

Mathematics and the Internet: A Source of Enormous Confusion and Great Potential

Walter Willinger, David Alderson, and John C. Doyle

For many mathematicians and physicists, the Internet has become a popular real-world domain for the application and/or development of new theories related to the organization and behavior of large-scale, complex, and dynamic systems. In some cases, the Internet has served both as inspiration and justification for the popularization of new models and mathematics within the scientific enterprise. For example, scale-free network models of the preferential attachment type [8] have been claimed to describe the Internet's connectivity structure, resulting in surprisingly general and strong claims about the network's resilience to random failures of its components and its vulnerability to targeted attacks against its infrastructure [2]. These models have, as their trademark, power-law type node degree distributions that drastically distinguish them from the classical Erdős-Rényi type random graph models [13]. These "scale-free" network models have attracted significant attention within the scientific community and have been largely responsible for launching and fueling the new field of *network science* [42, 4].

To date, the main role that mathematics has played in network science has been to put the physicists' largely empirical findings on solid grounds

Walter Willinger is at AT&T Labs-Research in Florham Park, NJ. His email address is walter@research.att.com.

David Alderson is assistant professor at the Naval Postgraduate School in Monterey, CA. His email address is dalders@nps.edu.

John C. Doyle is John G. Braun Professor of Control & Dynamical Systems, Electrical Engineering, and BioEngineering at Caltech. His email address is doyle@cds.caltech.edu.

by providing rigorous proofs of some of their more highly publicized claims [14, 15, 16, 23, 11, 25]. The alleged scale-free nature of the Internet's topology has also led to mathematically rigorous results about the spread of viruses over scale-free graphs of the preferential attachment type, again with strong and unsettling implications such as a zero epidemic threshold [11, 25]. The relevance of the latter is that in stark contrast to more homogeneous graphs, on scale-free networks of the preferential attachment type, even viruses with small propagation rates have a chance to cause an epidemic, which is about as bad as it can get from an Internet security perspective. More recently, the realization that large-scale, real-world networks such as the Internet evolve over time has motivated the mathematically challenging problem of developing a theory of graph sequences and graph limits [17, 19, 20]. The underlying idea is that properly defined graph limits can be expected to represent viable models for some of the enormous dynamic graph structures that arise in real-world applications and seem too unwieldy to be described via more direct or explicit approaches.

The generality of these new network models and their impressive predictive ability notwithstanding, surprisingly little attention has been paid in the mathematics and physics communities to parallel developments in the Internet research arena, where the various non-rigorous and rigorous results derived from applying the scale-free modeling paradigm to the Internet have been scrutinized using available measurements or readily available domain knowledge. A driving force behind these Internet-centric validation efforts has been the realization that—because of its engineered architecture, a thorough understanding

of its component technologies, and the availability of extensive (but not necessarily very accurate) measurement capabilities—the Internet provides a unique setting in which most claims about its properties, structure, and functionality can be unambiguously resolved, though perhaps not without substantial efforts. In turn, models or theories that may appeal to a more mathematically inclined researcher because of their simplicity or generality, but result in incorrect, misleading, or wrong claims about the Internet, can and will be identified and labeled accordingly, but it may take considerable time (and efforts) to expose their specious nature.

In this article, we take a closer look at what measurement-based Internet research in general, and Internet-specific validation efforts in particular, have to say about the popular scale-free modeling paradigm and the flurry of mathematical developments it has inspired. In particular, we illustrate why and how in the case of the Internet, scale-free network models of the preferential attachment type have become a classic lesson in how errors of various forms occur and can add up to produce results and claims that create excitement among non-networking researchers, but quickly collapse under scrutiny with real data or when examined by domain experts. These opposite reactions have naturally been a source of great confusion, but the main conclusion is neither controversial nor should it come as a big surprise: the scale-free modeling paradigm is largely inconsistent with the engineered nature of the Internet and the design constraints imposed by existing technology, prevailing economic conditions, and practical considerations concerning network operations, control, and management.

To this end, we document the main sources of errors regarding the application of the scale-free modeling approach to the Internet and then present an alternative approach that represents a drastic departure from traditional network modeling. In effect, we motivate here the development of a novel modeling approach for Internet-like systems that (1) respects the highly designed nature of the network; (2) reflects the engineering intuition that exists about great many of its parts; (3) is fully consistent with a wide range of measurements; and (4) outlines a mathematical agenda that is more challenging, more relevant, and ultimately more rewarding than the type of mathematics motivated by an alluring but largely misguided approach to Internet modeling based on scale-free graphs of the preferential attachment type. In this sense, this article demonstrates the great potential that the Internet has for the development of new, creative, and relevant mathematical theories, but it is also a reminder of a telling comment attributed to S. Ulam [12] (slightly paraphrased, though), who said “Ask not what mathematics can do for [the Internet]; ask what [the Internet] can do for mathematics.”

The Scale-free Internet Myth

The story recounted below of the scale-free nature of the Internet seems convincing, sound, and almost too good to be true. Unfortunately, it turned out to be a complete myth, but has remained a constant source of enormous confusion within the scientific community.

Somewhat ironically, the story starts with a highly-cited paper in the Internet research arena by Faloutsos et al. [27]. Relying on available measurements and taking them at face value, the paper was the first to claim that the (inferred) node degree distributions of the Internet’s router-level topology as well as AS-level topology are power-law distributions with estimated α -parameters between 1 and 2. To clarify, by *router-level topology*, we mean the Internet’s physical connectivity structure, where nodes are physical devices such as routers or switches, and links are the connections between them. These devices are further organized into networks known as *Autonomous Systems (ASes)*, where each AS is under the administrative control of a single organization such as an Internet Service Provider (ISP), a company, or an educational institution. The relationships among ASes, when organized as a graph, produce what is known as the Internet’s *AS-level topology*. Note that a link between two nodes in the AS-level topology represents a type of business relationship (either peering or customer-provider). Also, in contrast to the router-level topology that is inherently physical, the AS topology is a logical construct that reflects the Internet’s administrative boundaries and existing economic relationships.

These reported power-law findings for the Internet were quickly picked up by Barabási et al., who added the Internet to their growing list of real-world network structures with an apparently striking common characteristic; that is, their vertex connectivities (described mathematically in terms of node *degrees*) “follow a scale-free power-law distribution” [8, 3]. This property is in stark contrast to the Poissonian nature of the node degrees resulting from the traditional Erdős-Rényi random graphs [13] that have been the primary focus of mathematical graph theory for the last 50 years. Naturally, it has fueled the development of new graph models that seek to capture and reproduce this ubiquitously reported power-law relationship, thereby arguing in favor of these models as more relevant for representing real-world network structures than the classical random graph models. In fact, much of the initial excitement in the nascent field of network science can be attributed to an early and appealingly simple class of network models that was proposed by Barabási and Albert [8] and turned out to have surprisingly strong predictive capabilities.

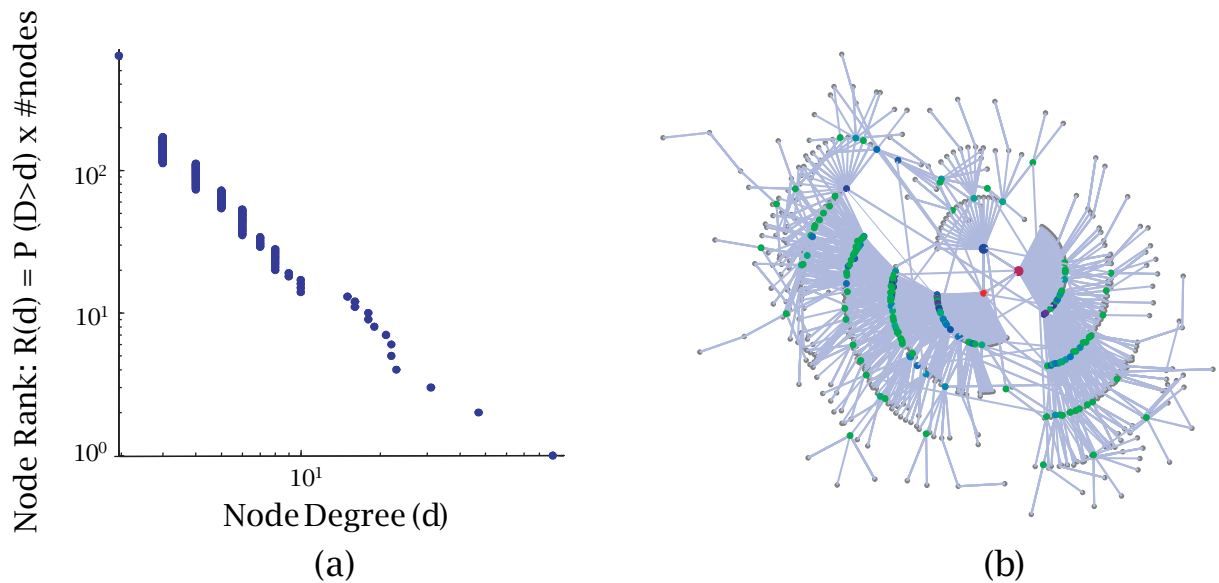


Figure 1.

In short, Barabási and Albert [8] described a network growth model in which newly added vertices connect preferentially to nodes in the existing graph that are already well connected. This *preferential attachment* mechanism had been studied over the previous 75 years by Yule [54], Luria and Delbrück [38], and Simon [49], but it was its rediscovery and application to networks by Barabási and Albert that recently popularized it. Although many variants of the the basic Barabási-Albert construction have been proposed and studied, we will focus in the following on the original version described in [8], mainly because of its simplicity and because it already captures the most important properties of this new class of networks, commonly referred to as *scale-free networks*. The term scale-free derives from the simple observation that power-law node degree distributions are free of scale—most nodes have small degree, a few nodes have very high degree, with the result that the average node degree is essentially non-informative. A detailed discussion of the deeper meanings often associated with scale-free networks is available in [34]. To avoid confusion and to emphasize the fact that preferential attachment is just one of many other mechanisms that is capable of generating scale-free graphs (i.e., graphs with power-law node degree distributions), we will refer here to the network models proposed in [8] as *scale-free networks of the preferential attachment (PA) type* and show an illustrative toy example with associated node degree distribution in Figure 1.

The excitement generated by this new class of models is mainly due to the fact that, despite being generic and largely oblivious to system-specific details, they share some key properties

that give them remarkable predictive power. These properties were originally reported in [2], put on mathematically solid footing by Bollobás and Rioridan in [14, 15, 16], and explain the key aspects of the structure and behavior of these networks. For one, a hallmark of their structure is the presence of “hubs”; that is, centrally located nodes with high connectivity. Moreover, the presence of these hubs makes these networks highly vulnerable to attacks that target the hub nodes. At the same time, these networks are extremely resilient to attacks that knock out nodes at random, since a randomly chosen node is likely to be one of the low-degree nodes that constitute the bulk of the nodes in the power-law node degree distribution, and the removal of such a node has typically minimal impact on the network’s overall connectivity or performance.

This property—simultaneous resilience to random attacks but high vulnerability to targeted worst-case attacks (i.e., attacks against the hub nodes)—featured prominently in the original application of scale-free networks of the PA type to the Internet [2]. The underlying argument follows a very traditional and widely-used modeling approach. First, as reported in [27], the Internet has node degrees that follow a power-law distribution or are scale-free. Second, scale-free networks of the PA type are claimed to be valid models of the Internet because they are capable of reproducing the observed scale-free node degree distributions. Lastly, when abstracted to a scale-free models of the PA type, the Internet automatically inherits all the emergent features of the latter, most notably the presence of hub nodes that are critical to overall network connectivity and performance and

are largely responsible for the network's failure tolerance and attack vulnerability. In this context, the latter property has become known as the "Achilles' heel of the Internet" and has been highly publicized as a success story of network science—the discovery of a fundamental weakness of the Internet that went apparently unnoticed by the engineers and researchers who have designed, deployed, and studied this large-scale, critical complex system.

The general appeal of such surprisingly strong statements is understandable, especially given the simplicity of scale-free networks of the PA type and the fact that, as predictive models, they do not depend on the particulars of the system at hand, i.e., underlying technology, economics, or engineering. As such, they have become the embodiment of a highly popular statistical physics-based approach to complex networks that aims primarily at discovering properties that are universal across a range of very diverse networks. The potential danger of this approach is that the considered abstractions represent simplistic toy models that are too generic to reflect features that are most important to the experts dealing with these individual systems (e.g., critical functionality).

Deconstructing the Scale-free Myth

Given that the scale-free story of the Internet is grounded in real measurement data and based on a widely-accepted modeling approach, why is it so far from the truth? To explain and trace the various sources of errors, we ask the basic question; i.e., *"Do the available measurements, their analysis, and their modeling efforts support the claims that are made in [2]?"* To arrive at a clear and simple answer to this question, we address below the issues of data hygiene and data usage, data analysis, and mathematical modeling (including model selection and validation).

Know your data

A very general but largely ignored fact about Internet-related measurements is that what we can measure in an Internet-like environment is typically not the same as what we really want to measure (or what we think we actually measure). This is mainly because as a decentralized and distributed system, the Internet lacks a central authority and does not support third-party measurements. As a result, measurement efforts across multiple ASes become nontrivial and often rely on engineering hacks that typically do not yield the originally desired data but some substitute data. Moreover, using the latter at face value (i.e., as if they were the data we originally wanted) and deriving from them results that we can trust generally involves a leap of faith, especially in the absence of convincing arguments or evidence that would support an "as-is" use of the data.

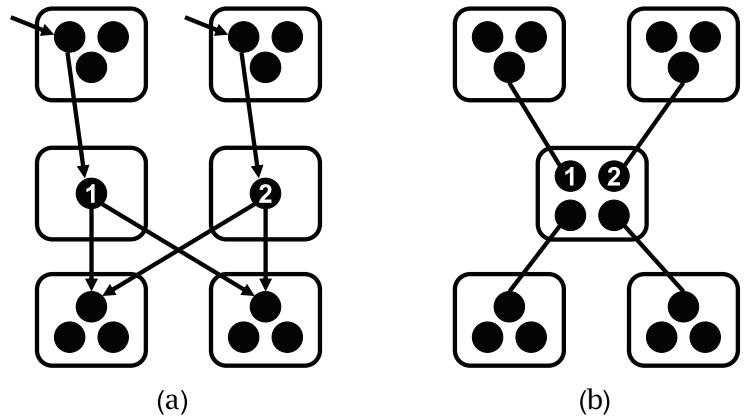


Figure 2. The IP alias resolution problem. Paraphrasing Fig. 4 of [50], traceroute does not list routers (boxes) along paths but IP addresses of input interfaces (circles) and alias resolution refers to the correct mapping of interfaces to routers to reveal the actual topology. In the case where interfaces 1 and 2 are aliases, (b) depicts the actual topology while (a) yields an "inflated" topology with more routers and links than the real one.

Internet-specific connectivity measurements provide a telling example. To illustrate, consider the data set that was used in [27] to derive the reported power-law claim for the (inferred) node degrees of the Internet's router-level topology.¹ That dataset was originally collected by Pansiot and Grad [44] for the explicitly stated purpose *"to get some experimental data on the shape of multicast trees one can actually obtain in [the real] Internet ..."* [44]. The tool of choice was traceroute, and the idea was to run traceroute between a number of different host computers dispersed across the Internet and glue together the resulting Internet routes to glean the shape of actual multicast trees. In this case, the engineering hack consisted of relying on traceroute, a tool that was never intended to be used for the stated purpose, and a substantial leap of faith was required to use Pansiot and Grad's data set beyond its original purpose and rely on it to infer the Internet's router-level topology [27].

For one, contrary to popular belief, running traceroute between two host computers does *not* generate the list of compliant (i.e., Internet Protocol (IP)-speaking) routers encountered en route from the source to the destination. Instead, since IP routers have multiple interfaces, each with its own IP address, what traceroute really generates is

¹While the arguments and reasons differ for the data sets used in [27] to derive the power-law claim for the Internet's AS-level topology, the bottom line is the same—the available measurements are not of sufficient quality for the purpose for which they are used in [27] (see for example [31]).

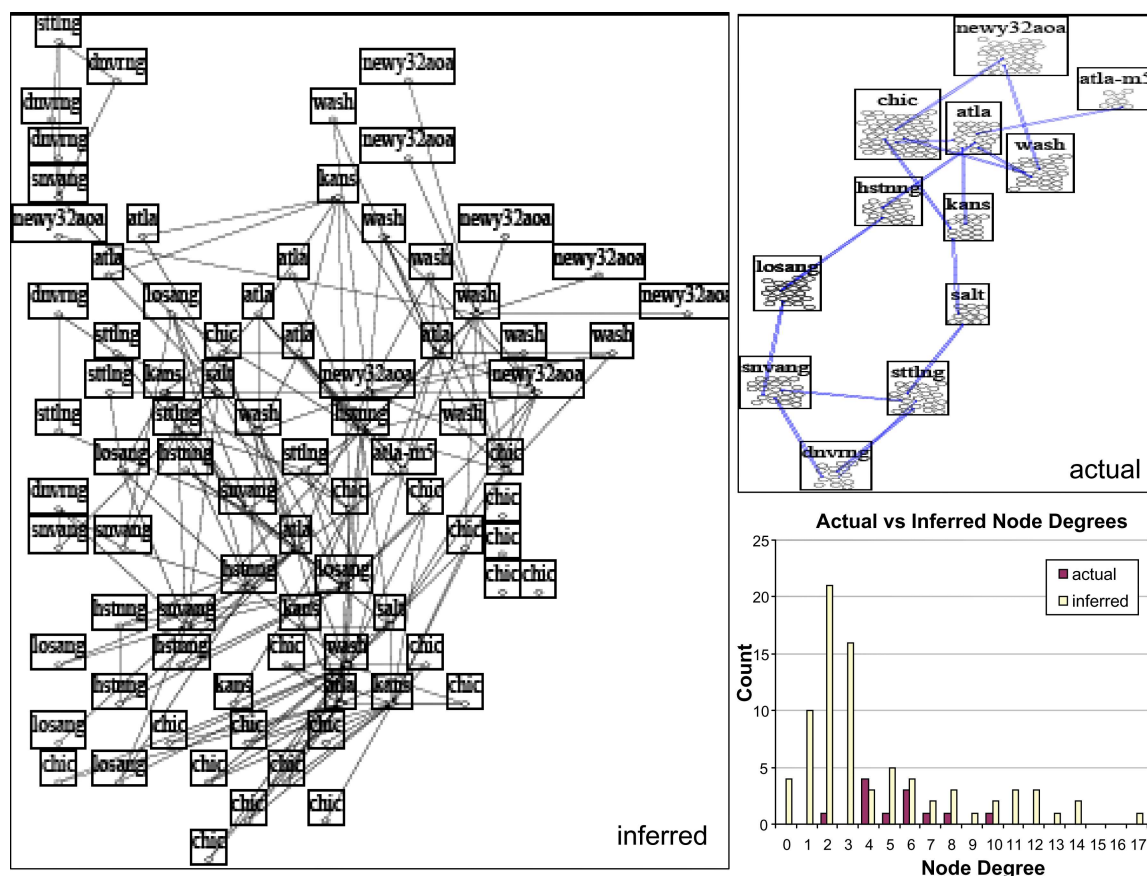


Figure 3. The IP alias resolution problem in practice. This is re-produced from [48] and shows a comparison between the Abilene/Internet2 topology inferred by Rocketfuel (a) and the actual topology (b). Rectangles represent routers with interior ovals denoting interfaces. The histograms of the corresponding node degrees are shown in plot (c).

the list of (input interface) IP addresses, and a very common property of `traceroute`-derived routes is that one and the same router can appear on different routes with different IP addresses. Unfortunately, faithfully mapping interface IP addresses to routers is a difficult open problem known as the *IP alias resolution problem* [51, 28], and despite continued research efforts (e.g., [48, 9]), it has remained a source of significant errors. While the generic problem is illustrated in Figure 2(a), its impact on inferring the (known) router-level topology of an actual network (i.e., Abilene/Internet2) is highlighted in Figure 2(b)—the inability to solve the alias resolution problem renders in this case the inferred topology irrelevant and produces statistics (e.g., node degree distribution) that have little in common with their actual counterparts.

Another commonly ignored problem is that `traceroute`, being strictly limited to IP or layer-3, is incapable of tracing through opaque layer-2 clouds that feature circuit technologies such as *Asynchronous Transfer Mode (ATM)* or *Multiprotocol Label Switching (MPLS)*. These technologies

have the explicit and intended purpose of hiding the network’s physical infrastructure from IP, so from the perspective of `traceroute`, a network that runs these technologies will appear to provide direct connectivity between routers that are separated by local, regional, national, or even global physical network infrastructures. The result is that when `traceroute` encounters one of these opaque layer-2 clouds, it falsely “discovers” a high-degree node that is really a logical entity—a network potentially spanning many hosts or great distances—rather than a physical node of the Internet’s router-level topology. Thus, reports of high-degree hubs in the core of the router-level Internet, which defy common engineering sense, can often be easily identified as simple artifacts of an imperfect measurement tool. While Figure 4(a) illustrates the generic nature of this problem, Figure 4(b) illuminates its impact in the case of an actual network (i.e., AS3356 in 2002), where the inferred topology with its highly connected nodes says nothing about the actual physical infrastructure of this network but is a direct consequence of

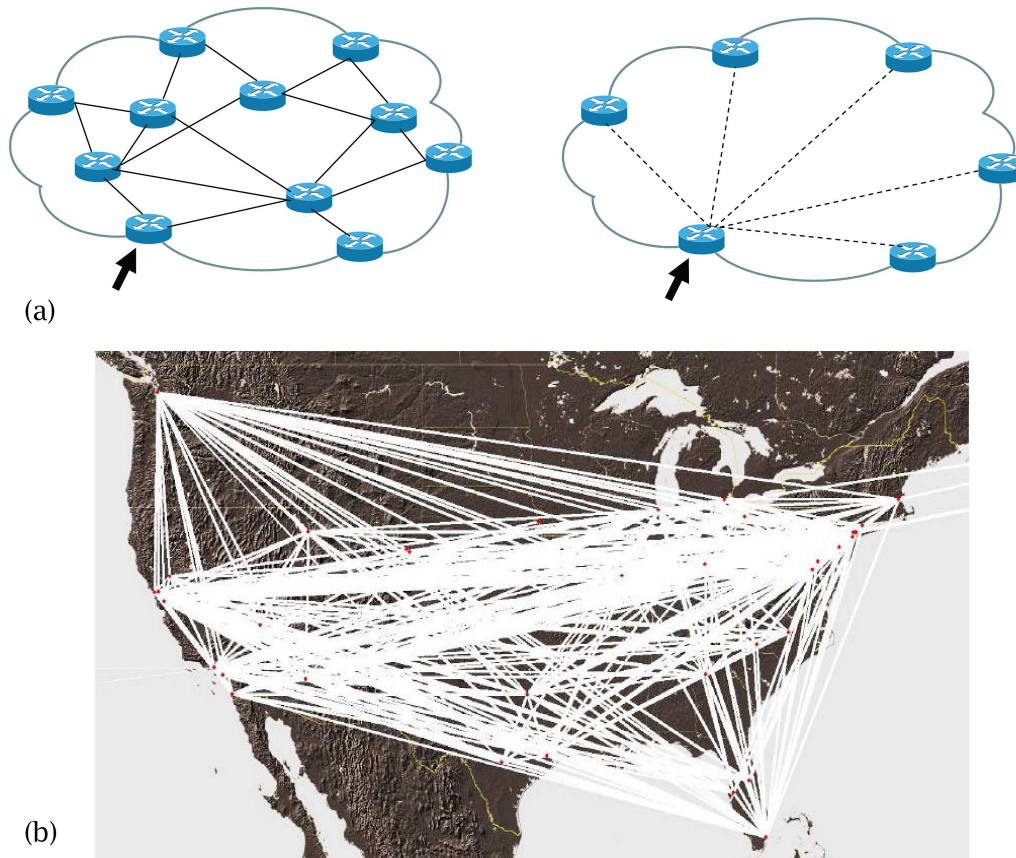


Figure 4. How traceroute detects fictitious high-degree nodes in the network core. (a) The actual connectivity of an opaque layer-2 cloud, i.e., a router-level network running a technology such as ATM or MPLS (left) and the connectivity inferred by traceroute probes entering the network at the marked router (right). (b) The Rocketfuel-inferred backbone topology of AS3356 (Level3), a Tier-1 Internet service provider and leader in the deployment of MPLS (reproduced from [50]).

traceroute's inability to infer the topology of an MPLS-enabled network.

We also note that from a network engineering perspective, there are technological and economic reasons for why high-degree nodes in the core of the router-level Internet are nonsensical. Since a router is fundamentally limited in terms of the number of packets it can process in any time interval, there is an inherent tradeoff in router configuration: it can support either a few high-throughput connections or many low-throughput connections. Thus, for any given router technology, a high-connectivity router in the core will either have poor performance due to its slow connections or be prohibitively expensive relative to other options. Conversely, if one deploys high-degree devices at the router-level, they are necessarily located at the edge of the network where the technology exists to multiplex a large number of relatively low-bandwidth links. Unfortunately, neither the original traceroute-based study of Pansiot and Grad nor any of the larger-scale versions that were subsequently performed by various

network research groups have the ability to detect those actual high-degree nodes. The simple reason is that these traditional traceroute studies lack access to a sufficient number of participating host computers in any local end-system to reveal their high connectivity. Thus, the irony of traceroute is that the high-degree nodes it detects in the network core are necessarily fictitious and represent entire opaque layer-2 clouds, and if there are actual high-degree nodes in the network, existing technology relegates them to the edge of the network where no generic traceroute-based measurement experiment will ever detect them.

Lastly, the nature of large-scale traceroute experiments also makes them susceptible to a type of measurement bias in which some points of the network are oversampled, while others are undersampled. Ironically, although this failure of traceroute experiments has received the most attention in the theoretical computer science and applied mathematics communities [32, 1] (most likely, because it is the most amenable to mathe-

matical treatment), it is the least significant from a topology modeling perspective.

In view of these key limitations of `traceroute`, it should be obvious that starting with the Pansiot and Grad data set, `traceroute`-based measurements cannot be taken at face value and are of no or little use for inferring the Internet's router-level topology. In addition, the arguments provided above show why domain knowledge in the form of such `traceroute`-specific "details" like IP aliases or layer-2 technology matters when dealing with issues related to data hygiene and why ignoring those details prevents us from deriving results from such data that we can trust. Ironically, Pansiot and Grad [44] detailed many of the above-mentioned limitations and shortcomings of their measurements. Unfortunately, [27] failed to revive these issues or recognize their relevance. Even worse, the majority of subsequent papers in this area typically cite only [27] and no longer [44].

Know your statistic

The inherent inability of `traceroute` to reveal unambiguously the actual node degree of any router (i.e., the number of different interfaces) due to the IP alias resolution problem, combined with the fundamental difficulties of the tool to correctly infer even the mere absence or presence of high-degree nodes (let alone their actual degrees) makes it impossible to describe accurately statistical entities such as node degree distributions. Thus, it should come as no surprise that taking `traceroute`-derived data sets "as is" and then making them the basis for any fitting of a particular parameterized distribution (e.g., power-law distribution with index α as in [27]) is statistical "overkill", irrespective of how sophisticated a fitting or corresponding parameter estimation technique has been used. Given the data's limitations, even rough rules-of-thumb such as a Pareto-type 80/20 rule (i.e., 80% of the effects come from 20% of the causes) cannot be justified with any reasonable degree of statistical confidence.

It is in this sense that the claims made in [27] and subsequent papers that have relied on this data set are the results of a data analysis that is not commensurate with the quality of the available data. It is also a reminder that there are important differences between analyzing high-quality and low-quality data sets, and that approaching the latter the same way as the former is not only bad statistics but also bad science, and doing so bolsters the popular notion that "there are lies, damned lies, and statistics." Unfortunately, the work required to arrive at this conclusion is hardly glamorous or newsworthy, especially when compared to the overall excitement generated by an apparent straight-line behavior in the easily obtainable log-log plots of degree vs. frequency.

Even if the available measurements were amenable to such an analysis, these commonly-used and widely-accepted log-log plots are not only highly non-informative, but have a tendency to obscure power-law relationships when they are genuine and fabricate them when they are absent (see for example [34]). In the case of the data set at hand, the latter observation is compounded by the unreliable nature of the `traceroute`-derived node degree values and shows why the power-law claims for the vertex connectivities of the Internet's router-level topology reported in [27] cannot be supported by the available measurements.

When modeling is more than data-fitting

We have shown that the data set used in [27] turns out to be thoroughly inadequate for deriving and modeling power-law properties for the distribution of node degrees encountered in the Internet's router-level topology. As a result, the sole argument put forward in [2] for the validity of the scale-free model of the PA type for the Internet is no longer applicable, and this in turn reveals the specious nature of both the proposed model and the sensational features the Internet supposedly inherits from the model.

Even if the node degree distribution were a solid and reliable statistic, who is to say that matching it (or any other commonly considered statistics of the data) argues for the validity of a proposed model? In the case of scale-free models of the PA type, most "validation" follows from the ability of a model to replicate an observed degree distribution or sequence. However, it is well-known in the mathematics literature that there can be many graph realizations for any particular degree sequence [47, 29, 35, 10] and there are often significant structural differences between graphs having the same degree sequence [6]. Thus, two models that match the data equally well with respect to some statistics can still be radically different in terms of other properties, their structures, or their functionality. A clear sign of the rather precarious current state of network-related modeling is that the same underlying data set can give rise to very different, but apparently equally "good" models, which in turn can give rise to completely opposite scientific claims and theories concerning one and the same observed phenomenon. Clearly, modeling and especially model validation has to mean more than being able to match the data if we want to be confident that the results that we derive from our models are valid.

At this point, it is appropriate to recall a quote attributed to G.E.P. Box, who observed that "*All models are wrong, but some models are useful.*" Without being more specific about which models are deemed useful and why, this comment is of little practical value. A more constructive piece of

advice that is more directly aligned with what we envision modeling should mean in the presence of imprecise data is from B.B. Mandelbrot [39], who observed *“If exactitude is elusive, it is better to be approximately right than certifiably wrong.”*

For complex network systems whose measured features suffer from the types of fundamental ambiguities, omissions, and/or systematic errors outlined above, we argue that network modeling must move beyond efforts that merely match particular statistics of the data. Such efforts are little more than exercises in data-fitting and are particularly ill-advised whenever the features of interest cannot be inferred with any reasonable statistical confidence from the currently available measurements. For systems such as the router-level Internet, we believe this to be a more scientifically grounded and constructive modeling approach. For one, given the known deficiencies in the available data sets, matching a particular statistic of the data may be precisely the wrong approach, unless that statistic has been found to be largely robust with respect to these deficiencies. Moreover, it eliminates the arbitrariness associated with determining which statistics of the data to focus on. Indeed, it treats all statistics equally. A model that is “approximately right” can be expected to implicitly match most statistics of the data (at least approximately).

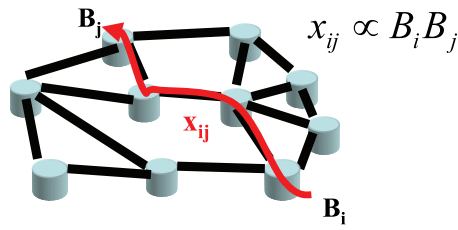
If we wish to increase our confidence in a proposed model, we ought also to ask what new types of measurements are either already available (but have not been used in the present context) or could be collected and used for validation. Here, by “new” we do not mean “same type of measurements as before, just more.” What we mean are completely new types of data, with very different semantic content, that have played no role whatsoever in the entire modeling process up to this point. A key benefit of such an approach is that the resulting measurements are used primarily to “close-the-loop”, as advocated in [53], and provide a statistically clean separation between the data used for model selection and the data used for model validation—a feature that is alien to most of today’s network-related models. However, a key question remains: *What replaces data-fitting as the key ingredient and driver of the model selection and validation process so that the resulting models are approximately right and not certifiably wrong?* The simple answer is: *rely on domain knowledge and exploit the details that matter when dealing with a highly engineered system such as the Internet.* Note that this answer is in stark contrast to the statistical physics-based approach that suggests the development of a system such as the Internet is governed by robust self-organizing phenomena that go beyond the particulars of the individual systems (of interest) [8].

A first-principles approach to internet modeling

If domain knowledge is the key ingredient to build “approximately right” models of the Internet, what exactly is the process that helps us achieve our goal? To illustrate, we consider again the router-level topology of the Internet, or more specifically, the physical infrastructure of a regional, national, or international Internet Service Provider (ISP).

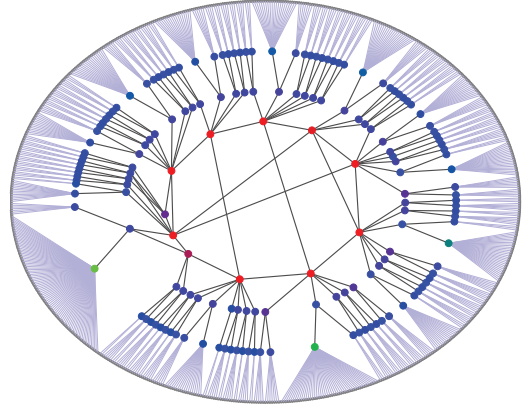
The first key observation is that the way an ISP designs its physical infrastructure is certainly not by a series of (biased) coin tosses that determine whether or not two nodes (i.e., routers) are connected by a physical link, as is the case for the scale-free network models of the PA type. Instead, ISPs design their networks for a purpose; that is, their decisions are driven by objectives and reflect trade-offs between what is feasible and what is desirable. The mathematical modeling language that naturally reflects such a decision making process is *constrained optimization*. Second, while in general it may be difficult if not impossible to define or capture the precise meaning of an ISP’s purpose for designing its network, an objective that expresses a desire to provide connectivity and an ability to carry an expected traffic demand efficiently and effectively, subject to prevailing economic and technological constraints, is unlikely to be far from the “true” purpose. In view of this, we are typically not concerned with a network design that is “optimal” in a strictly mathematical sense and is also likely to be NP-hard, but in a solution that is *“heuristically optimal”* in the sense that it results in “good” performance. That is, we seek a solution that captures by and large what the ISP can afford to build, operate, and manage (i.e., economic considerations), given the hard constraints that technology imposes on the network’s physical entities (i.e., routers and links). Such models have been discussed in the context of highly organized/optimized tolerances/tradeoffs (HOT) [18, 26]. Lastly, note that in this approach, randomness enters in a very specific manner, namely in terms of the uncertainty that exists about the “environment” (i.e., the traffic demand that the network is expected to carry), and the heuristically optimal network designs are expected to exhibit *strong robustness properties* with respect to changes in this environment.

Figure 3 shows a toy example of an ISP router-level topology that results from adopting the mathematical modeling language of constrained optimization and choosing a candidate network as a solution of an heuristically optimal network design problem. Despite being a toy example, it is rich enough to illustrate the key features of our engineering-driven approach to network modeling and to contrast it with the popular scale-free network models of the PA type. Its toy nature is mainly due to a number of simplifying assumptions we



$$\begin{aligned} \max_{\alpha} \sum_{i,j} x_{ij} &= \max \sum_{i,j} \alpha B_i B_j \\ \text{s.t.} \quad \sum_{i,j: k \in r_{ij}} x_{ij} &\leq B_k, \forall k \end{aligned}$$

(a)



(b)

Figure 5. Generating networks using constrained optimization. (a) Engineers view network structure as the solution to a design problem that measures performance in terms of the ability to satisfy traffic demand while adhering to node and arc capacity constraints. (b) A network resulting from heuristically optimized tradeoffs (HOT). This network has very different structural and behavioral properties, even when it has the same number of nodes, links, and degree distribution as the scale free network depicted in Fig. 1.

make that facilitate the problem formulation. For one, by simply equating throughput with revenues, we select as our objective function the maximum throughput that the network can achieve for a given traffic demand and use it as a metric for quantifying the performance of our solutions. Second, considering an arbitrary distribution of end-user traffic demand x_i , we assume a *gravity model* for the unknown traffic demand; that is, assuming shortest-path routing, the demands are given by the traffic matrix X , where for the traffic X_{ij} between routers i and j we have $X_{ij} = \rho x_i x_j$, for some constant ρ . Lastly, we consider only one type of router and its associated technologically feasible region; that is, (router degree, router capacity)-pairs that are achievable with the considered router type (e.g., CISCO 12416 GSR), and implicitly avoid long-haul connections due to their high cost.

The resulting constrained optimization problem can be written in the form

$$\begin{aligned} \max_{\rho} \quad & \sum_{i,j} X_{i,j} \\ \text{s.t.} \quad & RX \leq B \end{aligned}$$

where X is the vector obtained by stacking all the demands $X_{ij} = \rho x_i x_j$; R is the routing matrix obtained by using standard shortest path routing and defined by $R_{kl} = 1$ or 0, depending on whether or not demand l passes through router k ; and B is the vector consisting of the router degree-bandwidths constraints imposed by the technologically feasible region of the router at hand. While all the simplifying assumptions can easily be relaxed to allow for more realistic objective functions, more heterogeneity in the constraints, or more accurate

descriptions of the uncertainty in the environment, Figure 3 illustrates the key characteristics inherent in a heuristically optimal solution of such a problem. First, the cost-effective handling of end user demands avoids long-haul connections (due to their high cost) and is achieved through traffic aggregation starting at the edge of the network via the use of high-degree routers that support the multiplexing of many low-bandwidth connections. Second, this aggregated traffic is then sent toward the “backbone” that consists of the fastest or highest-capacity routers (i.e., having small number of very high-bandwidth connections) and that forms the network’s mesh-like core. The result is a network that has a more or less pronounced backbone, which is fed by tree-like access networks, with additional connections at various places to provide a degree of redundancy and robustness to failures.

What about power-law node degree distributions? They are clearly a non-issue in this engineering-based first-principles approach, just as they should be, based on our understanding illustrated earlier that present measurement techniques are incapable of supporting them. Recognizing their irrelevance is clearly the beginning of the end of the scale-free network models of the PA type as far as the Internet is concerned. What about replacing power-laws by the somewhat more plausible assumption of high variability in node degrees? While the answer of the scale-free modeling approach consists of tweaks to the PA mechanism to enforce an exponential cut-off of the power-law node degree distribution at the upper tail, the engineering-based approach

demystifies high-variability in node degrees altogether by identifying its root cause in the form of high variability in end-user bandwidth demands (see [33] for details). In view of such a simple physical explanation of the origins of node degree variability in the Internet's router-level topology, Strogatz' question, paraphrasing Shakespeare's Macbeth, "... power-law scaling, full of sound and fury, signifying nothing?" [52] has a resounding affirmative answer.

Great Potential for Mathematics

Given the specious nature of scale-free networks of the PA type for modeling Internet-related connectivity structures, their rigorous mathematical treatment and resulting highly-publicized properties have lost much of their luster, at least as far as Internet matters are concerned. Considering again our example of the router-level Internet, neither the claim of a hub-like core, nor the asserted error tolerance (i.e., robustness to random component failure) and attack vulnerability (i.e., Achilles' heel property), nor the often-cited zero epidemic threshold property hold. In fact, as illustrated with our HOT-based network models, intrinsic and unavoidable tradeoffs between network performance, available technology, and economic constraints necessarily result in network structures that are in all important ways exactly the opposite of what the scale-free models of the PA type assert. In this sense, the HOT toy examples represent a class of network models for the Internet that are not only consistent with various types of measurements and in agreement with engineering intuition, but whose rigorous mathematical treatment promises to be more interesting and certainly more relevant and hence more rewarding than that of the scale-free models of the PA type.

The Internet's robust yet fragile nature

Because high-degree nodes in the router-level Internet can exist only at the edge of the network, their removal impacts only local connectivity and has little or no global effect. So much for the widely-cited discovery of the Internet's Achilles' heel! More importantly, the Internet is known to be extremely robust to component failures, but this is *by design*² and involves as a critical ingredient the Internet Protocol (IP) that "sees failures and routes traffic around them." Note that neither the presence of protocols nor their purpose play any role in the scale-free approach to assessing the robustness properties of the Internet. At the same time, the Internet is also known to be very fragile, but again in a sense that is completely different from and has nothing in common with either the sensational Achilles' heel claim or the zero epidemic threshold

property, both of which are irrelevant as far as the actual Internet is concerned. The network's true fragility is due to an original trust model³ that has been proven wrong almost from the get-go and has remained broken ever since. While worms, viruses, or spam are all too obvious and constant reminders of this broken trust model, its more serious and potentially lethal legacy is that it facilitates the malicious exploitation or hijacking of the very mechanisms (e.g., protocols) that ensure the network's impressive robustness properties. This "robust yet fragile" tradeoff is a fundamental aspect of an Internet architecture whose basic design dates back some 40 years and has enabled an astonishing evolution from a small research network to a global communication infrastructure supporting mission-critical applications.

One of the outstanding mathematical challenges in Internet research is the development of a theoretical foundation for studying and analyzing this robustness-fragility tradeoff that is one of the single most important characteristics of complexity in highly engineered systems. To date, this tradeoff has been largely managed with keen engineering insights and little or no theoretical backing, but as the Internet scales even further and becomes ever more heterogeneous, the need for a relevant mathematical theory replacing engineering intuition becomes more urgent. The difficulties in developing such a theory are formidable as the "typical" behavior of a system such as the Internet is often quite simple, inviting naive views and models like the scale-free network models of the PA type that ignore any particulars of the underlying system, inevitably cause confusion, result in misleading claims, and provide simple explanations that may look reasonable at first sight but turn out to be simply wrong. Only extreme circumstances or rare accidents not easily replicable in laboratory experiments or simulations reveal the enormous internal complexity in systems such as the Internet, and any relevant mathematical theory has to respect the underlying architectural design and account for the various protocols whose explicit purpose is in part to hide all the complexity from the user of the system [5].

Network dynamics and system function

Real networks evolve over time in response to changes in their environment (e.g., traffic, technology, economics, government regulation), and currently proposed network models such as the scale-free models of the PA type cannot account for such interactions. They either ignore the notion of network *function* (i.e., the delivery of traffic) altogether or treat networks as strictly open-loop systems in which modeling exists largely as an

²Being robust to component failures was the number one requirement in the original design of the Internet [24].

³The original Internet architects assumed that all hosts can be trusted [24].

exercise in data-fitting. In stark contrast to the scale-free models of the PA type, the proposed HOT-based network models make the dependence of network structure on network traffic explicit. This is done by requiring as input a traffic demand model in the form of a traffic matrix (e.g., gravity model). A particular network topology is “good” only if it can deliver traffic in a manner that satisfies demand. When viewed over time, changes in the environment (e.g., traffic demands), constraints (e.g., available technologies), or objectives (e.g., economic conditions) are bound to impact the structure of the network, resulting in an intricate feedback loop between network traffic and network structure.

The task at hand is to develop a mathematical framework that enables and supports modeling network evolution in ways that account for this feedback loop between the structure of the networks and the traffic that they carry or that gets routed over them. This new modeling paradigm for networks is akin to recent efforts to model the network-wide behavior of TCP or the TCP/IP protocol stack as a whole: the modeling language is (constrained) optimization; a critical ingredient is the notion of separation of time scales; heuristic solution methods (with known robustness properties to changes in the environment) are preferred over mathematically optimal solution techniques (which are likely to be NP-hard); and the overall goal is to transform network modeling from an exercise in data-fitting into an exercise in reverse-engineering. In this sense, relevant recent theoretical works includes *network utility maximization* (e.g., see [30, 36, 37]), *layering as optimization decomposition* (e.g., see [21, 22]), and *the price of anarchy* (e.g., see [45]).

In view of this objective, developing a mathematical framework for studying sequences and limits of graphs that arise in a strictly open-loop manner (e.g., see [17, 19, 20]), while of independent mathematical interest, is of little relevance for studying and understanding real-world networks such as the Internet, unless it is supported by strong and convincing validation efforts. This difference in opinions is fully expected: while mathematicians and physicists tend to view the enormous dynamic graph structures that arise in real-world applications as too complex to be described by direct approaches and therefore invoke randomness to model and analyze them, Internet researchers generally believe to have enough domain knowledge to understand the observed structures in great detail and tend to rely on randomness for the sole purpose of describing genuine uncertainty about the environment. While both approaches have proven to be useful, it is the responsibility of the mathematician/physicist to convince the Internet researcher of the relevance or usefulness of their modeling effort. The scale-free models of the PA

type are an example where this responsibility has been badly lacking.

Multiscale network representations

Multiscale representations of networks is an area where Ulam’s paraphrased quote “*Ask not what mathematics can do for [the Internet]; ask what [the Internet] can do for mathematics.*” is highly appropriate. On the one hand, there exists a vast literature on mathematical multi-resolution analysis (MRA) techniques and methodologies for studying complex objects such as high-dimensional/semantic-rich data and large-scale structures. However, much less is known when it comes to dealing with highly irregular domains such as real-world graph structures or with functions or distributions defined on those domains. In fact, from an Internet perspective, what is needed is an MRA specifically designed to accommodate the vertical (i.e., layers) and horizontal (i.e., administrative or geographic domains) decompositions of Internet-like systems and capture in a systematic manner the “multi-scale” nature of the temporal, spatial, and functional aspects of network traffic over corresponding network structures. In short, the mathematical challenge consists of developing an MRA technology appropriate for dealing with meaningful multi-scale representations of very large, dynamic, and diverse Internet-specific graph structures; for exploring traffic processes associated with those structures; and for studying aggregated spatio-temporal network data representations and visual representations of them.

The appeal of an Internet-specific MRA is that the Internet’s architecture supports a number of meaningful and relevant multi-scale network representations with associated traffic processes. For example, starting with our example of the router-level Internet (and associated hypothetical traffic matrix), aggregating routers and the traffic they handle into Points-of-Presences, or PoPs, yields the PoP-level Internet and PoP-level traffic matrix. Aggregating PoPs and the traffic they handle into Autonomous Systems (ASes) or domains produces the AS-level Internet and corresponding AS-level traffic matrix. Aggregating even further, we can group ASes that belong to the same Internet Service Provider (ISP) or company/institution and obtain the ISP-level Internet. While the router- and PoP-level Internet are inherently physical representations of the Internet, the AS- and ISP-level structures are examples of logical or virtual constructs where nodes and links say little or nothing about physical connectivity. At the same time, the latter are explicit examples that support a meaningful view of the Internet as a “network of networks” (see below). With finer-resolution structures and traffic matrices also of interest and of possible

use (e.g., BGP prefix-level, IP address-level), the expectations for an Internet-specific MRA technique are that it is capable of recovering these multiple representations by respecting the architectural, administrative, and technological aspects that give rise to this natural hierarchical decomposition and representation of the Internet. While traditional wavelet-based MRA techniques have proven to be too rigid and inflexible to meet these expectations, more recent developments concerning the use of *diffusion wavelets* (e.g., see [40, 41]) show great promise and are presently explored in the context of Internet-specific structures.

Networks of networks

Changing perspectives, we can either view the Internet as a “network of networks” (e.g., AS-level Internet) or consider it as one of many networks that typically partake in the activities of enterprises: transportation of energy, materials, and components; power grid; supply chains, and control of transportation assets; communication and data networks. The networks’ activities are correlated because they are invoked to support a common task, and the networks are interdependent because the characteristics of one determine the inputs or constraints for another. They are becoming even more correlated and interdependent as they shift more and more of their controls to be information intensive and data network-based. While this “networks of networks” concept ensures enormous efficiency and flexibility, both technical and economical, it also has a dark side—by requiring increasingly complex design processes, it creates vastly increased opportunities for potentially catastrophic failures, to the point where national and international critical infrastructure systems are at risk of large-scale disruptions due to intentional attacks, unintentional (but potentially devastating) side effects, the possibility of (not necessarily deliberate) large cascading events, or their growing dependence on the Internet as a “central nervous system.”

This trend in network evolution poses serious questions about the reliability and performability of these critical infrastructure systems in the absence of an adequate theory [46]. Thus the long-term goal of any mathematical treatment of networked systems should be to develop the foundation of a nascent theory in support of such a “networks of networks” concept. To this end, the Internet shows great promise to serve as a case study to illustrate how early verbal observations and arguments with deep engineering insight have led via an interplay with mathematics and measurements to increasingly formal statements and powerful theoretical developments that can be viewed as a precursor of a full-fledged “network of networks” theory.

Conclusion

Over the last decade, there has been a compelling story articulated by the proponents of network science. Advances in information technology have facilitated the collection of petabyte scale data sets on everything from the Internet to biology to economic and social systems. These data sets are so large that attempts even to visualize them are nontrivial and often yield nonsensical results. Thus the “Petabyte Age” requires new modeling approaches and mathematical techniques to identify hidden structures, with the implication that these structures are fundamental to understanding the systems from which the vast amounts of measurements are derived. In extreme cases, this perspective suggests that the ubiquity of petabyte scale data on *everything* will fundamentally change the role of experimentation in science and of science as a whole [7].

In this article we have presented a retrospective view of key issues that have clouded the popular understanding and mathematical treatment of the Internet as a complex system for which vast amounts of data are readily available. Foremost among these issues are the dangers of taking available data “at face value” without a deeper understanding of the idiosyncracies and ambiguities resulting from domain-specific collection and measurement techniques. When coupled with the naive but commonly-accepted view of validation that simply argues for replicating certain statistical features of the observed data, such an “as is” use of the available data reduces complex network modeling to mere “data fitting,” with the expected and non-informative outcome that given sufficient parameterization, it is always possible to match a model to any data set without necessarily capturing any underlying hidden structure or key functionality of the system at hand.

For systems whose measured features are subject to fundamental ambiguities, omissions, and/or systematic errors, we have proposed an alternative approach to network modeling that emphasizes data hygiene (i.e., practices associated with determining the quality of the available data and assessing their proper use) and uses constrained optimization as modeling language to account for the inherent objectives, constraints, and domain-specific environmental conditions underlying the growth and evolution of real-world complex networks. We have shown that in the context of the router-level Internet, this approach yields models that not only respect the forces shaping the real Internet but also are robust to the deficiencies inherent in available data.

In this article, the Internet has served as a clear case study, but the issues discussed apply more generally and are even more pertinent in contexts of biology and social systems, where measurement is inherently more difficult and more

error prone. In this sense, the Internet example serves as an important reminder that despite the increasing ubiquity of vast amounts of available data, the “Garbage In, Gospel Out” extension of the phrase “Garbage In, Garbage Out” remains as relevant as ever; no amount of number crunching or mathematical sophistication can extract knowledge we can trust from low-quality data sets, whether they are of petabyte scale or not. Although the Internet story may seem all too obvious in retrospect, managing to avoid the same mistakes in the context of next-generation network science remains an open challenge. The consequences of repeating such errors in the context of, say, biology are potentially much more grave and would reflect poorly on mathematics as a discipline.

References

- [1] D. ACHLIOPTAS, A. CLAUSET, D. KEMPE, and C. MOORE, On the bias of traceroute sampling, *Journal of the ACM* (to appear), 2008.
- [2] R. ALBERT, H. JEONG, and A.-L. BARABÁSI, Error and attack tolerance of complex networks, *Nature* **406** (2000).
- [3] R. ALBERT and A.-L. BARABÁSI, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74** (2002).
- [4] D. L. ALDERSON, Catching the “Network Science” Bug: Insight and Opportunity for the Operations Researcher, *Operations Research* 2008 (in press).
- [5] D. L. ALDERSON and J. C. DOYLE, Contrasting views of complexity and their implications for network-centric infrastructures, *IEEE Trans. on SMC-A* (submitted), 2008.
- [6] D. ALDERSON and L. LI, Diversity of graphs with highly variable connectivity, *Phys. Rev. E* **75**, 046102, 2007.
- [7] C. ANDERSON, The end of theory, *Wired Magazine* **16**, July 2008.
- [8] A.-L. BARABÁSI and R. ALBERT, Emergence of scaling in random networks, *Science* **286** (1999).
- [9] A. BENDER, R. SHERWOOD, and N. SPRING, Fixing ally’s growing pains with velocity modeling. *Proc. ACM IMC*, 2008.
- [10] E. BENDER and E. R. CANFIELD, The asymptotic number of labeled graphs with given degree sequences, *J. of Comb. Theory A* **24** (1978), 296–307.
- [11] N. BERGER, C. BORGS, J. T. CHAYES, and A. SABERI, On the spread of viruses on the Internet, *Proc. SODA’05*, 2005.
- [12] W. A. BEYER, P. H. SELLERS, and M. S. WATERMAN, Stanislaw M. Ulam’s contributions to theoretical theory, *Letters in Mathematical Physics*, **10** (1985), 231–242.
- [13] B. BOLLOBÁS, *Random Graphs*, Academic Press, London, 1985.
- [14] B. BOLLOBÁS and O. RIORDAN, Mathematical results on scale-free random graphs, *Handbook of Graphs and Networks* (S. Bornholdt and H. G. Schuster, eds.), Wiley-VCH, Weinheim, 2002.
- [15] ———, Robustness and vulnerability of scale-free graphs, *Internet Mathematics*, **1**(1) (2003), 1–35.
- [16] ———, The diameter of a scale-free random graph. *Combinatorica*, **24**(1) (2004), 5–34.
- [17] C. BORGS, J. T. CHAYES, L. LOVÁSZ, V. T. SÓS, B. SZEGEDY, and K. VESZTERGOMBI, Graph limits and parameter testing, *Proc. STOC’06*, 2006.
- [18] J. M. CARLSON and J. C. DOYLE, Complexity and robustness, *Proc. Nat. Acad. of Sci. USA*, **99** (2002), 2538–2545.
- [19] J. T. CHAYES, C. BORGS, L. LOVÁSZ, V. T. SÓS, and K. VESZTERGOMBI, Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing, preprint, research.microsoft.com/ˆchayes, 2006.
- [20] ———, Convergent sequences of dense graphs II: Multiway cuts and statistical physics, preprint, research.microsoft.com/ˆchayes, 2007.
- [21] M. CHIANG, S. H. LOW, A. R. CALDERBANK, and J. C. DOYLE, Layering as optimization decomposition, *Proc. of the IEEE* **95** (2007).
- [22] L. CHEN, S. H. LOW, M. CHIANG, and J. C. DOYLE, Cross-layer congestion control, routing and scheduling design in ad-hoc wireless networks, *Proc. IEEE INFOCOM’06*, 2006.
- [23] F. CHUNG and L. LU, The average distance in a random graph with given expected degrees, *Internet Math.* **1** (2003), 91–113.
- [24] D. D. CLARK, The design philosophy of the Darpa Internet protocol, *ACM Computer Communication Review*, **18**(4) (1988), 106–114.
- [25] R. DURRETT, *Random Graph Dynamics*, Cambridge University Press, New York, 2007.
- [26] A. FABRIKANT, E. KOUTSOUPAS, and C. PAPADIMITRIOU, Heuristically optimized trade-offs: A new paradigm for power-laws in the internet, *Proc. ICALP*, 2002, 110–122.
- [27] M. FALOUTSOS, P. FALOUTSOS, and C. FALOUTSOS, On power-law relationships of the Internet topology, *ACM Comp. Comm. Review*, **29**(4) (1999).
- [28] M. H. GUNES and K. SARAC, Resolving IP aliases in building traceroute-based Internet maps, *IEEE/ACM Trans. Networking* (to appear) 2008.
- [29] S. HAKIMI, On the realizability of a set of integers as degrees of the vertices of a linear graph, *SAM. J. Appl. Math.* **10** (1962), 496–506.
- [30] F. P. KELLY, A. MAULLOO, and D. TAN, Rate control in communication networks: Shadow prices, proportional fairness stability, *Journal of the Operational Research Society*, **49** (1998), 237–252.
- [31] B. KRISHNAMURTHY and W. WILLINGER, What are our standards for validation of measurement-based networking research? *Proc. HotMetrics’08*, 2008.
- [32] A. LAKHINA, J. W. BYERS, M. CROVELLA, and P. XIE, Sampling Biases in IP topology Measurements, *IEEE INFOCOM 2003*.
- [33] L. LI, D. ALDERSON, W. WILLINGER, and J. C. DOYLE, A first principles approach to understanding the Internet’s router-level topology, *Proc. ACM SIGCOMM’04*, **34**(4) (2004), 3–14.
- [34] L. LI, D. ALDERSON, J. C. DOYLE, and W. WILLINGER, Towards a theory of scale-free graphs: Definitions, properties, and implications. *Internet Mathematics*, **2**(4) (2005), 431–523.
- [35] S. Y. R. LI, Graphic sequences with unique realization, *J. Combin. Theory B* **19** (1975,) 42–68.
- [36] S. H. LOW, A duality model of TCP and queue management algorithms, *IEEE/ACM Trans. on Networking*, **11**(4) (2003), 525–536.

- [37] J. WANG, L. LI, S. H. LOW, and J. C. DOYLE, Cross-layer optimization in TCP/IP networks, *IEEE/ACM Trans. on Networking*, **13** (2005), 582–595.
- [38] S. E. LURIA and M. DELBRÜCK, Mutations of bacteria from virus sensitivity to virus resistance, *Genetics*, **28** (1943), 491–511.
- [39] B. B. MANDELBROT, *Fractals and Scaling in Finance*, Springer-Verlag, New York, 1997.
- [40] M. MAGGIONI, A. D. SZLAM, R. R. COIFMAN, and J. C. BRENNER, Diffusion-driven multiscale analysis on manifolds and graphs: top-down and bottom-up constructions, *Proc. SPIE Wavelet XI*, 5914, 2005.
- [41] M. MAGGIONI, J. C. BRENNER, R. R. COIFMAN, and A. D. SZLAM, Biorthogonal diffusion wavelets for multiscale representations on manifolds and graphs, *Proc. SPIE Wavelet XI*, 5914, 2005.
- [42] National Reserch Council Report, *Network Science*, National Academies Press, Washington, 2006.
- [43] R. OLIVEIRA, D. PEI, W. WILLINGER, B. ZHANG, and L. ZHANG, In search of the elusive ground truth: The Internet’s AS-level connectivity structure, *Proc. ACM SIGMETRICS*, 2008.
- [44] J.-J. PANSIOT and D. GRAD, On routes and multicast trees in the Internet, *ACM Computer Communication Review*, **28**(1) (1998).
- [45] C. H. PAPADIMITRIOU, Algorithms, games, and the Internet, *Proc. STOC’01*, 2001.
- [46] President’s Commission on Critical Infrastructure Protection. Tech. Report, The White House, 1997.
- [47] H. J. RYSER, Combinatorial properties of matrices of zeroes and ones, *Canad. J. Math.* **9** (1957), 371–377.
- [48] R. SHERWOOD, A. BENDER, and N. SPRING, Dis-Carte: A disjunctive Internet cartographer, *Proc. ACM SIGCOMM*, 2008.
- [49] H. A. SIMON, On a class of skew distribution functions, *Biometrika* **42** (1955), 425–440.
- [50] N. SPRING, R. MAHAJAN, and D. WETHERALL, Measuring ISP topologies with Rocket fuel, *Proc. ACM SIGCOMM*, 2002.
- [51] N. SPRING, M. DONTCHEVA, M. RODRIG, and D. WETHERALL, How to Resolve IP Aliases, *UW CSE Tech. Report* 04-05-04, 2004.
- [52] S. STROGATZ, Romanesque networks, *Nature*, **433** (2005).
- [53] W. WILLINGER, R. GOVINDAN, S. JAMIN, V. PAXSON, and S. SHENKER, Scaling phenomena in the Internet: Critically examining criticality, *Proc. Nat. Acad. Sci.*, **99** (2002), 2573–2580.
- [54] G. YULE, A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, *F.R.S. Philosophical Transactions of the Royal Society of London (Series B)*, **213** (1925), 21–87.