A Case For Fractal Traffic Modeling

Ashok Erramilli and Walter Willinger
Bellcore
Morristown, NJ 07960-6438
Email: {ashok, walter}@bellcore.com

Abstract— **Developments in telecommunications technology have outpaced progress in teletraffic methods and practices. As a result, the planning and dimensioning of even the latest information-age services is, by default, commonly done on the basis of classical teletraffic methods developed for circuit switched voice networks. In this paper, we will discuss recent traffic measurement studies which have demonstrated that measured traffic streams from working packet switched networks exhibit variations and fluctuations over a wide range of time scales. We motivate the application of fractal models to parsimoniously represent this apparent complexity of actual network traffic, illustrate some immediate benefits that arise from moving beyond traditional traffic modeling concepts and accepting the idea of the fractal nature of network traffic dynamics, and discuss the use of fractal models to develop traffic management methods that are accurate, practical, and based on a solid understanding of the traffic observed in "real-life" packet switched networks.**

I. Introduction

Telecommunications companies worldwide are currently investing billions of dollars in deploying broadband and wireless networks that will support large-scale access to information services. While the switching and transmission technologies to enable this deployment are available, network management methods to efficiently manage large-scale deployment lag behind these technological capabilities. For example, effective traffic management is essential for the development of efficient telecommunications networks that can cost-effectively meet the performance requirements of a full range of emerging services. However, traffic management methods – at least to the extent they are practiced – do not adequately address the complexity of the traffic in packet-based networks. As a result, the planning, design, and engineering of the latest information-age services, is by default, often done on the basis of classical teletraffic methods developed for circuit switched voice networks.

The challenge for the teletraffic community is to close the existing gaps between current teletraffic theory and practice, and to provide the theoretical foundations needed to support large-scale deployment of emerging networks in practice. In this paper, we will discuss (i) the existing gap between theory and practice, as well as the gap between theory and real-world complexity, and (ii) what is required of teletraffic theory if these gaps are to be bridged. In dealing with the complexity of traffic flows in emerging networks (as indicated by recent traffic measurement studies), traditional modeling approaches implicitly emphasize tractability of the models over other considerations. Here we argue that if the gaps between theory, practice and reality are to be closed, *parsimony* of model descriptions is an essential consideration. In particular, we argue here that it is not necessary to resort to highly parameterized models to describe the complex nature or burstiness of broadband traffic; models based on stochastic and deterministic fractal processes allow for parsimonious descriptions of the empirically observed complexity or burstiness of measured traffic. It should be noted that bursty or irregular phenomena have been observed in virtually all branches of science and engineering, and fractal geometry can be viewed as the mathematics of bursty or irregular phenomena. Although fractal models have been used with some success in many branches of science and engineering, they are new to teletraffic theory and represent a recent addition to the already large class of alternative models for describing traffic in packet switched networks. The objective of this paper is to (i) summarize some of the fundamental ideas behind fractal traffic modeling, (ii) provide evidence for the viability of fractal processes in the teletraffic setting, (iii) motivate their application in teletraffic practice, and (iv) illustrate some of the engineering insights gained from fractal traffic modeling.

II. Traffic Modeling: Theory versus Practice

A. Traditional Teletraffic Theory

Teletraffic analysis is arguably one of the most successful applications of mathematical modeling in industry. Teletraffic theory and practice have enabled enormous efficiencies in the deployment of telecommunications networks, thereby facilitating universal telephony through much of the industrialized world. There are several reasons for the success of teletraffic theory in traditional telephony. First, the resulting models are tractable, and can be readily analyzed to accurately predict most performance measures of interest. A second reason is the well-known insensitivity or robustness property of key teletraffic results, that is, model predictions appear to be accurate in practice, even when many of the modeling assumptions underlying the analysis do not hold. The third (and not the least) reason is that the most widely applied models require only a few inputs, such as arrival and service rates, which can be readily collected and estimated in practice.

It is widely recognized that traffic characteristics in packet networks differ substantially from those in telephone networks, and that in spite of the appealing robustness property of many key teletraffic theory results, these differences seriously question the straightforward use of traditional traffic models and conventional traffic engineering methods for managing packet switched networks; in fact, they make many of the conventional approaches inapplicable (e.g., see [9], [22]). At the same time, newer methods which account for these differences in traffic characteristics have not yet supplanted methodologies that are firmly based on decades worth of well-founded teletraffic theory. This is not for lack of theoretical developments; on the contrary, there have been major advances in teletraffic theory over the last four decades, and we are at a point where we can analyze – in principle – complex queueing behavior, from arrival patterns, to resource usage, call admission policies, network controls, routing, network structures and so on. It is instructive to reexamine the reasons for the earlier successes to understand why major advances on the theoretical front have been largely ignored by network engineers and have typically not been reflected in methods that are used in practice for designing, managing and controlling packet switched networks.

Traffic flows in broadband networks are far more complex than in the telephone network, mainly due to a combination of the wide range of applications and services envisaged for these networks and of the inherent burstiness of the traffic generated by many of these applications. Recent advances in teletraffic theory have almost exclusively emphasized the tractability of the ensuing models, so that one can now analyze models of considerable complexity and sophistication. Accuracy of the models is also emphasized, though establishing the validity of a proposed model against actual traffic data from working packet networks has been hampered in the past by a lack of available data from emerging networks and services. Less attention has been paid to *model parsimony* (i.e., the desire to describe and explain complex traffic processes in as economical a way as possible), perhaps because of the widespread belief that the increase in model complexity required for accurately dealing with the increasingly complicated nature of network traffic will be offset by the sophisticated and powerful analysis techniques that have been developed. However, the bottleneck in applying these methods has not been tractability, but our inability in practice to specify the inputs required by sophisticated theory. Contrast this with traditional teletraffic theory, which is based on simple inputs (arrival and service rates) that can be readily specified, collected and estimated in practice.

B. Measurement Studies of Actual Network Traffic

While historically, the area of traffic modeling has suffered from a permanent "drought" of actual traffic measurements, in the recent past it has benefited from a tremendous "flood" of high-resolution, high-quality, and high-volume traffic data from a wide range of "live" packet networks that carry real applications generated by real users. Subsequent measurement studies (e.g., see [20], [7], [15], [5], [13], [23], [11], [1], [28]) have contributed greatly to new insights into the nature of traffic in actual networks, and have indicated that there exists a considerable gap between traditional traffic models and empirically observed traffic processes. The tractability that is so attractive in traditional teletraffic analysis is implicitly based on the assumption that the traffic has variations or fluctuations over a (pre)specified time scale (as, for example, with the Poisson process) or over a limited range of time scales (e.g., low order phase-type or Markovian processes). In contrast, measured traffic processes consistently show variations or fluctuations over a wide range of time scales. This empirical finding is in full agreement with observations made in many other branches of science and engineering, namely that underlying bursty phenomena typically exhibit variations over many length or time scales (e.g., see [19]).

Variability over many time scales can occur in several contexts in real traffic. First, many distributions of interest are *heavy-tailed*, which is to say that the tails of these distributions often decay so slowly that the variance (or even the mean) may not exist. This is in stark contrast to traffic distributions commonly assumed in theory, all of which are *light-tailed*, i.e., have asymptotically exponential decays so that all moments exist. Examples of traffic processes with heavy-tailed distributions are: burst lengths, sojourn times in active/inactive states, resource holding times, inter-arrival times etc. This discrepancy between theory and reality can have many obvious, as well as subtle, consequences for traffic modeling, engineering and management, which are described later.

Secondly, autocorrelations in traffic processes can span many time scales, i.e., they display *long-range dependence* or equivalently, they exhibit a power-law decay and are non-summable. In contrast, Markovian processes will generate traffic streams that are *short-range dependent*, i.e., have asymptotically exponential decays. Time series of counts of packets/ bytes/cells are typically consistent with the statistical characteristic of long-range dependence and do *not* support the prevailing modeling assumption that packet traffic is Markovian, or more general, exhibits short-range dependence.

From experiences in other disciplines, phenomena that span many length or time scales pose significant challenges to the modeler. A popular example is the problem of estimating the length of a coastline; because coastlines exhibit features over a very wide range of length scales, any length measurement depends sensitively on the length of the yardstick. To this extent, the length of a coastline is arbitrary, and the coastline is better represented by the notion of a *fractal dimension* which is a parametric representation of this dependence. A traffic analog of this example is the problem of estimating the "peakedness" (strictly speaking, the asymptotic value of the *Index of Dispersion of Counts*) of actual packet traffic; because the traffic exhibits variability over many time scales of engineering interest, the peakedness of a given trace keeps increasing with the length of the observation interval. To this extent, traditional teletraffic notions such as peakedness are inapplicable to describe packet traffic. The *Hurst parameter* (which is commonly used as a measure of the degree of long-range dependence or persistence in a given data set) can be viewed as a parametric representation of the peakedness functional and is a more appropriate description of the burstiness in packet traffic. The challenge for teletraffic theorists is then to capture the complexity of actual and future network traffic without a commensurate increase in model complexity and in the number of input parameters that must be independently specified. It is this need to model complexity efficiently that motivates an application of fractal models in the teletraffic arena.

III. Old versus New: Alternative Traffic Descriptions

In principle, the characteristics observed in measured network traffic data can be described in a number of ways. It is useful to separate fractal phenomena (i.e., empirically observed characteristics spanning many time scales) from models used to describe them. Regarding the modeling of empirically observed phenomena, one should always keep in mind that "... no model is ever correct – all are but better or worse approximations of reality" (see [12]).

A. Extending Traditional Approaches

The prevailing approach in the teletraffic literature to dealing with and describing the increasingly complex dynamics of today's traffic streams has been to model the observed traffic dynamics *one time scale at a time*, which typically results in Markovian models with a large number of states. While this approach has the obvious advantage of retaining, at least in principle, model tractability, we consider in the following some of the fundamental problems associated with using extensions of traditional Markov models to modeling highly variable (i.e., heavy-tailed) and persistent (i.e., long-range dependent) phenomena. In principle, any empirical distribution, whether heavy-tailed or light-tailed, can be represented by mixtures of exponentials. An early example of describing bursty processes in this fashion is the approach commented on in [18] of modeling inter-error times in digital transmission systems. In his criticism of this approach, Mandelbrot [18] points out several limitations and practical problems with the proposed method. For one, the resulting models are highly arbitrary and generally highly parameterized, with the number of parameters required to obtain a satisfactory fit increasing with the size of the data set. In the limit, one is effectively computing the Laplace transform of the empirical distribution, which can be viewed as an alternate representation of the data, but not a very useful one (as Mandelbrot [18] points out, a model should not explicitly include in its input all of the features it hopes to observe in its output). The resulting models yield little insight into the nature of the data, may be misleading in their performance predictions, and beyond the tractability of these models, there is no physical basis or explanation underlying them. In addition, there are several practical issues which limit the usability of these models. Fitting highly parametrized models requires the collection, processing, transport and storing of large amounts of operational data. In practice, the number and frequency of operational measurements is drastically limited by considerations of measurements overhead, and the capacity of operational systems. Secondly, even if the data were available, the inference problem of fitting a large (unknown) number of exponentials to empirical data is known to be ill-conditioned. In practice, it is such considerations, and not tractability, that determine the practical utility of these models. Similar limitations and practical reservations exist for models that attempt to capture the long-range dependence phenomenon using Markovian processes that incorporate a large numbers of states. While work on these models has emphasized the issue of tractability, the problem of statistical inference remains largely an open issue and stands in the way to check the validity and accuracy of these models against real traffic data.

B. Where Traditional Approaches Fail

In the recent past, the availability of high-volume/high-quality traffic measurements from working packet networks has seriously questioned the traditional traffic modeling approach that has been so successful for today's telephone network. In sharp contrast to the voice world, data networks are highly dynamic entities and undergo constant changes (e.g., network topology, user population, services and applications, network technologies, protocols). For example, a given WAN or Ethernet LAN, monitored 1 year or just a few months apart, can experience a significant change in the traffic due to the emergence and popularity of new "killer-applications" (e.g., WWW, Mbone). The traffic volume generated by these killer-applications is a perfect example that shows exponential or even faster-than-exponential growth rates. This observation makes life for the traffic modeler difficult and challenging: how to describe and come up with traffic descriptors (either at the source, application or aggregate level) that (i) are *robust* under the dynamics of "live" and evolving networks that continuously experience new users asking for new services and applications, (ii) are *simple, accurate and useful in practice* to support the design, engineering and operation of these networks, and (iii) have a *physical* meaning and hence contribute to a better understanding of the nature of traffic in modern packet networks.

Traditional traffic modeling has dealt with this challenging task by essentially avoiding the robustness issues and by treating traffic modeling on a "case by case" basis (e.g., scene and codec specific modeling of variable-bit-rate video traffic), by concentrating almost exclusively on model tractability and by largely ignoring issues related to model parsimony, and by not asking in general any questions regarding the physical basis of a proposed model. Consequently, our understanding to date of the nature of real network traffic based on the traditional modeling approaches is rather limited and often based on preconceptions, apparently inherent (i.e., invariant or robust) network traffic characteristics (e.g., Markovian properties, exponential tails) are confused at times with technical assumptions made at the modeling stage mostly for mathematical convenience, and there is a common belief that modeling the unquestionably complex nature of modern network traffic means automatically complicated and equally complex traffic models. We outline below an alternative approach to traffic modeling based on fractal processes, that is based on traffic measurements from working networks, provides new and profound insights into the nature of actual network traffic, and identifies robustness properties of the models that are closely related to the physical explanations that exist for these models at the level of individual sources as well as at the application level.

C. Beyond Traditional Approaches: Fractal Models

As has been demonstrated in numerous other areas in science and engineering, it is possible to represent processes that exhibit fluctuations and variability over a wide range of time scales without resorting to highly parametrized models. For example, heavy-tailed phenomena can be modeled using distributions that exhibit *power-law tail behavior*; in many cases, the performance phenomena of interest are dominated by the properties of the tail of the distribution. The *Pareto family of distributions*, for example, can parsimoniously match the power-law decay observed in many practical traffic processes with two parameters:

$$P(T > t) = (t/k)^{-\alpha},\ k > 0,\ \alpha > 0,\ t \geq k. \qquad (1)$$

For $\alpha < 2$, the variance and higher moments of the distribution are unbounded. Analysis of Ethernet traces at the level of individual source-destination pairs (see [28]) strongly favors a revision of the traditional *ON/OFF* source model in which the sojourn times are not modeled by the familiar exponential distribution, but instead by Pareto distributions with finite means and infinite variances. An equivalent description in terms of *chaotic maps* is also feasible (see [10],

[25], [24]). In the chaotic map formulation, the source state is represented by a continuous variable whose evolution in discrete time is described by a low order, nonlinear dynamical system. The packet generation process is now modeled by stipulating that a source generates a batch of packets at the peak rate when the state variable is above a threshold, and is idle otherwise. Realistic *ON/OFF* behavior can now be described in terms of a small number of parameters associated with a suitably chosen map. Either of these methods enables us to model directly what is a plausible physical basis of the self-similarity phenomenon observed in network traffic, namely the aggregation of heavy-tailed or highly-variable *ON/OFF* sources.

Likewise, self-similar models can parsimoniously capture autocorrelations that exhibit a power-law tail behavior, i.e., dependencies that range over a wide range of time scales. The Fractional Brownian Motion (FBM) traffic model, introduced by [15], [22] can capture the second-order properties of bursty traffic processes over many time scales with three parameters:

$$A(t) = mt + \sqrt{am}Z(t), \quad t \geq 0, \tag{2}$$

where $A(t)$ is the cumulative work up to time t, $Z(t)$ is a Fractional Brownian Motion (a Gaussian process which extends the familiar notion of Brownian Motion, characterized by independent Gaussian increments, to dependent increments with a power law autocorrelation), and where each of the three parameters m, H, a has a distinct physical interpretation - m is the arrival rate, the Hurst parameter H characterizes the decay of the auto-correlation function (or equivalently, the degree of self-similarity), and a is a "peakedness" term describing the magnitude of fluctuations.

Although the FBM model has been shown in an Ethernet LAN environment to be fully consistent with actual traffic data, not only at the level of aggregate network traffic traces but also at the level of individual source-destination traces (see [15], [28]), the model has several obvious drawbacks (shared by traditional diffusion models); for example, the "increments" of FBM can actually be negative. However, under heavy traffic conditions, this does not significantly impact the usefulness of the model and the accuracy of model predictions. The conditions under which the FBM model is expected to be valid in practice are described in [9] and are (i) the "time scales of interest" for the problem coincide with the scaling region – note that the FBM model is self-similar on all time scales, whereas measured network traffic exhibits self-similarity over a wide but limited range of time scales (bounded by upper and lower cut-offs), (ii) the traffic is aggregated from a large number of independent users so that a purely second-order description is adequate, and (iii) the effect of flow controls on any one user is negligible. For environments in which these conditions are satisfied, such as with Bellcore's internal Ethernet LANs, the FBM model provides reasonable agreement with simulation results for several performance measures of interest (e.g., see [9]). In other environments, for example, Bellcore's connection to the outside world, or the Australian FASTPAC network (see [21]), the number of aggregated sources does not appear to be large enough for a purely second-order description to be sufficient. In both cases, the reported utilizations are quite low; as utilizations increase, presumably as a result of increased aggregation, agreement with the FBM model can be expected to improve. Note that under conditions of limited aggregation, *no* purely second-order description, whether based on fractal or conventional Markov models, is expected to be valid. Currently, one can therefore parsimoniously model realistic traffic under two extreme conditions: at the single source level, and the limiting case of a large number of aggregate sources. The interim regime may not be readily susceptible to a parsimonious description, though descriptions in terms of aggregate *ON/OFF* sources may be more promising than others that would explicitly model higher-order statistics.

A second caveat deals with the choice of the performance measures of interest. Cell losses are extremely bursty processes, and longer term measures such as cell loss rates do not completely describe the cell loss process. In particular, the problem of predicting/maintaining/ controlling cell loss rates for a given utilization rate may well be ill-conditioned. Small variations in parameter values (such as utilizations, traffic arrival parameters) can cause large variations in cell loss rates. For such reasons, in analysis and engineering, it is important to select robust measures, such as setting safe operating points and statistical multiplexing gains.

Finally, the most significant criticism of fractal processes is that the resulting models, while powerful in their descriptive capabilities, are difficult to analyze. Currently, there are few techniques available that would allow for the routine analysis of fractal traffic models with the myriad possible variations of queueing systems. This is the experience in other disciplines as well. Engineering analysis can be divided into three distinct phases: description, analysis and control. Most applications of fractal models in science and engineering have focused on the descriptive phase, whereas all three aspects are necessary in teletraffic modeling. The difficulties in analyzing fractal models are not necessarily a reflection on fractal models per se, but are representative of the problems inherent in analyzing phenomena that involve a wide range of time scales. In the absence of methods that can permit the routine analysis of fractal queueing systems, simulation methods are suggested for routine analysis. The generation and statistical analysis of traffic with long-range dependence/heavy-tails is an active area of current research (see for example [14], [27], [28], [10], [24]). There are nevertheless several analytical results which provide considerable insights into the engineering impacts of fractal traffic. These are discussed next.

IV. Engineering Impacts

In this section, we will briefly review current insights into the performance and engineering impacts of the fractal nature of measured network traffic. It is now recognized that the asymptotic queue length distributions of traffic arrival processes with long-range dependence can be significantly heavier than the exponential decay predicted by finite state Markov models. Note that this statement refers to the nature of the decay, and does not imply, as some people argue, that fractal models necessarily predict "worse" performance! The relative performance predicted by a Markov or a fractal model will depend on specific parameter choices, and time scales of interest in the queueing system. The FBM model predicts a *Weibullian* or "stretched exponential" decay of the form

$$P(V > x) \sim \exp[-cx^{2-2H}], \tag{3}$$

where c is a decay coefficient that depends on the parameters of the arrival process and the capacity of the system (see, for example [22], [4]). This result is consistent with empirical simulation studies based on actual traffic traces (see [9]). The same result has been obtained directly from a superposition of heavy-tailed *ON/OFF* sources at high loads by [2]. The "heavier-than-exponential" decay of the queue length distributions has significant impacts on buffer dimensioning, connection admission controls, as well as on setting statistical multiplexing gains and safe operating points.

Even heavier queueing behavior is obtained with single *ON/OFF* sources with infinite variance sojourn time distributions in each state. The asymptotic queue length behavior exhibits a power-law such that the average queue length is unbounded. The same is true of aggregations of such sources, provided the peak rate of an individual source exceeds the capacity of the server (e.g., see [26] and also [17]). This suggests that in practice traffic shapers and policing mechanisms that attempt to shape the peak traffic rates of heavy-tailed *ON/OFF* sources will incur large backlogs. Sources with heavy-tailed *OFF* behavior, but light-tailed *ON* behavior also exhibit correlations spanning many time scales, but the queue length distribution decays exponentially (see for example [24], [25]). Under aggregation, both light/heavy- and heavy/heavy- tailed *ON/OFF* sources approach a purely second-order process with similar low-frequency structure. Asymptotically, both aggregates generate Weibullian tails over the bulk of the queue length distribution [24]. The heavy-tailed queueing behavior yields insights into the source of statistical multiplexing gains within networks. The exponential decay of queue length distributions driven by short-range dependent processes can be associated with an *equivalent bandwidth* (e.g., see [6]) that is less than the peak rate of the source. This corresponds to the gain obtained "within a source" by buffering the traffic during episodes when it arrives at a rate greater than the capacity of the server. In effect, this is an attempt to exploit the fact that traffic peaks do not persist in a short-range dependent source, and that the peaks can be averaged out over intervals of low traffic. The gains within a source may be substantial enough to ignore considerations of gains "across sources", where one attempts to exploit the fact that independent users may not be simultaneously active. In contrast, heavy-tailed queueing behavior suggests that there may not be much scope to obtain multiplexing gains within a source, though there is considerable potential to achieve substantial multiplexing gains across sources ([3]).

Heavy-tailed service time distributions can also have significant performance and engineering impacts, even in situations in which classical teletraffic results are known to be highly robust. An example is the blocking performance of a group of servers whose holding times can span many time scales (see [8]). Network scenarios under which this may occur are: circuit switched data applications, full service ATM networks supporting Video on Demand services. The Erlang B and Engset formulas are known to be insensitive to the service time distribution, beyond its mean value. Nevertheless, there are a number of reasons why classical tables will not correctly predict blocking performance: convergence to the steady state results requires of the order of 5-10 average holding times, which means that the system will not reach steady state over an engineering period. Secondly, the blocking for reattempts can be very high, because the state of the system evolves slowly. Over the span of an engineering period, the arrival rates may vary, so that nonstationary effects may have to be accounted for.

Finally, cell loss processes may also be best described by heavy-tailed inter-loss distributions. Cell losses are so bursty that the long-term average rate is believed to be an insufficient indicator of degradations in application performance.

V. Conclusion

Traffic in emerging high-speed networks is likely to be complex, because of the wide range of applications envisaged, as well as the inherent burstiness in the arrival and resource usage patterns of packet-based services. For teletraffic methods and practices to keep up with the rapid pace of technological advances, this complexity must be addressed in ways that can be applied in practice. This implies the need for traffic models that can capture complexity realistically, without a proportional increase in model complexity. Fractal traffic models based on fractal processes offer some promise in this regard. While these models do allow parsimonious descriptions of complex traffic processes, methods to allow routine analyses of these models, beyond simulations, are not as yet available. To this extent, fractal and Markovian models have complementary strengths and weaknesses. A promising and at the same time highly challenging direction for future research is the development of novel methodologies that allow for direct analyses of fractal traffic and queueing models. In the absence of such methods, a temporary solution to dealing with fractal models in practice requires new advances in the area of statistical inference for Markov models and new research into hybrid approaches that map fractal model descriptions onto Markov models for analysis.

Acknowledgment

We would like to acknowledge useful discussions with Darryl Veitch and Parag Pruthi of the Royal Institute of Technology, Stockholm, and our colleagues K.R. Krishnan, Gopal Meempat, Arnie Neidhardt, Don Smith, and Jonathan Wang.

References

[1] J. Beran, R. Sherman, M.S. Taqqu and W. Willinger. Long-Range Dependence in Variable-Bit-Rate Video Traffic. *IEEE Trans. Commun.*, Vol. 43, No. 2/3/4, pp. 1566–1579, 1995.

[2] F. Brichet, J. Roberts, A. Simonian, and D. Veitch. Heavy traffic analysis of a fluid queue fed by on/off sources with long-range dependence. Preprint, 1995.

[3] N. G. Duffield. Economies of scale in queues with sources having power-law large deviation scalings. Preprint, 1994.

[4] N. G. Duffield and N. O'Connell. Large deviations and overflow probabilities for the general single-server queue, with applications. *Proc. Cambridge Phil. Soc.*, 1995 (to appear).

[5] D.E. Duffy, A.A. McIntosh, M. Rosenstein and W. Willinger. Statistical Analysis of CCSN/SS7 Traffic Data from Working Subnetworks. *IEEE JSAC*, Vol. 12, No. 3, pp. 544–551, 1994.

[6] A. I. Elwalid and D. Mitra. Effective bandwidths of general Markovian sources and admission control of high speed networks. *IEEE/ACM Transactions on Networking*, Vol. 1, pp. 329–343, 1993.

[7] A. Erramilli and W. Willinger. Fractal Properties in Packet Traffic Measurements. *Proc. Regional ITC Seminar*, pp. 144–158, St. Petersburg, Russia, 1993.

[8] A. Erramilli, J. Gordon and W. Willinger. Applications of Fractals in Engineering for Realistic Traffic Processes. in: *The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks (Proc. of ITC-14, Antibes Juan-les-Pins,*

France, June 1994), J. Labetoulle and J.W. Roberts (Eds.), Elsevier, Amsterdam, 1994, pp. 35–44.

[9] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent traffic. Preprint, 1994.

[10] A. Erramilli, R.P. Singh and P. Pruthi. An Application of Deterministic Chaotic Maps to Model Packet Traffic. *Queueing Systems* (to appear).

[11] M.W. Garrett and W. Willinger. Analysis, Modeling and Generation of Self-Similar VBR Video Traffic. *Proc. of the ACM Sigcomm'94*, pp. 269–280, London, UK, 1994.

[12] F.R. Hampel. Data Analysis and Self-Similar Processes. *Proc. of the 46th Session of ISI, Tokyo, Japan*, Book 4, pp. 1–20, 1987.

[13] S. M. Klivansky, A. Mukherjee, and C. Song. On Long-Range Dependence in NSFNET Traffic. Preprint, 1994.

[14] W.-C. Lau, A. Erramilli, J.L. Wang and W. Willinger. Self-Similar Traffic Generation: The Random Midpoint Displacement Algorithm and Its Properties. *Proc. of the ICC'95*, Seattle, WA, pp. 466–472, 1995.

[15] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of Ethernet traffic. *ACM/SIGCOMM Computer Communications Review*, Vol. 23, pp. 183–193, 1993. Proceedings of the ACM/SIGCOMM'93, San Francisco, September 1993.

[16] 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*, Vol. 2, pp. 1–15, 1994.

[17] N. Likhanov, B. Tsybakov and N.D. Georganas. Analysis of an ATM Buffer with Self-Similar ("Fractal") Input Traffic. *Proc. of the INFOCOM'95*, Boston, MA, 1995.

[18] B.B. Mandelbrot. Self-Similar Error Clusters in Communication Systems and the Concept of Conditional Stationarity. *IEEE Trans. Communications Technology COM-13*, pp. 71–90, 1965.

[19] B. B. Mandelbrot. *The Fractal Geometry of Nature.* Freeman, New York, 1983.

[20] K. Meier-Hellstern, P. E. Wirth, Y.-L. Yan, and D. A. Hoeflin. Traffic models for ISDN data users: Office Automation application. In A. Jensen and V. B. Iversen, editors, *Teletraffic and Datatraffic in a Period of Change, Proc. of ITC13, Copenhagen*, pp. 167–172, Amsterdam, 1991, Elsevier Science Publishers B. V.

[21] T. Neame, R. Addie, F. Huebner and M. Zukerman. Towards a Simple Traffic Model for High-Speed Networks. *Proc. of ATNAC*, 1995.

[22] I. Norros. A storage model with self-similar input. *Queueing Systems*, Vol. 16, pp. 387–396, 1994.

[23] V. Paxson and S. Floyd. Wide-area traffic: The failure of Poisson modeling. *Proceedings of the ACM/SIGCOMM'94*, pp. 257–268, 1994.

[24] P. Pruthi. An Application of Chaotic Maps to Packet Traffic Modeling. Ph.D. thesis, Royal Institute of Technology, Stockholm, Sweden, 1995.

[25] P. Pruthi and A. Erramilli. Heavy-tailed on/off source behavior and self-similar traffic. *Proc. IEEE ICC'95*, Seattle, June 1995.

[26] A. Simonian and D. Veitch. A Storage Model with Long-Range Dependent *ON/OFF* Sources With Large Peak Rates Preprint 1995.

[27] W. Willinger, M.S. Taqqu, W. Leland and D.V. Wilson. Self-Similarity in High-Speed Packet Traffic: Analysis and Modeling of Ethernet Traffic Measurements. *Statistical Science*, Vol. 10, pp. 67–85, 1995.

[28] 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, pp. 100–113, 1995; Proceedings of the ACM/SIGCOMM'95, Cambridge, MA, August 1995.