ON THE MULTIRESOLUTION STRUCTURE OF INTERNET TRAFFIC TRACES

Konstantinos Drakakis  
University of Edinburgh

Dragan Radulovic  
Florida Atlantic University

ABSTRACT

Internet traffic exhibits a rich autocorrelation behavior, responsible for curving the Energy/Averaging function. We show that the traffic exhibits variations of its details in many different time scales (multiresolution structure), which can account for this feature. We relate the curving of the Averaging function to such “interesting” time scales, show that it is possible to “read” these time scales (levels) directly off the Averaging function, and propose some methods to accomplish that.

KEYWORDS

Internet traffic, multiscale model, level, Slow Start, self-similar, spikiness, RTT, Energy function, Averaging function

1. INTRODUCTION

1.1 The problem

The object of this paper is to investigate how the structure of Internet traffic varies across different time scales, propose tools that can monitor such a multiresolution structure, and finally propose a model, the Multiresolution Model (MM), that can simulate this structure. By “Internet traffic” we denote here the sum of a collection of user sessions, each of which is of the form \( \sum_{i=1}^{N} d_i \), where the \( d_i \)s are information quantities (e.g. in bytes), and the \( t_i \)s are the times at which these quantities are sent (for details see [33]). It is more convenient to consider the traffic in binned form: we choose a time bin \( \Delta t \), and we define \( X_{j}^{t} = \sum_{(j-1)\Delta t < t_j < j\Delta t} d_j \), i.e. as the sum of all packet sizes of the packets whose emission times fall in the bin.

Our work in this paper will build upon our previous work in Internet traffic modeling, and more precisely on the Discretized Self-Similar Model (DSSM) [32], and on the Averaging and Energy function [34]; we will expect our readers to be familiar with the terminology and the basic results contained therein. As the spikiness of the traffic is, up to a good approximation, independent from its autocorrelation, we will not need to deal with spikiness, and hence with the Discretized Self-Similar Model with Slow Start (DSSMSS) [33], before the very end of our paper, where we offer a model combining all the features of the aforementioned models. Note that the version of the Energy and Averaging functions we will use are improvements (more stable numerically) of the well known Energy function [15, 16].

We will see that network activity at a particular time scale causes the Averaging function to curve around this time scale, so that, by measuring the curvature, the Averaging function can reveal which time scales a traffic session is active at. This permits the simulation of traffic active at particular scales; it also has further practical applications, e.g. in network security. It is quite remarkable that the time scales which individual user sessions are active in are common for all sessions, and therefore do not vanish in the aggregate traffic, but are, on the contrary, one of its most salient features. This makes these time scales (levels) a feature of the network rather than of individual users.

1A preliminary version of this paper appeared in Proceedings of SPIE Vol. 4868 (2002) under the title: “Some results on the multiresolution structure of Internet traffic traces”.

2The author is a scholar of the Lilian Boudouris Foundation. Web page: www.maths.ed.ac.uk/~drakakis

3The largest portion of the research for this paper was conducted while the author was affiliated to Princeton University, and was partially supported by AT&T Research Center.
1.2 Description of the data sets and the simulations

The data sets used in this paper are the same as the ones we have used in our previous work (see [33] for details): we use 8 data sets denoted by 89, 94, 97, U8, L4, L5, M6, M7. A set of simulated traces will also be used: 0-1, ARR (standing for ARRival process), RH (standing for Random Heights), and and RH HT (standing for Random Heights + Heavy Tails) denote simulations according to the SSM [32], the DSSM [32], the Rewards Model [10, 11], and the α-β Traffic Model [23], respectively; ARRRH (standing for ARRival process with Random Heights) is a combination of RH and ARR; EXP IID is a sequence of i.i.d. exponential variables, and, similarly, HT IID is a sequence of heavy-tailed i.i.d. variables.

Simulations according to the MM will be labeled by the time scales of their levels, and an S will be added if the levels are “sharp”, e.g. 8, 12S, 7S/12/17 etc (all relevant definitions will be given below).

1.3 Existing models and the MM

A feature of the traffic which guided a substantial part of our work is the curving of the Energy function [15, 16]. Most of the existing models (see [33] for a small list), with the notable exception of the Multifractal Models [21, 22, 15, 3, 30], produce linear or piecewise linear Energy functions. As far as we know, there exists no model that produces traces a) that have the spikiness and long-range dependence of real traffic traces, b) whose marginals match the behavior of the observed marginals, which are sometimes close to and at other times deviate far from Gaussianity, and c) with a curving Energy function. In this paper, we will first work on c), and the result of our work will be the MM; subsequently, we will combine the MM and the DSSMSS into one model, the Multiresolution Model with Slow Start (MMSS), which will satisfy all three requirements.

2. THE MM

2.1 Introduction

The DSSM and DSSMSS were able to capture and explain crucial features of Internet traffic, such as its long range dependence, its different behavior in different time scales, and its spikiness (see [33]). Unfortunately, the Energy function of the DSSMSS is still far from reality (compare Fig. 4 and 1). This suggests that pieces of the puzzle are still missing.

2.2 A closer look into Averaging functions

So far, the simulations failed to introduce a substantial curvature into the Averaging or Energy function (Fig. 1, 5), although the processes used spanned a variety of attributes (heavy/light-tailed marginals, independence/long-range dependence, etc.), and even combined them (as in the α-β Traffic Model). It was precisely this characteristic that stimulated research of “multifractality”, questioned the accuracy of traffic description through Poisson or self-similar processes, and ended up demonstrating that the traffic is a fundamentally different process from these two [23, 16, 15].

The Averaging functions of the data traces (Fig. 5) are truly remarkable: recall (from [34]) that a slope \( \alpha \) at time scale \( \log_2(n) \) indicates an order of magnitude of \( n^{-\alpha} \) for the difference between two mean value estimators of \( n \) terms; if \( \alpha \) depends on \( n \), the process is not self-similar. This is the case for all of the traces. Moreover, one expects, if the \( X_i \) are asymptotically independent, that the Averaging function will decrease monotonically as a function of \( n \), at least for large \( n \). This is not the case for our data. For instance, the Averaging function of trace 94 is constant in the region from \( n = 2^8 = 256 \) till \( n = 2^{15} = 32768 \). In other traces the Averaging function even increases with \( n \): for example, trace 97 exhibits a little “bump” at time scale 13 (Fig. 5(a)); simulations carried out later in this section achieve an even more spectacular result (Fig. 8). We now know that the \( X_i \) in our traces cannot really be considered asymptotically independent, at least relatively to the trace duration, and we interpret the “deviant behavior” of the Averaging function as a
manifestation of this lack of independence. Why should it show up much more at some scales than in others? This is the question we address in this section.

2.3 Averaging functions and multiscale structure of the traffic — levels

There are still two important aspects of Internet traffic which we have overlooked throughout the discussion of the DSSMSS:

On the one hand, the exact attributes of time intervals between packet emissions were hitherto supposed to be of little importance: in the theorems, they were assumed to follow any distribution of finite variance, and simulations were performed with an exponential one. The exact distribution has indeed been an object of study, and it turned out that the Poisson model for a single user fails [1], and that (bi)Pareto, Weibull, and empirical (TcpLib [1]) distributions come into play [8].

On the other hand, traffic seems to comprise many phenomena that are hierarchically structured: for example, users organize texts in chapters, chapters in paragraphs, paragraphs in sentences etc., thus one expects a variety of durations of pauses and activity, which will certainly have an effect on traffic. Also, during Web browsing, users read, process the information, and act accordingly. Finally, protocol activity in the network spans many time scales: congestion control is performed every fraction of a second, but routing information is exchanged every half an hour, or so.

It appears then that many time scales are of importance, each for a different reason, but all because some kind of activity is taking place there. Their exact attributes are still unclear, to the best of our knowledge, although they are generally accepted to exist [1, 3, 18], as the result of both user behavior and protocols. This activity in multiple time scales, however, can alter completely the autocorrelation of the traffic, introducing strong correlation in some time scales and weak in others (see Fig. 2 and 3). This, in turn, will impact the Averaging function as well (see [34]). Consequently, it seems worth investigating whether the curving of the Averaging function has its roots in this multiscale activity.

We now proceed to give a formal definition of levels. Notice that the definition makes some assumptions, which need to be validated on a particular trace, before we can conclude that it contains levels. Our framework will be a discrete-time trace, originating from a binned trace, in which we associate the value 0 to a bin if no packet emission takes place within it, and 1 otherwise, and then map the bins to discrete points; the terms “ON-” and “OFF-interval” are used here as in the DSSM, so that an ON/OFF-interval is a maximal time interval of consecutive 1/0 values.

**Definition**: Assume that there exists \( n > 0 \), so that each OFF-interval in a discrete-time trace is the realization of one among \( n + 1 \) PDFs, which have finite first moment. Order the PDFs in order of increasing mean, and index them starting with 0. Assume furthermore that intervals of the 0th PDF are separated by ON-intervals of length 1. Then, define the \( i \)th level, \( i = 1, \ldots, n \), to be an ordered pair of PDFs, of which the first is the \( i \)th PDF of our assumption, and the second is the empirical PDF of the length of the ON-intervals, which are present in the trace after the substitution of all values of 0 belonging to OFF-intervals generated by the first \( i \) PDFs of our assumption by 1s. As a special case, the 0th level will comprise the 0th PDF of our assumption and a trivial delta distribution on 1, and will be called the “RTT” level.

This definition may sound strange to a mathematician, but should not surprise a computer scientist: indeed, this is how we would build the “Level” object in an object-oriented language. The key part of the definition is the two PDFs for each level; in order to determine them, we will use the Interval Detection Algorithm (see the Appendix). From now on, we will be quite liberal with the use of the term “level”, identifying it occasionally with any of its constituent PDFs (more often with the first), when the other is not of interest to us, or even with their (or any of their) mean values.

2.4 Evidence for multiscale activity

A useful and simple tool for the detection of these time scales (levels) in the traffic, in the original continuous setting, is the characteristic function of a connection⁴. A connection is defined to consist of the entries of the trace corresponding to a particular value of sender host, receiver host, sender port, and receiver port (see [33]). The characteristic function of a connection is a function of time, equal to 1 if that particular time is a time stamp belonging to the connection, and 0 otherwise.

⁴Similar diagrams exist in [1].
Characteristic functions seem to have a fractal structure (see Fig. 6 and 7); moreover, analysis of the characteristic functions of the 50 largest connections of trace 94, and connections of trace M6, leads to the conclusion that levels of connections do not average out in the aggregate traffic, but, on the contrary, have a strong influence on it. Capturing this effect will be the main goal of this section.

2.5 A Theorem on the MM

The theorem below implies that the covariance structure of total traffic is strongly dependent on its levels. The reader is referred to our previous work [32, 33, 35] for the rigorous definition of all the quantities not defined in this paper.

**Theorem.** Let \( A^2(W) = \varDelta^{-1}\left[\text{Var}\left(\int_0^\Delta W(t)dt\right) - \text{Cov}\left(\int_0^\Delta W(t)dt, \int_0^{2\Delta} W(t)dt\right)\right]^{1/2} \) and let \( \tilde{W}_{d,i}(t) \) be the same as \( W_{b,i}(t) \) but with \( O_{b,j}(i) = 0 \) a.s. and \( O_{b,n}(i) = \infty \) a.s. Also, take \( \tilde{W}_{d,i}(t) \) to be the same as \( W_{a,i}(t) \) (notice that \( \tilde{W}^d \) represents a continuous traffic (no spikes traffic), while \( \tilde{W}^d \) represents a discrete traffic, where RTT is present (spikes traffic)). Then:

- \( A^2(\tilde{W}^d_1) \approx \Delta^{-1/2} \rightarrow 0, \) as \( \Delta \rightarrow \infty \)
- \( A^2(\tilde{W}^d_1) \rightarrow 0, \) as \( \Delta \rightarrow \infty \) or \( \Delta \rightarrow 0, \) and there exists a constant \( K > 0 \) such that \( A^2(\tilde{W}^d_1) \geq K, \) for \( D = E(O^{b,n} + O^{b,f})/2 \)

2.6 Explanation of the curvature of the Averaging/Energy function

The following argument explains the link between the theorem and the curving of the Averaging/Energy function. Namely, assume there is a level at time scale \( j_0 \), and no other levels nearby. It will be assumed that

- For \( j < j_0 \) and \( 2^j \gg RTT \), the ON-intervals appear continuous, and can be modeled by no spikes traffic \( \tilde{W}^d \). Consequently, part b) suggests that the Averaging function increases.
- For \( j > j_0 \), the ON-intervals again appear continuous, so part b) suggests that the Averaging function decreases, as \( j \rightarrow \infty \).

Because the Averaging function is strictly positive, the points above suggest that it must have a maximum around the level \( j_0 \). This has indeed been confirmed by simulations (Fig. 8(b)).

However, the Averaging function of real traces initially decays, unlike the simulations in Fig. 8 (b). Part a) of the theorem explains this behavior as well:

- For \( j_0 < j \) and \( 2^j \approx RTT \), the ON-intervals appear discrete and can be modeled by spikes traffic \( \tilde{W}^d \). Part a) takes over and ensures White noise-like behavior, i.e. \( \log_2 A_j \approx -j/2 \).

In case more than one levels are present, the argument above can be repeated, the role of RTT being now played by the size of the previous (finer) level. Therefore, the Averaging function should initially decrease, then change slope around the level, and have a “bump” around every level after the first. Simulations indeed verify this statement (Fig. 8(a)). The exact form of these “bumps” has to do with the “sharpness” of the level, i.e. with the variance of the durations of ON- and OFF-intervals. Very sharp levels, i.e. with small variance, seem to lead, at the time scale of the level, to the phenomenon of superconvergence, i.e. convergence to the mean value faster than what the CLT predicts, which manifests itself through a slope less than -0.5 (Fig. 8).

\(^5\text{In real traces, we observed this behavior in the fine and middle scales of traces M6, M7, U8, and 89. See also 8S in Fig. 8.}\)
We would like to emphasize that we have introduced levels in our model consistent with our wish for “local” models (in the sense that the sessions are constructed by putting together appropriate building blocks, rather than the “global” approach of e.g. Random Cascades.) Our inspiration for the model was taken from direct inspection of observed traces; it does however differ from our earlier models (DSSM, DSSMSS) in that we cannot identify a particular feature in user behavior or network protocol as their cause. The scales of the levels, their sharpness, and their possible evolution in time may well be of interest to network engineers. For this reason we propose tools to extract this info, in section 4, that are faster than the IDA. Note also that heavy-tailed distributions were not needed in this construction: long-range dependence and levels thus appear to be completely decoupled.

Now that we have understood the role levels play, we can explain the shape of the Energy and Averaging functions in the case of Slow Start simulations (Fig. 4), where the two functions are actually piecewise linear, not linear. The change of slope occurs between scales 7 and 9. But the mean value of the connection size was 5, and the mean RTT was 100, which implies that the mean total duration of the session was $92 \approx 500$, and this is exactly the level of the traffic.

One could argue that if the simulations of Fig. 5 and 1 were repeated with different parameters, some sort of curving might be achieved. This is true, to some extent (Fig. 9), but the curvature achieved is very slight and still far from real traces.

Finally, let us reexamine Fig. 5, armed with this qualitative understanding of levels. Trace 94 seems to have several levels between the time scales 8 and 15, and maybe around 20. Trace 97 has probably just one strong level at 13; trace L4 has one level at 8 and one at 15; trace L5 seems to have levels all over the place, but mainly one at some small scale (between 0 and 4), one at 8 and one at 15. In the next session, tools will be developed to quantify such guesswork.

### 2.7 Levels with unequal OFF- and ON-intervals

Although the theorem does not require the OFF- and ON-intervals to follow the same distributions, the discussion in section 2.6 considered exclusively equally distributed OFF- and ON-intervals, in order to enhance clarity. As we have seen so far, though, (for example, see Fig. 6 and 7), different distributions seem to be the rule rather than the exception, with the ON-intervals being typically quite larger than the OFF-intervals, so we need to determine how the Averaging function will respond to such a situation. Fig. 10(a) shows the Averaging function of a simulated session containing just one level whose ON-intervals have lengths around $2^{15}$, but whose OFF-intervals have average lengths that are 1, 2, 4,... times smaller. Observe that the Averaging function appears to “lock” on the OFF-intervals, the existence of the larger ON-intervals being only alluded to by the size of the “bumps”, which are more prolonged than usual. Observe also (in Fig. 10(b)) that, if the lengths of the OFF-intervals are much smaller than the lengths of the ON-intervals, and their distributions sufficiently diffused, it is possible for the bump to become negligible, and thus for the Averaging function to give the impression, by visual inspection, that no level exists.

### 3. THE COMPLETE MODEL: THE MULTiresOLUTION MODEL WITH SLOW START (MMSS)

It was mentioned earlier that spikiness and levels are two phenomena highly uncorrelated, and that, as such, they can be studied independently. A successful simulation, though, should contain both, so our model would not be complete without an algorithm that combines spikiness and levels in a simulated trace.

The independence of the two features suggests that this combination can be done trivially. Here is an algorithm:

1. Generate a vector of isolated 1s separated by intervals of 0s; these intervals represent the RTTs and should be independent and identically distributed.
2. Set the Slow Start maximum at M, and consider two indices, $i$ and $j$. Set $j \leftarrow 1$. Then:
   (a) Generate a heavy tailed integer random variable $N$.
   (b) Move $i$ to the next 1 value of the vector (generated by Step 1). If end of vector found, then stop.
(c) Set the vector value at \( i \) equal to \( j \), set \( N \leftarrow N - j \), set \( j \leftarrow \min(2j, M) \). If \( N \leq 0 \), go to Step (a) and set \( j \leftarrow 1 \).

(d) Go to Step (b).

3. Generate a vector of alternating intervals of value 1 and 0. These intervals should also be identically and independently distributed, but also cross-independent. The two generating distributions, though (for intervals of value 1 an 0), need not be the same.

4. Repeat the previous step as many times as necessary, by choosing distributions with progressively smaller mean. Simulated traffic with \( n \) levels will result, if the step is repeated \( n \) times.

5. Multiply the vectors obtained above.

All Step 2 does is to supply the algorithm with a Slow Start background. The shortcoming of this construction, though, is that the simulated Slow Start sessions in the final simulated trace do not necessarily correspond to sessions that could have possibly been generated in practice; indeed, the later steps of the algorithm may “chop” the beginning or the middle of a Slow Start session, and obviously such behavior cannot exist in practice. So, at this point, our construction deviates slightly from the actual network mechanisms, but this does not affect the results at all.

It is possible, however, to modify the algorithm slightly and remove these artifacts. Just remove Step 2, and add the following step at the end this time:

6. Set the Slow Start maximum at \( M \), and consider two indices, \( i \) and \( j \). Set \( j \leftarrow 1 \). Choose a number of consecutive levels, starting with the finest one, and label them as “RTT levels”. Then:
   
   (a) Generate a heavy tailed integer random variable \( N \).
   
   (b) Move \( i \) to the next 1 value of the vector (produced at the end of Step 5). If end of vector found, then stop.
   
   (c) If the gap between the previous and the current value of \( i \) does not belong to an RTT level, generate a new heavy tailed integer random variable \( N \) and set \( j \leftarrow 1 \).
   
   (d) Set the vector value at \( i \) equal to \( j \), set \( N \leftarrow N - j \), set \( j \leftarrow 1 \). If \( N \leq 0 \), go to Step (a) and set \( j \leftarrow 1 \).
   
   (e) Go to Step (b).

Here, the “RTT level” OFF-intervals play the role of additional RTTs that a session may encounter. This additional step in the algorithm essentially identifies whether a gap should be considered a RTT or an OFF-interval: in the former case, the Slow Start session continues over it normally; in the latter, the current Slow Start session ends, and a new one starts.

4. BURSTINESS AND LEVEL DETECTION TOOLS

The theorem proved earlier states that the activity of the traffic in different time scales leaves a trace on the Averaging function; therefore, the Averaging function can be used to detect the level structure in a particular data set. This would allow for monitoring of user behavior without resorting to a connection-wise study of the trace, which is a very inefficient technique. We propose the following two algorithms:

Given the binary logarithms of the Averaging function: \( \log_2(A(0)), \ldots, \log_2(A(n-1)) \), we compute slopes \( s_i = \log_2(A(i)) - \log_2(A(i-1)) \), \( i = 1, \ldots, n-1 \) and slope changes \( |S_{i+1} - S_i|, i = 1, \ldots, n-2 \). These last quantities are equivalent to second derivatives, which, as is well known, measure the local curvature: as we argued in our discussion of the theorem, the Averaging function has “bumps” around its levels, thus any curvature indicates the presence of levels. There is, though, one more point: it is not really important how much adjacent slopes differ, but how much they differ with respect to their absolute magnitudes. It seems then more appropriate to consider: \( \bar{Sc}_i = \frac{|S_{i+1} - S_i|}{\max(\min(\frac{|S_{i+1}|}{|S_i|}, \frac{|S_i|}{|S_{i+1}|}), \epsilon)}, i = 1, \ldots, n-2 \), where \( \epsilon \) is a small positive number, which will safeguard against division by 0, or an extremely small number (we used \( \epsilon = 0.01 \)).

**Tool 1: (Levels detector)** \( L(i) = \bar{Sc}_i, i = 1, \ldots, n-2 \). The local maxima are considered to be levels.
Since White Noise leads to an Averaging function whose binary logarithm is just a straight line of slope -0.5 (see [32]), a much simpler idea would be to declare that a level exists where the slope becomes too large, as a signed quantity (ideally, it should become positive). How large it should be exactly is, in view of the lack of an accurate method to determine this, a matter of taste. We think that, due to the “noise” induced by the finite number of measurements available, it is very hard to distinguish a decay of type \( \text{type } \{ t \}^{0.1} \) from no decay at all, thus we set a threshold at -0.1.

**Tool 2: (Detecting flat regions)** \( F(i) = 1, i \geq -0.1, i = 1, ..., n - 1 \). Levels are considered to exist within flat (or concave) regions where \( F(i) = 1 \).

The results of Tools 1 and 2 are shown in Fig. 111, 12 and 13 for the real traces and some of the simulations, respectively.

Tools 1 and 2 seem to reveal more or less the same information, as to where the levels are, but Tool 1 is a bit more detailed than Tool 2, in showing how “prominent” levels are (Fig. 111 and 12). The accuracy of the tools was tested by their results on simulated data (Fig. 13): they nicely match the curvature of the Averaging function for real traces, and they perform reasonably well with the simulations, where levels were artificially induced. Before closing this section, let us submit Tool 2 to a final (and hard) test: can it detect correctly the levels of the trace 94 session presented in Fig. 6? This test is hard because of the substantial difference between OFF- and ON-interval lengths (see section 2.7). The Averaging function does not show any curvature comparable to the levels observed, but Tool 2 actually performs well, as Fig. 14 demonstrates: it detects levels near 2\(7 \text{ ms} \), 2\(11 \text{ ms} \), and 2\(16 \text{ ms} - 2\(19 \text{ ms} \), which are very close to the ones mentioned in Fig. 6.

5. **DISCUSSION AND CONCLUSIONS**

In this paper we established a connection between the Energy/Averaging function of Internet traffic and the activity of the traffic in various time scales. As the Averaging function is a linear transformation of the autocorrelation [34], the same is true for the autocorrelation. It then follows that, by controlling the activity of the traffic in the different time scales, we can produce simulated traffic possessing any desired autocorrelation. We also provided tools which can read such an activity off the Averaging function; these readings are not as complete as they could be, but at least they can identify the time scales of interest. We finally proposed a model, the MM, which produces traffic with any desired autocorrelation, and furthermore a combination of the MM and the DSSMSS, the MMSS, which produces traffic indistinguishable from the real one, at least by means of the tools hitherto presented: the Averaging function, the marginal distributions, and the autocorrelations yield identical results for simulations and real traffic. The fact that this was achieved without including explicitly correlation of user activity and network dynamics in the model suggests that the effect of these two factors is, in reality, less important than it is believed to be, or at least not detectable by these measurements.

6. **ACKNOWLEDGEMENTS**

The authors would like to thank Dr. Walter Willinger and Dr. Andy Ogielski of AT&T for their comments, ideas, and support, Anja Feldmann of the Computer Science Department, University of Saarbrücken, for the procuring data set 97, and Prof. Ingrid Daubechies of the Department of Mathematics, Princeton University, for contributing to the shaping of the key ideas, and improving decisively this treatise both aesthetically and scientifically, working as hard as the authors themselves did. K. Drakakis would also like to thank the Lilian Boudouris Foundation for the scholarship it granted him. Finally, the authors would like to thank NSF, AFOSR, ONR, DARPA, and AT&T for their financial support, as well as the reviewers, whose clear and extensive comments resulted in a better paper.

**References**

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

[2] A. Erramilli, O. Narayan, W. Willinger. *Experimental Queueing Analysis with Long-Range Dependent Packet Traffic*  IEEE/ACM Transactions on Networking Vol. 4, No. 2, pp. 209–223, 1996
[3] V. Ribeiro, R. Riedi, M. Crouse, R. Baraniuk. *Simulation of non-Gaussian Long-Range-Dependent Traffic using Wavelets* Proceedings of ACM SIGMETRICS 2001

[4] W. Willinger, M. Taqqu, R. Sherman, D. Wilson. *Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level* IEEE/ACM Transactions on Networking, Vol. 5, No. 1, February 1997

[5] M. Taqqu, V. TeVerovsky, W. Willinger. *Is network traffic self-similar or multifractal?* Fractals. 5 (1997) 63-73

[6] Y. Zhang, V. Paxson, S. Shenker. *The Stationarity of Internet Path Properties: Routing, Loss, and Throughput* ACIRI Technical Report, May 2000

[7] B. Ryu, A. Elwalid. *The importance of Long-Range Dependence of VBR Video Traffic in ATM Traffic Engineering: Myths and Realities* Proc. ACM SIGCOMM '96, Stanford University, CA, Aug. 1996

[8] C. Nuzman, I. Saniee, W. Sweldens, A. Weiss. *A compound Model for TCP Connection Arrivals* ITC workshop, September 2000

[9] J. Cao, W. Cleveland, D. Lin, D. Sun. *On the nonstationarity of Internet Traffic* Performance Evaluation Review: Proc. ACM Sigmetrics 2001, 29:102-112

[10] J. Levy, M. Taqqu. *Renewal-reward processes with heavy-tailed interrenewal times and heavy-tailed rewards* Bernoulli. 6 (2000) 23-44

[11] V. Pipiras, M. Taqqu. *The limit of a renewal-reward process with heavy-tailed rewards is not a linear fractional stable motion* Bernoulli. 6 (2000) 607-614

[12] A. Veres, M. Boda. *The Chaotic Nature of TCP Congestion Control* Proc. IEEE INFOCOM 2000, Tel Aviv, March 2000

[13] M. Crovella, C. Lindermann. *Internet Performance Modeling: The State of the Art at the Turn of the Century* Performance Evaluation, March 2000

[14] V. Paxson, S. Floyd. *Why We Don’t Know How To Simulate The Internet* Proceedings of the 1997 Winter Simulation Conference, December 1997

[15] A. Feldmann, A. Gilbert, W. Willinger. *Investigating the multifractal nature of Internet WAN traffic* Computer Communication Review, Vol. 28. No. 4, Proceedings of the ACM/SIGCOMM ’98, September 1998, Vancouver, Canada pp. 12–55, 1998

[16] A. Feldmann, A. Gilbert, P. Huang, W. Willinger. *Dynamics of IP traffic: A study of the role of variability and the impact of control* Proceedings of the ACM/SIGCOMM’99, August 29–September 1, 1999, Cambridge, MA

[17] K. Park, G. Kim, M. Crovella. *On the relationship between file sizes, transport protocols, and self-similar network traffic* ICNP ’96

[18] A. Feldmann, A. Gilbert, W. Willinger, T. Kurz. *The changing nature of network traffic: Scaling phenomena* Computer Communication Review, April 1998

[19] W. Leland, M.S. Taqqu, W. Willinger, D. Wilson. *On the self-similar nature of Ethernet traffic (extended version)* ACM/SIGCOMM ’93. Computer Communication Review, 23 (1993), 183-193

[20] R. Riedi, J. Lévy-Véhel. *Multifractal Properties of TCP Traffic: a Numerical Study* INRIA research report 3129, February 1997

[21] R. Riedi. *Introduction to Multifractals* Long range dependence: theory and applications, eds. Doukhan, Oppenheim and Taqqu, Birkhauser 2002

[22] R. Riedi. *Multifractal Processes* Long range dependence: theory and applications, eds. Doukhan, Oppenheim and Taqqu, Birkhauser 2002

[23] S. Sarvotham, R. Riedi, R. Baraniuk. *Network Traffic Analysis and Modeling at the Connection Level* Proceedings IEEE/ACM SIGCOMM Internet Measurement Workshop 2001, San Francisco, CA

[24] M. Taqqu, W. Willinger, R. Sherman. *Proof of a Fundamental Result in Self-Similar Traffic Modelling* Computer Communication Review, 27 (1997) 5-23

[25] L. Peterson, B. Davie. *Computer Networks: A System Approach (Second Edition)* Morgan–Kaufmann, 2000

[26] A. Araujo, E. Giné. *The central Limit theorem for Real and Banach Valued Random Variables* Wiley, New York (1980)

[27] D. Pollard. *Convergence of Stochastic Processes* Springer, New York (1984)

[28] M. Ledoux, M. Talagrand. *Probability in Banach Spaces* Springer–Verlag (1991)
[28] T. Mikosch, S.I. Resnick, H. Rootzén, A. Stegeman. *Is network traffic approximated by stable Levy motion or fractional Brownian motion?* Ann. Appl. Probab. (2002), to appear.

[29] M. Taqqu, J. Levy. *Using renewal processes to generate LRD and high variability* Progress in probability and statistics, E. Eberlein and M. Taqqu eds., Vol. 11, Birkhauser Boston 1986

[30] R. Riedi, M. Crouse, V. Ribeiro, R. Baraniuk. *A Multifractal Wavelet Model with Application to Network Traffic* IEEE Transactions on Information Theory, Vol. 45(3), April 1999

[31] J. Kilpi, I. Norros. *Testing the Gaussian approximation of aggregate traffic* IMW Proceedings 2002

[32] K.Drakakis, D. Radulovic, *A Discretized Version of the Self-Similar Model for Internet Traffic* (submitted for publication)

[33] K.Drakakis, D. Radulovic, *On the spikiness of Internet traffic* (submitted for publication)

[34] K.Drakakis, D. Radulovic, *An improvement of the Energy function* (submitted for publication)

[35] K. Drakakis. *A detailed mathematical study of several aspects of the Internet* Ph.D. Thesis, Princeton University, June 2003

---

![Energy function for data](image1)

**Figure 1:** Energy functions for data (a) and simulations (b).

![Energy function for simulations](image2)

**Figure 2:** Trace autocorrelations at four different time bins (1ms, 10ms, 100ms, 1s): the long-range dependence in coarse time scales is obvious. 97 (a) seems to have very strong periodic components, probably due to some monitoring protocol.

![Autocorrelation US](image3)

**Figure 3:** Trace autocorrelations at four different time bins: 10µs, 100µs, 1ms, 10ms for U8, and 1ms, 10ms, 100ms, 1s for 89. Trace 89 shows a relatively strong correlation irrespectively of the bin used. On the contrary, the newer traces seem to be uncorrelated, if binned with a bin other than 10ms.
Figure 4: Energy (a) and Averaging (b) functions for the Slow Start Simulations: although some slight curving is present in these simulations, it is much more modest than what is observed in real traffic (see Fig. 1).

Figure 5: Averaging functions for data (a) and simulation (b) traces.

Figure 6: The characteristic function of connection 1 341 513 1022 from trace 94, in 6 detail levels. Each figure presents a “zoom-in” in the boxed segment of the previous one, illustrating a fractal-like behavior. Here, 6 levels can be detected: the (approximate) lengths of their ON-intervals are 2000, 500, 100, 40, 5, and 0.1s, whereas the lengths of their OFF-intervals are much smaller in each case: 400, 15, 10, and 4s. Note that 0s are not plotted.

Figure 7: The characteristic function of connection 167782718 167772271 60363 119 from trace M6, in 3 detail levels. Here, 3 levels can be detected: the “approximate” length of their ON-intervals is 2, 0.44, and 0.01, whereas the corresponding OFF-intervals are much smaller. Notice that coarser levels are not present in the data set, due to its short duration. Notice also that 0s are not plotted.
Figure 8: Average functions of some MM simulations: 8 has a level at $2^8$, 12 at $2^{12}$ and 7/12/17 three levels at $2^7$, $2^{12}$, and $2^{17}$. ON- and OFF- intervals of the levels are exponentially distributed, the mean being the level. In 8S, ON/OFF-intervals are uniformly distributed within ±10% around the mean. In (a), RTT is present: the ON-intervals of the finest level consist of spikes of height 1, separated by exponentially distributed intervals with mean 4. In (b), RTT is not present, and the ON-intervals of the finest level are continuous.

Figure 9: Some more simulations: curvature is present, since there is one single change of slope, the smoothness of which varies. Hence, there is one level in these traces.

Figure 10: Averaging function of a simulated session with only one level whose ON-intervals live in time scale 15, but whose OFF-intervals live in time scale $15-i$, where, for each curve, $i$ can be read from the legend. In (a), as a rule, the local maximum of the bump "locks" on $15-i$, the scale of the OFF-intervals. In (b), using a more diffused interval length distribution, the bump is negligible.
Figure 11: Tools 1 and 2 applied to real traces. $\varepsilon=0.01$ was used.

Figure 12: Tools 1 and 2 applied to real traces. $\varepsilon=0.01$ was used.

Figure 13: Tools 1 and 2 applied to two model D simulations: 8 (No spikes) and 7/12/17 (Spikes)

Figure 14: Tool 2 applied on session 94/13415141022: it clearly identifies levels near $2^7\text{ms}$, $2^{11}\text{ms}$, and $2^{16}\text{ms}-2^{19}\text{ms}$

Results of Tools 3 and 4 for real traces:

| Trace   | 94  | 97  | L4  | L5  | M6  | M7  | U8  | 89  |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|
| Tool 3  | 0.274 | 0.084 | 0.272 | 0.289 | 0.016 | 0.017 | 0.145 | 0.125 |
| Tool 4  | -0.479 | 0.101 | -0.541 | -0.425 | -0.084 | -0.123 | -0.340 | -0.797 |

Table 1: The results of Tools 3 and 4 when applied to the real traces: notice that both of them indicate that trace 97 is very close to Gaussianity.