



This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



CC creative commons  
COMMONS DEED

**Attribution-NonCommercial-NoDerivs 2.5**

**You are free:**

- to copy, distribute, display, and perform the work

**Under the following conditions:**

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.

 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

**Your fair use and other rights are in no way affected by the above.**

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:  
<https://creativecommons.org/licenses/by-nc-nd/2.5/>

# Applying Wavelets for the Controlled Compression of Communication Network Measurements

Konstantinos G. Kyriakopoulos, David J. Parish

## Abstract

Monitoring and measuring various metrics of high-speed networks produces a vast amount of information over a long period of time making the storage of the metrics a serious issue. Previous work has suggested stream aware compression algorithms, among others, i.e. methodologies that try to organise the network packets in a compact way in order to occupy less storage. However, these methods do not reduce the redundancy in the stream information. Lossy compression becomes an attractive solution, as higher compression ratios can be achieved. However, the important and significant elements of the original data need to be preserved. This work proposes the use of a lossy wavelet compression mechanism that preserves crucial statistical and visual characteristics of the examined computer network measurements and provides significant compression against the original file sizes.

To the best of our knowledge, the authors are the first to suggest and implement a wavelet analysis technique for compressing computer network measurements. In this paper, wavelet analysis is used and compared against the Gzip and Bzip2 tools for data rate and delay measurements. In addition this paper provides a comparison of eight different wavelets with respect to the compression ratio, the preservation of the scaling behavior, of the long range dependence, of mean and standard deviation and of the general reconstruction quality. The results show that the Haar wavelet provides higher peak signal-to-noise ratio (PSNR) values and better overall results, than other wavelets with more vanishing moments. Our proposed methodology has been implemented on an on-line based measurement platform and compressed data traffic generated from a live network.

## Index Terms

Wavelet transforms, data compression, quality control, communication systems, measurements

## I. INTRODUCTION

**T**HE performance monitoring and measurement of communication networks is an increasingly important area as demands on and threats to such networks increase. Monitored network data allows network managers and operators to gain valuable insight into the health and status of a network, and if interpreted correctly, can assist in planning upgrades and remedial action to keep the network operating in a near optimum manner.

Whilst such data is useful for real-time analysis, there is often a need to post-process historical network performance data. Such analysis is useful if ongoing long-term problems or if more detailed analysis of a previous situation needs investigation. Storage of the monitored data then becomes a serious issue as network monitoring activities generate significant quantities of data over a long period of time.

The authors are with the Department of Electronic and Electrical Engineering, Loughborough University, Loughborough, LE11 3TU, U.K. e-mail: {elkk, d.j.parish}@lboro.ac.uk.

Manuscript received Jan, 2009.

Storing all this network performance information is practically inefficient as a large amount of storage capacity is required. A possible solution to this problem is to reduce the amount of data or, in other words, apply compression. In the case of lossless compression, the achieved compression ratios are not satisfactory, making the choice of lossy compression mandatory.

The work described in this paper examines compression schemes that offer compression ratios which are driven from the need to preserve the quality and statistical characteristics of the decompressed measurement signal. Such characteristics are, for example, the PSNR, the scaling behavior and the long range dependency. The measurement signal could be data rate, delay or any other time series signal generated from packet data. Thus, the use of the Wavelet analysis technique offers a variable compression ratio between the original and the reconstructed signals depending on the examined signal's statistical characteristics rather than a specified amount of storage capacity.

The justification of using wavelets instead of other transform analysis is the following. 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 [1] [2].

Furthermore, 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 patterns are statistically self-similar. Self-similar processes are by definition non-stationary [3] [4]. Wavelets have been widely applied to non computer network areas such as image and speech denoising and compression [5] [6] [7].

To date, in the area of computer networks, 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 round trip time (RTT) through the examination of the energy function of the detail coefficients [8]. They have also been used for de-noising one-way delay signals in order to detect shared congestion between different flows [9]. [10] shows that wavelet filters are quite effective at exposing the details and characteristics of ambient and anomalous traffic. [11] analyzes 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. In [12], [3] authors analyse and examine the scaling behaviour of network traffic using wavelets. Regarding compression techniques in the computer networks area some notable work is the following.

[13] describes an *off-line* methodology complimenting to sampling approaches for the reduction of data trace sizes while preserving mean, standard deviation and temporal structures (such as long range dependency and scaling behavior). The proposed methodology is appropriate for network traffic modelling purposes rather than network traffic monitoring. This method does not preserve the actual time based information of the measurements (i.e. it does not preserve the specific data rate and time information) but preserves the above mentioned statistics for the whole measurement set.

[14] tries to achieve compaction of the network packet data by sorting IP traffic as flow traces instead of packet traces. The proposed technique stores the IP data in a compact manner by creating per flow records for common characteristics and then per packet records for per packet information. One of the drawbacks of this work is that the presence of SYN network attacks

and port scans could prohibit the capture board to operate efficiently under these critical scenarios. Furthermore, such network attacks would dramatically increase the number of flows, which would result in increase of the storage required for the flows, thus decreasing the data reduction.

[15] describes a lossless tool, named IPzip, which has two modes of operation. Firstly, it provides an on-line algorithm for compressing data packets in real-time for reducing bandwidth load and secondly, a mechanism for efficient storage of the network traffic. The key idea of this work is that the algorithms identify the intra and inter-packet correlations and reorder the data in order to exploit the correlations. The actual compression process is performed by the Gzip tool. In the offline mode, responsible for storage purposes, IPzip offers 20% and 15% better results than Gzip alone for payload and header information respectively.

[16] proposes the use of wavelets for compressing databases containing counters related to mobile network traffic. Mobile networks and computer networks, though related, are two distinct areas and their traffic exhibits different characteristics. The communication performance in the packet-switching nature of computer networks depends on a much more diverse set of parameters and conditions than in circuit-switching networks where the performance is dependant on the call setup [16].

Previous work by the authors includes a comparison of two known wavelet threshold estimation techniques [17] and two threshold application methods [18] for compressing communication network measurements while controlling the reconstructed signal quality. The authors have shown that the hard threshold is more appropriate for compressing computer network measurements as it gives considerably better quality results [18]. In addition, the advantages of the Haar wavelet against other wavelets in terms of general reconstruction quality (PSNR) and signal energy were presented in [19].

To the best of our knowledge, the authors are the first to suggest and implement a wavelet analysis technique for compressing computer network derived measurements and there is no other literature regarding this specific application of wavelets. In contrast to streaming based compression algorithms [14], [15], which provide lossless compression of IP data by organising the information in a compact manner, our approach compresses the time based measurements provided by the IP packets and removes the inherent redundancy exhibited when storing the actual data packet. The proposed approach introduces the additional benefit of utilising the wavelet analysis for other tasks not just restricting to compression purposes such as event detection etc. as has been described in the related work above.

This paper presents a detailed description of the results produced from compressing communication network measurements with various wavelets of different characteristics with respect to the compression ratio, the scaling behavior, the long range dependence and of the general reconstruction quality in terms of PSNR and percentage relative error of mean and standard deviation.

This paper is divided into seven Sections. In the following Section a theoretical analysis of the wavelet transform is given with a mathematical treatment. In the third Section the characteristics that the proposed compression scheme preserves are presented and discussed. The fourth section presents two methodologies. The first is the proposed methodology for the compression scheme and the second is a methodology for estimating the long range dependence. The fifth section presents the results in terms of compression ratio, signal energy, scaling behavior and long range dependence preservation, of the general visual quality and of the mean and standard deviation of the reconstructed signals. In addition a comparison against dictionary based compression

tools Gzip [20] and Bzip2 [21] is given. The sixth section introduces an implementation of the proposed technique for an on-line system and finally conclusions are given in the last section.

## II. WAVELET ANALYSIS THEORY

### A. Continuous Wavelet Transform

In a similar way to the Continuous Fourier transform, we can define the Continuous Wavelet Transform (CWT) as “the sum over all time of the signal multiplied by scaled, shifted versions of the wavelet function  $\psi$ ”. This sum results in a set of wavelet coefficients that are a function of scale and position [1].

$$C_{scale,shift} = \int_{-\infty}^{\infty} f(t) \times \psi(scale, shift, t) dt \quad (1)$$

In order to produce the CWT coefficients, the following steps take place:

- 1) A wavelet is chosen and is compared against the analysed signal at the starting point.
- 2) A coefficient is generated that indicates how similar the wavelet is with the corresponding part of the examined signal.

The higher the value of the coefficient, the higher the resemblance between the two signals.

- 3) The wavelet is shifted to the right and steps 1 and 2 are performed until the whole signal is examined.
- 4) The wavelet is scaled and steps 1-3 are repeated.
- 5) Steps 1-4 are repeated for each scale.

When we multiply each of the coefficients with the specified scaled and shifted version of the wavelet we produce the constituent wavelet components of the original signal [1].

### B. Discrete Wavelet Transform

In order to make the Continuous Wavelet Transform more efficient but preserve it as accurate, the Discrete Wavelet Transform (DWT) is used. The difference from the CWT is that only scales and positions based on powers of two are used.

DWT maps the information of a time series  $x(t)$  into approximation  $a_x(J, k)$  and detail  $d_x(j, k)$  coefficients in the wavelet domain. In the above expressions,  $j$  is the scale level and takes values from  $[1 \dots J]$ , where 1 is the lowest, most detailed scale, and  $J$  is the crudest;  $k$  is the position of the wavelet in relation to the signal [22].

$$x(t) \rightarrow \{ \{ a_x(J, k) \}, \{ d_x(j, k), j = 1, 2, \dots, J \} \} \quad (2)$$

The approximation coefficients are defined as the inner product of signal  $x(t)$  with shifted and scaled versions of the scaling (a.k.a. father) function  $\phi(t)$ . For this reason the approximation coefficients are also called scaling coefficients. Similarly, the detail coefficients are defined as the inner product of the signal  $x(t)$  with shifted and scaled versions of the wavelet (a.k.a. mother) function  $\psi(t)$ . Thus, the detail coefficients are also referred to as *wavelet* coefficients [22].

$$a_x(j, k) = \langle x | \phi_{j,k} \rangle \quad (3)$$

$$d_x(j, k) = \langle x | \psi_{j,k} \rangle \quad (4)$$

The scaled and shifted versions of the scaling and wavelet functions can be defined as

$$\phi_{j,k}(t) = 2^{-\frac{j}{2}} \phi_0(2^{-j}t - k) \quad (5)$$

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi_0(2^{-j}t - k) \quad (6)$$

The scaling and the wavelet function definitions depend on the choice of wavelet. The simplest wavelet is the Haar wavelet and the scaling and wavelet functions for Haar are given by [16]:

$$\phi(t) = \begin{cases} 1, & \text{if } 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

$$\psi(t) = \begin{cases} 1, & \text{if } 0 \leq t < 1/2 \\ -1, & \text{if } 1/2 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

Wavelets have a mean value of zero and one characteristic property is the number of vanishing moments  $N$ , which can be defined as

$$\int t^k \psi(t) d(t) = 0 \quad k = 0, 1, \dots, N - 1 \quad (9)$$

If the average value of  $\int t^k \psi(t) d(t)$  is zero for  $k = 0, 1, 2, \dots, N - 1$ , then the wavelet has  $N$  vanishing moments and polynomials of degree  $N - 1$  are suppressed by this wavelet. In general a wavelet becomes smoother and more regular as the degree of vanishing moments is increased [1].

The computational complexity of the DWT procedure for an input signal of length  $N = 2^N$  is an operation of the  $O(N)$ , faster than the Fast Fourier Transform, whose computational complexity is  $O(N \log(N))$  [16][23].

### C. Multi-resolution Analysis

Assuming a time series  $x(t)$ , the Multi-resolution analysis at scale  $J$  decomposes the information in signal  $x$  into a collection of a series of detail signals, ranging from the finest (lowest) scale up to scale  $J$  (highest), and a low-resolution approximation signal at scale  $J$  [22].

$$x(t) = \text{approx}_J(t) + \sum_{j=1}^{j=J} \text{detail}_j(t) \quad (10)$$

The approximation signal,  $approx_J(t)$  is generated by convolution of the approximation coefficients (equation 3) and the scaling function  $\phi(t)$ . The approximation signal is the coarsest approximation of signal  $x$  that can occur during the wavelet analysis meaning that the scaling function behaves as a low-pass filter [22].

The collection of detail signals  $\sum_{j=1}^{j=J} detail_j(t)$  is generated similarly by convolution of the detail coefficients (equation 4) with the wavelet function  $\psi(t)$  at each scale. The detail signal at each scale is essentially the difference between two scales  $j$  and  $j + 1$ . In other words, the wavelet function behaves as a band-pass filter, which looks like a wave and therefore the name “wavelet” [22].

Thus, equation 10 can be written as

$$x(t) = \sum_k a_x(J, k) \phi_{J,k}(t) + \sum_{j=1}^{j=J} \sum_k d_x(j, k) \psi_{j,k}(t) \quad (11)$$

### III. GOALS OF LOSSY COMPRESSION

By adjusting the reduction of information in an automatic and dynamic manner, important features in the original time series are preserved, while “expected” features are subjected to higher reduction of information due to their less significant nature. In this way, interesting events that describe the network performance keep their detail and their original characteristics are better preserved. As a result, it is easier for the administrator or a researcher who runs experiments on the network under test, to draw conclusions regarding the effect of their experiments on the network with higher accuracy.

Furthermore, apart from high compression ratios and good reconstruction signal quality, several factors of a compressed signal should be investigated in order to determine the effect of compression on them. The following list shows the goals of compression including statistical and quality aspects that should be preserved.

- 1) Offer high compression ratios
- 2) Preserve Quality characteristics: peaks and dips, and provide high PSNR values
- 3) Preserve Statistical Characteristics: Mean and standard deviation, Energy and Scaling Behavior, Long Range Dependence

#### A. Spikes and overall quality

When times of very high network usage occur, spikes in the data traces are reported. On a public network, such spikes would be of considerable importance as they indicate periods of high load and their further investigation may indicate if changes to the network are required. On a research network such as JANET Lightpath [24], analysis of the spikes may indicate what impact a particular experiment has had on a given network path. Preservation of these spikes for later analysis is therefore of great importance.

We compared the quality of the reconstruction signals with the original by using the PSNR value calculated from

$$PSNR = 10 * \log \left( \frac{MAX^2}{MSE} \right) \quad (12)$$

where MAX is the maximum value of the original signal and MSE is the mean square error calculated from

$$MSE = \frac{1}{N} \sum_{i=0}^{N-1} |x_i - \bar{x}_i|^2 \quad (13)$$

where  $x_i$  is the  $i^{th}$  sample from the original signal,  $\bar{x}_i$  is the  $i^{th}$  sample of the reconstructed signal and  $N$  is the total number of samples.

### B. Energy and Scaling Behavior

The *energy function plot* was introduced by [12], for determining the range of scales of a process that exhibit self-similarity and for estimating the Hurst parameter for stationary or stationary-increment data. They used a statistic known as the *energy function*  $E_j$  which indicates the average energy of the arrival process contained at scale  $j$  and is defined by [12]:

$$E_j = \frac{1}{N_j} \sum_k |d_{j,k}|^2, \quad j = 1, 2, \dots, n \quad (14)$$

where  $N_j$  is the number of wavelet (“detail”) coefficients at scale  $j$ .

The energy plot can be generated by plotting  $\log(E_j)$  against the scale  $j$  from finer to coarser scales. Intuitively, this plot illustrates the scaling behaviour of the underlying time series (such as a traffic arrival process) at different time-scales.

This tool has been used by many researchers [25], [26], [27], [28] for evaluating their traffic models with respect to the correct reproduction of the scaling structure of the modeled traffic. By comparing the burstiness of the synthesized and original traffic at a variety of scales, researchers can evaluate how closely their model matches the correlation structure of the modeled network.

### C. Long Range Dependence

Leland et al., with their well-known publication in 1994 [4], showed that Ethernet LAN traffic does not follow the Poisson model as was commonly assumed. They proved that Ethernet LAN traffic is self-similar i.e. the network traffic is bursty across a wide range of time scales.

Self-similarity can be described as the property of a data set to look or behave the same when viewed at different time or space scales [29]. Some self-similar processes are given in [30]. Self-similarity of network traffic depends on the network utilization level and can be described by the Hurst parameter [4].

The long-range dependence (LRD) is one of the properties that self-similar processes have and it is also described by the Hurst parameter [30]. It is a way to measure the memory of a process or, in other words, how correlated distant events of a process are.

Many studies have concluded that LRD significantly increases queuing delays and leads to packet loss [31]. Furthermore, LRD and self-similarity are important concepts for network modeling, traffic synthesis, and queue analysis [32], [4]. The above issues have led the network community to become involved in identifying the Hurst parameter and consider it as a key concept in signal analysis. Thus, its preservation after applying any compression is an important issue.

## IV. METHODOLOGY

### A. Compression scheme

Wavelet analysis is not a compression tool but a transformation to a domain that provides a different view of the data that is more suitable to compression than the original data itself.



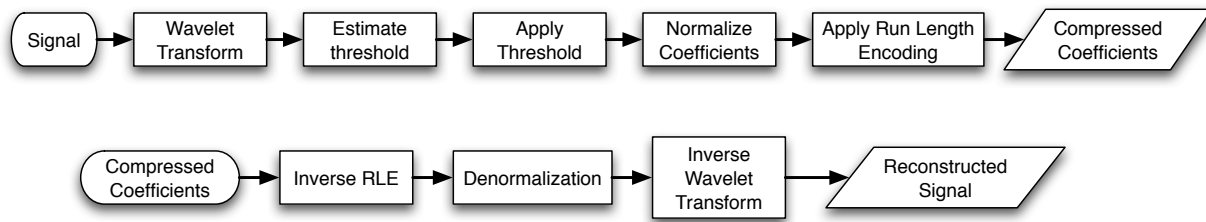


Fig. 1. Operation flow chart

Fig. 1 shows the flow chart of the algorithm used for analysis and compression of each examined signal and the reverse process for reconstructing each signal. First the signal is transformed into the wavelet domain where a threshold is calculated and applied to the wavelet coefficients. Then, the thresholded coefficients are normalized so that each coefficient can be stored in 1 byte and finally run length encoding is applied. It should be noted that compression occurs by taking advantage of sequential thresholded coefficients using the run length encoding and not because of the thresholded coefficients themselves.

Many of the wavelet coefficients produced from the wavelet analysis have an absolute value close to zero. These small coefficients are likely to attribute only small variations of the signal and contain a small percentage of the signal's total energy. These small coefficients can be discarded without a significant loss in the quality of the signal and more importantly of the interesting features. Thus, a threshold is required below which all coefficients will be discarded.

[33] proposed an adaptive thresholding technique (referred to in this paper as GK technique) that is calculated from the absolute value of the wavelet coefficients. This scheme is not based on signal denoising but rather tries to statistically identify significant coefficients. Specifically, the standard deviation ( $\sigma$ ) and mean ( $\mu$ ) of the absolute value of non-zero detail coefficients are first calculated. If the standard deviation is larger than the mean, then the threshold is set to two times the mean ( $2 * \mu$ ), otherwise it is equal to the mean minus the standard deviation ( $\mu - \sigma$ ).

$$T = \begin{cases} 2 * \mu, & \text{if } \sigma > \mu \\ \mu - \sigma, & \text{if } \sigma \leq \mu \end{cases} \quad (15)$$

In order to estimate the percentage of preserved energy after applying the proposed threshold (as discussed above), the approximation and detail coefficients need to be preserved and the flow chart in Fig. 2 was followed. The percentage of kept energy for each signal was calculated by the  $L^2$ -norm recovery criterion [1] given in percentage terms by:

$$PERFL2 = 100 * \left( \frac{l^2\text{-norm}(CXC)}{l^2\text{-norm}(C)} \right) \quad (16)$$

where  $CXC$  are the coefficients of the thresholded decomposition and  $C$  the coefficients of the original decomposition.

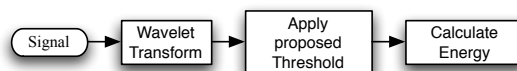


Fig. 2. Flow chart for Energy calculation in thresholded coefficients

To evaluate this algorithm, a series of off-line experiments using thirty delay and thirty data rate signals of 1024 points

each was conducted. The delay signals were measured on a research test bed network. 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 [34]. The data rate signals are from a real commercial network that generates around 25 TB of data per day and has data rates between 500 Mbps and 1.4 Gbps. In a window size of 2 seconds 30000 unique IP addresses may be observed in this network [35].

### B. Estimating the Hurst parameter

A variation of the classic Rescaled Range ( $R/S$ ) statistic block algorithm was used as a Hurst estimation algorithm. The  $R/S$  algorithm flowchart is presented in Fig. 3. The algorithm involves dividing the whole data set of  $n$  samples into  $\lfloor n/N \rfloor$  non-overlapping sample blocks; where  $1 \leq N \leq n$  is the sample block size. The calculations for the rescaled range value presented in Fig. 3 apply recursively to each sample block and ultimately produce  $\lfloor n/N \rfloor$  intermediate  $R/S$  values. These intermediate values are summed and divided by the number of sample blocks in order to find the average  $R/S$  value for the current sample block size. The whole process repeats for all values that the sample block size can take ( $1 \leq N \leq n$ ) [4].

One of the steps of the algorithm involves calculating the quantity  $W_k$  where

$$W_k = (X_1 + X_2 + \dots + X_k) - k\bar{X}(n) \quad (17)$$

and  $k = 1, 2, 3, 4, \dots, n$

For the regression analysis the edge values should not be considered [30], [36]. Thus, in order to exclude the edge values, a window that includes 90% of the data trace values was utilized. The window slides through the data trace and for each position of the window a regression line is calculated. The Hurst parameter is estimated based on the gradient of the regression line with the best goodness of fit value.

## V. RESULTS

### A. Wavelets and Compression Performance

The mother-wavelet chosen to analyze the network measurement signals is of prime importance as some wavelets offer better reconstruction quality and different compression ratios than others. However, there is no wavelet that gives the best results for all kinds of signals.

Eight wavelets were chosen and compared against each other in order to find out which one offers better reconstruction results after thresholding with the GK technique and hard threshold as described in section IV-A. In all experimental results presented, the decomposition was performed up to level 6 for both data rate and delay signals. Level 6 was chosen because previous work by the authors has observed that compression ratio stabilises at Level 6 of decomposition. Higher decomposition levels do not offer any significant increase in compression. Furthermore, Level 6 offers a reasonable balance between compression ratio and preservation of signal quality for the examined signals [17].

The following wavelets were compared: Haar, Meyer, Biorthogonal 3.9 and Daubechies D4, D6, D8, D10, D12. The index of the Daubechies wavelets indicates the number of coefficients. The number of vanishing moments each Daubechies wavelet has is half of the number of coefficients, i.e. D4 has 2 vanishing moments.

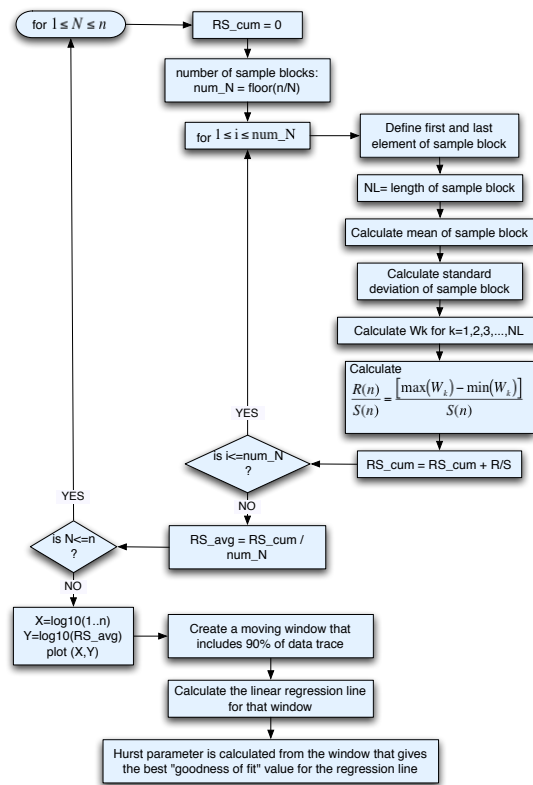


Fig. 3. Rescaled Range Statistic flow chart

Regarding speech and audio signals; [2], [37], among others have found that for better reconstruction quality, wavelets with many vanishing moments should be utilized, as they introduce less distortion and concentrate more signal energy in the approximation coefficients. Wavelets with many vanishing moments are described with many coefficients in the scaling and wavelet functions, thus increasing the computation overhead of the wavelet transform, the complexity of the algorithm and most importantly the output file size.

Tables I and II show the average compression ratio results for each of the above wavelets for the thirty delay and data rate signals described in section IV-A. The Haar wavelet with the least vanishing moments gives the best results for delay signals and has the third place for data rate signals. Even though Haar does not have first place for data rate signals, its compression ratio is very close to D6's that is in first place.

To further examine the performance of the Haar wavelet compression the authors compared the Haar analysis at level 6 of the decomposition compression results against two dictionary based lossless tools, Bzip2 and Gzip. An overall visual overview can be seen in Fig. 4 for all 30 delay signals and Fig. 5 for all 30 data rate signals.

Tables III and IV show the percentage reduction achieved with the Haar wavelet at level 6 of decomposition against the original file size and the Gzip and Bzip2 compressed versions of the original file. On average, for delay signals the original file is reduced by 93.3 %, which is 7.5 (84.9% reduced) and 6.4 (82.4% reduced) times better than Gzip and Bzip2 could achieve respectively. For data rate signals, on average, the original file is reduced by 90.8 %, which is 5.4 (80.8% reduced) and 4.8 (78.3% reduced) times better than Gzip and Bzip2 could achieve respectively.

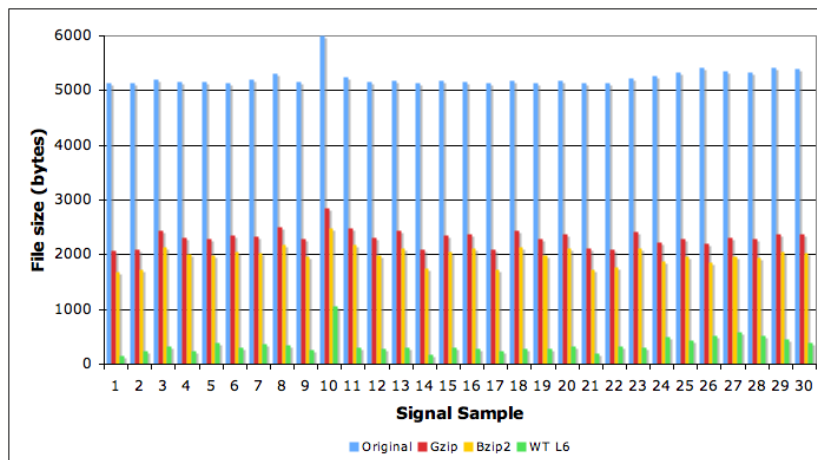


Fig. 4. Comparison of the original file size against the compressed Gzip, Bzip2 and the proposed algorithm at Layer 6 of decomposition for 30 delay signals

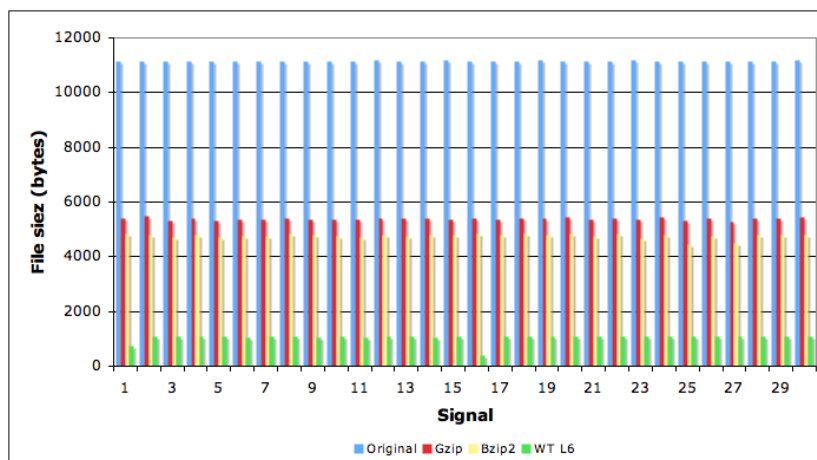


Fig. 5. Comparison of the original file size against the compressed Gzip, Bzip2 and the proposed algorithm at Layer 6 of decomposition for 30 data rate signals

### B. Wavelet Energy Attributes

The reconstruction quality depends largely on the approximation coefficients but also, less significantly, on the detail coefficients preserved after applying the suggested threshold. The signals described in section IV-A were used as inputs in the flow chart described in Fig. 2 and all eight wavelets were examined one by one.

After applying the proposed threshold, the Haar wavelet ranks third for delay signals and second for data rate signals in preserving the reconstructed signal's energy. In both cases, with the Haar wavelet, the average preserved energy is very high; 98.44 % for delay and 99.99 % for data rate signals [19].

### C. Wavelets and Scaling Behavior

In this section an energy plot (see section III-B) was generated for all 30 delay and data rate signals using *all* wavelets. This was done to examine how closely different wavelets preserve the scaling behavior of the original signals.

In order to quantify the findings, Tables V and VI show the average relative percentage error of energy per scale for the 30 delay and 30 data rate signals respectively. In most cases (L1, L3, L4 for Table V and L1, L3, L4, L5, L6 for Table VI ) the

*Haar* wavelet gives the smallest error per scale in the energy plots.

As a representative sample of most examined signals, Fig. 6 shows that the scaling behaviour between the original and compressed (reconstructed) signals is very similar when using the Haar wavelet. In Fig. 7 and 8 are shown the worst case scenarios for delay and data rate signals respectively. Table VII shows the relative percentage error of energy per scale for delay signal 26 and data rate signal 16 (worst cases).

The reconstruction very closely preserves the scaling characteristics of both types of signals and even in the worst case scenarios, it offers better results than some advanced network traffic generation models. As a comparison, reference [25] provides a model with the aim of reproducing the scaling nature of real traffic. Even though the authors do not provide numerical results, it can be seen from the figures that the error in some cases is several dB, much greater than that in figures 8 and 7 (worst cases) where the error is less than 0.5 dB.

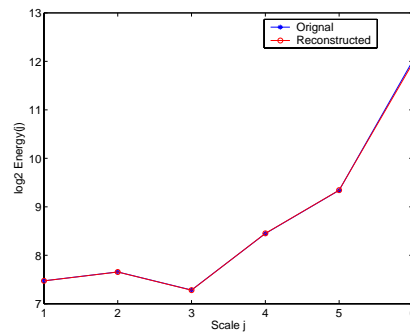


Fig. 6. Energy scaling behavior for data rate signal 22 with Haar analysis

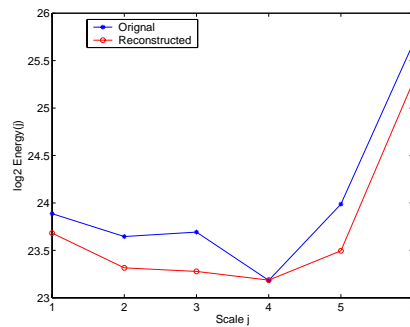


Fig. 7. Energy scaling behavior for delay signal 26 with Haar analysis

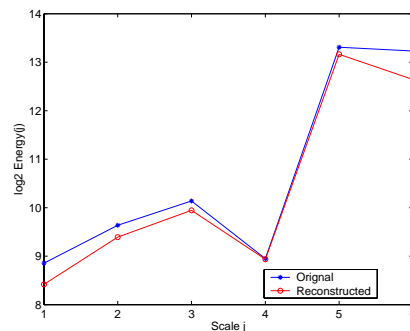


Fig. 8. Energy scaling behavior for data rate signal 16 with Haar analysis

Wavelet	Haar	D4	D6	D8	D10	D12	Meyer	Bio 3.9
Avg.	16.98	16.29	16.15	15.44	14.78	14.29	10.26	15.02
Min.	5.66	5.44	5.34	5.24	5.14	5.04	4.28	5.18
Max.	32.44	32.65	29.63	27.41	25.89	23.09	16.7	24.29
Rank	1	2	3	4	6	7	8	5

TABLE I  
COMPRESSION RATIO FOR 30 DELAY SIGNALS AT LEVEL 6 WITH VARIOUS WAVELETS

Wavelet	Haar	D4	D6	D8	D10	D12	Meyer	Bio 3.9
Avg.	11.20	11.23	11.33	10.84	10.53	10.34	8.45	10.47
Min.	10.32	10.08	9.87	9.66	9.44	9.23	7.68	9.30
Max.	27.43	33.44	33.54	34.48	32.56	31.37	21.88	33.44
Rank	3	2	1	4	5	7	8	6

TABLE II  
COMPRESSION RATIO FOR 30 DATA RATE SIGNALS AT LEVEL 6 WITH VARIOUS WAVELETS

	Original	Gzip	Bzip2
Min.	82.33	62.81	57.34
Average	93.32	84.89	82.42
Max.	96.92	92.34	90.58

TABLE III  
MIN., AVERAGE, AND MAX. PERCENTAGE REDUCTION ACHIEVED WITH THE HAAR WAVELET AT LEVEL 6 OF DECOMPOSITION AGAINST THE ORIGINAL, GZIP AND BZIP2 FILE SIZES FOR ALL 30 DELAY SIGNALS

	Original	Gzip	Bzip2
Min.	90.31	79.77	75.80
Average	90.75	80.83	78.28
Max.	96.35	92.50	91.57

TABLE IV  
MIN., AVERAGE, AND MAX. PERCENTAGE REDUCTION ACHIEVED WITH THE HAAR WAVELET AT LEVEL 6 OF DECOMPOSITION AGAINST THE ORIGINAL, GZIP AND BZIP2 FILE SIZES FOR ALL 30 DATA RATE SIGNALS

Wavelet	L1	L2	L3	L4	L5	L6
Haar	<b>3.860</b>	6.047	<b>7.355</b>	<b>7.573</b>	13.300	12.364
db4	5.241	6.848	9.883	12.588	13.231	12.793
db6	6.011	8.161	10.224	14.442	14.096	11.546
db8	6.011	7.790	10.004	14.609	15.570	11.821
db10	7.087	7.970	11.499	12.865	11.815	9.722
db12	7.040	8.289	11.005	15.303	14.037	11.175
dmey	6.557	6.704	12.609	13.549	12.768	<b>7.472</b>
bior3.9	7.070	<b>4.967</b>	9.056	12.179	<b>11.671</b>	7.763

TABLE V  
AVERAGE RELATIVE PERCENTAGE ERROR OF ENERGY PER SCALE FOR 30 DELAY SIGNALS. BOLD VALUES INDICATE THE MINIMUM VALUE PER COLUMN

#### D. Long Range Dependence Preservation

The Hurst parameters of the original and reconstructed data rate signals were estimated by following the algorithm described in the previous section and presented in Figure 3. Only data rate signals were examined as data rate expresses a traffic arrival process and long range dependence is usually examined in arrival processes.

Fig. 9 shows the relative percentage error of the reconstructed Hurst values for all 30 data rate signals. The average percentage error is 0.146%. From the graph it can be seen that the maximum error is less than 5% (4.185%) and occurs when examining signal 16.

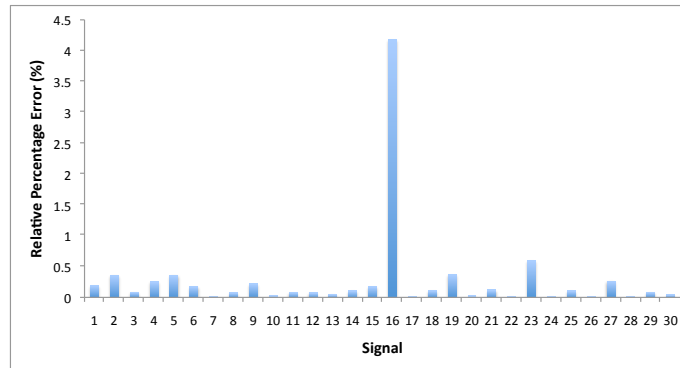


Fig. 9. Relative percentage error of the reconstructed Hurst values for 30 data rate signals

#### E. Wavelet Quality Attributes

For the task of examining which wavelet offers best reconstruction results, two stages were followed; decomposition and reconstruction (see Fig. 1). In order to measure the quality of the reconstructed signals with different wavelets, the inverse wavelet transform was applied on the produced coefficients after thresholding.

The quality of the reconstruction signal was compared with the original by using the PSNR value as described in equations 12 and 13. Tables VIII and IX show the average PSNR value after reconstruction at level 6 for the thirty delay and data rate signals respectively. Tables X and XI show the average relative percentage error between the original and the reconstructed delay and data rate signals for the mean and standard deviation respectively.

As can be inferred from the tables, wavelets with more vanishing moments do not provide higher PSNR values neither lower errors for mean and standard deviation for the reconstructed signals. The Haar wavelet ranks first on average in both delay and data rate signals in terms of quality of results.

There is a trade-off between thresholded coefficients and compression. The more coefficients are thresholded and removed the more likely it is that the reconstructed signal will lose visual characteristics and statistical values. However, for the compression to take effect, the thresholded coefficients need to be sequential in order to be utilised by the run length encoding.

The Haar wavelet achieves better results in preserving the reconstruction quality for the network measurements considered than other wavelets because it itself is more similar to the examined signals. Therefore, the Haar wavelet coefficients have higher values (as discussed in section II-A. For this reason, when using the Haar wavelet less coefficients are thresholded (relatively to other wavelets) during the compression process, which leads to higher reconstruction quality.

Wavelet	L1	L2	L3	L4	L5	L6
Haar	<b>1.024</b>	1.748	<b>1.024</b>	<b>0.592</b>	<b>1.046</b>	<b>2.219</b>
db4	2.202	<b>1.077</b>	1.356	1.563	6.572	5.141
db6	2.032	2.367	2.457	2.238	2.844	5.971
db8	1.474	2.701	1.977	1.062	4.053	9.932
db10	1.385	2.158	1.909	2.224	4.325	6.997
db12	1.162	1.693	2.021	3.064	3.254	6.013
dmey	1.135	1.167	1.784	1.809	1.164	4.187
bior3.9	1.485	2.225	1.492	2.174	5.297	8.500

TABLE VI

AVERAGE RELATIVE PERCENTAGE ERROR OF ENERGY PER SCALE FOR 30 DATA RATE SIGNALS. BOLD VALUES INDICATE THE MINIMUM VALUE PER COLUMN

Wavelet	L1	L2	L3	L4	L5	L6
Delay signal 26	13.237	20.464	24.993	0.217	28.89	24.353
Data rate signal 16	26.119	15.525	12.53	0.948	9.671	33.504

TABLE VII

RELATIVE PERCENTAGE ERROR OF ENERGY PER SCALE FOR FOR DELAY SIGNAL 26 AND DATA RATE SIGNAL 16 [WORST CASES]

Wavelet	Haar	D4	D6	D8	D10	D12	Meyer	Bio 3.9
Avg. PSNR	39.6	38.2	37.6	37.5	37	37	37.1	37.3
Rank	1	2	3	4	7	8	6	5

TABLE VIII

PSNR VALUES FOR DELAY SIGNALS AFTER RECONSTRUCTION

Wavelet	Haar	D4	D6	D8	D10	D12	Meyer	Bio 3.9
Avg. PSNR	53.4	51.6	51.7	49.7	50.8	51.5	53.0	50.4
Rank	1	3	4	8	6	5	2	7

TABLE IX

PSNR VALUES FOR DATA RATE SIGNALS AFTER RECONSTRUCTION

	Haar	D4	D6	D8	D10	D12	Meyer	Bio 3.9
Delay	0.03	0.08	0.11	0.16	0.15	0.15	0.10	0.11
Data Rate	0.02	0.03	0.03	0.03	0.04	0.04	0.03	0.04

TABLE X

AVERAGE RELATIVE PERCENTAGE ERROR OF MEAN VALUES FOR DELAY AND DATA RATE SIGNALS AFTER RECONSTRUCTION

	Haar	D4	D6	D8	D10	D12	Meyer	Bio 3.9
Delay	2.57	3.23	3.72	3.82	4.00	4.24	3.60	1.50
Data Rate	0.16	0.20	0.17	0.18	0.18	0.18	0.34	0.16

TABLE XI

AVERAGE RELATIVE PERCENTAGE ERROR OF STANDARD DEVIATION VALUES FOR DELAY AND DATA RATE SIGNALS AFTER RECONSTRUCTION



On the other hand, wavelets with many vanishing moments may have more thresholded coefficients but this is not a condition to offer higher compression as the thresholded coefficients are sporadic and not sequential so cannot take advantage of the run length encoding step. Besides, wavelets with many vanishing moments increase the number of produced coefficients making it more difficult to achieve high compression ratios.

Conclusively, the Haar wavelet is the simplest wavelet algorithm that can give perfect reconstruction (when no coefficients are removed) and has the following additional advantages [38] [39]: It is conceptually simple, fast, memory efficient, offers higher compression than other wavelets and is exactly reversible without producing edge effects.

## VI. ON-LINE EXPERIMENTS

The proposed algorithm (Fig. 1) has been implemented by the authors as a module for CoMo which is a passive monitoring platform developed for the purpose of monitoring network links at high speeds and replying to real-time queries regarding network statistics [40]. CoMo has various modules that calculate one or more network measurements. There are some built-in modules that come with CoMo but it can also accept third-party modules. The proposed algorithm can be imbedded in the modules and compress the calculated measurements. In this way, CoMo stores the output of the run length encoding step rather than the original signal itself. When CoMo receives a query, the information is first decompressed and then send to the end user [19].

CoMo is responsible for capturing the arriving packets at the monitoring node. Depending on the task of each module, information from each packet (for example packet length) are aggregated and binned per second thus producing measurements per second, for example data rate (bytes/sec). Therefore, the on-line analysis and compression processes are independent of the bandwidth of the monitored link. Regardless of the bandwidth of the monitored link, a few milliseconds of processing time are adequate for the algorithm to perform an on-line analysis and compression of a block of measurements. This means that the implementation of our proposed compression methodology can scale on high data rate links.

The on-line compression mechanism has been evaluated using traffic from a research group's live network. Results are similar to those presented in the section covering off-line experiments. Typical results for a 17 minute data rate sample set show errors in the reconstructed mean of -0.5%; of -1% in the standard deviation; and of -1.9% for the energy.

## VII. CONCLUSIONS

This paper has presented a solution to the problem of compressing large amounts of time based measurements taken from high speed computer networks. The compression methodology, even though lossy, preserves the statistical and visual characteristics of the examined signal measurements. In this paper, eight wavelets have been compared for four different characteristics: compression ratio, energy and scaling behavior, long range dependence preservation and quality of reconstruction. Due to the fact that each wavelet has its own characteristics and different wavelets approximate parts of the analyzed signal with different degree of accuracy, the results vary with each wavelet.

The Haar wavelet, however, shows a consistency in better describing the analyzed signals and on average has the best performance with respect to both compression and reconstruction quality. Other advantages of the Haar wavelet are its conceptual

simplicity, its fastness and memory efficiency and that it does not produce artifacts in the reconstructed signal. In contrast to wavelets with more vanishing moments, the Haar wavelet also produces less coefficients and thus smaller output file sizes, which increase its compression capabilities. The Haar wavelet ranks first on average in both delay and data rate signals in terms of quality of results (PSNR, mean and standard deviation).

On average, for delay signals the original file is reduced by 93.3 % and for data rate signals the original file is reduced by 90.8 %. Using the Haar wavelet, the averaged preserved energy is very high; 98.4 % for delay and 99.9 % for data rate signals. Regarding the scaling behavior preservation, the proposed mechanism offers better results than some advanced network traffic generation models, giving an error of less than 0.5 dB in the worst case.

In this paper the rescaled adjusted range statistic algorithm has been described, which is used for the estimation of the Hurst parameter. The Hurst value itself is of no direct concern in this research. However, the error in the Hurst parameter after analyzing and compressing each signal is of concern as this value should be preserved. The experiments run on data rate signals have a maximum relative percentage error of 4.2%. Finally, the proposed algorithm has been implemented by the authors as a module for an on-line passive monitoring tool.

#### REFERENCES

- [1] M. Misiti, Y. Misiti, G. Oppenheim, and J. Poggi, "Matlab wavelet toolbox," Tech. Rep., The MathWorks, Inc., 1997-2004.
- [2] Johnson Ihyeh Agbinya, "Discrete wavelet transform techniques in speech processing," in *Proceedings of the 1996 IEEE Region 10 TENCON - Digital Signal Processing Applications Conference*, Perth, Aust, Nov 26-29 1996, vol. 2, pp. 514–519, IEEE.
- [3] Patrice Abry, Richard Baraniuk, Patrick Flandrin, Rudolf Riedi, and Darryl Veitch, "Multiscale nature of network traffic," *IEEE Signal Processing Magazine*, vol. 19, no. 3, pp. 28–46, May 2002.
- [4] Will E. Leland, Murad S. Taqqu, Walter Willinger, and Daniel V. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, pp. 1–15, 1994.
- [5] Ralph Algazi and Jr. Robert R. Estes, "Analysis based coding of image transform and subband coefficients," *Proceedings of the SPIE: In Applications of Digital Image Processing*, vol. 2564, no. XVII, pp. 11–21, 1995.
- [6] S. Grace Chang, Bin Yu, and Martin Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532 – 1546, 2000.
- [7] Byung-Jun Yoon and P. P. Vaidyanathan, "Wavelet-based denoising by customized thresholding," in *International Conference on Acoustics, Speech and Signal Processing*, 2004, vol. 2, pp. 925–928.
- [8] Polly Huang, Anja Feldmann, and Walter 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.
- [9] 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*, 30 August through 3 September 2004, vol. 34, pp. 293–305.
- [10] P. Barford, J. Kline, D. Plonka, and A. Ron, "A signal analysis of network traffic anomalies," in *Proceedings of the 2nd Internet Measurement Workshop (IMW 2002)*, 2002, pp. 71–82.
- [11] Seong Soo Kim, A. L. Narasimha Reddy, and Marina Vannucci, "Detecting traffic anomalies through aggregate analysis of packet header data," in *NETWORKING*, Nikolas Mitrou, Kimon P. Kontovasilis, George N. Rouskas, Ilias Iliadis, and Lazaros F. Merakos, Eds. 2004, vol. 3042 of *Lecture Notes in Computer Science*, pp. 1047–1059, Springer.
- [12] Patrice Abry and Darryl Veitch, "Wavelet analysis of long-range-dependent traffic," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 2–15, January 1998.
- [13] Antonio Pescapè, "Entropy-based reduction of traffic data," *Communications Letters*, vol. 11, no. 2, pp. 3, February 2007.
- [14] Gianluca Iannaccone, Christophe Diot, Ian Graham, and Nick McKeown, "Monitoring very high speed links," in *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*, New York, NY, USA, 2001, pp. 267–271, ACM.

- [15] Antonio Nucci Su Chen, Supranamaya Ranjan, "Ipzip: A stream-aware ip compression algorithm," in *Data Compression Conference*. 2008, pp. 182–191, IEEE.
- [16] Claudio A. Ardagna, Elisa Bernardoni, Ernesto Damiani, and Salvatore Reale, "Mobile network traffic data compression by means of wavelet decomposition," in *2nd IEEE International Conference on Digital Ecosystems and Technologies*. 2008, pp. 274–280, IEEE.
- [17] Konstantinos G. Kyriakopoulos and David J. Parish, "Wavelet compression techniques for computer network measurements," in *Proceedings of the Fourth IASTED International Conference on Signal Processing, Pattern Recognition, and Applications 2007 (SPPRA 2007)*, Innsbruck, Austria, 14-16 February 2007, IASTED, pp. 109–115, Acta Press.
- [18] Konstantinos G. Kyriakopoulos and David J. Parish, "Wavelet compression of network delay measurements," in *Proceedings of the 3rd IADAT International Conference on Telecommunications and Computer Networks (IADAT-tn 2006)*, Portsmouth, U.K., 27-29 September 2006, pp. 115–119, IADAT.
- [19] Konstantinos G. Kyriakopoulos and David J. Parish, "A system for on-line compression of high speed network measurements," *International Journal of Internet Protocol Technology*, vol. 3, no. 2, pp. 95 –106, 2008.
- [20] "The gzip home page," Website. <http://www.gzip.org/>.
- [21] "The bzip2 home page," Website. <http://www.bzip.org/>.
- [22] Konstantinos G. Kyriakopoulos, *Wavelet Analysis for Compression and Feature Extraction of Network Performance Measurements*, Ph.D. thesis, Loughborough University, 2008.
- [23] Wayne O'Brian Jarrett, *Congestion Detection within Multi-Service TCP/IP Networks using Wavelets*, Ph.D. thesis, University of London, 2004.
- [24] "Janet lightpath," Website, 2008, Available on <http://www.ja.net/services/lightpath/index.html>. Page visited on 15/1/2009.
- [25] Kashi Venkatesh Vishwanath and Amin Vahdat, "Realistic and responsive network traffic generation," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 111–122, 2006.
- [26] Joel Sommers and Paul Barford, "Self-configuring network traffic generation," in *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, New York, NY, USA, 2004, pp. 68–81, ACM Press.
- [27] J. Ridoux, A. Nucci, and D. Veitch, "Seeing the difference in IP traffic: Wireless versus wireline," in *Proceedings of IEEE INFOCOM 2006*, Barcelona, Spain, April 2006.
- [28] C. Rolland, J. Ridoux, and B. Baynat, "Litgen, a lightweight traffic generator: application to p2p and mail wireless traffic," in *Proceedings of the Passive and Active Measurement Conference (PAM 2007)*, Louvain-La-Neuve, Belgium, April 2007, vol. 4427 of *LNCSS*, pp. 52–62, Springer.
- [29] Arun Avudainayagam, Prasanna Krishnamoorthy, and Krithi Rao, "Jitter analysis in atm networks handling self-similar traffic," Website, Available on <http://members.tripod.com/aron-10/selfsim/selfsim.htm>. Page last visited 15/01/2009.
- [30] O. Rose, "Estimation of the hurst parameter of long-range dependent time series," Tech. Rep., University of Würzburg, Institute of Computer Science, February 1996, Research Report Series 137.
- [31] Ashok Erramilli, Onuttom Narayan, and Walter Willinger, "Experimental queueing analysis with long-range dependent packet traffic," *IEEE/ACM Transactions on Networking*, vol. 4, no. 2, pp. 209–223, 1996.
- [32] Vinay J. Ribeiro, Rudolf H. Riedi, and Richard Baraniuk, "Wavelets and multifractals for network traffic modeling and inference," in *Proceedings of ICASSP*, 2001.
- [33] Savita Gupta and Lakhwinder Kaur, "Wavelet based image compression using daubechies filters," in *8th National conference on communications, I.I.T.*, 2002.
- [34] Mark J. Sandford, *Detecting Changes in Network Performance from Low Level Measurements*, Ph.D. thesis, Loughborough University, July 2001.
- [35] Peter Sandford and David J. Parish, "Traffic summary analysis for network abuse detection in high date-rate ISP networks traffic summary analysis for network abuse detection in high date-rate isp networks," in *Proceedings of the Sixth International Conference on Networking (ICN'07)*. 2007, IEEE.
- [36] David Nawrocki, "R/S analysis and long term dependence in stock market indices," *Managerial Finance*, vol. 21, no. 7, pp. 78–91, 1995.
- [37] Najih A.M.A.M., Ramli A.R., Ibrahim A., and Syed A.R., "Comparing speech compression using wavelets with other speech compression schemes," in *Proceedings of Student Conference on Research and Development, SCORED*, August 2003, pp. 55–58.
- [38] Yves Nievergelt, *Wavelets Made Easy*, Birkhauser, May 1999.
- [39] Ian Kaplan, "The wavelet lifting scheme," Website, [http://www.bearcave.com/misl/misl\\_tech/wavelets/haar.html](http://www.bearcave.com/misl/misl_tech/wavelets/haar.html). Page last visited on 15/1/2009.
- [40] Gianluca Iannaccone, Christopher Diot, Derek McAulley, Andrew Moore, Ian Pratt, and Luigi Rizzo, "The como white paper," Tech. Rep., INTEL, 2004, Available on <http://como.sourceforge.net/pubs/como.whitepaper.pdf>. Page last visited 15/01/09.