WAVELET SPECTRUM FOR INVESTIGATING STATISTICAL CHARACTERISTICS OF UDP-BASED INTERNET TRAFFIC

Jusak Jusak[1]and Richard J. Harris[2]

[1] STMIK STIKOM Surabaya, Jl. Raya KedungBaruk 98, Surabaya, Indonesia
jusak@stikom.edu
[2] School of Engineering and Advanced Technology, Massey University, Palmerston North, New Zealand
R.Harris@massey.ac.nz

ABSTRACT

In this paper, we consider statistical characteristics of real User Datagram Protocol (UDP) traffic. Four main issues in the study include(i) the presence of long rangedependence (LRD) in the UDP traffic,(ii) the marginal distribution of the UDP traces,(iii) dependence structure of wavelet coefficients,(iv) and performance evaluation of the Hurst parameter estimation based on different numbers of vanishing moments of the mother wavelet. By analyzing a large set of real traffic data, it is evident that the UDP Internet traffic reveals the LRD properties with considerably high non-stationary processes.Furthermore, it exhibits non-Gaussian marginal distributions. However, by increasing the number of vanishing moments,it is impossible to achieve reduction fromLRD to become a short range dependence. Thus, it can be shown that there is no significant difference in performance estimation of the Hurst parameter for different numbers of vanishing moments of the mother wavelet.

KEYWORDS

Long range dependence, User Datagram Protocol, Wavelet, Internet traffic

1. INTRODUCTION

We have witnessed significant proliferation of Internet usage for more than two decades. There is no doubt that since its emergence, the Transmission Control Protocol (TCP) has served as the transport protocol of choice for a large portion of Internet traffic applications [1],[2]. However, increasing demand on multimedia Internet applications with streaming as well as P2P (Point to Point) protocols [3] today has shown that utilization of UDP protocols is growing progressively. For example, a study based on CAIDA traces reported the escalation UDP-based traffic in terms of the number of packets, bytes and flows was approximately double or more between the year of 2002 and 2009 [4]. Our observations on Internet traces over trans-Pacific backbone links (collected by the MAWI[1] working group) reveal that, in 2002, the average rate for UDP-based applications was approximately 1 Mbps over 1 year of observation, while in 2008 the UDP average rate had increased significantly to become 7.19 Mbps. The reports on the growth of the UDP traffic have prompted us to look more carefully at statistical characteristics of traffic traces that employ UDP protocol. The main motivation is that by understanding the statistical behavior of this Internet traffic, it allows us to study the effect of various model parameters on network performance.

[1]http://mawi.wide.ad.jp

DOI : 10.5121/ijcnc.2012.4505

A multitude of references demonstrated the presence of the scale-dependent properties in both local area [5] and wide area traffic [6], as well as in World Wide Web (WWW) traffic [7]. What we mean by the term "scale-dependent properties" is that in these traces, high variability across a wide range of time scales was observed for packets or byte arrivals that range from a few hundreds of milliseconds up to hundreds of seconds. Under such conditions, the traffic looks self-similar and has non-negligible correlations on the arrival counts over a very long time interval. Hence, the term long-range dependence that is largely used in the literature and that will be adopted in this paper. However, we believe that these conclusions were drawn based on the general observation of Internet traffic whereby the TCP protocol has dominated the traces [8].

One of the most popular tool for studying the long-range dependence of Internet traffic is the wavelet-based estimator that was proposed in [9],[10],11]. This estimator is considered as the most successful method among other methods. This is mainly due to the fact that the wavelet spectrum is naturally able to capture self-similar scaling that is present in data that has long-range dependent statistical properties. In addition, under non-bias conditions and for Gaussian processes, the vanishing moments of the mother wavelet successfully convert the long dependence range into becoming a short dependence range. This has been the most important feature of the underlying wavelet estimator.

We confine our study in this paper on the following objectives: (i) To investigate whether the UDP traffic exhibits the long-range dependent (LRD) properties; (ii) To validate the use of the wavelet-based estimator by capturing the marginal distribution of real traces obtained from UDP traffic. The observation will be done for different time aggregations to accommodate the effects of scaling in the data. It should be noted that the key feature of the wavelet-based estimator holds for Gaussian processes. In contrast to this, a recent study suggested that Gamma distribution captures best the marginal distribution of Internet traffic for different aggregation levels [12]; (iii) To investigate the marginal distribution as well as the dependence structure of the wavelet coefficients under conditions where the scaling parameter of the UDP traffic with long-range dependent data needs to be estimated; (iv) To analyze the performance of the Hurst parameter estimation based on different numbers of vanishing moments of the mother wavelet.

The goals of this paper are:(1) to explore the statistical characteristics of Internet traffic traces, in terms of its dependent structure that are generated from UDP traffic only; (2) to study the statistical performance of the wavelet-based estimatorover UDP traffic traces.

Our primary findings show that UDP traffic traces generally can be categorized as non-Gaussian long-range dependent processes with significantly high non-stationary artifacts. Therefore, the fact that their marginal distributionsare non-Gaussian calls for further investigation of the wavelet estimator performance and, hence, the Hurst parameter estimation of the traces. In opposite to this, previous theoretical studies of the wavelet-based estimator assumed a Gaussian distribution for the traffic traces. In addition, careful observation of the statistical properties for the wavelet coefficients suggests that reduction of the long dependence range into becoming a short dependence range cannot be achieved by increasing the number of vanishing moments. Accordingly performance of the Hurst parameter estimation might not be affected by increasing the number of vanishing moments of the mother wavelet.

The remainder of this paper is organized according to the following sections. Section 2 recalls the statistical concepts underpinning the long-range dependent processes as well as the wavelet transform method. In Section 3, the real traffic data sets that have been used in this study will be briefly reviewed. It is then followed by a description of UDP traffic traces and their protocol characteristics that eventually shape their behavior in terms of scaling dependent structures. Section 4 provides a marginal distribution analysis of UDP traces based on the maximum likelihood estimation (MLE) to estimate parameters of the Gamma distribution for the real traffic traces. Furthermore, the Kolmogorov-Smirnov test will be employed to check the

Gaussianity of the marginal distribution. We shall observe the marginal distribution repeatedly for different aggregation intervals. In the same section, we also report analysis for the quasi-whitening effect in the wavelet estimator and the performance of the Hurst parameter estimation. Finally, conclusions on our study will be drawn in Section 5.

2. THEORETICAL BACKGROUND

This section gives an overview for the long-range dependence (LRD) and a summary of the wavelet spectrum for estimating the Hurst parameter by means of the weighted linear regression. The most important attribute of the wavelet-based estimator is underlined in this section as a basis for correlation structure analysis of the wavelet coefficients in Section 4.

2.1. Definition of Long-Range Dependence (LRD)

Consider a wide sense stationary time-series $Y = (Y_\tau : \tau = 0,1,2,\cdots)$be a discrete time with LRD. In our study, Y could be the time-series of byte counts or packet counts. The series Y is said to be LRD if its covariance function can be represented as [13]

$$\gamma_Y(\tau) \sim c_\gamma |\tau|^{-(2-2H)}, \qquad \tau \to +\infty, \tag{1}$$

with$c_\gamma > 0$ is a constant and $H \in (0.5,1)$. It can be seen that the covariance function in Eq. (1) appears to be a limiting function, this is because the value $\gamma_Y(\tau)$ will be equal to $c_\gamma |\tau|^{-(2-2H)}$ for large τ. An equivalent statement of Eq. (1) but stated in the frequency domain is

$$\Gamma_Y(\nu) \sim c_f |\nu|^{1-2H}, \qquad \nu \to 0, \tag{2}$$

wherec_f is a constant.

Based on Eq. (1) and Eq. (2), LRD can be defined as a slow power-law decrease of the covariance function of a wide sense stationary process with $H > 0.5$. The Hurst parameter, H, in these equations control the degree of LRD. For example, when H gets larger the temporal dependence becomes stronger as the covariance function of such a time series decays more slowly at infinity.

2.2. Wavelet Spectrum for LRD Analysis

We consider a reference function,$\psi(t)$,that is characterized by a strictly positive integer $M_\psi \geq 1$. The function $\psi(t)$ is called the mother wavelet and the parameter$M_\psi$is referred to as the vanishing moments.The mother wavelet is said to have $\psi(t)$ vanishing moments if

$$\int_{\mathbb{R}} t^m \psi(t) \mathrm{d}t = 0 \text{for } m = 0,1,2,\cdots, M_\psi - 1. \tag{3}$$

This equation means that $\psi(t)$ is orthogonal to any polynomial of degree $M_\psi - 1$. In the context of LRD analysis, the vanishing moments, $\psi(t)$, plays an important role for the estimation of the Hurst parameter variance [9]. Theoretically it has been shown that the larger $\psi(t)$ is, the better the estimation. We shall observe the impact of the vanishing moments on the Hurst parameter estimation for UDP traffic traces in Section 4.

Let us now concentrate on a specific discrete wavelet function called the dyadic grid wavelet, which will be used in this study. It has the following form

$$\psi_{j,k}(t) = 2^{-j/2} \psi\left(2^{-j} t - k\right), \qquad j,k \in \mathbb{Z}, \tag{4}$$

signifying dilations to scales 2^j and translation to a time position $2^j k$ of the mother wavelet, $\psi(t)$, for every integer number j and k.

The discrete wavelet transform (DWT) of a function $g(t)$can be obtained through the following linear operation:

$$d_g(j,k) = \int_{\mathbb{R}} \psi_{j,k}(t)\, g(t)\mathrm{d}t\,,$$

$$= \int_{\mathbb{R}} 2^{-j/2}\psi(2^{-j}t-k)g(t)\mathrm{d}t\; j,k \in \mathbb{Z}, \quad (5)$$

where$d_g(j,k)$ is referred to as the wavelet coefficient [14].

Taking Eq. (5) for the case of a second-order stationary LRD process Y, we have the wavelet coefficients $d_Y(j,k)$. The mean energy of the wavelet coefficients at scale j can be defined as:

$$\mathcal{E}_j = \mathbb{E}d_Y(j,k)^2 \approx c_f C 2^{j(2H-1)}, \qquad 2^j \to \infty \quad (6)$$

provided that the integral $C = \int_{\mathbb{R}} |\nu|^{1-2H}|\Psi_0(\nu)|^2\, \mathrm{d}\nu$ converges. It can be observed that the wavelet coefficients on scale j contain the frequency spectrum of the process Y around frequency 2^j. The most interesting feature of the wavelet-based estimator is that, for the Gaussian case, the correlation structure of the wavelet coefficients is not LRD (while the original data is still LRD). In other words, correlation of the $d_Y(j,k)^2$ at all scales 2^j is short-range dependent provided that the non-bias condition $M_\psi > H + 0.5$ is satisfied. This specific feature is called the quasi-whitening effect. Hence, the reduction on the dependence structure of the wavelet coefficients allows us to use the standard sample variance for energy estimation.

Based on a previous study in [15], a spectral estimator for the mean energy can be derived by taking a time average of the $|d_Y(j,k)^2|$ at fixed scale 2^j as follows

$$S_j = \frac{1}{N_j}\sum_{k=1}^{N_j} |d_Y(j,k)^2|, \quad (7)$$

whereN_j is the number of available wavelet coefficients at scale 2^j. It can be expected that $N_j = N/2^j$, where N is the length of Y. However, it should be noted that Eq. (7) is generally valid for the case of Gaussian processes [16].

Now, following the procedure in [9] and looking at Eq. (6) and Eq. (7), the Hurst parameter, H, can be estimated by using a weighted linear regression of $\log_2 S_j$ on scale j as

$$\widehat{H} = \frac{1}{2}\left(\sum_{j=j_1}^{j_2} w_j \log_2 S_j\right) + \frac{1}{2} \quad (8)$$

with $1 \le j_1 < j_2 \le J$, and the weights w_j are chosen such that $\sum_{j=j_1}^{j_2} w_j = 0$and $\sum_{j=j_1}^{j_2} j w_j = 1$.

3. DATA SETS AND WAVELET SPECTRUM OF UDP TRACES

3.1. Data Sets for the Study

The Internet traffic data sets in this study were retrieved from the Measurement and Analysis on the WIDE Internet (MAWI) working group site that resides in Japan, and the Waikato Internet Traffic Storage (WITS), University of Waikato, New Zealand. There are four data sets involved in the study as shown in Table 1. $\widehat{H}_b$and$\widehat{H}_p$ are the global estimated Hurst parameters for byte and packet counts, respectively. The estimated Hurst parameters in this table were obtained using the wavelet estimator.

The MAWI traffic repository records a collection of traffic data from WIDE backbone networks, which is a Japanese academic network that connects universities and research institutes. Our main data sets in the analysis herein were captured at Sample point-F between 18

March 2008 and 20 March 2008 as part of a Day in the Life of the Internet project. The traces were anonymized in their IP addresses and their payloads had been removed. The Sample point-F started to operate in June 2006 substituted for Sample point-B that was operated from January 2001 to June 2006. The Sample point-B had 100Mbps bandwidth with an 18Mbps Committed Access Rate (CAR), while the link for Sample point-F has over-provisioned bandwidth of 1Gbps, with 150Mbps CAR. In our study, the choice of the traces `mawi1`, `mawi2` and `mawi3` represents morning, afternoon and night sessions. By utilizing this variation in time, we are able to capture different behavior of the Internet traffic. We also chose them in such a way that anomalies in the traffic traces during the observation time were minimized. See [17] and its corresponding website (`http://www.fukuda-lab.org/mawilab/`) to assess the traffic anomalies. All applications that employ UDP as their transport protocol have been filtered out for the purposes of our study. In this case there are almost no anomalies found in the UDP-based traffic traces, most of the detected abnormalities are TCP related attacks. In addition, the UDP trace called `wits` was collected from the WITS archive. The trace used in this study was called Auckland IX as it was taken at the end point of the University of Auckland, New Zealand[2] in March 2008.

Table 2. Summary of the UDP traces data.

Trace	Date	Duration	$\widehat{H}_b$	$\widehat{H}_p$
mawi1	18 March 2008	9.00am-11.00am	0.92	1.01
mawi2	19 March 2008	14.00pm-16.00pm	0.99	0.96
mawi3	20 March 2008	19.00pm-21.00pm	0.90	0.93
wits	28 March 2008	08.00am-10.00am	0.89	0.72

3.2. Wavelet Spectrum of the UDP Traces and Hurst Parameter Estimation

The wavelet spectrum for all traces (i.e., `mawi1`, `mawi2`, `mawi3` and `wits`) in terms of byte arrival counts are shown in Figure 1. The plots depict the log-scale diagram of the wavelet spectrum, S_j, as a function of scale j.As shown in Section 2.2., the statistic S_j represents the energy of the traffic trace concentrated at around a frequency range related to the scale j. In Figure 1, the vertical line at each point signifies the confidence interval of $S_j$at a given scale j and a line that is fitted to the wavelet spectrum at large scales(in each trace) is the fitting line that is used for estimating the Hurst parameter, $\widehat{H}$ (i.e. the slope of the line).

In theory, a trace statistically exhibits LRD when $0.5 < \widehat{H} < 1$ [13]. It can be seen in Figure 1 that the wavelet spectrum for all traces show consistency with long-range dependent structure particularly at the coarse scale [18],[10]. The estimated Hurst parameter $\widehat{H}_b$ for byte counts and $\widehat{H}_p$ for packet counts, for each data set is shown in Table 1. It is clear that the values of the $\widehat{H}_b$ and the $\widehat{H}_p$for all traces are larger than 0.5, yet smaller than 1. However, variability in thespectrum, which is represented by a bumpy look of the wavelet spectrum, is clear. The non-linier wavelet spectrum lines in the figure indicate the presence of non-stationary effects in the traces. Because of these non-stationary effects, estimation of the Hurst parameters can be larger than 1, for example as in the $\widehat{H}_p$of`mawi1` trace. In the presence of the non-stationarity effects, estimation may only be valid if they are taken in the coarse scales. Hence, in Table 1, the Hurst parameters were estimated from the slope of the fitted lines from j=12 to j=16.

The sources of non-stationarity might be as a result of local mean-shifts in the traffic traces and the presence of high frequency oscillation. For example, Figure 2 reports the magnitude

[2]http://wand.cs.waikato.ac.nz

fluctuations of the periodogram for all UDP traces.It can be noticed thatspikes are noticeable at a range of frequencies. An empirical study in [5] shows that the bumpy look in the wavelet spectrum starts to appear as soon as$\nu \gg M_{\psi}$.

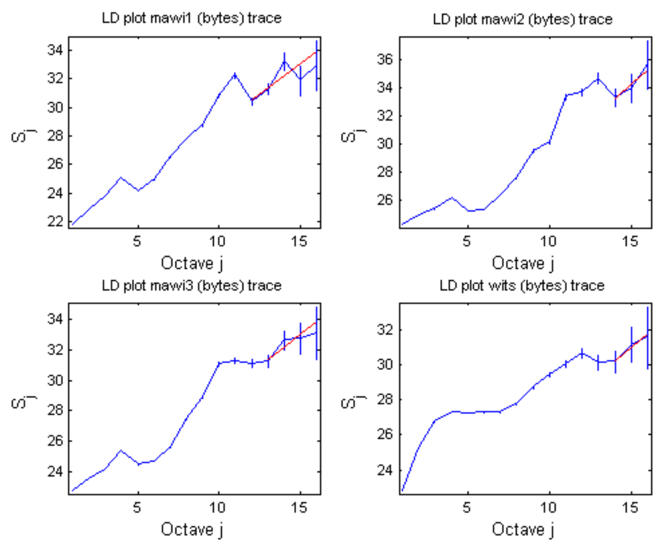

Figure 1.The log-scale wavelet spectrum of the traffic traces in the study in terms of byte arrival counts. The plot depicts the log-scale wavelet spectrum, S_j, as a function of the scale j.

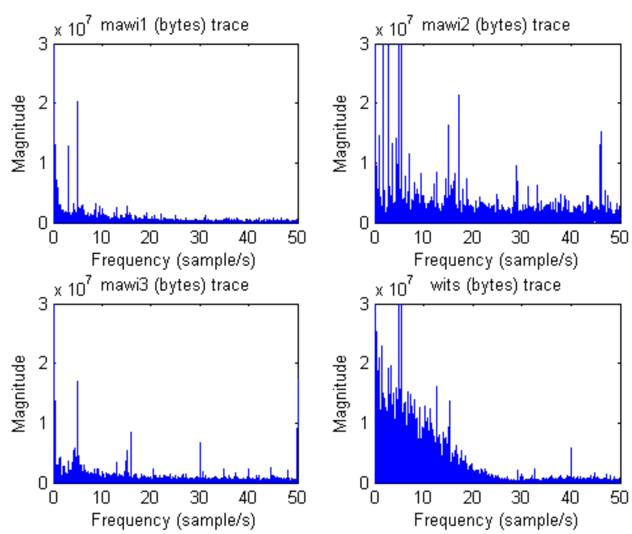

Figure 2.Periodogram of two hours UDP traces based on byte counts that arrive on a link in 10ms time intervals.

In addition, due to the presence of the non-stationarity artifacts in the UDP traffic traces, estimation the Hurst parameter of the LRD becomes a difficult task. Nevertheless, it is worth mentioning that a rather detailed analysis of time-series in the presence of non-stationary effects including some statistical tools, were proposed and outlined in [19][20]. Given this limitation, our focus in this paper is not specifically on the methods for estimating the $\widehat{H}$ parameter, we only use the available tools whenever they are necessary to produce the estimated Hurst parameter, $\widehat{H}$. For example, the local analysis of self-similarity outlined in [20] can accurately estimate the Hurst parameter for data that contains the non-stationarity elements.

In this study, the Hurst parameter estimation was carried out by firstly localizing the analysis of the time-series into smaller blocks of data instead of utilizing the whole block of the time-series, and then, $\widehat{H}$is estimated by averaging those local Hurst parameters. The methodof local analysis of self-similarity has advantage of washing-out the effects of the non-stationarity that are present in the time-series. The Hurst parameter estimation and further discussions will be elaborated in Section 4.

4. RESULTS AND DISCUSSIONS

This section covers the main results and some discussions of the study. It firstly focuses on the marginal distribution characterization of the UDP traces. It is then followed by an analysis of the wavelet coefficient statistics in the light of the quasi-whitening effect (mentioned in Section 2). Lastly, performance of the Hurst parameter estimation for different numbers of the vanishing moments, M_ψ, and different scales will be discussed.

4.1. Marginal Distributions of the UDP Traces

Marginal distributions of the UDP traces in terms of byte and packet counts are analyzed by means of empirical histograms for different aggregation levels. In order to represent a wide range of aggregation levels, we have used aggregation levels of $a = 1,2,\cdots,10$, that correspond to aggregation times of $2^a \times 10$ms. However, in Figure 3 and Figure 4 we only show empirical histograms for the case of $a = 3,4,6,7$ that signify aggregation time, Δ, of 80ms, 160ms, 640ms and 1.28s, respectively.

During the study, we tried different types of distribution models including the Negative Binomial distribution as suggested in [22], the Weibull distribution and the Gamma distribution [12] to be fitted to the marginal distributions of the UDP traces for both byte and packet counts. However, we found that the Negative Binomial and the Weibull did not fit the real UDP traffic marginal distributions very well for a wide range of aggregation levels. Thus, to account for the non-Gaussianity marginal distributions, we utilized the Gamma distribution, which had been considered to be a satisfactory model for traffic traces for both small and large aggregation levels [12].

The fitted lines in Figure 3 and Figure 4 represent Gamma laws that are characterized by two parameters: the shape parameter $\alpha > 0$ and the scale parameter β.Probability density function (pdf) of the Gamma law is defined as follows:

$$\Gamma_{\alpha,\beta}(y) = \frac{1}{\beta\Gamma(\alpha)}\left(\frac{y}{\beta}\right)e^{\left(-\frac{y}{\beta}\right)} \tag{10}$$

where the $\Gamma(.)$ represents the standard Gamma function. Varying the shape parameter from $\alpha = 1$ to $\alpha = +\infty$ corresponds to exponential to Gaussian law distributions. Hence, the distance between $\Gamma_{\alpha,\beta}(y)$ and Gaussian law is controlled by $1/\alpha$. To be specific, the skewness of the $\Gamma_{\alpha,\beta}(y)$ is defined as $2/\sqrt{\alpha}$ and its kurtosis is espresed as $3 + 6/\sqrt{\alpha}$. In contrast to this, Gaussian has skewness=0 and kurtosis=3.

The parameters of α_j and $\beta_j$correspond to the estimated value of $\alpha$and βat an aggregation level j, respectively. For each UDP traces, the parameterswere systematically estimated using the maximum likelihood estimator (MLE) with 95% confidence interval.They are shown in Table 2 together with their estimated skewness and kurtosis. We also checked the closeness of the UDP trace marginal distributions to the Gaussian law using the Kolmogorov-Smirnov test as shown in right hand side of Table 2. The result is $h = 1$ if the test rejects the hypothesis that the particular UDP trace under investigation has a Gaussian distribution at the 5% significance level.

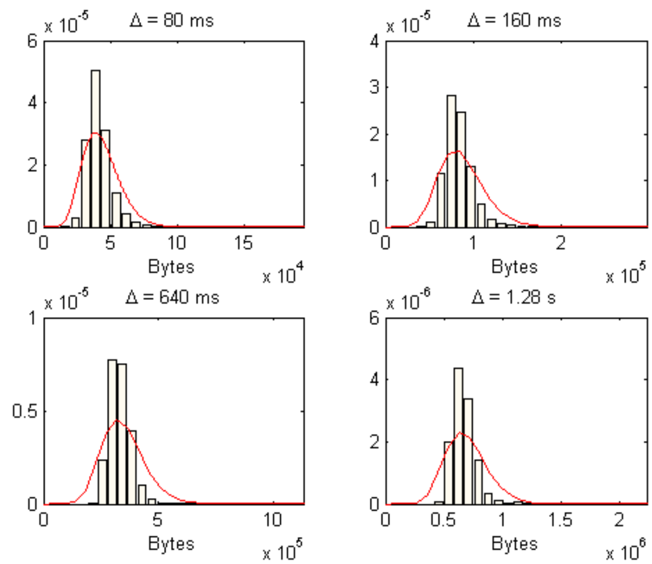

Figure 3. Marginal distributions of the UDP trace `mawi3` in terms of byte counts.

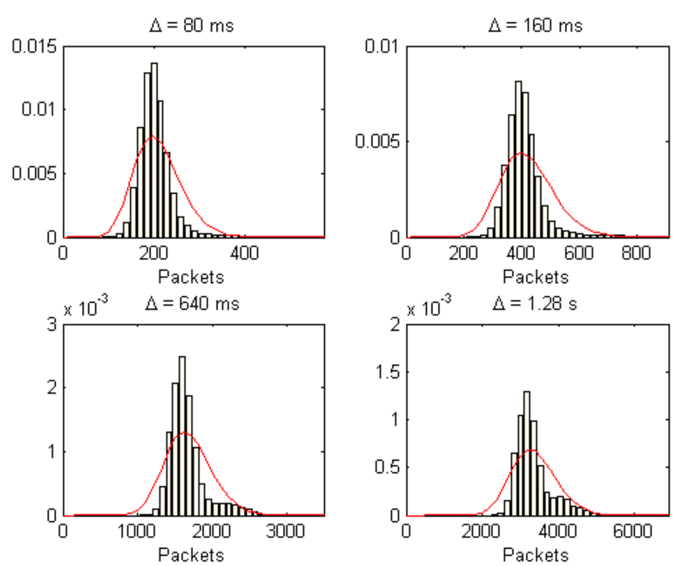

Figure 4. Marginal distributions of the UDP trace `mawi3` in terms of packet counts

By analyzing Table 2, it can be concluded that the marginal distributions of the UDP traffic traces in the study are shown to be non-Gaussian. It is also very interesting to see that increasing the aggregation level may result in shifting the distribution towards the Gaussian distribution (i.e., the skewness gets close to 0 and the kurtosis gets close to 3). However, due to data limitations, it is not achievable in this study. On the other hand, the Kolmogorov-Smirnov test plays an important role in confirming the non-Gaussianity characteristic.It is particularly useful for the case of the `mawi3` trace where estimation of the shape parameter, α, produces very large values. Theoretically, a specific Gamma distribution with large value of α corresponds to the Gaussian laws. It can be seen in the table, although the `mawi3` tracesobey the Gamma laws with very large values of the shape parameter (tend to move toward the Gaussian law), the Kolmogorov-Smirnov test confirms that they cannot be categorized as Gaussian.

Table 2. Gamma distribution parameters and Gaussianity check for the UDP traces data

Trace	$\boldsymbol{\alpha_j}$	$\boldsymbol{\beta_j}$	**Skewness**	**Kurtosis**	**KS-Check**
mawi1(byte)	$\alpha_3 = 5.40$	$\beta_3 = 4335.6$	0.86	5.58	$h = 1$
	$\alpha_4 = 6.41$	$\beta_4 = 7307$	0.79	5.37	$h = 1$
	$\alpha_6 = 7.96$	$\beta_6 = 23525$	0.71	5.13	$h = 1$
	$\alpha_7 = 8.50$	$\beta_7 = 44063$	0.69	5.06	$h = 1$
mawi1(packet)	$\alpha_3 = 9.95$	$\beta_3 = 13.69$	0.63	4.90	$h = 1$
	$\alpha_4 = 12.44$	$\beta_4 = 21.89$	0.57	4.70	$h = 1$
	$\alpha_6 = 18.11$	$\beta_6 = 60.16$	0.47	4.41	$h = 1$
	$\alpha_7 = 20.19$	$\beta_7 = 107.93$	0.45	4.36	$h = 1$
mawi2(byte)	$\alpha_3 = 4.38$	$\beta_3 = 12339$	0.96	5.87	$h = 1$
	$\alpha_4 = 5.16$	$\beta_4 = 20947$	0.88	5.64	$h = 1$
	$\alpha_6 = 5.93$	$\beta_6 = 72862$	0.82	5.46	$h = 1$
	$\alpha_7 = 6.12$	$\beta_7 = 141290$	0.81	5.43	$h = 1$
mawi2(packet)	$\alpha_3 = 9.55$	$\beta_3 = 21.03$	0.65	4.94	$h = 1$
	$\alpha_4 = 11.12$	$\beta_4 = 36.12$	0.60	4.80	$h = 1$
	$\alpha_6 = 13.63$	$\beta_6 = 117.85$	0.54	4.63	$h = 1$
	$\alpha_7 = 14.37$	$\beta_7 = 223.50$	0.53	4.58	$h = 1$
mawi3(byte)	$\alpha_3 = 9.89$	$\beta_3 = 4383.7$	0.64	4.91	$h = 1$
	$\alpha_4 = 11.89$	$\beta_4 = 7294.5$	0.58	4.74	$h = 1$
	$\alpha_6 = 14.63$	$\beta_6 = 23715$	0.52	4.57	$h = 1$
	$\alpha_7 = 15.47$	$\beta_7 = 44834$	0.51	4.53	$h = 1$
mawi3(packet)	$\alpha_3 = 16.45$	$\beta_3 = 12.80$	0.49	4.48	$h = 1$
	$\alpha_4 = 20.63$	$\beta_4 = 20.40$	0.44	4.32	$h = 1$
	$\alpha_6 = 29.41$	$\beta_6 = 57.72$	0.37	4.11	$h = 1$
	$\alpha_7 = 32.60$	$\beta_7 = 103.29$	0.35	4.05	$h = 1$
wits(byte)	$\alpha_3 = 3.51$	$\beta_3 = 8560$	1.07	6.20	$h = 1$
	$\alpha_4 = 4.12$	$\beta_4 = 14561$	0.98	5.95	$h = 1$
	$\alpha_6 = 6.15$	$\beta_6 = 39082$	0.81	5.42	$h = 1$
	$\alpha_7 = 7.97$	$\beta_7 = 60288$	0.71	5.13	$h = 1$
wits(packet)	$\alpha_3 = 7.78$	$\beta_3 = 20.850$	0.72	5.15	$h = 1$
	$\alpha_4 = 10.35$	$\beta_4 = 31.35$	0.62	4.87	$h = 1$
	$\alpha_6 = 17.87$	$\beta_6 = 72.61$	0.47	4.42	$h = 1$
	$\alpha_7 = 22.86$	$\beta_7 = 113.49$	0.42	4.25	$h = 1$

4.2. The Wavelet Coefficients Statistics

Theoretically, the correlation structure of the wavelet coefficients$d_Y(j,k)^2$ are short-range dependent at all scales 2^j provided that the non-bias requirement$M_\psi > H + 0.5$ is satisfied. However, it should be noted that this condition is valid only for Gaussian process. It has been shown that the UDP traffic traces in our study inherit the LRD characteristics (in Section 3.2) as well as the non-Gaussian marginal distribution (in Section 4.1).This sub-section investigates the statistical properties of the wavelet coefficients for the non-Gaussian long-range dependence that appears from our UDP traffic traces. The statistics of the wavelet coefficients will be observed through marginal distributions and the covariance structure for different number of vanishing moments, M_ψ.

Figure 5 and Figure 6 present the estimated skewness and kurtosis of the wavelet coefficients for different numbers of vanishing moments. Inorder to overcome the non-stationarity effects in the traces, estimation of the skewness and kurtosis was performed by dividing the two-hour UDP traces into 24 blocks (each of the mcorresponds to a 5 minute time interval). Then the wavelet coefficients were obtained from Eq. 5 in Section 2, where the UDP traffic traces samples substitute the function $g(t)$.In this way, the focus of the study is on the local behavior of the time series instead of the global one. Similar analysis for the traces that were distorted by then on-stationarity effects was discussed in [20].The figures also show the estimation of the skewness and kurtosis for different $M_\psi$with 95%confidence intervals.The confidence interval arerepresented by vertical lines at each point.On the other hand, the horizontal red lines serve as reference lines, i.e. the Gaussian skewness=0 and the Gaussian kurtosis=3.

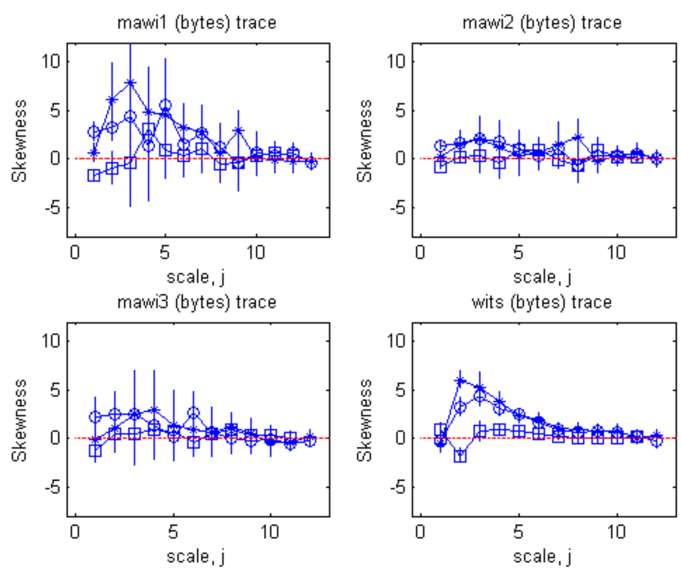

Figure 5. Estimated skewness of the wavelet coefficients for 2 hours UDP traces based on bytecounts.$M_\psi = 2$ represented by a line with '*',$M_\psi = 3$ represented by a line with 'o', and $M_\psi = 5$represented by a line with square.

Looking at these figures, it can be clearly seen that the marginal distributions of the wavelet coefficients for all UDP traces in the study are approaching the Gaussian lawat large scale, i.e., $2^j \to \infty$. It is noticeable that the skewness of the traces tends to converge to 0 while the kurtosis of all traces decreases slowly to 3. The second important characteristic of the wavelet coefficientsis that both the skewness and kurtosis of the traces, for different values of M_ψ, have tendencies to converge to a certain value at large scale. This is an indication that increasing the number of vanishing moments of the mother wavelet does not affect the performance of the wavelet estimator. This result is consistent with empirical studies shown in [23] andthis finding confirms that for UDP traces data that possess the non-Gaussian LRD behavior, the wavelet coefficients eventually becomes Gaussian only at large scale. Hence, it consequently verifies the validity of the non-bias condition of the wavelet-based estimator used in this study.

The next observation concerns an important feature of awavelet estimator called the quasi-whitening effect, which hasan impact on transforming the dependence structure of the waveletcoefficients from long to short by increasing the degree of theM_ψ[24]. Theoretically, it was stated that thewavelet transform with a higher degree vanishing moments willdisentangle the dependence structure of the traffic from longdependence to become short dependence irrespective of the marginaldistributions as long as $M_\psi > H + 0.5$[24].Hence, the case for $M_\psi = 1$ was not used in the study.

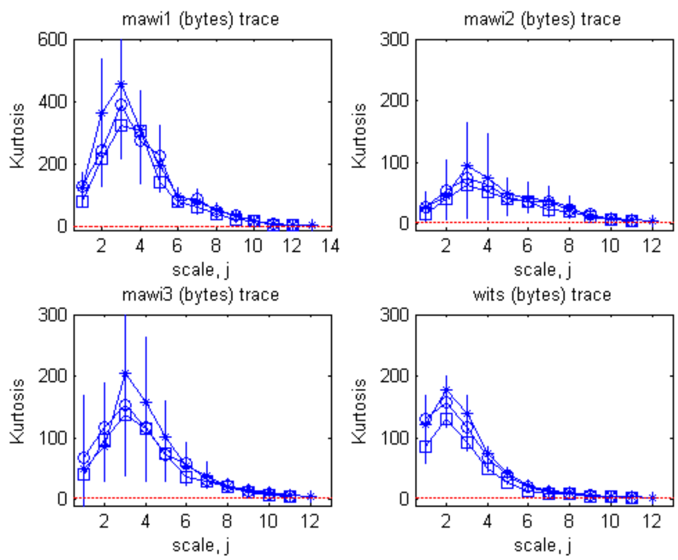

Figure 6. Estimated kurtosis of the wavelet coefficients for 2 hours UDP traces based on bytecounts.$M_\psi = 2$ represented by a line with '*', $M_\psi = 3$ represented by a line with 'o', and $M_\psi = 5$ represented by a line with square.

Figure 7 provides the dependence behaviour of the`mawi2` traces observed through covariance of the wavelet coefficients. Although examinations were also carriedout for all other traces, due to space limitations they are notshown here. The top part of the figure shows the covarianceat scale $j = 2$ and the bottom part shows the covariance for scale$j = 9$, while the left column depicts the covariance with$M_\psi = 2$ and the right column shows the covariance with$M_\psi = 5$. It can be observed that there is no significantimpact on the long range dependence structure either at the smallscale $j = 2$ or the large scale $j = 9$. A careful look from left toright in the figure also reveals that increasing the degree ofvanishing moments does not provide any impact on decreasing thedependence structure. Therefore, the fact that thequasi-whitening effect appears to be valid for Gaussianprocesses, it does not seem to apply for the non-Gaussian processeslike the UDP traces in this study. This result is again in agreementwith previous work in [23].

4.3. Performance of the Hurst Parameter Estimation

The Hurst parameter is estimated by employing methods outlined in Section 2.2, specifically using Eq. (9). Various values of $M_\psi$spanning from 2 to 5 were used in order to evaluate the effect of increasing the number of vanishing moments of the mother wavelet towards the performance of the Hurst parameter estimation. Results are also presented for different scale regimes as can be seen in Table 3. Fine scalesare observed at range $j_1 = 3$to $j_2 = 6$, coarse scales at range $j_1 = 8$ to $j_2 = 12$ and global scales at range $j_1 = 3$ to $j_2 = 12$.

The estimated Hurst parameter, $\widehat{H}$, and its variance, $\sigma_{\widehat{H}}$, were taken from the local analysis of the time-series.There are approximately7.2×10^5samples of data (representing 2 hours of UDP traces)that had been subdivided into 24 blocks. Thus, each block of data signifies 3.0×10^4samples of data. The wavelet spectrum estimator was then applied for each block of this data to estimate the Hurst parameter and the variance.

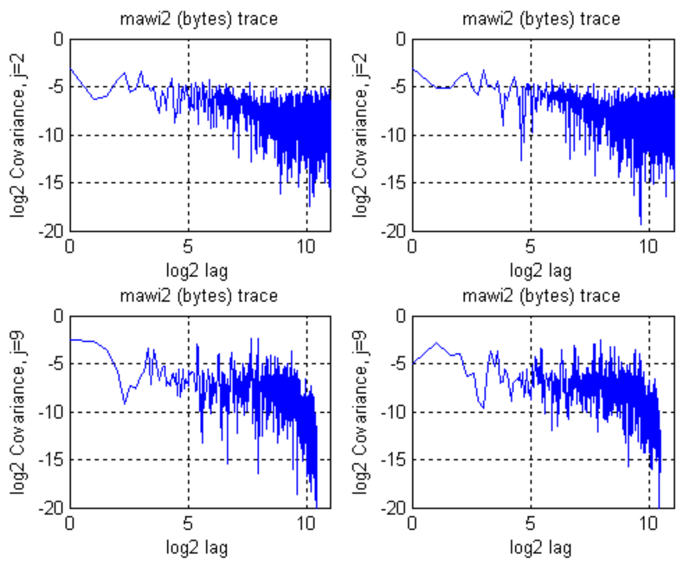

Figure 7. Covariance of the wavelet coefficients for the `mawi2` traces. Top part: covariance at scale $j = 2$,bottom part: covariance for scale $j = 9$. Left column: covariancewith $M_\psi = 2$, Right column: covariance with$M_\psi = 5$.

As can be seen in Table 3, a relatively small value of sample variance for different kinds of time-series data and for a range value of M_ψ indicates that the statistical discrepancy between each block of data is fairly small. The only exception is for the case of the `mawi2` trace for both byte and packet counts, where large values of its variance for different M_ψ result in significantly diverse sample means $\widehat{H}$. This is likely due to the statistical characteristic of the `mawi2` traces in away that the non-stationarity does not affect the global observation of time-series only, but also distorts the time-series up to a small block of observations.

Evaluation of Table 3 clearly shows that for the non-Gaussian processes like UDP traces, there may seem to be no difference in the performance of the estimated Hurst parameter $\widehat{H}$ as well as its varian, $\sigma_{\widehat{H}}$, as the number of vanishing moments were increased. This can be linked to the previous examination in Section 4.2., whereby setting the M_ψ to large values does not contributeto transforming the dependence structure of the wavelet coefficients from long to short.

Inspection at several scales shows significant difference for $\widehat{H}$. This result agrees with common analysis of the wavelet estimator for Internet traffic. See, for example in [25],[10]. However, the coarse scale asymptotic analysis of the wavelet coefficients in Eq. (8) suggests that it is necessary to confine the linear regression to estimating $\widehat{H}$ only on the coarse scale. Comparing $\widehat{H}$ in the coarse scale of Table 3 to the previous estimation in Table 1, it is clear that global analysis time-series gives an overestimate of the values for the Hurst parameter. For example, for the `wits` traffic, in Table 1 the estimated $\widehat{H}_b = 0.89$ and $\widehat{H}_p = 0.72$ for byte arrivals and packet arrivals, respectively, while local analysis in Table 3 only gives approximately $\widehat{H}_b = 0.54$ and $\widehat{H}_p = 0.66$ for byte counts and packet counts, respectively.

Furthermore, observation on local analysis UDP traffic in the coarse scale, i.e. between $j_1 = 8$ and $j_2 = 12$, shows that the estimated Hurst parameter values in Table 3 lie between 0.5 and 1. This is again an indication that the UDP traffic under investigation statistically exhibits LRD. Consequently, the LRD characteristics that has been proved to be existingin the UDP-based Internet traffic has important implications on the performance, design and dimensioning of the current network. See for example in [18].

Table 3. Sample mean and sample variance of $\widehat{H}$

	$M_\psi = 2$		$M_\psi = 3$		$M_\psi = 5$	
	$\widehat{H}$	$\sigma_{\widehat{H}}$	$\widehat{H}$	$\sigma_{\widehat{H}}$	$\widehat{H}$	$\sigma_{\widehat{H}}$
`mawi1 (byte)`						
$(j_1, j_2) = (3,6)$	0.62	0.060	0.64	0.055	0.66	0.045
$(j_1, j_2) = (3,12)$	0.63	0.050	0.65	0.060	0.64	0.050
$(j_1, j_2) = (8,12)$	0.78	0.061	0.74	0.068	0.73	0.098
`mawi1 (packet)`						
$(j_1, j_2) = (3,6)$	0.63	0.026	0.65	0.024	0.65	0.020
$(j_1, j_2) = (3,12)$	0.63	0.034	0.64	0.036	0.63	0.032
$(j_1, j_2) = (8,12)$	0.72	0.046	0.69	0.033	0.69	0.081
`mawi2 (byte)`						
$(j_1, j_2) = (3,6)$	0.51	0.007	0.50	0.009	0.52	0.008
$(j_1, j_2) = (3,12)$	0.57	0.013	0.56	0.013	0.55	0.013
$(j_1, j_2) = (8,12)$	0.85	0.104	0.86	0.103	0.71	0.220
`mawi2 (packet)`						
$(j_1, j_2) = (3,6)$	0.58	0.009	0.58	0.008	0.59	0.008
$(j_1, j_2) = (3,12)$	0.63	0.013	0.62	0.011	0.62	0.011
$(j_1, j_2) = (8,12)$	0.83	0.113	0.81	0.123	0.70	0.230
`mawi3 (byte)`						
$(j_1, j_2) = (3,6)$	0.55	0.012	0.58	0.021	0.58	0.016
$(j_1, j_2) = (3,12)$	0.58	0.019	0.59	0.027	0.58	0.023
$(j_1, j_2) = (8,12)$	0.87	0.070	0.86	0.096	0.82	0.162
`mawi3 (packet)`						
$(j_1, j_2) = (3,6)$	0.58	0.007	0.59	0.009	0.59	0.008
$(j_1, j_2) = (3,12)$	0.61	0.012	0.61	0.015	0.61	0.013
$(j_1, j_2) = (8,12)$	0.85	0.067	0.84	0.093	0.84	0.143
`wits (byte)`						
$(j_1, j_2) = (3,6)$	0.62	0.003	0.64	0.002	0.62	0.001
$(j_1, j_2) = (3,12)$	0.59	0.002	0.60	0.001	0.60	0.001
$(j_1, j_2) = (8,12)$	0.54	0.023	0.55	0.068	0.65	0.047
`wits (packet)`						
$(j_1, j_2) = (3,6)$	0.62	0.002	0.63	0.001	0.62	0.001
$(j_1, j_2) = (3,12)$	0.60	0.001	0.61	0.001	0.60	0.001
$(j_1, j_2) = (8,12)$	0.66	0.023	0.66	0.040	0.65	0.047

5. CONCLUSIONS

Internet traffic traces that utilized the UDP transport protocol have been observed. In this work, we investigated the performance of a wavelet-based estimator employed for the non-Gaussian long-range dependent data.

It is evident that samples of UDP traffic that arrived on a link during 10ms time intervals taken from various data sets exhibit non-Gaussian distributions for both byte and packet counts. A careful examination through their marginal distribution properties showed that their distributions are highly skewed to the left and well-modeled by the Gamma law for different aggregation levels.

Evaluation of the wavelet coefficient marginal distributions for all UDP traces was done in terms of estimated skewness and kurtosis for different numbers of vanishing moments, M_{ψ}. It showed that for the UDP trace data that possess non-Gaussian marginal distributions, the wavelet coefficients marginal distribution eventually become Gaussian at a large scale only. However, increasing the degree of vanishing moments does not result in any impact on decreasing the dependence structure of the wavelet coefficients. Therefore, the fact that the quasi-whitening effect appears to be valid for Gaussian processes, it does not seem to apply for non-Gaussian processes like the UDP traces in this study. As a result of this behavior, there may seem to be no difference on the performance of the $\widehat{H}$ as well as the $\sigma_{\widehat{H}}$ as the number of vanishing moments were increased.Numerical simulations on the real UDP traffic traces indicated that increase on the number vanishing moments of the wavelet estimator, did not change the performance of the $\widehat{H}$and the $\sigma_{\widehat{H}}$ significantly.

We close our conclusions with observations on local analysis UDP traffic in the coarse scale, i.e. between $j_1 = 8$and $j_2 = 12$. It showed that the UDP traffic under investigation statistically exhibits LRD behavior, where $0.5 < \widehat{H} < 1$. Hence, the LRD characteristics that is present in the UDP-based Internet traffic gives important implications on the performance, design and dimensioning of the current network.

6. REFERENCES

[1] Fomenkov, M., Keys, K., Moore, D., & Claffy, K., (2004) "Longitudinal Study of Internet Traffic in 1998 – 2003", *Proceedings of the Winter International Symposium on Information and Communication Technologies (WISICT)*, pp. 16.

[2] Borgnat, P., Dewaele, G., Fukuda, K., Abry, P., & Cho, K, (2009) "Seven Years and One Day:Sketching the Evolution of Internet Traffic", *Proceedings of the IEEE INFOCOM*, pp. 711-719.

[3] Mizanian, K., Vasef, M., & Analoui,M., (2010) "Bandwidth Modelling and Estimation in Peer to Peer Networks", *International Journal of Computer Networks & Communications (IJCNC)*, Vol 2, No. 3, pp. 65-83.

[4] Zhang, M., Dusi, M., John, W., & Chen, C., (2009) "Analysis of UDP Traffic Usage on Internet Backbone Links", *Proceedings of the IEEE Ninth Annual International Symposium on Applications and the Internet*, pp. 280-281.

[5] Leland, W., Taqqu, M.S., Willinger, W., & Wilson, D.V., (1994) "On the self-similar nature of Ethernet traffic", IEEE/ACM *Transactions on Networking*, Vol. 2, No. 1, pp. 1-15.

[6] Paxson, V. & Floyd, S., (1995) "Wide Area Traffic: The Failure of Poisson Modelling", IEEE/ACM *Transactions on Networking*, Vol. 3, No. 3, pp. 226-244.

[7] Crovella, M., &Bestavros, A., (1997) "Self-similarity in World Wide Web Traffic: Evidence and Possible Causes", IEEE/ACM *Transactions on Networking*, Vol. 5, No. 6, pp. 835-846.

[8] Fukuda, K., Takayasu, M., &Takayasu, H., (2005) "A Cause of Self-Similarity in TCPTraffic", *International Journal of Communication Systems*, Vol. 18, pp. 603-617.

[9] Abry, P., &Veitch, D., (1998) "Wavelet Anaylisis of Long-Range Dependent Traffic", IEEE *Transactions on Information Theory*, Vol. 44, No. 1, pp. 2-15.

[10] Veitch, D., &Abry, P., (1999) "A Wavelet-Based Joint Estimator of the Parameters of Long-Range Dependence", IEEE *Transactions on Information Theory*, Vol. 45, No. 3, pp. 878-897.

[11] Veitch, D., Abry, P., &Taqqu, M.S., *On the Automatic Selection of the Onset of Scaling*, Fractals, **11**(4), pp. 377-390, 2003.

[12] Scherrer, A., Larrieu, A., Borgnat, P., Owezarski, P., &Abry, P., (2007) "Non Gaussian and Long Memory Statistical Modeling of Internet Traffic", IEEE *Transaction on Dependable and Secure Computing*, Vol. 4, No. 1, pp. 56-71.

[13] Beran, J., (1994) *Statistics for Long Memory Processes*, London, UK,Chapman and Hall.

[14] Addison, P., (2002) *The Illustrated Wavelet Transform Handbook*, London, UK,Institute of Physics Publishing.

[15] Abry, P., Gonçalvès, P., Flandrin, P., (1993) "Wavelet-Based Spectral Analysis of 1/f Processes", *Proceedings of IEEE ICASP-93*, pp. III.237-III.240.

[16] Abry, P., Flandrin, P., Taqqu, M.S., &Veitch, D., (2000) *Wavelets for Analysis, Estimation and Synthesis of Scaling Data, in: Self-Similar Network Traffic and Performance Evaluation*, New York, Wiley.

[17] Fontugne, R., Borgnat, P., Abry, P., & Fukuda, K., (2010) "MAWILab: Combining Diverse Anomaly Detectors for Automated Anomaly Labeling and Performance Benchmarking", *Proceedings of the 6th International Conference Co-NEXT 2010*, pp. 8:1-8:12.

[18] Loiseau, P., Gonçalvès, P., Dewaele, G., Borgnat, P., Abry, P., &Primet, P.V., (2010) "Investigating Self-Similarity and Heavy-Tailed Distributions on a Large Scale Experimental Facility", IEEE/ACM *Transactions on Networking*, Vol 18, No. 4, pp. 1261-1274.

[19] Stoev, S., Taqqu, M.S., Park, C., &Marron, J.S., (2005) "On the Wavelet Spectrum Diagnostic for Hurst Parameter Estimation in the Analysis of Internet Traffic", *Computer Networks*, Vol. 48, pp. 423-445.

[20] Stoev, S., Taqqu, M.S., Park, C., Michailidis, G., &Marron, J.S., (2006) "LASS: a Tool for the Local Analysis of Self-Similarity", *Computational Statistics and Data Analysis*, Vol. 50, pp. 2447-2471.

[21] Park, C., Campos, F.H., Marron, J.S., & Smith, F.D., (2003) "Long-Range Dependence in a Changing Internet Traffic Mix", *Computer Networks*, Vol. 48, pp. 401-422.

[22] Liu, Z., Almhana, J., Choulakian, V., &McGorman, R., (2005) "Modeling Internet Traffic Using nonGaussian Time Series Models", *Proceedings of the 3rd Annual Communication Networks and Services Research Conference*, pp. 99-104.

[23] Abry, P., Helgason, H., &Pipiras, V., (2011) "Wavelet-Based Analysis of Non-Gaussian Long-Range Dependent Processes and Estimation of the Hurst Parameter", *Lithuanian Mathematical Journal*, pp. 1-16.

[24] Flandrin, P., (1992) "Wavelet Analysis and Synthesis of Fractional Brownian Motion", IEEE *Transactions on Information Theory*, Vol. 38, pp. 910-917.

[25] Zhang, Z., Ribeiro, V.J., Moon, S., &Diot, C., (2003) "Small-TimeScalingBehaviors of Internet BackboneTraffic: An EmpiricalStudy", *Proceedings of IEEE INFOCOM 2003*.

Authors

Short Biography

JusakJusak is a Senior Lecture in the School of Informatics Management and Computer Engineering STMIK STIKOM Surabaya. He graduated with B.Eng from Brawijaya University, Indonesia. He has a PhD in Electronic and Communication Engineering from RMIT University, Australia in 2005. During 2009-2011, He took an appointment as a Postdoctoral Research Fellow at Massey University, New Zealand. His research interests include Internet and Networks protocolstechonology as well as teletraffic engineering.

Richard Harris graduated with a B.Sc. (Hons) from the Uuniversity of Adelaide in 1971 and this was followed by a Ph.D. (in the area of optimal design of telephone networks) under the supervision of Professor R.B. Potts in 1974. He became Deputy Head of the Network Analysis Section in the Switched Networks Research Branch, Telecom Australia Research Laboratories, in 1987. In April 1989, he took up an appointment with Bond University as an Associate Professor in the School of Information and Computing Sciences and was Deputy Director of the Centre for Telecommunication Network Research. He moved to RMIT in early 1993 as Director of the Centre for Advanced Technology in Telecommunications(CATT) and Professor of Communication Systems. In 2005 he took up an appointment asChair in Telecommunications and Network Engineering at Massey University in New Zealand. He is currently teaching students in the fields of teletraffic engineering, communication switching, Internet engineering and conducting research in a variety of topics including network design and optimisation, network traffic management, network performance and security. He has over 200 publications in scientific journals, conference proceedings and research reports. He has served on the editorial board of the International Telecommunications Systems Journal and numerous Technical Program Committees for international and national conferences such as ATNAC, ITC Congresses, NETWORKS, LCN, SATNAC, VTC, ICETE, DELTA and ICARA.