Lessons from "On the Self-Similar Nature of Ethernet Traffic"

Walter Willinger
NIKSUN, Inc., USA
wwillinger@niksun.com

Murad S. Taqqu
Boston University, USA
murad@bu.edu

Daniel V. Wilson
Unaffiliated, USA
dvw@DanielVWilson.com

This article is an editorial note submitted to CCR. It has NOT been peer reviewed.
The authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

ABSTRACT

This editorial is an outgrowth of our research efforts that resulted in the SIGCOMM'93 paper [1] entitled *On the self-similar nature of Ethernet traffic.* We discuss some lessons we have learned as we have watched the published findings being absorbed by the scientific community in general and the networking community in particular. We focus on aspects that have remained relevant today, especially at a time when, with the proliferation of Artificial Intelligence and Machine Learning, networking research has become increasingly data-dependent and data-driven.

CCS CONCEPTS

• **Networks** → **Network performance evaluation**; **Network measurements**;

KEYWORDS

Ethernet LAN measurements, Self-similarity, Heavy-tailed distributions

1 INTRODUCTION

Our Sigcomm'93 paper "On the self-similar nature of Ethernet traffic" [1] was published at a time when a fledgling *packet-switched network*, which became known as the "Internet", was getting increasingly complex. It was experiencing an exponential growth in size (e.g., users, networks, infrastructure) and was required to support an ever-growing number of protocols and services. In spite of this, the engineers responsible for operating and managing the different bits and pieces of this network had to rely almost exclusively on knowledge involving *circuit-switched networks* which provided plain old telephone service (POTS).

POTS predated the Internet and was well-studied. It benefitted from decades of research and experience with designing, managing and operating circuit-switched networks. Modifying these POTS-based traffic models to study packet-switched networks had become a minor industry around 1990. But how well these newly proposed POTS-based traffic models and techniques fared in practice was anyone's guess because measurements of actual traffic over this emerging Internet were basically non-existent.

Our work described in [1] put an end to this guessing game. Using a first-of-its-kind dataset consisting of hundreds of millions of high-quality packet-level traffic measurements that were collected from real-world local-area networks (LAN) over a period of some three years, we demonstrated that (i) measured Ethernet LAN traffic is statistically self-similar, (ii) none of the commonly-used traffic models is able to capture this fractal-like behavior, and (iii) such behavior has practical implications for the design, control, and performance analysis of high-speed, packet-switched networks. Our findings were supported by a rigorous statistical analysis of the available data and were accompanied by a discussion of the underlying mathematical and statistical properties of self-similarity and their relationship with actual network behavior. They also hinted at a likely root cause of the observed self-similarity behavior of aggregate LAN traffic, namely the "high variability" behavior in the online activity of a typical LAN user.

An extended version of our Sigcomm'93 paper was subsequently published in the journal IEEE/ACM Transactions on Networking [2] where it not only garnered the 1995 W.R. Bennett Prize Paper Award from the IEEE Communications Society (for the most outstanding paper reporting original work in the IEEE/ACM Transactions on Networking in 1994) but also the 1996 IEEE W.R.G. Baker Prize Award from the IEEE Board of Directors (for the most outstanding paper reporting original work in all the Transactions, Journals, and Magazines of the IEEE Societies or in the Proceedings of the IEEE in 1994). In addition, the paper was also included in "The Best of the Best: Fifty Years of Communications and Networking Research", a 2007 compilation of the 50 most outstanding papers published by the IEEE Communication Society in its various journals in the society's 50-year history [3].

2 LESSONS LEARNED

While the reported discovery of the self-similar nature of packet-switched network traffic surprised network researchers and engineers alike, what really stunned many of them was the ability to clearly distinguish the measured traffic from traditional telephony traffic or from traces generated by commonly-used traffic models. At the same time, they found the ubiquity of the observed behavior across datasets collected at different points in the network and over a period of three years convincing and viewed the mathematical and statistical properties of self-similarity as being both intriguing and puzzling, especially in view of the common understanding of network traffic that prevailed at that time. Below, we discuss a number of key lessons that this research effort has taught us as we have watched the published findings being absorbed by the scientific community in general and the networking community in particular. Among the main lessons learned are the benefits of a fortuitous collaboration, the importance of timing, the significance of reproducibility, and an appreciation for mathematical models that are not merely evocative (i.e., able to reproduce a phenomenon of interest) but in fact explanatory (i.e., able to capture key causal mechanisms).

"Alone we can do so little, together we can do so much." (H. Keller, 1880-1968)

The pre-Web Internet of the late 1980s/early 1990s was a general-purpose research network (i.e., NSFNET) that used packet switching technology and connected five supercomputer centers along with the National Center for Atmospheric Research (NCAR) to each other and to the regional research and education networks that in turn connected to campus networks. Access within this three-tier network structure was provided using the TCP/IP protocol stack that had been successfully deployed on the ARPANET starting in 1969. As the research community started to rely more and more on this rapidly expanding "internetwork" or Internet, by 1990, Ethernets had arguably become the most successful and widely used LAN technology for campus networks or corporate research organizations.[1] Around that time, network engineers and operators had also become more critical of past traffic modeling work that generally ignored empirical evidence and relied instead on theoretical constructions that had little in common with the type of traffic that they observed in their networks. In view of these developments, Ethernet LANs were widely considered to be prime targets for collecting real-world traffic measurements that could be used to examine what measured packet-switched network traffic actually looks like.[2]

Against this backdrop of growing interest in all aspects concerning this emerging "Internet", the work that resulted in [1] started in mid-1987 when one of the authors (DW) who worked as a researcher as well as a network administrator in the Applied Research Area within Bell Communications Research (Bellcore) repurposed available hardware to collect high-quality traces of actual Ethernet traffic. Seeing first-hand how real network packet traffic behaves, aware of an apparent mismatch between existing mathematical traffic models and measured traffic, and frustrated with the limitations and poor quality of the network monitoring/diagnosis tools that were available at that time, he decided to build from scratch a first-of-its-kind piece of equipment that would allow the collection of complete (i.e., no losses), high time-resolution (i.e., time stamp accuracy of about 10-100 microseconds), and detailed (i.e., time stamp, status, length and 60 bytes of header information for every packet) packet-level traces from some of Bellcore's 10 Mbps Ethernet LANs for a number of different week-long periods.

After extensive stress-testing of the equipment, actual data collection began in 1989. The data was recorded on 8 mm tapes (about one tape for each full day of recording), and by the end of 1991, the tapes had filled multiple shoe boxes (literally!), and the datasets became known as the "Bellcore traces". By that time, DW had already teamed up with another author of [1] (WL) who was a networking researcher and colleague of his at Bellcore. As some of their findings from a preliminary analysis of some of the early segments of these Bellcore traces started to appear in print [5–7], they caught the attention of the other two authors of [1] (MS, WW), one being a colleague of WL and DW at Bellcore, and the other a professor at Boston University and one of the foremost experts on the theory of self-similar processes and their applications.

In particular, what initiated the ensuing fortuitous collaboration between two networking researchers and two researchers in the mathematical sciences was the use in [5] of the metaphor "traffic spikes ride on longer-term ripples, that in turn ride on still longer-term swells" to describe the most surprising finding of an initial analysis of these Bellcore traces. From the perspective of self-similar processes, this metaphor succinctly captured the intuition behind (statistical) self-similarity; that is, burstiness or intermittent behavior exists over a wide range of time scales and, contrary to conventional wisdom, applying simple low-pass filters to the original traffic trace does not result in smoother traffic. It was the use of this metaphor that established an initial connection between a new type of network data in the form of the Bellcore traces (i.e., high-volume, high-quality, and information-rich packet-level traffic measurements) and a relatively old but largely unknown topic in the mathematical sciences – the theory of self-similar processes and their applications.

In fact, self-similar processes had been extensively studied by various probabilists since the days of Kolmogorov [9]. During the 1960s and 1970s, they were brought to the attention of some statisticians and applied scientists by B. Mandelbrot and his co-workers. Widely recognized as "the father of fractals," Mandelbrot focused on and popularized the essence behind self-similarity, namely the notion of "scale-invariance," and he demonstrated the relevance of scaling phenomena in a wide range of areas in the physical, social and biological sciences [10]. For example, although not directly related to data networks, Mandelbrot's paper [11] was one of the first applications of the concept of self-similarity to communication systems. With regard to the problem of statistical inference for self-similar phenomena, a number of heuristic graphical methods had already been known and used for decades, including variance-time-type analysis, R/S-analysis, and some periodogram-based techniques (e.g., see [12] and references therein). These and other more recently considered methods got soon replaced by wavelet-based techniques, mainly because the latter were shown to be computationally feasible even for large datasets and to have desirable robustness properties (e.g., see [13] and references therein). Nevertheless, in the early 1990s when we performed our analysis of the Bellcore traces, it was difficult to think of any other area in the physical, social or biological sciences where the available data was so voluminous and high-quality and provided such detailed information about so many different facets of behavior as our Bellcore traces. This made the "mining" of these data especially challenging.

In the process of our collaboration and in an attempt to help building an effective "bridge" between data-analytic efforts for a new generation of datasets like the Bellcore traces and modeling-related efforts that leveraged the promising but still underdeveloped mathematical and statistical theory of self-similar processes, we were in constant contact with researchers in the areas of networking, stochastic processes and statistics. As a result, we had to overcome a number of theoretical and practical challenges, including (i) how to go about dealing with the unheard of amounts of available high-quality data, (ii) how to present the results of our analysis in a visually intuitive manner (knowing our audience includes network

[1] The intended use of the early NSFNET was "primarily for research and education in the sciences and engineering" and its use for commercial purposes was in general not allowed [4].

[2] Interest in real-world Ethernet LANs and their performance was further fueled by theoretical studies at that time that alleged that "Ethernet works in practice, but not in theory", a myth that was subsequently dispelled in [8].

engineers) without sacrificing statistical rigor, and (iii) how to reduce the observed self-similarity phenomenon to a level where it could be explained and validated in simple networking terms. For example, it was an all-out collaborative effort to express our "let the data do the talking" approach to data analysis in a form that ultimately resulted in the "a picture is worth a thousand words" plot sequences in [2, Figure 4][3] that demonstrated how simple and easy it is to clearly distinguish between measured network traffic and synthetic traffic generated by some of the most popular and widely-used packet traffic models at that time.

"Observe due measure, for right timing is in all things the most important factor." (Hesiod, around 700 BC)

New discoveries or observations only rarely occur in isolation. Instead, they are often a response to changing conditions which question traditionally-made (or previously validated) assumptions and are sufficiently interesting and challenging, both from a theoretical and practical perspective, so as to attract similarly-minded researchers in different parts of the world, with potentially very different backgrounds. In this sense, the discovery of the self-similar nature of packet-switched network traffic was no exception. The overall level of interest in and awareness of the problem (i.e., how does real-world packet-level network traffic look like, and are existing traffic models consistent with measured traffic?) was sufficiently high, the ingredients needed to move beyond well-established boundaries (i.e., plenty of data) were becoming available, and the theoretical groundwork (i.e., self-similar stochastic processes and their applications) had already been established. In short, the time was right!

With respect to measurements, the activities at Bellcore around 1990 were by no means an isolated effort. In fact, around the same time and even prior to 1990, a number of networking research groups were involved in projects that were intended to examine the relevance of existing traffic models in view of the emergence and popularity of packet-switched networks and performed their own traffic measurements and trace collections. For example, [14] describes monitoring a token ring network at MIT in 1984 and collecting some 11 million packets over a week-long period. This dataset formed the basis for the analysis reported in [15] that concluded that measured packet arrival times are neither Poisson nor Compound Poisson but are well described by a so-called "packet train" model. Other noteworthy efforts include the pioneering measurement study of an Ethernet LAN at Xerox PARC around 1980 [16], an examination of measured Ethernet LAN traffic generated by diskless workstations [17], early work on characterizing the observed traffic on the NSFNET national backbone [18], and initial efforts on measuring and studying wide-area TCP/IP traffic [19]. However, compared to these and similar efforts, the Bellcore traces were unique and exceptional in terms of their high volume, high quality, rich amount of meta-data, and diversity in time (over a period of three years) and space (different Ethernet LANs). The collection of recorded packet-level traffic traces at Bellcore was arguably one of the first example of "big data" in the Internet measurement area and set the standard for a field that would come into its own some ten years later.

In terms of the self-similarity discovery itself, our work in early 1990 turned out to coincide with three different research efforts that were going on completely independent from one another – one in Japan, one in Finland, and the third one in a different part of Bellcore. The research effort in Japan was carried out in the early 1990's at the NTT Telecommunication Network Laboratories and was brought to our attention around 1995 by H. Saito. The work involved measured ATM traffic that was generated by a video application and was found to exhibit characteristics consistent with fractal or self-similar behavior. Unfortunately, the work describing this study was published in hard-to-find places (even for Japanese researchers, e.g. [20, 21]) and remained unknown outside of Japan. In Finland, the research activity was carried out by I. Norros, a mathematician working for VTT Information Technology. Based on a plot in [6] that suggested that measured LAN traffic had a variance-time behavior consistent with a self-similar process, he proposed in 1993 a new traffic model based on fractional Brownian motion that was capable of accounting for the empirically observed behavior and also resulted in a drastically different queuing performance compared to existing and widely-used traffic models [22]. A third independent effort was pursued by another group of researchers at Bellcore around 1990 that included A. Erramilli and P. Pruthi. In an attempt to develop accurate and parsimonious new models of packet traffic that can be used by traffic engineers and network designers, that group demonstrated that by modeling packet traffic using deterministic nonlinear chaotic maps, it was possible to capture the reported wide range of burstiness in measured packet-level traffic. A complete account of that work was published in 1995 [23] (see also [24]), at which time we already had combined forces to develop measurement-based traffic engineering techniques that accounted for the observed self-similar behavior and replaced traditional POTS-based methods.

Although all three of these efforts concluded that in view of some apparent fractal- or self-similar-like behavior of measured packet-level traffic, performance modeling for modern communication networks had to be revisited if not completely overhauled, our self-similarity discovery was supported by superb measurements, our data-analytic effort was comprehensive and fully exploited the available "big data", and with our subsequent efforts [30], we were able to move beyond the "descriptive phase" and provided a physical understanding of the causes and effects of self-similarity in measured network traffic.

"Non-reproducible single occurrences are of no significance to science." (K. Popper, 1902-1994)

In modern science in general and in the area of Internet measurement in particular, reproducibility is widely considered to be a necessary condition for determining whether or not the results reported by an author are an artifact of the measurement procedures or the statistical analysis techniques used by that author in the course of her experiment. Here, reproducibility refers to the ability of researchers other than that author to obtain the same results when performing the author's original experiment under different conditions and using their own independently-developed

[3] One half of that plot sequence was already included in [1, Figure 1].

artifacts [25]. In short, a reported experimental result is not considered fully established unless it has been independently reproduced.

Just as our collection of Bellcore traces was an early instance of truly "big (Internet) data", our work was also one of the first examples in the area of Internet measurement that both inspired and benefitted from extensive and early reproducibility efforts. In fact, while the speed with which the networking community tried to reproduce the measurement and analysis work we had done at Bellcore, but in very different networking environments and under very different conditions, was impressive, it was also an indication of how much the networking community valued "reproducibility" in the context of gaining a better understanding of the nature the network traffic as modern packet-switching technology replaced traditional circuit-switching. In practice, reproducibility offers a viable alternative to conventional statistical model validation methods that put a strong emphasis on a model's data-fitting ability but are at a complete loss when it comes to dealing with the large number of large data sets collected from today's packet networks.

One piece of "reproducibility" work that stands out as one of the most important validations of our Ethernet analysis studies is the WAN traffic study performed by V. Paxson and S. Floyd. Within one year of the original publication of our results [1], Paxson and Floyd had analyzed packet-level traffic traces from a number of different WANs and described their results in a paper published at Sigcomm'94 [26].[4] Their work not only confirmed our self-similarity discovery in the context of their WAN traffic traces, but they also demonstrated that self-similarity comes in different "shades". That is, while Ethernet LAN traffic tends to be exactly self-similar, measured WAN traffic has a tendency to be asymptotically self-similar in the sense that it only exhibits self-similar scaling for a range of sufficiently large time scales. Their work set the standards for future WAN traffic studies, and their papers became landmark papers in their own right.

Another important piece of "reproducibility" work that has to be mentioned in this context appeared right after the emergence of the Web in 1994 and the resulting explosive growth of WWW-related traffic on the Internet. Using traces collected from both Web clients and Web servers, M. Crovella and A. Bestravos showed in [28] (see also [29]), among other findings, that the emergence of a new "killer application" such as the Web did not change the self-similar nature of Internet traffic at the packet level. Their work contributed to the present understanding that self-similarity is an ubiquitous property (i.e., an invariant) of modern Internet traffic, irrespective of what applications or protocols are responsible for the bulk of the traffic.

"Every solution to a problem is a new problem." (J. W. von Goethe, 1749-1832)

Discovering a phenomenon like self-similarity in measured network traffic is one problem, describing and quantifying that phenomenon using mathematics and statistics is another problem, but what about providing insight into and understanding of the observed phenomenon by explaining it in terms of the underlying root cause(s)? In addition to the "big data" aspects of our study and the subsequent reproducibility efforts it motivated, what ultimately set our work apart from prior network traffic analysis and modeling endeavors was the ability to explain self-similarity in simple networking terms that were intuitively appealing, coincided with engineering experience, agreed with an analysis of the measurements that more fully exploited the information available in the collected data, and could be shown mathematically to cause exactly the self-similar behavior that we observed in the actual packet traces.

To illustrate this physical-based explanation that goes after the root cause of the observed self-similarity phenomenon in measured network traffic, consider for example a typical LAN setting, where the aggregate traffic consists of the superposition of a large number of individual host-host packet traffic streams or connections. It turns out that self-similarity of the aggregate traffic (i.e., packets from all flows or connections) is caused by an "on-off" behavior of the individual host-host streams where during the "on" or burst-periods, packets are sent at a constant rate, where no packets are sent during the "off" or idle-periods, and where, more importantly, the burst- or idle-periods exhibit high variability.[5] While some initial results of and conjectures concerning this physical explanation of the self-similar nature of network traffic can already be found in the original papers on the self-similar nature of Ethernet LAN traffic [1] and wide-area traffic [26], in the case of the Ethernet LAN settings considered in [1], we presented a complete explanation in a subsequent paper that appeared at Sigcomm'95 [30]. That work described the appropriate mathematical framework, the type of host-host traffic measurements that needed to be extracted from the Bellcore traces, and the statistical analysis of those measurements that confirmed the high variability phenomenon in measured traffic of individual host-host packet-level traces (see also [31]). In the case of Internet WAN traffic, a similar, yet slightly different explanation holds and has been discussed in, for example, [26, 28, 32].

Subsequently, this physical explanation of the self-similar nature of aggregate network traffic in terms of high variability of the burst- or idle-periods at the level of individual host-host traffic or connections motivated further work aimed at "digging deeper"; that is, explaining the root cause(s) for the observed high variability associated with individual host traffic. For example, as part of their effort to demonstrate the self-similar nature of measured Web traffic [28], M. Crovella and A. Bestavros also provided empirical evidence that self-similarity in measured Web traffic can be explained based on Web file transfer times having high variability, a property that is primarily due to the distribution of available Web file sizes being heavy-tailed with infinite variance. They further hypothesized that the deeper root cause for Web traffic self-similarity lies in the basic characteristics of human information organization and retrieval.

This hypothesis was made more formal in [33] where the authors considered Web layout design in the spirit of source coding for data compression and rate distortion theory. Leveraging this novel approach to Web layout and access, they showed that minimizing the average size of files downloaded during Web browsing sessions subject to certain navigability constraints with respect to the Web layout design produces heavy-tailed distributions for Web file transfers. As such, their approach provides a natural and plausible explanation for the origin of observed Web traffic self-similarity.

[4] A more detailed version of their paper appeared subsequently in *IEEE/ACM Transactions on Networking* [27].

[5] Mathematically speaking, we require the distributions of the burst/idle periods to be heavy-tailed with infinite variance.

Together, [28] and [33] make it clear that self-similar Web traffic is neither a machine-induced artifact nor the result of particular user behavior. While particular changes in protocol processing or Web document display are not likely to ever eliminate the observed self-similarity of Web traffic, they can be expected have an impact on the precise nature of self-similarity (e.g., intensity, strict vs. asymptotic self-similarity). Importantly, these findings suggest that the self-similar nature of Internet traffic may indeed be intrinsic to any application which organizes information for human consumption.

Given these various successful efforts in establishing the main root cause(s) of the self-similar nature of aggregate network traffic, it is natural to expect a gradual decrease in interest in the phenomenon itself – a case of a once “hot topic” going “cold” [34]. After all, while early on, self-similarity was considered to be a surprising or even esoteric property of aggregate network traffic, these subsequent efforts on explaining the phenomenon demonstrated that it was just an outward sign of a more fundamental and ubiquitous property of packet network traffic – high variability (as parsimoniously captured by heavy-tailed distributions with infinite variance) in the sizes (e.g., number of bytes) or the durations (e.g., number msec) of data transfers at the level of individual users, connections or flows. Moreover, with the further understanding that self-similarity in Internet traffic may be intrinsic to any application which organizes information for human consumption, why waste time and effort to analyze more traffic traces, unless they show a drastic shift towards a new generation of Internet applications or services that have little to do with humans or with the way humans organize, access, or consume information, at which time the topic of self-similarity and long-range dependence may experience some “re-heating” and become “hot” again [35].

“In theory there is no difference between theory and practice ...” (Y. Berra?)

Ever since the pioneering work of B. Mandelbrot, many of the theoretical studies on self-similar processes have been concerned with mathematical constructions of stochastic processes capable of modeling stylized features of observational time series from areas such as networking, finance, and turbulence. A key goal of these efforts has been to find models with analytically and stochastically tractable correlation structures displaying either weak or strong dependence and also having marginal distributions that are infinitely divisible and hence related to either Poisson, Gaussian, or stable processes. A number of these efforts have been pursued in the context of packet-switched networks where the observed traffic typically results from a large number of individual communication exchanges or connections that involve different hosts on the network and are governed by the TCP/IP protocol stack. This setting is especially appealing in view of Mandelbrot’s original idea of using sums of renewal-reward processes with heavy-tailed distributions for the inter-renewal times to study self-similarity phenomena in measured time series data [36]. His idea was subsequently implemented for general renewal-reward processes in [37] and for network traffic-specific “on”-periods and “off”-periods in [38] and shown to yield either fractional Brownian motion or a stable Lévy motion in the limit, depending on the precise nature of the rescaling scheme used.

Motivated in part by its applications to the networking area, since the early 2000s, the mentioned line of research has seen significant advances, including the introduction of the family of “telecom processes” to describe intermediate limiting regimes between fractional Brownian motion and stable Lévy motion [39, 40]. More recently, another family of processes that shares a structure similar to these telecom processes has been obtained as limits of the so-called “Ornstein- Uhlenbeck (OU) processes” driven by Lévy noise and their superpositions (i.e., supOU processes) [42]. An attractive feature of these new processes is that they allow the marginal distribution and the dependence structure to be modeled independently from each other. Moreover, they offer a flexible choice of different forms of correlation functions, with the class of finite variance stationary supOU processes containing examples where the correlation function decreases like a power function as the lag increases (i.e., long-range dependence). In fact, depending on the conditions on the underlying supOU process, four different limiting processes may be obtained after suitable normalization; namely, Brownian motion, fractional Brownian motion, stable Lévy process, and stable process with dependent increments akin to the one that appears in the previously considered setting of telecom processes [43].

From a networking perspective, knowledge of these theoretical developments is important in case the Internet as we know it today should undergo significant changes that have the potential of drastically changing the nature of its traffic. For example, such a change could be due to a move towards a future scenario where the traffic is dominated by machine-to-machine communications and no longer by human interactions or by applications in support of how humans organize, access, and consume information. Another example of such a change could be the design and implementation of new protocols that are capable of fully exploiting more bandwidth-rich future infrastructures by almost instantaneously blasting packets into the network at possibly very high rates and potentially for significant periods of time. Such possible future scenarios may be more consistent with some of the limiting regimes that this line of research is exploring and that differ from the fractional Brownian motion limiting behavior that has been an ubiquitous property of aggregate Internet traffic to date. At the same time, it will be important that these theoretical studies be accompanied by advances in the development of statistical inference and modeling methods suitable for these new processes.

“... In practice there is [a difference between theory and practice].” (Y. Berra?)

A rather obvious but rarely mentioned implication of the empirical finding of the ubiquity of self-similarity in measured Internet traffic is that, from a practical perspective, it brought to light a dramatic difference between modern packet-switched network traffic and traditional circuit-switched telephony traffic . In a nutshell, while monitoring a compact set of known features (e.g., call arrival rate, call duration) suffices to accurately describe and capture the inherent uncertainty in actual telephony traffic so as to detect and alert on “atypical” behavior with high confidence, actual Internet traffic affords no such convenience. In fact, monitoring any finite set of metrics derived from observed Internet traffic is in general insufficient when dealing with the seemingly chaotic fluctuations exhibited by fractal or self-similar dynamic processes. That is, real Internet traffic has always ways to "surprise" even the most ardent and diligent observers or measurers, precluding therefore the sort of telephony-like traffic monitoring and data analytics that can successfully identify "typical" behavior so as to inform operators of any "anomaly" or deviation from what is considered to be normal or typical.

This observation has profound implications for network monitoring and also for network management in general and network security and network performance in particular in packet-switching networks such as the Internet. In fact, understanding that the root cause of self-similarity in aggregate network traffic is high-variability behavior (expressed mathematically in the form of heavy-tailed distributions with infinite variance) at the level of individual connections, flows, or sessions, the observed scale-invariant property is not limited to the packet level (i.e., layers 2 or 3 of the TCP/IP protocol stack) but also manifests itself all the way up to the application level (i.e., layer 7) where it appears in the form of scale-invariance with respect to the size or duration of sessions, connections, or flows. In effect, this type of scale invariance at the higher layers implies that in terms of size (e.g., number of bytes) or duration, there are no "typical" sessions, connections, or flows so that even at the higher layers, actual Internet traffic is essentially unpredictable and full of "surprises". Thus, any diligent network operator has to accept the fact that in reality, never-seen-before events, be they nefarious or benign, or performance-impairing conditions with potentially catastrophic consequences can occur with non-negligible probability and can manifest themselves at any time at any of the different layers across the entire TCP/IP protocol stack.

In view of the operational challenges posed by such a holistic and practical view of modern network traffic, it is not a coincidence that two of the pioneers of the early work on the self-similar nature of network traffic, P. Pruthi and V. Paxson, sooner or later turned their attention to commercializing their deep understanding of the defining features and properties of real-world network traffic, with an eye towards developing products that could help network operators with network security-related problems or network-performance-specific issues. While a common theme in both of their efforts has been the importance of network monitoring for obtaining all the "right" data to keep a network safe and performant, their efforts differ in how to obtain that data and what the "right" data ought to be. In particular, soon after finishing his PhD at the KTH Royal Institute of Technology, P. Pruthi founded NIKSUN, Inc. [44] and has led the company from a garage start-up in 1997 to a successful enterprise with over a thousand customers worldwide, including Fortune 500 companies, government agencies, financial institutions, and service providers. Over the years, by commercializing its patented real-time analysis and recording technology, the company has develop an industry-leading suite of scalable, forensics-based cyber security and network performance monitoring products that is capable of capturing, inspecting, mining, correlating, and storing every piece of data traversing an enterprise's network at rates through 100 Gbps and beyond. With this capability that is exclusive to NIKSUN, network operators and cybersecurity analysts are able to inspect and analyze all network activity, continuously in time, network-wide, and across the entire TCP/IP protocol stack, providing them with a systematic and practical means for dealing with the "surprises" that are intrinsic to modern network traffic. V. Paxson, whose career also started in the mid-1990s, took a less direct path to co-founding a company. After finishing his PhD at Berkeley in 1997 where he was also a staff member at the Lawrence Berkeley National Laboratory (LBNL), he began developing the Bro Intrusion Detection System, an open-source network security monitoring tool [45] that enables users to tailor Bro's analysis of the monitored traffic to the specifics of the local environment. Over time, Bro has attracted a large user community and been used to monitor networks of major universities, large research labs, supercomputing centers, and open-science communities around the country. During that time, V. Paxson joined the International Computer Science Institute (ICSI) at Berkeley in 1999, and in 2007 he was appointed a professor of the Electrical Engineering and Computer Sciences Department of UC Berkeley, all the while keeping Bro up-to-date and improving its performance and its integration into operational deployments. At long last, in 2013, he co-founded Corelight [46], a start-up that specializes in developing network security monitoring solutions that leverage Bro, now named Zeek. Corelight products simplify Zeek deployment and are used by customers worldwide, including Fortune 500 companies, major government agencies, and large research universities.

Acknowldegements

Murad S. Taqqu's work was supported by a grant from the Simons Foundation/569118MT at Boston University.

REFERENCES

[1] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of Ethernet traffic. *ACM SIGCOMM Comput. Commun. Rev. (Proc. ACM SIGCOMM'93)*, 23(4):183–193, 1993.
[2] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of Ethernet traffic (Extended version). *IEEE/ACM Trans. on Networking*, 2(1):1–15, 1994.
[3] W. H. Tranter, D. P. Taylor, R. E. Ziemer, N. F. Maxemchuk, and J. W. Mark (Editors). The Best of the Best: Fifty Years of Communications and Networking Research. *IEEE Communications Society*, Wiley-IEEE Press, 2007.
[4] NSF. The launch of the NSFNET. *https://web.archive.org/web/20170202190224/https://www.nsf.gov/about/history/nsf0050/internet/launch.htm*
[5] W. E. Leland. LAN traffic behavior from milliseconds to days. *Proc. 7th ITC Specialist Seminar*, Morristown, NJ, 1990.
[6] W. E. Leland and D. V. Wilson. High time-resolution measurements and analysis of LAN traffic: Implications for LAN interconnections. *Proc. IEEE Infocom'91*, Bal Harbor, FL, 1991.
[7] H. J. Fowler and W. E. Leland. Local area network traffic characteristics, with implications for broadband network congestion management. *IEEE Journal on Selected Areas in Communication*, 9(7):1139–1149, 1991.
[8] D. R. Boggs, J. C. Mogul, and C. A. Kent. Measured capacity of an Ethernet: Myths and Reality. *ACM SIGCOMM Comput. Commun. Rev. (Proc. ACM SIGCOMM'88)*, 18(4):222–234, 1988.
[9] A. N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *C. R. Acad. Sci. URSS (N.S.)*, 30:301–305, 1941. (Translation in *Turbulence*. S. K. Friedlander and L. Topper (Eds.), Interscience, New York, 1961.)
[10] B. B. Mandelbrot. *The Fractal Geometry of Nature.* W. H. Freeman and Co., New York, 1983.
[11] B. B. Mandelbrot. Self-similar error clusters in communication systems and the concept of conditional stationarity. *IEEE Trans. on Communication Technology*, 13(1):71-99, 1965.
[12] J. Beran. *Statistics for Long-Memory Processes.* Chapman & Hall, New York, 1994.
[13] P. Abry and D. Veitch. Wavelet analysis of long-range dependent traffic. *IEEE Trans. on Information Theory*, 44(1):2–15, 1998.
[14] D. C. Feldmeier. Empirical Analysis of a Token Ring Network. 1984.
[15] R. Jain and S. Routhier. Packet Trains-Measurements and a New Model for Computer Network Traffic. *IEEE Journal on Selected Areas in Communications*, 4(6):986–995, 1986.

[16] J. F. Shoch and J. A. Hupp. Performance of the Ethernet local network. *Commun. ACM*, 23(12):711–721, 1980.
[17] R. Gusella. A measurement study of diskless workstation traffic on an Ethernet. *IEEE Trans. on Communications*, 38(9):1557–1568, 1990.
[18] S. A. Heimlich. Traffic characterization of the NSFNET national backbone. *ACM SIGMETRICS Performance Evaluation (Proc. ACM SIGMETRCS'90)*, 18(1):257–258, 1990.
[19] R. Caceres, P. B. Danzig, S. Jamin, and D. J. Mitzel. Characteristics of wide-area TCP/IP conversations. *ACM SIGCOMM Comput. Commun. Rev. (Proc. ACM SIGCOMM'91)*, 21(4):101–112, 1991.
[20] S. Shimogawa, S. Nojo, and T. Betchaku. A fractal property in an ATM video statistical multiplexing. *The 1992 Fall National Conference of ORSJ*, 1-C-6, 1992 (in Japanese).
[21] S. Shimogawa, S. Nojo, and T. Betchaku. Self-similar phenomena in an ATM video cell flow. *Proc. Symp. Performance Models in Information and Communication Networks*, M. More et al. (Eds.), Namazu, Japan, 1993.
[22] I. Norros. A storage model with self-similar input. *Queueing Systems*,16(3-4):387–396, 1994.
[23] A. Erramilli, R. P. Singh, and P. Pruthi. An application of deterministic chaotic maps to model packet traffic. *Queueing Systems*, 20(1-2):1771–206, 1995.
[24] P. Pruthi. *An Application of Chaotic Maps to Packet Traffic Modeling.* Ph.D. Thesis, Kungliga Tekniska Hogskolan, Stockholm, Sweden, 1995.
[25] ACM. Artifact Review and Badging. *https://www.acm.org/publications/policies/artifact- review- badging*, 2018.
[26] V. Paxson and S. Floyd. Wide-area traffic: The failure of Poisson modeling. *ACM SIGCOMM Comput. Commun. Rev. (Proc. ACM SIGCOMM'94*, 24(4):257–268, 1994.
[27] V. Paxson and S. Floyd. Wide-area traffic: The failure of Poisson modeling. *IEEE/ACM Trans. on Networking*, 3(3):226-244, 1995.
[28] M. E. Corovella and A. Bestravos. Self-similarity in World-Wide-Web traffic - evidence and passible causes. *Proc. ACM SIGMETRICS'96*, Philadelphia, PA, 1996.
[29] M. E. Crovella and A. Bestravos. Self-similarity in World-Wide-Web traffic - evidence and possible causes. *IEEE/ACM Trans. on Networking*, 5(6):835–846, 1997.
[30] 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. *ACM SIGCOMM Comput. Commun Rev. (Proc. ACM SIGCOMM'95)*, 25(4):100-113, 1995.
[31] 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. *IEEE/ACM Trans. on Networking*, 5(1):71 - 86, 1997.
[32] T. G. Kurtz. Limit theorems for workload input models. *Stochastic Networks: Theory and Applications*, F. P. Kelly, S. Zachary, and I. Ziedins (Eds.), Clarendon Press, Oxford, UK, 1996.
[33] X. Zhu, J. Yu, and J. Doyle. Heavy Tails, Generalized Coding, and Optimal Web Layout *Proc. IEEE Infocom'01*, 2001.
[34] J. Crowcroft. Cold Topics in Networking. *ACM SIGCOMM Comput. Commun. Rev.*, 38(1):45–47, 2008.
[35] J. Crowcroft. Reheating Cold Topics (in Networking). *ACM SIGCOMM Comput. Commun. Rev.*, 39(2):48–49, 2009.
[36] B. B. Mandelbrot. Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International Economic Review*, 10:82-113, 1969.
[37] J. B. Levy and M. S. Taqqu. Renewal reward processes with heavy-tailed inter-renewal times and heavy-tailed rewards. *Bernoulli*, 6(1):23–44, 2000.
[38] M. S. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *ACM SIGCOMM Comput. Commun. Rev.*, 27(2):5–23, 1997.
[39] R. Gaigalas and I. Kaj. Convergence of scaled renewal processes and a packet arrival model. *Bernoulli*, 9:671–803, 2003.
[40] M. Lifshits. *Random Processes by Example.* World Scientific, Singapore, 2014.
[41] V. Pipiras and M. S. Taqqu. *Long-Range Dependence and Self-Similarity.* Cambridge University Press, Cambridge, UK, 2017.
[42] O. E. Barndorff-Nielsen, F. E. Benth, and A. E. D. Veraart. *Ambit Stochastics.* Springer International Publishing, 2018.
[43] D. Grahovac, N. N. Leonenko, and M. S. Taqqu. Limit theorems, scaling of moments and intermittency for integrated finite variance supOU processes. *arXiv:1711.09623*, 2017.
[44] NIKSUN, Inc. *https://www.niksun.com*
[45] V. Paxson. Bro: A System for Detecting Network Intruders in Real-Time. *Proc. 7th USENIX Security Symposium*, San Antonio, TX, 1998.
[46] Corelight. *https://www.corelight.com.*