Fractals, Heavy-Tails, and the Internet

By Dr. Martin J. Fischer, Dr. Thomas B. Fowler

In a recent *Business Week* article (April 2001), a mathematical concept known as fractals was listed as one of the top ten technologies of the future. The article referenced many potential applications for fractals, including data compression and the Internet. In this paper, we discuss the impact fractals have on the Internet and how fractals relate to heavy-tailed probability distributions. Fractals are applicable when the underlying process being mathematically modeled has a similar appearance regardless of the time or observation scale. It turns out that much of the traffic riding the Internet can be modeled using fractals. Fowler (1999) presents a tutorial on fractals and the Internet. Here we review his tutorial and develop the relationship to heavy-tailed probability distributions and problems they cause in analyzing Internet congestion.

In the good old days of traffic engineering, modeling congestion in telecommunications networks was simpler than today. There was only one kind of traffic of any significance, and that was voice. It had (and has) well-known characteristics; namely, Poisson interarrival rates (time between calls reaching the central office switch) and exponential call length distribution. There was no need to worry about things such as network layers, as they did not exist. It was easy to measure critical values of the important parameters. Queueing theory permitted analysis of voice networks to meet any desired performance characteristics; for example, call-blocking probability. Erlang's (1917) pioneering work and his famous loss formula and its extension were in wide use from the early 20th century.

With the growth of data networks, those days are rapidly shrinking. Multilayer protocols are here to stay and many complicated traffic statistics are creating greater difficulty in analyzing network traffic. There are also a far greater number of applications (not just voice conversation), each with its own traffic characteristics, and new applications can arise at any time. There are more varieties of network connectivity, architecture, and equipment, and, accordingly, different types of traffic flow. There are no standard network topologies around which all design efforts can be based, and the topologies that exist are subject to constant change.

Perhaps the most serious and most surprising characteristic that packet networks exhibit is burstiness. Burstiness is a qualitative concept, but it can be described analytically as self-similar on multiple time scales. Such behavior observed over a wide range of time scales suggested that fractals are the most appropriate mathematical tool to describe certain aspects of network behavior. This was the astonishing discovery of Bellcore researchers in the late 1980s and early 1990s. These results, originally observed in local area networks (LANs) (Leland et al, 1994), were later extended to wide area networks (WANs) (Paxson and Floyd, 1995). The implications of this discovery are only beginning to be explored.

Figure 1 was taken from Willinger and Paxson (1998). It shows the fractal or self-similar behavior of Internet traffic (right column) compared to that of voice (Poisson) traffic. The problem of self-similarity can be readily understood from this figure, which shows the effect of looking at telephone traffic, which has Poisson arrival rates, with Internet traffic, which does not. As shown in

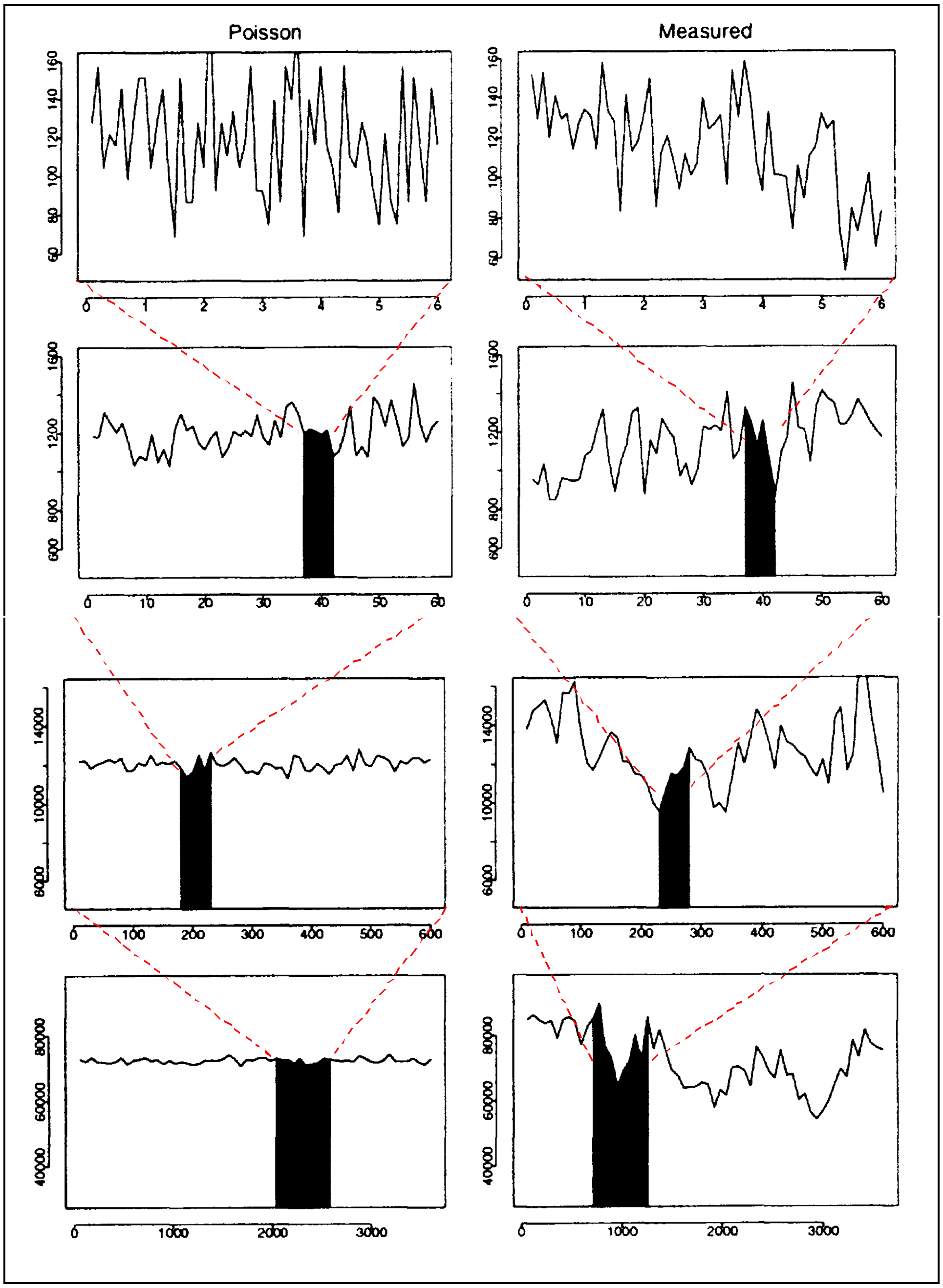

Figure 1. Self-Similarity of Internet Traffic (Measured) and Not In Poisson Or Ordinary Telephone Traffic

Figure 1, packets per unit time are counted (vertical scale) for a given time interval (horizontal scale). The time interval is then increased by a factor of 10 (for the second and third graphs) and by a factor of 7 (for the bottom graphs). The unit time interval is increased by this same factor. Traffic here is measured at the link layer. In effect, the averaging intervals become longer from top to bottom.

As the figure shows, averaging Poisson or voice traffic over longer intervals smoothes out the burstiness, whereas Internet traffic shows the same burstiness regardless of time scale. This means that traffic peaks in voice networks are limited in frequency of occurrence and severity, and a voice network can be engineered to reduce ill effects below any desired threshold. Such is not the case with the Internet, however, where the burstiness cannot be averaged out, nor can ill effects be handled by buffering. Indeed, the probability of loss cannot, in general, even be estimated, as with the voice traffic. The similar appearance of the graphs on the right-hand side of Figure 1, regardless of time scale, is the telltale sign of fractal behavior.

Figure 2 comes from Fowler (1999); it shows the traffic characteristic (Invariant) and the related probabilility distribution that can be used to model the invariant.

The listed references are contained in Fowler (1999). It is clear that two probability distributions that play an important role in modeling these traffic characteristics are the Pareto and Lognormal distributions. They, along with the Weibull, come from a class of distribution that is known as heavy-, power-, long- or fat-tailed. For a precise definition of each type of distribution (Fischer et al, 2000). It is immediate that if one wants to analyze Internet congestion using queueing theory, then one has to able to deal with the Pareto, Lognormal, and Weibull distributions. In fact, Paxson and Floyd (1995) suggested that self-similar traffic processes can be generated by a sequence of independent and identically distributed Pareto observations.

Why do these distributions limit the use of queueing theory results and what is being done about it? Many of the available results from queuing theory require the existence of the Laplace transform of the underlying interarrival or service time distributions. The Pareto, Lognormal, and Weibull do not possess closed form mathematical expressions for their Laplace transform. Thus, standard queueing results cannot be applied. Researchers have been addressing the problem on three fronts. First, they have tried to fit the heavy-tailed distribution with a phase-type distribution that is amenable to using results from queueing theory (Feldmann and Whitt, 1998 or Grenier et al. 1999). This approach has met with some success, but is limited because fitting the distribution can get complicated. Another approach being investigated is called the Transform Approximation Method (TAM) (Harris et al, 2000, Brill et al, 2001, and Fischer et al, 2000). With this method, the transform is approximated and then standard queueing results applied. Research to date has been quite positive using TAM. The final method has been in the area of using available simulation packages like GPSSH and Arena on these problems. Initial analysis (Brill et al, 2000 and Gross et al, 2001) has shown some potential, but requires excessive computer runtimes.

Invariant	Protocol level	Distribution	Parameters	Reference
Connection size	-	**Lognormal**		Paxson (1997)
Connection duration	-	**Lognormal**		Paxson (1997)
Requested file popularity	Application	**Zipf**		Pitkow, p. 6
Requested file sizes (overall)	Application	Hybrid: **Lognormal** body, **Pareto** tail (Heavy-tailed)	HTML Size μ=4-6KB Median: 2KB Images: 14 KB	Pitkow, p. 6
FTP transfers Number of Page Requests/Site	Application Application	**Pareto** tail (Heavy tailed) **Inverse Gaussian** (Heavy-tailed)	$0.9 \leq \alpha \leq 1.1$ μ=3 σ=9 mode=1	Crovella, 1997, p. 14 Pitkow, p. 6
Reading time/page (sec)	Application	**Heavy-tailed**	μ=30 median=7 σ=100	Pitkow, p. 6
Sessions (arrivals)	Session	**Poisson**		Feldman [1998], p. 7
Session duration	Session	**Pareto** (Heavy-tailed)		Feldman [1998], p. 7
Session size WAN traffic at TCP level	Session Transport	**Pareto** (Heavy-tailed) **Self-similar** (fractal, multifractal)		Feldman [1998], p. 7 Crovella, 1997, pl 21
TCP connections/Web session	Transport	**Heavy-tailed**		Crovella, 1997, p. 23
Interarrival time of packets	Network	**Heavy-tailed** (LRD, fractal)	Cox model	Crovella, 1997, p. 15
Interarrival (generation) time of packets generated by user at keyboard Interarrival time of Ethernet frames	Network Data Link	**Pareto** (body) **Pareto** (upper tail) **Self-similar** (fractal)		Paxson & Floyd, 1997, p. 6 Crovella & Bestavros, 1997, p.

Figure 2. Internet Traffic Invariant, Protocol Level and Related Probability Distribution

Figure 3 demonstrates the differences in Internet waiting times when using the old rules (Erlang models) as compared with the use of the Transform Approximation Method for characterizing the self-similar nature of Internet traffic. Orders in magnitude differences can result in using the two methods and thus significant care must be taken.

In this paper, we have presented the concept of fractals and described how they apply to Internet traffic analysis. We have related fractals and heavy-tailed distributions; and we have explained way these distributions severely limit the use of available queuing theory results. Finally, current research to overcome these problems was summarized. Interested readers may contact the authors of this article or read the appropriate references provided below.

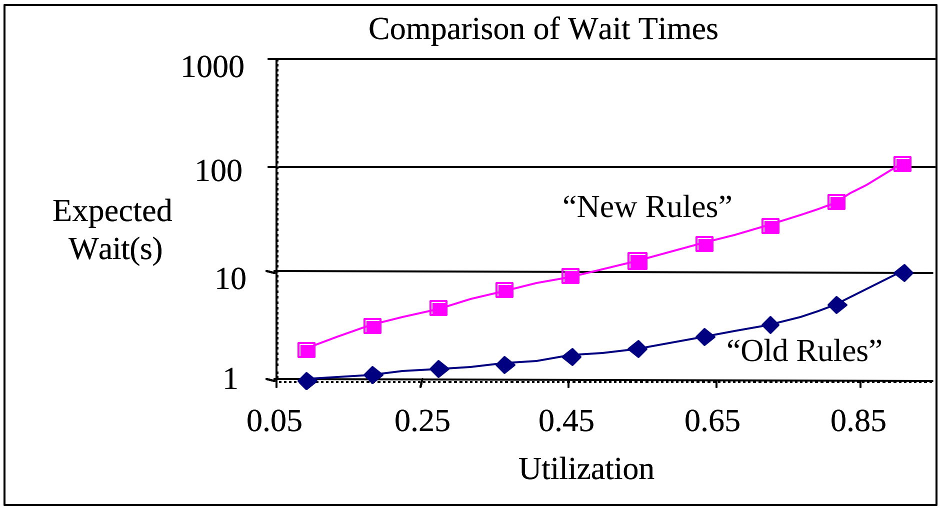

Figure 3. Significant Longer Waiting Times Using New Rules for Internet Traffic Characterization as Compared with Old Rules

References

Brill, P. H., M. J. Fischer, D. Gross, D. M. Masi, "Internet-Type Queues With Heavy-Tailed Service Times and Their Computational Solution," Submitted to the *INFORMS Journal on Computing*, 2001.

Business Week, 7 April 2001.

Erlang, A. K., "Solution of Some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges," *Post Office Electrical Engineer's Journal*, 10, pp. 189-197, 1917-1918.

Feldman, A. and W. Whitt, "Fitting Mixtures of Exponentials to Long-Tail Distributions to Analyze Network Performance Models," *Performance Evaluation* 31, pp. 245-279, 1998.

Fowler, T. B., "A Short Tutorial on Fractals and Internet Traffic," *The Telecommunications Review*, Volume 10, pp. 1-15, 1999.

Fischer, M. J., D. Gross, J. F. Shortle, D. M. Masi, "Analyzing the Waiting Time Process in Internet Queueing Systems with the Transform Approximation Method," Submitted to *The Telecommunications Review*, *Volume 11*, 2001.

Greiner, M., M. Jobmann, L. Lipsky, "The Importance of Power-Tail Distributions for Modeling Queueing Systems," *Operations Research* 47, pp. 313-326, 1999.

Gross, D., J. F. Shortle, M. J. Fischer, D. M. Masi, "Using Quantile Estimates in Simulating Internet Queues with Heavy-Tailed Service Times," *5th World Multi-Conference on Systemics, Cybernetics and Informatics*, Orlando, Florida," 2001.

Harris, C. M., P. H. Brill, M. J. Fischer, "Internet-Type Queues with Power-Tailed Interarrival Times and Computational Methods for Their Analysis," *INFORMS Journal on Computing* 12, pp. 261-271, 2000.

Leland, W., M. Taqqu, W. Willinger, and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEEE/ACM Transactions on Networking*, 2(1), pp. 1-13, 1994.

Paxson, V. and S. Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling," *IEEE/ACM Transactions on Networking*, Vol. 3, No. 3, pp. 226-244, 1995.

Willinger, W. and V. Paxson, "Where Mathematics Meets the Internet," *Notices of the American Mathematical Society*, Vol. 45, No. 8, pp. 961-970, 1998.

Marty Fischer is a Senior Fellow working a wide variety of network and information technology problems for Mitretek's clients. Marty can be reached at 703-610-1684 or mfischer@mitretek.org.

Tom Fowler is a Senior Principal working a wide variety of network and information technology problems for Mitretek's clients. Tom can be reached at 703-610-2944 or tfowler@mitretek.org.