Colloquium

Scaling phenomena in the Internet: Critically examining criticality

Walter Willinger*[†], Ramesh Govindan[‡], Sugih Jamin[§], Vern Paxson[‡], and Scott Shenker[‡]

*AT&T Labs–Research, Florham Park, NJ 07932-0971; [‡]International Computer Science Institute (ICSI), Berkeley, CA 94704-1198; and [§]Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI 48109-2122

Recent Internet measurements have found pervasive evidence of some surprising scaling properties. The two we focus on in this paper are self-similar scaling in the burst patterns of Internet traffic and, in some contexts, scale-free structure in the network's interconnection topology. These findings have led to a number of proposed models or "explanations" of such "emergent" phenomena. Many of these explanations invoke concepts such as fractals, chaos, or self-organized criticality, mainly because these concepts are closely associated with scale invariance and power laws. We examine these criticality-based explanations of self-similar scaling behavior—of both traffic flows through the Internet and the Internet's topology—to see whether they indeed explain the observed phenomena. To do so, we bring to bear a simple validation framework that aims at testing whether a proposed model is merely evocative, in that it can reproduce the phenomenon of interest but does not necessarily capture and incorporate the true underlying cause, or indeed explanatory, in that it also captures the causal mechanisms (why and how, in addition to what). We argue that the framework can provide a basis for developing a useful, consistent, and verifiable theory of large networks such as the Internet. Applying the framework, we find that, whereas the proposed criticality-based models are able to produce the observed "emergent" phenomena, they unfortunately fail as sound explanations of why such scaling behavior arises in the Internet.

Today's Internet is a prime example of a large-scale, highly engineered, yet highly complex system. It is characterized by an enormous degree of heterogeneity any way one looks and continues to undergo significant changes over time. In terms of size, by mid-2001, the Internet consisted of about 120 million hosts, or endpoints, and more than 100,000 distinct networks, totaling millions of routers and links connecting the hosts to the routers and the routers to one another. These links differ widely in speed (from slow modem connections to high-speed "backbone" links) as well as technology (e.g., wired, wireless, satellite communication).

At the largest scale, the Internet is divided into Autonomous Systems (ASs). Each AS is a collection of routers and links under a single administrative domain. The global Internet currently consists of several thousand separate ASs, interlinked to give users the illusion of a single seamlessly connected network capable of providing a universal data-delivery service. The foundation of the ubiquitous connectivity is a datagram (packet) delivery mechanism termed the Internet Protocol, or IP.

Despite all of the efforts devoted to understanding today's Internet, it is still surprising how often networking researchers observe "emergent phenomena"—measurement-driven discoveries that come as a complete surprise, cannot be explained or predicted within the framework of the traditionally considered mathematical models, and rely crucially on the large-scale nature of the Internet, with little hope of encountering them when considering small-scale IP networks.

One example of such a discovery was that measured traffic rates on links in the Internet (i.e., number of packets or bytes that traverse a given link per time unit) exhibit self-similar (or "fractal-like") behavior: a segment of the traffic rate process measured at some time scale looks or behaves like an appropriately scaled version of the traffic rate process measured over a different time scale [see ref. 1 and followup studies (refs. 2 and 3)]. These empirical studies describe pertinent statistical characteristics of the temporal dynamics of measured traffic rate processes and provide ample evidence that these traces are consistent with asymptotic second-order self-similarity or, equivalently, long-range dependence (LRD); i.e., with autocorrelations that decay like a power for sufficiently large lags. These empirical findings were in stark contrast to what traditionally used models assumed about actual Internet traffic, namely exponentially fast decaying autocorrelations and, in turn, a classical white noise behavior when measuring traffic rates over large time scales.

A more recent example of an emergent phenomenon concerns the Internet AS graph, an aspect of the Internet's topology that describes network connectivity at the level of individual ASs. A recent empirical study (4) of the Internet's AS topology reported that the vertex degree distribution of snapshots of the measured AS-connectivity graph follows a power law. This finding implies that, whereas most of the ASs have a vertex degree of one or two, the probability of encountering a few ASs that are highly connected is significant. This data-driven observation is again in sharp contrast to the traditionally considered types of topology models (5, 6), which yield vertex degree distributions that decay exponentially fast, essentially ruling out the occurrence of high-degree vertices and giving most probability to "typical" node degrees on the order of the average node degree of the graph. In contrast, no such preferred or "typical" node degree can be identified for power-law vertex degree distributions and, because of this absence of a characteristic scale, the resulting structures are termed scale-free networks.

The discovery of the self-similar scaling behavior of Internet traffic over large time scales, the claim of scale-free characteristics of the Internet topology at the AS level, as well as other emergent phenomena such as the apparent intermittent nature of Internet congestion (7), and the reported multifractal scaling properties of Internet traffic over small time scales (8), have triggered renewed interest in Internet modeling. The ensuing research activities and resulting mathematical modeling efforts

can be roughly separated into two different categories: evocative or "descriptive" modeling vs. explanatory or "structural" modeling. Evocative models can produce (or approximate) the phenomena in question. Not only do explanatory models produce the phenomena, but also their applicability can be verified by "closing the loop," that is, by further measurements that test for the presence of the causes that are proposed to explain the phenomena of interest.

Evocative models are valuable in that they can shed light on how certain phenomena might arise. They are typically used to synthetically generate and statistically describe the phenomena of interest, and they can suggest what further measurements to take to test whether they are, in fact, explanatory. The ultimate scientific challenge, however, consists of developing explanatory models, because by demystifying emergent phenomena, these models provide a solid foundation for a useful, consistent, and verifiable theory for immensely complex systems such as today's Internet.

Our main purpose with this paper is to demonstrate that, because of its highly engineered nature and the many different facets of the available measurements, the Internet offers unprecedented opportunities for successful explanatory modeling. To this end, we provide a concrete framework for checking whether a proposed Internet-related model is indeed explanatory. We apply this framework to a number of recently introduced Internet traffic models (at the packet level) and Internet topology models (at the AS level). In the process, we find that the proposed statistical mechanics models of Internet traffic, where self-similarity is a signature of criticality caused by a phase-transition phenomenon (9, 10), and the recently considered scale-free graph models of the Internet topology, where power-law vertex degree distributions arise as a signature of self-organized criticality‖ (11, 12), are only evocative; they are not explanatory.

In particular, as a result of not "closing the loop," these models tend to be too generic in nature. Because they ignore important networking-specific details and fail to exploit the rich semantic content of the available measurements, they can lead to incorrect conclusions about the causes and origins of the emergent phenomena at hand.

On the other hand, we also show that certain mathematical models of Internet traffic, originally due to Mandelbrot (13) and Cox (14), are genuinely explanatory. Because they can "close the loop," these models lead to a fundamental understanding of emergent phenomena in the Internet context, thereby better advancing our knowledge about how such a large-scale and highly engineered man-made system works.

Modeling the Internet

Highly Engineered Yet Complex. Fundamental to the Internet's architecture is its design as a series of layers (15). Each layer relies on the next lower layer to execute more primitive functions and provides services to the next higher layer. Two hosts with the same layering architecture communicate with one another by having the corresponding layers in the two systems talk to one another. The latter is achieved by means of formatted blocks of data that obey a set of rules or conventions known as a protocol.

As briefly discussed earlier, the fundamental building block is a packet of data, routing and delivery of which is provided by IP. All information exchanges—whether a short e-mail message, a large file transfer, or a complicated Web transaction—are broken down into these basic building blocks. Each packet of each connection is self-contained in the sense that its header contains complete "addressing" information. The routers along the packet's path need only inspect the header of the packet to determine its next-hop destination and forward it through the network to its destination. Each packet is transmitted independently from the other packets that have already been transmitted or still await transmission; the routers do not keep track of which packets belong to which active connections. Thus, a router can forget about a packet as soon as it has been forwarded. This feature buys robustness in the sense that the network can transparently route packets around failed network components (e.g., links, routers) without perturbing active connections—it can continue to operate and successfully deliver data even in the face of major equipment failure.

IP's packet-oriented service buys efficiency and flexibility over the traditional connection-oriented service used in telephone networks. For example, instead of reserving a fixed amount of bandwidth for the exclusive use of communication between two end nodes (i.e., even if the two end nodes are silent, the resources cannot be shared by a third party), in a packet-oriented service network, each packet competes with all of the others. If there happens to be little competing traffic along a particular path, then a connection using the path can enjoy essentially the entire capacity. On the other hand, if many connections compete along the same path, each one of them will receive a (perhaps unfair) portion of the capacity. Furthermore, if packets arrive at a given point at too high a rate, such that they exhaust the router's finite buffer capacity for holding them pending further transmission, then the router will discard or drop the excess, a phenomenon termed congestion.

It is IP that provides the mechanism for unifying thousands of different networks, operating under diverse administrations. IP's main task is to adequately implement all of the mechanisms necessary to knit together divergent networking technologies and administrative domains into a single virtual network (an "internet") to enable data communication between sending and receiving hosts, irrespective of where in the network they are. The abstraction of end-to-end connectivity provided by IP serves as a layer that hides the underlying physical technologies. Further abstractions (e.g., reliable delivery, access to Web URLs) are then layered above IP. Thus, IP ensures a critical separation between the constantly evolving physical network infrastructure at lower levels and an ever-increasing user demand for more abstract services and applications at higher levels.

The layer above IP is termed the transport layer, where the most commonly used transmission control protocol (TCP) provides a number of additional services for end-to-end communication beyond those provided by IP: reliable delivery in the presence of lost packets; a "byte stream" abstraction that hides the underlying packetization; error recovery; flow control (ensuring that the sender does not overrun the receiver's ability to accept new data); and congestion control. This last means that TCP automatically adapts the rates at which data are transmitted depending on whether congestion is detected. Its additive-increase/multiplicative-decrease congestion control mechanism gives rise to traffic that dynamically adapts to changing networking conditions and does so on time scales of a few round-trip times.** In large part, it is the finite link capacity that drives the dynamics of protocols such as TCP and couples the different simultaneous connections sharing the link in intricate ways, introducing significant and complicated correlations across time, among active connections, and between the different layers in the protocol hierarchy.

Finally, the top layer in the Internet's suite of protocols is the

---

‖Following widespread (but perhaps overly general) practice, we use the term self-organized criticality (SOC) to refer to highly interactive self-organized systems that display power law behavior.

**In addition, TCP's flow control leads to a "self-clocking" structure that also introduces structure on the time scales of round-trip times, but one that is in this case separate from current network conditions.

application layer. It contains a range of protocols that directly serve the user; e.g., Telnet (remote log-in), file transfer protocol (FTP), simple mail transfer protocol (SMTP) for e-mail, HTTP (Web), and hundreds more. The applications also induce patterns of communication (e.g., keystrokes for Telnet, a control session coupled with multiple data transfer sessions for FTP, transfers interrupted by "think time" for HTTP) that echo downward into the dynamics of the underlying TCP and IP layers.

We find complex structure elsewhere in the set of Internet protocols, too. The routers internal to the network run distributed algorithms coordinated via routing protocols to discover paths from any given router to any given network node. For our purposes, the most interesting of these is the border gateway protocol (BGP) that maintains connectivity between the ASs. It is the glue that ties the ASs together, ensuring seamless communication across AS boundaries. Being a variant of the class of "distance-vector" routing protocols, each BGP-speaking router selects the "next hop" to use in forwarding packets to a given destination based on paths to those destinations advertised by the routers at the neighboring ASs. Routers exchange paths to destinations to facilitate route selection based on policy: ASs apply individual local policies when selecting their preferred routes, usually based on the series of ASs that a given route transits. This feature enables an administratively decentralized Internet—using these policies, ASs can direct traffic to ASs with which they have business relationships, where traditional network routing protocols would have selected the shortest path. However, this feature also introduces complex and subtle dynamics that can have global implications. Internet research is just beginning to unravel some of these interactions and their impact on the network's overall traffic characteristics and stability (16, 17).

The Rich Semantic Content of Measurements. From a scientific viewpoint, a crucial—and perhaps unique—facet of studying Internet measurements is their very high semantic content. Individual measurements, such as time-stamped IP packet headers or BGP routing table dumps, contain a wealth of information, both because the tools for measuring them often can capture them with perfect fidelity and because the measurements include the full structure of the layers relevant to network behavior. For example, a traffic trace collected from a link within the Internet is not merely a simple uni- or multivariate time series of packets but manifests itself at the different networking layers in a variety of different forms:

- At the application layer, we can describe the traffic in terms of session arrivals, session durations, and session sizes (volume in bytes). Examples of sessions are remote login, file transfer, e-mail delivery, or web surfing.
- At the transport layer, the overall traffic can be characterized in terms of TCP connection arrivals, durations, and sizes. Other components of the traffic that use transport protocols other than TCP have their own characterization. A single session might correspond to a single transport connection (e.g., remote login via Telnet, e-mail) or a group of either consecutive or concurrent connections (Web surfing, file transfer, Ssh remote login).
- At the internetwork layer, traffic descriptions focus on either individual IP packets, or on IP flows: their arrival patterns, sizes, origination, and destination addresses. A single transport connection might comprise a single IP flow, or a series of flows separated by significant lulls, depending on both the application driving the transport, and the network conditions encountered. In addition, we can view network traffic as an aggregate of IP packets generated by many host–host pairs.

- At the link layer, traffic can be dealt with by treating the individual packets as black boxes, i.e., by focusing on the mere existence of a measured packet (time stamp, packet size) and not on its "meaning" as revealed by its header.

Thus, as a result of the architecture of the Internet, actual network traffic—i.e., the flow of packets across a link inside the Internet—is the result of intertwined mechanisms, pronounced and often unexpected modes, and complex interactions that exist at and between the different networking layers.

Similar observations of rich semantics apply when trying to infer Internet connectivity from BGP measurements. BGP reachability information is obtained as a result of route updates exchanged between AS border routers. These messages are stored in the routers' routing tables and provide information about Internet connectivity at the AS level. However, because of the prevalence of policy, it is in general very difficult to know how complete or incomplete the AS-level connectivity information is that can be derived from the routing tables of a relatively small number of BGP routers. Additional difficulties arise because of the highly dynamic nature of BGP, that is, the high frequency with which changes in routing information occur.

Model Validation Framework. The ability to collect high-volume and high-quality Internet-related measurements, and the subsequent discoveries of intriguing statistical characteristics in the data, have led to much interest in Internet modeling and analysis. Whereas the majority of research efforts have focused on evocative models that succeed in synthetically generating, statistically describing, or formally reproducing the statistical characteristics of interest, explanatory modeling remains rare. Moreover, when competing Internet-related models (evocative or explanatory) have been proposed, the technical standards for validating them against the full power of the available measured data have generally been low or even nonexistent. That the same phenomenon can give rise to a number of different highly context-specific models motivates looking into scientific approaches for favoring one model over another.

We now formulate a framework for checking whether a proposed Internet model is indeed explanatory or only evocative. The proposed procedure for identifying an Internet-related modeling approach as explanatory consists of the following three steps:

(i) *Discovery.* Begin with a data-driven finding or "emergent" phenomenon that defies conventional modeling.

(ii) *Construction.* Devise a mathematical construction that reproduces the "emergent" phenomenon of interest, is given in terms of elementary networking concepts or mechanisms, and reflects the highly engineered structure of the Internet.

(iii) *Validation.* Revisit the available measurements, extract from them the data necessary to study the elementary networking concepts or mechanisms that have been identified in step *ii*, and check whether the proposed elementary concepts or mechanisms are indeed consistent with these data.

A critical feature of the proposed framework is that it "closes the loop" between the discovery of a network-related empirical phenomenon on the one hand and its proposed explanation in terms of a structural model on the other, where the structural model identifies a set of more elementary mechanisms as the main cause of the phenomenon.[††] This "closing of the loop" is achieved by requiring that the proposed model conforms to measured data not only at the level where the discovery was originally made but also at the level where the more elementary mechanisms are observable and verifiable.

[††]We are not claiming that "closing the loop" is a new concept; for example, much of physics has been all about "closing the loop." We simply argue here for applying the "closing the loop" concept in the context of Internet modeling, too.
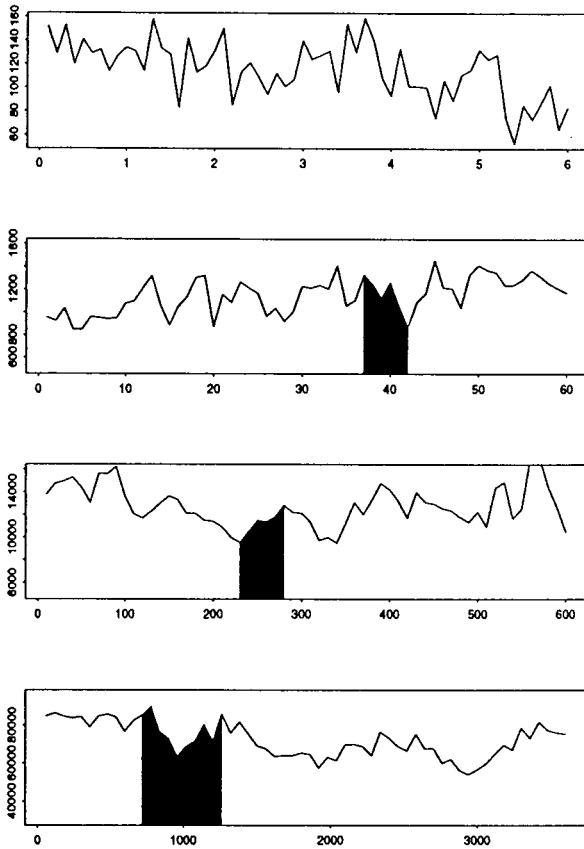
Fig. 1. Internet traffic viewed over four orders of magnitude.

final plot, by another factor of 6, such that now the plot spans the entire hour of the collected trace.

Similar plots produced from a synthesized trace generated from some traditionally assumed Internet traffic model tend to "smooth out" very quickly as the time scale increases, with hardly any variability left on the coarser time scales. In contrast, measured Internet traffic is highly bursty—as depicted in Fig. 1—and remains so even on quite coarse time scales. In fact, Fig. 1 suggests that measured Internet traffic is invariant under some judicious scaling of time and space—a trademark of self-similar or fractal-like objects.

More precisely, consider a second-order stationary and zero mean stochastic process $X = (X_k : k \geq 1)$ with autocorrelation function $(r(k), k \geq 0)$ and define the family of aggregated processes $(X^{(m)} : m \geq 1)$, where for $m = 1, 2, \ldots, X^{(m)} = (X^{(m)}(i) : i \geq 1)$ is given by $X^{(m)}(i) = (X_{(i-1)m+1} + \cdots + X_{im})/m$. Following ref. 14, X is called asymptotically second-order self-similar (with self-similarity or Hurst parameter $0 < H < 1$), if (i) $\lim_{m \to \infty} \mathrm{Var}(m^{1-H} X^{(m)}) = \sigma^2$, where $0 < \sigma^2 < \infty$ is a finite positive constant, and (ii) $\lim_{m \to \infty} r^{(m)}(k) = ((k + 1)^{2H} - 2k^{2H} + (k - 1)^{2H})/2$, where $r^{(m)} = (r^{(m)}(k), k \geq 0)$ denotes the auto-correlation function of the aggregated process $X^{(m)}$. It is in this sense that Internet traffic exhibits self-similar scaling properties, and the form of the autocorrelation function appearing in the above definition implies (and is implied by) the presence of long-range correlations in Internet traffic. To this end, X is said to exhibit LRD if for $1/2 < H < 1$,

$$r(k) \sim c_1 k^{2H-2}, \text{ as } k \to \infty,$$

where c_1 is a finite positive constant.[§§] Note that the power law decay of the autocorrelations of a long-range dependent process implies $\Sigma_k |r(k)| = \infty$. Even though the high-lag autocorrelations are individually small, their cumulative effect is of importance and gives rise to a behavior of the underlying stochastic process that is markedly different from that of the conventionally considered short-range dependent processes. Here, a second-order stationary stochastic process $X = (X_k : k = 1, 2, \ldots)$ is called short-range dependent (SRD) if for some $0 < \rho < 1$,

$$r(k) \sim c_2 \rho^k, \text{ as } k \to \infty,$$

where c_2 is a finite positive constant. Thus, in contrast to LRD, SRD is characterized by an autocorrelation function that decays geometrically fast and satisfies $\Sigma_k |r(k)| < \infty$; it is this difference between the autocorrelations of an LRD and SRD process that captures the surprising and distinctive difference between the actually observed and commonly assumed temporal behavior of Internet traffic.

The original finding of self-similar scaling behavior in measured network traffic was reported in ref. 1 and was based on an extensive statistical analysis of traffic measurements from Ethernet local-area networks (LAN) over a 4-year period from 1989 to 1993. A number of important followup studies provided further evidence of the prevalence of self-similar traffic patterns in measured traffic from wide-area networks (2, 3).

An important point to note, however, is that the above definition of self-similarity allows for various shades of "bursti-ness"—from highly bursty all the way to very smooth—depending on the relative magnitude of the overall mean, the variability σ^2, and the Hurst parameter H. But the basic fact of the presence of self-similarity has been an invariant (18) of Internet traffic for the past 10 or so years, despite the sometimes drastic changes the network has undergone during that period.

This ubiquity has intrigued traffic modelers and Internet researchers alike. However, whereas the former typically re-

When conformance is verified, the model in question can be expected to provide considerable new insights into the structure and dynamics of networks such as the Internet. These, in turn, can be exploited for various engineering purposes. On the other hand, if the verification fails—that is, the proposed elementary mechanisms turn out to be inconsistent with the measured data—the proposed Internet model should be deemed evocative, not explanatory (although it may still be explanatory when used in a context other than the Internet).

Self-Similar Internet Traffic

Measurement-Driven Discovery. Fig. 1 (see also ref. 18) captures the intuition behind the discovery that measured Internet traffic exhibits self-similar scaling properties. The plots were generated on the basis of an hour-long trace of Internet traffic collected from a network link connecting a large corporation to the Internet and consisting of IP packet headers timestamped to an accuracy on the order of a few milliseconds.[‡‡] The top plot in Fig. 1 shows a randomly selected subset of the trace on a time scale of 100 msec; that is, each observation represents the number of packets recorded on the link during a 100-msec interval for a total of 6 sec. The second plot shows a time scale (x axis) that is a factor of 10 larger and a y axis that has also been scaled up by a factor of 10; now each observation represents the number of packets per 1 sec, spanning 60 sec in total. The black-shaded region indicates from where the plot in the row above was chosen. Repeating this process, for the third plot, we have again increased the scale in both x and y by a factor of 10 and, in the

---

‡‡The measurements were gathered by J. Mogul in 1995 and are available from the Internet Traffic Archive, http://www.acm.org/sigcomm/ITA/.

§§The symbol $\sim$ means "behaves asymptotically as."

sponded with a series of increasingly refined evocative models of Internet traffic—where the networking context in which the data had been generated and collected in the first place has often been disregarded—the latter were interested mainly in explanatory models, models that make sense in the networking context and can be phrased and—more importantly –validated in terms of elementary traffic-related entities.

A Criticality-Based Explanation. Numerous papers have appeared during the last few years, mainly in the physics literature, arguing that the self-similar scaling behavior of measured Internet traffic can be explained simply as a phase-transition phenomenon from statistical mechanics. To illustrate this claim, consider the recently studied types of network traffic models (9, 10) that are not atypical of the proposed statistical mechanics approach to networking. The network topology is modeled by a square lattice with the usual neighborhood relationship (i.e., four nearest neighbors) and with periodic boundary conditions. A fraction ρ of the nodes are assumed to be hosts that can generate and receive packets, with the rest of the nodes serving as routers, where packets can only be stored or forwarded. Each node is assumed to maintain a queue with infinite buffer space. The local interactions among the nodes can be of two types. In case the node is a host, it injects (randomly) at some rate λ new packets into this network and does so independently from other hosts, and only another host can serve as the final destination of a packet. If the node is a router, it selects the packet at the front of its queue and decides according to a fixed rule which link to use to forward the packet to the next-hop router. For the time evolution of the resulting network traffic model or interacting particle system, a time step is defined to consist of one update (according to the above mechanisms) at all nodes.

Through simulations, these models have been shown to exhibit a phase transition as the packet injection rate parameter λ varies from 0 to 1. At the critical point, i.e., for $\lambda = \lambda_c$, efficiency measures such as the total number of delivered packets are maximized. More importantly—as far as this paper is concerned—at criticality, the time series describing the number of packets in a given node's queue show self-similar scaling behavior in the sense of Fig. 1 and exhibit $1/f$-type power spectra. Accordingly, the argument has been made that the proposed simple network traffic model identifies self-similarity of Internet traffic as a phase-transition phenomenon; that is, the network (hosts, routers) self-organizes to run at criticality, where it achieves maximum information transfer and efficiency (10). Moreover, because some of the key features of this network traffic model are shared by highway traffic models (19), it is claimed that there exist some deep connections between the dynamics displayed by traffic on highways and computer networks close to criticality.

Although it is interesting and educational to know that self-similarity can arise from such a simple process and can be elegantly described as a phase-transition phenomenon from statistical mechanics, the question we ask here is: Is self-similarity in the Internet indeed the signature of this type of criticality-based dynamic? That is, is the proposed statistical mechanics model explanatory or simply evocative? To answer this question, we expose the proposed model to the validation framework outlined in *Model Validation Framework*. Although step *i* applies trivially, step *ii* already reveals serious problems with the basic model. For one, instead of exploiting the highly engineered structure of the Internet, it ignores essentially all aspects of the Internet architecture described in *Highly Engineered Yet Complex* (e.g., no layering, no feedback, infinite buffers). Furthermore, being void of any networking-specific concept, the only mechanism to study is the packet injection rate λ, which roughly reflects link utilization. However, what really identifies this phase-transition-based explanation as irrelevant as

far as the self-similar scaling behavior of Internet traffic is concerned is step *iii*; that is, self-similar scaling has been observed in networks with low, medium, or high loads, and any notion of a ''magical'' load scenario where the network has to run at critical rate λ_c to show self-similar traffic characteristics is inconsistent with the measurements.

A Networking-Based Explanation. Next, consider the following mathematical construction that fits in well with the layering architecture of the Internet. At the application layer, sessions (i.e., FTP, HTTP, Telnet) arrive at random (i.e., according to some stochastic process) on the link and have a ''lifetime'' or session length during which they exchange information. This information exchange manifests itself at the IP layer, too, where from the start until the end of a session, IP packets are transmitted in some bursty fashion. Thus, at the IP layer, the aggregate link traffic measured over some time period (e.g., 1 h) is made up of the contributions of all of the sessions that, during the period of interest, actively transmitted packets.

Mathematically, this construction, referred to as Cox's construction, is known to give rise to LRD or, equivalently, asymptotic second-order self-similarity, provided the session arrivals follow a Poisson process and, more importantly, the distribution $F(x)$ of the session sizes T (i.e., number of packets or bytes per session) are heavy-tailed with infinite variance (14). That is, as $x \to \infty$,

$$1 - F(x) = P[T > x] \sim c_3 \, x^{-\alpha},$$

where $1 < \alpha < 2$. The main ingredient of Cox's construction (also known as an immigration-death process or $M/G/\infty$ queuing model) is the heavy-tailedness of the session sizes, where the index α is related to the self-similarity or Hurst parameter of the aggregate traffic and satisfies the relation $H = (3 - \alpha)/2$. Intuitively, the heavy-tailedness property implies that there is no ''typical'' session size, but instead the session sizes are highly variable (i.e., exhibit infinite variance) and fluctuate over a wide range of scales, from bytes to kilobytes to megabytes and beyond. It is this basic characteristic at the application layer that causes the aggregate traffic at the IP layer to exhibit self-similar scaling behavior. A closely related earlier construction, originally due to Mandelbrot (13), relies on the notion of a renewal-reward process but uses the same basic ingredient of heavy-tailedness to explain the self-similarity property of the aggregate link traffic (20).

To see how this networking-based explanation holds up against our proposed model validation framework, observe that Cox's construction (or, equivalently, Mandelbrot's) passes step *ii* with ease—it clearly identifies the data sets that need to be extracted from the available IP packet-header traces to check step *iii*, namely session arrivals and session sizes.

For FTP and Telnet, the session structures have been shown to be consistent with Cox's construction (2). For HTTP sessions (i.e., Web surfing), which are responsible for the bulk of today's Internet traffic, an individual session is typically made up of many individual HTTP connections, and obtaining session information is generally more involved. Indirect evidence can be obtained by analyzing the durations or sizes of individual HTTP connections. For example, Fig. 2 plots $P[T > x]$ vs. x on a log–log scale for the empirical distribution of measured HTTP request sizes for a 1996 24-hour measurement period, resulting in 226,386 observations. The linear relationship over more than three orders of magnitude is strong evidence that the data are consistent with the crucial heavy-tailedness assumption underlying Cox's construction, and the slope values between 1.2 and 1.4 give rise to self-similar aggregate traffic with H-values between 0.8 and 0.9. See refs. 3, 21, and 22 for further empirical studies in support of the ubiquitous nature of heavy-tailedness in measured Internet traffic.
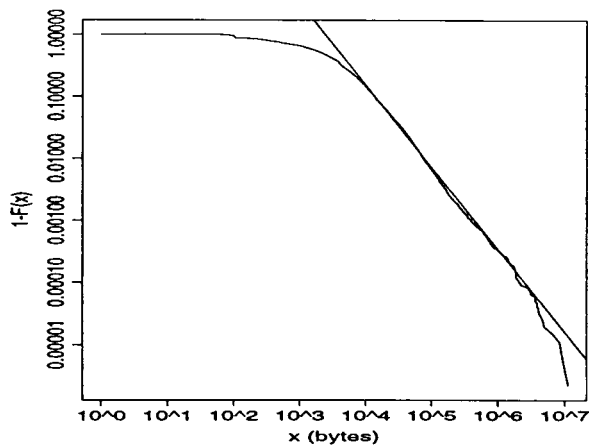
Fig. 2. Log–log plot of $1 - F(x)$ vs. x for HTTP connection sizes.

Fig. 3. Log–log plot of $1 - F_d$ vs. d for the vertex degree distributions F_d for three different BGP-derived AS map snapshots.

That this networking-based explanation of the self-similarity phenomenon successfully passes our model validation framework has far-reaching implications:

On the one hand, that we can explain self-similar scaling in terms of the statistical properties of the individual sessions that make up the aggregate link traffic suggests that the LRD nature of network traffic is mainly caused by user/application characteristics (i.e., Poisson arrivals of sessions, heavy-tailed session sizes as a result of transmitting heavy-tailed files, or Web documents). This in turn reveals that self-similarity has little to do with network-specific aspects such as the protocol-related mechanisms that determine the actual flow of packets as they traverse the Internet.[¶¶] Consequently, self-similarity is likely to remain with us [assuming the way humans tend to organize information does not change drastically (23)].

On the other hand, the fact that LRD leaves the smaller time-scale behavior essentially unspecified has in turn motivated researchers to focus investigations into the fine-grained structure of network traffic. Here, the goal is to relate the observed complex and highly time-localized traffic patterns to the most important features of the common protocols (8).

Scale-Free Internet Topology

Measurement-Driven Discovery. As mentioned in *Highly Engineered Yet Complex*, border routers exchange BGP route updates to propagate reachability information. This reachability information is stored in routing tables in each of the BGP routers. Starting in November 1997, the National Laboratory for Applied Network Research (NLANR) has collected BGP routing tables once per day from the route server route-views.oregon-ix.net, whose sole purpose is to connect to several operational routers and obtain their routing tables. After processing these routing tables, NLANR provides, among other things, daily "AS connectivity maps" that have been used to infer and reproduce snapshots of the Internet AS graph.[‖‖]

Relying on three snapshots of such BGP-derived AS maps (November 1997, April 1998, and December 1998), one of the surprising findings, originally reported in ref. 4, concerns the vertex degree distribution of the resulting AS graphs, namely the

observation that f_d, the number of nodes with outdegree d, follows a power-law. That is, $f_d \propto d^{-\alpha}$, where the symbol $\propto$ means "proportional to." We can equivalently express the relationship in terms of the complementary cumulative distribution function, $1 - F_d = 1 - \Sigma_{i=1}^{i=d} f_i$ ($d = 1, 2, \ldots$). In this case, we have $1 - F_d \propto d^{-(\alpha-1)}$.

For each of three snapshots (see Fig. 3), we find $\alpha \approx 2.1$. Intuitively, the significance of this discovery is that the vertex degrees observed in the Internet AS graph are highly variable. In fact, such highly variable vertex degrees have been unheard of in the context of the traditional and well studied Erdös–Rényi-type random graph models (24) or the more hierarchical graph structures that have been proposed as realistic topology models in the networking literature (5, 6). In both of these cases, the vertex degree distribution tends to be sharply concentrated around some "typical" node degree (e.g., the mean of the distribution), with essentially negligible probability for encountering vertex degrees that deviate by, say, more than one order of magnitude from the mean. Because of the absence of any such "typical" node degrees in graphs that exhibit power-law vertex degree distributions, these power-law graphs are also called scale-free graphs.

The Barabasi–Albert Model. Several recent papers in the physics and complex systems literature have attempted to uncover the mechanisms that cause graphs to be scale-free. Among these efforts, the papers by Barabasi, Albert, and colleagues (11, 12) have attracted the most attention in the networking community, as their authors propose a very appealing construction of network topologies [henceforth the Barabasi–Albert (BA) construction or model] that is claimed to explain the observed scale-free nature of the Internet's AS graph. Inspired by the concept of self-organization, models resulting from the BA construction explain and reproduce a number of the empirically observed power-law relationships reported in ref. 4. The models rely on three generic mechanisms to drive the evolution of such graph structures over time: incremental growth, preferential connectivity, and rewiring. Incremental growth follows from the observation that most networks develop over time by adding new nodes and new links to the existing graph structure. Preferential connectivity expresses the frequently encountered phenomenon that there is higher probability for a new or existing node to connect or reconnect to a node that already has a large number of links (i.e., high-vertex degree) than there is to (re)connect to a low-degree vertex. More formally, in the case of the prefer-

---

[¶¶] We hasten to note, however, that without TCP, or some other form of congestion control, the highly variable session workloads are capable of creating aggregate link traffic that would be very different from what we observe in today's Internet.

[‖‖] It is important to note, however, that because of BGP-specific features such as address aggregation and policy routing, the NLANR-generated AS connectivity maps may provide a very incomplete picture of the actual AS connectivity in the Internet. We ignore in the following this largely unsolved problem, but resolving it is an area of active research.
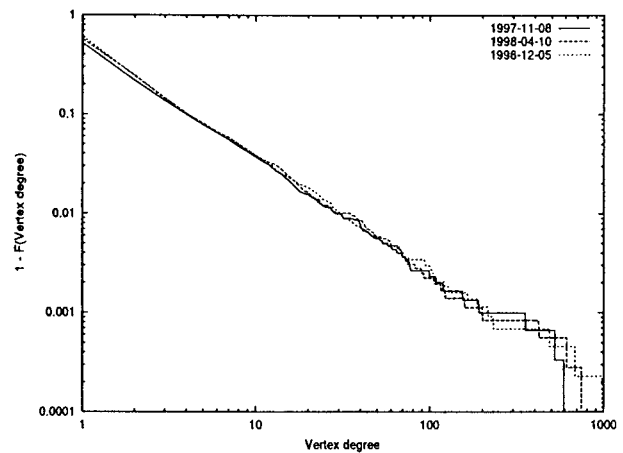
ential connectivity mechanism underlying the BA model, when a new AS joins the network, the probability of the new node to connect to each existing node (henceforth, "target node" or "peer") is given by $k_i/\Sigma k_j$, where k_i is the vertex degree of the target node, and Σk_j is the sum of the vertex degrees of all nodes in the graph before the addition of the new node. Finally, rewiring allows for some additional flexibility in the formation of networks by removing links connected to certain nodes and replacing them by new links in what effectively amounts to a local reshuffling.

Constructing a graph according to these elementary mechanisms, the authors of refs. 11 and 12 showed that the resulting graph attains a steady state, where, for example, the distribution of the node degree—after reaching steady-state—follows a power law with an exponent that is a function of the input parameters. Given the appeal and simplicity of the BA model, the question we ask here is again: Is the scale-free nature of AS graphs in the Internet indeed a signature of self-organized criticality? That is, is the BA model explanatory or simply evocative?

To pursue this question, note that the BA model is an ideal test case for the model validation framework proposed in *Model Validation Framework*. For one, the construction is explicit, relies on some elementary concepts (i.e., incremental growth, preferential connectivity, and rewiring), and reproduces the emergent phenomenon at hand. Thus, to test whether the BA model is explanatory or evocative, the remaining step is to "close the loop"—validating the elementary concepts against AS-level measurements.

Gathering these AS-level measurements takes considerable care, as we need to ensure they form a consistent BGP-based view of the Internet's AS connectivity. Using such a set of measurements spanning November 1998 through November 2000, we first extract information about basic events associated with a graph structure that grows over time (e.g., node birth, node death, link birth, link death). Accounting for the dead ASs and links, the BGP-derived Internet AS graph can be shown to be indeed consistent with the incremental growth condition assumed by the BA model.

However, when checking how new ASs connect to the existing AS graph, Fig. 4 illustrates a distinctly different mechanism at work than that predicted—the preferential connectivity assumption fails to hold, as follows. Starting with the AS map of November 1998, consider the next AS (node u) that joins the network. Node u joins the network with initial vertex degree m_u. Before we actually let node u join the network, we simulate the addition of node u with target AS(s) selected by sampling from the linear preferential model. We record the vertex degrees of the m_u target ASs so chosen and label them $\hat{k}_i^u$, $1 \le i \le m_u$. Next we actually add node u to the network, connect it to those target ASs it actually connected to in the real Internet, and record the vertex degrees of those target ASs, labeling them k_i^u, $1 \le i \le m_u$. We repeat the above process for the 1,000 new nodes added to the Internet between November 1998 and May 1999. The top plot of Fig. 4 shows the k_i^u's of the 1,000 nodes, and the bottom plot depicts the corresponding $\hat{k}_i^u$'s. Clearly, the preferential connectivity assumption underlying the BA construction is not consistent with the Internet's actual AS connectivity: in the real Internet (top plot), new ASs have a much stronger preference to connect to high vertex degree ASs and a significantly smaller preference to peer with low vertex degree ASs than predicted by the linear preferential model.

Finally, we check the validity of the rewiring concept. This third part of the BA construction is crucial for making the power-law exponent a function of the input parameters of the model. However, the data provide strong empirical evidence
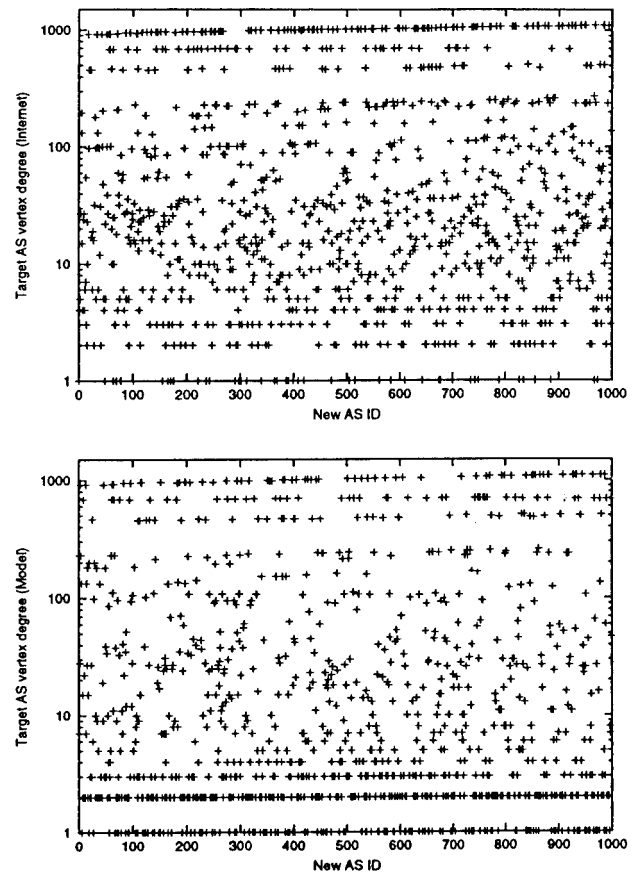

Fig. 4. New ASs target vertex degree(s) in the Internet (*Upper*) and according to the BA model (*Lower*).

(not shown here) that rewiring may not at all be a significant factor in the actual time-evolution of the Internet topology at the AS level.

In summary, exposing the proposed modeling approach to a validation framework that requires "closing the loop" reveals that the original BA model fails to be a genuinely explanatory model. As a result, findings about the behavior of Internet-related connectivity that rely on assumptions of the BA model that are not consistent with available measurements have to be very cautiously assessed (12). Moreover, the failure of the original BA model to "close the loop" motivates pursuing new approaches for modeling the time evolution of Internet-related topologies to demystify emergent phenomena such as scale-free AS graphs and to lead to a deeper understanding of how an administratively decentralized Internet evolves over time. On the one hand, a number of such new approaches can be expected to focus on more highly parameterized BA-type models that will likely result in an improved fit with the dynamical data. However, such approaches would still seek to explain the scale-free phenomenon in terms of the detailed dynamics of network growth. On the other hand, an alternative approach would be to ignore the growth dynamics altogether and instead explain the scale-free nature of the Internet topology at the AS level by linking the degree distribution to, for example, the underlying AS size distribution, which also appears to exhibit high variability, irrespective of how "size" is measured (25). This latter approach would be similar to how the Cox model explains the self-similar nature of Internet traffic by linking it to a ubiquitous, well-documented, but

largely unexplained high-variability phenomenon (i.e., heavy-tailed distribution of connection sizes).

Conclusion

We argue in this paper that, because of its highly engineered nature and the highly structured networking-specific semantic context of the available measurements, the Internet is an example of a large-scale complex system that offers unique opportunities for successfully distinguishing between two classes of models: evocative and explanatory. Only the latter can provide a sound scientific basis for the origins of such emergent phenomena as the self-similar dynamic of Internet traffic or the scale-free nature of the Internet AS graph.

To this end, we provide a concrete framework for checking whether a proposed Internet-related model is explanatory or simply evocative and illustrate its applicability to Internet modeling with a number of examples. In particular, we examine a number of recently proposed dynamical models of Internet traffic and Internet topology at the AS level that explain the entirely unexpected scaling behavior in terms of critical phenomena. In the process, we offer conclusive evidence that even though the models produce the self-similar scaling phenomena of interest, they do not explain why or how the observed phenomena arise in the Internet. Some of these criticality-based explanations can still be put to good use, however. For one, by teaching us how certain emergent phenomena might arise, they can serve as simple "null hypothesis" models to compare with. Moreover, they often suggest what further measurements to take for testing whether they are indeed evocative, and in this sense, whether they can lead to an improved understanding of the Internet.

To contrast, we also show that a class of mathematical models, originally due Cox (14) and Mandelbrot (13), readily passes our model validation framework. These models explain why and how Internet traffic exhibits self-similar scaling behavior and provide novel insights into (and new questions about) the dynamics of actual Internet traffic. In short, aiming for explanatory Internet models that successfully "close the loop" in the sense of *Model Validation Framework* calls for novel approaches that explicitly account for systems with such nongeneric, specialized, and highly structured architectures as the Internet and which are optimized (or suboptimized) through explicit design (26). When successful, these approaches can be expected to significantly advance our understanding of large-scale complex systems such as the Internet, where engineering design plays a central role and cannot simply be abstracted away.

1. Leland, W. E., Taqqu, M. S., Willinger, W. & Wilson, D. V. (1994) *IEEE/ACM Trans. Network.* **2,** 1–15.
2. Paxson, V. & Floyd, S. (1995) *IEEE/ACM Trans. Network.* **3,** 226–244.
3. Crovella, M. & Bestavros, A. (1997) *IEEE/ACM Trans. Network.* **5,** 835–846.
4. Faloutsos, M., Faloutsos, P. & Faloutsos, C. (1999) *Comput. Commun. Rev.* **29,** 251–262.
5. Waxman, B. M. (1988) *IEEE J. Select. Areas Comm.* **6,** 1617–1622.
6. Calvert, K., Doar, M. B. & Zegura, E. W. (1997) *IEEE Comm. Mag.* **35,** 160–163.
7. Huberman, B. A. & Lukose, R. M. (1997) *Science* **277,** 535–537.
8. Feldmann, A., Gilbert, A. C., Huang, P. & Willinger, W. (1999) *Comput. Commun. Rev.* **29,** 301–313.
9. Ohira, T. & Sawatari, R. (1998) *Phys. Rev. E* **58,** 193–195.
10. Sole, R. V. & Valverde, S. (2001) *Physica A* **289,** 595–605.
11. Albert, R. & Barabasi, A.-L. (2000) *Phys. Rev. Lett.* **85,** 5234–5237.
12. Albert, R., Jeong, H. & Barabasi, A.-L. (2000) *Nature (London)* **406,** 378–382.
13. Mandelbrot, B. B. (1969) *Int. Econ. Rev.* **10,** 82–113.
14. Cox, D. R. (1984) in *Statistics: An Appraisal*, eds. David, H. A. & David, H. T. (Iowa State Univ. Press, Ames, IA), pp. 55–74.
15. Clark, D. D. (1988) *Comput. Commun. Rev.* **18,** 106–114.
16. Griffin, T. G. & Wilfong, G. (1999) *Comput. Commun. Rev.* **29,** 277–288.
17. Labovitz, C., Ahuja, A., Bose, A. & Jahanian, F. (2000) *Comput. Commun. Rev.* **30,** 175–187.
18. Willinger, W. & Paxson, V. (1998) *Not. AMS* **45,** 961–970.
19. Nagel, K. & Schreckenberg, M. (1992) *J. Physique I* **2,** 2221–2229.
20. Willinger, W, Paxson, V., Taqqu, M. S. & Riedi, R. (2001) in *Long-Range Dependence: Theory and Applications*, eds. Doukhan, P., Oppenheim, G. & Taqqu, M. S. (Birkhäuser, Boston), in press.
21. Willinger, W., Taqqu, M. S., Sherman, R. & Wilson, D. V. (1997) *IEEE/ACM Trans. Network.* **5,** 71–86.
22. Feldmann, A., Gilbert, A. C., Willinger, W. & Kurtz, T. G. (1998) *Comput. Commun. Rev.* **28,** 5–29.
23. Zhu, X., Yu, J. & Doyle, J. (2001) *Proc. IEEE INFOCOM 2001* (IEEE Computer Society Press, Los Alamitos, CA).
24. Erdös, P. & Rényi, A. (1960) *Publ. Math. Hung. Acad. Sci.* **5,** 17–61.
25. Tangmunarunkit, H., Doyle, J., Govindan, R., Jamin, S., Shenker, S. & Willinger, W. (2001) *Comput. Commun. Rev.* **31,** 7–10.
26. Carlson, J. M. & Doyle, J. (1999) *Phys. Rev. E* **60,** 1412–1427.