

Searching for Self-Similarity in GPRS

Roger Kalden, Sami Ibrahim

Ericsson Research, Ericsson Eurolab Deutschland GmbH, Aachen, Germany
Roger.Kalden@ericsson.com, phone +49 2407 575 7831

Abstract. Based on measurements in live GPRS networks, the degree of self-similarity for the aggregated WAP and WEB traffic is investigated by utilizing six well established Hurst parameter estimators. We show that in particular WAP traffic is long-range dependent and its scaling for time scales *below* the average page duration is not second order self similar. WAP over UDP can also determine the overall traffic scaling, if it is the majority traffic. Finally we observe that the minor traffic exhibits a larger Hurst value than the aggregated traffic, in case of WAP as well as in case of WEB traffic.

1. Introduction

Based on live GPRS traffic measurements, we investigate the packet arrival process and the data volume arrival process of WAP and WEB traffic on statistical self-similarity.

Many studies have been looking at various network types and found evidence for self-similarity (e.g., [1],[2],[3]). This property is regarded as an invariant of network traffic and has serious performance implications. In the case of self-similar traffic the applied statistics for performance analysis and network dimensioning are different from those when applied to statistically more simple traffic, which can be modeled with Markovian processes ([4],[5]). For instance the queue tail behavior is heavy-tailed in the case of self-similar input traffic [6]. This leads to heavy-tailed packet delay distributions, which can influence the TCP round-trip time estimations. Furthermore, the traffic does not smooth out in the case of aggregation, leading to congestion situations and packet-drops due to the burstiness of the traffic. Consequently, it is important to understand the self-similar nature of the traffic in a network in order to apply the right statistical methods for performance investigations and network dimensioning. In previous studies the reason for self-similarity has been identified as the heavy-tailedness of many statistical properties of Internet traffic, on the user and application level. In [7] and [18] the authors showed that heavy-tailed sessions and file size lengths lead to self-similarity for large aggregation scales (long-range dependency). The authors in [8] and [14] have furthermore shown that self-similarity (in particular the multiscaling) behavior at small timescales (e.g., smaller than the average round-trip time) is due to protocol interactions of TCP.

But GPRS traffic and in general cellular access network traffic has not yet been investigated on its self-similarity property. GPRS is not merely a new access technology, it also introduces novel applications such as WAP and MMS, and provides Inter-

net access in a mobile or nomadic usage environment. This yields a special traffic composition, different from wireline Internet traffic. We have investigated GPRS networks with more than 60% of the traffic volume consisting of UDP traffic. This is in sharp contrast to the usual 80% of TCP traffic in the fixed Internet (cf. [9] for earlier figures on GPRS traffic). Additionally, WAP and MMS file sizes are in general much shorter than WEB and FTP file sizes [17]. For these reasons we counter-check the property of self-similarity in GPRS. Based on our results, which show that GPRS and in particular WAP traffic, is asymptotically second order self-similar (long-range dependent), we propose for further research to study the scaling nature of GPRS traffic and to explore in particular the reasons for this for WAP traffic.

The remainder of the paper is structured as follows. Firstly, we give a brief overview of self-similarity together with the commonly used methods to test self-similarity. Next, we describe our measurement set-up and the investigated traces. In our main section we present the results of the Hurst parameter estimation methods to assess the degree of self-similarity in GPRS. Finally, in the concluding section we list open issues and future work items.

2. Self-Similarity

Self-similarity, in a strict sense, means that the statistical properties (e.g., all moments) of a stochastic process do not change for all aggregation levels of the stochastic process. That is, the stochastic process “looks the same” if one zooms in time “in and out” in the process. Let $X = (X_t : t = 0, 1, 2, \dots)$ be a covariance-stationary stochastic process, with constant and finite mean and finite variance and autocorrelation function $\rho(k)$, which only depends on k . Further, let $X^{(m)} = (X_k^{(m)} : k = 1, 2, 3, \dots)$ be a new time series generated by averaging the original time series X over non-overlapping blocks of size m . That is, for each $m = 1, 2, 3, \dots$ the series is given by $X_k^{(m)} = 1/m(X_{km-m+1} + \dots + X_{km})$, $k = 1, 2, 3, \dots$ and $\rho^{(m)}(k)$ the corresponding autocorrelation function. The stochastic process $X = (X_t : t = 0, 1, 2, \dots)$ is then called *exactly second-order self-similar* with self-similarity parameter $H = 1 - \beta/2$, if the autocorrelation functions $\rho^{(m)}(k)$ of the processes $X^{(m)}$ are all equal to the autocorrelation function $\rho(k)$ of the original process X . That is: $\rho^{(m)}(k) = \rho(k)$, $k = 1, 2, 3, \dots$ for all $m = 1, 2, 3, \dots$. X is called *asymptotically second-order self-similar* with self-similarity parameter $H = 1 - \beta/2$ if the correlation structure of the aggregated time series $X^{(m)}$ is indistinguishable from the correlation structure of the original time series X as $m \rightarrow \infty$. That is, the above holds asymptotically for large aggregation levels.

H expresses the degree of self-similarity; large values indicate stronger self-similarity. If $H \in (0.5, 1)$ X is called long-range dependent (LRD). Both, exactly and asymptotically second-order self-similar traffic can include long-range dependency, however, as it is often the used case, we will use long-range dependent for asymptotically second-order self-similar traffic.

Asymptotical second-order self-similar processes have a few interesting characteristics. For instance, the autocorrelation function of an LRD process is decaying hyperbolically. That is (with β the same as in H, above):

$$\lim_{k \rightarrow \infty} \rho(k) \sim |k|^{-\beta} \quad (0 < \beta < 1).$$

Also the variance of the aggregated time series is very slowly decaying. That is:

$$\text{Var}[X^{(m)}] \sim m^{-\beta} \quad (0 < \beta < 1).$$

Another property is that the power spectrum is singular as the frequency is approaching 0. That is:

$$S(w) \underset{w \rightarrow \infty}{\sim} 1/|w|^{(1-\beta)} \quad (0 < \beta < 1).$$

3. Estimation Methods for Self-Similarity

Various methods for estimating the Hurst parameter H exist for deducing self-similarity or long-range dependency [15]. The estimation methods can be grouped into time-based and frequency-based methods. We will briefly provide an overview of the methods used to estimate the value of the Hurst parameter.

The first four methods are time-based:

R/S method

This method is based on empirical observations by Hurst and estimates H are based on the R/S statistic. It indicates (asymptotically) second-order self-similarity. H is roughly estimated through the slope of the linear line in a log-log plot, depicting the R/S statistics over the number of points of the aggregated series.

Variance Method

The Variance Method is based on the slowly decaying variance property as stated above. It indicates long-range dependency. The slope β of the straight line in a log-log plot, depicting the sample variance over the block size of each aggregation, is used for roughly estimating H. H is given by $H = 1 - \beta/2$.

Absolute Moment Method

This method is related to the variance method computed for the first moment. The slope α of the straight line in a log-log plot, depicting the first moment of the aggregated block over the block size, provides an estimator for H, by $H = 1 + \alpha$.

Ratio of Variance of Residuals

This uses the empirical observation, that is, the sample variance of residuals, plotted over the aggregation level, yields a slope equivalent to roughly 2H. It indicates some self-similarity.

The next two methods are frequency-based:

Periodogram method

This method is based on the power-spectrum singularity at 0-property as stated above. The slope of the straight line, approximating the logarithm of the spectral density over the frequency as the frequency approaches 0, yields H.

Abry-Veitch method

This method is based on the multi-resolution analysis and the discrete wavelet transformation. H is estimated by fitting a straight line to the energy in the series over octave j (expressing the scaling level in the time and the frequency domain) in a log-log plot.¹ This method is the most comprehensive and robust method for determining the scaling behavior of traffic traces. Its strength follows from the fact that the multi-resolution analysis itself has a structural affinity to the scaling process under study. That is, multi-resolution analysis itself exploits scaling, but transfers the complex scaling process to a much simpler wavelet domain, in which short range dependent (SRD) statistics can be applied to infer answers on the scaling of the process [10].

We show results from our traces for all mentioned methods. In particular we will use the Abry-Veitch method to derive the scaling nature of the process.

All used methods provide some intermediate statistics, which is used to derive the Hurst value. For instance, in the case of the Variance Method, these are the aggregated variance values for each aggregation level; or, in the case of the Abry-Veitch method, the intermediate statistics used are the wavelet coefficients. Based on those values linear regression is used to fit a straight line to derive the Hurst value.

Important to consider is that typically the linear regression should not consider all of the values from the intermediate statistic. In case of the R/S method, the Variance Method and the Absolute Moment method, it is recommended not to use the results of the first few aggregations levels and neither the last few aggregation levels. The reason for this is that these values are not very reliable because either the aggregation level is too low (sampling too few points per block) or it is too high (sampling all points in just a few blocks). In the case of the Periodogram Method it is recommended only to use approximately the first 10 percent of the results, close to the frequency 0. This is justified by the asymptotic LRD property close to the frequency 0. The Abry-Veitch method is the most robust of all estimation methods [10]. It actually shows the scaling of the process over all aggregation levels (octave j), which allows to optimally select an appropriate starting point for the regression. This starting point is indicated by a χ^2 -goodness-of-fit test. In the case of assumed LRD traffic, the regression line is fitted from this starting point to the largest available octave in the data.

We use the SELFIS tool [11] to derive the intermediate statistics for all but the Abry-Veitch method. Additionally, we use the LDestimate-code which implements the Abry-Veitch method [12]. In the case of the SELFIS tool we applied our own linear regression on the intermediate results obtained from SELFIS, as explained above. This is necessary as SELFIS estimates H , always based on a linear regression over all available points, which can heavily bias the results due to outliers at the end points. For this reason we shortened the intermediate statistical results to all but the first 2 and the last 2 aggregation levels. In all cases, applying the manual regression, the fit is better than the SELFIS tool directly provides. The differences in the values of H , between manually applied linear regression and the final results of SELFIS, are some-

¹ More precisely, based on the wavelet coefficient $d_{j,k}$ the amount of energy $|d_{j,k}|^2$ in the signal at about the time $t=2^j k$ and about the frequency $2^{-j} \lambda$ is measured. $E[|d_{j,k}|^2]$ is the amount of energy at octave j . If the initial resolution of the time series is t_0 (bin size), the time resolution at each scaling level j is $t_j=2^j t_0$.

times quite large, which stresses the importance to apply manual post-processing. The LD estimator function for the Abry-Veitch test suggests an optimal starting point for j which we always used. It furthermore plots the scaling behavior over all octaves j , allowing to judge on the type of scaling. We discuss those results as well.

4. Data Traces

In cooperation with Vodafone we conducted measurements in two live GPRS networks. We captured the IP packet headers at the Gi interface for all users in a geographical region of the operator's GPRS network. The Gi interface connects the mobile network to external packet switched networks (like the Internet, a corporate Intranet or Email and WAP/MMS proxies). All IP packets from or to mobile terminals traverse this interface. The traces we focus on were taken during summer 2003.

We collect for every packet crossing the network monitor interface a time stamp, with $1\mu\text{s}$ accuracy, with the total length of the packet in bytes. For our further investigations we generate from the original trace the Packet Arrival Process (PAP) as a discrete time-series process by counting the number of packets and the Data Volume Process (DVP) as the total number of bytes within a time interval (bin) or 100 ms.

We are interested in the scaling behavior for the aggregated traffic, WAP oriented traffic and WEB oriented traffic. For this purpose we look at three "sup"-sampled traces. Firstly, we investigate the total aggregated traffic (up and downlink traffic), which we measured on the Gi interface. Next, we have split-up the traffic into WEB oriented traffic and WAP oriented traffic. We do this by splitting the data according to the APN (Access Point Names) they are belonging to. GPRS allows users to attach to different APNs provided by the operator. Typically, the APNs are used for different types of traffic and split logically the traffic on the Gi interface. We checked our measurements for the used applications and found that most of the traffic on one APN consists of WEB like applications, including HTTP, FTP, Email, etc. On the other APN we see mostly WAP like traffic, comprising WAP and MMS.² However, the APNs do not filter for applications, hence it occasionally happens that we see WEB traffic on the WAP APN and vice versa.

On one of the network measurement points (Vfe1) the traffic splits up into 25% WEB-APN and 75% WAP APN traffic, while in the other network (Vfe2) the split is 70% WEB APN and 30% WAP APN traffic.

We traced several 24-hour periods and investigated appropriate busy-hour periods spanning several hours, for each of the networks. Our results will be presented only for one selected day for each network. In the case of Vfe1 we chose a 110-minute busy-hour period in the afternoon, in the case of Vfe2 we chose a 430-min (7-hour) busy-hour period covering the whole afternoon. Since all estimator methods (with exception of the Abry-Veitch method) require stationarity and often are very sensitive to underlying trends or correlations in the traffic process [13], we investigated the chosen busy-hour periods on trends by plotting the moving average, on periodicity by investigating the Fourier transformation, and on stationarity by applying the average run

² We used the same investigation method we have used in [9] for differentiating the applications.

test. All tests indicated the suitability of the chosen periods for the estimation methods, i.e., they appear to be stationary and they have no visible periodicities or trends.

5. Results on Self-Similarity

We present for all processes detailed results obtained by the Abry-Veitch method, together with Hurst estimations acquired by the other described methods. Table 1 and Table 2 show the results for the estimated Hurst values by the Abry-Veitch method for the packet arrival process. The second row 'H' contains the Hurst values and in the third row marked 'conf.' the confidence values for the estimated Hurst values are listed. Figure 1 and Figure 2 illustrate the Hurst values obtained from the different methods, for comparison. All values are very similar with Hurst values of about 0.8 and higher. This strongly suggests long range dependency for the PAP for both networks Vfe1 and Vfe2.

Table 1. Results for PAP in Vfe1

Vfe1	Agg	WEB	WAP
H	0.86	0.83	1.06
Conf.	[0.76,0.95]	[0.74,0.92]	[0.96,1.14]
Scaling	Fig. 8	Fig. 8	Fig. 7

Table 2. Results for PAP in Vfe2

Vfe2	Agg	WEB	WAP
H	0.90	1.02	0.89
Conf.	[0.81,0.98]	[0.93,1.11]	[0.79,0.97]
Scaling	Fig. 5	Fig. 5	Fig. 5

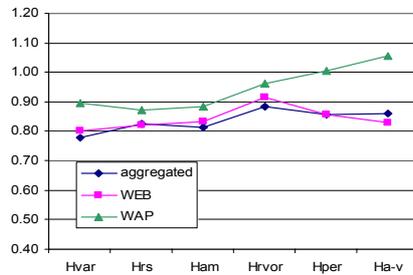


Fig. 1. All Hurst values for PAP and Vfe1

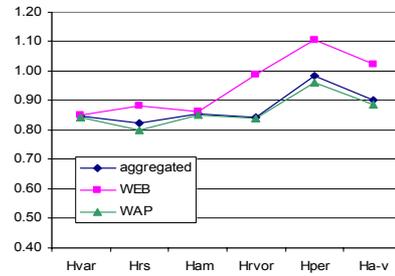


Fig. 2. All Hurst values for PAP and Vfe2

'Hvar' stands for Variance Method, 'Hrs' for R/S Method, 'Ham' for Absolute Moment Method, 'Hrvor' for Variance of Residuals, 'Hper' for Periodogram Method and 'Ha-v' for Abry-Veitch Method

In some cases the Hurst value is above 1, actually precluding LRD, but in the case of the Abry-Veitch method all confidence intervals include also a value of below 1. Furthermore, as explained next, inspection of the output-plots suggests in such cases still asymptotic self-similarity.

For the DVP, Table 3 and Figure 3 list the Hurst values for Vfe1 and Table 3 and Figure 4 list the Hurst values for Vfe2. Again all results strongly indicate LRD.

All estimation methods, except for the Abry-Veitch method, assume an LRD model beforehand. That is, the Hurst estimation value can only be regarded as correct if the assumption of an LRD process holds. In contrast, the Abry-Veitch test is not based on such assumptions. It shows the scaling of the process for all time scales in the diagram. Only by interpreting the results in this Logscale Diagram, the true nature of the process is determined (e.g., self-similarity, long-range dependency, multiscaling) [10]. For our traces we have encountered four basic log-scale diagram types, depicted in Figure 5 to Figure 8. In Table 1 to Table 4 we list in the 4th row for each process the plot that comes closest, respectively. The individual plots looked very similar to the exemplary plots, but with different scales on the y-axes.

Table 3. Results of DVP for Vfe1

Vfe1	Agg	WEB	WAP
H	0.69	0.68	0.92
Conf.	[0.65,0.72]	[0.64,0.71]	[0.88,0.96]
Scaling	Fig. 6	Fig. 6	Fig. 7

Table 4. Results of DVP for Vfe2

Vfe2	Agg	WEB	WAP
H	0.82	1.07	0.81
Conf.	[0.73,0.90]	[0.98,1.15]	[0.72,0.89]
Scaling	Fig 5	Fig. 7	Fig. 5

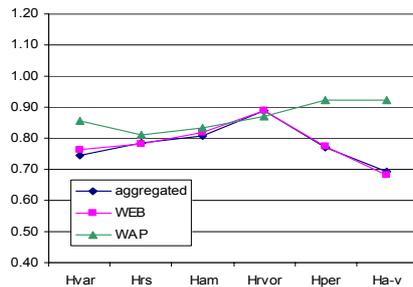


Fig. 3. All Hurst values for DVP and Vfe1

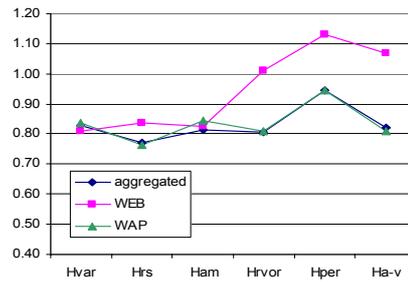


Fig. 4. All Hurst values for DVP and Vfe2

'Hvar' stands for Variance Method, 'Hrs' for R/S Method, 'Ham' for Absolute Moment Method, 'Hrvor' for Variance of Residuals, 'Hper' for Periodogram Method and 'Ha-v' for Abry-Veitch Method

Figure 5 and Figure 7 both show a typical plot for LRD traffic. The second-order scaling starts from a certain scaling point on and continues until the largest available scale in the trace. For small scales we do not see second-order scaling behavior. We have found this scaling behavior for all WAP processes.

Figure 6 also exhibits LRD scaling, but actually has two scaling regions. One from approximately 1 to 8 and one from 8 to maximum scale. This is called bi-scaling. We have found this in the case of DVP for Web traffic in Vfe2. Figure 8 depicts the case where the process has second-order scaling over all scales. This indicates strictly second-order self-similarity. We see this in the case of PAP for WEB traffic also in Vfe2.

We point out some interesting observations (cf. Table 1 to Table 4). Firstly, the Hurst value of the aggregated traffic is always very close to the Hurst value of the ma-

jority of the traffic. This is in agreement with [14]. In the case of Vfe1 the major part of the traffic is WAP traffic, in the case of Vfe2 it is WEB traffic. More in detail, even the whole scaling behavior, as depicted by the Logscale Diagram, is very similar between the majority traffic and the aggregated traffic. This implies, by knowing the scaling of the majority traffic, one obtains also the scaling of the aggregated traffic. Secondly, the minor traffic has always slightly higher Hurst values, with changing roles of WAP and WEB in the cases of Vfe1 and Vfe2. We do not have an explanation for this, yet. One reason might be that even so we applied the estimation methods on separated traces per APN, it is not possible to separate them truly: they have been both traveling together through the GPRS network, thereby most likely affecting each other.

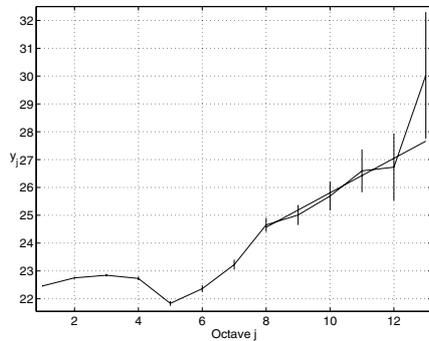


Fig. 5. Typical plot for processes showing long range dependency. Below a certain scale no regular linear scaling exists. The linear part is at $j=8$ divided in two scaling regions

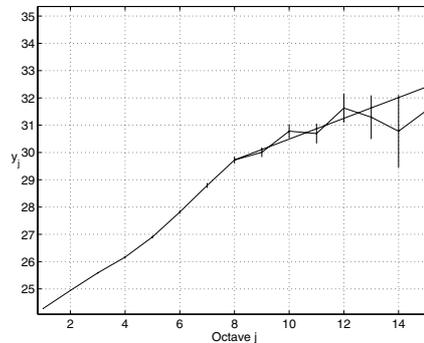


Fig. 6. Typical plot for processes showing bi-scaling. The second scaling region starts at $j=8$. It also implies long range dependence

An interesting question arises whether it is actually possible to truly identify the Hurst value for each type of traffic separately. In [14] the authors have shown that non-self-similar UDP traffic is affected by self-similar TCP traffic. But this effect is only strong if the self-similar traffic is the major traffic. In our case we observe high Hurst values even in the case the WAP (UDP) traffic is dominating. This suggests that WAP traffic itself is strongly long-range dependent.

Furthermore, we see that WAP traffic has a very different scaling behavior for small scales compared to WEB traffic. The reason for the different small scaling behavior can be assumed to be the very different transport mechanism of TCP and WAP over UDP. As we explained in the footnote to the Abry-veitch method, the octave j in the logscale diagram also expresses the timescale of the aggregated process. Based on this insight it is possible to determine over which timescales scaling occurs. We start with an initial bin size of 100 ms, which leads to values of $t_j=0.1, 0.2, 0.4, 0.8, 1.6, 3.2, \dots$ for $j=1, 2, 3, 4, 5, \dots$, respectively. For all processes in which we observed a scaling like the one depicted in Figure 5 and Figure 7, the knee-point at which the linear scaling starts is at about $j=5$ or $j=6$, i.e., respectively 1.6 seconds or 3.2 seconds. In [16] the authors show that the average download time per WAP page, including

embedded objects, is in the order of 1.5-3 seconds. Hence, this timescale marks the demarcation line between the WAP transport layer protocol and the user behavior.

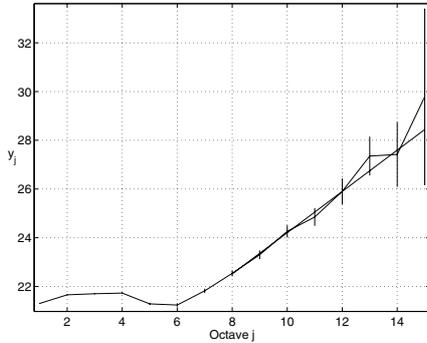


Fig. 7. Typical plot for processes showing long-range dependency. Below a certain scale no regular linear scaling exists

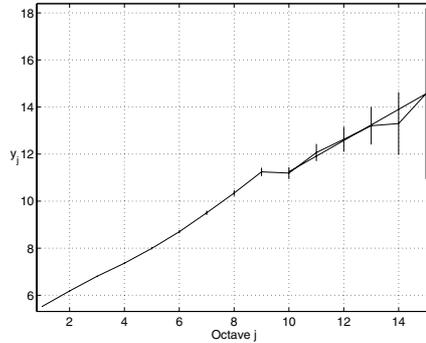


Fig. 8. Linear scaling over the whole range. Although there is a step at $j=9$, the slope on both sides is almost the same. This indicates second-order self-similarity

6. Conclusion

Based on live GPRS measurements, we applied 6 different established Hurst estimation methods, including the comprehensive and robust Abry-Veitch test on packet arrival and data volume processes. We showed the Hurst value as well as the scaling behavior for the busy-hour for 3 different traffic types and two different networks. In the case of aggregated traffic and also in the case of individual WAP and WEB traffic traces, the results strongly suggest long-range dependency.

We confirmed that the dominant traffic type (WAP or WEB) determines the degree of self-similarity of the aggregated traffic. In our case however, the minor traffic always exhibits a Hurst value larger than the Hurst value of the major traffic. This is particularly interesting as WAP traffic is based on UDP.

We showed that WAP traffic has a very different scaling behavior compared to WEB traffic for small scales. We identified the demarcation line between small and larger scales for WAP to coincide with the average page duration. Larger scales can probably be accounted to user behavior.

As future research we wish to investigate the reason(s) why the minor traffic fraction exhibits higher Hurst parameter values for all tests.

Further research also includes investigating the reason(s) of self-similarity of WAP traffic and the exact nature of scaling; whether it is multifractional, monofractional, strictly self-similar or ‘just’ long-range dependent.

Acknowledgements

We thank Vodafone for providing us with live GPRS traces from various networks in Europe. We also thank Thomas Karagiannis, Michalis Faloutsos, the authors of the

SELFIS tool, and Darryl Veitch, the author of the LD estimator-code, for making their software available on the Internet.

References

- [1] W. Willinger, M. S. Taqqu, and A. Erramilli, "A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks," *Stochastic Networks: Theory and Applications*. In Royal Statistical Society Lecture Notes Series, Oxford University Press, 1996, vol. 4, pp.339–366.
- [2] V. Paxson and S. Floyd, "Wide Area Traffic: The Failure of Poisson Modeling," *Proceedings of ACM SIGCOMM'94*, 1994.
- [3] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEEE/ACM Transactions on Networking*, 1994, vol.2 no.1, pp.1-15.
- [4] A. Erramilli, O. Narayan, and W. Willinger, "Experimental queueing analysis with long-range dependent packet traffic," *IEEE/ACM Transactions on Networking*, New York, NY, Apr. 1996, vol.4, pp.209–223.
- [5] A. Erramilli, O. Narayan, A. L. Neidhardt, and I. Saniee, "Performance impacts of multi-scaling in wide-area TCP/IP traffic," *Proc.*, IEEE INFOCOM 2000, Tel Aviv, Israel, 2000, vol.1, pp.352–359.
- [6] I. Norros, "A storage model with self-similar input," *Queueing Systems*, 1994, vol.16, pp.387–396
- [7] M.E. Crovella and A. Bestavros A., "Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes," *Proceedings of ACM SIGMETRICS '96*, 1996.
- [8] A. Feldmann, A.C. Gilbert, P. Huang, and W. Willinger, "Dynamics of IP traffic: A study of the role of variability and the impact of control," *ACM SIGCOMM '99*, 1999.
- [9] R. Kalden, T. Varga, B. Wouters, B. Sanders, "Wireless Service Usage and Traffic Characteristics in GPRS networks", *Proceedings of the 18th International Teletraffic Congress (ITC18)*, Berlin, 2003, vol.2, pp.981-990.
- [10] P. Abry, Flandrin, M.S. Taqqu, D. Veitch, "Wavelets for the analysis, estimation and synthesis of scaling data." In: K. Park and W. Willinger, Eds., "Self Similar Network Traffic Analysis and Performance Evaluation," Wiley, 2000.
- [11] T. Karagiannis, M. Faloutsos, "SELFIS: A Tool For Self-Similarity and Long-Range Dependence Analysis," 1st Workshop on Fractals and Self-Similarity in Data Mining: Issues and Approaches (in KDD), Edmonton, Canada, July 23, 2002.
- [12] Darryl Veitch's home page: <http://www.cubinlab.ee.mu.oz.au/~darryl/>
- [13] T. Karagiannis, M. Faloutsos, "Long-Range Dependence: Now you see it, now you don't!" CSE Dept., UC Riverside, Rudolf H. Riedi, ECE Dept., Rice University. In: *IEEE Global Internet*, 2002.
- [14] K. Park, G.T. Kim, and M.E. Crovella, "On the Relationship Between File Sizes, Transport Protocols, and Self-Similar Network Traffic," *Proceedings of IEEE International Conference on Network Protocols*, 1996, pp 171-180.
- [15] A. Popescu, "Traffic Self-Similarity", *ICT2001*, Bucharest, 2001.
- [16] Internal communication, Ericsson ETH, Budapest, 2003
- [17] R. Kalden, H. Ekström, "Searching for Mobile Mice and Elephants in GPRS Networks," submitted for publication, 2004.
- [18] W. Willinger, V. Paxson, M. Taqqu, "Self-Similarity and Heavy Tails: Structural Modeling of Network Traffic", in "A practical Guide To Heavy Tails: Statistical Techniques and Applications", Birkhauser, Boston, 1998