Discussion of "Heavy Tail Modeling and Teletraffic Data" by S.R. Resnick*

Walter Willinger and Vern Paxson
AT&T Labs – Research and Lawrence Berkeley National Laboratory

As network researchers who have spent many long days in the past few years poring over Gigabytes of network-related measurements to try to get some basic understanding of network traffic, we are very pleased to see this article and to be able to comment on it. Historically, the field of networking and communications research has suffered from a severe and constant shortage of traffic measurements [11]; however, during the past 5-10 years, an abundance of enormous-sized data sets of high-quality network traffic measurements have become available and keep on being collected in ever-increasing quantities and from ever-faster networks. Unfortunately, the development of statistical techniques and methods for efficiently dealing with this "flood" and for effectively handling phenomena such as heavy tails and long-range dependence, hitherto unknown in the networking arena, has not kept up with the rate at which newer and larger data sets of traffic measurements are being captured, nor with the pace at which the networks themselves change. Prof. Resnick's paper is therefore extremely timely and highly welcome by traffic analysts who are stymied by the many unfamiliar features that appear in networking-related data sets, and typically lack even the most basic tools to sensibly deal with them.

At the same time, Prof. Resnick's paper also serves as useful reminder that collected data often determine the utility and relevance of one data analysis or modeling approach over another. To illustrate this, we first point out to the interested statisticians and data analysts some of the features that make data sets of traffic measurements from today's networks unique and challenging, especially when compared to available data sets from other areas of science and engineering (e.g., hydrology, medicine, biophysics, economics, finance). Given the special nature of the data at hand, we then argue – in contrast to Resnick – for abandoning the *black box modeling approach* from traditional time series analysis and focusing instead on *structural models* that take into account the context in which the data arose in the first place, namely the highly intertwined hierarchies of networking functions that form the basis of modern computer communication. While we readily admit that black box models can be and are useful in other contexts, we strongly believe that they are essentially of no use for our main purposes of trying to understand the dynamic and complex nature of traffic in today's packet networks and, subsequently, of exploiting this understanding to design, manage and control these networks.

* *The Annals of Statistics*, 25(5), pp. 1805–1869, 1997. Includes this and other discussion articles, and rejoinder.

1 Not your usual data sets

Although Prof. Resnick points out that today's communication networks have recently become highly prolific providers of large amounts of traffic data, he does not dwell further into what kinds of data have actually been collected. Below, we provide a more detailed account of the data sets recently measured on current communication networks, such as Common Channel Signaling Networks (CCSNs) used by the telephone system, Ethernet Local Area Networks (LANs) ubiquitous in shared computing environments, and IP links comprising the explosively-growing global Internet. These are not the usual data sets that statisticians have been dealing with in the past. Our goal is to point out their unique features and properties, since these drive the entire subsequent discussion.

Thanks to sophisticated measurement devices, themselves often full-blown computers, the data sets collected from these networks are not only unique with respect to size, but also in terms of the quality and amount of information recorded for every observation. Depending on the network under consideration, an observation can be a *telephone call*, a *message*, or a *packet*. Today's properly designed and well-tested measurement devices record exactly the bits and bytes that are transported over the network, resulting in complete, accurate and error-free data. The only uncertainties come from time stamp resolution and the possibility of “dropping” a measurement because the recording device is too slow. Both of these can be controlled by the diligent researcher. Furthermore, once such a device has been built and tested, there are no limitations on the length of the measurement periods, that is, on the amount of data that can be collected from a “live” network (in practice, though, available disc space often *does* impose a limit on the length of the period over which traffic can be recorded continuously).

In this context, some numbers might be telling: half a day of monitoring a single 56 kilobit/sec channel of an A-link to an end office in a CCS subnetwork resulted in about 500,000 calls (January 1993); one hour of measuring traffic on a 10 megabit/sec Ethernet LAN at Bellcore (August 1989) yielded about 1,500,000 packets; a 37 minute trace of traffic at the busy Internet exchange point FIX-WEST consisted of more than 35,000,000 packets (June 1995); and a week's worth of Internet traffic into and out of the University of California at Berkeley consisted of 439,000,000 packets and 89 Gigabytes of data (January 1995).

In terms of the information recorded, for a CCS network a typical calling record consists of call arrival time, call holding time, caller and callee number, and possibly further data associated with the caller profile. For an Ethernet or Internet setting, the recorded information on a per-packet basis consists of a time stamp and the full packet “header” information, including source and destination address of the packet, size, protocol, protocol-specific information such as sequencing position, and possibly the entire packet data contents, giving all of the application details.

These high-quality and high-volume measurements result in data sets that can easily extend into the Giga- or even Terabyte range. On the one hand, such voluminous data sets pose obvious challenges for researchers interested in exploratory data analysis or “data mining,” where the latter phrase is used to denote “sensible digging into data to try to reveal what they are saying” rather than “torturing the data till they confess” (e.g., see the discussion in [2]). On the other hand, they offer data analysts and modelers unique opportunities, mainly because of the richness of available information at all levels of interest (e.g., at the level of aggregate network traffic, at the level of individual sources and destinations, at the transport protocol level, at the application level). Examples of some data sets from existing networks can be found in the *Internet Traffic Archive* at `http://www.acm.org/sigcomm/ITA/` .

2 Not your usual modeling world

Given the recent abundance of traffic data, traffic modeling is faced with a number of pressing new problems. Solving them will require a close collaboration between data analysts, applied mathematicians and networking experts. To avoid any misunderstanding: for us, the main objective of traffic analysis and modeling is to try to gain a good understanding of the actual dynamics of network traffic and to make use of this know-how when designing, managing and controlling existing or future networks.

For the present discussion, it is also crucial to keep in mind that modern communication networks are highly dynamic entities that undergo constant changes (e.g., network topology, user population, services and applications, network technologies, protocols). Just consider today's Internet and compare it with the network that existed 3 years ago (when it had practically no WWW traffic), or 1 year ago (just before the decommissioning of the NSFNET), or just a month or week ago (when it was still possible to connect to your favorite sites). Such change is nothing new, either. For example, the USENET Internet "news" system has exhibited striking, sustained exponential growth of 75%/year *since 1984* [12]. Even at fixed points in time, connection characteristics vary significantly from site to site [13]. Furthermore, not only do networks quickly change in many ways, so too can patterns of use of well-established applications. For example, in October 1992 the *median* size of an FTP file transfer measured at a large research institute was 4,500 bytes. This statistic is presumably highly robust, as it was drawn from a sample of more than 60,000 transfers. Yet only five months later, the median fell by more than a factor of two, to 2,100 bytes, in a sample of more than 80,000 transfers [13].

Clearly, life for traffic data analysts and modelers is extremely challenging and difficult. What does it mean in this setting to pick the "best-fitting" model for a given traffic trace, when the underlying data set keeps changing over time, across the same network, across different networks, and so on? The task is comparable to chasing a moving target – blindfolded! That is, the traffic data analyst is asked to detect some unknown characteristic features in enormous-sized empirical records, where the records are constantly augmented by new traffic measurements from new (and generally faster) networks that carry new (and typically more bandwidth-intensive) services and applications, and are being clogged up by an exponentially growing (and increasingly diverse) user population.

Successful traffic modeling for modern communications networks also has to face the popular belief that since actual network traffic is commonly considered to be highly complex in nature, only complicated and highly parameterized models are likely to result in accurate approximations of reality. However, in network engineering practice, which is the ultimate application of traffic modeling, such models are viewed as essentially useless, because there is no hope of forming solid estimates for numerous parameters in such a changing world. Instead, a traffic model is considered useful only if it (i) is simple and accurate, (ii) has a physical explanation in the network context and hence provides new insights into the dynamics of network traffic, (iii) can be inferred from operational measurements, and (iv) has measurable and practical impact on system performance. Unfortunately, these practical criteria for traffic modeling have all but been ignored in the past, where traffic modeling was traditionally done without access to any data and where the resulting models were primarily judged by how well they could be analyzed mathematically and hardly ever by how relevant they are in engineering practice. In contrast, the ever-increasing size of the latest data sets of traffic measurements from today's networks, the dynamic nature of these networks, and the complexity of the traffic they carry, all argue strongly in favor of modeling network traffic based on the *principle of parsimony*, also known as *Ockham's Razor* (e.g., see [6]).

The idea behind parsimonious modeling is to explain facts in as economical a way as possible. As a result, the quest for parsimonious traffic models forces the data analysts to concentrate on features that are common among the many large data sets and hence robust under changing networking conditions (the quest for "traffic invariants"); it offers the traffic modelers the opportunity to build traffic models around these invariants and to come up with plausible physical explanations for the empirically observed phenomena; and it provides network researchers with an improved understanding of the complex nature of network traffic, giving them tools to explore network performance without being hit by the "curse of dimensionality" (i.e., a large number of parameters that have little physical meaning). Naturally, this quest for parsimony will be successful the more that data analysts, traffic modelers and network engineers interact with one another.

2.1 Searching for traffic invariants: Heavy tails

Despite its great diversity, measured network traffic has revealed a number of surprises (and candidate invariants) that make it specially interesting to data analysts and modelers. One of these surprises, which is to a large degree the motivation behind Prof. Resnick's article, is the prevalence of the *infinite variance phenomenon* (also known as the *Noah Effect*), or more generally, of heavy tailed distributions in networking-related activities (e.g., CPU processor time, file size, video frame size, call holding times, inter-packet times for interactive applications, burst sizes in data bulk transfer, WWW item sizes). On the one hand, this observation is good news for network researchers who have seen traditional (i.e., telephone network-based) traffic invariants either vanish or change into highly variable and unpredictable quantities in the context of other data networks (such as the Internet), where none of the following are any longer constant: network size, topology, transfer rates, congestion levels, dominant applications, traffic mixes across sites and over time, and connection characteristics. Traffic analysts and modelers are *desperate* for traffic invariants because they provide the basis for understanding the dynamic nature of network traffic and are crucial in narrowing down the search space for models both consistent with measurements and useful in engineering practice. To this end, heavy tails are a god-send: something not changing in a sea of change.

On the other hand, the empirically observed omnipresence of heavy tails has caused serious concerns among network researchers, not only because of their inexperience with heavy-tailed phenomena, but more importantly because of an almost complete lack of readily available and practically useful tools and techniques for dealing with them. Here is where Prof. Resnick's paper (especially Section 3) not only fills a real need, but does so admirably well and in a surprisingly(?) frank manner. The paper not only surveys the growing body of techniques for inferring and modeling heavy tailed phenomena, but equally important, it emphasizes the limitations and illustrates the pitfalls associated with a "blind"

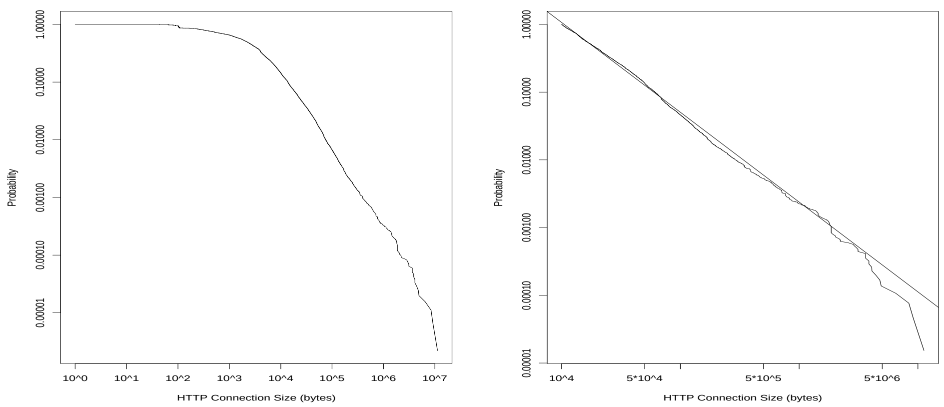

Figure 1: Log-log complementary distribution plot of WWW transfer sizes: full dataset (left), upper tail (right).

belief in these statistical techniques. Such a combined good-news/bad-news survey is especially (but not exclusively!) valuable for the field of engineering sciences, where one often encounters a strong belief (blame the statisticians!?) in the notions of universally applicable statistical techniques and absolute numbers. Resnick's discussions detailing the conditions under which the different Hill estimator techniques "work" in theory and exploring the "gray area" where they should work in theory but don't work in practice, are as valuable for practitioners as are the mathematical results (and his comment that the underlying assumptions are essentially unverifiable in practice!) showing that the Hill estimator has nice theoretical properties. As practitioners, we cannot overemphasize Prof. Resnick's implicit warnings that (i) no statistical technique works in all cases, (ii) you have to know what you are doing, and (iii) relying on a portfolio of different techniques is always preferable (often much more work, though) to sticking with a single method. We wish that more importance would be given in the statistical literature and in statistical teaching to exploring scenarios under which a given technique works well or doesn't, the pitfalls in using the technique, and possible alternative methods for dealing with the same problem (and their advantages and shortcomings).

Finally, our only criticism concerning Section 3 is that none of the presented methods really exploits or makes use of the availability of the enormous-sized data sets mentioned earlier. This obscures the different statistical regimes in which network research operates. For example, Figure 1 shows on the left a log-log complementary distribution plot for a data set of WWW transfer sizes (226,386 observations), and on the right the upper 14% tail consisting of those transfers exceeding 10 kilobytes (32,630 points). The line fitted to these argues for a Pareto tail with $\alpha \approx 1.35$.[1] Surely, there must be ways of taking advantage of the fact that for measured network traffic data, the upper heavy tails themselves contain more data points than the entire data sets considered in this paper!

2.2 Black box models vs. structural models

Besides heavy tails, another invariant feature that has been consistently observed in recent traffic data sets is *long-range dependence* (LRD), also known as the *Joseph Effect*. The mere mention of LRD usually causes strong (predominantly

[1] We have 900 additional WWW data sets like this one available for analysis.

negative) reactions from statisticians as well as mathematical modelers and is typically accompanied by endless philosophical discussions about the (im)possibility of inferring "true" LRD from a finite data set (e.g., see [7, 5, 10]). In sharp contrast, in the present context of network traffic analysis, there exist very pragmatic reasons for moving beyond philosophical discussions towards the challenging problem of providing phenomenological explanations. For real-world analysis, the philosophical points are moot: the traditional theoretical framework based on Poisson assumptions of fleeting correlations lies in shambles [9, 14], and networking practitioners find themselves giving serious consideration to LRD models as *good approximations* if nothing more.

Searching for plausible physical explanations for empirically observed phenomena such as LRD is intimately related to abandoning the *black box models* from traditional time series analysis in favor of *structural models*. The unique feature of structural traffic models is that they are capable of explicitly accounting for the hierarchical nature of today's network architectures and hence can capture the intertwined mechanisms and modes (at the different layers in the networking architecture) that determine the flow of packets across a network. For example, above the physical layer (i.e., the raw media such as copper or fiber over which data is sent) is the network interface or "link" layer with its mechanisms and protocols controlling how packets are sent over the media. Above it is the internetworking layer responsible for getting a packet through a series of different networks (e.g., the Internet Protocol, IP). At a yet higher level, the "transport" layer provides the functions needed to exchange data between two applications (e.g., the Transport Control Protocol, TCP), such as assuring that packets are reliably delivered and/or arrive at the application in the order they were sent. And above it, the application layer incorporates services designed to support a range of applications (e.g., Telnet, FTP, WWW).

Accurate accounting of this multi-level hierarchy in measured network traffic is possible because all the relevant information can be obtained unambiguously from looking "inside" the collected packets – checking the header of each recorded packet usually suffices. In contrast, (univariate) time series models typically treat these packets as black boxes. That is, by focusing on the mere existence of these packets (i.e., the corresponding time stamps, or the resulting time series of counts of packets) and not on their "meaning" as revealed by their headers, they ignore most of the gathered information. Even when replacing univariate by multivariate time series models, where the latter can account for covariates associated with the individual packets, it is not clear how the hierarchical and interconnected structure of network traffic can be captured. On the other hand, successful structural modeling makes use of potentially all of the recorded information at all layers of the network architecture. It results in traffic models that are consistent with that data at all levels of interest (e.g., internetworking layer, transport layer, application layer). And it offers practical answers to

- the data analysts' dilemma of making sense of the ever-increasing size of available network traffic measurements,
- the network researchers' quest for traffic invariants,
- the traffic modelers' emphasis on model parsimony, and
- the network engineers' insistence on models that are simple, physically meaningful, and relevant to engineering practice.

It does so by "carefully mining the available data" in the sense of Cox (see [2]), i.e., determining to what extent important conclusions from the data lie on the surface; by "broadening the basis" in the sense of Tukey [15, p. 277] (see also [4]), that is, a traffic invariant is verified by checking against a wide range of similar data sets; by applying the principle of parsimony and abstracting out features of the data that do not significantly contribute to our understanding of network traffic; by fully exploiting subject-matter considerations (achieved by close collaborations between network researchers and data analysts); and by iterating these procedures, if deemed necessary.

The appeal for structural models in the network setting stems from the observation that network hosts (e.g., computers, routers) behave in a highly predictable fashion in a given mode (e.g., request/response vs. bulk transfer vs. interactive vs. multimedia applications, file servers vs. clients). What can change drastically is the behavior in the different modes. Yet *all you have to do to recognized the different modes is look inside the packet header!* – structural modeling at its simplest. The data set "Packet" considered in the paper is an example where all information concerning the different modes that generated this traffic trace has been ignored and where only the time series of arrival times of the successive packets is considered. As such, the data set is not terribly interesting for actual network engineering. On the other hand, with structural modeling that makes use of all available information on a per-packet basis, ISDN packet arrivals are known in advance to have a non-trivial dependence structure, due to the nature of the dominant applications (e.g., word-processing) and the transport protocol used. In particular, the jump at lag 6 in the autocorrelation function is almost certainly a result of the transport protocol used or of fragmentation boundaries (if we were given the packet headers we would immediately know which!), and is of only minor interest for network engineering ("statistically significant but practically irrelevant"). For examples of successful structural modeling approaches in the context of network traffic, we refer to the recent papers by Willinger et al. [16, 18, 17], Kurtz [8], and Crovella and Bestavros [3]. These papers also elaborate on the connection between the empirically observed phenomena of heavy tails and LRD at different levels of the network structure and show, as alluded to by Prof. Resnick, how structural modeling is able to provide a unified framework within which both phenomena can be explained.

Clearly, our strong preference for structural traffic models over black box models clashes with Resnick's decision to stick with the more traditional black box modeling approach from time series analysis in Section 5. At the same time, his arguments in favor of pursuing black box models (tradition; requires no subject matter expertise; applicable across a broad range of disciplines) are not all that convincing, especially when recalling the title of the article, which explicitly mentions the application area of interest. In fact, as network researcher, we are more than willing to apply our networking know-how when it comes to constructing sensible structural models; moreover, we are much less concerned about the models' potential applicability in other areas of science, but worry immensely about the models' usefulness and practical relevance in the network context. This leaves Prof. Resnick with tradition as the remaining argument for black box models, and we would love to see him argue for tradition in front of a group of Internet researchers.

3 Not your usual topics for future work

Judging from the recent past, it is a safe to predict that network researchers will continue to be faced with practically unlimited amounts of traffic measurements from the latest packet networks. Our experience has been that traditional traffic modeling (including conventional time series analysis) has little to offer for effectively and sensibly dealing with this situation, and it has not been for want of trying. Instead, we have found that the rather vaguely defined concepts (see above) of (i) careful data mining, (ii) broadening the basis, (iii) insistence on model parsimony, and (iv) reliance on subject-matter expertise do a much better job of addressing the task of practical network traffic analysis and modeling. In the process, one garners both feasible alternatives for dealing with the available data and some basic understanding of the nature of actual network traffic (and as a result: useful traffic models).

All four concepts deal in one way or another with the basic problem of dealing effectively with large, diverse data sets, and in all four cases, the main research challenges center around the quest for techniques that scale (in the number of available observations). Beran and Terrin's [1] recent work is an example that shows how a technique (i.e., Whittle's method for estimating the degree of LRD) that is computationally infeasible in the presence of a large number of observations can be turned into a method that scales (and hence can be applied even to large data sets) by proving new central limit-type results and by exploiting modern high-performance computing and communication capabilities. Similarly, Willinger et al. [18] present an algorithm for quickly generating fractional Gaussian noise that also relies on a combination of new convergence results for stochastic processes and a highly parallel computing architecture. We strongly believe that such combinations of new probabilistic results and advanced parallel or distributed computing capabilities will be a driving force behind future progress in areas associated with the four mentioned concepts.

Another development in the area of network traffic analysis and modeling that results from the ease with which large amounts of network traffic measurements can be obtained these days is a steady shift away from *statistical inference* with its traditional emphasis on a single data set and on testing models toward *scientific inference*. Scientific inference (as discussed for example in [2]) typically involves many data sets (e.g., traffic collected on the same network over different periods in time and at different points in the network) and attempts to identify features in the data that generalize to different conditions (e.g., different network loads, different mix of users or applications). As a result, in modern network research, reproducibility (also referred to as replication or repetition) of traffic measurements studies by other researchers, in other *different* networks, and at different times, is becoming an essential part of model specification, model selection and model verification. In this sense, network research is starting to adopt a concept that has a long tradition in the physical sciences but has been all but ignored in the social sciences and in the mainstream statistics literature. Put succinctly: how do we deal with huge data sets? How do we deal with huge numbers of huge data sets? How can we do reliable estimation in the presence of strong correlations and infinite variance? How can we more formally assess heuristic arguments of similarity in inference across different data sets?

Although it is worthwhile to always heed Hampel's [5] advice that “no model is ever correct – all are but better or worse approximations of reality”[2], our decision as network researchers is obvious when given a choice between a

[2]The same idea is often expressed in the well-known saying “All models are wrong, but some are useful.”

black box model that fits a single data set "perfectly," and a model that "works" under a wide range of networking conditions and makes sense on physical grounds. However, we emphasize that our criticism of the black box modeling approach and the underlying "Box-Jenkins machinery" is subject-specific (but so is the paper under discussion) and should not be interpreted as an argument in favor of writing off time series analysis all together. At the same time, it is very disconcerning to see that despite the drastic changes in terms of available network traffic data, much work in this area is still done "the same old way," where data was a scarce resource and where "squeezing the available data set dry" (see [2]) made sense. One cannot help but be highly critical of such work when there is now plenty of data to go around and when squeezing a single data set dry would become a researcher's life time job. Indeed, things change so fast on the networking scene that spending too much time studying a particular dataset brings with it a genuine risk of rendering the ultimate findings obsolete. Surely, there must be other ways for dealing with the currently available data! In this sense, the ideas put forward by Tukey in [15] that "data analysis, like calculations, can profit from repeated starts and fresh approaches" and that "there is not just one analysis for a substantial problem" are highly relevant to network traffic analysis and modeling, especially now that the data at hand gives new meaning to statistical inference and makes structural modeling a viable alternative to traditional time series analysis.

References

[1] J. Beran and N. Terrin. Estimation of the long memory parameter, based on a multivariate central limit theorem. *Journal of Time Series Analysis*, 15:269–278, 1994.

[2] C. Chatfield. Model uncertainty, data mining and statistical inference. *J. R. Statist. Soc.* **A**, 158:419–466, 1995.

[3] M. Crovella and A. Bestavros. Self-similarity in World Wide Web traffic: Evidence and possible causes. *Proc. ACM SIGMETRICS '96*, Philadelphia, 1996.

[4] D. Draper, J. S. Hodges, C. L. Mallows and D. Pregibon. Exchangeability and data analysis. *J. R. Statist. Soc.* **A**, 156:9–37, 1993.

[5] F. R. Hampel. Data analysis and self-similar processes. In *Proceedings of the 46th Session of the International Statistical Institute*, Tokyo, Japan, September 1987. International Statistical Institute.

[6] W. H. Jefferys and J. O. Berger. Ockham's razor and Bayesian analysis. *American Scientist*, 80:64–72, 1992.

[7] V. Klemes. The Hurst phenomenon: a puzzle? *Water Resources Research*, 10:675–688, 1974.

[8] T. G. Kurtz. Limit theorems for workload input models. in: *Stochastic Networks; Theory and Applications* (eds. F. P. Kelly, S. Zachary and I. Ziedins), Oxford University Press, Oxford, 1996.

[9] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2:1-15, 1994.

[10] L. S. Liebovitch. Testing fractal and markov models of ion channel kinetics. *Biophysics Journal*, 55:373–377, 1989.

[11] P. Pawlita. Two decades of data traffic measurements: A survey of published results, experiences and applicability. in: *Teletraffic Science for New Cost-Effective Systems, Networks and Services, ITC-12* (ed. M. Bonatti), Elsevier Science Publishers B.V. (North-Holland), 1989.

[12] V. Paxson. Growth trends in wide-area TCP connections. *IEEE Network*, 8(4):8-17, 1994.

[13] V. Paxson. Empirically-derived analytic models of wide-area TCP connections. *IEEE/ACM Transactions on Networking*, 2:316-336, 1994.

[14] V. Paxson and S. Floyd. Wide-area traffic: The failure of Poisson modeling. *IEEE/ACM Transactions on Networking*, 3:226-244, 1995.

[15] J. W. Tukey. Data analysis and behavioral science or learning to bear the quantitative man's burden by shunning badmandments. in: *The Collected Works of John W. Tukey* (ed. L. W. Jones), Vol. III. Monterey: Wadsworth.

[16] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *Computer Communications Review*, 25:100–113, 1995. Proceedings of the ACM/SIGCOMM'95, Boston, August 1995.

[17] W. Willinger, V. Paxson and M. S. Taqqu. Self-similarity and heavy tails: Structural modeling of network traffic. in: *A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions* (eds. R. Adler, R. Feldman and M. S. Taqqu), Birkhauser Verlag, Boston (to appear).

[18] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level (Extended Version). *IEEE/ACM Transactions on Networking*, 5:71–86, 1997.

AT&T Labs – Research, Florham Park, New Jersey 07932
Email: walter@research.att.com

Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720
Email: vern@ee.lbl.gov