Segmenting LRD Traffic with Wavelets and the Schwarz Information Criterion

David Rincón[1], Flaminio Minerva[2], and Sebastià Sallent[1,2]

[1] Technical University of Catalonia (UPC), 08860 Castelldefels, Spain
[2] i2CAT Foundation, 08034 Barcelona, Spain
drincon@mat.upc.es

Abstract. Network traffic is expected to change its fractal parameters over time. An algorithm for segmenting traffic into regions with homogeneous characteristics would be useful for the development of fractal-aware network algorithms. Starting from the wavelet-based Abry-Veitch estimator, we have added a change-point detection algorithm based on information theory concepts, a redundant wavelet transform, and an automatic alignment procedure based on the Hough transform in order to detect changes in the variance structure of network traffic. We present the results obtained with some synthetic and real traffic traces.

1 Introduction

Network traffic exhibits fractal properties such as self-similarity, long-range dependence (LRD) and multifractality [1, 2]. This has implications for network performance, including increased buffer losses compared with those predicted by classical teletraffic models based on Poisson arrivals [3, 4]. The Abry-Veitch estimator [5] is widely accepted as one of the best and most efficient estimators of fractal parameters. It is based on the Discrete Wavelet Transform (DWT) and can detect the characteristic variance structure of self-similar and LRD processes.

While real traffic is expected to change its behavior over time, few studies have addressed time-varying fractal parameters ([6–11] among them). Identifying the boundaries of segments with homogeneous behavior could be useful for the real-time application of algorithms that exploit the fractal properties of network traffic, such as the congestion control in [12], the predictive bandwidth control in [13] and the effective bandwidth estimator in [14].

Changes in the the scaling (fractal) properties of traffic alter the variance structure. Therefore, an algorithm for detecting the variance change point at each scale analyzed by the wavelet transform would be able to detect changes in the fractal parameters. In our experiments we have used the Iterated Cumulative Sum of Squares (ICSS) and the more powerful Schwarz Information Criterion (SIC) for the change-point detection. More accurate localization of changes can be achieved with a time-redundant version of the DWT, known as the Stationary Wavelet Transform (SWT).

This paper describes the DWT-SIC and SWT-SIC algorithms and some related issues. Section 2 reviews basic concepts and previous work; Section 3

presents the DWT-SIC and SWT-SIC algorithms, together with an alignment technique based on the Hough Transform, some comments on the wavelet family, and the results of tests using synthetic and real traffic traces. Conclusions are presented in Section 4.

2 Previous Work

2.1 Estimation of the LRD Scaling Phenomenon Using Wavelets

Long-range dependence (LRD) is considered to occur in a discrete-time stationary stochastic process $x(n)$ if $S_x(f) \sim c_f/|f|^\alpha$ when $|f| \to 0$, $0 < \alpha < 1$ [2]. In other words, the process exhibits LRD if it has strong low frequency components and its autocorrelation function decays more slowly than a negative exponential. This is known as the *long memory* property. Normalization of the autocorrelation function has been used to extend this definition to non-stationary LRD processes with shifts in the mean and variance [9]. An example of a Gaussian LRD process is Fractional Gaussian Noise (FGN), which we use in our experiments [15, 16].

The scaling parameter α is a qualitative measure related to the intensity of the LRD phenomenon and is usually expressed in terms of the Hurst parameter $H = (1 + \alpha)/2$, while c_f has dimensions of variance and can be interpreted as a quantitative measure of LRD. Although H has received more attention, c_f is also important, since it appears in the expression of queue loss probability when the buffer is fed with LRD traffic, and also in the variance of the sample mean of LRD processes (thus determining the confidence intervals for its estimation [5]).

2.2 Wavelet Transforms

The Discrete Wavelet Transform (DWT) is a powerful tool that provides a fast, efficient and precise estimation of LRD parameters. Given a discrete-time signal $x(n)$, the coefficients of the DWT are denoted by $d_x(j,k) = \langle x, \psi_{j,k} \rangle$, where $\psi_{j,k}$ is a function (typically with finite support) generated from translated and dilated versions of the *mother wavelet* $\psi_{0,0}$: $\psi_{j,k}(n) = 2^{-j/2}\psi_{0,0}(2^{-j}n - k)$ for $j = 1, \ldots, J,\ k \in \mathcal{Z}$, being j and k the scale and the location at which the analysis is performed. The choice of the mother wavelet has implications for the properties of the transform, including linear phase and time invariance [5, 17]. Examples include the *Daubechies* family (`dbN`), which has N vanishing moments and is insensitive to polynomial trends of order up to $N - 1$, and the Least Asymmetric or *Symlets* family, which approximates a linear phase filter.

The resulting transform is orthogonal and can be efficiently implemented using a bank of quadrature-mirror low-pass and high-pass filters ($h(n)$ and $g(n)$, respectively) followed by a downsampling step [18]. In the frequency domain, the DWT causes decomposition in subbands whose spectra are halved at each step, giving rise to a multiresolution analysis in which the original signal is decomposed into a low-pass approximation at scale J, $a_x(J,k)$ and a set of high-pass details $d_x(j,k)$ (the *wavelet coefficients*) for each scale $j = 1, \ldots, J$. The downsampling

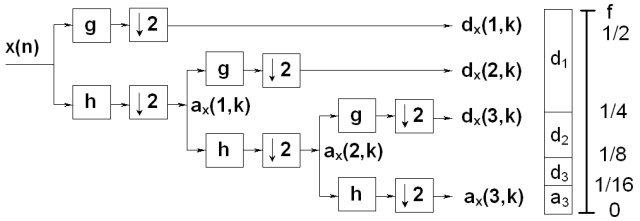

Fig. 1. Left: the DWT as a filter bank (J=3). Right: the normalized frequency subband decomposition, with the approximation a_3 and details $d_{1,\ldots,3}$.

step causes the number of DWT coefficients to be approximately halved at each stage. Figure 1 illustrates the algorithm and the subband decomposition.

The Stationary Wavelet Transform (SWT) is identical to the DWT, except that no downsampling is performed. This results in an increased number of coefficients at each scale and more accurate localization of signal features.

2.3 DWT Analysis of LRD Traffic: the Abry-Veitch Estimator

The Abry-Veitch estimator [5] is an unbiased, efficient estimator of LRD parameters. It uses a graphical tool known as the *LogScale Diagram* to estimate the power of the subbands of the wavelet decomposition in terms of the sample variance of the coefficients at each branch of the filter, μ_j. Given the $1/f$ spectrum,

$$\mu_j = E[d_x^2(j,k)] = 2^{j\alpha} c_f C(\alpha, \psi_{0,0}), \quad j = 1, \ldots, J \tag{1}$$

and taking logarithms on both sides,

$$\log_2(\mu_j) = j\alpha + \log_2[c_f C(\alpha, \psi_{0,0})] + g_j, \quad j = 1, \ldots, J \tag{2}$$

where g_j are bias-correction terms that depend on n_j, the number of wavelet coefficients at scale j. The parameters α and $c_f C$ can be estimated from expression (2) by performing a weighted linear regression on $\log_2(\widehat{\mu_j})$, in which the weight is proportional to n_j. Confidence intervals for the estimation can be derived under Gaussian assumptions. Similar expressions are derived for the SWT [17].

2.4 On-line Estimation of the Scaling Parameter

There is an on-line version of the Abry-Veitch estimator that performs progressive (cumulative) computation of subband variance and returns updated estimates using all available samples [7]. The same authors also studied the robustness of the method when mean and variance shifts are present [9]. Elsewhere, changes in the Hurst parameter were analyzed for a LAN trace in a 1-hour or 4-hour window [8]. A statistical test of the constancy of the scaling parameter over constant-length windows has also been developed and is able to detect changes in H but is insensitive to changes in the variance structure [10]. Furthermore,

a wavelet-based study of TCP flow arrivals revealed nonstationarities, found no strict scaling, and generalized the study to high-order scaling [19]. Finally, an accuracy-based approach has been reported for the real-time estimation of multifractality, where the window length is related to the confidence interval of the estimation [11].

Our approach is more general than the constant window-based methods, since we can localize the variance-transition points to any position and scale. In addition, we can monitor the nonstationarities of the time series (both for the whole process and scale-by-scale). This is important, since LRD estimators can be fooled by variance changes or shifts in traffic volume [6, 20].

2.5 Change-point Detection with the Schwarz Information Criterion

Change-point detection is a classical problem in statistics and several algorithms have been developed to address it. In the past we have used the Iterated Cumulative Sum of Squares (ICSS) [21]. However, that algorithm lacks flexibility in the selection of the significance level (critical values are computed using Monte Carlo simulation) and it is difficult to build a real-time implementation.

The Schwarz Information Criterion (SIC) [22] is a more powerful approach. The logic behind SIC is that a sequence with a variance change point has higher entropy (disorder) than a sequence with constant variance. Therefore, a way to simultaneously detect the presence and location of a single change point at position k is to compute the entropy of the sequence (of length n) and of the pairs of subsequences ($i = 1, \ldots, k$ and $i = k+1, \ldots, n,\ 1 < k < n$), compare their values, and decide if there is a change at k according to whether the entropy of the subsequences is significantly lower than the entropy of the sequence. A binary segmentation procedure is used to detect multiple change points: wherever a change point is detected, the sequence is split into two subsequences and the process is iterated until no more changes are detected. This technique reduces the problem to testing the null hypothesis H_0 (no change is present) against the alternative H_1 (a single change is present). Assuming Gaussianity and independence, the SIC statistic for the two hypotheses is given by

$$H_0 : SIC(n) = n \log 2\pi + n \log \hat{\sigma}^2 + n + \log n \tag{3}$$

$$H_1 : SIC(k) = n \log 2\pi + k \log \hat{\sigma_1}^2 + (n-k) \log \hat{\sigma_2}^2 + n + 2 \log n \tag{4}$$

where $1 < k < n$, and $\hat{\sigma}^2$, $\hat{\sigma_1}^2$, $\hat{\sigma_2}^2$ are the unbiased Maximum Likelihood Estimators (MLEs) of the variances of the entire sequence and of the first and second subsequence, respectively. The introduction of a critical value c_α allows us to adjust the statistical significance α of the estimation:

$$1 - \alpha = Pr\{SIC(n) < \min_{1<k<n} SIC(k) + c_\alpha | H_0\} \tag{5}$$

The asymptotic expression of c_α for the Gaussian case can be found in [22]. Another free parameter is n_{min}, the minimum length for the binary segmentation. This is not the constant-length window used by other authors, since n_{min} can be greater than the minimum distance between two detected change points.

3 DWT-SIC and SWT-SIC Algorithms

3.1 Linking Wavelet Transforms to SIC

Our main contribution is to apply the SIC change-point estimator to the output of each branch of the wavelet filter bank. An alignment of variance changes across several scales will indicate a change in variance structure at that position.

Certain wavelet-related issues must be taken into account [17]. The choice of the mother wavelet is important, since it affects the estimation in several ways. As with any filtering process, there are border effects due to the convolution of finite-length sequences that influence the accuracy with which the wavelet variance is estimated, meaning that boundary coefficients (those affected by border effects) should not be used. Another effect is the phase shift introduced by the mother wavelet. This can be corrected or minimized by the choice of the wavelet family (for example, *Symlets*). There is also a trade-off between wavelet filter length (L), computational complexity, and accuracy. Increasing L allows a better estimation of wavelet variance but requires a greater number of computations and involves fewer non-boundary coefficients.

The choice of the transform is also important. The downsampling causes the number of DWT samples available at each scale to decrease with increasing scale, while the SWT maintains the same number of samples at each scale. An SWT-based algorithm provides higher accuracy for the detection of change points. At higher scales, where LRD is actually detected, DWT-SIC performs badly because of the limited number of samples available. SWT-SIC corrects this problem at the cost of a larger number of computations and greater memory requirements. More problematically, at higher scales there is also correlation, contradicting the assumption of independence in SIC. We balance correlation by decreasing the significance level α and increasing n_{min} with increasing scale. Since the correlation effect increases with the length of the mother wavelet, the shortest possible wavelet should be used. In our experiments, we have used the *Daubechies* family, starting with `db1` and increasing the number of vanishing moments until the analysis is stabilized. In most cases, `db3` or `db4` is sufficient.

3.2 Tests Using Synthetic Traces

Both DWT-SIC and SWT-SIC algorithms have been validated using synthetic traces of Fractional Gaussian Noise (FGN) with constant variance and regions with different values of H. The samples were generated as described in [16].

The results presented here refer to a trace with 131072 samples, 3 regions with $H = 0.8, 0.9$ and 0.7, and change points at positions 65536 (*smooth* change $0.8 \rightarrow 0.9$) and 98304 (*abrupt* change $0.9 \rightarrow 0.7$). In this preliminary phase of the study, α and n_{min} were chosen empirically. As shown in Figure 2 (left), DWT-SIC clearly identifies both change points at lower scales ($j = 1 \ldots 7$ for the smooth change and $j = 1 \ldots 9$ for the abrupt change), with the exception of $j = 5$ for the first change and $j = 4$ for the second change. These two exceptions are what we call *blind points*, i.e. the scales where the LogScale Diagram alignments

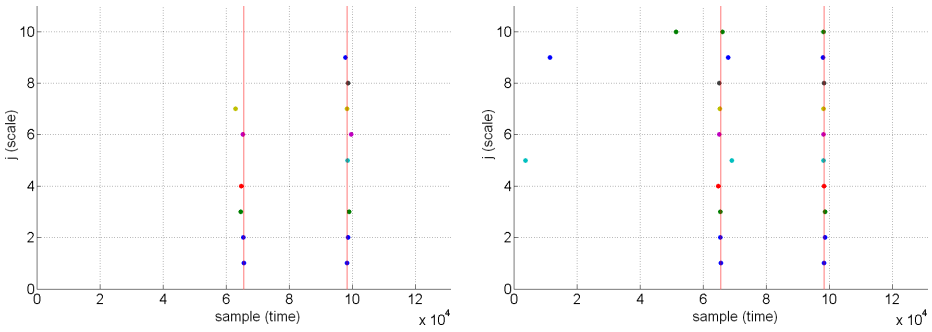

Fig. 2. Change point candidates at each scale, with vertical lines at the true change positions. Left: DWT-SIC at J=10 levels, `db2`, $\alpha = 10^{-2}$, $n_{min} = 256$. Right: SWT-SIC at J=10 levels, `db2`, $\alpha_{j=1,\ldots,10} = [10^{-1}\ 10^{-2}\ 10^{-3}\ 10^{-4}\ 10^{-5} \ldots 10^{-5}]$, $n_{min_{j=1,\ldots,10}} = [256\ 1024\ 4096\ 16384\ 65536 \ldots 65536]$.

cross and, therefore, the variances are similar on either side of the change point. This reveals the inability of DWT-SIC to resolve the changes when variances are too close. At the higher scales, the downsampling of DWT masks the changes in the variance, making them smoother with time; the abrupt change lasts for two scales more than the smooth one, thanks to the larger difference in the variance.

The SWT-SIC algorithm outperforms the DWT-SIC (Figure 2, right). It works well even at the highest scales and is not easily fooled by the blind-point phenomenon. In other tests, however, variances at the blind point were so close that even SWT was either unable to detect them or incorrectly shifted the position (such as scale 5, first change point). Furthermore, SWT-SIC is prone to false points such as those that can be seen to the left of the first change point at scales 5, 9, and 10. These false points reveal some irregularities in the synthetic trace and are easily detected and eliminated during the alignment-detection procedure described in the next section. Although DWT-SIC also exhibits false points, the rate of appearance is considerably lower than that of SWT-SIC.

In summary, the DWT-based algorithm gives good results at lower scales but suffers from a lack of accuracy at higher scales (where LRD is to be found), while SWT-SIC offers the same resolution at all scales, at a cost of a higher computational load, and is more error-prone. With these conclusions in mind, we thought that a joint algorithm that uses DWT-SIC at the lower scales and SWT-SIC at the higher scales would exploit the advantages of both. An example of such an algorithm is provided in Section 4. More details and testing of the joint and disjoint algorithms are provided in [23].

3.3 Alignments of Detected Change Points Across Scales

After detection of change points, an alignment-detection algorithm is required to detect whether a change is present in enough scales and sufficiently close in time to consider it a change in variance structure.

The Hough transform is used in image processing to detect features of a particular shape through a polar representation of lines [24]. The transform maps a line in the space domain (x, y) to a point in the Hough domain (θ, ρ), where $\rho > 0$ is the length of the normal segment from the origin to this line and $\theta \in [0, \pi)$ is the orientation of ρ with respect to the x axis. The dual interpretation is that a point (x_0, y_0) in the space domain is mapped in the Hough domain to a cosinusoid. Collinear points in the Cartesian image space are detected in the Hough space as curves that intersect at a common (ρ, θ) point.

As far as the implementation is concerned, the Hough space is quantized into finite intervals or accumulator cells. Each point in the space domain is transformed into a cosinusoid in the Hough domain and the cells that lie along this curve are incremented. Peaks in the accumulator matrix represent strong evidence that a line exists in the image. A voting process with a minimum *quorum* (the number of aligned points across scales) can then be performed.

3.4 Interpreting the Results

Interpreting the results (i.e., detecting changes in the scaling parameter or non-stationarity of the variance) is easy when one is working with synthetic traces. A level shift in the traffic volume leads to a variance change at every scale and all changes are simultaneously positive (increase) or negative (decrease), while a change in the scaling parameter produces different signs in the variance change (the scales above the blind point will increase while the others will decrease, or vice versa) and possibly a blind point in the middle scales. If we assume a self-similar model, we will look for strict variance alignments; in contrast, an LRD model only needs alignments at higher scales. Real traffic traces are usually more complex to interpret as we can find simultaneous changes of variance and scaling parameter, or alignments in limited ranges of scales.

Since we assume an LRD model in our analysis, we lose generality due to our focus on monofractality rather than multiscaling or multifractality. However, recent studies have shown that a Poisson model for sub-second scales can be employed with core network traffic, while at higher scales piecewise-linear non-stationarity and long-range dependence seem to fit well with measurements of real networks [25]. Nevertheless, we can always look for less-rigorous alignments, like the typical effect produced by congestion or rate limitation of TCP traffic on WANs that increases the value of the variances over selected scales (usually at scales below round-trip time, where TCP rate adaption occurs).

3.5 Testing the Algorithms With Real Traffic

We now use the DWT/SWT-SIC method to analyze the series of cumulated work per time interval (bytes/slot) aggregated at $\delta = 10$ msec (W_δ) in the `BC-pAug89` Bellcore trace. This trace has been extensively studied and is known to have a Hurst parameter of around 0.8 when analyzed as a whole [1]. The stationarity of its scaling parameter has been studied in [26]. In that study, the authors detected variations of H in the range $0.7 - 0.85$ (approximately). In a more recent study,

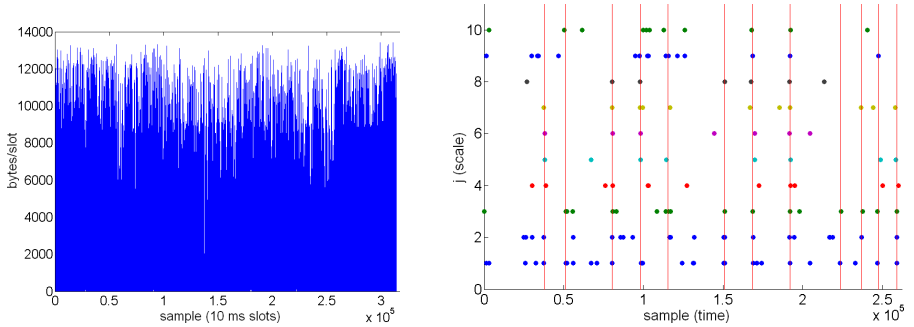

Fig. 3. *Left*: `BC-pAug89` Bellcore trace aggregated at $\delta = 10$ msec. *Right*: candidate change points with the DWT/SWT-SIC algorithm and `db3` wavelet. SIC parameters: $\alpha^{dwt}_{j=1,\ldots,4} = 10^{-2}$, $\alpha^{swt}_{j=5,\ldots,10} = 10^{-5}$, $n^{dwt}_{minj=1,\ldots,4} = 25000$, $n^{swt}_{minj=5,\ldots,10} = 58000$. Hough parameters: resolution = 1000 samples, quorum = 3 points.

although no statistical evidence to reject the null hypothesis of a stationary H was found, the actual values of H varied in the aforementioned range [10].

We performed a joint DWT/SWT-SIC analysis with `db3` at scales $j = 1, \ldots, 4$ (DWT) and $j = 5, \ldots, 10$ (SWT), and a combination of the free parameters that provides a good trade-off between resolution of the detected change points and readability of the diagram. Figure 3 shows the change points diagram for the first 262144 samples (due to a limitation of our SWT implementation).

We found that the variance structure changes considerably throughout the trace. Table 1 shows the segments and the corresponding H estimations (assuming LRD scaling). The observation of little or no overlap between neighboring confidence intervals in most cases increases our confidence in the method. The values of H range from 0.664 to 0.830 and are consistent with the *average* value of $\widehat{H} = 0.812$ with 95% confidence intervals [0.804, 0.820] calculated for the trace. Our results for the segmentation algorithm exhibit reasonable agreement with the constant-length window analysis in [26], since we find approximately the same temporal evolution of H but can define the boundaries precisely.

4 Conclusions

We have described a family of wavelet-based algorithms capable of segmenting LRD traffic traces into segments with homogeneous behavior. This includes two disjoint algorithms based on the Schwarz Information Criterion (DWT-SIC and SWT-SIC) and a joint method. After validating it with synthetic traces, the method was applied to a real traffic trace from the Bellcore dataset. The results obtained confirmed the findings of other authors, improving the accuracy in the identification of change locations. Although the analysis assumed a strict scaling model, the methods are actually capable of finding any change in the variance structure of the analyzed trace. Consequently, they are also applicable to more complex models.

Table 1. DWT/SWT-SIC results for the `BC-pAug89` trace aggregated at $\delta = 10$ msec, using `db3` wavelet.

Segment boundaries	Estimated $\widehat{H}$ (95% CI)	Regression scales
1-37819	0.830 (0.814,0.847)	$j_1 = 3, j_2 = 9$
37820-50907	0.744 (0.715,0.773)	$j_1 = 3, j_2 = 9$
50908-80524	0.777 (0.732,0.822)	$j_1 = 5, j_2 = 9$
80525-98233	0.829 (0.805,0.853)	$j_1 = 3, j_2 = 10$
98234-115296	0.762 (0.721,0.804)	$j_1 = 4, j_2 = 8$
115297-150765	0.820 (0.796,0.844)	$j_1 = 4, j_2 = 10$
150766-168693	0.755 (0.715,0.796)	$j_1 = 4, j_2 = 8$
168694-191862	0.804 (0.779,0.829)	$j_1 = 3, j_2 = 7$
191863-223759	0.706 (0.692,0.719)	$j_1 = 2, j_2 = 7$
223760-237003	0.799 (0.758,0.841)	$j_1 = 4, j_2 = 11$
237004-247535	0.718 (0.681,0.755)	$j_1 = 3, j_2 = 7$
247536-259053	0.664 (0.642,0.687)	$j_1 = 2, j_2 = 7$
258228-262144	0.742 (0.729,0.754)	$j_1 = 3, j_2 = 11$

Several issues require further investigation, including calculation of the confidence intervals of the estimations, fine tuning a real-time version of the algorithms [27] and evaluating its computational cost, and studying how departures from the ideal situation (non-Gaussianity, nonstationarities) bias detection of the change point. The Generalized Gaussian Distribution (GGD) seems to be the most appropriate for the non-Gaussian distributions of the wavelet coefficients of real traffic traces, as reported recently in a study in which the wavelet coefficients of real traffic traces obeyed a mixture of Laplacian and Gaussian distributions [28]. Finally, it may be possible to generalize our method to multifractal or multiscaling processes by studying the changes of the higher-order moments. Our efforts are focused on real-time characterization of fractal traffic, which may be useful for the development on new, fractal-aware algorithms.

Acknowledgments

This work was partially supported by funds from the Spanish Government and FEDER (TSI2005-06092), and by the i2CAT Foundation. The work was carried out as part of the EuroNGI Network of Excellence. The authors are grateful to the referees for their insightful comments.

References

1. Leland, W., Taqqu, M., Willinger, W., Wilson, D.: On the Self-similar Nature of Ethernet Traffic. IEEE/ACM Transactions on Networking **2** (1994) 1–15
2. Park, K., Willinger, W., eds.: Self-similar Traffic and Network Performance. John Wiley & Sons (2000)
3. Norros, I.: A Storage Model with Self-similar Input. Queueing Systems **16** (1994) 387–396

4. Dang, T.D., Molnar, S., Maricza, I.: Some Results on Multiscale Queueing Analysis. In: Proceedings of ICT 2003. Volume 2. (2003) 1631–1638
5. Veitch, D., Abry, P.: A Wavelet-based Joint Estimator of the Parameters of Long-range Dependence. IEEE Transactions on Information Theory **45** (1999) 878–897
6. Duffield, N., Lewis, J., O'Connell, N., Russell, R., Toomey, F.: Statistical Issues Raised by the Bellcore Data. In: 11th IEE UK Teletraffic Symposium. (1994)
7. Roughan, M., Veitch, D., Abry, P.: On-line Estimation of the Parameters of Long-range Dependence. In: Proceedings of Globecom 98. Volume 6. (1998) 3716–3721
8. M.Roughan, D.Veitch: A Study of the Daily Variation in the Self-similarity of Real Data Traffic. In: Proc. of the International Teletraffic Congress 16. (1999) 67–76
9. M.Roughan, D.Veitch: Measuring Long-range Dependence under Changing Traffic Conditions. In: Proc. of IEEE INFOCOM'99. (1999) 1513–1521
10. Veitch, D., Abry, P.: A Statistical Test for the Time Constancy of Scaling Exponents. IEEE Transactions on Signal Processing **49** (2001) 2325–2334
11. Atzori, L., Aste, N., Isola, M.: Estimation of Multifractal Parameters in Traffic Measurement: an Accuracy-based Real-time Approach. Procs. ICC 2005 21–25
12. He, G., Gao, Y., Hou, J., Park, K.: A Case for Exploiting Self-similarity of Network Traffic in TCP Congestion Control. In: Procs. IEEE ICNP 2002. (2002) 34–43
13. Ouyang, Y.C., Yeh, L.B.: Predictive Bandwidth Control for MPEG video: a Wavelet Approach for Self-similar Parameters Estimation. ICC 2001. 1551–1555
14. Yu, Xiang Yu; Li-Jin Thng, I., Jiang, Y.: Measurement-based Effective Bandwidth Estimation for Long Range Dependent Traffic. Procs. of IEEE Region 10 Intl. Conf. on Electrical and Electronic Technology (TENCON). Volume 1. (2001) 359–365
15. Paxson, V.: Fast Approximation of Self-similar Network Traffic. Technical report, LBL-36750/UC-405 (1995)
16. Stoev, S.: Generation of FGN. http://math.bu.edu/people/sstoev/ (2005)
17. Percival, D., Walden, A.: Wavelet Methods for Time Series Analysis. Cambridge University Press (2002)
18. Mallat, S.: A Theory for Multiresolution Signal Decomposition: the Wavelet Representation. IEEE Pattern Anal.and Machine Intell. **11** (1989) 674–693
19. Uhlig, S.: Nonstationarity and High-order Scaling in TCP Flow Arrivals: a Methodological Analysis. ACM SIGCOMM Comp. Comm. Review **34** (2004) 9–24
20. Molnar, S., Dinh Dang, T.: Pitfalls in Long Range Dependence Testing and Estimation. In: Proceedings of IEEE Globecom 2000. Volume 1. (2000) 662–666
21. Rincón, D., Sallent, S.: Characterizing Fractal Traffic with Redundant Wavelet-based Transforms. In: Proceedings of the First EuroNGI Workshop on New Trends in Modelling, Quantitative Methods and Measurements. (2004) 361–371
22. Chen, J., Gupta, A.: Testing and Locating Variance Change Points with Application to Stock Prices. Journal of the American Statistical Assoc. **92** (1997) 739–747
23. Minerva, F.: Wavelet Analysis of Long-range Dependent Traffic. Master Thesis, Faculty of Telecommunications Engineering, University of Pisa (2005)
24. Russ, J.: The Image Processing Handbook (Third Edition). CRC Press (1999)
25. Karagiannis, T., Molle, M., Faloutsos, M., Broido, A.: A Nonstationary Poisson View of Internet Traffic. In: IEEE INFOCOM 2004. Volume 3. (2004) 1558–1569
26. Veitch, D., Abry, P.: Wavelet Analysis of Long-range-dependent Traffic. IEEE Transactions on Information Theory **44** (1998) 2–15
27. Rincón, D., Sallent, S.: On-line Segmentation of Non-stationary Fractal Network Traffic with Wavelet Transforms and Log-likelihood Statistics. In: Proceedings of QoS-IP 2005, LNCS 3375. (2005) 110–123
28. Glomb, P.: Analysis of FGN and HTTP Requests Traces Using Localized Multiscale H Parameter Estimation. In: Procs. of IEEE SAINT 2005. (2005) 288–291