forthcoming wavelet coefficient clustering and operating with neural network learning capabilities, where decomposed real data traces are classified in the face 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.

As a result we have tested 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 wavelet function for NN operating. 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 function are not subject of this research and are mentioned as a ground for the future research.

**References.**


Submitted for publication 2008 02 19