Automated detection of changes in computer network measurements using wavelets

Automated Detection of Changes in Computer Network Measurements using Wavelets

Konstantinos G Kyriakopoulos and David J Parish
Department of Elecronic and Electrical Engineering
Loughborough University
United Kingdom
Email: kostas7@ieee.org

***Abstract*—Monitoring and measuring various metrics of high speed and high capacity networks produces a vast amount of information over a long period of time. For the collected monitoring data to be useful to administrators, these measurements need to be analyzed and processed in order to detect interesting characteristics such as sudden changes. In this paper wavelet analysis is used along with the universal threshold proposed by Donoho - Johnstone in order to detect abrupt changes in computer network measurements. Experimental results are obtained to compare the behaviour of the algorithm on delay and data rate signals. Both type of signals are measurements from real networks and not produced from a simulation tool. Results show that detection of anomalies is achievable in a variety of signals.**

***Index Terms*—Computer networks, measurements, wavelet analysis, anomaly detection**

I. Introduction

Monitoring and measuring various metrics of high speed and high capacity networks produces a vast amount of information over a long period of time. These metrics describe the status and performance of the network in terms of utilization, congestion, packets lost, etc and help operators to identify potential problems.

For the collected monitoring data to be useful to administrators, these measurements need to be analysed and processed in order to detect interesting characteristics such as sudden changes. Identifying such characteristics in large amounts of data is a not an easy task and has been an interest of network researchers for many years.

Changes in networks cause changes in their performance and this is reflected in the collected measurements. These changes may occur due to change of load in the network, fault, or planned alterations in the infrastructure.

An automated tool for the data analysis and change detection phases would reduce costs required by the training and retaining of human resources. An example for the need of this tool comes from research conducted by the authors for the UKLight network.

The UKLight initiative is a 10 Gb/s, high capacity research network facility that interconnects JANET, the UKs research and educational network, with several other continental research networks. Researchers running experimental protocols over that network and administrators require reports of significant changes in the data.

The aim of this work is to analyze the collected measurements of network performance, which can be represented as time series processes and estimate when a sudden change takes place. Wavelet analysis is used for multi-resolution analysis and then a threshold is applied that filters the wavelet domain coefficients and keeps only those that represent a significant change in the time domain.

So far, wavelets have been generally used to detect network performance problems. They have been applied to traffic rate signals in order to infer the time scale associated with the dominant RTT through the examination of the energy function of the detail coefficients [1]. They have also been used for de-noising one-way delay signals in order to detect shared congestion between different flows [2]. [3] shows that wavelet filters are quite effective at exposing the details and characteristics of ambient and anomalous traffic. [4], [5] analyze the correlation of destination IP addresses of outgoing traffic at an egress router. Based on statistical historical margins, estimated after using wavelet analysis, sudden changes are detected.

The proposed algorithm is applied to network delay and data rate signals. Experimental results are obtained to examine how well the applied method detects the changes in signals.

The rest of the paper is structured as follows. In section 2 wavelet analysis is discussed. In section 3 the methodology followed for producing off-line results is presented and the results are given in section 4. Finally, conclusions and ideas for future work are given in Section 5.

II. Wavelet Analysis

A. Wavelet Analysis Advantages

A common methodology for detecting events in a network involves using historical data to estimate the mean and the variance and then flagging events outside the third standard deviation as anomalous [4] [5] [6].

However, the time varying nature of a network should be taken into consideration. The performance of a network varies with respect to the time of day, day of the week, or season of the year. Thus, for a system to properly detect anomalies it should adapt to the dynamic nature of the network [6].

In this work wavelets are used to adapt to the time varying environment of a network and detect any abrupt changes that are included in the measurements taken from that network.

1-4244-1251-X/07/$25.00 2007 IEEE.

In contrast with other signal analysis techniques that use a constant window size to analyze a section of a signal (for example DCT, STFT), wavelet analysis has the benefit of varying the window size. This means that wavelets can efficiently trade time resolution for frequency resolution and vice versa. For this reason, wavelets can adapt to various time-scales and perform local analysis. In simple words, wavelets can reveal both the forest and the trees [7] [8].

Wavelets have the ability to detect characteristics of non-stationary signals due to their finite nature that describes local features better than say sinusoids. Non-stationary signals are stochastic signals whose statistical properties change with time. A lot of research in network traffic analysis shows that packet switched data traffic patters are statistically self-similar. Self-similar processes are by definition non-stationary [9] [10].

B. Multi-resolution Signal Decomposition

The wavelet analysis transforms a given signal s[n] of n samples into n/2 approximation (scaling) and n/2 detail (wavelet) coefficients. The approximation coefficients represent the smoothed version of the signal (low frequency bands), while the detail coefficients represent the detailed version (high frequency bands).

The basic idea of wavelet analysis is that an average of two samples of signal s at scale j produces an approximation coefficient at the next higher scale j+1. The difference between those samples produces a detail coefficient at scale j+1. Thus, for a specific scale, approximation coefficients are associated with the averages, whereas detail coefficients represent the change of averages. In the case of the simplest wavelet function, Haar, this can be expressed as:

$$\alpha_{j+1} = \frac{S_j[1] + S_j[2]}{2} \qquad d_{j+1} = \frac{S_j[1] - S_j[2]}{2}$$

Wavelet Analysis can be used as a Multi-resolution Signal Decomposition (MSD) tool, decomposing a signal into scales of varying time and frequency resolution. Initially, the first level (or scale) of decomposition of the multi-resolution analysis takes place. The same process can be applied again on the previously produced (n/2) approximation coefficients yielding n/4 detail and approximation coefficients and so on for higher scales. The total group of detail coefficients from scale 1 up to J and the approximation coefficients at scale J compose the wavelet decomposition tree at scale J.

C. Quadrature Mirror Filter

In 1988, Mallat [11] developed a Fast Wavelet Transform (FWT) algorithm that became well known in the signal processing community as a two channel subband coder using conjugate filters or quadrature mirror filters (QMF).

For the decomposition phase, two finite impulse filters are used. The high-pass filter (HPF) produces the detail coefficients and the low-pass filter (LPF) the approximation of the signal. The output of the LPF becomes the input of the next pair of filters for further decomposition at higher scales.

The QMF pair divides the input signal into low-frequency and high-frequency components. The dividing frequency is between 0 Hz and the maximum frequency of the analyzed signal, which according to the Nyquist theorem is half of the data sampling frequency.

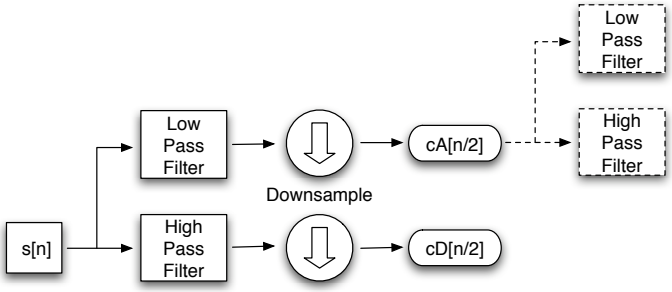

Fig. 1. Quadrature Mirror Filter Bank

III. Methodology

A. Calculating the threshold

Because detail coefficients are actually the changes of the average, those coefficients with large magnitude reveal a change in the original signal. In order to filter those coefficients that have a large enough magnitude to infer change in the original signal, a threshold is required.

For this task, a threshold based on the Donoho - Johnstone universal threshold (aka VisuShrink) [12], [7] is utilized. For each level of decomposition the threshold is rescaled by a level-dependent estimation of the level's noise σ_{lev}.

Thus, the level dependent threshold is of the following form:

$$T_{lvl} = \sigma_{lvl} \times \sqrt{2 \log_e n} \tag{1}$$

Where n is the number of the total wavelet domain coefficients and σ_{lev} is the level-dependent noise standard deviation. As suggested by [12], the median absolute deviation is used as a robust estimation for the noise standard deviation.

$$\hat{\sigma}_{lvl} = \frac{median(|cDetail_{lvl}|)}{0.6745} \tag{2}$$

where cDetail are the detail coefficients for level lvl.

B. Algorithm

For the analysis part, the Haar wavelet was used as the mother wavelet for the analysis because it has the following advantages [13]:

1) It is conceptually simple
2) It is fast
3) It is memory efficient

The methodology flow chart is presented below in Fig. 2. After applying wavelet analysis on the examined signal, the threshold (estimated as described above) was applied on each level. This step filters all coefficients that do not represent a significant change.

C. Estimating the position of a change

Some changes may appear at more than one scale but some others may appear only at one scale. This section describes the algorithm (Fig. 3) for estimating the position of a change depending on the number of scales that it is revealed.

In this phase, the notion of descendants, borrowed from Shapiro's Embedded Zero-Tree Wavelet algorithm [14], is used. A coefficient at a coarse scale is named a parent and all coefficients at the next finer scale describing the same time location are its children. These children similarly may have other children coefficients that are referred to as descendants of the parent coefficient. Coefficients at the highest level are not descendants by definition.

The coefficients are scanned in a decreasing level order. First coefficients at the highest level are scanned. If a coefficient is found with a value not equal to zero, its value and position in the decomposition tree are stored.

Afterwards, its descendants are scanned. If a descendant is non-zero and has larger value than the parent, then its value and position replace the previous entries. Otherwise, the value and position variables remain the same. This process of checking the descendants is named "check descendants" in Fig. 3 (the flow chart on the left of the brace). The whole process iterates for checking all the descendants progressively down the decomposition tree.

When a descendant coefficient is scanned and has non-zero value, then it gets recorded so it will not be scanned again as the scanning of coefficients continues in lower scales.

It should be noted that coefficients that belong in the lowest two scales of decomposition (Level 1 and Level 2) are not scanned during the above process in order to minimize false positive detections.

IV. Off-line results

In this section, we present off-line experimental results, i.e. the test signals are already captured from test beds and real networks as discussed below and are later fed into the algorithm developed using the methodology and techniques discussed above. For implementation of the methodology, MATLAB® and the Wavelet toolbox were used [7].

In order to examine how well the algorithm performs, 30 delay and 30 data rate signals of 1024 samples each were used. The delay signals were measured on a research test bed. Traffic generators were used to emulate a time of day profile similar to that of a commercial network. Delay signals are usually smooth with sudden discrete bursts spread over the data.

The data rate signals are from a real commercial network that generates around 25 TB of data per day and has data rate between 300 Mbps and 1.4 Gbps. In a window size of 2 seconds 30.000 unique IP addresses may be observed in that network.

The proposed procedure detects the anomalies in the examined delay and data rate signals. In the following figures, the original examined signal is presented on top and the detected changes on the bottom of the figures. The significant coefficients produced after the thresholding, described in section III, are normalized and then plotted in the time instance that they represent.

Fig. 4 shows a data rate signal with four instances of significant change. All changes have been detected and plotted in the graph. The third change in the signal lasts for a longer period of time than the rest. However this is also reflected in the graph with high detection values in the time axis around samples 600 - 630.

In Fig. 5 a bursty data rate signal is presented that includes a big spike along some time samples (810 - 820). The anomaly is captured and plotted along the time samples that the change appears.

Fig. 6 shows another case of a bursty data rate signal with two spikes. Both spikes are captured and accurately represented in the graph.

Fig. 7 shows a very bursty delay signal. The output of the detection algorithm is not identifying any instance as suspicious. However, around samples 450 and 800 there are two instances that are set apart from the rest of signal. Those two instances can be identified if the lowest two scales were examined. On the other hand, examining coefficients in these scales could increase false positives in the experiments.

Fig. 8 shows a delay signal with several spikes in the beginning and a region of burstiness later on. The algorithm

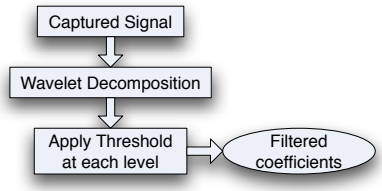

Fig. 2. Algorithm Flow Chart

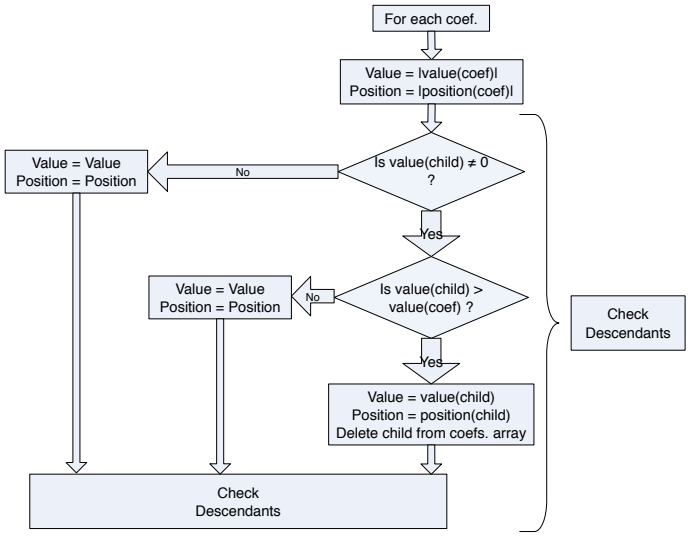

Fig. 3. Algorithm for estimating position of a change

detects the whole region of burstiness along with the most significant spikes.

In Fig. 9 another delay signal is pictured with eight spikes of different size. All of them are detected and accurately plotted both in time and size.

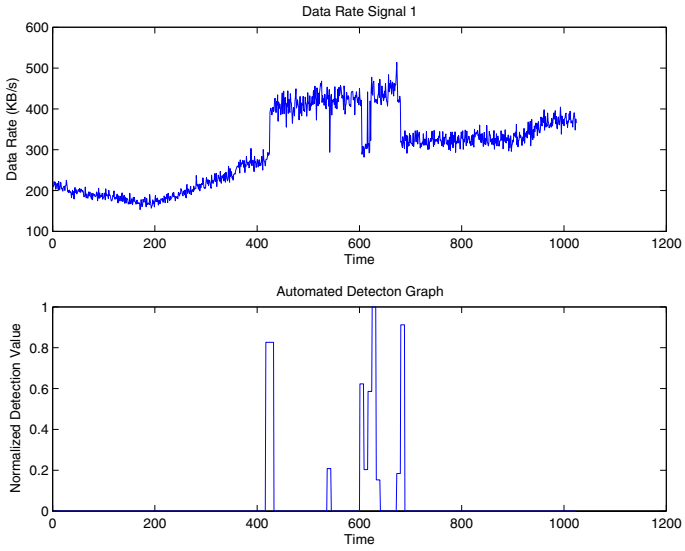

Fig. 4. Detecting Changes in Data Rate Signal 1

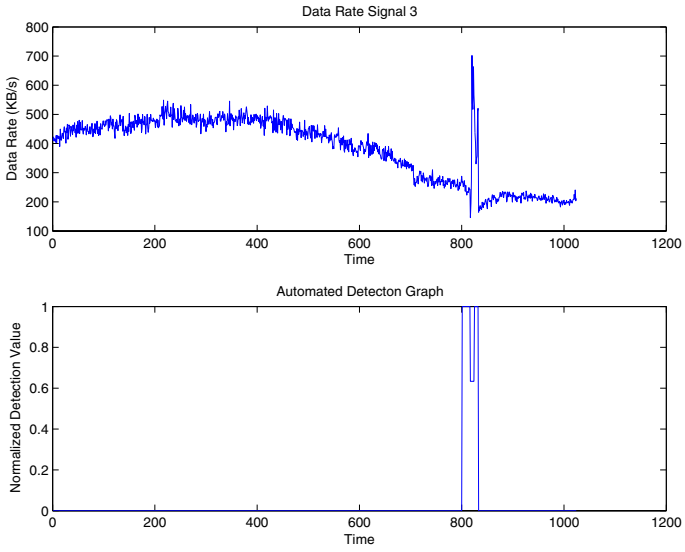

Fig. 5. Detecting Changes in Data Rate Signal 2

V. Conclusion

In this paper a wavelet transform based signal analysis is used along with a threshold proposed by Donoho - Johnstone for detecting abrupt changes in computer network measurements such as delay and data rate. The signals examined were from real computer networks and not from simulation tools.

The time adaptive characteristic of wavelet analysis makes it a suitable tool for examining an environment that is time varying such as the computer network. Additionally, wavelet analysis can perform a local analysis and provide both frequency and time resolutions, which are necessary for the anomaly detection procedure. This could not be possible with the global representation offered from Fourier analysis.

After using the multi-resolution analysis capability of wavelets, the universal threshold is applied to filter those coefficients with a value large enough to indicate a significant change in the original signal. In order to determine as accurately as possible the position of a change, the coefficients are scanned in a progressive way from the largest to the smallest scale. The duration of the anomaly is also indicated by the algorithm.

As for future work, the algorithm will be implemented in a real-time computer network-monitoring tool. This would allow

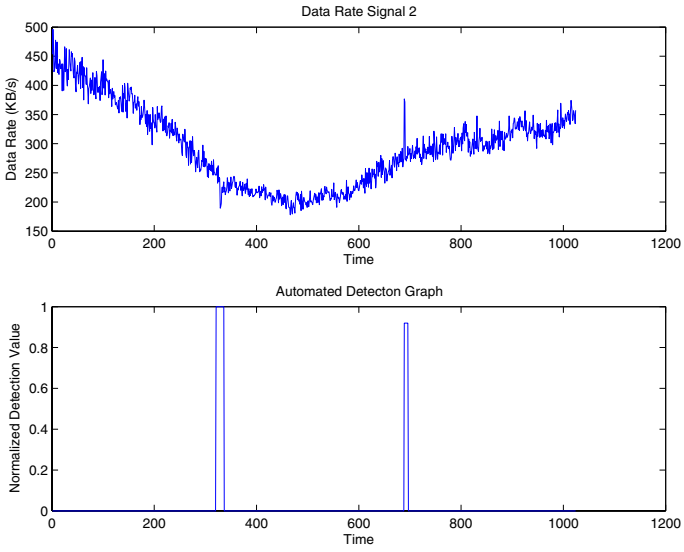

Fig. 6. Detecting Changes in Data Rate Signal 3

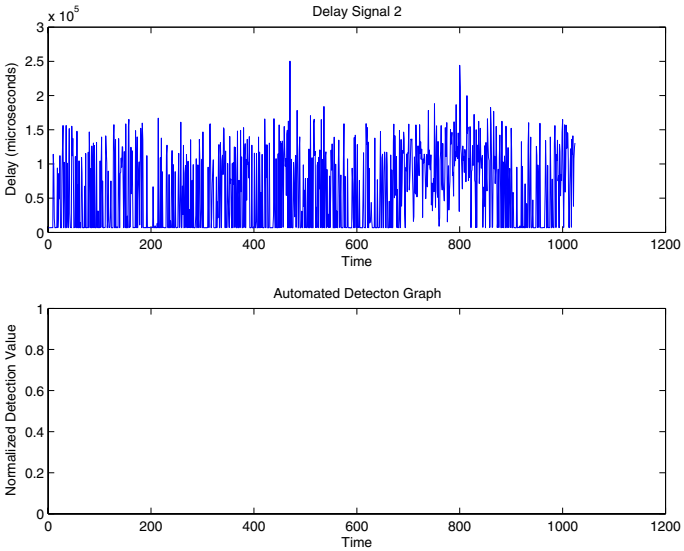

Fig. 7. Detecting Changes in Delay Signal 1

the detection of anomalies in an on-line manner. A promising candidate seems to be CoMo, a passive monitoring platform developed for the purpose of measuring performance metrics of high speed links and replying to real time queries [15].

The proposed algorithm can be implemented as a module in the CoMo platform. CoMo will be responsible for capturing data packets and producing measurements of the network, while the module of the proposed algorithm will detect anomalies in the analyzed captured signal.

The calculation time of the algorithm will not be an issue for the on-line implementation of the anomaly detection algorithm. This is because when analyzing a data rate or delay signal where each sample is per second, a few milliseconds of processing time are adequate. Capturing such a signal of, say, 1024 measurement points would require 1024 seconds. Thus, there is a window of around 17 minutes for the analysis and detection phases to complete.

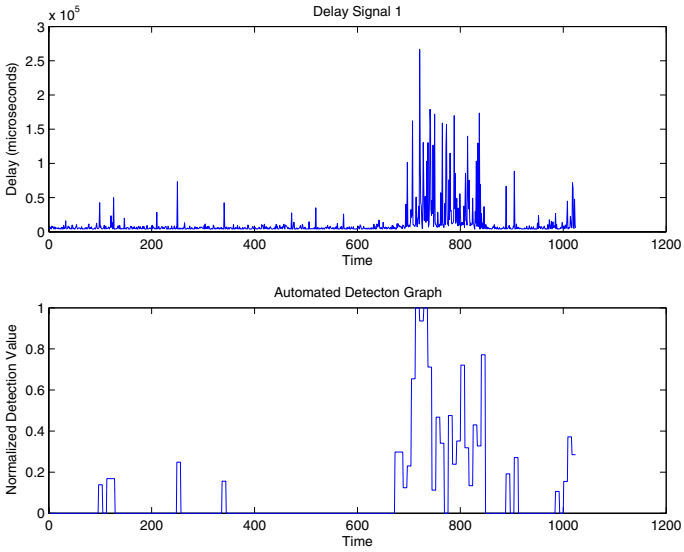

Fig. 8. Detecting Changes in Delay Signal 2

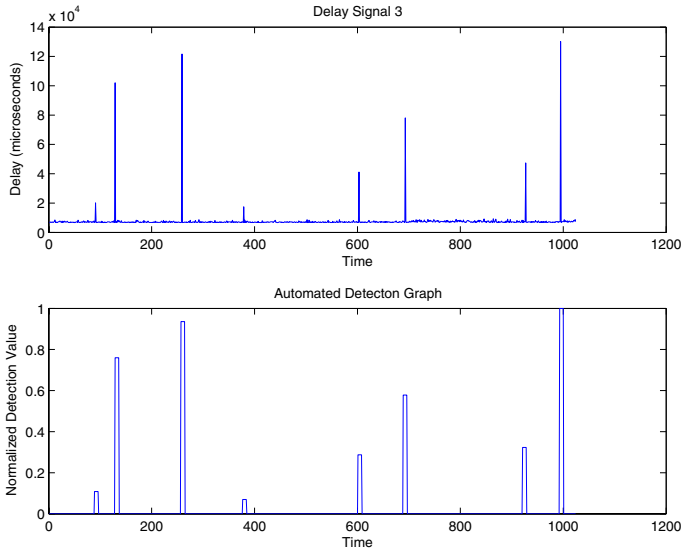

Fig. 9. Detecting Changes in Delay Signal 3

REFERENCES

[1] P. Huang, A. Feldmann, and W. Willinger, "A non-intrusive, wavelet-based approach to detecting network performance problems," in *Proceedings of the ACM SIGCOMM Internet Measurement Workshop*, November 1-2 2001, pp. 213–227.

[2] M. S. Kim, T. Kim, Y. Shin, S. S. Lam, and E. J. Powers, "A wavelet-based approach to detect shared congestion," in *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, 30 August through 3 September 2004, pp. 293–305.

[3] P. Barford, J. Kline, D. Plonka, and A. Ron, "A signal analysis of network traffic anomalies," pp. 71–82, 2002.

[4] S. S. Kim, A. L. N. Reddy, and M. Vannucci, "Detecting traffic anomalies through aggregate analysis of packet header data." in *NETWORKING*, ser. Lecture Notes in Computer Science, N. Mitrou, K. P. Kontovasilis, G. N. Rouskas, I. Iliadis, and L. F. Merakos, Eds., vol. 3042. Springer, 2004, pp. 1047–1059.

[5] ——, "Detecting traffic anomalies using discrete wavelet transform." in *ICOIN*, ser. Lecture Notes in Computer Science, H.-K. Kahng, Ed., vol. 3090. Springer, 2004, pp. 951–961.

[6] F. Feather, D. P. Siewiorek, and R. A. Maxion, "Fault detection in an ethernet network using anomaly signature matching." in *Proceedings of the Conference on Communications Architectures, Protocols and Applications*. San Francisco, CA, USA: ACM, Sep 13-17 1993, pp. 279–288. [Online]. Available: http://dx.doi.org/10.1145/166237.166264

[7] M. Misiti, Y. Misiti, G. Oppenheim, and J. Poggi, "Matlab wavelet toolbox," The MathWorks, Inc., Tech. Rep., 1997-2004. [Online]. Available: http://www.mathworks.com/access/helpdesk/help/pdf_doc/wavelet/wavelet_ug.pdf

[8] J. I. Agbinya, "Discrete wavelet transform techniques in speech processing," in *Proceedings of the 1996 IEEE Region 10 TENCON - Digital Signal Processing Applications Conference*, vol. 2. Perth, Aust: IEEE, Nov 26-29 1996, pp. 514–519.

[9] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, "Multiscale nature of network traffic," *IEEE Signal Processing Magazine*, vol. 19, no. 3, pp. 28–46, May 2002. [Online]. Available: http://dx.doi.org/10.1109/79.998080

[10] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, pp. 1–15, 1994. [Online]. Available: http://dx.doi.org/10.1109/90.282603

[11] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1998.

[12] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994. [Online]. Available: citeseer.ist.psu.edu/donoho93ideal.html

[13] K. G. Kyriakopoulos and D. J. Parish, "A live system for wavelet compression of high speed computer network measurements," in *Passive and Active Measurements*, ser. Lecture Notes in Computer Science. Springer, April 2007.

[14] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 1, pp. 3445–3462, 1993. [Online]. Available: http://dx.doi.org/10.1109/78.258085

[15] G. Iannaccone, C. Diot, D. McAulley, A. Moore, I. Pratt, and L. Rizzo, "The como white paper," INTEL, Tech. Rep., 2004, page last visited 22/08/05. [Online]. Available: http://www.cambridge.intel-research.net/como/pubs/como.whitepaper.pdf